Model Based Signal Processing Techniques for Nonconventional Optical Imaging Systems

Daniele Picone

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : SIGNAL IMAGE PAROLE TELECOMS

Arrêté ministériel : 25 mai 2016

Présentée par

Daniele PICONE

Thèse dirigée par **Mauro DALLA MURA**, Université Grenoble Alpes
et co-encadrée par **Laurent CONDAT**, CNRS

préparée au sein du **Laboratoire Grenoble Images Parole Signal Automatique**
dans **l'École Doctorale Electronique, Electrotechnique, Automatique, Traitement du Signal (EEATS)**

Techniques de traitement du signal basées modèles pour systèmes d'imagerie optique non conventionnels

Model Based Signal Processing Techniques for Nonconventional Optical Imaging Systems

Thèse soutenue publiquement le **25 novembre 2021**,
devant le jury composé de :

Monsieur Mauro DALLA MURA
MAITRE DE CONFERENCE HDR, GRENOBLE INP, Directeur de thèse
Monsieur Andrés ALMANSA
DIRECTEUR DE RECHERCHE, CNRS ILE-DE-FRANCE VILLEJUIF, Rapporteur
Monsieur Magnús Örn ÚLFARSSON
PROFESSEUR, University of Iceland, Rapporteur
Madame Valérie PERRIER
PROFESSEUR DES UNIVERSITES, GRENOBLE INP, Présidente
Monsieur Enrico MAGLI
PROFESSEUR, Politecnico di Torino, Examinateur
Monsieur Etienne LE COARER
INGENIEUR HDR, UNIVERSITE GRENOBLE ALPES, Examinateur

Model Based Signal Processing Techniques for Nonconventional Optical Imaging Systems

---

Daniele Picone

November 25, 2021
Version: Camera ready

GIPSA-lab
Laboratoire Grenoble Images Parole Signal Automatique

ED EEATS
l'École Doctorale Électronique, Électrotechnique, Automatique, Traitement du Signal

A thesis submitted for the degree of
Doctor of Philosophy
at the Université Grenoble Alpes
Specialisation: Signal, Image, Parole, Télécoms (SIPT)

Model Based Signal Processing Techniques for Nonconventional Optical Imaging Systems

Presented and defended by
Daniele Picone

Committee Members:

President: **Valérie Perrier**, Professeur des universités
Université Grenoble Alpes, CNRS, Grenoble INP, LJK, Grenoble, France

Reviewers: **Andrés Almansa**, Directeur de recherche au CNRS
MAP5, CNRS, Université Paris Descartes, Paris, France

Magnús Örn Úlfarsson, Professor
University of Iceland, Reykjavík, Iceland

Examiners: **Enrico Magli**, Professore ordinario
Politecnico di Torino, Turin, Italy

Etienne le Coarer, Ingénieur de recherche
IPAG/UGA-CNRS, Université Grenoble Alpes, Grenoble, France

Supervisor: **Mauro Dalla Mura**, Maître de conférences
Université Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France
Institut Universitaire de France, Paris, France
Tokyo Tech WRHI, School of Computing, Tokyo Institute of Technology, Tokyo, Japan

Co-supervisor: **Laurent Condat**, Research scientist
King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

November 25, 2021

Daniele Picone
Model Based Signal Processing Techniques for Nonconventional Optical Imaging Systems
Doctor of Philosophy, November 25, 2021
l'École Doctorale Électronique, Électrotechnique, Automatique, Traitement du Signal
Specialisation: Signal, Image, Parole, Télécoms (SIPT)
Reviewers: Andrés Almansa and Magnús Örn Úlfarsson
Examiners: Valérie Perrier, Enrico Magli and Etienne le Coarer
Supervisors: Mauro Dalla Mura and Laurent Condat

Université Grenoble Alpes
Equipe SIGnal iMAge PHYsique (SIGMAPHY)
Laboratoire Grenoble Images Parole Signal Automatique (GIPSA-lab)
Department of Image and Signal (DIS)
11 rue des Mathématiques, Grenoble Campus BP46
38402 Saint Martin d'Hères

Abstract

There is an increasing demand for images with higher spectral and spatial resolution for applications in several domains such as health, environment, quality checking and natural disasters monitoring. Hyperspectral imagery provides the necessary spectral diversity to recover the composition of materials on site for applications such as the detection of fires, anomalies, chemical agents, targets and changes in the scene. The requirement for cheaper and more compact devices (e.g. to be embarked on low cost satellites and airborne platform) which are capable of capturing this information has led to the development of nonconventional innovative design concepts to overcome the technological limitations of traditional cameras. Data acquired by such novel imaging devices following the computational imaging paradigm are typically not readily exploitable for the final application. A computational phase is hence needed for extracting useful information from the raw acquisitions.

This thesis addresses this issue by setting up an inversion problem. The general approach is to characterize the data fidelity term with a physical model, describing the underlying optical transformations performed by the device. The challenge is then shifted on the regularization step to properly characterize the features of the quantities of interest and improve the accuracy of the estimation, which can be tackled with variational techniques.

The analysis is applied on two novel concepts for nonconventional optical devices. The first one is a novel compressed acquisition imaging system based on color filter arrays, which embeds information from sensors with different spatial and spectral characteristics into a single mosaiced product. As opposed to existing compressed sensing based devices, the goal is not to recover the original uncompressed multiresolution sources, but instead to directly recover a synthetic fused image with both high spatial and spectral resolutions. The proposed solution relies on the total variation regularization and is the subject of a detailed analysis, comparing its compressive power with straightforward software alternatives, evaluating its performances as the amount of channels changes, and validating its efficiency in comparison to state of the art methods when applied to classical fusion or mosaicing algorithms separately. The second class of devices is based on the ImSPOC patent, a design concept for a low finesse snapshot imaging spectrometer based on the interferometry of Fabry-Pérot. Its ideal behaviour follows the principle of the Fourier

Transform Spectroscopy, as its acquisition can be interpreted as a sampled version of an interferogram, arranged across different sub-images distributed on the same focal plane. After defining a physical model based on optical geometry, its validity is evaluated over real acquisitions by setting up a Bayesian inference problem to determine its parameters, with approaches based on maximum likelihood estimators, regular-grid searches and nonlinear regression. A variety of preliminary tests are then carried out on the inversion method, with approaches based on singular value decomposition and sparse-inducing regularizers, accompanied by an analysis of their robustness to model mismatches.

Résumé

Il existe une demande croissante d'images avec une résolution spectrale et spatiale plus élevée pour des applications dans plusieurs domaines tels que la santé, l'environnement, le contrôle qualité et la surveillance des catastrophes naturelles. L'imagerie hyperspectrale fournit la diversité spectrale nécessaire pour récupérer la composition des matériaux sur site pour des applications telles que la détection d'incendies, d'anomalies, d'agents chimiques, de cibles et de changements de scène. L'exigence de dispositifs moins chers et plus compacts (par exemple, pour être embarqués sur des satellites à faible coût et une plateforme aéroportée) capables de capturer ces informations a conduit au développement de concepts de conception innovants non conventionnels pour surmonter les limitations technologiques des caméras traditionnelles.

Les données acquises à partir de ces nouveaux dispositifs d'imagerie suivant le paradigme d'imagerie informatique ne sont généralement pas facilement exploitables pour l'application finale. Une phase de calcul est nécessaire pour extraire des informations utiles des acquisitions brutes.

Cette thèse aborde cette question en mettant en place un problème d'inversion. L'approche générale consiste à caractériser le terme de fidélité des données avec un modèle physique, décrivant les transformations optiques sous-jacentes effectuées par le dispositif. Le défi est ensuite déplacé vers l'étape de régularisation pour bien caractériser les caractéristiques des quantités d'intérêt et améliorer la précision de l'estimation, ce qui peut être abordé avec des techniques variationnelles. L'analyse est appliquée à deux nouveaux concepts de dispositifs optiques non conventionnels.

Le premier est un nouveau système d'imagerie d'acquisition compressé basé sur des matrices de filtres de couleur, qui intègre des informations provenant de capteurs avec différentes caractéristiques spatiales et spectrales dans un seul produit mosaïqué. Contrairement aux dispositifs existants basés sur la détection compressée, l'objectif n'est pas de récupérer les sources multirésolutions non compressées d'origine, mais plutôt de récupérer directement une image fusionnée synthétique avec une résolution spatiale et spectrale élevée. La solution proposée repose sur la régularisation de la variation totale et fait l'objet d'une analyse détaillée, comparant sa puissance de compression avec des alternatives logicielles simples, évaluant ses performances au fur et à mesure que le nombre de canaux change, et validant son efficacité par rapport aux méthodes de l'état de l'art lorsque appliqué séparément aux algorithmes classiques de fusion ou de mosaïquage. La deuxième classe d'appareils considérée dans ce travail est basée sur le brevet ImSPOC, un concept de conception pour un spectromètre imageur instantané de faible finesse basé sur l'interférométrie de Fabry-Pérot. Son comportement idéal suit le principe de la spectroscopie à transformée de Fourier, car son acquisition peut être interprétée comme une version échantillonnée d'un interférogramme, disposée sur différentes sous-images réparties sur le même plan focal. Après avoir défini un modèle physique basé sur la géométrie optique, sa validité est évaluée sur des acquisitions réelles en mettant en place un problème d'inférence bayésienne pour déterminer ses paramètres, avec des approches basées sur des estimateurs du maximum de vraisemblance, des recherches en grille régulière et une régression non linéaire. Divers tests préliminaires sont ensuite menés sur la méthode d'inversion, avec des approches basées sur la décomposition en valeurs singulières et les régularisations creuses, accompagnées d'une analyse de leur robustesse aux mésappariements de modèles.

Abstract intended to a wider audience

There is a continuous quest for finer resolution images. The limits of traditional imaging systems (e.g., RGB cameras) are constantly pushed as applications demand increasingly spatially and spectrally resolved images for better sensing the phenomena of interest. Moreover, increasingly more compact and less expensive prototypes are also in high demand, opening new applications when mounted on space or airborne platforms.

Novel acquisition strategies, such as those obtained by "computational imaging devices", attempt to answer these needs by overcoming the technical limitations of

traditional imaging devices. This comes at the expenses of a heavier computational phase. For example, the raw acquisitions of these nonconventional imaging systems are often unintelligible to the final user and require a signal processing to extract useful information.

This thesis addresses the development of signal and image processing approaches for the analysis of acquisitions obtained by two nonconventional optical imaging systems. The first explores a novel strategy for the compressed acquisition of a high spatial resolution monochromatic image and a lower resolution multispectral image of the same scene. The retrieval of a high spatial multispectral image of the scene requires to perform a joint fusion and reconstruction of the raw acquisitions. The second prototype is a novel snapshot Fourier Transform imaging spectrometer developed in Grenoble based on a matrix of Fabry-Pérot interferometers. A hyperspectral image of the scene can be obtained from the processing of the raw acquisitions which are composed of a series of interferograms. This requires to perform a characterization of the imaging device and an inversion of the raw acquisitions.

The computational techniques proposed in this manuscript rely on physical models representing the acquisition process of the two imaging prototypes considered. Image reconstruction is then addressed as an inverse problem and tackled with variational techniques.

Résumé destiné à un public plus large

Il y a une quête continue pour des images à résolution plus fine. Les limites des systèmes d'imagerie traditionnels (par exemple, les caméras RVB) sont constamment repoussées car les applications exigent de plus en plus des images résolues spatialement et spectralement pour mieux détecter les phénomènes d'intérêt. De plus, des prototypes de plus en plus compacts et moins chers sont également très demandés, ouvrant de nouvelles applications lorsqu'ils sont montés sur des plateformes dans l'espace ou aéroportées.

Des nouvelles stratégies d'acquisition, telles que celles obtenues par les dispositifs baseés su la co-conception computationelle/optique, tentent de répondre à ces besoins en surmontant les limitations techniques des dispositifs d'imagerie traditionnels. Cela se fait au prix d'une phase de calcul plus lourde. Par exemple, les acquisitions brutes de ces systèmes d'imagerie non conventionnels sont souvent

inintelligibles pour l'utilisateur final et nécessitent un traitement du signal pour en extraire des informations utiles.

Cette thèse porte sur le développement d'approches de traitement du signal et de l'image pour l'analyse des acquisitions obtenues par deux systèmes d'imagerie optique non conventionnels. La première explore une nouvelle stratégie pour l'acquisition compressée d'une image monochromatique à haute résolution spatiale et d'une image multispectrale à plus faible résolution de la même scène. La récupération d'une image multispectrale spatiale élevée de la scène nécessite de réaliser une fusion et une reconstruction conjointes des acquisitions brutes. Le deuxième prototype est un nouveau spectromètre imageur instantané à transformée de Fourier développé à Grenoble à partir d'une matrice d'interféromètres de Fabry-Pérot. Une image hyperspectrale de la scène peut être obtenue à partir du traitement des acquisitions brutes qui sont composées d'une série de interférogrammes. Cela nécessite de réaliser une caractérisation de l'appareil d'imagerie et une inversion des acquisitions brutes.

Les techniques de calcul proposées dans ce manuscrit reposent sur des modèles physiques représentant le processus d'acquisition des deux prototypes d'imagerie considérés. La reconstruction d'images est ensuite abordée comme un problème inverse et abordée avec des techniques variationnelles.

Acknowledgement

First and foremost I am extremely grateful to my supervisors, Prof. Mauro Dalla Mura and Prof. Laurent Condat for their priceless advices, the creative environment that they created for a mature development of the ideas to this thesis, and their availability at any stage of my PhD study. I would also like to thank all the members of the working groups linked to the ImSPOC project, which are spread among the GIPSA-lab, IPAG and ONERA laboratories, especially Dr. Silvère Gousset, Prof. Etienne le Coarer, Dr. Aneline Dolet, Prof. Sylvain Douté, Prof. Stéphane Mancini, Prof. Didier Voisin, Dr. El Mehdi Abdali, and Prof. Yann Ferrec, for sharing their knowledge in different fields of expertise, necessary for a deeper understanding of the topics touched in this thesis.

Very special thanks to the members of my CSI (Comité Suivi de Thèse) committee, Prof. Michel Desvignes and Prof. Hacheme Ayasso, for their patience and precious suggestions during this long journey, as well as the members of the PhD committee for having accepted to review my work and for their precious feedback.

I cannot forget to mention Prof. Kuniaki Uto and all the members of the Shinoda Laboratory of the Tokyo Institute of Technology, for the precious opportunity that they granted to visit their work group and for their hospitality during my time in Japan.

I would also like to thank all the people I met during the time of this thesis, in particular Dr. Mohamad Jouni, for the wide variety of topics and ideas we discussed while sharing our office, together with all the people that made my time in France stimulating and enjoyable. I also include in this circle the people that weren't present in my life in person, but were always ready to emotionally support me and push me forward to complete this work.

I would finally express my most heartfelt gratefulness to my parents, for all their time, as well as their emotional and financial support that they provided me during these years; as a person that gets completely enthralled by scientific discoveries, I commend them, together with my brother and my sister, for their understanding and patience, especially in the moments I am the most concentrated on my work.

Contents

List of Figures

List of Tables

1 Introduction

1.1 General overview

In recent years, both the industry and the scientific community have shown increasing interest for high-quality images that complement the information with finely sampled spectra, which allow to extract features of a given portion of a scene that were previously unavailable with traditional cameras.

Hyperspectral imaging (HSI) addresses these demands by measuring the light spectrum in a contiguous set of wavelengths, thus providing novel information for the analysis of the scene, which includes change detection, identification of materials, chemical imaging, vegetation monitoring [52].

The target of making HSI cheaper, more compact, easy to interpret and readily available for the final user is an open goal in many fields of research [42, 99].

In this context, **computational imaging** is a quickly growing field that provides a novel approach to face such issues; this domain encompasses all digital image capture and processing techniques that combine computation and acquisitions from optical imaging devices.

Devices designed following this principle exploit the advantages of nonconventional cutting-edge technologies to reach better performances with respect to traditional cameras, e.g. in terms of signal to noise ratio (SNR), resolution, compactness, and cost. Unfortunately, the acquired raw products are most of the times not intelligible for the final user, leaving the bulk of the work to the algorithmic side to recover the quantities of interest. These quantities are described, in general terms, by a synthetic datacube which emulates an acquisition taken by a standard HSI device.

This thesis focuses on the analysis of nonconventional optical devices from the perspective of the software engineer, viewed under the framework of **inverse problems**. In principle, the pipeline of operations of this study can be separated in two steps. In the **characterization** phase, the task is reduced to develop a physical model that is able to accurately model the optical transformations performed by the device. In

the **inversion** phase, the target is to recover a robust estimation of the desired final product.

While computational imaging systems cover a broad range of applications [143], our analysis will focus on novel concepts of snapshot optical devices: the multiresolution color filter array acquisition (MRCA), an original compressed acquisition device based on color filter arrays (CFAs) capable of capturing multi-modal remotely sensed data at different resolutions [187], and the image spectrometer on chip (ImSPOC) [104], a snapshot image spectrometer based on Fabry-Pérot (FP) interferometry [124].

The MRCA is a proposed novel approach for a compressed acquisition system of images at different resolutions; the measured samples constitute a partial representation of the scene, which demands both a superresolution and demosaicing to reconstruct the quantities of interest, which we address through variational techniques. The target of this thesis is a proof of concept of the feasibility of the novel design, for which we set up a very versatile mathematical model, which jointly deals with the problem of fusion and reconstruction, which is applicable to any filter array pattern over the focal plane. This joint inversion algorithm explores state-of-the-art inversion techniques based on total variation (TV) regularization, and its performances are compared to the results of classic fusion and demosaicing algorithms, applied separately and in cascade to one another.

For the ImSPOC project, which relies on real prototypes, the target is to provide a solid mathematical description of the physical operations performed by the device and set up a solid formulation of the inversion problem based on it. The pipeline of the data processing is then described in detail, providing some baseline results for each of the operations that characterize the device. For this device, the description of the model can be fine-tuned by analyzing real acquisitions in well-known setups; the inversion is approached with different techniques, favoring a low computational complexity as they should eventually be run on board of embedded platforms.

1.2 Context

This thesis is aimed at providing novel algorithmic approaches for the treatment of acquisitions taken with nonconventional cameras. Their design, in comparison to standard devices (e.g. RGB cameras), is aimed at surpassing the current technological limitations, in terms of either resolution, size, cost, or storage power.

The scope of this work is aimed at modeling prototypes for compact, low-cost nonconventional hyperspectral (HS) imaging devices, and at reconstructing the desired products from their raw acquisitions.

In this chapter, we provide a quick introduction to the related fields of applicability of miniaturized snapshot devices, able to be embarked on low-cost satellites or airborne vehicles. We also describe the characteristics of HS images in Section 1.2.1, presenting the related applications and acquisition methods. **Computational imaging** involves a series of conceptual and practical optical devices designs that aim to indirectly measure those quantities, addressing their reconstruction with an increased computational effort. This is the topic of Section 1.2.2. Finally, a brief very high level description and classification of the mathematical models which are employed in this work are the topic of Section 1.2.3.

1.2.1 Hyperspectral imaging (HSI)

Image representation

In the context of data processing, a natural image can be seen as an ordered collection of intensity values, which contains both some spatial information, e.g. the position of a determined object on the scene, and some spectral information, e.g. its color component. The amount N_p of bits used to describe these intensity levels is known as **bit depth** (or sometimes **radiometric resolution**, since it is typically related to the sensitivity of the radiance), so that each pixel is described by an integer in the range $\left[0, \dots, 2^{N_p} - 1\right]$.

In its **natural representation**, a digital image is represented by a 3-way array, denoted in this thesis by the tensor notation $\boldsymbol{\mathcal{U}}$. A letter contained as superscript in square brackets specifies which kind of image we are considering.

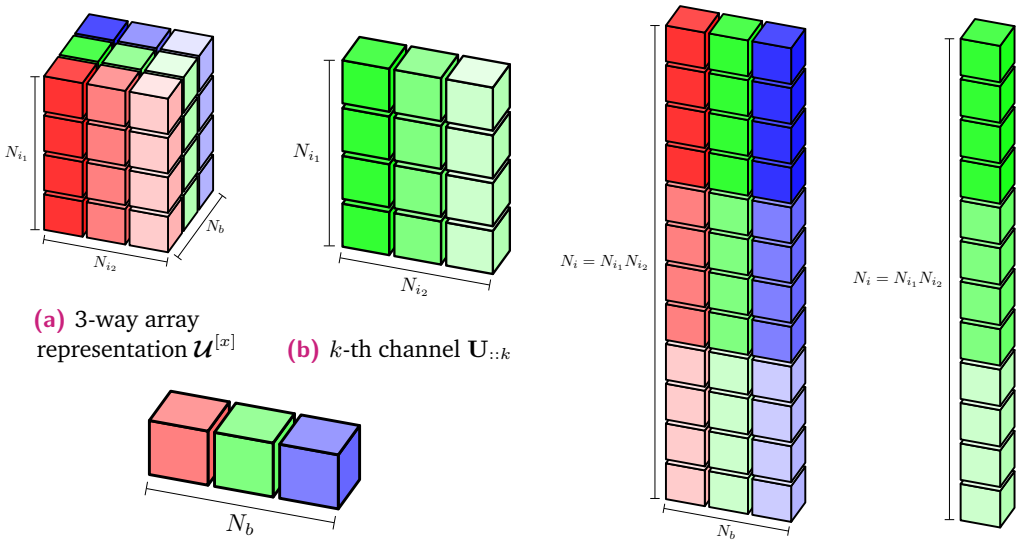
I.e., the label x is used in this thesis to define a desired image product and the associated tensor is denoted by $\boldsymbol{\mathcal{U}}^{[x]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_b}$, where N_{i_1} and N_{i_2} are the amount of column and row pixels, respectively, and N_b is the amount of **channels** or **bands**.

A visual representation of $\boldsymbol{\mathcal{U}}^{[x]}$ is given in Fig. 1.1a. Its bands are color coded as red green blue (RGB) channels, although, depending on the technology, the spectral information may be also associated with other wavelengths, e.g. near infrared (NIR) or ultraviolet (UV).

In Chapter 3 and 4, we also employ a different representation of the image, which we define as **lexicographic order**. In this representation, the image is reshaped as a 2-dimensional matrix, so that the first and second dimensions represent the spatial and spectral information, respectively. This is denoted by a bold uppercase letter, corresponding to associated label. E.g., $\boldsymbol{\mathcal{U}}^{[x]}$ can be rewritten in lexicographic order as $\mathbf{X} \in \mathbb{R}^{N_i \times N_b}$ where $N_i = N_{i_1} N_{i_2}$, as shown in Fig. 1.1d. This reshaping operation is denoted by $\mathbf{X} = \text{matr}\left(\boldsymbol{\mathcal{U}}^{[x]}\right)$.

With this formalism, the k-th column $\mathbf{x}_{:k}$ of $\mathbf{X}$ contains the information associated with the k-th channel (Fig. 1.1e), while the i-th row is the spectrum associated with the i-th pixel (Fig. 1.1c).

In Chapter 6, the 4-dimensional array $\boldsymbol{\mathcal{U}}^{[x]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_b \times N_a}$ denotes a set of N_a images captured at different times. In that context, it is convenient to define a permuted lexicographic order $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{N_b \times N_a \times N_i}$. With this representation, the i-th frontal slice $\mathbf{X}_{::i}$ is a list of N_a spectra, ordered along the columns and relative to the i-th pixel. This operation is denoted by $\boldsymbol{\mathcal{X}} = \text{reshape}(\boldsymbol{\mathcal{U}}^{[x]})$. More details are given in Section 6.1.1.



(a) 3-way array representation $\boldsymbol{\mathcal{U}}^{[x]}$

(b) k-th channel $\mathbf{U}_{::k}$

(c) Lexicographic i-th spectrum $\mathbf{x}_{i:}$

(d) Lexicographic order $\mathbf{X} = \text{matr}\left(\boldsymbol{\mathcal{U}}^{[x]}\right)$

(e) Lexicographic k-th channel $\mathbf{x}_{:k}$

Fig. 1.1. Different representations of an image with N_{i_1} column and N_{i_2} row pixels, for which N_b channels are available. Each channel is represented with a different color hue. Other then the natural representation (Fig. 1.1a), we also give its representation in lexicographic order (Fig. 1.1d). The remaining figures show various slicing operations.

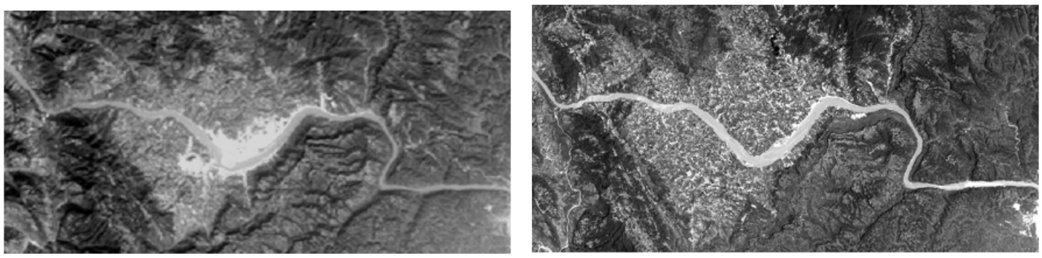
Spectral and spatial resolution

For the qualitative assessment of image products, especially in the field of remote sensing, it is a common practice to classify them in terms of their resolution in the spatial and spectral domains.

Formally, we define:

- **Spatial resolution**: The ability of the imaging system to separate objects spatially adjacent in the target scene and resolve them as distinct. This characteristic usually depends on the spatial sampling rate (e.g., the number of pixels), the type and quality of the optical system. Some examples of factor that may factor in the quality of the systems include its magnification power and its point spread function (PSF), which describes the response of the imaging system to a point source.
- **Spectral resolution** or **spectral diversity**: The ability to resolve different spectra as distinct. This depends on the number of bands and their full width at half maximum (FWHM).

A visual representation of these definitions are provided in Fig. 1.2 and 1.3. An additional property, **radiometric resolution**, denotes the available amount of intensity levels per pixel, and is expressed in units of bits. The described properties allow to link the image to the characteristics of the scene.

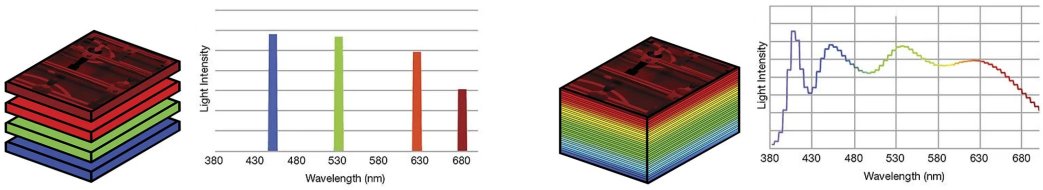


(a) Low spatial resolution **(b)** High spatial resolution

Source: [216]

Fig. 1.2. Visual representation of images acquired with different spatial resolutions. In the high spatial resolution acquisition, it is possible to distinguish more details on the scene.

The fusion problem, which will be detailed in sec. 3.5, is a good example to show the utility of these concepts. This problem consists in combining multimodal data with different resolutions in order to generate a single synthetic image with the best spectral and spatial resolution available.



(a) 4-band datacube **(b)** Low resolution Spectrum **(c)** Hypercube **(d)** High resolution spectrum

Source: Smart Vision Lights [@11]

Fig. 1.3. Visual representation of acquisitions with different spectral resolution. The depicted spectra refer to a single pixel of their respective datacube.

Acquisition technologies

We define as **hypercube** a data collection of both spectral and spatial information related to the scene under target; the capture of a hypercube is the main task of HS imaging devices.

It is possible to distinguish between three different acquisition techniques [28] for the scanning of the scene:

- **Spatial scanning**: which is performed on a pixel-by-pixel basis. This configuration is typical of spectrometers whose field of view is limited to a fixed direction, which are specialized in resolving very finely sampled spectra and are consequently able to provide very high spectral resolution. To acquire a full hypercube, the target image is scanned sequentially by varying the orientation of the instrument with respect to the scene, which results in relatively slow acquisition times. Different technologies may be employed to distinguish the spectral components of the scene, such as: 1) Elements of spectral dispersion (e.g. prisms), 2) Static filters on the detectors, 3) Linear variable filters.
- **Spectral scanning**: which is specialized in resolving very fine spatial detail on the scene. The 2-dimensional scene is recorded at a fixed angle of view and then scanned sequentially with filters with different bandwidths. The amount of acquisitions defines the spectral diversity that the user wants to achieve; this generally results in better performances than equivalent larger, more expensive and more complex instantaneous detectors, but it is really dependant on changes in the scene. These devices typically exploit one of the following two technologies: 1) interferometry (Fig. 1.4a), 2) tunable filters. In terms of spectral resolution, better performance is typically achieved in interferometry-based devices compared to tunable filters.

- **Snapshot or instant imaging**: They involve a simultaneous acquisition of both the spatial and spectral component of the hypercube. In this configuration, there is no need for any scanning, which greatly enhances the speed of acquisition and reduces the temporal sensitivity [107, 108]. On the other hand, the quality of the image is generally degraded and the amount of available bands is limited to around 30, with the complexity of the system increasing with the number of desired channels. Two of the most widespread technologies for snapshot systems are: a) based on dispersive elements or b) based on a mosaic of filters (such as multispectral filter array (MSFA)) with different spectral responses for each detector. The latter principle is commonly used for RGB cameras (Fig. 1.4b).

In the context of satellite platforms, where the flight direction is defined as **along-track direction** and the perpendicular direction as **across-track direction**, two main scanning techniques are available.

In the **whisk broom scanning**, the detection system focuses on a subsection of the swath width, so that its whole length is sequentially scanned in the across-track direction, while in the **push broom scanning**, each line of sensor statically captures the whole swath sequentially as it moves in the along-track direction.

Applications

HS imaging is a field of research that studies the techniques aimed at capturing a scene with the highest spectral diversity. The acquisitions typically involve a few dozens to hundred channels, each providing information over a different set of wavelengths. They typically cover the domain of visible (VIS), NIR, infrared (IR) or UV, with either narrow or wide bandwidths, depending on the application [42, 132]. The availability of such diversity, compared to classical RGB cameras [146], allows to discriminate the chemical composition of the components of the scene through its spectral signature, to either discriminate or classify them. [132]

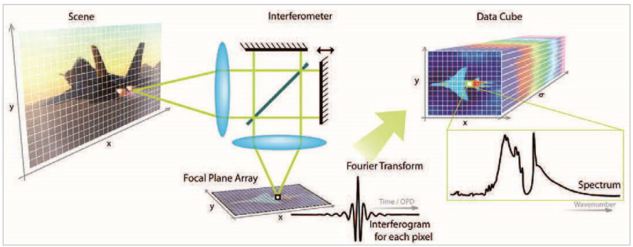
Remote sensing makes vast use of this kind of imagery; the currently commercially available high performance HS cameras are generally bulky (more than 15 kg) and expensive (from 50 to 500 thousands of euros for Telops Hypercam cameras) [28].

The spread of such systems are thus still limited in many real life applications, especially if the aim is to embark them over airplanes/drones. This situation has stemmed, in the last 20 years, an increased interest in developing HS detectors which employ nonconventional technologies [194, 208], so that a wide range of

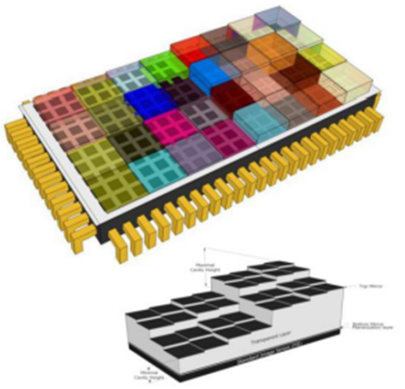
alternatives is currently available, with different specifics with regard to the desired spectral, spatial or temporal resolution.

HS imaging enables a wide variety of applications [132] and its scope spans across multiple domains which include:

- **Industry**: Their main application is for non-invasive monitoring of products, such as in the case of automated waste recycling [207, 151]. Additionally, the technology has been efficiently employed for agricultural [74, 37] and pharmaceutical quality control [199]
- **Agriculture and environment**: Historically, these domains were the first to adopt the advantages of HS imagery. Miniaturized devices would allow to both enhance the precision of the measurements and to allow for multi-view acquisitions through the employment of drones. These setups can be employed for **environment monitoring** [2, 23], **gas detection** [135] or **cultivation monitoring** [220]. The detection of pollution (e.g. asbestos in the domain of the IR [27])) is also a prominent field of application. The reader can also consult the related works [17, 16] for a review on the most recent applications in remote sensing.
- **Health**: HS imagery has also emerged in the medical domain for the diagnosis and monitoring of pathologies [37, 138], other than as support for surgery [28].



Source: ABB Ltd [@1]



Source: IMEC [@6]

(a) Schematized principle of the Fourier transform spectrometer (FTS)

(b) MSFA pattern over a matrix of detectors

Fig. 1.4. Two different technologies for snapshot acquisitions, requiring ad-hoc algorithms to recover user-ready HS products [28].

1.2.2 Computational imaging

In recent years, the increasing availability of computing power and the recent advances in terms of algorithms opened novel possibilities for the design of more sophisticated optical systems, as computational-intensive methods could be exploited for the recovery of the images of interest.

As imaging acquisition systems are physically limited by the current rate of traditional technological growth, computational imaging aims at overcoming those limitations for increased precision, focusing on cutting-edge data processing protocols of measurements taken with nonconventional approaches. As some physical phenomena are inherently coded, their raw acquisitions, which are theoretically able to provide better performances in terms of either SNR, resolution or some other practical advantage (cost, weight, etc.), require a post-processing to transfer them to the desired domain.

This transfer of domain can sometimes be also exploited to retrieve information from the measurements that would otherwise be unavailable or by obtaining higher performances on the final acquisition (e.g. better resolution or lower SNR).

Computational imaging techniques are constantly gaining more attention in the field of consumer cameras, cell phone cameras, vehicle camera systems, surveillance, medical imaging, remote sensing, human computer interaction [154]. In remote sensing in particular, the synthetic aperture radar (SAR) systems are a prominent example, as the synthesis of the antennas of the arrays is performed with computational techniques [81].

A very closely related field, known as **optical co-design** (or sometimes with its French definition "co-conception") has the additional goal to design both the optical components and the data processing simultaneously. The mutual interaction across the two point of view can at the same time simplify the inversion algorithms and compensate eventual flaws due to the intrinsic optical properties.

We do not aim in this section to provide a detailed list of researches and works related to computational imaging, which spans across various fields; the associated approaches also vary wildy in relation to the given capturing device. For a review of some of the applicable fields of computational imaging, which include pinhole cameras and more, one can check the work of Mait et al. [154]; we just list here a couple of examples that are relevant to the devices that are investigated in this work:

- **Compressed sensing**: According to the theory of compressed sensing imaging [69], it is possible to capture all the relevant desired information from sparse signals, reducing the amount of necessary measurements which would be usually needed in traditional cameras. In this context, the desired product is only available through partial measurements, according to a defined encoding system, requiring an ad hoc computational analysis to infer the missing information. Various encoding approaches have been proposed, so that the compressive measurements contains the least redundancy with respect to the full spectral and spatial content of the scene from the acquisition. I.e., one of such compressed acquisition systems is the compressive coded aperture spectral imaging (CASSI) [14], a device employing **coded aperture** spatial encoding and spectral dispersion to encode the information over a shared focal plane.

- **Fourier transform spectrometers (FTSs)**: The FTS describes a group of techniques aimed at the estimation of spectra through the detection of radiation of coherent sources. This class includes instruments for optical spectroscopy, IR spectroscopy (widespread in chemistry and known as Fourier transform infrared spectroscopy (FTIR)), nuclear and magnetic resonances, mass spectrometers [186]. Perhaps the most known example of such devices is the **Michelson interferometer**, a device able to split the incident light into two interfering beams and combine them over a common detector; the acquisition obtained by varying the path difference between the two beams, known as interferogram, can be interpreted, under particular constraints, as the Fourier transform (FT) of the spectrum, hence the name. The processing of the acquisitions is particularly well suited for computational imaging strategies, as a deeper knowledge of the model allows a more precise reconstruction of the spectrum and the image.

1.2.3 Mathematical modeling of the imaging process

The main approach that we employ for the analysis of nonconventional devices consists of constructing a physical model of the optical system, able to link the quantities of interest to the readout of the instrument under test.

The final goal is to reconstruct a digital representation of spectra, with a user-specified target spectral resolution. Each spectrum may be associated their direction of incidence of the incoming radiance, forming a multiband image that provides the

full characterization of the scene. For the sake of exposition, the spectra in this set are arranged over columns of a matrix $\mathbf{X}$.

In a snapshot-type acquisition, the observation is given by a set of measurements of a group of sensors, typically determined by the photo-detectors over a focal plane array (FPA), which we can temporarily denote with a column vector $\mathbf{y}$. The snapshot characteristic of the investigated devices implies that all the information about the quantities to reconstruct are obtained at a given instant, so there is no need to consider any acquisition at different time.

In general terms, we can now describe the optical transformation that characterizes the device with a general purpose relationship, such as $\mathbf{y} = \mathbb{A}(\mathbf{X})$. No particular constraint is imposed at this point on the structure of $\mathbb{A}$, other than describing the imaging process.

In this framework, we are interested in three different scenarios:

- **Simulation** ($\mathbf{y}^{[sim]} = \mathbb{A}(\mathbf{X})$): In the design phase for a certain novel technology, before the prototypes are available or if they are not fully finalized, it is interesting to verify the viability of given components, parameters and operating conditions. For this purpose, it is interesting to generate some simulated responses $\mathbf{y}^{[sim]}$, based on a given model of the acquisition system $\mathbb{A}$ and a set of known realistic inputs $\mathbf{X}$.

- **Model characterization** ($\mathbf{y} = \mathbb{A}^{[\widehat{\beta}]}(\mathbf{X})$): Its target is to estimate the properties of the acquisition system $\mathbb{A}(\cdot) \equiv \mathbb{A}^{[\beta]}(\cdot)$ under test. The latter is described as function of a discrete set of parameters β, of which we aim to find an estimation $\widehat{\beta}$, given a set of example pairs of $\mathbf{X}$ and $\mathbf{y}$. I.e., if $\mathbb{A}$ is described by a linear operator, the parameters to infer may be the coefficients of the characteristic matrix.

  In the context of machine learning this problem is commonly known as **supervised learning**, yet considering the small cardinality of available experimental data, this interpretation will be considered out of the scope of this thesis. Instead, the main approach is based on the so-called **geometrical optics** [202]. With this framework, the light rays within the optical system are assumed to propagate in straight lines in homogeneous media, while bending and splitting occurs only at the interface between dissimilar media. The energetic balance between the incident and the detected rays allows to build a theoretical model which links the readout of the instrument to the directional input spectrum that we want to reconstruct.

- **Inversion** ($\widehat{\mathbf{X}}$ s.t. $\mathbf{y} \approx \mathbb{A}(\widehat{\mathbf{X}})$): If $\mathbf{y}$ and $\mathbb{A}$ are given, the problem reduces to estimate the spectra from the measurements. This problem, if approached as a mathematical inversion of the deterministic part of the model, is generally ill-posed. In Hadamard's sense [106], this means the reconstruction is either not unique or does not vary continuously with the observation. This demands to impose some regularization, or in other words, to impose some prior information on the solution to compensate for such drawbacks.

When we analyse a novel acquisition prototype, where no or very few real data are available, the robustness of the inversion protocol may only be tested by imposing a model for the expected behaviour of the system and simulating an acquisition, given a series of realistic inputs. As soon as a prototype is available, the mismatches with the real behaviour of the system can then be tested with an ad hoc calibration to refine the original model; this procedure can be iterated until a certain targeted accuracy is reached.

1.3 Investigated devices

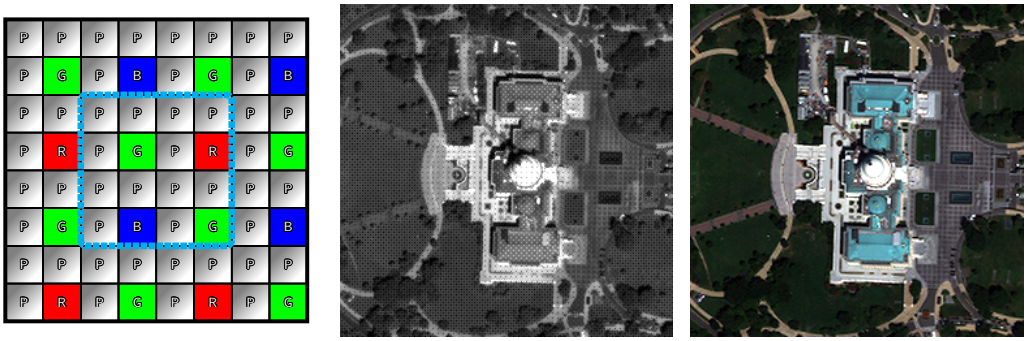
The proposed approach for the inversion of acquisitions to reconstruct a HS virtual image was applied to two nonconventional design concept, which we denote with the acronyms MRCA and ImSPOC. This section provides a brief overview of these two optical devices, describing the specific challenges which are associated with the data processing of their raw acquisitions.

1.3.1 Multiresolution color filter array acquisition (MRCA)

The MRCA defines an acquisition strategy based on compressed acquisitions of multiresolution remotely sensed data. The envisioned design is defined by an array of mosaiced sensors with different characteristics, following the same principle of CFA cameras.

More specifically, the focal plane is composed of a mosaic of sensors with different spatial and spectral characteristics, disposed in a user-defined pattern (Fig. 1.5a); such sensors provide a partial knowledge of the characteristic of the scene, that the data processing is in charge to extend to the whole scene [187]. An example of such acquisition is shown in Fig. 1.5b, which shows a monochromatic product that we aim to process to ideally recover the information of the scene of Fig. 1.5c.

Two main challenges are associated with this novel concept. Firstly, we need to setup a framework to recover the desired products that takes into account the correlation between the information provided by each of the readouts. For the case of sensors with the same characteristics, this problem is known in the literature as **demosaicking** [145, 134]. The case under study additionally involves the fusion of information with different characteristics; if two separate acquisitions were available on the whole scene, this problem is commonly known in the literature as **pansharpening** [224, 147], which is a particular instance of data fusion. An additional challenge is given by the choice of the distribution of the sensors over the focal plane. This pattern has to provide the least amount of redundant information, to facilitate the process of inversion.



(a) Example sensor pattern **(b)** MRCA compressed acquisition **(c)** Ideal reconstruction product

Fig. 1.5. Visual representation of the MRCA concept. In Fig. 1.5a, an example disposition of the sensors over the focal plane, with the P label referring to wideband monochromatic sensors, and R,G, and B referring to CFA with a RGB spectral response. The compressed acquisition (Fig. 1.5b) can be seen as a degraded resolution mosaiced version of the ideal acquisition to reconstruct (Fig. 1.5c).

1.3.2 Image spectrometer on chip (ImSPOC)

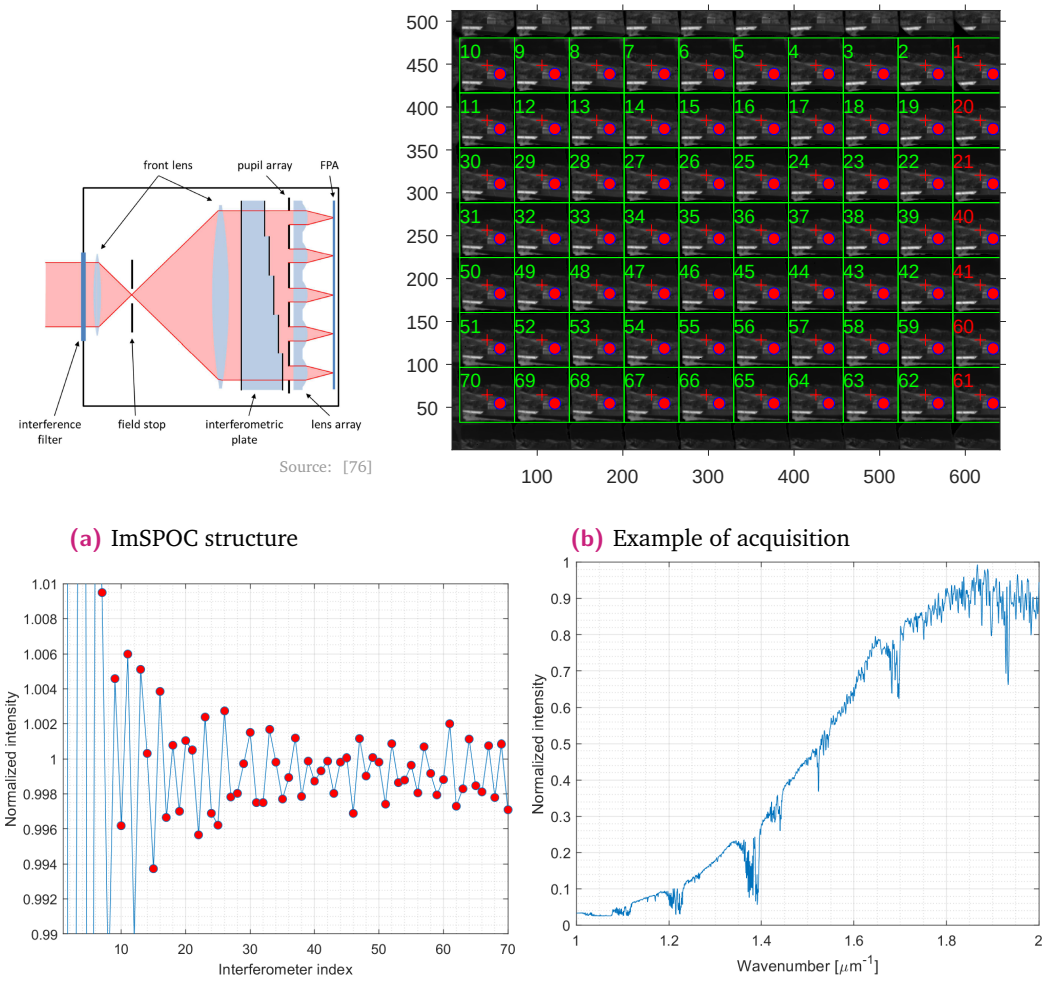
The concept of ImSPOC [104] defines a snapshot imaging spectrometer based on the interferometry of Fabry-Pérot (FP). A set of different FP etalons with different thickness is disposed as a matrix in a staircase shaped structure. This structure allows to capture multiple simultaneous acquisitions at once by focusing the response of each interferometer on separated portions of the FPA. This separation is obtained through an array of small lenses, one for each interferometer, which leads to a division of aperture. A possible implementation of this design is shown in Fig. 1.6a. The different replicas of the scene are commonly defined as **subimages** (or, equivalently, as thumbnails).

If we fix the angle of incidence, it is possible to identify an associated set of readings across the different subimages (Fig. 1.6b), which is a sampled version of a continuous interferogram (Fig. 1.6c). According to the principle of FTSs, the interferogram can be inverted to reconstruct the associated spectrum (Fig. 1.6d).

This project is currently in its prototyping stage, and a series of test devices based on this concept were recently made available for testing.

Various challenges are associated with the reconstruction of the input spectra:

- It is necessary to rigorously define a mathematical model that describes the optical transformations operated by the device, attempting to include all the possible causes of nonideality, to increase the precision of the reconstruction;
- The different subimages are usually not perfectly co-registered, so a proper procedure to align them is required to have consistent samples across different representations of the scene;
- The expression of the transfer function which describes the system cannot be based just on the physical structure of the system under test, as there is a mismatch between the design and the manufacture stage. This mismatch has to be properly compensated in the characterization stage;
- The inversion protocols have to be robust enough to adapt to deviations of the real behaviour of the system from our calibrated transfer model.



(a) ImSPOC structure

(b) Example of acquisition

(c) Extracted interferogram

(d) Ideal spectrum to reconstruct

Fig. 1.6. Visual representation of the ImSPOC acquisition principle. In Fig. 1.6a, an example for a prototype structure of a device based on the ImSPOC concept, together with an associated acquisition (Fig. 1.6b). The acquisition shows the different subimages associated with the same scene, framed within green rectangles. If we isolate the red spots related to a given angle of incidence and arrange them as a sequential array, we obtain a sampled version of the interferogram (Fig. 1.6c). This graph can be inverted to obtain the desired input spectrum associated with the incoming radiance (Fig. 1.6d)). The spectrum is expressed in terms of wavenumbers (reciprocal of wavelengths), while the interferogram in terms of optical path difference (OPD), which is proportional to the interferometers' thicknesses.

1.4 Manuscript structure

Each chapter of this manuscript is self contained, and can be read independently from the other ones.

Chapter 2 is a brief introduction of the inversion framework employed in this work, Chapter 3 describes a set of operations to deal with remotely sensed images with different resolutions.

Chapter 4 introduces the novel concept of MRCA, a compressed acquisition system based on CFAs that simultaneously captures scene samples with different resolution, including a detailed analysis of the joint demosaic and fusion of the raw data, based on TV regularization, aimed at generating a synthetic image that simultaneously features high spatial and spectral resolution. Chapter 5 provides some background on optical concepts to develop an encompassing model that describes the operating principle of ImSPOC. Chapter 6 is devoted to the pipeline of processes for the treatment of raw acquisitions with the ImSPOC device, involving the registration of subimages, the inference of the parameter of the transfer model and some preliminary results for the interferogram inversion.

The manuscript can also be read from a thematic point of view. If the reader is more interested in inversion problems, chapters 2, 4, and 6 provide a theoretical and practical overview of their application in image processing. If the reader is instead more interested in how to model an optical system, chapters 4 and 5 offer an in depth explanation on how to describe the optical transformation of the two prototypes under test. Finally, if the reader wants an analysis from the point of view of the fusion of multimodal data, chapters 3 and 4 provide an overview of the standard techniques, together with an application to the MRCA concept.

1.5 Scientific contributions

1.5.1 Original contributions of this PhD

Given the multidisciplinary nature of this thesis, it was decided to provide a sufficiently exhaustive introduction of the concepts which are discussed. Nonetheless, for the reader's convenience, we deem appropriate to provide a list the original contributions of this manuscript. The section where those original contributions are detailed is also indicated below.

The original contribution related to the MRCA include:

- the MRCA concept itself, which describes a novel compression acquisition for multiresolution sensors, inspired by the CFA [176] (Section 4.2);
- a Bayesian formulation of the posteriori framework for the inversion of said compressed acquisitions, which jointly addresses the problem of fusion and demosaicing of multiresolution data by exploiting state-of-the-art variational techniques (Section 4.4);
- a comparison of the quality of the products generated by a selection of software image compression algorithms and of the reconstructed acquisition of coded aperture based compressed acquisition imaging systems, such as the CASSI and the MRCA (Section 4.6.3);
- a comparison between the reconstruction results of our proposed joint approach and of an approach based on classic fusion and demosaicing techniques, applied separately in cascade (Section 4.6.4);
- a preliminary analysis of the combined filter array patterns (Section 4.3) developed ad hoc for the proposed device, analyzing the effectiveness of periodic and pseudo-random designs (Section 4.6.6);
- preliminary considerations on the description of a selection of demosaic algorithms with the framework of classic pansharpening techniques (Section 3.7).

The original contributions linked to the ImSPOC concept include:

- a detailed physical model that describes the optical transformations of the device, based on Airy's distribution (Section 5.5);
- three novel algorithms for the estimation of the center points of subimages over a focal plane, applicable both to point and extended sources, based on mathematical morphology and nonlinear regression (Section 6.2);
- a proposed procedure for the coregistration of the sub-images, based on a point mapping protocol. The procedure is based on a point mapping calibration, measuring the shifts across different sub-images, and applying a polynomial geometric transformation and a regular grid resampling to synchronize their geometry with respect to a reference one (Section 6.3);
- the definition of three different novel approaches for the characterization of the parameters of the acquisition system, based on maximum likelihood, exhaustive search and a nonlinear regression, comparing the accuracy of the estimated parameters in the case of real calibration datasets taken with 3 different prototypes (Section 6.4);
- a comparison of single pixel inversion techniques based on singular value decomposition (SVD) and least absolute shrinkage and selection operator

(LASSO) frameworks for the reconstruction of interferograms generated by solar spectra, with an analysis of the degradation of the products in the case of parametric mismatches between the acquisition system and its model (Section 6.5).

For all the contributions, we developed a toolbox of algorithms, which we plan to release to the scientific community, in order to support the reproducibility of the results and any further research development.

1.5.2 Publications

Part of the content of this PhD manuscript has been presented in the following publications:

- "Pansharpening of images acquired with color filter arrays" [189];
- "Image Fusion and Reconstruction of Compressed Data: a Joint Approach" [187];
- "Analysis of masks for compressed acquisitions in variational-based pansharpening" [188];
- "Gas characterization based on a snapshot interferometric imaging spectrometer" [59];
- "Characterization of a Snapshot Fourier Transform Imaging Spectrometer Based on an Array of Fabry-Perot Interferometers" [190].

1.5.3 Other contributions

Some contributions developed during the course of this PhD were not included in this manuscript, as they slightly deviate from the topics that are discussed in this dissertation. Those include:

- a collaboration with doctor Bouthayna Msellmi, of the University of Manouba, on the topic of **sub-pixel mapping**. The published works, related to this subject, propose to solve the problem with two novel regularization approaches based on the isotropic total variation (ITV) [173] and on sparse dictionary decomposition [174];
- an ongoing collaboration with doctor Aneline Dolet and professor Didier Voisin, of the Université de Grenoble Alpes, France, in the context of project ImSPOC-ultraviolet (ImSPOC-UV) (Appendix A.2), employing a different strategy to

process the data acquired with the ImSPOC devices, based on the differential optical absorption spectroscopy (DOAS) [60, 61];

- an ongoing collaboration with doctor El Mehdi Abdali and professor Stéphane Mancini of the Université de Grenoble Alpes, France, also in the context of the ImSPOC-UV, aimed at the implementation of the proposed inversion algorithms on a field programmable gate arrays (FPGAs);
- a collaboration in a Precursory Research for Embryonic Science and Technology (PRESTO) project [@14], directed by professor Kuniaki Uto of the Tokyo Institute of Technology, aimed at the estimation of crop vitality through the analysis of aerial acquisitions. This project led to a 3-month visit to Japan, where I participated in campaigns for in situ acquisitions of images for precision agriculture with unmanned aerial vehicles (UAVs).

2 Inverse problems theory

Signal reconstruction constitutes one of the most challenging problems related to data processing for computational imaging systems.

As the operating principle of nonconventional imaging devices can be seen as a set of optical transformations up to the final measurements, those may in fact generally span in a domain other than the physical quantity of interest.

Mathematically, this process of estimating the quantities that act as causal factor for the observations is formally known as an **inverse problem** [26], which can be addressed as a discrete **optimization problem**, a field of research that investigates methods for the choice of the best solution among all the feasible ones [29].

An inverse problem is typically solved by imposing an objective function to minimize, which, in the formulation of Tikhonov [170], can be separated into a weighted linear combination of two metrics: a data fidelity term and a nonnegative functional that acts as regularization. These quantities can also be respectively interpreted, in the framework of a Bayesian inference, as a maximum likelihood (ML) estimation and a nonnegative functional that models the prior information, which in this thesis is imposed through both sparsity-inducing and variational regularizers.

This chapter provides a brief introduction to the abstract concepts of inversion problems, decoupled from any application. In particular, Section 2.1 introduces the general framework of inverse problems, Section 2.2 describes some regularization approaches that address the ill-posedness of the problem and Section 2.3 briefly describes some widespread algorithms for their solution.

2.1 The inversion framework

In general terms, an inverse problem consists in estimating a set of quantities of interest that have a causal relationship with given observations [26, 112]. The term inversion stems from the intention to produce the inverse result of a given forward model of a physical phenomenon, which in our context defines the acquisition system.

This forward model is in general a stochastic process, where the input variables to infer are realizations of a random variable (r.v.). These variables are indirectly measured through a series of observations and physically obtained by a series of complex transformations in a typically noisy environment.

The transformations are partially known to the user through a direct model, whose expression may be obtained by describing the underlying physical phenomena, either through a data-assisted calibration or with data-driven models (such as deep neural networks) [15].

More formally, in its standard stochastic formulation, the goal of an inverse problem is to obtain an estimation $\widehat{\mathbf{x}} \in \mathtt{E}_x$ of the desired quantities $\mathbf{x}$, captured through a set of indirectly measured data $\mathbf{y} \in \mathtt{E}_y$, modeled as:

$$\mathbf{y} = \mathbb{A}(\mathbf{x}) + \mathbf{e}\,. \tag{2.1.1}$$

Here, $\mathbf{e}$ is a particular realization of additive noise, $\mathtt{E}_x$ and $\mathtt{E}_y$ are respectively the desired product and observation spaces. $\mathbb{A}(\cdot) : \mathtt{E}_x \to \mathtt{E}_y$ is an operator that describes a direct model, which we assumed to be known. This relationship is also summarized in Fig. 2.1.

Traditionally, inverse problems can be analyzed both with a deterministic and statistical interpretation and many of the results of each of the two framework can find a correspondence in the other domain. Section 2.1.1 and 2.1.2 provide a brief introduction to these two interpretations.

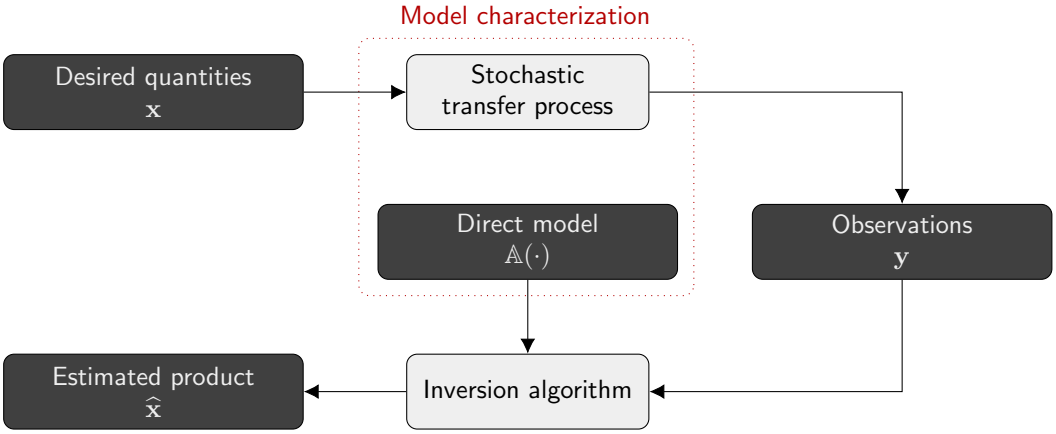


Fig. 2.1. Block diagram representation of a direct and an inverse problem. In the forward sense, an unknown realization of the input is available to the user as a series of indirect measurements. In the inverse sense, the input data has to be inferred from such observations, given a deterministic model that describes the direct transformation.

2.1.1 Ill-posed problems

To avoid ambiguities with the stochastic model, we denote with $\mathbf{y}^{[sim]}$ a "simulated" observation, obtained as a deterministic component of eq. (2.1.1):

$$\mathbf{y}^{[sim]} = \mathbb{A}(\mathbf{x})\,. \tag{2.1.2}$$

The problem of generating a vector $\mathbf{y}^{[sim]}$ that most closely matches the true instrument acquisitions, given a known input $\mathbf{x}$, is commonly known as **direct problem** or **forward problem**.

Under the formalism proposed by Hadamard [106], an **inversion operator** $G_{\mathbb{A}} : \mathtt{E}_G \subseteq \mathtt{E}_y \rightarrow \mathtt{E}_x$ is defined as a series of procedures whose goal is to find an estimation of the input $\mathbf{x}$, given $\mathbf{y}^{[sim]}$, whose respective domains $\mathtt{E}_x$ and $\mathtt{E}_y$ are generic Banach spaces with norms $\|\cdot\|_x$ and $\|\cdot\|_y$, respectively.

$G_{\mathbb{A}}$ specifically defines a **well-posed problem** if:

- a solution exists;
- the solution is unique;
- its solution is stable to perturbations of $\mathbf{y}^{[sim]}$, i.e. $G_{\mathbb{A}}$ is defined on all $\mathtt{E}_y$ and is continuous.

More specifically, if the problem is well-posed, let the observation $\mathbf{y}_{\check{\delta}}$ be known within a given error [1] $\check{\delta}$, or in other words that its distance $\|\mathbf{y}_{\check{\delta}} - \mathbf{y}\|_y \leq \check{\delta}$, then the candidate solution $G_{\mathbb{A}}(\mathbf{y}_{\check{\delta}})$ exists and verifies the condition:

$$\lim_{\check{\delta}\rightarrow 0} \|G_{\mathbb{A}}(\mathbf{y}_{\check{\delta}}) - \mathbf{x}\|_x = 0\,. \tag{2.1.3}$$

Conversely, if any of Hadamard conditions is not verified, the problem is **ill-posed**. This is the case of many real life problems, as the direct model is typically a physical phenomenon defined in a continuous space, while the observations are limited in number.

The mathematical formalism of Tikhonov [219] shows that, if $G_{\mathbb{A}}$ is ill-posed, it is sometimes possible to generate new operators that are well-posed, which a Cauchy convergence to a mapping of $G_{\mathbb{A}}$ if the observation is not in the domain of $G_{\mathbb{A}}$.

[1] In our thesis we are concerned with both the domain of inverse problems and of optics. In the standard literature of both of these domains, it is common practice to employ Greek letters to define certain standard variables. To distinguish between the two domains, we add a $\check{}$ symbol over Greek letters to denote variables related to inverse problems, in comparison with those of the domain of optics, which are kept without. E.g., the value $\check{\lambda}$ denotes a regularization parameter in the inversion domain, while λ denotes a wavelength in the domain of optics.

In this case $G_{\mathbb{A}}$ is called **regularizable** and the novel mapping functional is called **regularizing operator**.

A well-posed problem does not however assure that the numerical implementation of the operator leads to stable solutions, as machines work with finite precision. Such instances are commonly known as **ill-conditioned problems**. Roughly speaking, in ill-conditioned problems, small variations on the observation map to very large deviations of the desired products. This effect is quantified by the **condition number**, defined as:

$$\kappa(\mathbb{A}) = \lim_{\check{\epsilon}\to 0} \sup_{\check{\delta}\leq\check{\epsilon}} \frac{\|G_{\mathbb{A}}(\mathbf{y}_{\check{\delta}}) - G_{\mathbb{A}}\left(\mathbf{y}^{[sim]}\right)\|_x}{\|\mathbf{y}_{\check{\delta}} - \mathbf{y}^{[sim]}\|_y} \tag{2.1.4}$$

The most common case study is the one in which the direct model is a linear operator represented by a multiplication by a matrix $\mathbf{A}$, which operates between two real vector spaces.

In fact, if $\mathbf{A}$ is nonsingular matrix describing a deterministic process, then the inverse matrix $\mathbf{A}^{-1}$ describes a well-posed problem. However, its condition number (2.1.4) is given [51] by the product $\kappa(\mathbf{A}) = \|\mathbf{A}\|_{op}\|\mathbf{A}^{-1}\|_{op}$, where the $\|.\|_{op}$ defines the so called **operator norm**.

For any given operator $\mathbb{A} : \mathtt{E}_x \to \mathtt{E}_y$, its operator norm is defined as the largest scalar by which $\mathbb{A}$ stretches the elements of $\mathtt{E}_x$:

$$\|\,\mathbb{A}\,\|_{op} = \inf\left\{\check{\alpha} \geq 0 : \|\,\mathbb{A}(\mathbf{x})\|_y \leq \check{\alpha}\|\mathbf{x}\|_x, \forall \mathbf{x} \in \mathtt{E}_x\right\}\,, \tag{2.1.5}$$

which in the case of a matrix multiplication the operator norm $\|\mathbf{A}\|_{op} = \check{\zeta}_{\mathbf{A}}^{max}$ is given by the largest singular value of the matrix $\mathbf{A}$ (i.e. the square root of the largest eigenvalue associated with the Gram matrix $\mathbf{A}^*\mathbf{A}$).

2.1.2 Statistical description

An alternative framework for the estimation of $\widehat{\mathbf{x}}$ requires a statistical description of the optical transformations, taking into account the sources of uncertainties in the model description and the unavoidable noise phenomena.

Two classical approaches are widespread for the statistical formulation of the problem. Following the formalism of eq. (2.1.1), the Bayesian approach attempts to

maximize the a posteriori probability of the observation, given the statistical description of both the noise $\mathbf{e}$ and the desired input $\mathbf{x}$. If $\mathbf{x}$ is considered deterministic, this results into the special case known as maximum likelihood estimation (MLE).

Maximum likelihood estimation

The ML method is based on the assumption that the every possible input $\mathbf{x}$ is "equally likely" (uniformly distributed). With this assumption, one can imagine to fix a particular input $\mathbf{x}$ and analyze the conditional statistical distribution of the observation. This allows to construct a family of possible distributions, among which we can select the most likely to have generated the measured realization $\mathbf{y}$.

Mathematically, let $\boldsymbol{\psi}$ be the r.v. associated with the observation and let the **likelihood function** $p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x})$ denote the probability mass function (pmf) of the observation conditioned to a fixed deterministic input $\mathbf{x}$.

The MLE $\widehat{\mathbf{x}}$ consists then in selecting the input that maximizes $f_L(\mathbf{x})$:

$$\widehat{\mathbf{x}} = \arg\max_{\mathbf{x}\in\mathbf{E}_x} p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x})\,. \tag{2.1.6}$$

As a main case study, let us suppose that the noise $\mathbf{e}$ is described as a multi-variate Gaussian noise with zero mean and covariance matrix $\mathbf{C}$ (its elements c_{ij} are the covariances between its i-th and j-th component). In the ML framework, the statistical representation of the r.v. ψ is as well a multi-variate Gaussian but with mean $\mathbb{A}(\mathbf{x})$; in mathematical terms:

$$p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\frac{1}{2}\left((\mathbf{y}-\mathbb{A}(\mathbf{x}))^{*}\,\mathbf{C}^{-1}(\mathbf{y}-\mathbb{A}(\mathbf{x}))\right)\right) \tag{2.1.7}$$

where $\propto$ stands for "proportional to" and $\mathbf{A}^{*}$ denotes the complex conjugate of $\mathbf{A}$.

The criterion of ML is equivalent to minimizing the so called log-likelihood $\log f_L(\mathbf{x})$:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbf{E}_x} \log p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x}) \tag{2.1.8a}$$

$$= \arg\min_{\mathbf{x}\in\mathbf{E}_x} \frac{1}{2}\left((\mathbf{y}-\mathbb{A}(\mathbf{x}))^{*}\,\mathbf{C}^{-1}(\mathbf{y}-\mathbb{A}\,(\mathbf{x}))\right)\,. \tag{2.1.8b}$$

In the case in which the noise is composed of independent and identically distributed (i.i.d.) acquisitions the covariance matrix is proportional to an identity matrix, so the criterion is equivalent to minimization of the mean square error:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbf{E}_x} \frac{1}{2}\| \mathbb{A}(\mathbf{x}) - \mathbf{y}\|_2^2 \,, \tag{2.1.9}$$

where $\|\cdot\|_2$ denotes the ℓ_2-norm.

Bayesian estimator

In the Bayesian estimation, the model of the previous section is extended to consider that the input itself is modelled as a r.v. $\boldsymbol{\chi}$, with its own characteristic distribution.

In the formulation of eq. (2.1.1), the r.v. $\boldsymbol{\psi}$ of the observation can then be expressed as a function of two r.v.:

$$\boldsymbol{\psi} = \mathbb{A}(\boldsymbol{\chi}) + \boldsymbol{\nu} \,, \tag{2.1.10}$$

where $\boldsymbol{\nu}$ is the r.v. associated with the noise.

With this framework, the stochastic description of $\boldsymbol{\psi}$ is given by a joint pmf $p_{\boldsymbol{\psi},\boldsymbol{\chi}}(\mathbf{x},\mathbf{y})$, which, according to Bayes' theorem, is a function of both the marginal $p_{\boldsymbol{\chi}}(\mathbf{x})$ and the likelihood function $p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x})$:

$$p_{\boldsymbol{\psi},\boldsymbol{\chi}}(\mathbf{x},\mathbf{y}) = p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x}) p_{\chi}(\mathbf{x}) \,. \tag{2.1.11}$$

The target of the **Bayesian estimator** is to maximize the **a posteriori probability** $p_{\boldsymbol{\chi}}(\mathbf{x}|\mathbf{y})$, which acts as a measure for how likely a certain representation of the input is, given a particular realization $\mathbf{y}$ of the output. The a posteriori probability can be expressed in terms of the **a priori** probability $p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x})$ as:

$$p_{\boldsymbol{\chi}}(\mathbf{x}|\mathbf{y}) = \frac{p_{\boldsymbol{\psi},\boldsymbol{\chi}}(\mathbf{x},\mathbf{y})}{p_{\boldsymbol{\psi}}(\mathbf{y})} = \frac{p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x}) p_{\psi}(\mathbf{x})}{p_{\boldsymbol{\psi}}(\mathbf{y})} \tag{2.1.12}$$

Consequently, given a certain realization $\mathbf{y}$ of $\boldsymbol{\psi}$, the criterion of estimation of the input becomes:

$$\widehat{\mathbf{x}} = \arg\max_{\mathbf{x}\in\mathbf{E}_x} p_{\boldsymbol{\chi}}(\mathbf{x}|\mathbf{y}) = \arg\max_{\mathbf{x}\in\mathbf{E}_x} \frac{p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x})p_{\chi}(\mathbf{x})}{p_{\boldsymbol{\psi}}(\mathbf{y})} \tag{2.1.13a}$$

$$= \arg\max_{\mathbf{x}\in\mathbf{E}_x} p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x})p_{\psi}(\mathbf{x}) \tag{2.1.13b}$$

$$= \arg\min_{\mathbf{x}\in\mathbf{E}_x} -\log p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x}) - \log p_{\boldsymbol{\chi}}(\mathbf{x}) \tag{2.1.13c}$$

It can be easily shown that, if $\boldsymbol{\chi}$ is uniformly distributed the term $\log p_{\boldsymbol{\chi}}(\mathbf{x})$ is constant and the estimator reduces to the MLE approach of the minimization of the log-likelihood $p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x})$.

The Bayesian estimator can be seen as a minimization of a function $J(\mathbf{x}) = -\log p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x}) - \log p_{\chi}(\mathbf{x})$ which is composed as the sum of the log-likelihood and a contribution given by the prior information of the input. $J(\mathbf{x})$ usually known as **objective function** or **cost function**.

In the case the noise components are i.i.d. realizations of zero mean Gaussian distributions, the objective function can be simplified as follows:

$$J(\mathbf{x}) = \frac{1}{2}\| \mathbb{A}(\mathbf{x}) - \mathbf{y}\|_2^2 + \check{\lambda} g'(\mathbf{x}) \tag{2.1.14}$$

where the **regularization function** or **penalization function** $g'(\mathbf{x})$ can be seen as a Bayesian interpretation of the regularizing operation popularized by Tikhonov's formalism [219] to impose a well-posedness condition over a standard mean square error (MSE).

2.2 Regularization approaches

The role of the regularization is to modify an objective function associated with an ill-posed problem to enforce the uniqueness and continuity of a solution. This section provides a brief introduction to the regularization techniques that will be employed in this thesis, which include approaches based on the penalized matrix decomposition (PMD) (Section 2.2.1), sparsity-inducing operators (Section 2.2.2), and variational methods (Section 2.2.3).

2.2.1 Penalized matrix decomposition

In this section we assume that the linear operator $\mathbb{A}(\cdot)$ that describes the direct model is a linear operator represented by a multiplication by a matrix $\mathbf{A} \in \mathbb{R}^{N_y \times N_x}$.

We are concerned in this section with methods that make use of the singular value decomposition (SVD). For a given matrix $\mathbf{A} \in \mathbb{R}^{N_y \times N_x}$, the SVD of $\mathbf{A}$ is given by:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^* \,. \tag{2.2.1}$$

In its economized (also known as compact) form, the matrices $\mathbf{U} \in \mathbb{R}^{N_y \times N_r}$ and $\mathbf{V} \in \mathbb{R}^{N_x \times N_r}$ are semi-unitary and $\mathbf{\Sigma} \in \mathbb{R}^{N_r \times N_r}$ is diagonal; here, $N_r \leq \min(N_x, N_y)$ defines the rank of $\mathbf{A}$.

The elements $\{\check{\zeta}_{\mathbf{A}}^{(}i)\}_{i\in[1,\dots,N_r]}$ of the main diagonal of $\mathbf{\Sigma}$ are known as **singular values (s.v.)** of $\mathbf{A}$. Without loss of generality, those are typically sorted in descending order of magnitude (i.e., $\check{\zeta}_{\mathbf{A}}^{(1)} \geq \check{\zeta}_{\mathbf{A}}^{(2)} \geq \dots \geq \check{\zeta}_{\mathbf{A}}^{(N_r)} > 0$).

As described in Section 2.1.2, the most naive approach for the inversion is given by the minimization of the MSE between the simulated noiseless output $\mathbf{y}^{[sim]}$ of eq. (2.1.1) and the observation $\mathbf{y}$:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{E}_x} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 = \mathbf{A}^\dagger\mathbf{y} = \mathbf{U}^*\mathbf{\Sigma}^{-1}\mathbf{V}\mathbf{y}\,, \tag{2.2.2}$$

where $\mathbf{A}^\dagger$ denotes the Moore-Penrose pseudo-inverse of $\mathbf{A}$. If the SVD of $\mathbf{A}$ is known, the SVD $\mathbf{A}^\dagger = \mathbf{U}^*\mathbf{\Sigma}^{-1}\mathbf{V}$ can be efficiently computed, as $\mathbf{\Sigma}^{-1}$ is also diagonal and the generic i-th element $\check{\zeta}_{\mathbf{A}^\dagger}{}^{(i)}$ of its main diagonal is given by $\check{\zeta}_{\mathbf{A}^\dagger}{}^{(i)} = 1/\check{\zeta}_{\mathbf{A}}^{(i)}$.

In the majority of practical scenarios, the problem at hand is ill-conditioned, as the condition number associated with the inverse problem is generally too large to be computationally stable. The singular values associated with $\mathbf{A}$ are generally upper bounded, as they are associated with physical processes which are limited in energy.

However, if certain s.v. of $\mathbf{A}$ are below a certain threshold, the associated s.v. of $\mathbf{A}^\dagger$, and consequently, the condition number associated with the inverse problem, may get too large. Consequently, the estimation itself is sensitive to errors which are either introduced by the finite precision or by uncertainties introduced by noise.

In the literature, the approaches based on defining a new set of s.v. are commonly known as **penalized matrix decomposition (PMD)**; among those, the most widespread approaches are:

- **Truncated singular value decomposition (TSVD)**: The TSVD approach consists in setting all the s.v. below a certain threshold equal to zero. I.e. if we let $N'_r \leq N_r$ denote the amount of nonzero s.v., then:

$$\check{\zeta}^{(i)}_{\widetilde{\mathbf{A}}^\dagger} = \begin{cases} 1/\check{\zeta}^{(i)}_{\mathbf{A}} & \text{if } i \leq N'_r\,, \quad (2.2.3a) \\ 0 & \text{if } i > N'_r\,. \quad (2.2.3b) \end{cases}$$

- **Ridge regression (RR)**: Following the framework described in Section 2.1.1, the approach known as **Tikhonov regularization** was popularized, as it was proven to satisfy the requirements of a regularizing operator to be robust to errors both in the observations and the direct operator $\mathbf{A}$ [219]. It consists in the following estimation:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbf{E}_x} J(\mathbf{x}) = \arg\min_{\mathbf{x}\in\mathbf{E}_x} \|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2 + \|\mathbf{\Lambda}\mathbf{x}\|_2^2 \quad (2.2.4a)$$

$$= \arg\min_{\mathbf{x}\in\mathbf{E}_x} \|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2 + \check{\lambda}^2\|\mathbf{x}\|_2^2\,, \quad (2.2.4b)$$

where $\mathbf{\Lambda}$ is a domain transformation matrix, known as **Tikhonov matrix**. The expression (2.2.4b) is a special case known as ridge regression (RR), for which the Tikhonov matrix $\mathbf{\Lambda} = \check{\lambda}\mathbf{I}$ is an identity matrix $\mathbf{I}$, multiplied by a scalar factor $\check{\lambda}$, known as **ridge parameter**. For this choice, the expression of the gradient $\nabla_J(\mathbf{x})$ of the objective function $J(\mathbf{x})$ is given by:

$$\nabla_J(\mathbf{x}) = 2\left(\mathbf{A}^*\left(\mathbf{A}\mathbf{x}-\mathbf{y}\right) + \check{\lambda}^2\mathbf{x}\right) \quad (2.2.5a)$$

$$= 2\left(\mathbf{A}^*\mathbf{A} + \check{\lambda}^2\mathbf{I}\right)\mathbf{x} - 2\mathbf{A}^*\mathbf{y} \quad (2.2.5b)$$

$$= 2\mathbf{V}\left(\mathbf{\Sigma}\mathbf{\Sigma}^* + \check{\lambda}^2\mathbf{I}\right)\mathbf{V}^*\mathbf{x} - 2\mathbf{V}\mathbf{\Sigma}^*\mathbf{U}^*\mathbf{y}\,, \quad (2.2.5c)$$

where we have used the property of semi-orthogonality of $\mathbf{V}$ to express the identity matrix as $\mathbf{I} = \mathbf{V}^*\mathbf{V}$. The condition $\nabla_J(\widehat{\mathbf{x}}) = 0$ is then equivalent to:

$$\widehat{\mathbf{x}} = \mathbf{V}\left(\mathbf{\Sigma}\mathbf{\Sigma}^* + \check{\lambda}^2\mathbf{I}\right)^{-1}\mathbf{\Sigma}^*\mathbf{U}^*\mathbf{y}\,, \quad (2.2.6)$$

thus the influence matrix is characterized by $\widetilde{\mathbf{\Sigma}} = \left(\mathbf{\Sigma}\mathbf{\Sigma}^* + \check{\lambda}^2\mathbf{I}\right)^{-1}\mathbf{\Sigma}^*$ and consequently:

$$\check{\zeta}_{\widetilde{\mathbf{A}}^\dagger}{}^{(i)} = \frac{\check{\zeta}^{(i)}_{\mathbf{A}}}{\left(\check{\zeta}^{(i)}_{\mathbf{A}}\right)^2 + \check{\lambda}^2}\,, \qquad \forall i \in [1, \dots, N_r]\,. \quad (2.2.7)$$

Choice of the regularization parameter

The choice of the parameter N_r' in the TSVD or $\check{\lambda}$ in the RR is not a trivial problem and many possible approaches have been proposed in the literature to automatize its choice and a noncomprehensive selection of the most widespread methods is provided here.

Let $\widetilde{\mathbf{x}}_\beta$ denote the estimation obtained with a given parameter β (either $\check{\lambda}$ or N_r' depending on the method). Three widespread strategies are available:

- **Morozov's Discrepancy Principle** [171]: Let us suppose the observation $\mathbf{y}$ is known with a certain uncertainty such that $\|\mathbf{y} - \mathbf{y}^{[sim]}\| < \check{\delta}$, then the discrepancy principle consists in picking an the largest β such that $\|\mathbf{A}\widehat{\mathbf{x}}_\beta - \mathbf{y}\|_2 \leq \delta$ [102].
- **Generalized cross validation (GCV)** [93]: For the case of RR, a common procedure for the estimation of $\check{\lambda}^2$ is given by the the following minimization criterion:

$$\arg\min_{\check{\lambda}^2} \frac{\frac{1}{N_r}\|\mathbf{A}\widehat{\mathbf{x}}_{\check{\lambda}}\|_2^2}{\left(\frac{1}{N_r}\operatorname{tr}\left(\mathbf{I} - \mathbf{A}\left(\mathbf{A}\mathbf{A}^* + \check{\lambda}^2\mathbf{I}\right)^{-1}\mathbf{A}^*\right)\right)^2}, \tag{2.2.8}$$

  where $\operatorname{tr}(\cdot)$ evaluates the trace of the square matrix it takes as argument.
- **L-curve criterion** [113]: the parameter β is chosen, such that maximizes the curvature of the parametric curve $\log(\|\mathbf{A}\mathbf{x}_\beta - \mathbf{y}\|_2)$ /$\log(\|\mathbf{x}_\beta\|_2)$.

If applied to real world scenarios, each of these methods has a series of drawbacks. I.e. the discrepancy methods require the knowledge of the noise energy in the system, the characteristic functional GCV is sometimes too small (below the machine precision) within the decision range and that L-curve decision does not converge in the case of low-noise acquisitions [110].

With some adjustments [197, 55], these techniques can also be applied to the sparsity-inducing regularizations that will be presented in the following section.

More recently, the **hierarchical Bayesian inference technique** [184] was proposed for an iterative algorithm for successively refining the choice of the regularization parameter $\check{\lambda}$. The method applies to the solution of all objective functions in the form $J(\mathbf{x}) = f(\mathbf{x}) + \check{\lambda}g(\mathbf{x})$, where the term $f(\mathbf{x}) = -\log p_{\boldsymbol{\psi}}(\mathbf{y}|\mathbf{x})$ is the ML estimator

and $\check{\lambda} g(\mathbf{x}) = -\log p_{\boldsymbol{\chi}}(\mathbf{x})$ is the prior. Let us suppose that the regularizer term λ has its own stochastic formulation, so that the a priori probability can be expressed as:

$$p_{\boldsymbol{\chi}}(\mathbf{x}) = \int_0^{\infty} p_{\boldsymbol{\chi}}(\mathbf{x}|\lambda) p_{\lambda}(\check{\lambda}) d\check{\lambda} \tag{2.2.9}$$

where $p_{\lambda}(\check{\lambda})$ is the probability density function (pdf) associated to the regularization parameter and:

$$p_{\boldsymbol{\chi}}(\mathbf{x}|\check{\lambda}) = \frac{\exp(-\check{\lambda} g(\mathbf{x}))}{C\left(\check{\lambda}\right)} \tag{2.2.10a}$$

$$C\left(\check{\lambda}\right) = \int \mathbf{x} \in \mathbb{R}^{N_x} \frac{\exp(-\check{\lambda} g(\mathbf{x}))}{d} \mathbf{x} \tag{2.2.10b}$$

is the expression of prior conditioned to a certain value of λ, which has to be normalized over the space of the input parameter $\mathbf{x}$ through the term $C\left(\check{\lambda}\right)$. If $g(\mathbf{x})$ is N_m-homogeneous (that is, if $g(\eta\mathbf{x}) = \eta^{N_m} g(\mathbf{x})$, which is the case of most norms), then $C\left(\check{\lambda}\right) = C(1)\check{\lambda}^{-N_x/N_m}$, and assuming that $\check{\lambda}$ is distributed with an exponential distribution ($p_{\lambda}(\check{\lambda}) = \exp -\beta\check{\lambda}$ for $\check{\lambda} > 0$), the proposed method proves [184], that if the majorization-minimization (MM) algorithm [212] is applied to the cost function $J(\mathbf{x})$, given a first guess $\check{\lambda}^{(0)}$ for the regularization parameter, the q-th update $\widehat{\mathbf{x}}_{\check{\lambda}^{(q)}}$ of the desired product is obtained as:

$$\widehat{\mathbf{x}}_{\check{\lambda}^{(q)}} = \arg\min_{\mathbf{x}} f(\mathbf{x}) + \check{\lambda}^{(q-1)} g(\mathbf{x}) \tag{2.2.11a}$$

$$\check{\lambda}^{(q)} = \frac{N_x/N_m - 1}{g(\widehat{\mathbf{x}}_{\check{\lambda}^{(q)}}) - 1} \tag{2.2.11b}$$

2.2.2 Sparsity-inducing regularizers

In the context of image processing, it is a common scenario to work with data that contains a large amount of redundancy, which allows for natural images to be represented in **sparse domains**, i.e. domains with a limited amount of nonzero

elements. This property has allowed to develop very popular algorithms for image compression, such as the Joint Photographic Experts Group (JPEG) [@7].

The sparsity condition is properly measured with the pseudo-norm ℓ_0, which, in the formalism of Donoho [63], denotes the amount of nonzero elements in an array. The problem is theoretically formalized by:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbf{E}_x} \| \mathbb{L}(\mathbf{x})\|_0 \quad \text{s.t. } \| \mathbb{A}(\mathbf{x}) - \mathbf{y}\|_2^2 \leq \check{\epsilon}^2 \,, \tag{2.2.12}$$

where $\| \cdot \|_0$ denotes the ℓ_0-norm, $\mathbb{L}(\cdot)$ is an operator that represents the argument in a sparse domain and $\check{\epsilon}^2$ is a given scalar threshold.

This problem is not trivial to solve exactly in the form of (2.2.12), which led to a series of techniques of **sparse approximation** to make it mathematically approachable. Two main strategies exist: the **orthogonal matching pursuit (OMP)** [222], a greedy iterative algorithm which sequentially locates all the nonzero elements of the solution, and the **basis pursuit** [62], where the ℓ_0-norm is substituted with an ℓ_1-norm.

In this second approach, the estimation is given by the functional known as **least absolute shrinkage and selection operator (LASSO)** [217]:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbf{E}_x} \frac{1}{2}\| \mathbb{A}(\mathbf{x}) - \mathbf{y}\|_2^2 + \check{\lambda}\| \mathbb{L}(\mathbf{x})\|_1 \,, \tag{2.2.13}$$

and a visual representation of its sparse-inducing effect (for $\mathbb{L}$ equal to an identity operator) in comparison to the RR is provided in Fig. 2.2.

In this thesis, two orthogonal domain transformations will be employed to induce sparsity in the data: the **discrete cosine transform (DCT)** and the **discrete wavelet transform (DWT)**.

If $\mathbf{x} \in \mathbb{R}^{N_x}$, the most common form of DCT, known as DCT-II, is a multiplication by a matrix $\mathbf{W} \in \mathbb{R}^{N_x \times N_x}$, whose elements $\{w_{kl}\}_{k\in[1,\dots,N_x],l\in[1,\dots,N_x]}$ are:

$$w_{kl} = \sqrt{\frac{2}{N_x}} \cos\left(\frac{\pi}{N_x}\left(k+\frac{1}{2}\right)l\right) \,. \tag{2.2.14}$$

The obtained coefficients are related to the discrete Fourier transform (DFT)[198] and can be seen as a frequency representation of the signal. Multidimensional versions of this transform are also available, i.e. in the case it is applied to both spatial coordinates of an image.

The DWT is the topic of a vast and mature literature, and the work of Mallat [155] provides an in depth introduction. In this context, we just want to recall the main ideas behind it, with no claim of mathematical rigorousness. The DWT's goal is to discretely sample a (typically dyadic) wavelet transform and the obtained coefficients provide **time-scale representation** of the original signal, i.e. embedding information relative both to the self-similarity of the signal and in the original domain.

The DWT is based on the multiresolution analysis (MRA), which consists in segmenting the space $\mathtt{E}_x$ into nested subspaces [155], which maps into subsequent levels of decomposition, whose coefficients describe the evolution of the signal in the original domain at each scale.

For each level, the signal is simultaneously decomposed by passing through a low and high pass convolution filter, whose kernels are related and form a **quadrature mirror filter (QMF)** [9]; subsequently, the coefficients are decimated by a factor of 2. If there is no source of error, biorthogonal wavelets define a particular class of invertible DWT, for which case the decomposition and the associated reconstruction are shown in Fig. 2.3 and employ a complementary set of synthesis and analysis filters. If perfect reconstruction is achieved with the same filters, the wavelet is called orthogonal, so that the inverse and the adjoint operator of the DWT are the same, as seen in eq. (A.1.1).

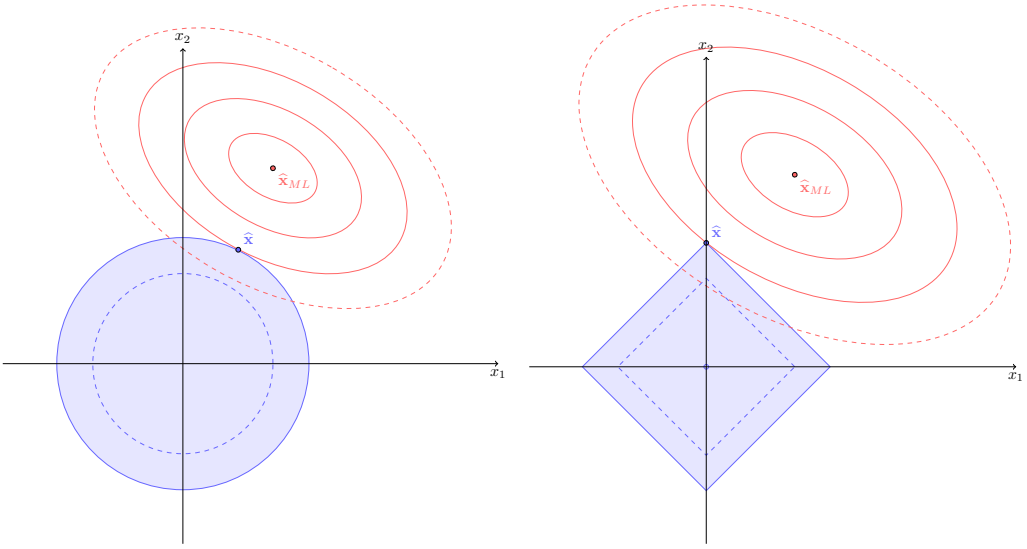
The DWT can be straightforwardly extended to multiple dimensions, and an example of the product for monochromatic images is shown in Fig. 2.4. A variant algorithm, known as stationary wavelet transform (SWT), proposed by Holschneider et al. [121], is sometimes employed in image processing to allow for the transformation to verify the property of translation invariance; this result is obtained by skipping the decimation and comes at the cost of introducing redundancy in the transformation.

2.2.3 Variational methods

In their seminal paper, the authors Rudin, Osher and Fatemi [200] propose a regularization model for images $s(\mathbf{r}) : \mathtt{E}_r \subseteq \mathbb{R}^2 \to \mathbb{R}$ defined on an ideally continuous plane $\mathbf{r} = [r_1; r_2]$ and for which it is possible to define a gradient $\nabla_{\mathbf{r}} s = \left[\frac{\partial s}{\partial r_1}; \frac{\partial s}{\partial r_2}\right]$.

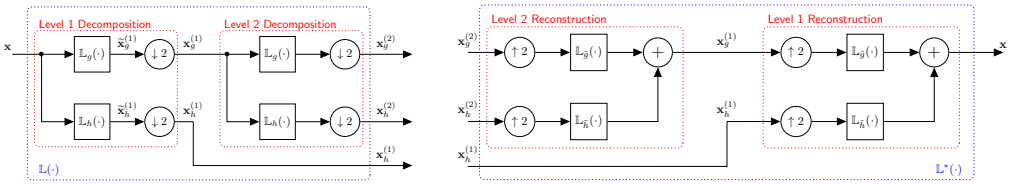
The regularization criterion consists in minimizing the term:

$$\iint_{\mathtt{E}_r} \|\nabla_{\mathbf{r}} s(\mathbf{r})\|_2 \, d\mathbf{r} \tag{2.2.15}$$



(a) ℓ_2-norm regularization **(b)** ℓ_1-norm regularization

Fig. 2.2. Representation of the ℓ_1 and ℓ_2-norm regularization effect for arrays of length 2, whose space spans in the coordinates $[x_1; x_2]$. The MLE estimation is denoted with $\widehat{\mathbf{x}}_{ML}$. Along the red curves, the data fidelity term has constant magnitude and the regularized estimation $\widehat{\mathbf{x}}$ is obtained at the intersection between the ℓ_2 and ℓ_1 ball, whose sizes depend on the regularization parameter. The ℓ_1 regularizer has a sparsity-inducing effect, since it has a single non-zero component.



(a) Decomposition **(b)** Reconstruction

Fig. 2.3. Block scheme for a two-level biorthogonal wavelet decomposition and reconstruction through QMFs. The synthesis filters $\mathbb{L}_g(\cdot)$ and $\mathbb{L}_h(\cdot)$ implement a high and a low pass convolution product, respectively. $\mathbb{L}_{\tilde{g}}(\cdot)$ and $\mathbb{L}_{\tilde{h}}(\cdot)$ denote the associated analysis filters to allow for perfect reconstruction. They are identical to $\mathbb{L}_g(\cdot)$ and $\mathbb{L}_h(\cdot)$ in the case of orthogonal wavelets. $\downarrow 2$ and $\uparrow 2$ denote the operations of decimation and expansion by a factor of 2, respectively, which are bypassed for SWTs.

(a) Original image **(b)** 2-level DWT

Fig. 2.4. Example of a 2-level bidimensional wavelet applied on a square image. In this case the DWT transformation is applied to both dimension, so at each step, 4 different classes of coefficients are generated, according to the combination of low-pass and high-pass filters. The successive step is just performed on the low-pass component (here, depicted in the top left corner). The sparsity-inducing effect is also shown, as the high-pass components feature a high amount of low-intensity (black) pixels.

in conjunction with the data fidelity term that characterizes the specific problem (e.g. denoising, deblurring, etc.). The signal $s(\mathbf{r})$ can be seen as an ideal monochromatic input image with infinite spatial resolution; the prior constraint given by eq.(2.2.15) penalizes fast oscillations in sufficiently homogeneous zones, while keeping sharp transitions across intensity edges. This criterion promotes sparsity across the gradient of the image, flattening the variation of intensity in regions bounded by those sharp transitions.

The regularization criteria based on this principle were then identified as **variational methods** [2], and the approaches aimed at adapting the Rudin-Osher-Fatemi (ROF) model (2.2.15) to discrete spaces became popular with the name **Discrete Total Variation**, or simply total variation (TV).

This section provides a brief introduction to the approaches developed in the literature [31, 48]. The employed formalism separates between strategies to evaluate the gradient and metrics to weight their combined contribution, based on the formalism of the collaborative total variation (CTV) [65, 66].

[2] In the classical literature of variational methods, the term is usually referred to any optimization method that requires an analysis to any kind of variation of the quantities to estimate, such as in the case of analysis of the gradients. In this work, we instead refer to variational methods as an analysis of variations of the spatial intensities.

Notation

In the following, we denote with $\boldsymbol{\mathcal{U}}^{[x]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_b}$ the conventional 3-dimensional array representation of an image, composed of N_b frontal slices $\left\{ \mathbf{U}^{[x]}_{::k} \right\}_{k \in [1,\dots,N_b]}$, each relative to a particular channel.

For our purposes, it will be also useful to represent the same images in **lexicographic order**; in this representation, denoted with $\mathbf{X} = \operatorname{matr}\left(\boldsymbol{\mathcal{U}}^{[x]}\right)$, the original image $\boldsymbol{\mathcal{U}}^{[x]}$ is reshaped into a horizontal matrix $\mathbf{X} \in \mathbb{R}^{N_i \times N_b}$, with $N_i = N_{i_1} N_{i_2}$. The concatenation operation $[\cdot;\cdot]$, $[\cdot,\cdot]$, and $[\cdot,\cdot]_3$ denotes a concatenation over the first, second and third dimension, respectively. $\|\cdot\|_{p_1 p_2 p_3}$ denotes the ℓ_{p_1}, ℓ_{p_2} and ℓ_{p_3}-norm, applied in order on the third, second and first dimension of the 3-dimensional array, respectively [3] I.e., the operator $\left\|\boldsymbol{\mathcal{U}}^{[x]}\right\|_{\infty 21}$ applies the ℓ_∞-norm over each frontal slice (each channel), then the ℓ_2-norm over the rows, and finally the ℓ_1-norm over the columns.

A generic coefficient of the 3-dimensional array is denoted with the associated lowercase letter, i.e. $u^{[x]}_{i_1,i_2,k}$ (or sometimes simply $u^{[x]}_{i_1 i_2 k}$) represents the pixel index coordinates (i_1, i_2) of the k-th channel of $\boldsymbol{\mathcal{U}}^{[x]}$.

Total variation

In the most classic formulation of the TV, the main building blocks for the formulation of a discrete regularizer are the gradients $\boldsymbol{\mathcal{U}}^{[h]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_b}$ and $\boldsymbol{\mathcal{U}}^{[v]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_b}$ in the horizontal and vertical direction of the image, respectively. They are defined as:

$$u^{[v]}_{i_1 i_2 k} = u^{[x]}_{i_1+1,i_2,k} - u^{[x]}_{i_1,i_2,k}\,, \tag{2.2.16a}$$

$$u^{[h]}_{i_1 i_2 k} = u^{[x]}_{i_1,i_2+1,k} - u^{[x]}_{i_1,i_2,k}\,, \tag{2.2.16b}$$

$$\tag{2.2.16c}$$

paying attention to replace the values outside of the domain of $\boldsymbol{\mathcal{U}}^{[x]}$ with zeroes. The TV operator, which we will denote with $\mathbb{L}_{TV}(\mathbf{X})$ can then be obtained by simply concatenating the two building blocks over a common singleton dimension:

$$\mathbb{L}_{TV}(\mathbf{X}) = \left[\operatorname{matr}\left(\boldsymbol{\mathcal{U}}^{[v]}\right), \operatorname{matr}\left(\boldsymbol{\mathcal{U}}^{[h]}\right)\right]_3 . \tag{2.2.17}$$

[3] We opted in this work to invert the order with respect to the dimensions to prioritize the order in which each norm is applied first.

The regularizers $g_{ATV}(\mathbf{X})$ and $g_{ITV}(\mathbf{X})$ associated with the anisotropic and isotropic TV, respectively, are then defined as:

$$g_{ATV}(\mathbf{X}) := \sum_{i_1=1}^{N_{i_1}} \sum_{i_2=1}^{N_{i_2}} \sum_{k=1}^{N_b} \left(\left| u_{i_1 i_2 k}^{[h]} \right| + \left| u_{i_1 i_2 k}^{[v]} \right| \right) \qquad = \| \mathbb{L}_{TV}(\mathbf{X}) \|_{111} \,, \tag{2.2.18a}$$

$$g_{ITV}(\mathbf{X}) := \sum_{i_1=1}^{N_{i_1}} \sum_{i_2=1}^{N_{i_2}} \sum_{k=1}^{N_b} \sqrt{\left(u_{i_1 i_2 k}^{[h]} \right)^2 + \left(u_{i_1 i_2 k}^{[v]} \right)^2} \qquad = \| \mathbb{L}_{TV}(\mathbf{X}) \|_{211} \,. \tag{2.2.18b}$$

The anisotropic total variation (ATV) can be seen as a sparsity-inducing regularizer in the domain of the gradient of the image, not in a dissimilar way to what was shown in Section 2.2.2, but is known to be a poor discrete approximation of the ROF criterion, as it over-estimates the gradient over oblique contours. The isotropic total variation (ITV) partly improves by imposing more isotropy across oblique directions, however it can be shown that both regularization functionals (2.2.18) do not compute to the same value after the image is flipped horizontally/vertically or rotated by 90 or 180 degrees [48].

To increase the isotropy of the representation, the approach of Chambolle et al. [39], known as upwind total variation (UTV), consists in evaluating 4 different gradients:

$$u_{i_1 i_2 k}^{[v^+]} = \max \left(u_{i_1+1, i_2, k}^{[x]} - u_{i_1, i_2, k}^{[x]}, 0 \right) \,, \tag{2.2.19a}$$

$$u_{i_1 i_2 k}^{[v^-]} = \max \left(u_{i_1, i_2, k}^{[x]} - u_{i_1-1, i_2, k}^{[x]}, 0 \right) \,, \tag{2.2.19b}$$

$$u_{i_1 i_2 k}^{[h^+]} = \max \left(u_{i_1, i_2+1, k}^{[x]} - u_{i_1, i_2, k}^{[x]}, 0 \right) \,, \tag{2.2.19c}$$

$$u_{i_1 i_2 k}^{[h^-]} = \max \left(u_{i_1, i_2, k}^{[x]} - u_{i_1, i_2-1, k}^{[x]}, 0 \right) \,, \tag{2.2.19d}$$

and the associated linear operator

$$\mathbb{L}_{UTV}(\mathbf{X}) = \left[\text{matr} \left(\boldsymbol{\mathcal{U}}^{[v^+]} \right), \text{matr} \left(\boldsymbol{\mathcal{U}}^{[v^-]} \right), \text{matr} \left(\boldsymbol{\mathcal{U}}^{[h^+]} \right), \text{matr}(\boldsymbol{\mathcal{U}}^{[h^-]}) \right]_3 \tag{2.2.20}$$

results invariant to flipping the image around a central vertical or horizontal axis.

More sophisticated solutions are available in the literature, such as the Shannon total variation (STV) [1], that tries to reconciliate the theory of TV with the Shannon interpolation theory, a new definition of TV with very highly isotropic behaviours [48], and the total generalized variation (TGV) [31], which considers higher order derivatives of $\boldsymbol{\mathcal{U}}$.

Collaborative total variation

The principles of TV described up to now considers the contribution on the regularizer of each channel separately; the framework proposed by Duran et al. [65, 66] allows to take into account the cross-correlation across channels. In this formalism, known as collaborative total variation (CTV), the regularizer operates over a transformed domain $\mathbb{L}(\mathbf{X})$, where one can identify three different dimensions relative to the derivatives (gradients), the channels and the pixels, and the associated regularization function is a combination of the norms operating over the three dimensions. The estimation problem then reduces to this minimization problem:

$$\mathbf{X} = \arg \min_{\mathbf{X} \in \mathbb{R}^{N_i \times N_b}} \frac{1}{2} \| \mathbb{A}(\mathbf{X}) - \mathbf{y} \|_2^2 + \check{\lambda} \| \mathbb{L}(\mathbf{X}) \|_{p_1 p_2 p_3} \,. \tag{2.2.21}$$

The operator $\mathbb{L}(\mathbf{X})$ is a generic TV operator, which, i.e., could be substituted with either eq. (2.2.17) or (2.2.20).

The norms ℓ_{p_1}, ℓ_{p_2} and ℓ_{p_3} are the norms relative to the gradients, channels and pixels, respectively, with the norm $\| \cdot \|_{221}$ being a particularly widespread choice, known as **vector total variation (VTV)** [32].

The norms in eq. (2.2.21) can also operate jointly across multiple dimensions, such as in the case of the **Shatten norm** S_p defined as the ℓ_p norm of the s.v. of a matrix. The case $p = 1$, known in the literature as **nuclear norm**, has been shown to yield particularly good reconstruction performances of red green blue (RGB) images if the regularizer is in the form $\| \mathbb{L}_{TV}(\mathbf{X}) \|_{S_1, \ell_1}$ where the nuclear norm is applied jointly on the derivatives and the channels [203, 66] (the subsequent ℓ_1-norm over all pixels gives an equal penalization weight in every local area of the image).

2.3 Algorithms for inverse problems

The minimization of objective functions that will be addressed in this thesis can be generally seen as the solution of a general problem in the form:

$$\widehat{\mathbf{x}} = \arg \min_{\mathbf{x}} J(\mathbf{x}) = \arg \min_{\mathbf{x}} f(\mathbb{A}(\mathbf{x})) + \langle \mathbf{x}, \mathbf{c} \rangle + g(\mathbb{L}(\mathbf{x})) \,, \tag{2.3.1}$$

where:

- $J(\mathbf{x})$ denotes the objective function,
- $\mathbf{x} \in \mathrm{E}_x$ is the data to reconstruct,

- $\mathbb{A}(\cdot) : \mathtt{E}_x \to \mathtt{E}_y$ is the direct model operator,
- $\mathbb{L}(\cdot) : \mathtt{E}_x \to \mathtt{E}_u$ is a generic domain transform operator,
- $f(\cdot) : \mathtt{E}_y \to \mathbb{R}^+$ is the metric for the data term,
- $g(\cdot) : \mathtt{E}_u \to \mathbb{R}^+$ is the metric for the regularization term,
- $\langle \cdot, \cdot \rangle : \mathtt{E}_x \times \mathtt{E}_x \to \mathbb{R}^+$ is a scalar product defined in $\mathtt{E}_x$,
- $\mathbf{c} \in \mathtt{E}_x$ is a constant term.

The sets $\mathtt{E}_x$ and $\mathtt{E}_u$ are sometimes known as **primal** and **dual space**, respectively, and the joint minimization on both of these spaces can be exploited to improve the convergence speed of the algorithms aimed at solving (2.3.1) [29, 40].

In the case when $\mathbb{A}(\cdot)$ is a bounded **linear operator** (or a linearized version of some generically nonlinear process), it can be shown that the formalism of (2.3.1) is a generalization of eq. (2.1.14), with $g'(\cdot) = \frac{1}{\check{\lambda}} g(\mathbb{L}(\cdot))$. In fact:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathtt{E}_x} \frac{1}{2} \|\mathbb{A}(\mathbf{x}) - \mathbf{y}\|^2 + \check{\lambda} g(\mathbb{L}(\mathbf{x})) \tag{2.3.2a}$$

$$= \arg\min_{\mathbf{x} \in \mathtt{E}_x} \frac{1}{2} \|\mathbb{A}(\mathbf{x})\|^2 - \langle \mathbb{A}(\mathbf{x}), \mathbf{y} \rangle + \frac{1}{2} \|\mathbf{y}\|^2 + \check{\lambda} g(\mathbb{L}(\mathbf{x})) \tag{2.3.2b}$$

$$= \arg\min_{\mathbf{x} \in \mathtt{E}_x} \frac{1}{2} \| \mathbb{A}(\mathbf{x})\|^2 + \langle \mathbf{x}, -\mathbb{A}^*(\mathbf{y}) \rangle + \check{\lambda} g(\mathbb{L}(\mathbf{x})) \,, \tag{2.3.2c}$$

where $\mathbb{A}^*$ denotes the adjoint of $\mathbb{A}$, defined in Appendix A.1.1, and we used the symmetry of the scalar product in the step (2.3.2b) and the invariance of the minimization to $\|\mathbf{y}\|$ in the step (2.3.2c).

This section's goal is a brief introduction to possible algorithmical approaches for the solution of eq. (2.3.1) under different conditions, starting with the classical, but limiting, approach of the gradient descent (GD) and then expanding the solution to a more encompassing class of solvers, based on the proximal operator; finally, we present some classical approaches to manage the nonlinearity of the direct model.

2.3.1 Gradient descent algorithms

The **gradient descent (GD) algorithm** is a first-order optimization algorithm, attributed to Cauchy, but popularized by Hadamard [106]. It is applicable to any situation in which the objective function $J(\cdot)$ is convex and differentiable, so that it is possible to define an associated Fréchet gradient ∇_J.

The GD method consists in iterative refinements of the estimation of the minimum of J, obtained through successive descents in the direction of the gradient ∇_J. If the current estimate at the q-th step is denoted with $\mathbf{x}^{(q)}$, the next step update $\mathbf{x}^{(q+1)}$ is given by:

$$\mathbf{x}^{(q+1)} = \mathbf{x}^{(q)} - \check{\gamma}\nabla_J(\mathbf{x}^{(q)})\,. \tag{2.3.3}$$

If ∇_J is a $\check{\beta}$-Lipschitz continuous function:

$$\|\nabla_J(\mathbf{x}) - \nabla_J(\mathbf{x}_0)\|_x \leq \check{\beta}\|\mathbf{x} - \mathbf{x}_0\| \qquad \forall \mathbf{x}, \mathbf{x}_0 \in \mathtt{E}_x \tag{2.3.4}$$

where $\|.\|_x$ is the norm of the normed space $\mathtt{E}_x$, the convergence is assured for any coefficient $\check{\gamma} \leq 2/\check{\beta}$. I.e. if $J(\mathbf{x}) = \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2$, where $\mathbf{A}$ is a matrix, the condition is equivalent to $\check{\gamma} \leq 2/\|\mathbf{A}\|_{op}$.

If $J(\mathbf{x})$ is written as the quadratic expression (2.3.1) and both $f(\cdot)$ and $g(\cdot)$ are differentiable with gradients $\nabla_f(\cdot)$ and $\nabla_g(\cdot)$, respectively, the update step can be rewritten as:

$$\mathbf{x}^{(q+1)} = \mathbf{x}^{(q)} - \check{\gamma}\left(\mathbb{A}^*(\nabla_f(\mathbb{A}(\mathbf{x}))) + \mathbf{c} + \mathbb{L}^*(\nabla_g(\mathbb{L}(\mathbf{x})))\right)\,. \tag{2.3.5}$$

The **conjugate descent** is a very common variant of the GD method, for which the descent is not pursued in the direction of the local gradient of the function, but instead on one orthogonal to all previously explored directions, derived with a Gram-Schmidt decomposition [119].

2.3.2 Proximal gradient algorithms

In the context of inverse problems, it is very common to face the situation in which the objective function to minimize is not differentiable, such as in the case of the ℓ_1-norms employed in the LASSO framework (Section 2.2.2). To address this problem, we introduce here a particular class of iterative algorithms, known as **proximal gradient algorithms** or simply **proximal algorithms** [181] This class of algorithms relies on the use of **proximal operators**, which we describe in the following section; we also list some of its desirable properties and introduce two of the most widespread algorithms belonging to this class: the **Loris-Verhoeven** and the **Chambolle-Pock algorithms**.

Iterative nonexpansive algorithms

The basic idea of iterative algorithms is to perform a series of operations $\mathcal{T}$ on a given estimate of $\mathbf{x}^{(q)}$ to obtain a new estimate $\mathbf{x}^{(q+1)} = \mathcal{T}(\mathbf{x}^{(q)})$ for the minimizer of $J(\mathbf{x}) \in \Gamma_0(\mathtt{E}_x)$, real-valued bounded lower semicontinuous convex cost function defined in a Hilbert space $\mathtt{E}_x$. For the transformation $\mathcal{T}$ to be convergent, the necessary condition is given by:

$$\lim_{q\to\infty} J(\mathbf{x}^{(q)}) = \min_{\mathbf{x}} J(\mathbf{x})\,, \qquad \forall \mathbf{x}^{(0)} \in \mathtt{E}_x\,. \tag{2.3.6}$$

Additionally, if the iterative operator $\mathcal{T}$ is also $\check{\beta}$-Lipschitz continuous for a certain $\check{\beta} \in (0,1)$:

$$\|\mathcal{T}(\mathbf{x}_1) - \mathcal{T}(\mathbf{x}_2)\| \leq \check{\beta}\|\mathbf{x}_1 - \mathbf{x}_2\|, \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathtt{E}_x \tag{2.3.7}$$

the estimation converges to one of the **fixed points**, defined as the points $\widehat{\mathbf{x}} \in \mathtt{E}_x$: $\mathcal{T}(\widehat{\mathbf{x}}) = \widehat{\mathbf{x}}$. An effective design of an iterative algorithm $\mathcal{T}$ for the minimization of $J(\mathbf{x})$ is made up of two main steps:

- Check that the fixed points of $\mathcal{T}$ are also minimizers of $J(\mathbf{x})$;
- Assure that $\mathcal{T}$ is $\check{\beta}$-Lipschitz continuous for a certain $\check{\beta} \in (0,1)$.

If the operator $\mathcal{T}$ is nonexpansive (that is if it is 1-Lipschitz continuous), it is possible to generate a $\check{\beta}$-averaged new operator $\mathcal{T}' = \check{\beta}\mathcal{T} + (1-\check{\beta})\,\mathrm{Id}$, where Id is the identity operator. The so called $\check{\beta}$-averaged operator $\mathcal{T}'$ will be $\check{\beta}$-Lipschitz continuous. This result is known in the literature as **Krasnoselskii–Mann method** for nonexpensive mappings [44]. Equivalently, if the operator is already $\check{\beta}$-Lipschitz continuous for a certain $0 < \check{\beta} < 1$, one can over-relax the operator $\mathcal{T}'$ by $\check{\rho}$-averaging it with any $0 < \check{\rho} < 1/\check{\beta}$, so that the operator becomes $(\check{\rho}\check{\beta})$-Lipschitz continuous [50].

Proximal operator

Let $f : \mathtt{E}_x \subseteq \mathbb{R}^{N_x} \to \mathbb{R} \setminus [-\infty, \infty]$ be a lower semi-continuous convex function. The **proximal operator** associated with a function $\check{\gamma} f$, obtained by scaling f by a factor $\check{\gamma} > 0$, is defined by:

$$\mathrm{prox}_{\check{\gamma} f}(\mathbf{x}') = \arg\min_{\mathbf{x}\in\mathtt{E}_x} \left(f(\mathbf{x}) + \frac{1}{2\check{\gamma}}\|\mathbf{x} - \mathbf{x}'\|_2^2 \right)\,. \tag{2.3.8}$$

The function $M_{\check{\gamma} f}(\mathbf{x}') = \min_{\mathbf{x}\in\mathtt{E}_x} \left(f(\mathbf{x}) + \frac{1}{2\check{\gamma}}\|\mathbf{x} - \mathbf{x}'\|_2^2 \right)$ is usually known as **Moreau envelope** of $\check{\gamma} f$ and is, in non rigorous terms, a smoothed and regularized version

of f. $M_{\check{\gamma} f}(\mathbf{x}')$ is continuously differentiable and its domain is in $\mathbb{R}^{N_x}$, even if f is not. Fig. 2.5a shows a visual example of the Moshida envelope for a particular f.

With this interpretation, the proximal operator is a compromise between minimizing f and being close (in terms of the Euclidean metric) to $\mathbf{x}$, with the parameter $\check{\gamma}$ weighting each of the two contributions. Additionally if and only if $\widehat{\mathbf{x}}$ is a fixed point, then $\operatorname{prox}_{\check{\gamma} f}(\widehat{\mathbf{x}}) = \widehat{\mathbf{x}}$.

This effect is shown on Fig. 2.5b and the parameter $\check{\gamma}$ acts similarly to the step size of the GD method to converge over the minimizer, although the effect is extended outside the boundaries of the domain of f.

The proximal operator is also useful to address problems defined on complementary domains, such as in the primal/dual framework. Formally, for a topological space characterized by a scalar product $\langle \cdot, \cdot \rangle : \mathtt{E}_x \times \mathtt{E}_u \to \mathbb{R}^+$, it is possible to define the **Fenchel conjugate** $f^\star : \mathtt{E}_u \to \mathbb{R}$ as:

$$f^\star(\mathbf{u}) := \sup \left\{ \langle \mathbf{u}, \mathbf{x} \rangle - f(\mathbf{x}) : \mathbf{x} \in \mathtt{E}_x \right\} \tag{2.3.9}$$

and the following relation, known as **Moreau's identity** [169], holds:

$$\operatorname{prox}_{\check{\gamma} f^\star}(\mathbf{x}) = \mathbf{x} - \check{\gamma} \operatorname{prox}_{\frac{f}{\check{\gamma}}} \left(\frac{\mathbf{x}}{\check{\gamma}} \right) . \tag{2.3.10}$$

This identity allows to express some proximal operators in close form; one common example is in the case $f(\cdot) = \| \cdot \|_1$ is the ℓ_1-norm, for which $\operatorname{prox}_{\check{\gamma} f^\star}(\mathbf{x})$ is the projection of $\mathbf{x}$ over the ℓ_∞ ball, which is equivalent of a soft thresholding:

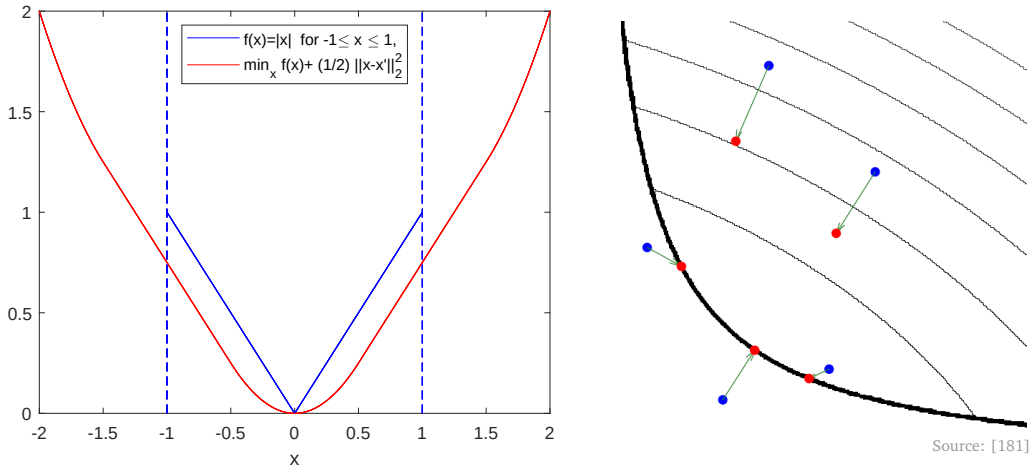
$$\operatorname{prox}_{\check{\gamma} f^\star}(\mathbf{x}) = \operatorname{proj}\{\|\mathbf{x}\|_\infty \leq \check{\gamma}\} = \max(\min(\mathbf{x}, \check{\gamma}), -\check{\gamma}) , \tag{2.3.11}$$

or for the case $f(\cdot) = \| \cdot \|_\infty$, where $\operatorname{prox}_{\check{\gamma} f^\star}(\mathbf{x}) = \operatorname{proj}\{\|\mathbf{x}\|_1 \leq \check{\gamma}\}$ is the projection over an ℓ_1-ball, which can also be implemented efficiently [49].

Loris-Verhoeven algorithm

Let us consider a special case of eq. (2.3.1), in which the function $h(\mathbf{x}) = f(\mathbb{A}(\mathbf{x})) + \langle \mathbf{x}, \mathbf{c} \rangle$ is Fréchet differentiable with a $\check{\beta}$-Lipschitz continuous gradient ∇_h for a certain $\check{\beta} > 0$. We are interested in solving the following **primal problem**:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} h(\mathbf{x}) + g(\mathbb{L}(\mathbf{x})) , \tag{2.3.12}$$



(a) Moreau envelope **(b)** Evaluation of the proximal operator

Fig. 2.5. In Fig. 2.5a, an example of the expression of the Moreau envelope (in red) applied to a function $f(x) = |x|$ defined in the domain $\mathtt{E}_x \equiv [-1, 1]$ (in blue). In Fig. 2.5b, the effect (in red) of applying the proximal operator to points (in blue) in the domain $\mathbb{R}^2$, when f is compact and whose boundary is represented by a thick black line. Points outside $\mathtt{E}_x$ are projected over the boundary of the domain, while points within the boundary are shifted towards the minimum of f.

for which we make use in this section of the so-called property of **duality**. This approach consists in defining a dual variable $\mathbf{u}$ and its associate **dual problem**:

$$\widehat{\mathbf{u}} = \arg\min_{\mathbf{u}} \; h^\star(-\mathbb{L}^*(\mathbf{u})) + g^\star(\mathbf{u})\,, \tag{2.3.13}$$

which allows to get rid of the dependency of the operator $\mathbb{L}$ in the function g. Here, $h^\star$ and $g^\star$ are the Fenchel conjugate of h and g, respectively, while $\mathbb{L}^*$ is the adjoint of $\mathbb{L}$.

The **primal-dual forward-backward iterations**, also known as **Loris-Verhoeven algorithm** [148] allows to solve both problems jointly. The update rule is in charge to provide a more refined estimation $\mathbf{u}^{(q+1)}$ to the dual variable, as well as one $\mathbf{x}^{(q+1)}$ for the primal one:

$$\begin{cases} \mathbf{u}^{(q+\frac{1}{2})} &= \mathrm{prox}_{\check{\sigma} g^\star}\left(\mathbf{u}^{(q)} + \check{\sigma}\,\mathbb{L}\left(\mathbf{x}^{(q)} - \check{\tau}\left(\nabla_h(\mathbf{x}^{(q)}) + \mathbb{L}^*(\mathbf{u}^{(q)})\right)\right)\right) \\ \mathbf{x}^{(q+1)} &= \mathbf{x}^{(q)} - \check{\rho}\check{\tau}\left(\nabla_h(\mathbf{x}^{(q)}) + \mathbb{L}^*(\mathbf{u}^{(q+\frac{1}{2})})\right) \\ \mathbf{u}^{(q+1)} &= \mathbf{u}^{(q+1)} + \check{\rho}\left(\mathbf{u}^{(q+\frac{1}{2})} - \mathbf{u}^{(q)}\right)\,. \end{cases} \tag{2.3.14}$$

The parameters $\check{\sigma} > 0$ and $\check{\tau} > 0$ must verify the condition $\check{\sigma}\check{\tau} \leq 1/\| \mathbb{L} \|^2_{op}$ to allow for convergence. $\check{\tau}$ is a free parameter that defines the speed of the convergence and is typically chosen to be $\check{\tau} = 1/\| \mathbb{A} \|^2_{op}$ [148]. $\check{\rho}$ is the so-called over-relaxation parameter and the convergence is assured for $\check{\rho} \in (0, 2)$.

These iterations can be expressed for the basic problem in eq. (2.3.1) as follows:

$$\begin{cases} \mathbf{u}^{(q+\frac{1}{2})} & = \text{prox}_{\check{\sigma} g^\star} \left(\mathbf{u}^{(q)} + \check{\sigma} \mathbb{L} \left(\mathbf{x}^{(q)} - \check{\tau} \left(\mathbb{A}^* (\nabla_f (\mathbb{A}(\mathbf{x}^{(q)}))) + \mathbf{c} + \mathbb{L}^* (\mathbf{u}^{(q)}) \right) \right) \right) \\ \mathbf{x}^{(q+1)} & = \mathbf{x}^{(q)} - \check{\rho}\check{\tau} \left(\mathbb{A}^* (\nabla_f (\mathbb{A}(\mathbf{x}^{(q)}))) + \mathbf{c} + \mathbb{L}^* (\mathbf{u}^{(q+\frac{1}{2})}) \right) \\ \mathbf{u}^{(q+1)} & = \mathbf{u}^{(q)} + \check{\rho} \left(\mathbf{u}^{(q+\frac{1}{2})} - \mathbf{u}^{(q)} \right) . \end{cases} \tag{2.3.15}$$

The Loris-Verhoeven algorithm is well suited to solve problems in the form of eq. (2.3.2a) (or eq. (2.3.2c)) simply by substituting:

$$h(\mathbf{x}) = f(\mathbb{A}(\mathbf{x})) + \langle \mathbf{x}, \mathbf{c} \rangle = \frac{1}{2} \| \mathbb{A}(\mathbf{x}) \|^2 + \langle \mathbf{x}, \mathbf{c} \rangle \tag{2.3.16}$$

as the gradient of $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$ is the identity operator $\nabla_f(\mathbf{x}) = \mathbf{x}$. The Fréchet gradient in this scenario has form: $\nabla_h(\mathbf{x}) = \mathbb{A}^*(\mathbb{A}(\mathbf{x}) - \mathbf{y})$, for which we can solve a variety of common cost function minimizations by substituting this value in eq. (2.3.14) accordingly.

Chambolle-Pock algorithm

In the more general case in which f in eq. (2.3.1) is not differentiable, but it still possible to define a proximal operator, the problem can be solved by the **generalized Chambolle-Pock** algorithm;

$$\begin{cases} \mathbf{u}^{(q+\frac{1}{2})} = \text{prox}_{\check{\sigma} g^\star} \left(\mathbf{u}^{(q)} + \check{\sigma} \mathbb{L} \left(\mathbf{x}^{(q)} \right. \right. \\ \qquad\qquad \left. \left. -\check{\tau} \left(\mathbb{A}^* \left(\mathbf{v}^{(q)} \right) + \mathbf{c} + \mathbb{L}^* \left(\mathbf{u}^{(q)} \right) \right) \right) \right) , & \text{(2.3.17a)} \\ \mathbf{x}^{(q+\frac{1}{2})} = \mathbf{x}^{(q)} - \check{\tau} \left(\mathbb{A}^* \left(\mathbf{v}^{(q)} \right) + \mathbf{c} + \mathbb{L}^* \left(\mathbf{u}^{(q+\frac{1}{2})} \right) \right) , & \text{(2.3.17b)} \\ \mathbf{v}^{(q+\frac{1}{2})} = \underset{\eta f^\star}{\text{prox}} \left(\mathbf{v}^{(q)} + \eta \, \mathbb{A} \left(2\mathbf{x}^{(q+\frac{1}{2})} - \mathbf{x}^{(q)} \right) \right) , & \text{(2.3.17c)} \\ \mathbf{u}^{(q+1)} = \mathbf{u}^{(q)} + \check{\rho} \left(\mathbf{u}^{(q+\frac{1}{2})} - \mathbf{u}^{(q)} \right) , & \text{(2.3.17d)} \\ \mathbf{x}^{(q+1)} = \mathbf{x}^{(q)} + \check{\rho} \left(\mathbf{x}^{(q+\frac{1}{2})} - \mathbf{x}^{(q)} \right) , & \text{(2.3.17e)} \\ \mathbf{v}^{(q+1)} = \mathbf{v}^{(q)} + \check{\rho} \left(\mathbf{v}^{(q+\frac{1}{2})} - \mathbf{v}^{(q)} \right) , & \text{(2.3.17f)} \end{cases}$$

whose convergence is assured in the case $\check{\sigma}\check{\tau} \leq 1/\|L\|_{op}^2$ and $\eta\check{\tau} \leq 1/\|A\|_{op}^2$; typically one can just impose $\check{\tau}$ as a single tuning parameter and choose $\check{\sigma}$ and η such that the above inequalities become equalities.

Let us assume $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$ is a quadratic function and let us impose $\check{\eta} = 1$. Under these constraints, we can prove that the third term (2.3.17c) in the generalized Chambolle-Pock expression can be rewritten as:

$$\mathbf{v}^{(q+\frac{1}{2})} = \frac{1}{2}\mathbf{v}^{(q)} + \mathbb{A}\left(\mathbf{x}^{(q+\frac{1}{2})} - \mathbf{x}^{(q)}\right), \tag{2.3.18}$$

which makes use of the property of quadratic functions that $\text{prox}_f(\mathbf{x}) = \text{prox}_{f^\star}(\mathbf{x}) = \frac{1}{2}\mathbf{x}$. Consequently, if we initialize the algorithm with $\mathbf{v}^{(0)} = \mathbb{A}\left(\mathbf{x}^{(0)}\right)$, the update becomes $\mathbf{v}^{(q)} = \mathbb{A}(\mathbf{x}^{(q)})$ for all $q \in \mathbb{N}$. It can be shown [50] that the Chambolle-Pock is a more general version of the widespread alternating direction method of multipliers (ADMM) algorithm.

2.3.3 Nonlinear regression

It is a common occurrence in inverse problems to find situations where the model $\mathbb{A} : \mathbb{R}^{N_x} \to \mathbb{R}^{N_y}$ is nonlinear.

A possible strategy in this case consists in linearizing such model through successive **Taylor decompositions** around the current estimation, which is applicable to the case in which $\mathbb{A}$ is differentiable. Let $\nabla_{\mathbb{A}}^{(q)}$ denote the gradient of $\mathbb{A}$ evaluated at a generic point $\mathbf{x}^{(q)} \in \mathtt{E}_x$, which can be obtained as $\{\frac{\partial\,\mathbb{A}(\mathbf{x}))}{\partial x_i}\Big|_{\mathbf{x}=\mathbf{x}^{(q)}}\}_{i\in[1,\dots,N_x]}$ where x_i denotes the i-th component of $\mathbf{x}$, then the Taylor decomposition of $\mathbb{A}(\mathbf{x})$ truncated at the first order derivative is given by:

$$\mathbb{A}(\mathbf{x}) \approx \mathbb{A}\left(\mathbf{x}^{(q)}\right) + \nabla_{\mathbb{A}}^{(q)}\left(\mathbf{x} - \mathbf{x}^{(q)}\right). \tag{2.3.19}$$

If $\mathbf{x}^{(q)}$ denotes the estimation of the desired quantity at the q-th iteration, the updated estimation $\mathbf{x}^{(q+1)}$ at the successive step can thus be obtained by solving a classical linear inverse problem; the procedure is repeated until convergence is reached.

More specifically, in the quadratic case of eq. (2.3.2a), the $(q+1)$-th iteration of the algorithm is equivalent to solve:

$$\mathbf{x}^{(q+1)} = \arg\min_{\mathbf{x}\in\mathbb{E}_x} f\left(\mathbb{A}\left(\mathbf{x}^{(q)}\right) - \mathbf{y} + \nabla_{\mathbb{A}}^{(q)}\left(\mathbf{x} - \mathbf{x}^{(q)}\right)\right) + g'(\mathbf{x}) \tag{2.3.20a}$$

$$= \arg\min_{\mathbf{x}\in\mathbb{E}_x} \frac{1}{2}\left\|\mathbb{A}\left(\mathbf{x}\right) - \mathbf{y} + \nabla_{\mathbb{A}}^{(q)}\left(\mathbf{x} - \mathbf{x}^{(q)}\right)\right\|_2^2 + \check{\lambda} g\left(\mathbb{L}\left(\mathbf{x}\right)\right) \tag{2.3.20b}$$

$$= \arg\min_{\mathbf{x}\in\mathbb{E}_x} \frac{1}{2}\left\|\nabla_{\mathbb{A}}^{(q)}\left(\mathbf{x}\right) - \left(\nabla_{\mathbb{A}}^{(q)}\left(\mathbf{x}^{(q)}\right) - \mathbb{A}\left(\mathbf{x}^{(q)}\right) + \mathbf{y}\right)\right\|_2^2 + \check{\lambda} g\left(\mathbb{L}\left(\mathbf{x}\right)\right) \tag{2.3.20c}$$

$$= \arg\min_{\mathbf{x}\in\mathbb{E}_x} \frac{1}{2}\left\|\nabla_{\mathbb{A}}^{(q)}\left(\mathbf{x}\right)\right\|_2^2 + \left\langle \nabla_{\mathbb{A}}^{(q)*}\left(\mathbf{x}\right), \nabla_{\mathbb{A}}^{(q)}\left(\mathbf{x}^{(q)}\right) - \mathbb{A}\left(\mathbf{x}^{(q)}\right) + \mathbf{y}\right\rangle + \check{\lambda} g(\mathbb{L}(\mathbf{x})) \tag{2.3.20d}$$

$$= \arg\min_{\mathbf{x}\in\mathbb{E}_x} \frac{1}{2}\left\|\nabla_{\mathbb{A}}^{(q)}\left(\mathbf{x}\right)\right\|_2^2 + \left\langle \mathbf{x}, -\nabla_{\mathbb{A}}^{(q)*}\left(\nabla_{\mathbb{A}}^{(q)}\left(\mathbf{x}^{(q)}\right) - \mathbb{A}\left(\mathbf{x}^{(q)}\right) + \mathbf{y}\right)\right\rangle + \check{\lambda} g(\mathbb{L}(\mathbf{x}))\,, \tag{2.3.20e}$$

so that the update $\mathbf{x}^{(q+1)}$ can be performed, for any $g \in \Gamma(0)$, via the Loris-Verhoeven algorithm with the constant $\mathbf{c} = -\nabla_{\mathbb{A}}^{(q)*}\left(\nabla_{\mathbb{A}}^{(q)}\left(\mathbf{x}^{(q)}\right) - \mathbb{A}\left(\mathbf{x}^{(q)}\right) + \mathbf{y}\right)$.

Of course, if less restrictive objective functions are used, such as in the case that g is equally zero, simpler updating methods can be employed. I.e. if g is uniformly equal to zero, the update can be performed with the well known **Gauss-Newton algorithm (GNA)** [78], for which eq. (2.3.20d) has a closed form solution:

$$\mathbf{x}^{(q+1)} = \mathbf{x}^{(q)} - \left(\nabla_{\mathbb{A}}^{(q)*}\nabla_{\mathbb{A}}^{(q)}\right)^{-1}\nabla_{\mathbb{A}}^{(q)*}(\mathbb{A}\left(\mathbf{x}^{(q)}\right) - \mathbf{y}) \tag{2.3.21}$$

The same problem can also be addressed by the GD algorithm, previously described in Section 2.3.1 and the update as:

$$\mathbf{x}^{(q+1)} = \mathbf{x}^{(q)} - \check{\gamma}^{(q)}\nabla_{\mathbb{A}}^{(q)*}(\mathbb{A}\left(\mathbf{x}^{(q)}\right) - y) \tag{2.3.22}$$

where $\check{\gamma}^{(q)} < 2/\|\nabla_{\mathbb{A}}^{(q)}\|_{op}$ is a used-defined parameter deciding the rate of convergence.

Another popular variant, commonly known as **Levenberg–Marquardt algorithm** [142, 156], introduces an extra penalization terms to impose the solution to not be too distant from the current estimation, by imposing $g'(\mathbf{x}) = \check{\lambda}\|\mathbf{x} - \mathbf{x}^{(q)}\|_2^2$. This work is not meant to provide a comprehensive dissertation on the vast topic of nonlinear optimization; for a more in depth discussion, including techniques to avoid

the nonconvexity of the objective function, the interested reader may refer to [22, 77].

3 Signal processing of multimodal data

3.1 Introduction

Multimodal data defines a class of remotely sensed acquisitions that may differ in imaging mechanism, spatial resolution, and coverage. Due to the improvement in the specificity of different sensor technology, multimodality allows for a wide degree of diversity of information of a given scene, whose complementarity can be exploited in various applications, e.g., precision agriculture, urban planning, and disaster responses [52].

Data fusion defines the process of of integrating multiple data sources to produce relevant and consistent information to the final user. Various applications are available that exploit the different characteristics of the image, which may include either differences in elevation, such as LIDAR systems, or in structure, such as in optical and synthetic aperture radar (SAR) system, or in the amount of available channels, such as in the case of multispectral and hyperspectral data [52]. Data fusion is not limited to the case of multimodal acquisitions, as different information may be captured by sensors at different angles of view or at different times [4, 191].

In this chapter, we consider scenes captured with imaging systems characterized by different spectral and spatial resolutions. The goal of a **sharpening** procedure is to generate a synthetic image featuring the best resolutions of each of the two sensor technologies.

The multimodality of the data is associated with a set of challenges to address, such as spectral/spatial variations, missing information, and sensor-specific issues. In particular, the necessity to link a diversified set of information to the same portion of the scene requires specific procedures, such as scaling and co-registration.

Additionally, in this thesis, the concept of multimodality is also considered extended to sensors with an overlaid **color filter arrays (CFAs)**, whose acquisitions can be

seen as multichannel samples with gaps in specific portions of the scene, which require a procedure of **demosaicing** to infer information where it is not available.

In this chapter we provide a brief introduction to the concepts related to multimodality, defining the different sources employed in this work (Section 3.2), the preprocessing required to coordinate different data sources (Section 3.4), image fusion (Section 3.5), and demosaicing techniques (Section 3.7). We also introduce the standard procedures for the validation of the reconstructed product (Section 3.8).

3.2 Multimodal acquisition systems

This section's goal is to describe the acquisition systems that can describe the multimodal data sources that are employed in this work, which in this context consist in either multiresolution or multichannel data.

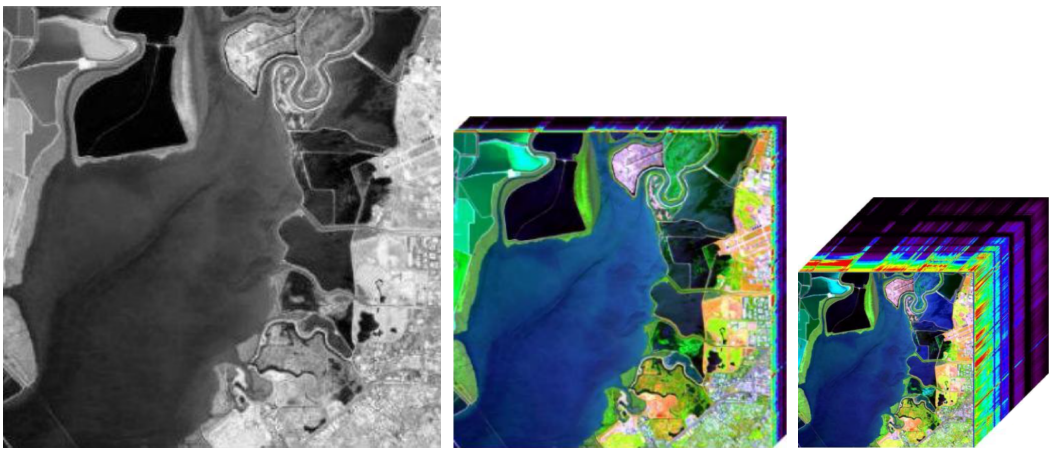
3.2.1 Multiresolution data

In the domain of remote sensing, it is quite common for a satellite platform to accommodate sensors with different technologies. In terms of spatial and spectral resolution, which were defined in Section 1.2.1, some specialized sensing technologies were developed which target a specific combination of those resolutions.

However, both technological and physical constraints do not allow to manufacture sensors capable of simultaneously capturing images with simultaneously very high spectral and spatial characteristics. In fact, if the noise level is kept constant, the relevant signal's energy can be raised by either targeting a larger field of view (FoV) (hence reducing the spatial resolution) or from a larger band coverage (hence reducing the spectral diversity).

In this regard, it is typical to distinguish between three different categories of sensors, listed below in descending order of spatial resolution:

- **Panchromatic (PAN)**: characterized by a single band (monochromatic) acquisition covering a wide spectral domain,
- **Multispectral (MS)**: characterized by a few (typically 3 to 15) spectral channels,
- **Hyperspectral (HS)**: characterized by hundreds of narrow and contiguous spectral bands.



(a) PAN (Panchromatic) **(b)** MS (Multispectral) **(c)** HS (Hyperspectral)

Fig. 3.1. Differences between different spectral and spatial resolution for multi-resolution imagery, with the depth proportional to the amount of spectral channels. The portrayed scale ratio between each class is just for illustration purposes, as it is typically bigger in commercial bundles.

Many high quality commercial sensors, such as QuickBird, IKONOS, WorldView-2 and WorldView-3 are equipped with a platform of PAN and MS sensors working in the optical range of wavelengths which simultaneously acquire the scene [@9].

The availability of HS data with a simultaneous matched multi-modal acquisition is somewhat rarer: the now-discontinued Earth Observing-1 provided HS imagery together with a matched PAN and MS sensor (the latter at the same spatial resolution of the HS). In the last years, another option has been made available by the spacecraft Hyperspectral Precursor of the Application Mission (PRISMA), whose platform is equipped with both an HS sensor and a medium resolution PAN camera [84].

Although the terminology and the application of this work are mostly targeted to remote sensing imaging systems, the theory is applicable with minor adjustments also to commercial cameras dedicated to the acquisition of natural images, even for portable devices that feature monochrome sensors (e.g. some smartphone models by Huawei) [144].

In this thesis, we typically consider two types of multiresolution sources, which we define as:

- **High resolution image (HRI)**: characterized by relatively higher spatial resolution and lower spectral resolution,
- **Low resolution image (LRI)**: characterized by relatively lower spatial resolution and higher spectral resolution,

and the goal of a **sharpening algorithm** is to produce a synthetic image with the spatial resolution of the HRI and the spectral resolution of the LRI. In the most widespread setup, known as **pansharpening**, the role of the HRI is played by the PAN, while the LRI is a MS image [224]. Other configurations are also available, such as PAN/HS (sometimes known as hypersharpening) [147] and MS/HS fusion [239].

3.2.2 Multi-channel acquisitions

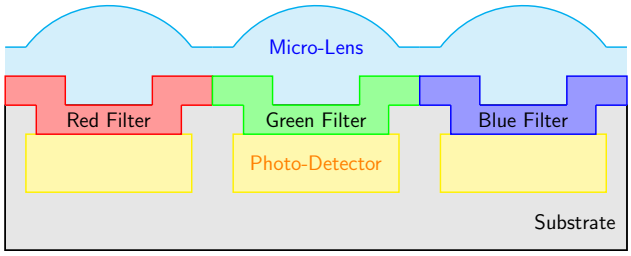
The most typical technologies of photodetectors for low to mid-cost commercial cameras only allow for the characteristic spectral response allowed by the photodetector technology, which does not allow to distinguish spectral information. To retrieve spectral information, most commercial devices are equipped with a set of spectral filters that limit the spectral response of the photodetector.

Two main strategies are available for filtering the acquisitions: in the **multishot** setup, the filter is placed as leading optic, so that the filtering effect is applied to the whole focal plane, while in the **snapshot** setup, separate filters are overlaid over each photosensor. Some examples of both approaches are described in the following section.

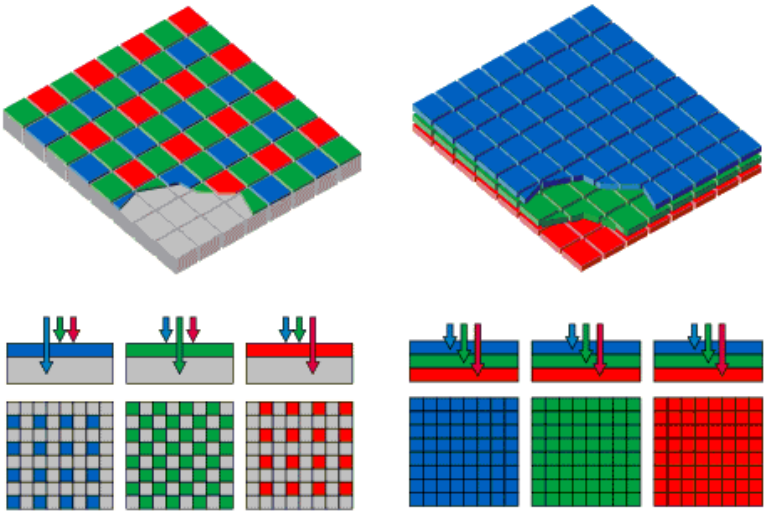
Snapshot setups

The most common solution for multichannel acquisitions, which makes up for the great majority of available commercial cameras, consists in assigning a specialized optical filter to each photodetector on the focal plane, as shown in Fig. 3.2a. Those dedicated filters are taken from a set of given sensitivity, whose response defines the channel, and arranged through patterns, commonly known as CFA, in the case of red green blue (RGB) cameras, or multispectral filter array (MSFA), in the most general case. The most widespread arrangement, known as Bayer pattern [21], is shown in Fig. 3.2b, and a more in depth analysis of the literature of such arrangements is provided in Section 3.6.2.

In the raw acquisition, each pixel holds information relative to a single channel, so the full spectral information is incomplete. This raw product can be modeled as a degraded version of the full color information, obtained by an element by element multiplication with a binary mask, as described in Section 3.6.1. The **demosaicing** (or **demosaicking**) is the field of study for the algorithms and techniques to recover



(a) CFA Physical Design



Source: SIGMA Corporation [@4]

Source: SIGMA Corporation [@4]

(b) CFA Operating Principle

(c) Foveon Operating Principle [@4]

Fig. 3.2. Physical scheme and operating principle of snapshot-based spectral filters.

the full spectral information from the available samples. An overview of the literature on demosaicing is given in Section 3.7.

A more recent snapshot technique was made available by the **Foveon X3 Sensor** technology designed by Foveon Inc. (now Sigma corporation [@4]). Its principle of operation, shown in Fig. 3.2a, consists of a matrix of photosites composed of three stacked photodiodes with different spectral sensitivity. The RGB information is separated thanks to different silicon penetration depths, which, as opposed to what happens in CFA-based structures, allows to capture the information of all channels on a single photosite. However, the cross-talk between each layer may cause issues in terms of color accuracy.

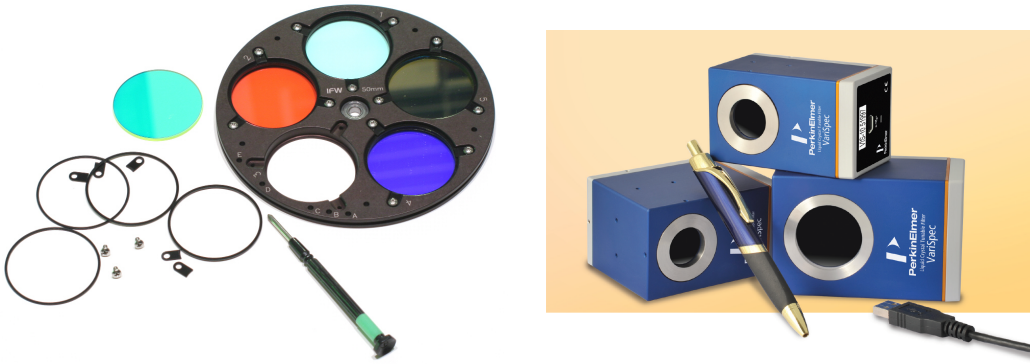
Multishot setups

In the multishot configuration, the filter is applied simultaneously on all the sensors of the focal plane. This setup requires a series of acquisitions, each of which corresponds to a specific spectral response.

The filter is positioned before the leading optic and different technologies are available to change the spectral response of the filter itself. Compared to the snapshot setup, which gives an intrinsically degraded product compared to the potentially achievable spatial resolution, no such limitation is present in the multishot setup. However, some time delay is introduced in order to switch filter between multiple acquisitions, making this setup sensitive to temporal changes in the scene. Consequently, this setup is limited to applications whose target is unaltered across different shots, i.e. if the scene is static and the relative position and angle of view of the instrument is kept constant. Between each acquisition, the filters can be switched either through mechanical devices, such as in the case of **dichroic filters** mounted on wheels (Fig. 3.3a) or sliders, or electronically, such as in the case of **tunable filters** (Fig. 3.3b).

This first setup is quite common for applications such as astronomical observations: the scene itself changes slightly, but the shots can be co-registered in post-processing. The second setup is often used for benchmarks, such as some publicly available HS image bundles: the CAVE dataset by the Columbia University [@2, 237] or the ones by Harvard [@5, 38], and by TokyoTech [@13, 167].

An alternative setup is given by the **colour co-site sampling** [141], which defines a **micro-scanning** of captures taken with a conventional camera. The micro-scanning is obtained by moving the camera along the scene so that the focal plane shifts



Source: Optec Inc. Source: Perkin Elmer

(a) Dichroic filter wheel **(b)** Varispec LCD tunable filter

Fig. 3.3. Different solutions for leading optic-type spectral filters.

by one pixel in one direction for every consecutive shot. If the scene is stable and the alignment across each shot is carefully controlled, this allows to get the color information over each pixel.

3.3 Notation

In the following, we denote with an uppercase $\mathcal{U}$ a 3-way tensor representation of an image, whose dimensions represent in order the two spatial and the spectral dimensions. A specific letter is given between square brackets as superscript to label the assigned function of each denotation; specifically we denote with:

- $\mathcal{U}^{[x]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_b}$ the ideal image to reconstruct;
- $\mathcal{U}^{[p]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_{b_p}}$ the HRI, which e.g. can be a PAN if $N_{b_p} = 1$;
- $\mathcal{U}^{[m]} \in \mathbb{R}^{(N_{i_1}/\rho) \times (N_{i_2}/\rho) \times N_b}$ the LRI, which typically represents a MS image;
- $\mathcal{U}^{[h]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_b}$ a compressed acquisition mask;
- $\mathcal{U}^{[b]} \in \mathbb{R}^{N_{d_1} \times N_{d_2} \times N_b}$ a generic blurring filter.

Here, N_{i_1} and N_{i_2} are the amount of pixel rows and columns, respectively, while N_b and $N_{b_p} < N_b$ are the amount of channels of the LRI and the HRI, respectively. Finally, ρ denotes the scale ratio between the HRI and LRI, which will be assumed to be the same in the along-track and across-track directions.

The k-th channel of any of the described variables, which can be seen with the tensor formalism as the k-th frontal slice, is denoted with $\mathbf{U}_{::k}$, so i.e. $\mathbf{U}^{[m]}_{::k}$ is the k-th

channel of the LRI. A given pixel of said channel is denoted with a lowercase letter, i.e. $u^{[m]}_{i_1,i_2,k}$ is the pixel at the (i_1, i_2)-th position of $\mathbf{U}^{[m]}_{::k}$.

For each of the listed quantities, the same coefficients can also be arranged in lexicographic order; this operation, denoted with $\mathrm{matr}(\cdot)$, consists in representing each frontal slice as a vector array, and then concatenating the results along each row. Its result is denoted with the representative letter of the associated variable, in bold uppercase, i.e. $\mathbf{X} = \mathrm{matr}\left(\boldsymbol{\mathcal{U}}^{[x]}\right)$ or $\mathbf{M} = \mathrm{matr}\left(\boldsymbol{\mathcal{U}}^{[m]}\right)$. Formally, this is equivalent to assign $x_{i_1+N_{i_1}(i_2-1),k} = u^{[x]}_{i_1,i_2,k}$; the inverse reshaping operation is denoted with matr^{-1}, i.e. $\boldsymbol{\mathcal{U}}^{[x]} = \mathrm{matr}^{-1}(\mathbf{X})$.

With this formalism, we obtain the following lexicographic representation of the same variables:

- $\mathbf{X} \in \mathbb{R}^{N_i \times N_b}$ for the ideal product to reconstruct,
- $\mathbf{M} \in \mathbb{R}^{\frac{N_i}{\rho^2} \times N_b}$ for the LRI,
- $\mathbf{P} \in \mathbb{R}^{N_i \times N_{b_p}}$ for the HRI,
- $\mathbf{H} \in \mathbb{R}^{N_i \times N_b}$ for the mask,
- $\mathbf{B} \in \mathbb{R}^{N_d \times N_b}$ for the filter.

where $N_i = N_{i_1} N_{i_2}$ and $N_d = N_{d_1} N_{d_2}$. With this formalism, the k-th channel is described as a vector column with a lowercase bold letter, i.e. $\mathbf{m}_{:k}$ is the k-th channel of $\mathbf{M}$, and its i-th element is denoted with the corresponding non-bold letter m_{ik} (or sometimes $m_{i,k}$).

If the $\mathrm{matr}(\cdot)$ operator is applied to an image already expressed in lexicographic order, the effect is to concatenate all bands together into a single column vector; so that, i.e., $\mathbf{v}^{[x]} = \mathrm{matr}(\mathbf{X})$ is a vector array such that $\mathbf{v}^{[x]} \in \mathbb{R}^{(N_i N_b)}$. The operation is however denoted as $\mathbf{v}^{[x]} = \mathrm{vec}(\mathbf{X})$, as the result is a vector, and is formally defined as: $v^{[x]}_{i+(k-1)N_i} = x_{i,k}$.

Given a generic vector $\mathbf{a} = \{a_i\}_{i \in [1,\dots,N_i]}$, $\overline{\mathbf{a}} = \frac{1}{N_i} \sum_{i=1}^{N_i} a_i$ denotes its mean and $\mathrm{std}(\mathbf{a}) = \frac{1}{N_i - 1} \sqrt{\sum_{i=1}^{N_i} (a_i - \overline{\mathbf{a}})^2}$ its standard deviation (STD).

When an image is extended or decimated by a factor of ρ, this is denoted with a $\uparrow$ and $\downarrow$ respectively, so that $\mathbf{M}^{\uparrow} = \mathrm{matr}\left(\boldsymbol{\mathcal{U}}^{[m^{\uparrow}]}\right)$ stands for the product of a spatial extension by a factor of ρ applied to $\mathbf{M}$ and $\mathbf{P}^{\downarrow} = \mathrm{matr}\left(\boldsymbol{\mathcal{U}}^{[p^{\downarrow}]}\right)$ stands for a decimation of the HRI $\mathbf{P}$ by the same factor. A low pass filtered version of an image is denoted with a tilde, so that $\widetilde{\mathbf{P}} = \mathrm{matr}\left(\boldsymbol{\mathcal{U}}^{[\widetilde{p}]}\right)$ is a low frequency version of $\mathbf{P}$.

Those operations can be also combined together so that $\widetilde{\mathbf{M}}^{\uparrow}$ is an upsampled version of $\mathbf{M}$ (a cascade of an extension and a low pass filtering), and $\widetilde{\mathbf{P}}^{\downarrow}$ is a downsampled version of $\mathbf{P}$ (a cascade of a low pass filtering and a decimation).

The symbols $\odot$ and $\otimes$ denote the Hadamard (element-wise) and Kronecker product, respectively, while $(\cdot)^{\boxdot}$ denotes the masked version of a variable, i.e. $\mathbf{X}^{\boxdot} = \mathbf{X} \odot \mathbf{H}$.

The operator $\circledast\!\!\circledast$ denotes a circular convolution product in the spatial domain, when it is applied between images in their natural representation, and its rigorous expression is given in eq. (A.1.3). However, as we commonly perform this operation between images in their lexicographic order we define a shorthand operator $\circledast$, such that:

$$\mathbf{x}_{:k} \circledast \mathbf{b}_{:k} = \operatorname{matr}\left(\operatorname{matr}^{-1}(\mathbf{x}_{:k}) \circledast\!\!\circledast \operatorname{matr}^{-1}(\mathbf{b}_{:k})\right) = \operatorname{matr}\left(\mathbf{U}^{[x]}_{::k} \circledast\!\!\circledast \mathbf{U}^{[b]}_{::k}\right) . \quad (3.3.1)$$

Additionally, the $[\cdot;\cdot]$ and $[\cdot,\cdot]$ operators respectively stand for column and row concatenation, while $[\cdot,\cdot]_p$ is a generic concatenation along the p-th dimension.

The operations of addition, difference, element by element multiplication $(\cdot \odot \cdot)$ and division $(\frac{\cdot}{\cdot})$ between arrays with different shapes assume that the array with shorter dimension is broadcast to match the dimension of the longer one. I.e., $\mathbf{m}_{:k} - \overline{\mathbf{m}}_{:k}$ denotes that the mean of $\mathbf{m}_{:k}$ is subtracted from each its elements.

$\mathbf{0}_{[N_1 \times N_2]}$ and $\mathbf{1}_{[N_1 \times N_2]}$ denote a $N_1 \times N_2$ matrices of all zeros and all ones, respectively. Given a generic vector $\mathbf{w}$, $\|\mathbf{w}\|_1$, $\|\mathbf{w}\|_2$ its ℓ_2-norm, $\operatorname{std}(w)$ is its standard deviation, while $\langle \mathbf{w}, \mathbf{w}' \rangle$ and $\operatorname{cov}(\mathbf{w}, \mathbf{w}')$ are its scalar product and its covariance with a vector $\mathbf{w}'$ of the same size, respectively. For any matrix $\mathbf{W}$, $\|\mathbf{W}\|_F$ denotes its Frobenius norm.

Given a generic 3-way tensor $\boldsymbol{\mathcal{U}}$ the notation $\|\boldsymbol{\mathcal{U}}\|_{p_1 p_2 p_3}$ denotes that the norm ℓ_{p_1} is applied over the third dimension, ℓ_{p_2} over the second one, ℓ_{p_3} over the first one, in this order. I.e, this notation was already employed for the collaborative total variation (CTV) in Section 2.2.3, for which each of the norms was applied to the dimensions associated with the gradients, to the channels and to the pixels, in this order. Similarly, $\|.\|_{S_{p_1}\ell_{p_2}}$ similarly denotes a nuclear (also known as Shatten) norm S_{p_1} jointly applied over the second and third dimensions, followed by an ℓ_{p_2}-norm operating on the first dimension.

The operators $\max(\cdot,\cdot)$ and $\min(\cdot,\cdot)$ denote the maximum and minimum between two scalar arguments, respectively, while $\max(\mathbf{w})$ and $\min(\mathbf{w})$ denote the maximum and minimum element of the vector $\mathbf{w}$, respectively.

3.4 Image preprocessing

The multimodality of the data introduces challenges related to the different characteristics of the acquisitions, which have to be related to one another. Many of such operation have often specific features which depend on the joint characteristics of each data source, however some basic operations, such as rescaling and co-registration, are commonly employed in many situations, and are the topic of the following sections.

3.4.1 Image scaling

In the context of image processing, we often face the situation in which it is necessary to either increase or decrease the scale of an image; in our context, e.g., it could be useful to resize a LRI to the scale of a HRI or to find a low-resolution equivalent of the HRI.

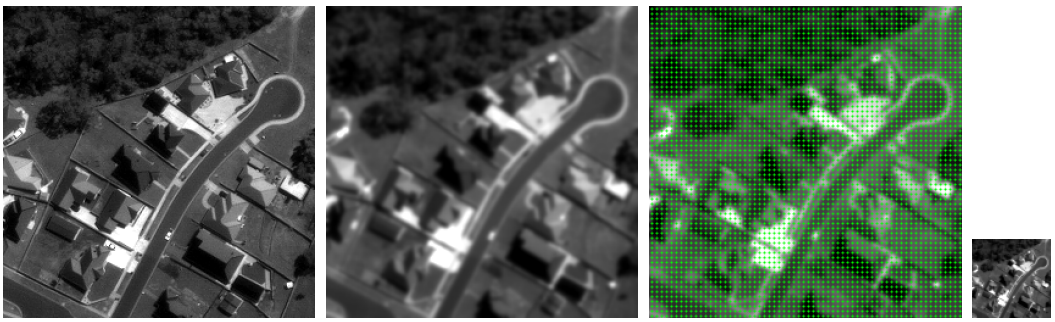
While different choices are available, we consider image resizing as a cascade of operations over monodimensional arrays; specifically, each of the described operation is first applied column by column and then row by row.

For the scale reduction, the main requirement is to avoid the aliasing effects which arise when the Shannon-Nyquist theorem is not verified, which in the context of image processing, typically shows as **moiré patterns**, that is, interferences in the image due to quickly varying intensity levels.

We define as **downsampling** by an integer factor ρ the cascade of the following two operations:

- **Low-pass filtering**: this operation allows to remove the high frequency components to satisfy the Shannon-Nyquist,
- **Decimation**: taking every ρ-th sample.

The main requirement for the low pass filter (LPF) is to have a monolateral cut-off frequency of $1/(2\rho)$, to avoid any overlap between the signals' replicas which arise when decimating the image, whose spectra are separated by a digital frequency of $1/\rho$. Various strategies can be employed for the choice of the LPF, such as a stationary wavelet transform (SWT) [94], a Gaussian filter (e.g. matching the modulation transfer function (MTF) of the sensor [5]) or classical finite impulse response (FIR) designs such as the ones based on Butterworth or Chebyshev responses. This process of image downsampling is also shown in Fig. 3.4.



(a) $\mathbf{P}$ - Original image (b) $\widetilde{\mathbf{P}}$ - LPF product (c) Decimation Samples (d) $\widetilde{\mathbf{P}}^{\downarrow}$

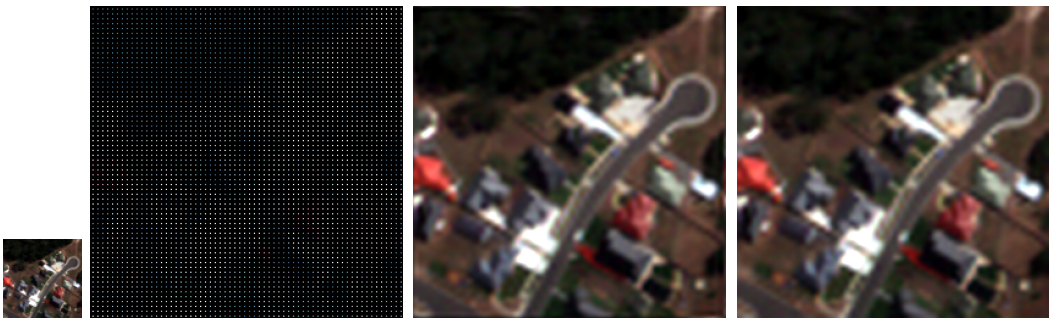
Fig. 3.4. Processing pipeline for the downsampling of a monochromatic image by a factor of 4. In this example, the LPF-effect is given by a Gaussian filter with cutoff frequency $1/4$. Fig. 3.4c shows, in green, the selected pixels for the decimation.

Similarly, the **upsampling** operation can also be split into two parts:

- **Expansion** (or **zero interlacing**): In this step a $\rho - 1$ zeros are interposed in between samples;
- **Interpolation**: The obtained image is smoothed via a LPF or by any other technique that allows to evaluate the samples in the empty spots that were introduced.

A visual representation of the upsampling process is shown in Fig. 3.5.

If the original signal is a discrete sampling of a band-limited continuous function, the **Whittaker–Shannon interpolation formula** allows the perfect recovery of the continuous function through upsampling. However, the use of this formula is limited in practice as the resulting filters are infinite lengths, hence the most common procedure involves convolution products with FIR filters. The most fitting ones for this situation have a symmetrical response and they are categorized as **type I** when they are even length, or **type II** when they are odd length.



(a) $\mathbf{M}$ (b) $\mathbf{M}^{\uparrow}$ - Expansion (c) $\widetilde{\mathbf{M}}^{\uparrow}$ - Lagrange interpolator (d) $\widetilde{\mathbf{M}}^{\uparrow}$ - Spline interpolator

Fig. 3.5. Processing pipeline for the upsampling of a multiband image by a factor of 4. The image is firstly expanded and then either interpolated by a Lagrange polynomial kernel of order 11 (Fig. 3.5c) or a cubic spline (Fig. 3.5d).

Various kernels can be used for the creation of FIR filters, such as the **Lagrange polynomials** [8], which consists in finding the coefficients of a polynomial of a given order that pass through the given coefficients.

Specifically, let $\{r_i\}_{i\in[1,\dots,N_i]}$ be the closest set of known sample points, $\{u_i\}_{i\in[1,\dots,N_i]}$ the associated intensity values, and r be the query point. Each couple of values (r_i, u_i) is also sometimes known as **breakpoint**. The Lagrange polynomial interpolator of order N_i is given by:

$$s(r) = \sum_{i=1}^{N_i} u_i s_i(r)\,, \tag{3.4.1a}$$

$$s_i(r) = \prod_{\substack{k\leq N_i \\ k'\neq i}} \frac{r - r_k}{r_i - r_k}\,, \tag{3.4.1b}$$

which results into a digital filter of length $\rho(N_i - 1) + 1$ if the image is upsampled by a factor of ρ. For example, the most common kernel in image processing is for the order $N_i = 3$, known as **bicubic interpolation**, but higher order interpolators are not uncommon if the computational speed is not limited and the image sizes are large enough not to generate relevant issues at the boundaries [5].

Computationally, it is also common to employ the so-called **piecewise polynomial interpolation**. In the monodimensional case, it consists in estimating the coefficients of an analytical function at each interval between breakpoints. For example, for the **cubic spline** the imposed condition for the interpolators $s_i(r)$ and $s_{i+1}(r)$ around a given breakpoint (r_i, u_i) are that they are cubic polynomial (in Hermite form) and they are continuous up to their second order derivatives. That is, if

$s_i(r_i) = s_{i+1}(r_i)$, $\left.\frac{\partial s_i(t)}{\partial t}\right|_{r=r_i} = \left.\frac{\partial s_{i+1}(t)}{\partial t}\right|_{r=r_i}$ and $\left.\frac{\partial^2 s_i(t)}{\partial t^2}\right|_{r=r_i} = \left.\frac{\partial^2 s_{i+1}(t)}{\partial t^2}\right|_{r=r_i}$ for any intermediate breakpoint; extra conditions have to be imposed for the boundary points [205]. Some alternatives are also available, such as the piecewise cubic Hermite interpolating polynomial (PCHIP), which is very similar to the cubic spline, but it does not impose the continuity on the second order derivative and instead it preserves the shape and monotonicity between consecutive breakpoints.

Irregular grid interpolation

When we need to perform an upsampling with the techniques described in last section, the data points have to be arranged over a regular grid, to allow for the procedure is to be applied in both direction. However, it is sometimes necessary to re-sample an image from data points scattered irregularly over a generic bidimensional grid.

To manage such situations, some different procedures have been developed through the years. I.e., the grid could be represented as a triangulated irregular network (TIN), that is, segmented into triangular meshes, whose vertices correspond to the given data points. Each of those regions can then be resampled with either nearest neighbour (NN) or linear interpolators.

Another possibility is to employ the **radial basis function (RBF)** interpolators [35]. Let $\mathbf{R} \in \mathbb{R}^{2\times N_i}$ be a matrix such that its i-th column $\mathbf{r}_{:i}$ is a 2 element vector denoting the (vertical and horizontal) coordinates of the i-th sample point, and let u_i be it associated intensities. The RBF defines a function $z(\cdot) : \mathbb{R}^+ \to \mathbb{R}$ that whose value depends only on the distance between the input and some fixed point. The associated interpolator function at any given coordinate $\mathbf{r}'$ is given by:

$$s(\mathbf{r}') = \sum_{i=1}^{N_i} w_i z\left(\|\mathbf{r}' - \mathbf{r}_{:i}\|_2\right) , \tag{3.4.2}$$

where the coefficients $\{w_i\}_{i\in[1,\dots,N_i]}$ can be obtained by solving the system of equations:

$$s(\mathbf{r}_{:i'}) = \sum_{i=1}^{N_i} w_i z\left(\|\mathbf{r}_{:i'} - \mathbf{r}_{:i}\|_2\right) , \qquad \forall i' \in [1, \dots, N_i] . \tag{3.4.3}$$

Typical choices for the RBF are, for a given coefficient $\check{\epsilon} > 0$ and a distance represented by the scalar $t \leq 0$:

$$z(t) = \exp(-(\check{\epsilon}t)^2) \qquad \text{Gaussian}\,, \tag{3.4.4a}$$

$$z(t) = \sqrt{1 + (\check{\epsilon}t)^2} \qquad \text{Multiquadric}\,, \tag{3.4.4b}$$

$$z(t) = \frac{1}{1 + (\check{\epsilon}t)^2} \qquad \text{Inverse Quadratic}\,, \tag{3.4.4c}$$

$$z(t) = t^2 \ln(t) \qquad \text{Thin plate spline (TPS)}\,. \tag{3.4.4d}$$

Other approaches for multivariate interpolation non included in this thesis are also available, such as **inverse distance weighting (IDW)**, **Kriging** [244] and **irregular sampling of band-limited images** [101].

3.4.2 Registration

The **registration** phase has the role to fit all sources to the same coordinate system. This issue is mostly relevant for multiplatform, multitemporal or generally inhomogeneous sources (such as for fusing SAR and optical images). Some pre-constructed commercial image bundles usually do not need this step, as in the last level of the processing pipeline, they are also available in **orthorectified** form, which is not only geometrically corrected but with the elevation of the terrain taken into account. Even for such user-ready products, however, some unpredictable effects, such as small variations in the viewpoints and elevation of the satellite, have to be corrected on a case by case basis.

Various techniques have been developed during the years for the joint alignment of images, which is known as **co-registration**. The co-registration typically involves the following steps [95]:

- **Preprocessing**: It is an initial preparation of the images to identify common features; i.e. when there is a LRI/HRI combination, it could be useful to rescale the LRI to match the ground sample distance (GSD) of the HRI;
- **Feature Selection**: Identification of a set of common features between the different sources; in the case of sharpening, the size of the features have to be large enough to be identifiable in the LRI, but small enough to allow to identify their extension on the scene with sub-pixel precision. I.e. the selected features may be corners of detected edges;

- **Feature Correspondence**: Identification of the the coordinates on both images associated with the shared features; this operation is straightforward for point features (e.g. a corner reflector), but can be more involved in other situation, such as for determining a centroid for extended sources, or the boundary points for line features (e.g. a bridge);
- **Determination of a transformation function**: Given the correspondence of the coordinates across the different images, it is necessary to identify a (typically analytic) transformation function to adapt the unregistered image to the geometry of the reference one. The choice of this function is strongly dependent on the characteristics of the image.
- **Resampling**: Given the transformation to the new coordinates, a resampling is often necessary to generate the new samples in the geometry of the reference image. I.e., the translated original points may have been translated to an irregular grid and an interpolation is necessary to obtain the contributions over a regular grid.

Point mapping is a special case of this framework, in which the matching features across the images to register are made of point coordinates. The geometry transformation functions are typically categorized as either **rigid transformations**, which consist of just translations and rotations, and **non-rigid transformations**. The latter category includes **affine transformations**, which preserve lines and parallelism, and **nonlinear transformations**, such as TPS or polynomial transformations of order at least equal to 2 [75].

3.5 Sharpening algorithms

The sharpening problem consists in estimating a synthetic image $\widehat{\mathbf{X}}$ that ideally matches the spatial resolution of a given HRI $\mathbf{P}$ and the spectral resolution of a LRI $\mathbf{M}$. This fusion scheme is necessary, as both technological and physical limitations (e.g., signal to noise ratio of the acquisitions) prevent the acquisition of a single image of both high spatial and spectra resolution. To simplify the exposition, we present in this section a brief introduction to the approaches developed in the literature for the classical problem of PAN/MS fusion, known as **pansharpening**, and $\mathbf{p} \equiv \mathbf{P}$ is a column vector which describes a monochromatic image.

In classical approaches, the generic k-th band of the fused signal $\widehat{\mathbf{x}}_{:k}$ is obtained as:

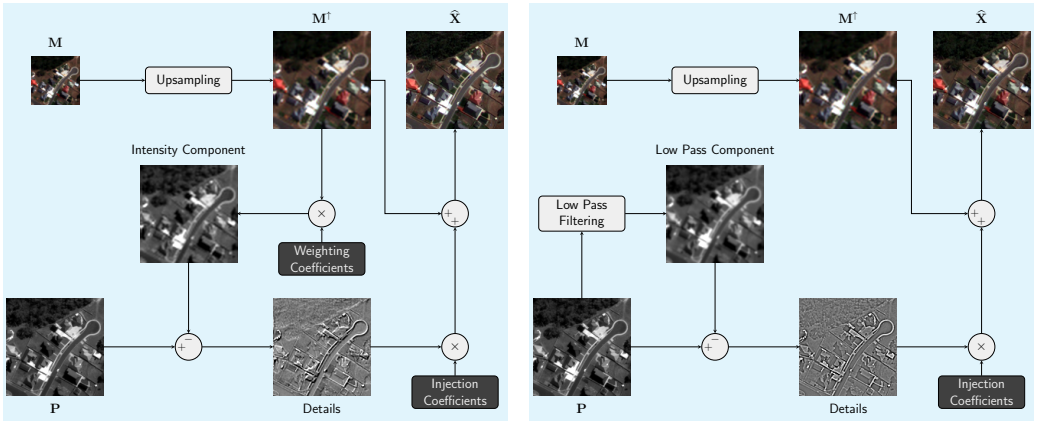
$$\widehat{\mathbf{x}}_{:k} = \widetilde{\mathbf{m}}^{\uparrow}_{:k} + \mathbf{g}_{:k} \odot (\mathbf{p} - \mathbf{j}_{:k}) \,, \qquad \forall k \in [1, \dots, N_b] \,, \tag{3.5.1}$$

where $\mathbf{g}_{:k} \in \mathbb{R}^{N_i}$ are known as **injection coefficients** and the difference $\mathbf{p} - \mathbf{j}_{:k}$ describes the **details** to be injected in the upsampled version $\mathbf{m}^{\uparrow}_{:k}$ of the LRI.

Depending on the method to obtain the term $\mathbf{j}_{:k}$, it is possible to define three different classes of approaches for classical pansharpening:

- **Component substitution (CS)**: if $\mathbf{j}_{:k}$ acts as an **intensity component** of the LRI, that is if it is a linear combination of all channels of the upsampled LRI,
- **Multiresolution analysis (MRA)**: if $\mathbf{j}_{:k}$ is a **low pass filtered version** of the HRI,
- **Hybrid methods**: if $\mathbf{j}_{:k}$ contains a combination of information both from the LRI and the HRI.

An illustrative flowchart for the CS and MRA approaches is shown in Fig. 3.6. A wide variety of approaches are available to define the implementation of the relevant coefficients [224], for which we provide a summary of the relevant literature in the following sections.



(a) CS (Component substitution) **(b)** MRA (Multiresolution analysis)

Fig. 3.6. Different pipeline of operations of classical pansharpening approaches.

A special case of the relation (3.5.1), which is typically used as baseline for the performances, is when the injection coefficients are identically equal to zero and the sharpened image is simply equal to the upsampled version of the LRI; this method is denoted with EXP.

The two main strategies of injection coefficients which will be used in this work are:

$$\mathbf{g}_{:k} = \frac{\widetilde{\mathbf{m}}^{\uparrow}_{:k}}{\mathbf{j}_{:k}} \qquad \text{High pass modulation (HPM)}\,, \tag{3.5.2a}$$

$$\mathbf{g}_{:k} = \frac{\mathrm{cov}(\mathbf{m}_{:k}, \mathbf{j}_{:k})}{\mathrm{std}^2(\mathbf{j}_{:k})} \qquad \text{Context based decision (CBD)}\,. \tag{3.5.2b}$$

The HPM method is also known as multiplicative scheme, as the resulting sharpened image is obtained as:

$$\widehat{\mathbf{x}}_{:k} = \widetilde{\mathbf{m}}^{\uparrow}_{:k} + \frac{\widetilde{\mathbf{m}}^{\uparrow}_{:k}}{\mathbf{j}_{:k}} \odot (\mathbf{p} - \mathbf{j}_{:k}) = \widetilde{\mathbf{m}}^{\uparrow}_{:k} \odot \frac{\mathbf{p}}{\mathbf{j}_{:k}}\,, \tag{3.5.3}$$

while the injection coefficients in CBD can be seen as obtained by a Gram-Schmidt (GSA) orthogonal decomposition of the details, if they are the same for each channel.

3.5.1 Component substitution methods

According to the rationale behind the CS approach, the associated intensity component is given by:

$$\mathbf{j}_{:k} = \sum_{k'=1}^{N_b} w_{kl} \widetilde{\mathbf{m}}^{\uparrow}_{:l}\,, \tag{3.5.4}$$

where $\mathbf{W} = \{w_{kl}\}_{k,l\in[\in,\dots,1]N_b}$ is a square matrix containing the **weighting coefficients**.

The two CS methods that will be employed in this thesis were chosen according to the best performances in terms of quality of the final product [224].

In the GSA-based method, the injection method is given by eq. (3.5.2b). For the **Gram-Schmidt adaptive (GSA)** method [6] in particular, the weighting coefficients are given by solving the following regression problem at reduced resolution:

$$\mathbf{w}_{k:} = \arg\min_{\mathbf{w}'} \left\| \widetilde{\mathbf{p}} - \sum_{l=1}^{N_b} w'_l \widetilde{\mathbf{m}}^{\uparrow}_{:l} \right\|_2\,, \qquad \forall k \in [1, \dots, N_b]\,, \tag{3.5.5}$$

where $\mathbf{w}' = \{w'_l\}_{l\in[1,\dots,N_b]}$.

In the **band-dependent spatial detail (BDSD)** method [86] it is typical to redefine the complementary variables $\mathbf{S} = \left[\mathbf{p}, \widetilde{\mathbf{m}}^{\uparrow}_{:1}, ..., \widetilde{\mathbf{m}}^{\uparrow}_{:N_b}\right]$ and its associated low resolution version $\mathbf{S}^{\downarrow} = [\widetilde{\mathbf{p}}^{\downarrow}, \widetilde{\mathbf{m}}_{:1}, ..., \widetilde{\mathbf{m}}_{:N_b}]$, for which the sharpened channel can be obtained as:

$$\hat{\mathbf{x}}_{:k} = \widetilde{\mathbf{m}}^{\uparrow}_{:k} + \mathbf{S}\mathbf{S}^{\downarrow\dagger}(\mathbf{m}_{:k} - \widetilde{\mathbf{m}}_{:k})\,, \tag{3.5.6}$$

where $\mathbf{S}^{\downarrow\dagger} \in \mathbb{R}^{(N_b+1)\times\frac{N_j}{\rho^2}}$ is the Moore-Penrose pseudo-inverse of $\mathbf{S}^{\downarrow}$.

CS methods tend to prioritize the accuracy of the spatial information compared to the spectral information, so that the product has a good visual appearance. They are also robust to misregistration errors and aliasing [19].

3.5.2 Multiresolution analysis methods

In the multiresolution analysis (MRA) methods, the intensity component $\mathbf{j}_{:k}$ is given as a LPF version of the HRI.

Specifically, for the à trous wavelet transform (ATWT) method [178], this low pass component is obtained as a SWT transformation of the original image (see Section 2.2.2), ignoring all high pass components. The specific implementation is obtained with a convolution product of each column and each row with a 5-tap filter $[1, 4, 6, 4, 1]/16$, repeating the process $\log_2(\rho)$ times.

In the MTF-matched generalized Laplacian pyramid (MTF-GLP) methods [7], the low pass filter is obtained with a convolution product by a Gaussian filter matched to the MTF. Typically, this filter is constructed with the constraint that the gain at the Nyquist cutoff frequency matches the one given by the specifics of the LRI sensor (typically around 0.3 for most commercial sensors). Two different versions of this algorithm are employed in this thesis, denoted with MTF-GLP-HPM and MTF-GLP-CBD, based on the injection scheme (either CBD or HPM, respectively).

MRA methods tend to preserve the accuracy of the spectral information over the spatial details and are robust to temporal misalignments (e.g. changes in the scene) [19].

3.5.3 Bayesian methods

The **Bayesian framework** provides a possible alternative method to model the fusion LRI and HRI as an inverse problem. The model was first proposed by Hardy et al. [116] and successively employed in a variety of fusion approaches [231, 209].

The method is based on describing the available acquisitions as degraded version of an ideal reconstruction $\mathbf{X}$. Mathematically, this is represented by the stochastic process:

$$\mathbf{P} = \mathbb{A}_p(\mathbf{X}) + \mathbf{E}^{[p]} \tag{3.5.7a}$$

$$\mathbf{M} = \mathbb{A}_{m\downarrow}(\mathbf{X}) + \mathbf{E}^{[m]} \tag{3.5.7b}$$

where $\mathbf{P}$ and $\mathbf{M}$ define the HRI and the LRI product, respectively, while $\mathbf{E}^{[p]}$ and $\mathbf{E}^{[m]}$ are additive realizations of a specific noise model relative to their associated acquisition systems.

The direct model operators are defined as follows:

- $\mathbb{A}_p(\cdot)$ is a **spectral degradation** operator. Typically this is obtained as a linear combination of channels, similarly to the description of the intensity component in eq. (3.5.4) for the CS pansharpening methods. Specifically the k-th channel $\mathbf{p}_{:k}^{[sim]}$ of $\mathbf{P}^{[sim]} = \mathbb{A}_p(\mathbf{X})$ is given by:

$$\mathbf{p}_{:k}^{[sim]} = \sum_{l=1}^{N_b} w_{kl}\mathbf{x}_{:l} \tag{3.5.8}$$

  where $\{w_{kl}\}_{\substack{k\in[1,\dots,N_{b_p}]\\ l\in[1,\dots,N_b]}}$ are once again weighting coefficients. The generic coefficients w_{kl} is commonly obtained as the relative overlap between the spectral response of l-th band of the LRI with respect to that of the k-th band of the HRI.

- $\mathbb{A}_{m\downarrow}(\cdot)$ is a **matched downsampling** operator. Similarly to what described in Section 3.4.1, this operation is typically obtained as a cascade of a low pass filter and a decimation by a factor ρ. The low pass filtering option is in charge of the **spatial degradation** and it can be performed with a convolution product with a filter whose expression matches the MTF of the sensor (e.g. a MTF-GLP filter). The decimation is then performed by taking every ρ-th value.

Both $\mathbb{A}_p(\cdot)$ and $\mathbb{A}_{m\downarrow}(\cdot)$ are linear operators, as $\mathbb{A}_p$ is simply a linear combination of given channels and $\mathbb{A}_{m\downarrow}$ is a cascade of a convolution and a decimation, which can both be represented as matrix multiplication, as as shown in Appendix A.1.2 and A.1.4.

If we assume that the noise is distributed as additive white Gaussian noise (AWGN), then, as described in Section 2.1.2, the problem (3.5.7) is equivalent to minimizing an objective function:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}\in\mathbb{E}_x} \frac{1}{2}\left\|\mathbb{A}_p(\mathbf{X}) - \mathbf{P}\right\|_F^2 + \frac{\check{\lambda}_m}{2}\left\|\mathbb{A}_{m\downarrow}(\mathbf{X}) - \mathbf{M}\right\|_F^2 + \check{\lambda}\, g(\mathbb{L}(\mathbf{X}))\,, \qquad (3.5.9)$$

where $\check{\lambda}_m$ and $\check{\lambda}$ are regularization coefficients, used to weight each term of the regularization, while $\|\cdot\|_F$ denotes the Frobenius norm.

Typically, in the context of PAN/HS fusion, the operator $\mathbb{L}(\cdot)$ is chosen to be a subspace transformation that reduces the redundancy of the data of the HS, such as the principal component analysis (PCA) [72].

Various strategies can be employed for the regularization, such as applying the vector total variation (VTV) (Section 2.2.3) in the transformed domain $\mathbb{L}(\mathbf{X})$, which is known as **hyperspectral superresolution (HySure)** [209]. The **Bayesian with naive regularization (BayesNaive)** [232] assumes instead that the regularizer function $g(\cdot)$ is quadratic and has a computationally efficient implementation through the solution of a Sylvester equation [233].

3.5.4 Other sharpening methods

A variety of other techniques and approaches are present in the literature for the sharpening of image, which we consider outside the scope of this thesis. Firstly, it is possible to extend classical methods to MS/HS fusions, i.e. with dedicated band selection approaches from the MS to act like a single band HRI [192].

Other methods include the **coupled nonnegative matrix factorization (CNMF)** [238], which is particularly suited for the sharpening of HS images; both the LRI and HRI are alternately unmixed through nonnegative matrix factorization [10] to estimate the spectral signatures of endmembers from the former and the high-resolution abundance maps from the latter.

More recently, more sophisticated techniques have been proposed based on deep learning techniques to extract features from the HRI and LRI and inject them in the fused image through a convolutional neural network [157, 241].

3.6 Color filter arrays

In this section we present a mathematical direct model that is able to describe the direct acquisition of the CFA (Section 3.6.1), which represent one of the most widespread solutions for multiband acquisition for commercial cameras. We also present the most common designs available commercially and in the literature (Section 3.6.2). Raw acquisitions of CFA do not simultaneously provide the information relative to all channels for every portion of the scene, which demands a reconstruction of the full image with techniques known as demosaic. These techniques are the topic of 3.7.

3.6.1 Direct acquisition model

In the context of digital cameras, a CFAs defines a matrix of spectral filters of photosensors, chosen from a subset of possible spectral responses. In physical terms, let us focus the analysis on a generic i-th detector, which captures the radiant intensity contained within a given solid angle Ω_i.

Let $\mathfrak{I}_\lambda^{[t]}(\Omega_i)$ denote the spectral radiance transmitted to the focal plane, which is a function of the wavelength λ. While more details on the physical significance of this quantity are provided on Section 5.2.3, at this stage it is sufficient to state each photosensor has an assigned filter. The spectral response of this filter is chosen within an assigned set $\{\xi_k(\lambda)\}_{k\in[1,\dots,N_b]}$ of loosely bandpass filters over non-overlapping wavelengths.

The measured radiant intensity y_i at the i-th photosensor is given by:

$$y_i = \int_0^\infty \mathfrak{I}_\lambda^{[t]}(\Omega_i)\xi_{\mathcal{H}(i)}(\lambda)\, d\lambda\,, \tag{3.6.1}$$

where $\mathcal{H}(i) : i \in [1,\dots,N_i] \to [1,\dots,N_b]$ is an assignation function that associates the generic class of spectral responses to the i-th photosensor. I.e., if the i-th pixel is assigned to a green filter, whose spectral response is $\xi_2(\lambda)$ is a passband filtered centered at the green wavelengths, then $\mathcal{H}(i) = 2$.

The vector $\mathbf{y} = \{y_i\}_{i\in[1,\dots,N_i]}$ can be interpreted as a selection of samples from an ideal multiband acquisition $\mathbf{X} \in \mathbb{R}^{N_i \times N_b}$, whose coefficients x_{ik} are given by:

$$x_{ik} = \int_0^\infty \mathfrak{I}_\lambda(\Omega_i)\xi_k(\lambda)\, d\lambda\,, \tag{3.6.2}$$

which consequently allows to interpret $\mathbf{y}$ as obtained from a masking operation such as:

$$\mathbf{y} = \sum_{k=1}^{N_b} \mathbf{x}_{:k} \odot \mathbf{h}_{:k}\,, \tag{3.6.3}$$

where $\mathbf{H} \in \mathbb{R}^{N_i \times N_b}$ is a binary mask, whose coefficients are:

$$h_{ik} = \begin{cases} 1\,, & \text{if } \mathcal{H}(i) = k\,, \\ 0\,, & \text{otherwise}\,. \end{cases} \tag{3.6.4}$$

From another perspective, the position of the 1 in the row vector $\mathbf{h}_{i:}$ defines which channel is selected by the i-th photosensor.

The 3-dimensional array $\boldsymbol{\mathcal{U}}^{[h]} = \mathrm{matr}^{-1}(\mathbf{H})$ is typically represented as a color-coded map: a unique color is assigned to each available class of sensors (or equivalently to a certain channel), and the pixels of that class are colored accordingly. An example of such representation is shown in Fig. 3.8.

In the literature it is also common to distinguish between **CFA**, in the case the spectral response is relative to RGB components, or **MSFA**, in the case they are associated with general spectra. It is possible to extend this framework to the case for which $\mathbf{H}$ is not binary, for which $\mathbf{y}$ can be seen as a linear combination of spectral responses from different channels.

The problem of **demosaicing** can then be seen as finding the most accurate estimation $\widehat{\mathbf{X}}$ of $\mathbf{X}$ given $\mathbf{y}$.

3.6.2 Mask design

Over the course of the years, a great variety of mask designs have arisen either in commercial venues and in scientific endeavors. In this section, we discuss some general design rules, based on periodic and pseudo-random patterns.

General design principles

The design phase of MSFA filters are typically not independent from the envisioned demosaicing algorithms to be implemented for the reconstruction of the full color image.

As the accuracy of the reconstructed product is a compromise between the spectral and spatial resolution, the mask design approaches can prioritize the optimization of either the former or the latter, depending on the final users' demands.

- Spectral resolution optimization: In the first case, the different set of filters are distributed as much as possible uniformly over the whole pixel matrix. More specifically if we define with L_k the minimum distance among elements of the grid belonging to the k-th spectral characteristic, one strategy could be to find an arrangement such that $\sum_{k=1}^{N_b} L_k$ is minimized. This approach was studied in [46] for three bands (3.8k) and can be extended to the pattern in Fig. 3.8j in the case of four bands, where the yellow denotes the additional band (e.g. a near infrared (NIR) channel). These patterns will be denoted as **maximum distance (MAXDIS)** in the following.
- Spatial resolution optimization: In the second scenario, a widespread approach is to define a **dominant band**, which appears the most frequently in the pattern. The main goal of this approach is that the dominant band allows for an easier recovery of the spatial component, which can be used to guide the recovery of the samples from the remaining bands, exploiting their spectral correlation [167].

The **binary tree (BT)** procedure [162] allows to design mask patterns that are versatile for either strategy. The procedure, whose algorithm is described in detail in the Algorithm 1 consists in sequentially splitting the available slots into two subsections, deciding the channel assignation based on a binary tree. The tree can either dictate for each channel to be approximately uniformly distributed, which is the approach of **uniform binary tree (UBT)** in Fig. 3.7a, or to keep a single dominating band to be assigned to the first split section such as in the **dominant binary tree (DBT)** in Fig. 3.7b. Fig. 3.7 offers a visual representation of some examples of BTs, together with their generated pattern.

Periodic masks

Periodic CFA patterns make up for the vast majority of existing digital cameras, with the most common design dating back to a patent from Bryce Bayer of Eastman

Algorithm 1: Binary tree (BT) method for the creation of masks [162].

Result: A square matrix $\mathbf{U}^{[h]}$, whose elements are in $[1, \dots, N_f]$ and define the index of a set of assignation classes $\{\xi_k\}_{k\in[1,\dots,N_f]}$

Inputs:

- A binary tree (e.g. as shown in 3.7) with N_k levels and N_f leaves (in a binary tree, each node has at most 2 children)

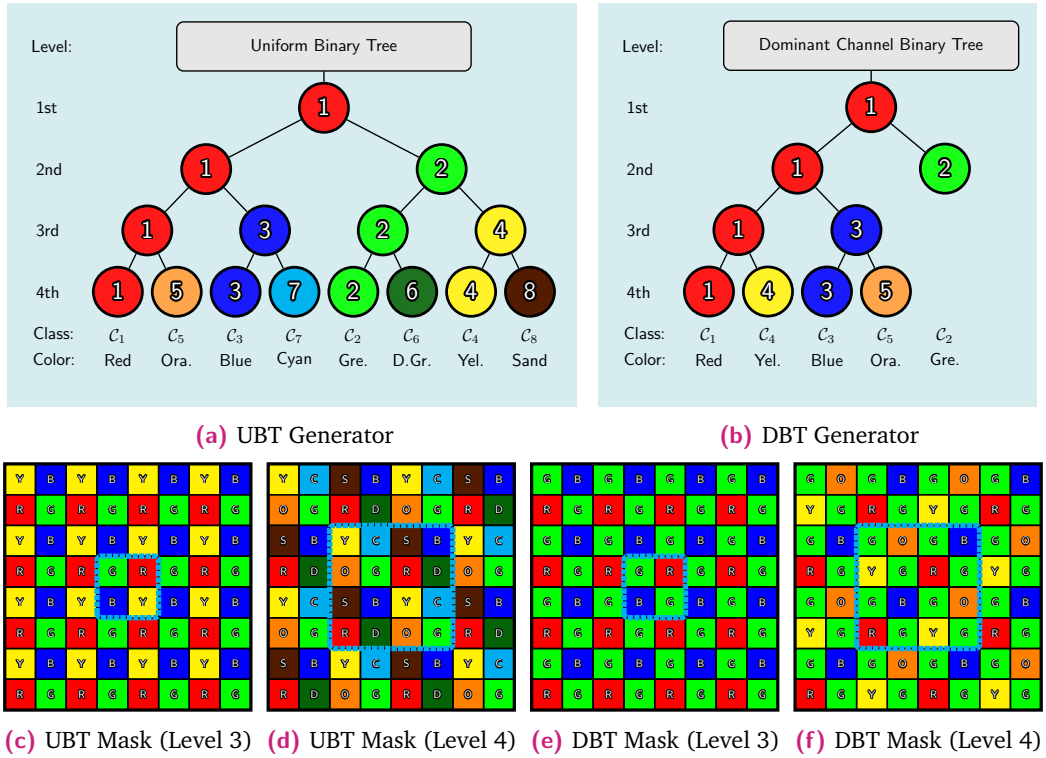
Definitions:

- Q_p: label of the parent node
- Q_{c_1}: the label of the first child node
- Q_{c_2}: the label of the second child node

Initialization:

- Assignation matrix: $\mathbf{F}^{(1)} \leftarrow 1$
- Label of the root node: $Q_p \leftarrow 1$
- Current node: $q \leftarrow 1$
- Current level: $k \leftarrow 1$

while $q \leq N_f$ **do**
  foreach *node at level* k **do**
    Define the current node as the parent node
    $Q_{c_1} \leftarrow Q_p$
    if *the parent node has 2 children* **then**
      $Q_{c_2} \leftarrow q$
      $\mathbf{F}' \leftarrow \mathbf{F}^{(q)}$
      if *at most one element of* $\mathbf{F}^{(q)}$ *is equal to* Q_p **then**
        Substitute the element of $\mathbf{F}'$ equal to Q_p with Q_{c_2}
        $\mathbf{F}^{(q+1)} \leftarrow \begin{pmatrix} \mathbf{F}' & \mathbf{F}^{(q)} \\ \mathbf{F}^{(q)} & \mathbf{F}' \end{pmatrix}$
      else
        Substitute one single element of $\mathbf{F}'$ equal to Q_p with Q_{c_2}
        $\mathbf{F}^{(q+1)} \leftarrow \mathbf{F}'$
      end
      $q \leftarrow q + 1$
    end
  end
  $k \leftarrow k + 1$
end
return $\mathbf{U}^{[h]} \leftarrow \mathbf{F}^{(N_f)}$



(a) UBT Generator **(b)** DBT Generator

(c) UBT Mask (Level 3) **(d)** UBT Mask (Level 4) **(e)** DBT Mask (Level 3) **(f)** DBT Mask (Level 4)

Fig. 3.7. Two examples of a binary tree and the associated masks generated truncating the tree at the third and at the fourth level.

Kodak [21]; the **Bayer filter** is a 2×2 periodic pattern featuring RGB bandpass filters arrangement shown in Fig. 3.8a. The green filters were originally chosen to have twice the representation compared to their red and blue counterparts to ensure that the green component may be reconstructed more accurately, in accordion to the wavelength sensitivity of the human eye. This design is a special case of the result of DBT pattern generation algorithm with three leaves. Since then, many other RGB designs have been proposed, both in patents (such as Yamanaka's [236] in Fig. 3.8b) and in scientific publications (e.g. the proposition from Lukac [152] in Fig. 3.8c).

The **Quad Bayer** pattern shown in Fig. 3.8e is a variant of the Bayer filters, but with 2×2 square patterns being assigned to the same filter. This design was recently manufactured by Sony to be included in the IMX250YMR sensor. While at first blush this design may seem not to obey the principles described in Section 3.6.2, its goal is to provide more flexibility for different acquisition modes. In conditions of low lighting, the photons incident on each 2×2 square may be combined to emulate a classical Bayer filter, and consequently raise the signal to noise ratio (SNR). Additionally, this designs allows to reduce the area on the silicon for each photosite, allowing for higher spatial resolution in normal conditions.

The **uniform mask design** are a common alternative for periodic MSFA mask designs, made up of non-redundant periodic rectangles (with no dominant band), such as in the case of Fig. 3.7c and 3.7d.

Many other patterns were proposed, e.g. by employing dyes that work in the cyan yellow magenta (CYM) color space instead of the classical RGB. Some designs include "white" pixels, characterized by a wide-band filter (or more commonly, no filter, so that the spectral response matches the one of the photosensor) [1]; some examples of this approach include the commercial cameras such as Teledyne Onyx (Fig. 3.8f) and various patents by Kodak (Fig. 3.8g to 3.8i).

Pseudorandom masks

While deterministic masks are the standard in commercial cameras (with the Bayer's mask being the most widespread), recent studies have shown potential in employing random patterns. In [12, 13], the authors investigate the effectiveness of random binary masks, by proposing an MSFA with a completely randomized mosaic and a customized demosaicing algorithm, which requires a training phase.

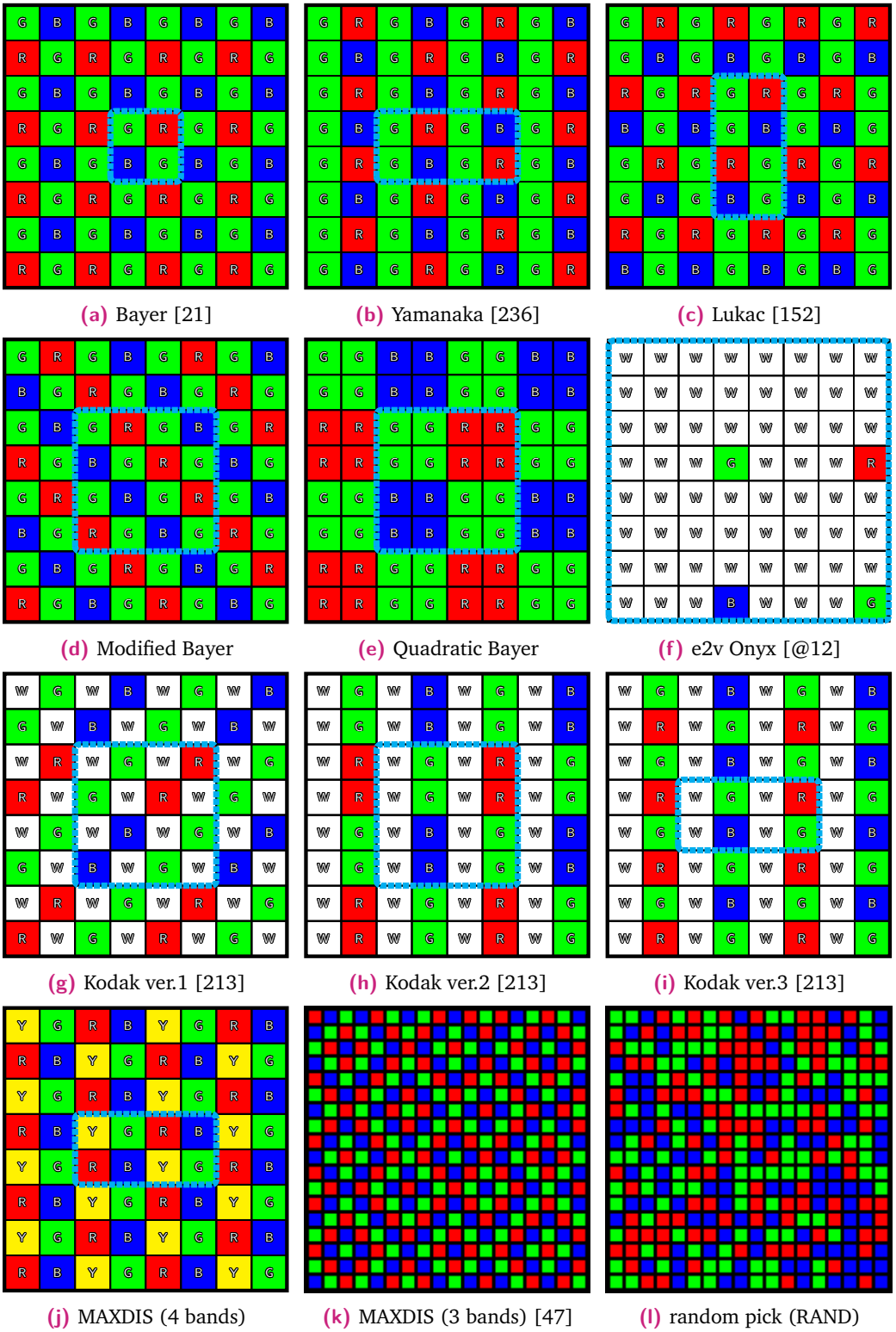
In the **compressive coded aperture spectral imaging (CASSI)** approach, specifically in its single dispersion version (SD-CASSI) shown in Fig. 3.9, each band is sequentially shifted by one pixel in the horizontal direction through a dispersive element (e.g. a prism) and then filtered through a coded aperture, typically realized with digital micromirror device (DMD), which emulates the behaviour of a binary mask. Each filtered channel is finally combined over a shared focal plane array (FPA).

In mathematical terms, this acquisition can be expressed as a modified version of eq. (3.6.3), as shown below:

$$\mathbf{U}^{[y]} = \sum_{k=1}^{N_b} \left(\left[\mathbf{U}_{::k}^{[x]} \odot \mathbf{U}_{::k}^{[h]}, \, \mathbf{0}_{[N_{i_1} \times (N_b - 1)]} \right] \right)_{\rightarrow (k-1)} , \tag{3.6.5a}$$

$$\mathbf{y} = \sum_{k=1}^{N_b} \left(\left[\mathbf{x}_{:k} \odot \mathbf{h}_{:k}; \, \mathbf{0}_{[N_{i_1}(N_b - 1) \times 1]} \right] \right)_{\downarrow ((k-1) N_{i_1})} , \tag{3.6.5b}$$

[1] Wideband pixels are commonly known as panchromatic in the CFA literature; nonetheless, we denote them as wideband or "white" pixels not to generate confusion with sensors characterized by higher spatial resolution.



(a) Bayer [21] (b) Yamanaka [236] (c) Lukac [152]

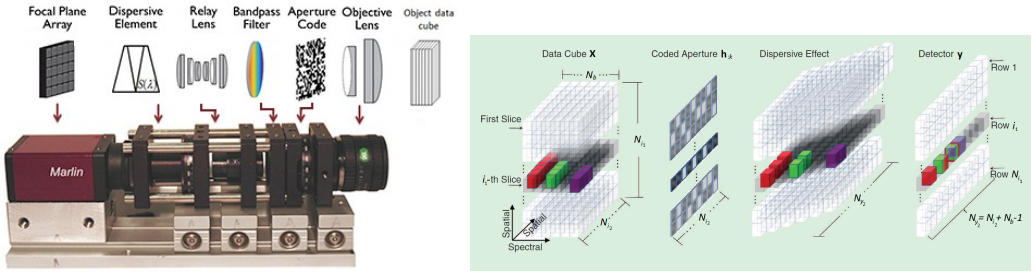
(d) Modified Bayer (e) Quadratic Bayer (f) e2v Onyx [@12]

(g) Kodak ver.1 [213] (h) Kodak ver.2 [213] (i) Kodak ver.3 [213]

(j) MAXDIS (4 bands) (k) MAXDIS (3 bands) [47] (l) random pick (RAND)

Fig. 3.8. Various choices for RGB CFA patterns. The pixels are labelled with the initial of their associated color filter. The masks in Fig. 3.8f to 3.8i include a wide-band sensor, labelled with "W". The cyan dotted outline delimits the periodicity.

in its natural and lexicographic representation, respectively. The final acquisition $\mathbf{U}^{[y]} \in \mathbb{R}^{N_{i_1} N_{y_2} N_b}$ features more rows than the product to reconstruct, due to the shifting effect introduced by the dispersive element, represented by the operators $(\cdot)_{\to k}$ and $(\cdot)_{\downarrow i}$, which denote a circular shift by k columns on the right and by i rows below, respectively. Typically, the masks $\mathbf{h}_{:k}$ are chosen to be random binary masks, and equal for each channel, as it allows for easy practical implementation, as the DMD does not have to switch their coded aperture in a single acquisition. The original paper proposes to reconstruct the full spectral image with a Bayesian framework, by employing a sparsity inducing regularizers, with a wavelet transformation on the spatial dimensions and a discrete cosine transform (DCT) transformation on the spectral one [14].



Source: [14] Source: Adapted from [14]

(a) Physical realization **(b)** Acquisition model

Fig. 3.9. Single dispersion CASSI prototype and its acquisition model.

3.7 Demosaicing algorithms

The goal of a demosaicing method is to recover a full band image, given an acquisition by an optical system involving either a CFA or a MSFA.

Mathematically, the instrument's readout $\mathbf{y} \in \mathbb{R}^{N_i}$ is taken over a FPA with N_b different kinds of filters, and the reconstructed product $\widehat{\mathbf{x}} \in \mathbb{R}^{N_i \times N_b}$ has to contain the full spectrum information at each pixel.

In this section we present a brief discussion of the demosaicing techniques proposed in the literature, without any claim of exhaustivity, as we give priority here on summarizing the general ideas behind various classes of approaches instead of detailing each procedure. This section also contains some preliminary considerations to establish a mathematical link between demosaicing methods and classical sharpening methodologies.

3.7.1 Basic operations

This section's goal is to describe some basic definitions and operations that are common to a variety of demosaic approaches.

The discussion is limited here to acquisitions taken with MSFAs patterns described by binary masks with a sum-to-1 condition in the spatial domain ($\sum_{k=1}^{N_b} \mathbf{h}_{ik} = 1$, for all $i \in [1, \dots, N_i]$), so that each pixel is uniquely assigned to a single filter.

We define **sparse channel** $\mathbf{Y}^{\boxdot} = \mathbf{y} \odot \mathbf{H}$ the Hadamard product between the acquisition and the mask. The operation is mathematically consistent as long as we broadcast the acquisition in the spectral dimension. In other words, the k-th channel $\mathbf{y}^{\boxdot}_{:k}$ of $\mathbf{Y}^{\boxdot}$ is rigorously defined as:

$$\mathbf{y}^{\boxdot}_{:k} = \mathbf{y} \odot \mathbf{h}_{:k} \,. \tag{3.7.1}$$

The result $\mathbf{y}^{\boxdot}_{:k}$ is sparse in the sense that it is zero everywhere, except for the pixels assigned to the k-th filter, where it is equal to the original acquisition.

In mathematical terms, as proven in the Appendix A.1.3, $\mathbf{Y}^{\boxdot}$ is equivalent to the adjoint of the masking operation (3.6.3), applied to $\mathbf{y}$.

If the direct acquisition model is fully deterministic, the sparse channel $\mathbf{y}_{:k}$ can be interpreted as a subsampled version of the ideal channel $\mathbf{x}_{:k}$ that we aim to reconstruct. By just exploiting the spatial correlation of each channel, it is hence possible to obtain a naive demoisaiced product through the interpolation of each of the sparse channels.

The **weighted bilinear (WB)** [30] is one of such methods, which implements a linear interpolator. Let us define a uniform mask be defined as the concatenation of non-redundant periodic rectangles of size $N_{b_1} \times N_{b_2}$ such that $N_b = N_{b_1} N_{b_2}$. In this case let:

$$\mathbf{c}_{(N_b)} = \frac{1}{2N_b - 1}[1; 2; \dots; N_b - 1; N_b; N_b - 1; \dots; 2; 1] \tag{3.7.2}$$

be a column array that implements a linear kernel. The bilinear interpolation filter is then given by the Kronecker product:

$$\mathbf{U}^{[b]}_{::k} = \mathbf{c}_{(N_{b_1})} \otimes \mathbf{c}^{\mathsf{T}}_{(N_{b_2})} \,, \qquad \forall k \in [1, \dots, N_b] \,. \tag{3.7.3}$$

The WB estimation $\widetilde{\mathbf{y}}^{\boxdot}_{:k}$ (denoted with a tilde because it is equivalent to a low pass filtering) is then given by the circular convolution product:

$$\widetilde{\mathbf{y}}^{\boxdot}_{:k} = \mathbf{y}^{\boxdot}_{:k} \ast \mathbf{b}_{:k} \,. \tag{3.7.4}$$

where $*$ is the operator defined in eq. (3.3.1) and $\mathbf{b}_{:k} = \mathrm{matr}\left(\mathbf{U}^{[b]}_{::k}\right)$.

In a more general case, the interpolation filter $\mathbf{b}_{:k}$ can be different for each channel and even not resulting from a Kronecker product; in the most widespread example, applicable to acquisitions taken by a standard Bayer filter camera, each missing sample is obtained as a weighted sum of the closest available neighbours, which yields the following filters [109]:

$$\mathbf{U}^{[b]}_{::1} = \frac{1}{4}\begin{bmatrix}1 & 2 & 1\\ 2 & 4 & 2\\ 1 & 2 & 1\end{bmatrix}, \quad \mathbf{U}^{[b]}_{::2} = \frac{1}{2}\begin{bmatrix}0 & 1 & 0\\ 1 & 2 & 1\\ 0 & 1 & 0\end{bmatrix}, \quad \mathbf{U}^{[b]}_{::3} = \frac{1}{4}\begin{bmatrix}1 & 2 & 1\\ 2 & 4 & 2\\ 1 & 2 & 1\end{bmatrix}, \tag{3.7.5}$$

relative to the red, green, and blue channel, respectively.

Another technique for band indipendent interpolation is the **binary tree-based edge-Sensing (BTES)** [161], which performs a progressive estimation of missing samples for masks generated with binary trees [162], although the interpolation weights are locally different as they follow a custom edge sensing approach.

Finally, if the available samples are unevenly distributed, such as in the case of the pseudo-random masks of Section 3.6.2, $\widetilde{\mathbf{y}}^{\boxdot}_{:k}$ can be obtained with any of the irregular grid interpolation methods shown in Section 3.4.1, such as the TPS RBF.

3.7.2 Spectral difference methods

The **spectral difference (SD)** techniques are a subset of demosaicing methods based on the injection of complementary spectral information from all interpolated sparse channels.

Let $\mathbf{d}^{[l]}_{:k}$ describe the difference between the k-th sparse channel and the spectral component of the l-th channel of the pixels assigned to the k-th channel:

$$\mathbf{d}^{[l]}_{:k} = \mathbf{y}^{\boxdot}_{:k} - \widetilde{\mathbf{y}}^{\boxdot}_{:l} \odot \mathbf{h}_{:k}\,, \tag{3.7.6}$$

and let $\widetilde{\mathbf{d}}^{[l]}_{:k} = \mathbf{d}^{[l]}_{:k} * \mathbf{b}_{:k}$ be its interpolated value. The SD estimation is then given by:

$$\widehat{\mathbf{x}}_{:k} = \sum_{l=1}^{N_b}\left(\mathbf{y}^{\boxdot}_{:l} + \widetilde{\mathbf{d}}^{[l]}_{:k} \odot \mathbf{h}_{:l}\right). \tag{3.7.7}$$

which can also be rewritten as:

$$\widehat{\mathbf{x}}_{:k} = \sum_{l=1}^{N_b} \mathbf{y}_{:l}^{\boxdot} + \left(\left(\mathbf{y}_{:k}^{\boxdot} - \widetilde{\mathbf{y}}_{:l}^{\boxdot} \odot \mathbf{h}_{:k} \right) \ * \ \mathbf{b}_{:k} \right) \odot \mathbf{h}_{:l} \tag{3.7.8a}$$

$$= \widetilde{\mathbf{y}}_{:k}^{\boxdot} + \left(\mathbf{y} - \sum_{l=1}^{N_b} \left(\left(\widetilde{\mathbf{y}}_{:l}^{\boxdot} \odot \mathbf{h}_{:k} \right) \ * \ \mathbf{b}_{:k} \right) \odot \mathbf{h}_{:l} \right) \tag{3.7.8b}$$

as a result of eq. (3.7.1) and the relationship $\mathbf{y} = \sum_{l=1}^{N_b} \mathbf{y}_{:l}^{\boxdot} \odot \mathbf{h}_{:l}$, due to the structure of the masks we are considering. The expression (3.7.8) showcases how the SD estimation can be seen as an injection of the details $(\mathbf{y} - \mathbf{j}_{:k})$ in the interpolated sparse channel $\widetilde{\mathbf{y}}_{:k}^{\boxdot}$; the term $\mathbf{j}_{:k} = \sum_{l=1}^{N_b} \left(\left(\widetilde{\mathbf{y}}_{:l}^{\boxdot} \odot \mathbf{h}_{:k} \right) \ * \ \mathbf{b}_{:k} \right) \odot \mathbf{h}_{:l}$ is a linear combination of the spectral information from the other channels, similarly to what the CS class performs in the context of fusion methods (Section 3.5.1).

An extension of this method, known as **iterative spectral difference (ItSD)** [165], consists in iterating this procedure by taking into account that bands with closer central wavelenghts are more likely to show higher correlation than the rest.

3.7.3 Residual interpolation methods

The residual interpolation (RI) methods are a particular class of demosaicing algorithms for the reconstruction of acquisitions with a **dominant band**, denoted with the index k'. I.e., for the case of a Bayer filter, for which the green filter has double the occurrence compared to the other two colors, we have $k' = 2$.

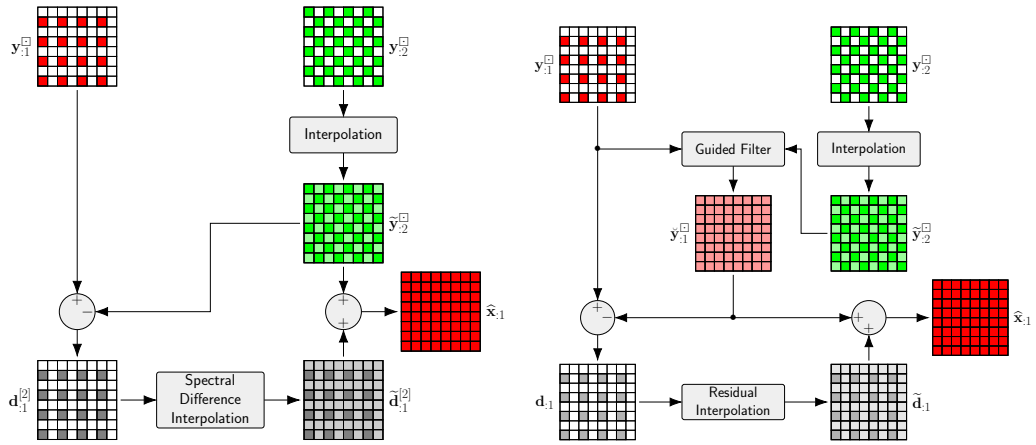
The RI method is very similar to the case of a SD demosaicing procedure, injecting in the naive spatial interpolated estimation, the complementary spectral information just from the dominant band, or in other words:

$$\widehat{\mathbf{x}}_{:k} = \left(\mathbf{y}_{:k'}^{\boxdot} + \widetilde{\mathbf{d}}_{:k}^{[k']} \odot \mathbf{h}_{:k'} \right) . \tag{3.7.9}$$

This procedure is shown in Fig. 3.10a.

The sparse channel difference $\mathbf{d}_{:k}^{[k']}$, which for SD was defined in eq. (3.7.6), is substituted in RI methods with the **residual**:

$$\mathbf{d}_{:k} = \mathbf{y}_{:k}^{\boxdot} - \breve{\mathbf{y}}_{:k} \, , \tag{3.7.10}$$



(a) Spectral difference **(b)** Residual interpolation

Fig. 3.10. Different procedures for the demosaic of a CFA camera acquisition with with a dominant band. The figure shows the demosaicing of the red channel with green used as guide.

which is the difference between the sparse channel and a **guided interpolation** $\breve{\mathbf{y}}_{:k} = g_{GF}(\mathbf{y}^{\boxdot}_{:k}, \widetilde{\mathbf{y}}_{:k'})$, obtained by setting up a **guided filter** [118]. The flowchart of this method is shown in Fig. 3.10b.

The specific definition of the guided filter has stemmed different variations on this algorithm in the literature, such as the minimized-Laplacian residual interpolation (MLRI) [133], where the guiding procedure is performed by minimizing the energy of the Laplacian of the image and the adaptative residual interpolation (ARI) [166], which evolves on the previous concept by introducing an iterative procedure. If $\mathbf{x}^{(0)}_{:k}$ is the product of the MLRI, the estimation $\mathbf{x}^{(q)}_{:k}$ is given by applying the same algorithm, except for the expression of the guided filter $g_{GF}(\mathbf{y}^{\boxdot}_{:k}, \widehat{\mathbf{x}}^{(q-1)}_{:k})$, with the estimation at the previous iteration $\widehat{\mathbf{x}}^{(q-1)}_{:k}$ as guidance.

The RI procedure was also proposed for the reconstruction of NIR channels through the removal of a subset of the filters from a standard Bayer pattern [215].

3.7.4 Intensity difference methods

The **intensity difference (ID)** methods are based on constructing a pseudo-panchromatic, containing the common details available in all sparse channels, to inject in the naive interpolation.

The pseudo-panchromatic is typically defined as a convolution product in the form:

$$\mathbf{p} = \mathbf{y} \ * \ \mathbf{b}' \tag{3.7.11}$$

where $\mathbf{b}'$ is an averaging filter that provides a smooth image from the mosaiced acquisition. For the sake of exposition, we illustrate here a simple procedure to construct the filter $\mathbf{b}'$ which is applicable to masks made up of periodic non-redundant rectangle of size $N_{b_1} \times N_{b_2}$, but more sophisticated approaches are available [164].

Let $\mathbf{c}'_{(N_b)}$ be a column array acting as averaging kernel

$$\mathbf{c}'_{(N_b)} = \begin{cases} \frac{1}{N_b}\mathbf{1}_{[N_b \times 1]} & \text{if } N_b \text{ is odd} \\ \frac{1}{N_b+1}[1; 2 \cdot \mathbf{1}_{[(N_b-1)\times 1]}; 1] & \text{if } N_b \text{ is even} \end{cases} \tag{3.7.12}$$

and the associated averaging filter $\mathbf{U}^{[b']} = \mathrm{matr}^{-1}(\mathbf{b}')$:

$$\mathbf{U}^{[b']} = \mathbf{c}'_{(N_{b_1})} \otimes \mathbf{c}'^{\mathsf{T}}_{(N_{b_2})} . \tag{3.7.13}$$

The resulting ID estimation is then given by:

$$\widehat{\mathbf{x}}_{:k} = \mathbf{p} + \left(\mathbf{y}^{\boxdot}_{:k} - (\mathbf{p} \odot \mathbf{h}_{:k})\right) \ * \ \mathbf{b}_{:k} \tag{3.7.14a}$$

$$= \widetilde{\mathbf{y}}^{\boxdot}_{:k} + (\mathbf{p} - (\mathbf{p} \odot \mathbf{h}_{:k}) \ * \ \mathbf{b}_{:k}) \tag{3.7.14b}$$

which can be interpreted as adding to the pseudo-panchromatic the sparse difference $\mathbf{y}^{\boxdot}_{:k} - (\mathbf{p} \odot \mathbf{h}_{:k})$ interpolated to estimate the missing samples. Alternatively, this can be interpreted as injecting into $\widetilde{\mathbf{y}}^{\boxdot}_{:k}$ the details $\mathbf{p} - (\mathbf{p} \odot \mathbf{h}_{:k}) \ * \ \mathbf{b}_{:k}$, which are obtained by subtracting from $\mathbf{p}$ its low pass component $(\mathbf{p} \odot \mathbf{h}_{:k}) \ * \ \mathbf{b}_{:k}$, similarly to what is performed in the MRA fusion methods (Section 3.5.2).

This iterative intensity difference (ItID) [163] is an extended version of the previous procedure, where the pseudo-panchromatic is re-evaluated at each successive iteration as a weighted average of the current full channel estimation.

3.7.5 Other methods

Other than the discussed methods, the vast literature on demosaicing includes various other alternatives. The **multiscale gradients (MSG)** [126] provides a quick noniterative procedure applicable to acquisitions which employ a Bayer mask, which

is based on the calculation of image gradients at different scales. In the **discrete wavelet transform (DWT)** method [228], each interpolated sparse channel is decomposed with a wavelet transform and its high-filter component is substituted with that of the dominant channel, if available, or generically the sharpest one. In another derived approach [164], the high filter component is instead injected by a pseudo-panchromatic image, such as the one obtained in eq. (3.7.11). Some authors have also proposed to solve this problem by setting up a Bayesian inference framework; in these approaches the reconstructed image is obtained by solving the problem:

$$\widehat{\mathbf{X}} = \arg\min_{\mathbf{x}} \frac{1}{2} \|\mathbb{A}(\mathbf{X}) - \mathbf{y}\|_2^2 + \check{\lambda}\, g(\mathbb{L}(\mathbf{X})) \tag{3.7.15}$$

where $\mathbb{A}(\cdot)$ is the direct model that describe the mosaicing process of eq.(3.6.3), $g(\cdot)$ is a generic regularization function, $\mathbb{L}(\cdot)$ is a linear operator and $\check{\lambda}$ is the regularization parameter. I.e., such approach was employed for the inversion of compressed acquisitions acquired by the CASSI, for which eq. (3.7.15) is set as a least absolute shrinkage and selection operator (LASSO) framework, with the sparsity inducing operator $\mathbb{L}(\cdot)$ being a wavelet transformation in the spatial domain and a DCT transformation in the spectral domain [14]. More recently, patch-based and convolutional neural network (CNN)-based approaches [34, 214, 67] showed even more impressive performances.

3.8 Validation

To evaluate the quality of a reconstruction technique, it is often necessary to compare the obtained product with a reference.

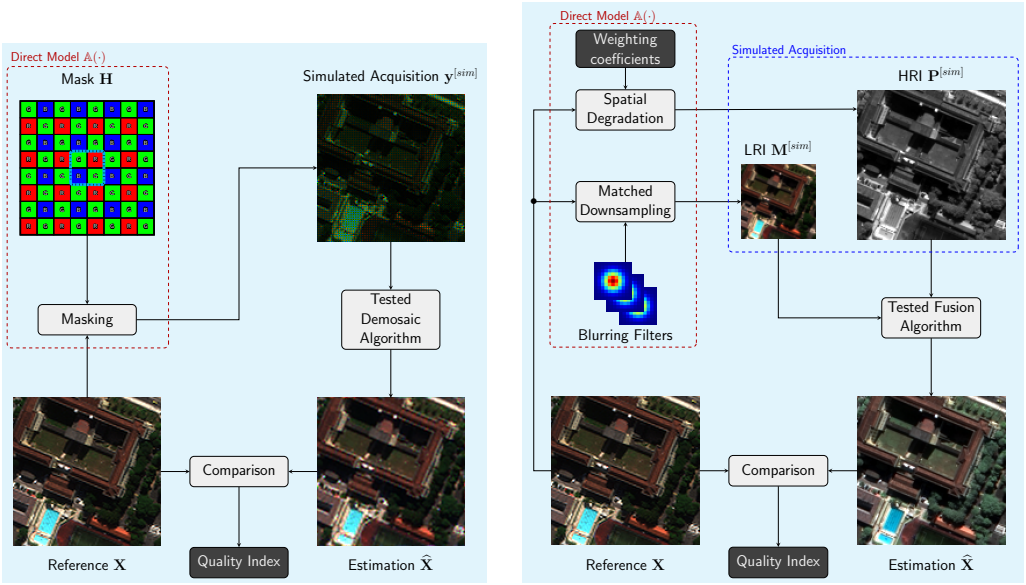
I.e. for a classic pansharpening problem such as the one described in Section 3.5, when the PAN and MS to fuse are given with a certain GSD L_g and ρL_g, respectively, the reference should be given by an MS whose GSD is equal to L_g. I.e. for cameras mounted on unmanned aerial vehicles (UAVs), those references could be obtained by taking a different acquisition at a lower altitude, assuming the scene has not changed.

However, it is a common occurrence that this reference is not available, so that the validation has to be performed over a simulated acquisition. Let $\mathbf{X} \in \mathbb{R}^{N_i \times N_b}$ be a multiband image that is representative of the products we would want to achieve, and $\mathbb{A}(\cdot)$ a given model of the acquisition system. In this validation framework,

the simulated acquisition is given by $\mathbf{y}^{[sim]} = \mathbb{A}(\mathbf{X})$, whose format depends on the specific acquisition system.

The reconstruction algorithm, that is the object of our test, is then performed as usual over the simulated acquisition $\mathbf{y}^{[sim]}$, ignoring the knowledge of $\mathbf{X}$, to obtain an estimation $\widehat{\mathbf{X}}$ of the desired product.

Finally, the performances of the algorithm are evaluated by comparing $\mathbf{X}$ and $\widehat{\mathbf{X}}$ with a set of quality indices that measure the similarities among the two, according to a given criterion, which are described in Section 3.8.1.



(a) Demosaicing **(b)** Fusion

Fig. 3.11. Block scheme pipeline of the operations for the validation of the reconstruction algorithms for the case of demosaicing and data fusion. The monochromatic acquisition $\mathbf{y}^{[sim]}$ is shown as a coloured mosaic for the sake of exposition.

This procedure can be adapted to any optical system with a known direct model, such as in the case of the assessment of demosaicing or fusion algorithms, which are shown in Fig. 3.11. In the case of fusion, the simulated acquisition consists in both a LRI $\mathbf{M}^{[sim]}$ and a HRI $\mathbf{P}^{[sim]}$, which are obtained as spatial and spectral degradations of $\mathbf{X}$, respectively.

However, in the context of fusion algorithms, an alternative procedure is also common, known as **Wald's protocol** or **reduced resolution quality assessment** [227]; in this case, the $\mathbf{P}^{[sim]}$ is not obtained from $\mathbf{X}$, but instead as downsampling from a HRI at a higher resolution, similarly to how $\mathbf{M}^{[sim]}$ is obtained from $\mathbf{X}$, but whose spatial degradation is obtained with filters match the point spread function (PSF) of the HRI instead of the LRI.

3.8.1 Quality indices

Let $\mathbf{X} = \{x_{ik}\}_{\substack{i\in[1,\dots,N_i]\\ k\in[1,\dots,N_b]}}$ and $\widehat{\mathbf{X}} = \{\widehat{x}_{ik}\}_{\substack{i\in[1,\dots,N_i]\\ k\in[1,\dots,N_b]}}$ be the reference and the reconstructed image, respectively. We present here a list of the quality indices employed in this thesis to compare their characteristics.

- **Mean absolute error (MAE)**: It is defined as the mean absolute error between $\widehat{\mathbf{X}}$ and $\mathbf{X}$:

$$\text{MAE} := \frac{1}{N_i N_b} \sum_{i=1}^{N_i} \sum_{k=1}^{N_b} |\widehat{x}_{ik} - x_{ik}| \tag{3.8.1}$$

- **Root mean square error (RMSE)**: It is defined as the STD of the error $\widehat{\mathbf{X}}$ and $\mathbf{X}$:

$$\text{RMSE} := \text{RMSE}\left(\mathbf{X}, \widehat{\mathbf{X}}\right) = \sqrt{\frac{1}{N_i N_b} \sum_{i=1}^{N_i} \sum_{k=1}^{N_b} (\widehat{x}_{ik} - x_{ik})^2} \tag{3.8.2}$$

- **Peak signal to noise ratio (PSNR)**: The PSNR is a derived indices from the RMSE, which evaluates the logarithmic difference between the maximum available power of a signal and the distortion (noise) power. It is defined as:

$$\text{PSNR} := 20 \log_{10} \left(\frac{\max(\mathbf{X})}{\text{RMSE}} \right) , \tag{3.8.3}$$

  where $\max(\mathbf{X})$ is the maximum possible intensity value of $\mathbf{X}$.

- **Relative dimensionless global error in synthesis (ERGAS)** [226]: It is a specific index for the quality assessment of multimodal fused images with a scale ratio ρ between the HRI and the LRI:

$$\text{ERGAS} := \frac{100}{\rho} \sqrt{\frac{1}{N_b} \sum_{k=1}^{N_b} \frac{\text{RMSE}\left(\mathbf{x}_{:k}, \widehat{\mathbf{x}}_{:k}\right)}{\overline{\widehat{\mathbf{x}}}_{:k}}} \tag{3.8.4}$$

- **Spectral angle mapper (SAM)** [242]: Let each pixel's spectral component be described by an N_b-dimension vector. The SAM is defined as the average angle between the vector associated with the reference and the estimated image:

$$\text{SAM} := \frac{1}{N_i}\sum_{i=1}^{N_i}\frac{\langle \mathbf{x}_{i:}, \widehat{\mathbf{x}}_{i:}\rangle}{\|\mathbf{x}_{i:}\|_2\,\|\widehat{\mathbf{x}}_{i:}\|_2} = \frac{1}{N_i}\sum_{i=1}^{N_i}\frac{\sum_{k=1}^{N_b} x_{ik}\widehat{x}_{ik}}{\sqrt{\left(\sum_{k=1}^{N_b} x_{ik}^2\right)\left(\sum_{k=1}^{N_b}\widehat{x}_{ik}^2\right)}}\,. \tag{3.8.5}$$

  Small values of the SAM correspond to closer matches between the spectral components.

- **Structural similarity (SSIM)** [229]: The SSIM is a comparison index between the humanly perceived differences in the image, and it is defined as:

$$\text{SSIM} := \frac{1}{N_b}\sum_{k=1}^{N_b}\frac{\left(2\overline{\mathbf{x}}_{:k}\overline{\widehat{\mathbf{x}}}_{:k} + c_1\right)\left(2\,\text{cov}\left(\mathbf{x}_{:k}, \widehat{\mathbf{x}}_{:k}\right) + c_2\right)}{\left(\overline{\mathbf{x}}_{:k}^2 + \overline{\widehat{\mathbf{x}}}_{:k}^2 + c_1\right)\left(\text{std}^2(\mathbf{x}_{:k}) + \text{std}^2(\widehat{\mathbf{x}}_{:k}) + c_2\right)}\,, \tag{3.8.6}$$

  with the default choices for the scalars $c_1 = (0.01\max(\mathbf{X}))^2$ and $c_2 = (0.03\max(\mathbf{X}))^2$. The SSIM can be interpreted as the product:

$$\begin{aligned}\text{SSIM} = \frac{1}{N_b}\sum_{k=1}^{N_b}&\frac{2\overline{\mathbf{x}}_{:k}\overline{\widehat{\mathbf{x}}}_{:k} + c_1}{\overline{\mathbf{x}}_{:k}^2 + \overline{\widehat{\mathbf{x}}}_{:k}^2 + c_1}\cdot\frac{2\,\text{std}\left(\mathbf{x}_{:k}\right)\text{std}\left(\widehat{\mathbf{x}}_{:k}\right) + c_2}{\text{std}^2(\mathbf{x}_{:k}) + \text{std}^2(\widehat{\mathbf{x}}_{:k}) + c_2}\\ &\cdot\frac{\text{cov}\left(\mathbf{x}_{:k}, \widehat{\mathbf{x}}_{:k}\right) + \frac{c_2}{2}}{\text{std}\left(\mathbf{x}_{:k}\right)\text{std}\left(\widehat{\mathbf{x}}_{:k}\right) + \frac{c_2}{2}}\,,\end{aligned} \tag{3.8.7}$$

  where each factor performs, in order, a comparison between the luminance, contrast, and structure of the two. The SSIM is an extended case of the **universal image quality index (UIQI)** [230], an index proposed by the same authors, which coincides with the SSIM in the case $c_1 = c_2 = 0$:

$$\text{UIQI} := \frac{1}{N_b}\sum_{k=1}^{N_b}\frac{4\overline{\mathbf{x}}_{:k}\overline{\widehat{\mathbf{x}}}_{:k}\,\text{cov}\left(\widehat{\mathbf{x}}_{:k}, \mathbf{x}_{:k}\right)}{\left(\overline{\widehat{\mathbf{x}}}_{:k}^2 + \overline{\mathbf{x}}_{:k}^2\right)\left(\text{std}^2(\widehat{\mathbf{x}}_{:k}) + \text{std}^2(\mathbf{x}_{:k})\right)}\,. \tag{3.8.8}$$

  Those index are also sometimes calculated locally, over overlapped square portions of the image of a given size, and a unique index is provided as the average over all considered partitions. The original definition [229] was given for monochromatic images, which is here extended here for multiband images by taking the average across all bands.

- $Q2^n$ **index (Q^2n)** [85]: The Q^2n is an extension of the UIQI, which was originally proposed for 4 bands [11]. In the proposed index, each pixel spectral component is represented by a hypercomplex number (specifically as N_b-ions, or quaternions if $N_b = 4$) and eq. (3.8.8) can be rewritten in terms of the definition of variances and covariances specific to hypercomplex numbers [11].

- **Spatial cross-covariance coefficient (sCC)**: The sCC is a measurement of the cross-covariance between images, after processing both of them with an edge-detection filter. I.e., let $\mathbf{X}^{[e]}$ and $\widehat{\mathbf{X}}^{[e]}$ be the reference and the estimated image, with each frontal slice processed with a Laplacian filter, then the sCC is defined as:

$$\text{sCC} := \frac{1}{N_b} \sum_{k=1}^{N_b} \frac{\operatorname{cov}\left(\widehat{\mathbf{x}}_{:k}^{[e]}, \mathbf{x}_{:k}^{[e]}\right)}{\operatorname{std}^2\left(\widehat{\mathbf{x}}_{:k}^{[e]}\right) \operatorname{std}^2\left(\mathbf{x}_{:k}^{[e]}\right)} \,. \tag{3.8.9}$$

4 Joint fusion and demosaicing of compressed multiresolution acquisitions

This chapter presents a novel design for a custom optical imaging device inspired by the color filter array (CFA) technologies, whose compressed acquisitions embed information from sources with different characteristics (e.g. with different spatial and spectral resolution) disposed over a common focal plane. The main application of this device is aimed at a constellation of low-cost satellites, whose optical on-board system is able to directly generate a compressed final acquisition in order to reduce the memory footprint and the downlink. The associated inversion algorithm is in charge of directly estimating an ideal fused product from the multimodal sources of information, containing the full spectrum information at every pixel. We propose to address this problem with a joint reconstruction of compressed images and data fusion. A full analysis is presented to demonstrate the flexibility and limitations of the proposed approach, to discuss viable optimizations for the acquisition system, and to identify the scenarios that lead to the highest quality reconstruction products.

4.1 Introduction

With the availability of lower budget small satellite carrying high-quality optical imagery [240], on-board image compression has become an increasingly interesting field to compensate for limited on-board resources in terms of mass memory and downlink bandwidth. Many strategies have been developed to address this issue, which either focus on software compression [122], or on the implementation of nonconventional optical devices. The latter approach is the focus of this chapter.

The main motivation behind this work is to propose an acquisition device operating with a single focal plane array (FPA), whose performances in terms of spectral and spatial resolution match those of the fused products of image bundles acquired by more sophisticated Earth observing satellites. In many of such cases, the satellite is equipped with two different classes of sensor. Each class is specialized in capturing

either a high resolution image (HRI) or a low resolution image (LRI), which are respectively characterized by a high spatial and spectral resolution. This separation is due to overcome technological and physical limitations, leaving the task of providing a synthetic image with the best characteristics of each of the two class to the ground segment.

In the case of sensors with the same spatial resolution, many commercial cameras employ CFA patterns, whose operating principle is based on mosaicing different spectral responses on the same focal plane.

We propose here the multiresolution color filter array acquisition (MRCA) concept as a novel design for a nonconventional optical device, based on the assumption that sensors with different characteristics can be accommodated on the same focal plane. With this design, the compressed acquisition contains partial information both from the LRI and the HRI. An example of such acquisitions was shown in the introduction in Fig. 1.5; as the MRCA registers the image over a single array of sensors, the raw product is monochromatic (Fig. 1.5b), but includes both samples from a panchromatic (PAN) and a multispectral (MS), according to the pattern in fig 1.5a.

The associated reconstruction algorithm requires then to solve two problems at once; it has to recover the missing information acquired by each of the two classes of sensors and it has to fuse them to reach the maximum available spatial and spectral resolution. This is a computational imaging problem, as the burden of the data downlink, storage, as well as the amount of dedicated sensors, is significantly reduced at the expense of added processing power performed on the ground segment. The main envisioned application for this setup is aimed at embarking devices based on this architecture over a constellation of low cost satellites.

We present in this chapter a Bayesian formulation of the problem, whose goal is to address the fusion and demosaicing of the compressed acquisition jointly, whose regularization is approached through variational techniques. The proposed model is not exclusively a demosaic-style image reconstruction, as the final product is not the two image sources, but their fused product, nor it is a simple fusion, as the products to process are not well-distinguished multi-modal sources, but rather a lossy compressed combination of the two.

We employ in this chapter the same notation of the previous one, detailed in Section 3.3.

The novel contributions of this chapter include:

- a compression scheme for multiresolution sensors sharing a common focal plane. The optical device can be manufactured with a cheap hardware implementation and easy to embark on board of satellite and avionic platforms. The design is inspired by the theory of CFA [176],
- an inversion framework capable of simultaneously addressing the problem of demosaicing and the fusion of partial multiresolution acquisition. For this scheme, we test a variety of regularization approaches based on the total variation (TV) [200],
- an analysis of the products, which we reconstruct both with our proposed joint approach and with a cascade of classical algorithms of fusion and demosaicing algorithms, applied separately. Their performances are comparing in different scenarios with respect of the number of channels,
- an analysis of the compression power of the proposed approach, in comparison with classical software compression schemes;
- a comparison of CFA-style masks patterns for the distribution of multimodal sensors over the focal plane array, which includes both periodic and pseudo-random designs inspired by the principles of compressed sensing [80].

The chapter is organized as follows: Section 4.2 describes the proposed acquisition system, Section 4.4 presents the proposed inversion model capable to estimate the desired product, Section 4.3 details some general design techniques to optimize the design of the joint compressed acquisition and Section 4.6 presents the related experiments.

4.2 Acquisition system

This section is devoted to the definition of the proposed acquisition system, both in terms of design, manufacture, and mathematical modeling. We describe a CFA-based theoretical model in Section 4.2.1 and some guidelines for the manufacture of an optical device to implements them in Section 4.2.2.

4.2.1 Multiresolution masking

The operating principle of the MRCA is mainly inspired by the CFA technology, whose physical model was detailed in 3.6.1. To briefly summarize the concepts

that are relevant for the current discussion, a CFA/multispectral filter array (MSFA) device acquisition can be modeled as a monochromatic compressed acquisition $\mathbf{U}^{[y]} \in \mathbb{R}^{N_{i_1} \times N_{i_2}}$, where N_{i_1} and N_{i_2} are the number of column and row pixels of the FPA. To simplify the notation, all image variables are reshaped in their lexicographic order so that $\mathbf{y} = \mathrm{matr}\left(\mathbf{U}^{[y]}\right)$ is a column vector of length $N_i = N_{i_1} N_{i_2}$.

This acquisition can be interpreted as a linear combination along the spectral dimension of an ideal demosaiced product $\mathbf{X} \in \mathbb{R}^{N_i \times N_b}$, so that:

$$\mathbf{y} = \sum_{k=1}^{N_b} \mathbf{x}_{:k} \odot \mathbf{h}_{:k} \,, \tag{4.2.1}$$

where $\mathbf{x}_{:k}$ and $\mathbf{h}_{:k}$ are the k-th column of $\mathbf{X}$ and of a given mask $\mathbf{H}$, respectively. Typically, each row $\mathbf{h}_{i:}$ is chosen such that it satisfies the sum-to-one condition:

$$\sum_{k=1}^{N_b} h_{ik} = 1 \,, \qquad \forall i \in [1, \dots, N_i] \,. \tag{4.2.2}$$

For the special case of a binary mask, $\mathbf{h}_{i:}$ is an all-zero vector except for one element, which is equal to one; this is equivalent to choose a single filter for the i-th pixel among a set of spectral responses $\{\xi_k\}_{k \in [1,\dots,N_b]}$.

We aim in this chapter to extend this concept to multiresolution images, by defining two different masks, $\mathbf{H}^{[m]} \in \mathbb{R}^{\frac{N_i}{\rho^2} \times N_b}$ and $\mathbf{H}^{[p]} \in \mathbb{R}^{N_i \times N_{b_p}}$, respectively associated with a LRI $\mathbf{M} \in \mathbb{R}^{\frac{N_i}{\rho^2} \times N_b}$ and a HRI $\mathbf{P} \in \mathbb{R}^{N_i \times N_{b_p}}$.

As the sensors are at a different resolution but target the same scene, it is necessary to consider a common system of spatial coordinates to project them both on the same focal plane array This is achievable by performing an extension by a factor of ρ of both the LRI and its associated mask, as described in Section 3.4.1. By properly choosing the shift in the extension operation, it is possible to keep the centers of the original samples of the LRI within half a pixel of misalignment in the target coordinates of the HRI.

The sparse channels $\mathbf{M}^{\boxdot} \in \mathbb{R}^{N_i \times N_b}$ and $\mathbf{P}^{\boxdot} \in \mathbb{R}^{N_i \times N_{b_p}}$, for the LRI and HRI are thus given with the same amount of pixels as follows:

$$\mathbf{M}^{\boxdot} = \mathbf{M}^{\uparrow} \odot \mathbf{H}^{[m]\uparrow} = \widetilde{\mathbf{M}}^{\uparrow} \odot \mathbf{H}^{[m]\uparrow} \,, \tag{4.2.3a}$$

$$\mathbf{P}^{\boxdot} = \mathbf{P} \odot \mathbf{H}^{[p]} \,, \tag{4.2.3b}$$

where $\mathbf{H}^{[m]\uparrow}$ and $\mathbf{M}^{\uparrow}$ are an extension (zero interleaving) by a factor of ρ. The zero samples in $\mathbf{M}^{\uparrow}$ can be filled with any value as they are masked out, but the most natural choice is to consider its interpolated version $\widetilde{\mathbf{M}^{\uparrow}}$, because, as it will be described in Section 4.4.1, this allows to simplify the direct model for the inversion.

The acquisition can then be modeled as the sum of the sparse channels associated with both the LRI and the HRI:

$$\mathbf{y} = \sum_{k=1}^{N_b} \mathbf{m}_{:k}^{\boxdot} + \sum_{l=1}^{N_{b_p}} \mathbf{p}_{:l}^{\boxdot} \,. \tag{4.2.4}$$

Let $\mathbf{S} = [\widetilde{\mathbf{M}^{\uparrow}}, \mathbf{P}]$ be the row concatenation of the LRI and the HRI, then eq. (4.2.4) can also be rewritten as:

$$\mathbf{y} = \sum_{k=1}^{N_b+N_{b_p}} \mathbf{s}_{:k} \odot \mathbf{h}_{:k} \,, \tag{4.2.5}$$

where $\mathbf{h}_{:k}$ is the r the k-th column of a generic mask $\mathbf{H} \in \mathbb{R}^{N_i \times (N_b+N_{b_p})}$, i.e. such that $\mathbf{H} = [\mathbf{H}^{[m]\uparrow}, \mathbf{H}^{[p]}]$, while $\mathbf{s}_{:k}$ can either describe a channel from the upsampled LRI or from the HRI, depending on the value of k.

If $\mathbf{H}$ is binary and verifies the sum-to-1 condition (4.2.2), then it is possible to give a color coded representation of the associated mask, some examples of which are shown in Fig. 4.3. Compared to the case discussed in Section 3.6, other than the colors for the channel of the LRI, it is necessary to allocate also a new set of colors for the sensors associated with the HRI.

If both the HRI and the LRI have the same pixel depth (i.e., each sample, regardless of the sensor, is represented with the same amount of bits), the **compression ratio** ρ_c, defined as the ratio between the total amount of storage space necessary to the acquisition and the uncompressed sources, is given by:

$$\rho_c = \frac{N_i}{N_i N_{b_p} + N_i N_b/\rho^2} = \frac{\rho^2}{N_b + \rho^2 N_{b_p}} \,. \tag{4.2.6}$$

According to the discussion in Appendix A.1.3, eq. (4.2.5) is a linear operator $\mathbf{y} = \mathbb{A}_h(\mathbf{S}) : \mathbb{R}^{N_i \times (N_b + N_{b_p})} \rightarrow \mathbb{R}^{N_i}$ that can also be described by a multiplication with a matrix $\mathbf{C}$ such as:

$$\mathbf{y} = \mathbf{C}\mathbf{v}^{[s]} = [\mathrm{diag}(\mathbf{h}_{:1}), \mathrm{diag}(\mathbf{h}_{:2}), ..., \mathrm{diag}(\mathbf{h}_{:(N_b+N_{b_p})}] \, \mathrm{vec}(\mathbf{S}) \tag{4.2.7}$$

where $\mathbf{v}^{[s]} = \mathrm{vec}(\mathbf{S})$ is a representation of $\mathbf{S}$ over a single column and $\mathrm{diag}(\mathbf{h}_{:k})$ is a diagonal matrix whose elements on the main diagonal are given by $\mathbf{h}_{:k}$.

4.2.2 Physical implementation

A physical prototype of the proposed scheme can be manufactured with various technologies; in principle, the ideal solution would involve a matrix of acquisition sensors exactly matching the type of CFA structures such as the ones shown in Fig. 4.1. Unfortunately, HRI and LRI sensors generally operate with different technologies, as they are optimized to overcome complementary physical constraints: the former captures more energy from larger ground areas, while the latter from wider bandwidths of wavelengths. For a matrix of exclusively MS sensors, the CFA can be implemented with optical filters realized with patterned optical coatings, made possible by recent advances in micro-lithography and coating technologies [68]. When considering a classical setup, such as in the configuration in which the PAN and MS sensors are separate (and with potentially distinct optical paths), the implementation could be done by separating each MS component with a dispersive element, usually a prism, and let each component pass through an assigned coded aperture, which is in charge of ideally realizing the operation of masking described in eq. (4.2.5). The PAN signal can go through a similar optical processing, with a coded aperture that could implement $\mathbf{H}^{[p]}$; the PAN sensor matrix can also intentionally feature holes in the places where the acquisition is not needed, avoiding the redundancy of acquisition and the need of a mask altogether. The results for each of those operations can be focused appropriately on FPA detector, which is in charge of integrating the post-processed scene to generate $\mathbf{y}$. It is worth noting that the reconstruction method (which will be described below) does not require that the position of each processed sample is the same as described in Section 3.6; any permutation of the samples is allowed, as long as the final position of each is known beforehand by the ground segment. That implies that the two sources can be kept separate and managed by two different FPAs, if the joint focusing would pose any challenge in the implementation [70].

A full analysis of the advantages and disadvantages of physical components capable of manufacturing such a prototype is outside the scope of this work. We just recall here that a certain cost balance has to be reached for the manufacture of MRCA devices, as it overall requires a reduced number of photosensors compared to the separate HRI/LRI acquisition setup, but on the other hand it requires an infrastructure to couple the different sensing technologies on the same FPA.

While technological constraints may still not allow imaging systems at different resolutions on the same focal plane, some recent technology advancements may make some alternative designs available in the near future.

Some promising designs include Onyx [@12] and certain prototypes by Kodak [213] are based on mosaics which mix both wide-band and MS spectral responses, although at the same spatial resolutions. Additionally, the Color SHADES© technology by Silios [@10] allows for a flexible design for the photosites sizes and their associated spectral responses. This solution may also allow to design sensors that realize whatever compromise between the spatial and spectral resolutions.

4.3 Mask design

In the MRCA the final acquisition can be interpreted as a linear combination of samples with a shared coordinate system; the inversion framework that will be proposed in Section 4.4 is in principle applicable to every design following this principle.

However, not every mask pattern is equally effective, as the general design principles described in Section 3.6.2 still apply. In this section we provide a tentative analysis for a general mask design based on compressed sensing (Section 4.3.1), which leads to a random mask pattern (Section 4.3.2); additionally, the design for periodic masks of Section 3.6.2 is extended to the case of multiresolution sensors (Section 4.3.3).

4.3.1 A compressed sensing interpretation

Let $\mathbf{y} = \mathbf{C}\mathbf{v}^{[x]}$ define a linear compression system, described by a multiplication with a matrix $\mathbf{C} \in \mathbb{R}^{N_y \times N_x}$. A proper design of $\mathbf{C}$ must preserve as much information as possible from the input $\mathbf{v}^{[x]} \in \mathtt{E}_x \subseteq \mathbb{R}^{N_x}$ within the acquisition $\mathbf{y} \in \mathtt{E}_y \subseteq \mathbb{R}^{N_i}$.

We discuss here an approach based on compressed sensing [62, 69]. As $\mathbf{v}^{[x]}$ represents a natural scene, all signals belonging to the sample space $\mathtt{E}_x$ of possible

acquisitions can be approximated by its sparse representation in a certain transformed domain (e.g., in its wavelet representation). Specifically, the signals of $\mathtt{E}_x$ are defined N_s-sparse if their representation $\mathbb{L}\left(\mathbf{v}^{[x]}\right)$ in a specific transformed domain have at most N_s nonzero samples ($\|\mathbb{L}(\mathbf{v}^{[x]})\|_0 \leq N_s$).

The target of the compression is to embed the acquisitions so that all signals are well distinguishable in the target domain $\mathbb{E}_y$. In mathematical terms, given any two signals $\mathbf{v}^{[x]}$ and $\mathbf{v}^{[\widehat{x}]} \in \mathtt{E}_x$, one would want to at least ensure that:

$$\mathbf{v}^{[x]} \neq \mathbf{v}^{[\widehat{x}]} \Rightarrow \begin{cases} \mathbf{C}\mathbf{v}^{[x]} \neq \mathbf{C}\mathbf{v}^{[\widehat{x}]}\,, & \forall \mathbf{v}^{[\widehat{x}]} \neq \mathbf{v}^{[x]}\,, \\ \mathbf{C}\mathbf{e} \neq \mathbf{0}_{[N_y \times 1]}\,, & \forall \mathbf{e} \neq \mathbf{0}_{[N_x \times 1]}\,, \end{cases} \tag{4.3.1}$$

where $\mathbf{e} = \mathbf{v}^{[x]} - \mathbf{v}^{[\widehat{x}]}$ defines an error in vector form.

The error $\mathbf{e}$ is a $(2N_s)$-sparse signal, as the N_s sparse samples of $\mathbf{v}^{[x]}$ and $\mathbf{v}^{[\widehat{x}]}$ may be in different positions; this implies that the necessary condition to perfectly reconstruct N_s-sparse signals is that $\mathbf{C}$ has at least $(2N_s)$ linearly independent columns. Given that in eq. (4.2.7) we are considering a compression matrix $\mathbf{C}$ with fixed dimensions, the best we can achieve is to impose that $\mathbf{C}$ is full rank. This is obtained by assuming that $\mathbf{C}$ has no row with all zeros, that is, if each pixel on the FPA is actually capturing a non-masked sample [69].

Unfortunately, the above condition may not be sufficient in most scenarios, as, in noisy environments, signals may have slight deviations on their energy content. A more strict definition is to preserve a certain distance in the observation space, if the inputs are sufficiently distant themselves.

This condition is mathematically formalized by the so called restricted isometry property (RIP) [36]; specifically, an observation matrix $\mathbf{C}$ satisfies the RIP of order N_s if it exists a scalar $\delta_{N_s} \in [0, 1]$ such that:

$$(1 - \delta_{N_s}) \left\|\mathbf{v}^{[x]}\right\|_2^2 \leq \left\|\mathbf{C}\mathbf{v}^{[x]}\right\|_2^2 \leq (1 + \delta_{N_s}) \left\|\mathbf{v}^{[x]}\right\|_2^2 . \tag{4.3.2}$$

In other words, if $\mathbf{C}$ satisfies the RIP of order $2N_s$, the distance (measured by the ℓ_2 norm) among N_s-sparse signals in $\mathtt{E}_x$ is approximately preserved in $\mathtt{E}_y$. The scope of this work does not include the analysis of the minimum amount of observations N_i to provide N_s-sparse signals satisfying the RIP; however, we remind here the well-known result in the literature that if $\delta_{2N_s} \leq 1/2$, the amount of observations which are necessary for $\mathbf{C}$ to satisfy the RIP of order $2N_s$ is $N_i \geq 0.28 N_s \log(N_x/N_s)$ [53].

In our analysis, we instead fix the amount of observations N_i and we aim to design a compression algorithm that maximizes the RIP property of order $2N_s$ with the highest possible N_s. In the next section, we discuss some tentative designs to achieve this goal.

4.3.2 Random mask patterns

Some deterministic patterns which allow to fulfill the RIP of a certain order have been proposed, however we will not employ them as those constructions require unreasonably large values for N_i [56].

However, a theorem [160] states that, if all elements of each row of $\mathbf{C}$ are chosen according to a subgaussian distribution with a minimum value of observations $N_i = k_1 N_s \log(N_x/ns)/\delta^2_{2N_s}$, then $\mathbf{C}$ satisfies the RIP with probability $1-2\exp(-k_2\delta^2_{2N_sN_i})$, where k_1 and k_2 are two constant values that depend only on the characteristics of the chosen distribution.

Subgaussian distributions include, among other ones, multi-variate Gaussian, bounded and Bernoulli distributions. Unfortunately, most of the elements of our matrix $\mathbf{C}$ in eq. (4.2.7) are null, since $\mathbf{C}$ is a representation of the masking operation (4.2.5), and we have no possibility to modify their values.

We are just allowed to adjust the elements of the mask $\mathbf{H}$, as long as the sum-to-1 condition of eq. (4.2.2) is verified; we review here three possible strategies to randomize these elements, which will be tested in Section 4.6.6.

- **Compressive coded aperture spectral imaging (CASSI)**: In this setup the vectors $\{\mathbf{h}_{:k}\}$ assigned to each channel are random binary masks; in other terms, for the k-th channel, the i-th element h_{ik} can be either 0 or 1, with a certain possibility. The elements of the mask $\mathbf{H}$ are consequently independent and identically distributed (i.i.d.) Bernoulli random variables. In this case, the matrix $\mathbf{C}$ of eq. (4.2.7) may not be full rank, as a certain vector $\mathbf{h}_{i:}$ assigned to the i-th pixel may be made up of all zeros. The CASSI mask design is inspired by the digital micromirror device (DMD) physical realization described in eq. (3.6.5b) in Section 3.6.2. In the experimental section, all the vectors $\mathbf{h}_{i:}$ are scaled by a factor of $1/(N_b + N_{b_p})$ to satisfy the sum-to-1 condition of eq. (4.2.2).
- **Random pick (RAND)**: In this configuration, the vector $\mathbf{h}_{i:}$ associated with the i-th pixel is made of all-zeros, except for a single position, which is equal to 1. This position is assigned randomly, so the RAND design is equivalent

to randomly assigning a spectral response from the set $\{\xi_k\}_{k\in[1,\dots,N_b]}$ to each pixel. The resulting matrix $\mathbf{C}$ is automatically fully rank, but its values are only softly randomized, as the corresponding mask $\mathbf{H}$ is binary. This randomized mask design, which is equivalent to an MSFA with completely randomized pattern, was proposed by various authors [152] and its effectiveness was investigated by Amba et al. [12].

- **Dirichlet distribution based (DIRI)**: In this approach, that we proposed in [188], each vector $\mathbf{h}_{i:}$ is filled with nonbinary weights generated according to a flat Dirichlet distribution [136], which enforces a uniform distribution on a $(N_b + N_{b_p} - 1)$-dimensional simplex. This constraint, which is equivalent to verify the condition (4.2.2), can be implemented by the Algorithm 2 [57]. As the matrix $\mathbf{H}$ is nonbinary, the spectral response associated with each pixel is a generally different linear combination of the ones available in the set $\{\xi_k\}_{k\in[1,\dots,N_b]}$.

On a practical level, this last design would correspond to having a matrix of sensors with vastly different spectral responses. For silicon based technologies, a possibility could be to filter a different set of wavelengths for each pixel from a wide-band response; e.g., this setup could be realized with **COLOR SHADES** [@10] by SILIOS technologies, which combines thin film deposition and micro/nano-etching processes onto a silica substrate to provide band-pass filters in the visible and near infrared range [140]. Since the resulting responses may be wider compared to the usual MS, the larger amount of incoming light could overcome some of the signal to noise ratio (SNR) limitations of available sensors, possibly increasing the spatial resolution of the sensor.

If the need arises, considering without loss of generality the HRI to be a PAN, it is possible assign a less/more prominent contribution of the weights assigned to the HRI. I.e., a generic Dirichlet distribution $\mathcal{D}([\mathbf{1}_{1\times N_b}, \alpha])$ generates random samples for the spectrum $\mathbf{h}_{i:}$ of the i-th pixel. As a result, the average of the elements in $\mathbf{h}_{:,N_b+1}$ is α times larger than the ones in $\{\mathbf{h}_{:k}\}_{k\in[1,\dots,N_b]}$.

4.3.3 Periodic patterns for combined masks

In this section we investigate some proposed deterministic design for the arrangements of multiresolution sensors on the FPA, which we aim to physically implement in the MRCA. For simplicity, we will assume that the HRI is a (monochromatic) PAN.

Algorithm 2: Pseudorandom generation of a vector of N_a i.i.d. r.v. uniformly distributed over a $(N_a - 1)$-simplex.

Result: Vector $\mathbf{v}$ of samples uniformly distributed over a $(N_a - 1)$-simplex
Generate N_a random variables (r.v.s) $v_i \sim \mathcal{U}[0, 1]$, with $i \in [1, \dots, N_a]$
Assign them to a column vector $\mathbf{v} \in \mathbb{R}^{N_a}$, whose elements are $\{v_i\}_{i \in [1, \dots, N_a]}$
$\mathbf{v} \leftarrow -\ln(\mathbf{v})$
$\overline{\mathbf{v}} = \sum_{i=1}^{N_a} v_i$
$\mathbf{v} \leftarrow \frac{\mathbf{v}}{\overline{\mathbf{v}}}$
return $\mathbf{v}$

The most straightforward design, which we denote with **panchromatic coverage (COVE)**, consists in selecting a classic MSFA mask $\mathbf{H}^{[m]} \in \mathbb{R}^{N_i/\rho^2 \times N_b}$ and performing an expansion $\mathbf{H}^{[m]\uparrow}$ by a factor of ρ (Section 4.2.1). The expansion $\mathbf{H}^{[m]\uparrow}$ is nothing else than an operation of zero interleaving, hence it introduces a certain amount of zero pixels in the mask, which in the COVE design are assigned to the PAN. Formally, the elements $h_i^{[p]}$ of the mask $\mathbf{H}^{[p]} \in \mathbb{R}^{N_i \times 1}$ are given by:

$$h_i^{[p]} = \begin{cases} 1 & \text{if } \sum_{k=1}^{N_b} h_{ik}^{[m]\uparrow} = 0\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{4.3.3}$$

The combined mask $\mathbf{H} = [\mathbf{H}^{[m]\uparrow}, \mathbf{H}^{[p]}]$ is shown in Fig. 4.1c for a scale ratio $\rho = 2$.

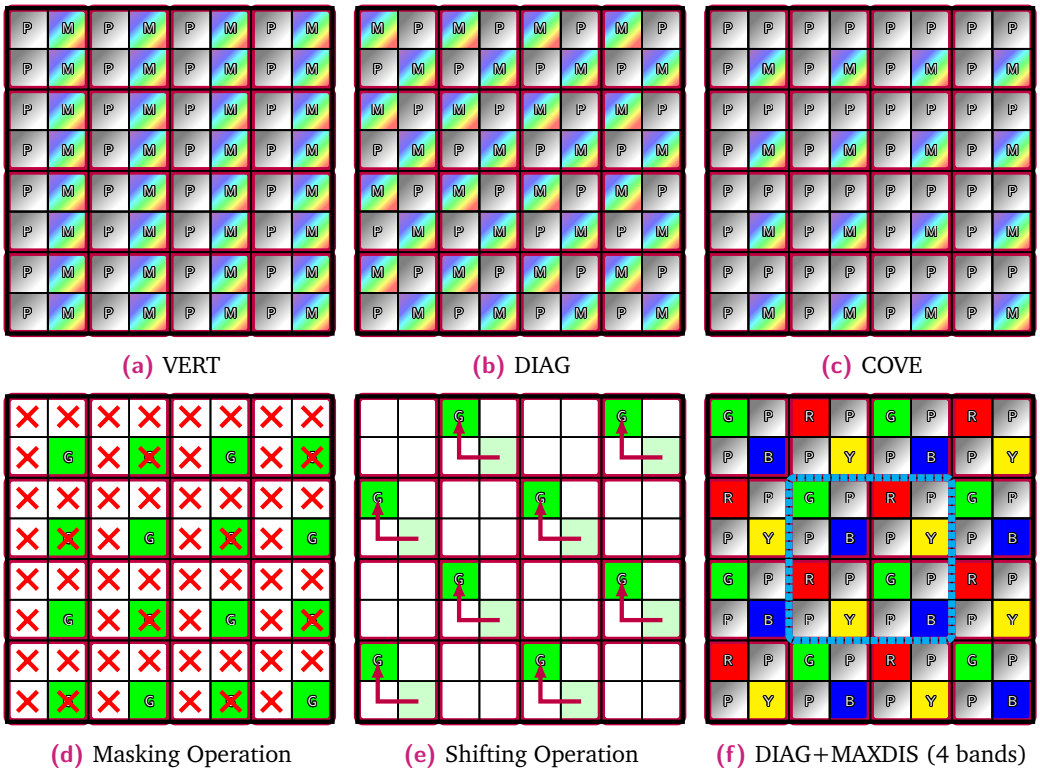
The COVE pattern can be interpreted as a dominating band pattern, where the occurrence of the HRI and LRI sensors are in a ratio of $(\rho^2 - 1) : 1$. If a larger amount of spectral information is necessary, some different designs may be useful to increase the amount of sensors assigned to the LRI. We propose here some alternatives, based on the patterns of commercial optical devices with dominant wideband sensors, such as the ones by Teledyne and Kodak [@12, 213], shown in Fig. 3.8f to 3.8i.

Let vertical (VERT) and diagonal (DIAG) define the patterns for the HRI mask $\mathbf{H}^{[p]}$ shown in Fig. 4.1a and 4.1b, respectively. If ρ is even, these patterns can be separated into square blocks of size $\rho \times \rho$, with a given amount of empty slots N_e. I.e., in Fig. 4.1b, where $\rho = 2$, each 2×2 square (delimited by red borders) has $N_e = 2$ empty slots. If the periodic mask $\mathbf{H}^{[m]}$ associated with the LRI follows the condition $\sum_{k=1}^{N_b} h_{ik}^{[m]} = N_e$ for all $i \in [1, \dots, N_i]$, then the filters associated with the i-th pixel can be associated with the empty slots in the associated block.

Mathematically, this operation of filling slots is equivalent to an appropriate horizontal and vertical shift from the associated expanded mask $\mathbf{H}^{[m]\uparrow}$, that yields the following description of the acquisition:

$$\mathbf{y} = \sum_{l=1}^{N_{b_p}} \mathbf{p}_{:l} \odot \mathbf{h}_{:l}^{[p]} + \sum_{k=1}^{N_b} \left(\mathbf{m}_{:k}^{\uparrow} \odot \mathbf{h}_{:k}^{\uparrow} \right)_{\downarrow r_k} \tag{4.3.4}$$

where $(\cdot)_{\downarrow r_k}$ defines an operation of circular shift by r_k, that is in charge to placing the sample in the appropriate position. As the mask is periodic, this shift is the same for every pixel assigned to the same class. This process is shown in Fig. 4.1d to 4.3c, for $\rho = 2$ and for a mask inspired by the 4-band maximum distance (MAXDIS) described in Section 3.6.2.



(a) VERT (b) DIAG (c) COVE

(d) Masking Operation (e) Shifting Operation (f) DIAG+MAXDIS (4 bands)

Fig. 4.1. The first row shows some basic patterns for the positioning of PAN samples, marked with "P" for periodic masks. The remaining rainbow colored spots, labelled with "M" denote the spots to be assigned to any sample from the uncompressed MS. The second row shows the operation of masking and shifting of the green MS pixels over empty slots, together with an example of the final combined PAN/MS mask. Dark red borders mark the $\rho \times \rho$ regions of the PAN which map to a single position of the original MS, while the cyan dotted cage shows the periodicity of the combined mask.

4.4 The inversion protocol

According to the Bayesian formulation described in Chapter 2, the inversion framework allows to find an estimation $\widehat{\mathbf{X}}$ of $\mathbf{X}$ by minimizing a certain cost function. This is composed of a data fidelity term, in charge of comparing the acquisition to an expected value of the transformation, and a regularization term, which takes into account the prior information on the desired product, aimed at resolving the ill-posedness of the problem.

This section is divided as follows: the direct model used for the data term is defined in Section 4.4.1 and is employed in the definition of the complete objective function in Section 4.4.2. The full inversion algorithm is then detailed in Section 4.4.4.

4.4.1 Direct model

The direct model is a description of the optical transformations linking the desired product $\mathbf{X}$ with the acquisition $\mathbf{y}$. This transformation is here considered linear and denoted by $\mathbb{A}(\mathbf{X})$.

In general terms, $\mathbb{A}$ may combine both the effect of the compression and the spectral/spatial degradation into a single functional; however, to simplify the analysis, these effects are described here as a cascade of systems, which are shown in Fig. 4.2.

The first leg of the transformation performs a spectral and spatial degradation, to obtain the k-th channel of a simulated HRI $\mathbf{P}$ and of an upscaled LRI $\widetilde{\mathbf{M}}^{\uparrow}$, with the same procedure described in Section 3.5.3:

$$\begin{cases} \mathbf{p}_{:k} = \displaystyle\sum_{l=1}^{N_b} w_{kl}\mathbf{x}_{:l} & \forall l \in \left[1, \dots, N_{b_p}\right] \quad \text{(4.4.1a)} \\ \widetilde{\mathbf{m}}^{\uparrow}_{:k} = \mathbf{x}_{:k} \ * \ \mathbf{b}_{:k} & \forall k \in [1, \dots, N_b] \quad \text{(4.4.1b)} \end{cases}$$

where $\mathbf{b}_{:k}$ is a Gaussian convolution kernel, matching the modulation transfer function (MTF) of the k-th LRI sensor, and $\{w_{kl}\}_{l\in[1,\dots,N_b]}$ denote the weighting coefficients. The expression of w_{kl} is obtained by evaluating the ratio of area overlap between the spectral responses of each LRI sensor and that of the l-th band of the HRI. In the expression (4.4.1), the LRI is not decimated for convenience, as the relevant samples can be ignored when masked.

If we also consider the compression step $\mathbb{A}_h(\cdot)$ described in Section 4.6.3, the full direct model is thus given by:

$$\mathbf{y}^{[sim]} = \mathbb{A}_h(\mathbb{A}_d(\mathbf{X}))\,. \tag{4.4.2}$$

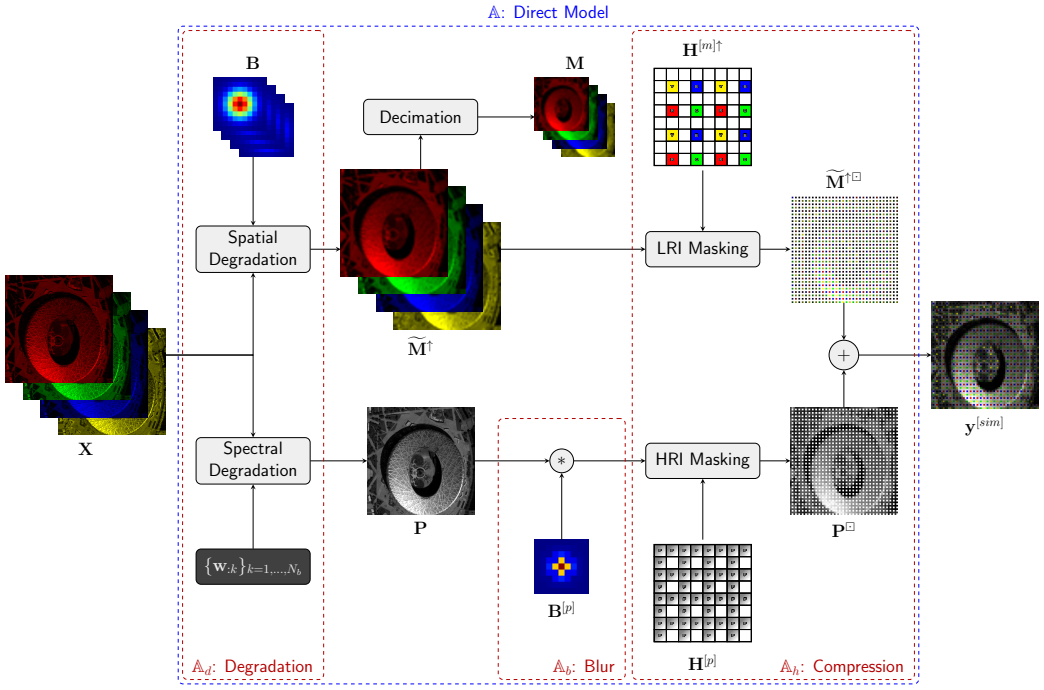


Fig. 4.2. Scheme of the direct model for the relationship between the ideal reconstruction product $\mathbf{X}$ and the acquisition, as described in Section 4.4.2. In this representation, the HRI and LRI (whose role is taken here by the PAN and MS) compression is completely separated. The products shown in the figure are color coded 4-bands bundles, and the associated positions on the MS mask take samples only from matching colors.

4.4.2 Cost function

The direct model (4.4.2) can be embedded in a classical Bayesian formulation for an inverse problem. We assume here that the noise component associated with the operation of either spatial or spectral degradation is an additive i.i.d. Gaussian distribution with zero mean. According to eq. (4.2.5), as each of the observation samples is given by a linear combination of different samples from the multiresolution sensor readouts, the noise associated with the acquisitions

themselves are still Gaussian and independent. Under these conditions, the inversion problem is equivalent to the following minimization:

$$\widehat{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2} \left\| \mathbb{A}_h(\mathbb{A}_d(\mathbf{X})) - \mathbf{y} \right\|_2^2 + g_R(\mathbf{X}) \,, \tag{4.4.3}$$

where $g_R(\mathbf{X})$ is a generic regularization function.

We aim here to show how the proposed model (4.4.2), denoted with **Demosaic+Fusion**, is a generic case of the Bayesian formulation of the sharpening and the demosaic problem, described in Section 3.5.3 and 3.7.5, respectively.

- **Fusion Only**: This case can be obtained by simply substituting the compression operator $\mathbb{A}_h$ with an identity; this describes a sharpening operation:

$$\widehat{\mathbf{X}} = \frac{1}{2} \left\| \mathbb{A}_d(\mathbf{X}) - \mathbf{S} \right\|_F^2 + g_R(\mathbf{X}) \tag{4.4.4a}$$

$$= \frac{1}{2} \left(\left\| \mathbb{A}_p(\mathbf{X}) - \mathbf{P} \right\|_F^2 + \left\| \, \mathbb{A}_m(\mathbf{X}) - \widetilde{\mathbf{M}}^{\uparrow} \right\|_F^2 \right) + g_R(\mathbf{X}) \,, \tag{4.4.4b}$$

  where $\mathbf{S} = \left[\widetilde{\mathbf{M}}^{\uparrow}, \mathbf{P} \right]$ is the concatenation of the LRI and interpolated HRI. This is similar to the formulation of eq. (3.5.9), except that the data fidelity term for the LRI is evaluated at the scale of the HRI; additionally, the coefficient $\check{\lambda}_m$ that weights each of the two Frobenius norms is set to 1.

- **Demosaic Only**: For this case, if we consider the degradation operator $\mathbb{A}_d$ as an identity, the masking operator $\mathbb{A}_h$ has to be slightly modified, as the dimensions of $\mathbf{X} \in \mathbb{R}^{N_i \times N_b}$ differ from those of $\mathbf{S}$. However, in principle, this operation may be performed by just considering a different mask $\mathbf{H}^{[x]}$, operating on N_b bands, such as the ones considered in Section 3.6.2. The equivalent cost function then becomes:

$$\widehat{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2} \left\| \sum_{k=1}^{N_b} \mathbf{x}_{:k} \odot \mathbf{h}_{:k}^{[x]} - \mathbf{y} \right\|_2^2 + g_R(\mathbf{X}) \tag{4.4.5}$$

4.4.3 Regularization approaches

The problem of eq. (4.4.3) is intrinsically ill-posed, as the amount of samples that represent the acquisition is lower than that of the desired product, which demands a regularization procedure. In our context, as the direct model is seen as a cascade of two operators, the choice of the regularizer must provide a good compromise to solve the two inversion problems associated with each operator.

In particular, the inversion problem (4.4.5) can be seen as an inpainting problem, as it is equivalent to recover missing samples over a regular grid. The regularization approach of this thesis employs variational methods, but this may be a limitation, as it was proven that the total variation is not able to preserve long elongated structures if applied to inpainting problems [41].

To alleviate this issue, we propose to reframe the inversion as a magnification problem. Specifically we introduce an additional blur operator $\mathbb{A}_b$ for the HRI $\mathbf{P}$, which is defined as:

$$\mathbb{A}_b(\mathbf{P}) = \mathbf{P} \ * \ \mathbf{B}^{[p]}\,, \tag{4.4.6}$$

where $\mathbf{B}^{[p]}$ is a low pass filtering kernel. In the experimental section, this is defined with a 21-tap square matrix kernel, which implements a Butterworth filter with normalized cutoff frequency equal to $1/d_b$. d_b is a (generally not integer) user defined parameter, which defines the radius of the blurring filter in the spatial dimension.

As the blurring operation is performed after the spatial degradation and before the masking, this allows allows some leeway to introduce some information from the masked pixels of the HRI in the data term, at the cost of partially sacrificing the spatial resolution of the final product.

The resulting objective function we propose is thus in the form:

$$\widehat{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2} \left\| \mathbb{A}_h(\mathbb{A}_b(\mathbb{A}_d(\mathbf{X})) - \mathbf{y} \right\|_2^2 + \check{\lambda} g(\mathbb{L}(\mathbf{X}))\,, \tag{4.4.7}$$

where $\check{\lambda}$ is a user-defined scalar, and $\mathbb{L}(\cdot) : \mathbb{R}^{N_i \times N_b} \to \mathtt{E}_u$ is a domain transformation operator and $g : \mathtt{E}_u \to \mathbb{R}^+$ is a metric function.

The different combination of $\mathbb{L}$ and $g(\cdot)$ defines the regularization approach.

In this thesis we will employ three different definitions for the $\mathbb{L}$ operator, which correspond to as many approaches for the discretizations of the Rudin-Osher-Fatemi (ROF) model [200]. These are the classic **TV**, defined in eq. (2.2.17), the **upwind total variation (UTV)** [39], defined in eq. (2.2.20), and the **Shannon total variation (STV)** [1].

With regard to the operator $g(\cdot)$, we employ the framework of the collaborative total variation (CTV) [65, 66]; in this interpretation, the function $g(\cdot)$ is a combination of norms $\|\cdot\|_{p_1p_2p_3}$, which are applied in order to the dimension of the gradients, of the channels and the pixels, respectively, as described in Section 2.2.3. As it will be discussed in detail in the experimental section, we experienced that the norm

ℓ_{221}, also known as vector total variation (VTV), is a good compromise between reconstructed image quality and computational time, while the norm $S_2\ell_1$ provides the best overall results. The nuclear norm S_2 whitens the correlated information of the spatial gradients and of the channels.

This framework will be compared to the least absolute shrinkage and selection operator (LASSO) protocol [218] employed for the inversion of the CASSI acquisitions [14]:

$$\widehat{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2} \|\mathbb{A}_h(\mathbb{A}_d(\mathbf{X})) - \mathbf{y}\|_2^2 + \check{\lambda} \| \mathbb{L}(\mathbf{X})\|_{111} \,, \tag{4.4.8}$$

where $\mathbb{L}$ is a Symlet-8 discrete wavelet transform (DWT) transformation in the spatial domain and a discrete cosine transform (DCT) in the spectral domain.

4.4.4 Implementation details

The expression of eq. 4.4.7 is composed of a differentiable data fidelity term and a regularization term whose metric function $g(\cdot)$ is a lower semi-continuous convex function.

These conditions are sufficient to obtain the estimation $\widehat{\mathbf{X}}$ with the iterative procedure known as Loris-Verhoeven algorithm [148], described in Section 2.3.2.

The implementation requires the knowledge of the adjoint and the operator norm of both linear operators $\mathbb{A}$ and $\mathbb{L}$.

As $\mathbb{A}(\cdot) = \mathbb{A}_h(\mathbb{A}_b(\mathbb{A}_d(\cdot)))$ is a cascade of operators, its adjoint $\mathbb{A}^*$ is given by a cascade of adjoint operators $\mathbb{A}^*(\cdot) = \mathbb{A}_d^*(\mathbb{A}_b^*(\mathbb{A}_h^*(\cdot)))$.

As a result of Cauchy inequality, its operator norm is bounded by the sum of the individual norms:

$$\| \mathbb{A} \|_{op} \leq \| \mathbb{A}_h \|_{op} + \| \mathbb{A}_b \|_{op} + \| \mathbb{A}_c \|_{op} \,. \tag{4.4.9}$$

The individual adjoint and operator norms are derived in Appendix A.1, specifically:

- $\mathbb{A}_h$ is a masking operator, described in Appendix A.1.3;
- $\mathbb{A}_b$ is a convolution operation, described in Appendix A.1.2;

- $\mathbb{A}_d(\cdot) = [\mathbb{A}_m(\cdot), \mathbb{A}_p(\cdot)]$ is made up of two components. $\mathbb{A}_m$ is also obtained by a convolution operation, while $\mathbf{P} = \mathbb{A}_p(\mathbf{X})$ from eq.(4.4.1a) can be rewritten as a masking operation, as the i-th pixel of the k-th band of the HRI can be rewritten as:

$$p_{:k} = \sum_{l=1}^{N_b} \mathbf{x}_{:l} \odot \left(w_{kl} \mathbf{1}_{[N_i \times 1]} \right) . \tag{4.4.10}$$

  The overall adjoint operator is then obtained as $\mathbb{A}_d^*(\cdot) = \mathbb{A}_m^*(\cdot) + \mathbb{A}_p^*(\cdot)$ and the operator norm is $\| \mathbb{A}_d \|_{op} = \| \mathbb{A}_m \|_{op} + \| \mathbb{A}_p \|_{op}$.

The definition of the adjoint and the operator norm of $\mathbb{L}$ is derived in Appendix A.1.5 in the case of the classic TV, while the reader may refer to the dedicated literature for the other variants [39, 1].

The expression of the proximal operators relative to the metric function $g(\cdot)$ are instead given in [66], for the norm that we employ in this work.

As Loris-Verhoeven is a primal-dual algorithm, it requires an initialization, both in the reconstruction space $\mathtt{E}_x$ and in the transformed space $\mathtt{E}_u$, although the algorithm converges to the same solution, as the objective function is convex.

In the thesis, the reconstruction variable is initialized with $\mathbb{A}^*(\mathbf{y})$, and the dual variable with $\mathbb{L}(\mathbb{A}^*(\mathbf{y}))$. When multiple values of the regularization parameter $\check{\lambda}$ are tested, the solution obtained with the most recent is used as initialization, instead, to speed up the convergence [82].

The algorithm can either terminate after a fixed number of iterations or by evaluating that the variation of the cost function at a certain iteration is below a certain threshold. The full procedure is described by the Algorithm 3.

For our problem we propose to employ this algorithm mostly for its enormous flexibility, as it allows to test for a great variety of regularizers, as long is is possible to define for them a proximal operator, which is in general a more relaxed condition than just having a gradient. We preferred the Loris-Verhoeven algorithm over more complicated alternatives, such as the alternating direction method of multipliers (ADMM) or the Chambolle-Pock as our fidelity term assumes Gaussian noise; if a different choice is taken, the problem can still be tackled with one of those solutions, but at the cost of slower computational time.

Algorithm 3: Inversion procedure described in Section 4.4.4.

Result: Estimation of the fused product $\hat{\mathbf{X}}$

Input:

- Acquisition: $\mathbf{y}$
- Masking Matrices: $\{\mathbf{h}_{:k}\}_{k\in[1,\dots,N_b+N_{b_p}]}$
- Spectral degradation coefficients: $\{w_{kl}\}_{\substack{k\in[1,\dots,N_{b_p}]\\ l\in[1,\dots,N_b]}}$
- Spatial degradation kernels: $\{\mathbf{b}_{:k}\}_{k\in[1,\dots,N_b]}$
- Diameter of the blurring matrix: d_b (default: 1)
- Regularization parameter: $\check{\lambda}$ (default: $10^{-3}\max(\mathbf{y})$)
- Over-relaxation parameter: $\check{\rho}$ (default: 1.9)
- Maximum amount of iterations: $q^{[max]}$ (default: 250)

Preprocessing:

- Direct model operator: $\mathbb{A}(\cdot)=\mathbb{A}_h(\mathbb{A}_b(\mathbb{A}_d(\cdot)))$ where:
  - $\mathbb{A}_d(\cdot)$ from eq. (4.4.1),
  - $\mathbb{A}_b(\cdot)$ from eq. (4.4.6),
  - $\mathbb{A}_h(\cdot)$ from eq. (4.2.5).
- Its adjoint $\mathbb{A}^*$ and its operator norm $\|\mathbb{A}\|_{op}$ as described in Section 4.4.4;
- Operator $\mathbb{L}$ to choose among:
  - Classic TV, from eq. (2.2.17),
  - UTV, from eq. (2.2.20),
  - STV [1].
- Their adjoint $\mathbb{L}^*$ and operator norm $\|\mathbb{L}\|_{op}$, as described in Section 4.4.4;
- Proximal operator $\mathrm{prox}_{\check{\gamma}g^\star}$ of the Fenchel conjugate $g^\star$ of g [66];
- Apply the histogram matching procedure of eq. (4.6.1)

Initialization:

- $\check{\tau}=0.99/\|\mathbb{A}\|_{op}^2$
- $\check{\sigma}=1/(\check{\tau}\|\mathbb{L}\|_{op}^2)$
- $\mathbf{X}^{(0)}=\mathbb{A}^*(\mathbf{y})$
- $\mathbf{U}^{(0)}=\mathbf{Y}$

while $q<q^{[max]}$ **do**

$\mathbf{U}^{(q+\frac{1}{2})}\leftarrow \mathrm{prox}_{\check{\sigma}(\check{\lambda}g^\star)}\left(\mathbf{U}^{(q)}+\check{\sigma}\,\mathbb{L}\left(\mathbf{X}^{(q)}-\check{\tau}\left(\mathbb{A}^*(\mathbb{A}(\mathbf{X}^{(q)}-\mathbf{y})+\mathbb{L}^*(\mathbf{U}^{(q)})\right)\right)\right)$

$\mathbf{X}^{(q+1)}\leftarrow \mathbf{X}^{(q)}-\check{\rho}\check{\tau}\left(\mathbb{A}^*(\mathbb{A}(\mathbf{X}^{(q)})-\mathbf{y})+\mathbb{L}^*\left(\mathbf{U}^{(q+\frac{1}{2})}\right)\right)$

$\mathbf{U}^{(q+1)}\leftarrow \mathbf{U}^{(q+1)}+\check{\rho}\left(\mathbf{U}^{(q+\frac{1}{2})}-\mathbf{U}^{(q)}\right)$

$q\leftarrow q+1$

return $\widehat{\mathbf{X}}=\mathbf{X}^{(q^{[max]})}$

4.5 Related works

While the concept of the MRCA, intended as a shared focal plane for multiresolution sensors is, to the best of our knowledge, a novelty, some compressed acquisition systems that employ multiresolution data were already proposed in the literature. Espitia et al. [70], in particular, proposed a fusion model in which both HRI and the LRI are available as two compressed acquisitions (respectively, $\mathbf{y}^{[p]}$ and $\mathbf{y}^{[m]}$) through two separate single dispersion CASSI devices.

The direct model for each of the available products is given by eq. (3.6.5b), applied to both the ideal HRI and the LRI:

$$\mathbb{A}_h^{[p]}(\mathbf{P}) = \sum_{k=1}^{N_{b_p}} \left(\left[\mathbf{p}_{:k} \odot \mathbf{h}_{:k}^{[p]};\, \mathbf{0}_{[(k-1)N_{i_1} \times 1]} \right] \right)_{\downarrow((k-1)N_{i_1})} , \tag{4.5.1a}$$

$$\mathbb{A}_h^{[m]}(\mathbf{M}) = \sum_{k=1}^{N_b} \left(\left[\mathbf{m}_{:k} \odot \mathbf{h}_{:k}^{[m]}; \mathbf{0}_{[(k-1)N_{i_1}/\rho \times 1]} \right] \right)_{\downarrow((k-1)N_{i_1}/\rho)} , \tag{4.5.1b}$$

where $(\cdot)_{\downarrow k}$ denotes a circular shift of the vector by k elements. In the natural image representation this is equivalent, for each of the two products, to an overlap of the sparse channels, after an appropriate shift in the horizontal direction.

The estimated inversion is then performed by minimizing the following cost function, inspired by the Bayesian formulation of the sharpening problem (3.5.9):

$$\widehat{\mathbf{X}} = \arg \min_{\mathbf{X} \in \mathbb{R}^{N_i \times N_b}} \frac{1}{2} \left\| \mathbb{A}_h^{[p]}(\mathbb{A}_p(\mathbf{X})) - \mathbf{y}^{[p]} \right\|_2^2 + \frac{\check{\lambda}_m}{2} \left\| \mathbb{A}_h^{[m]}(\mathbb{A}_{m\downarrow}(\mathbf{X})) - \mathbf{y}^{[m]} \right\|_2^2 + \check{\lambda} \left\| \mathbb{L}(\mathbf{X}) \right\|_{111} , \tag{4.5.2}$$

where $\mathbb{A}_h^{[p]}(\cdot)$ and $\mathbb{A}_h^{[m]}(\cdot)$ are given by eq.s (4.5.1a) and (4.5.1b), respectively, while (L) is a DWT in the spatial domain and a DCT in the spectral domain. $\mathbb{A}_p$ and $\mathbb{A}_{m\downarrow}$ are respectively the spectral degradation and matched downsampling defined in Section 3.5.3.

The availability of two separate products does not allow for a fair comparison with our compression scheme, as it reaches a different compression ratio. The model of eq. (4.5.1) can however be slightly modified to generate an acquisition $\mathbf{y}$ over a single plane:

$$\mathbf{y} = \sum_{k=1}^{N_b} \left(\left[\mathbf{s}_{:k} \odot \mathbf{h}_{:k};\, \mathbf{0}_{[(k-1)N_{i_1} \times 1]} \right] \right)_{\downarrow((k-1)N_{i_1})} , \tag{4.5.3}$$

where $\mathbf{s}_{:k}$ can be either a channel from the HRI or from the interpolated LRI $\widetilde{\mathbf{m}}^{\uparrow}_{:k}$. This design will be employed in the comparison of Section 4.6.3.

Another formulation was given by Fu et al. [83], which proposed a fusion framework for an uncompressed LRI with a CFA-based compressed acquisition HRI. The proposed estimation is given by:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2} \left(\left\| \mathbb{A}_h^{[p]}(\mathbb{A}_p(\mathbf{X})) - \mathbf{y}^{[p]} \right\|_2^2 + \|\mathbb{A}_m(\mathbf{X}) - \mathbf{M}\|_F^2 \right) + \check{\lambda} \, \|\mathbb{L}_{SK}(\mathbf{X})\|_{S_1 \ell_1} \,, \tag{4.5.4a}$$

where we can immediately notice that there is no compressed acquisition for the LRI $\mathbf{M}$.

The linear operator $\mathbb{L}_{SK}(\mathbf{X})$ denotes a segmentation of $\mathbf{X}$ into square blocks of user defined sizes $N_x \times N_x$, expressed in lexicographic order [1]. The penalization term induces a low-rank decomposition of each multichannel block, which is imposed by minimizing their singular values.

4.6 Experiments

To prove the effectiveness of our inversion framework, in the following sections we analyse the proposed model under different points of view, whose setup is described in Section 4.6.1. In particular we perform a comparison with naive software compression algorithms in Section 4.6.3, we compare the effectiveness of the proposed model considering the two problems both in cascade and separated in Section 4.6.4, we analyse the flexibility of the algorithm with different channel configurations in Section 4.6.5 and we explore some more advanced designs for masks in Section 4.6.6. An analysis of the parameters of is finally given in Section 4.6.7.

[1] With the formalism described in Section 6.1.1 and 6.2.1, the linear operator $\mathbb{L}_{SK}(\mathbf{X})$ can be formally described as:

$$\mathbb{L}_{SK}(\mathbf{X}) = \mathrm{matr}\left(\mathrm{stack}\left(\mathrm{matr}^{-1}(\mathbf{X}), \mathbf{R}^{[0]}, [N_x, N_x]\right)\right) \,, \tag{4.5.5}$$

where $\mathbf{R}^{[0]}$ is a list of center coordinates arranged over a regular grid, whose horizontal and vertical spacing is equal to N_x.

4.6.1 Experimental setup

In the following sections we analyse multiple scenarios aimed at the estimation $\hat{\mathbf{X}}$ of an ideal uncompressed image $\mathbf{X}$, as described in Section 4.6.4. We always consider the situation in which the role of the HRI is taken by a PAN and the role of the LRI by a MS. Each test configurations is detailed in the dedicated section, but the validation protocol is common to every experiment. The validation is based on Wald's protocol for reduced resolution validation [227] described in Section 3.8. In this setup, we assume to have an ideal MS acquisition $\mathbf{X}$, which we label **GT** and a simultaneously captured PAN.

The PAN and the GT are both downsampled with a blurring kernel, respectively realized with a bicubic filter and a set of filters matched to the MTF of the MS sensors. The obtained acquisitions are then processed with the compressed acquisition model $\mathbb{A}_h$, which depends on the compressed design under test, to generate a simulated acquisition $\mathbf{y}^{[sim]}$.

Each tested reconstruction algorithm is then applied to $\mathbf{y}^{[sim]}$ and the product $\widehat{\mathbf{X}}$ is compared with the GT, evaluating its quality through the set of indices described in Section 3.8.1.

Unless otherwise noted, the interpolated MS, denoted with interpolated image (EXP), is realized with a 11-order Lagrange interpolator. In the spectral degradation model $\mathbb{A}_p$, the weighing coefficients are always set as equal to $w_{1,k} = 1/N_b$, for all $k \in [1, \dots, N_b]$, so that the PAN is modeled as the average of the channels of the GT. [2]

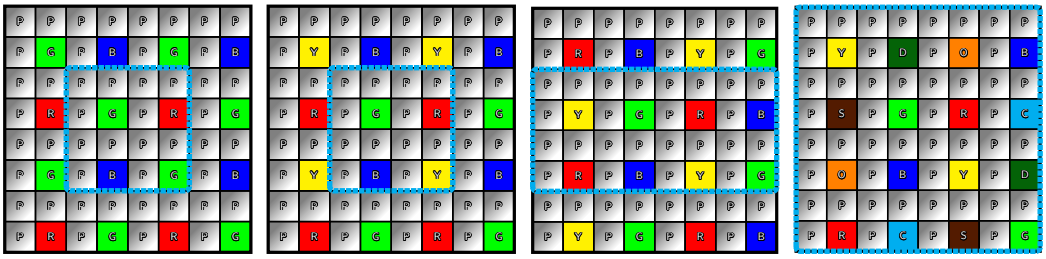
When Bayesian frameworks are employed, the results are labeled with `XXX+YYY+XXX` where `XXX` is the acquisition system, `YYY` is the regularization and `ZZZ` is the regularization norm (if not clear from the context). Whenever the blurring radius d_b is not specified, it is assumed that the operator $\mathbb{A}_b$ is set to identity. All iterative inversion algorithms were run for 250 iterations.

[2] For the San Francisco dataset, the employed spectral degradation weights for red green blue (RGB) and near infrared (NIR) were $[0.2113, 0.2903, 0.1934, 0.3048]$, respectively, instead of the default setting (0.25 to each), as QuickBird (QB) is more sensitive to a more accurate representation of the wide-band spectral response.

If the PAN and the MS samples are separable on the focal plane, their samples are histogram matched. Specifically the histogram matched version $\mathbf{m}_{:k}^{[HM]}$ of the MS samples from the k-th channel is obtained as:

$$m_{:k}^{\boxdot[HM]} = \frac{\mathrm{std}(\widetilde{\mathbf{p}}^{\boxdot\downarrow})}{\mathrm{std}(\widetilde{\mathbf{m}}_{:k})} \left(\mathbf{m}_{:k}^{\boxdot} - \overline{\widetilde{\mathbf{m}}}_{:k}^{\boxdot} \right) + \overline{\widetilde{\mathbf{p}}} \tag{4.6.1}$$

where $\mathbf{m}_{:k}^{\boxdot}$ is the k-th MS sparse channel, $\widetilde{\mathbf{m}}_{:k}^{\boxdot}$ is its weighted bilinear (WB) demosaiced version. Similarly $\mathbf{p}^{\boxdot}$ is the sparse channel of the PAN, with $\widetilde{\mathbf{p}}^{\boxdot}$ its WB demosaiced version and $\widetilde{\mathbf{p}}^{\boxdot\downarrow}$ its demosaiced and decimated version.



(a) Bayer (3 Bands) **(b)** PERI (4 Bands) **(c)** MAXDIS (4 bands) **(d)** UBT (8 bands)

Fig. 4.3. A series of PAN and MS mask patterns to be used for the experiments in Section 4.6. The masks are color coded to visualize the band they are assigned to: gray for the PAN, red, green, blue for the MS RGB, yellow for the its NIR. In the case of 8 bands, their brighter counterparts denote the remaining MS bands. The thick outline bounds the periodicity of each mask. The small captions just denote the MS pattern for simplicity.

4.6.2 Dataset description

Each reference dataset is composed of a PAN/MS image bundle acquired almost simultaneously, originally featuring a scale ratio of $1:4$, although the tests are performed at reduced resolution with a scale ratio of $\rho = 2$. They are acquired by a variety of commercial satellites, a selection of which is freely available for download by MAXAR Technologies [@9]. The characteristics of the obtained GT for each employed dataset are shown in Table 4.1; for two of the employed sensors, GeoEye-1 (GE-1) and QB, 4 MS channels are available, RGB and NIR, while WorldView-2 (WV2) and WorldView-3 (WV3) feature 8 bands in the visible (VIS) and NIR range. All the products are encoded with 11 bits; the spatial misalignment of the two PAN and MS was verified to be within half pixel via the rigid transformation alignment method described in [105].

When the acquisition is taken with the proposed MRCA system, unless otherwise noted, we assume that the HRI mask is the COVE pattern described in Section 4.3.3 (for each 2×2 region, 3 pixels are assigned to the PAN and the remaining one to the MS). The empty slots in the pan are filled with either a Bayer pattern in the case of 3 bands, or by a uniform binary tree (UBT) in the other cases [162]. The employed combined patterns are shown in Fig. 4.3.

As discussed in Section 4.6.6, these masks are not the most efficient for a Bayesian inversion. However, the reconstruction of the acquisitions with this simple design can be also approached with a variety of classic algorithms, to compare the performances.

The alignment between the PAN and an interpolated version of the MS is checked by evaluating the peak of the cross correlation in the Fourier domain [105]. The images are then co-registered with rigid translations by integer pixels on the HRI, in order to preserve the original intensity of the available samples.

Table 4.1. Characteristics of the GT of the datasets employed in the tests of Section 4.6.

Label	Country	Scene	Sensor	GSD [m]	Sizes
Beijing	China	Bird's Eye Nest	WV2	1.6	512×512
Hobart	Tasmania	Periphery Area	GE-1	2.0	512×512
Janeiro	Brazil	Bay Area	WV3	1.2	256×256
San Francisco	U.S.	Luxury Rental Area	QB	2.4	512×512
Stockholm	Sweden	Industrial Area	WV3	1.6	256×256
Washington	U.S.	Capitol Building	WV3	1.6	512×512

4.6.3 Compression

The objective of this section is to compare the quality of reconstructed product in a setting with fixed storage capabilities. To this end, the compression ratio ρ_c is set equal to eq.(4.2.6) for all the algorithms under test. For 4-channel MS images, this yields $\rho_c = 50\%$.

The analysis of this chapter is focused on an image quality comparison of the reconstructed products, generated from compressed images obtained with the two following approaches:

- **Software Compression**: In this category, both the PAN and MS are first compressed with a lossy software compression algorithm and these products

are then fused with a selection of classical image fusion protocols. Two compression approaches will be analyzed here:

  - **Radiometric binning (BIN)**: This approach is defined by a binning of the radiometric levels. That is, if the original image is represented with N_p bits, or equivalently over 2^{N_p} levels, a naive compression scheme consists in grouping together every $1/\rho_c$ level, so that the compressed image is instead represented over $\rho_c 2^{N_p}$ levels.

  - **Joint Photographic Experts Group (JPEG)**: This approach is based on the JPEG 2000 [@7] lossy protocol. To comply with the specifics of the algorithm, the original data is first reduced from 11 to 8 bits with the BIN method. As the JPEG operates over monochromatic or RGB bands, we just apply the algorithm on each band separately.

- **Hardware Compression**: in this category, the acquisition are simulated with the proposed multiresolution compressed acquisition model and the reconstructed products are tested with our proposed inversion scheme. In particular we investigate the following models of hardware compression:

  - **Multiresolution color filter array acquisition (MRCA)**: whose proposed design is described by eq. (4.2.4), with the combined mask setup of Fig. 4.3b.
  - **Compressive coded aperture spectral imaging (CASSI)**: whose adapted design is given by eq. (4.5.3).

  For those optical devices, the regularization is performed both with the TV-based inversion detailed in Algorithm 3 and with the protocol of eq. (4.4.8), which is simply denoted as LASSO. The latter is the original proposition for the inversion of the CASSI acquisitions [14, 70], but it is here employed for the inversion of both compressed acquisition systems.

Finally, we will provide a baseline for the potential results we could achieve if no compression step was considered, by simply fusing the unmodified PAN and MS with a selection of classical fusion techniques.

Table 4.2 shows the results of the reduced resolution quality assessment over three datasets, with the associated visual comparison is provided in Fig. 4.4, 4.5, and 4.6. If we limit the analysis to hardware compression techniques, the reconstruction with our proposed variational framework leads to fused product with superior quality compared to the literature. Additionally, the visual analysis of the CASSI reconstructed products (Fig. 4.4h to 4.4i, Fig. 4.5h to 4.5i, and Fig. 4.6h to 4.6i)

shows that the inversion is not able to fully recover the details, as the products remain quite blurry. Even in the case in which the LASSO regularization is employed on our custom masks, some texture pattern are not properly corrected, most likely because the contribution of sparsity inducing regularizers is limited within small areas.

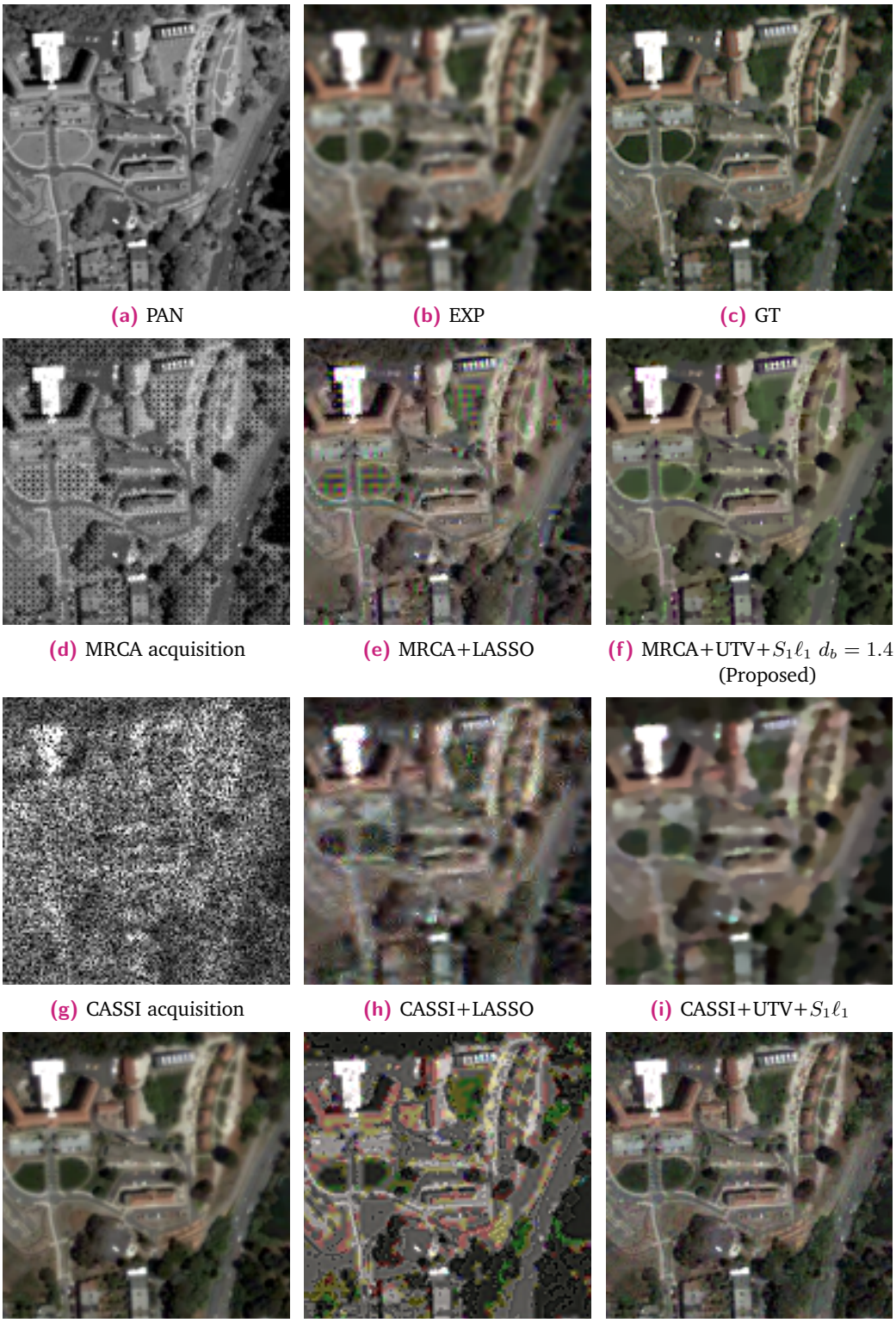
In general, hardware compression does not show a drastic reduction in performance compared to the baseline tests with uncompressed acquisitions, as long as an appropriate combination of on-board acquisition and regularization is selected. Some artifacts are although still present in our final product, especially in very thin thread-like zones such as the road limits in Fig. 4.5f. Nonetheless, except for some exceptions such as the blue field in Fig. 4.6f, the proposed algorithm is consistently better at smoothing homogeneous areas. This is the case even for small areas, like in the case of buildings, which show no visible texture artifact (Fig. 4.4f). On the other hand, software compression tends to outperform hardware compression in general, which is probably due to the fact that there is no information gap, especially in critical zones, such as along borders. Naive software compression may however be very limited in certain situations, such as in Fig. 4.4k, where many areas are patched together because of the radiometric binning.

It is also worth noting, however, that some more advanced strategies of software compression are available in the literature, such as those described by Consultative Committee for Space Data Systems (CCSDS) [240]. We did not investigate these more sophisticated approaches, as the currently employed approaches already outperform the hardware compression, as shown in Table 4.2. This result is not completely unexpected; in fact, our hardware compression solutions work on the basis that some samples are completely removed from our source pool, hence some information is completely lost. The software compression instead takes full advantage of the intrinsic redundancy of the image, from all the available samples, before generating the compressed product.

To provide better performances, an alternative setup could be to combine the hardware and software strategies, and achieve the compression advantages of both. I.e., one could think of a setup where the final acquisition $\mathbf{y}$ could be post-processed before being stored/transmitted. This step may anyway not be as straightforward, as the algorithm has to be adapted to the particular statistical description of the acquisition, which does not necessarily match that of a natural image.

Table 4.2. Compression comparison results for the tests described in Section 4.6.3 for the 4-band version of the Beijing, Hobart and San Francisco datasets. The dashed lines separate among the three considered test setups: No Compression, Software and Hardware Compression, respectively. For each of these classes, bold represents the best result. Only the two best performing classical fusion methods are shown.

		SSIM	PSNR	Q^2n	SAM	ERGAS	sCC
	Ideal (GT)	1	∞	1	0	0	1
Beijing	**EXP**	0.9099	25.24	0.7763	**4.394**	12.45	0.2883
	GSA	0.9655	28.46	0.9104	4.582	8.436	0.7395
	MTF-GLP-CBD	**0.9712**	**28.59**	**0.9116**	4.446	**8.291**	**0.7397**
	BIN+EXP	0.9030	25.19	0.7716	4.790	12.54	0.2776
	BIN+GSA	0.9555	28.19	0.8985	5.018	8.724	0.7142
	BIN+MTF-GLP-CBD	0.9628	28.36	0.9012	4.897	8.546	0.7152
	JPEG+EXP	0.9088	25.23	0.7753	**4.446**	12.47	0.2869
	JPEG+GSA	0.9644	28.42	0.9088	4.635	8.472	0.7363
	JPEG+MTF-GLP-CBD	**0.9702**	**28.56**	**0.9100**	4.502	**8.324**	**0.7366**
	CASSI+LASSO	0.8330	24.37	0.7273	13.68	6.861	0.2663
	CASSI+TV+ℓ_{221}	0.8537	24.69	0.7413	13.18	5.759	0.3222
	CASSI+UTV+$S_1\ell_1$	0.8744	24.95	0.7640	12.79	5.498	0.3603
	MRCA+LASSO	0.9029	25.71	0.8226	11.63	6.654	0.5597
	MRCA+TV+ℓ_{221}	0.9288	26.49	0.8508	10.64	5.493	0.6075
	MRCA+UTV+$S_1\ell_1$	**0.9445**	**27.02**	**0.8738**	**9.975**	**5.061**	**0.6475**
Hobart	**EXP**	0.9644	37.17	0.8835	2.987	6.393	0.5160
	MTF-GLP-CBD	0.9855	40.61	0.9489	3.053	3.915	0.7692
	BDSD	**0.9884**	**40.98**	**0.9605**	**2.867**	**3.743**	**0.7964**
	BIN+EXP	0.9489	36.41	0.8245	4.194	7.083	0.4030
	BIN+MTF-GLP-CBD	0.9658	37.85	0.8468	4.452	5.930	0.6200
	BIN+BDSD	0.9641	38.09	0.8653	4.656	5.729	0.6518
	JPEG+EXP	0.9611	37.01	0.8654	3.183	6.531	0.4914
	JPEG+MTF-GLP-CBD	0.9830	40.17	0.9310	3.257	4.196	0.7435
	JPEG+BDSD	**0.9853**	**40.46**	**0.9423**	**3.130**	**4.046**	**0.7699**
	CASSI+LASSO	0.9108	34.12	0.7421	8.813	5.516	0.2948
	CASSI+TV+ℓ_{221}	0.9325	35.05	0.7830	7.952	4.728	0.3973
	CASSI+UTV+$S_1\ell_1$	0.9396	35.49	0.8003	7.508	4.368	0.4452
	MRCA+LASSO	0.9465	35.96	7.059	0.8540	4.903	0.5410
	MRCA+TV+ℓ_{221}	0.9669	37.85	0.9062	5.526	3.768	0.6483
	MRCA+UTV+$S_1\ell_1$	**0.9737**	**38.35**	**0.9204**	**5.195**	**3.467**	**0.6633**
San Francisco	**EXP**	0.9904	45.91	0.9235	**2.352**	5.465	0.5487
	MTF-GLP-CBD	0.9964	48.90	0.9551	2.713	4.020	0.7534
	BDSD	**0.9967**	**49.88**	**0.9694**	2.414	**3.558**	**0.7835**
	BIN+EXP	0.9673	42.03	0.7983	5.930	8.655	0.2770
	BIN+MTF-GLP-CBD	0.9707	41.53	0.7961	6.470	9.294	0.4958
	BIN+BDSD	0.9661	41.92	0.8079	7.212	9.001	0.5135
	JPEG+EXP	0.9865	44.99	0.8949	**3.220**	6.094	0.4457
	JPEG+MTF-GLP-CBD	**0.9934**	47.24	0.9308	3.514	4.826	0.6948
	JPEG+BDSD	0.9931	**47.77**	**0.9429**	3.611	**4.562**	**0.7242**
	CASSI+LASSO	0.9806	43.48	0.8509	4.835	7.203	0.3152
	CASSI+TV+ℓ_{221}	0.9838	44.57	0.8837	4.084	6.375	0.4689
	CASSI+UTV+$S_1\ell_1$	0.9857	**45.10**	0.8941	**3.835**	**5.942**	0.5454
	MRCA+LASSO	0.9767	42.91	0.8340	5.558	7.985	0.5588
	MRCA+TV+ℓ_{221}	0.9857	44.47	0.8907	4.213	6.699	0.6283
	MRCA+UTV+$S_1\ell_1$	**0.9867**	44.70	**0.8969**	3.846	6.484	**0.6849**



(a) PAN (b) EXP (c) GT

(d) MRCA acquisition (e) MRCA+LASSO (f) MRCA+UTV+$S_1\ell_1$ $d_b = 1.4$ (Proposed)

(g) CASSI acquisition (h) CASSI+LASSO (i) CASSI+UTV+$S_1\ell_1$

(j) Uncompressed MTF-GLP-CBD (k) BIN+MTF-GLP-CBD (l) JPEG+MTF-GLP-CBD

Fig. 4.4. Visual comparison of the hardware and software compression for the 4-band San Francisco dataset (128×128 px cropped detail).

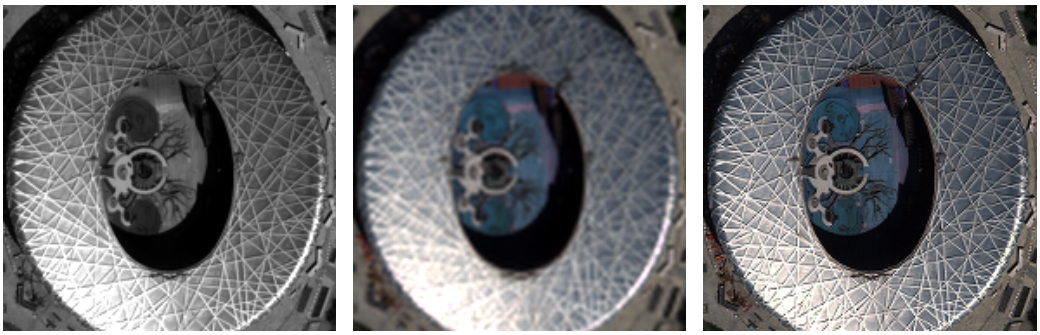
(a) PAN (b) EXP (c) GT

(d) MRCA Acquisition (e) MRCA+LASSO (f) MRCA+UTV+$S_1\ell_1$ $d_b = 1.3$ (Proposed)

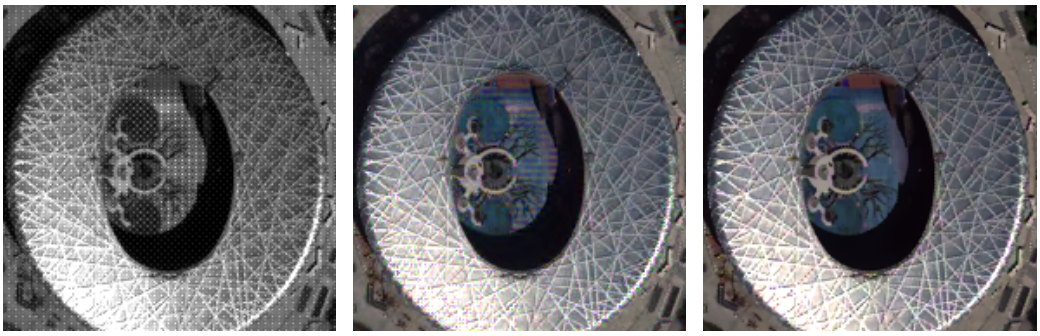
(g) CASSI Acquisition (h) CASSI+LASSO (i) CASSI+UTV+$S_1\ell_1$

(j) Uncompressed MTF-GLP-CBD (k) BIN+MTF-GLP-CBD (l) JPEG+MTF-GLP-CBD

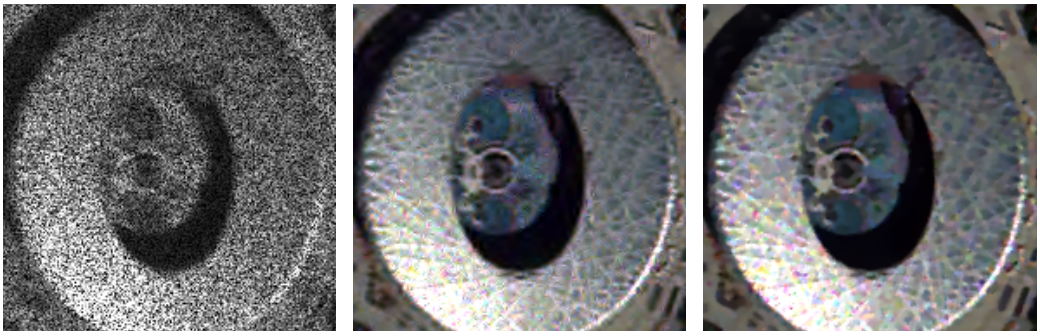
Fig. 4.5. Visual comparison of the hardware and software compression for the 4-band Hobart dataset (128×128 px cropped detail).



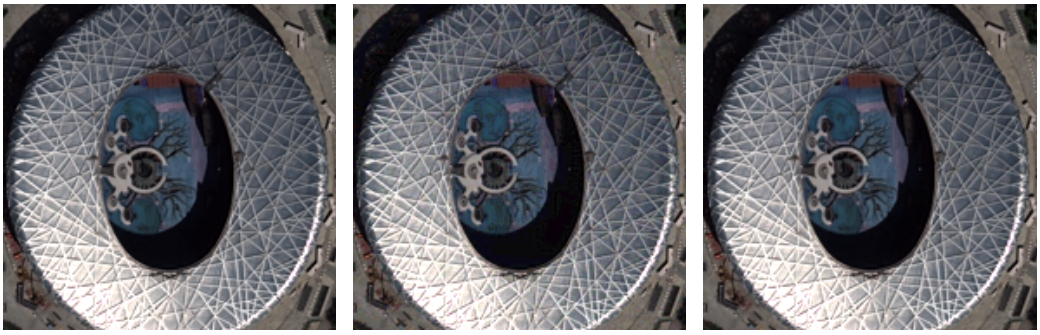
(a) PAN (b) EXP (c) GT



(d) MRCA acquisition (e) MRCA+LASSO (f) MRCA+UTV+$S_1\ell_1$ $d_b = 1.3$ (Proposed)



(g) CASSI Acquisition (h) CASSI+LASSO (i) CASSI+UTV+$S_1\ell_1$ $[d_b = 1.3]$



(j) Uncompressed MTF-GLP-CBD (k) BIN+MTF-GLP-CBD (l) JPEG+MTF-GLP-CBD

Fig. 4.6. Visual comparison of the hardware and software compression for the 4-band Beijing dataset (192×192 px cropped detail).

4.6.4 Separate and joint approach

The direct model proposed in Section 4.2 can be considered as a cascade of two systems, which we labeled as **Demosaic+Fusion**. As described in Section 4.4.2, this model can be specialized to a **Fusion Only** one, in charge of the fusion of the multiresolution acquisitions and a **Demosaic Only**, in charge of recovering the missing samples.

This section's aim is to evaluate the performance of the proposed inversion framework in each of these three situations. In particular, when our framework is applied to the "Fusion Only" and "Demosaic Only" problems, its performances can be compared to the fusion algorithms Section 3.5 or to the demosaic algorithms of Section 3.7, which were described in the analysis of the literature.

Additionally, the "Demosaic+Fusion" scenario can also be addressed, other than with our proposed inversion framework, by using a cascade of those same methods.

Specifically, if we can generate separate sparse channels $\mathbf{P}^{\boxdot}$ and $\mathbf{M}^{\boxdot}$, we can define the following procedure for the inversion of the MRCA acquisitions:

- The demosaiced PAN can be obtained with any of the irregular grid interpolation methods described in Section 3.4.1, among which we propose to employ a radial basis function (RBF) interpolator with a thin plate spline (TPS) kernel;
- The samples of the MS can be rearranged over a decimated grid and demosaiced with one of the techniques described in Section 3.7;
- The two demosaiced products can then be fused with one of the pansharpening techniques, described in Section 3.5.

To the best of my knowledge, this cascaded approach is a novel proposition, which can be used to investigate the effectiveness of noniterative approaches for the inversion of the MRCA.

The datasets "Washington" and "Janeiro" described in Section 4.6.2 are used for the experimental analysis. In particular, the GT is composed by cropping a section of $256 \times 256\text{px}$ from their RGB components.

The simulated MRCA acquisition for the "Demosaic+Fusion" setup is obtained with the mask shown in Fig. 4.3a, which is a combination of the COVE pattern for the PAN and a Bayer mask for the MS. The "Demosaic Only" setup is instead a obtained simulating a standard Bayer mask. As the reconstruction of Bayer mosaic has a very

mature literature (Section 3.7), the obtained benchmarks are expected to generally favor classical approaches.

Quality assessment

The objective quality assessment results are given in Table 4.3 for "Washington" and 4.4 for "Janeiro".

In these tables, we compared the best setup of our proposed inversion framework (MRCA+UTV+$S_1\ell_1$), with a variety of classic demosaic methods, such as WB, ID [163], ItID [164], ARI [166], MLRI [133], AP [149] and MSG [183]. We also compared those with a variety of classic fusion methods such as: GSA [6], BDSD [86], ATWT [178], MTF-GLP-HPM [7], MTF-GLP-CBD [7] and BayesNaive [232].

The three study cases all share the same GT, the indices relative to the three setups are directly comparable. Consequently, by focusing on the "Demosaic+Fusion" results, one can immediately assess the amount of distorted information in the recovered products due to the compression, if those are compared with the "Fusion Only" results, or due to the multiresolution sources, if compared with the "Demosaic Only" results.

The associated visual comparison is given in Fig. 4.7 for "Washington" and Fig. 4.8) for "Janeiro".

For the "Fusion Only" case, the recovered products with the proposed framework are barely indistinguishable from the generalized Laplacian pyramid (GLP) methods, which provide the best performances among the classical ones, as shown by comparing Fig. 4.8i and 4.8h.

For the "Demosaic Only" problem, the performances are mostly comparable to the best performing legacy methods although it does not typically outperforms the methods belonging to the family of residual interpolation (RI) [133]. The main drawback of our proposed solution is showcased in well confined thread-like region, such as in the outline of the Capitol Building in Fig. 4.7l or in regions with scattered single pixels objects, such as the pool in Fig. 4.8l.

In the "Demosaic+Fusion" framework, as one can see in Fig. 4.7f and 4.8f the color reconstruction of our proposed framework is sometimes not accurate (Fig. 4.7f and 4.8f). This is confirmed by the analysis of the spectral angle mapper (SAM) index, which typically is the best description of the accuracy of the spectral reconstruction.

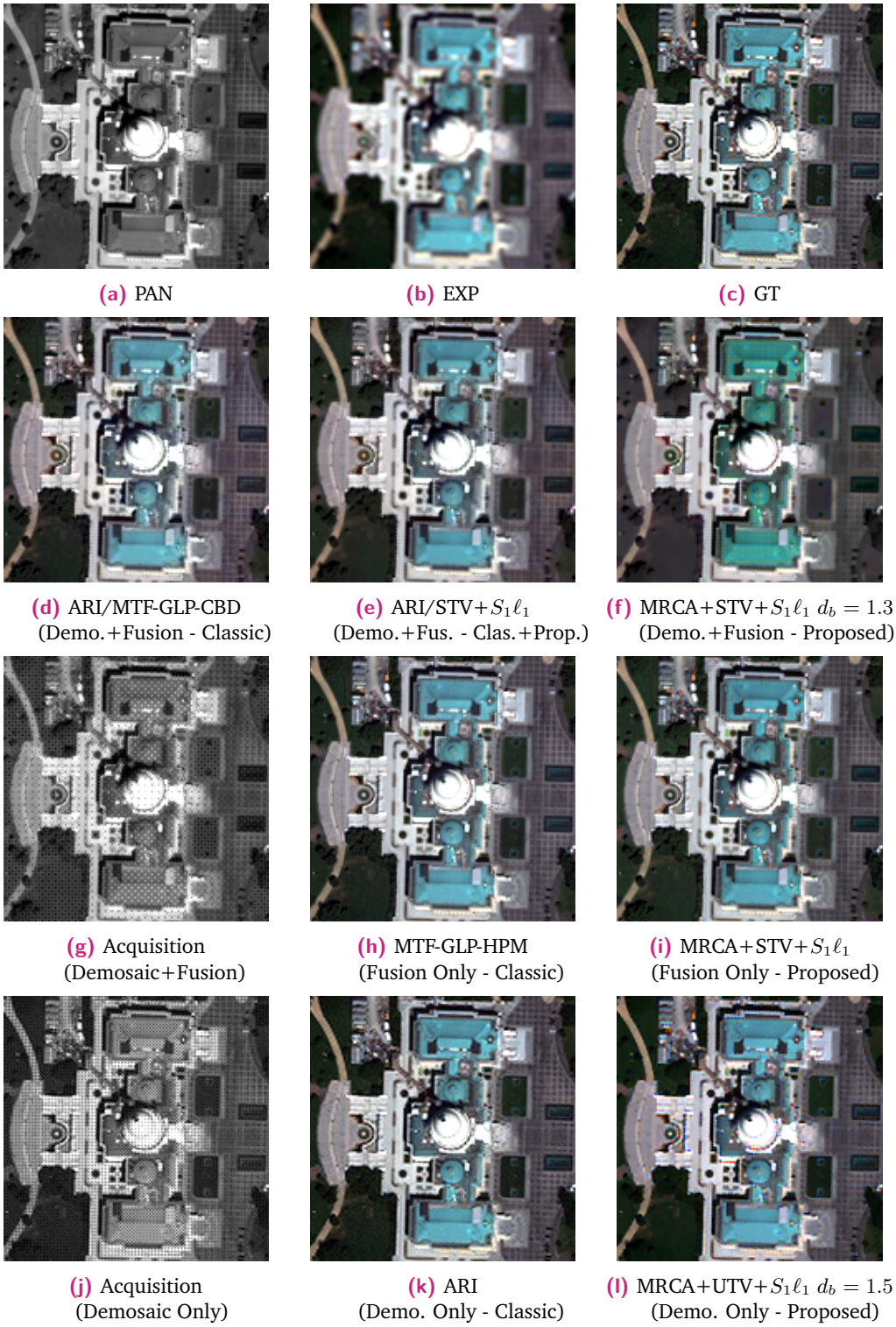
To address this issue, we investigated a mixed model, in which the demosaicing part is dealt with the ARI, while the fusion is performed with the proposed framework. With this choice, the visual results are close to indistinguishable with the overall best performing methods. E.g. the reader can compare Fig. 4.7e and 4.7f or Fig. 4.8e and 4.8f).

Table 4.3. Results for the test scenarios described in Section 4.6.4 for the 3-band Washington dataset with scale ratio $\rho = 2$. The PAN for the classic demosaic is interpolated via TPS RBF. The proposed method uses a (normalized) regularization parameter $\check{\lambda} = 1.93E-03$ and a blurring diameter $d_b = 1.3$.

	Washington (RGB)	SSIM	PSNR	Q^2n	ERGAS	SAM	sCC
	Ideal (GT)	1	∞	1	0	0	1
Demosaic+Fusion	MRCA+STV+$S_1\ell_1$	0.8934	30.75	0.8267	9.415	4.934	**0.7677**
	ARI/STV+$S_1\ell_1$	0.9311	34.28	0.9192	6.113	2.940	0.7640
	WB/EXP	0.7699	28.29	0.7720	12.22	3.929	0.2234
	ID/EXP	0.7923	28.81	0.8085	11.50	3.707	0.2560
	ItID/EXP	0.8298	29.67	0.8565	10.34	4.042	0.3669
	ARI/EXP	0.8480	30.48	0.8584	9.371	2.922	0.4423
	MLRI/EXP	0.8430	30.23	0.8573	9.647	3.025	0.4175
	AP/EXP	0.8316	29.67	0.8469	10.35	4.442	0.3675
	MSG/EXP	0.8407	30.02	0.8578	9.905	3.722	0.4016
	WB/GSA	0.8683	29.72	0.8065	10.35	5.085	0.7320
	MLRI/GSA	0.8934	31.21	0.8580	8.671	3.921	0.7510
	WB/BDSD	0.8088	28.96	0.8161	11.35	4.165	0.4685
	ARI/BDSD	0.8833	31.60	0.8922	8.256	2.976	0.6668
	WB/ATWT	0.9125	32.95	0.9084	7.224	3.867	0.7399
	ARI/ATWT	0.9293	34.54	0.9299	5.954	3.064	0.7617
	MLRI/ATWT	0.9283	34.44	**0.9308**	6.015	3.056	0.7584
	WB/MTF-GLP-HPM	0.9023	31.98	0.8897	8.086	3.791	0.7118
	ARI/MTF-GLP-HPM	0.9348	34.75	0.9298	5.814	3.062	0.7544
	MLRI/MTF-GLP-HPM	0.9330	34.48	0.9304	5.992	3.093	0.7493
	WB/MTF-GLP-CBD	0.8990	31.95	0.8883	8.119	3.838	0.7246
	ARI/MTF-GLP-CBD	**0.9351**	**34.97**	0.9288	**5.658**	**2.894**	0.7595
	WB/BayesNaive	0.8379	28.60	0.7890	11.66	7.338	0.6310
	ARI/BayesNaive	0.8982	30.29	0.8574	9.351	5.916	0.7437
	MSG/BayesNaive	0.8699	29.47	0.8281	10.64	7.268	0.6839
Demosaic Only	MRCA+UTV+$S_1\ell_1$	0.9603	36.57	0.9615	9.587	2.612	0.8386
	WB	0.9204	32.86	0.9221	14.80	3.505	0.6268
	ID	0.9464	34.58	0.9470	12.16	3.112	0.7374
	ItID	0.9546	36.15	0.9490	10.44	3.349	0.8529
	ARI	**0.9756**	**40.21**	**0.9727**	**6.311**	**2.181**	**0.9287**
	MLRI	0.9746	39.54	0.9715	6.814	2.248	0.9162
	AP	0.9458	35.29	0.9267	11.52	3.694	0.8241
	MSG	0.9578	37.15	0.9466	9.361	3.037	0.8762
Fusion Only	MRCA+STV+$S_1\ell_1$	0.9444	35.47	0.9372	5.285	**2.282**	0.8044
	EXP	0.8629	30.89	0.8767	8.923	2.435	0.4914
	GSA	0.9170	32.45	0.8866	7.471	3.209	0.7969
	BDSD	0.8986	32.14	0.9080	7.735	2.413	0.7256
	ATWT	0.9415	35.69	0.9443	5.186	2.505	**0.8091**
	MTF-GLP-HPM	0.9475	36.04	**0.9456**	4.991	2.580	0.8019
	MTF-GLP-CBD	**0.9484**	**36.40**	0.9450	**4.776**	2.350	0.8066
	BayesNaive	0.9159	30.77	0.8779	8.782	5.524	0.7972

Table 4.4. Results for the test scenarios described in Section 4.6.4 for the 3-band Washington dataset with scale ratio $\rho = 2$. The PAN for the classic demosaic is interpolated via Thin Plate RBF. The proposed method uses a (normalized) regularization parameter $\check{\lambda} = 1.93E-03$ and a blurring diameter $d_b = 1.3$.

	Janeiro (RGB)	SSIM	PSNR	Q^2n	ERGAS	SAM	sCC
	Ideal (GT)	1	∞	1	0	0	1
Demosaic+Fusion	MRCA+UTV+$S_1\ell_1$	0.8412	29.39	0.8281	12.53	3.941	0.5755
	ARI/UTV+$S_1\ell_1$	0.8830	31.44	0.8914	9.455	3.077	0.6209
	WB/EXP	0.7273	27.59	0.7190	15.03	4.089	0.2007
	ID/EXP	0.7483	28.01	0.7527	14.33	3.869	0.2320
	ItID/EXP	0.8041	29.07	0.8223	12.46	3.837	0.3642
	ARI/EXP	0.8174	29.48	0.8244	11.80	3.126	0.4215
	MLRI/EXP	0.8088	29.20	0.8177	12.22	3.233	0.3838
	AP/EXP	0.8080	29.00	0.8251	12.61	4.522	0.3647
	MSG/EXP	0.8129	29.23	0.8330	12.16	3.837	0.3851
	WB/GSA	0.8041	28.69	0.8037	12.98	5.002	0.5830
	ARI/GSA	0.8202	29.17	0.8288	12.36	4.278	0.5932
	WB/BDSD	0.7758	28.33	0.7753	13.81	4.476	0.4394
	ARI/BDSD	0.8527	30.33	0.8608	10.69	3.239	0.5980
	WB/ATWT	0.8557	30.45	0.8662	10.80	3.967	0.5878
	ARI/ATWT	0.8810	31.53	0.9015	9.436	3.251	0.6225
	WB/MTF-GLP-HPM	0.8430	29.83	0.8411	11.67	3.786	0.5702
	ARI/MTF-GLP-HPM	**0.8938**	**31.82**	**0.9053**	**9.122**	**2.957**	**0.6247**
	MLRI/MTF-GLP-HPM	0.8891	31.52	0.9013	9.442	3.032	0.6131
	WB/MTF-GLP-CBD	0.8377	29.79	0.8403	11.71	3.962	0.5696
	ARI/MTF-GLP-CBD	0.8883	31.75	0.9021	9.161	3.090	0.6190
	WB/BayesNaive	0.7999	28.14	0.7977	13.97	5.558	0.5479
	ARI/BayesNaive	0.8349	29.23	0.8270	12.23	3.802	0.5938
Demosaic Only	MRCA+UTV+$S_1\ell_1$	0.9407	34.59	0.9473	13.79	2.815	0.8098
	WB	0.8943	31.64	0.8977	19.56	3.757	0.6081
	ID	0.9234	33.08	0.9286	16.67	3.398	0.7055
	ItID	0.9566	36.04	0.9641	11.46	3.082	0.8830
	ARI	**0.9760**	**39.05**	**0.9810**	**8.561**	**2.105**	**0.9387**
	MLRI	0.9707	37.27	0.9747	10.03	2.382	0.9037
	AP	0.9447	34.41	0.9479	13.94	3.690	0.8545
	MSG	0.9523	35.60	0.9604	12.05	3.135	0.8792
Fusion Only	MRCA+UTV+$S_1\ell_1$	0.9153	**31.59**	**0.9378**	**7.944**	3.610	**0.6843**
	EXP	0.8391	29.99	0.8459	11.11	2.457	0.4759
	GSA	0.8540	30.05	0.8664	11.16	3.547	0.6497
	BDSD	0.8765	31.02	0.8830	9.844	2.567	0.6653
	ATWT	0.9024	32.53	0.9232	8.314	2.571	0.6817
	MTF-GLP-HPM	**0.9176**	32.85	0.9291	8.106	**2.250**	0.6734
	MTF-GLP-CBD	0.9110	32.86	0.9258	8.003	2.413	0.6762
	BayesNaive	0.8579	29.81	0.8529	11.41	2.966	0.6549



(a) PAN (b) EXP (c) GT

(d) ARI/MTF-GLP-CBD (Demo.+Fusion - Classic)

(e) ARI/STV+$S_1\ell_1$ (Demo.+Fus. - Clas.+Prop.)

(f) MRCA+STV+$S_1\ell_1$ $d_b = 1.3$ (Demo.+Fusion - Proposed)

(g) Acquisition (Demosaic+Fusion)

(h) MTF-GLP-HPM (Fusion Only - Classic)

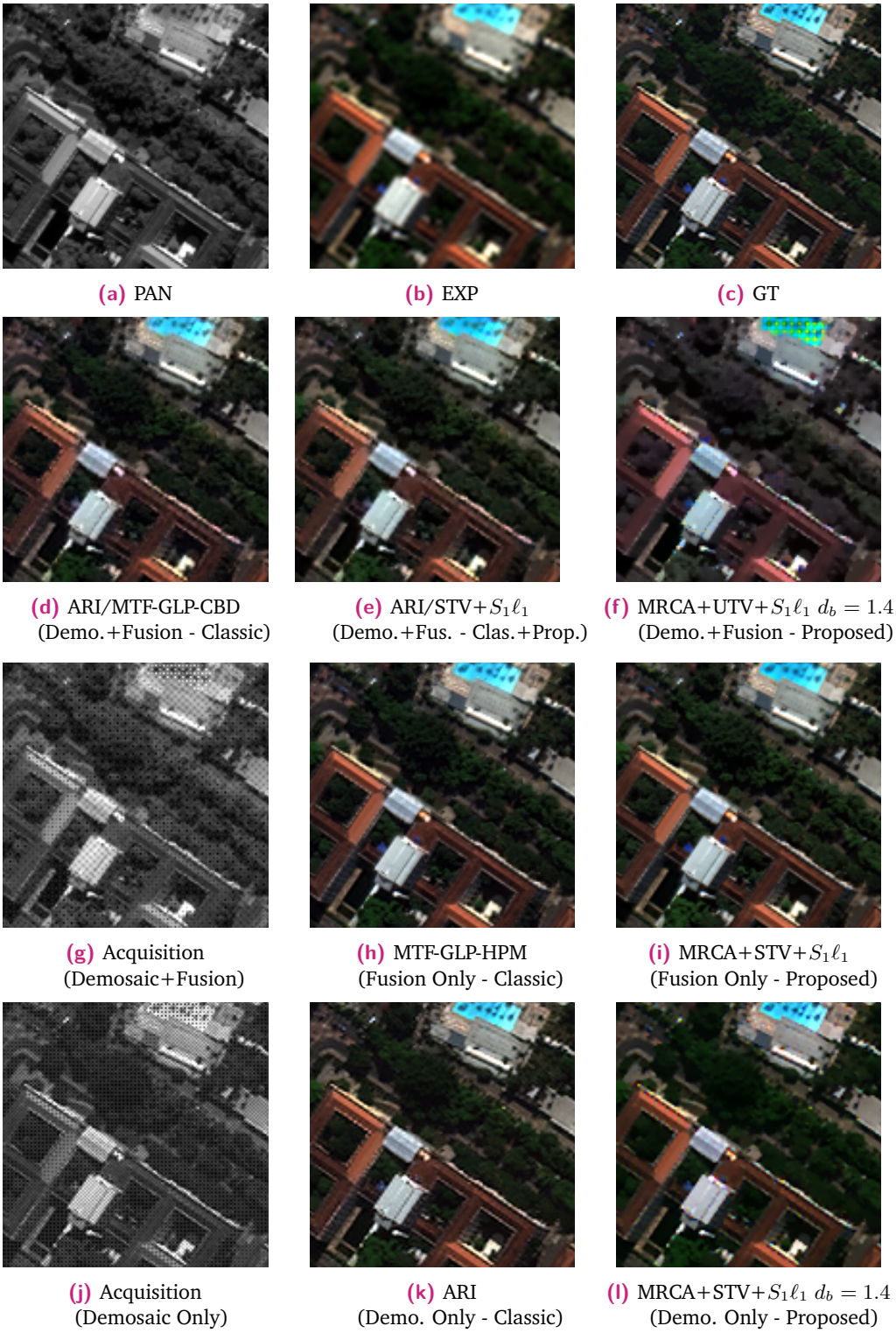
(i) MRCA+STV+$S_1\ell_1$ (Fusion Only - Proposed)

(j) Acquisition (Demosaic Only)

(k) ARI (Demo. Only - Classic)

(l) MRCA+UTV+$S_1\ell_1$ $d_b = 1.5$ (Demo. Only - Proposed)

Fig. 4.7. Comparison among joint and separate fusion and demosaicing algorithms for the RGB bundle of the Washington dataset (160×160 px cropped detail).The considered acquisition setup is indicated within parentheses. The employed mask for Demosaic+Fusion is shown in Fig. 4.3a.



(a) PAN (b) EXP (c) GT

(d) ARI/MTF-GLP-CBD (Demo.+Fusion - Classic)

(e) ARI/STV+$S_1\ell_1$ (Demo.+Fus. - Clas.+Prop.)

(f) MRCA+UTV+$S_1\ell_1$ $d_b = 1.4$ (Demo.+Fusion - Proposed)

(g) Acquisition (Demosaic+Fusion)

(h) MTF-GLP-HPM (Fusion Only - Classic)

(i) MRCA+STV+$S_1\ell_1$ (Fusion Only - Proposed)

(j) Acquisition (Demosaic Only)

(k) ARI (Demo. Only - Classic)

(l) MRCA+STV+$S_1\ell_1$ $d_b = 1.4$ (Demo. Only - Proposed)

Fig. 4.8. Comparison among joint and separate fusion and demosaic algorithms for the RGB bundle of the Janeiro dataset (128×128 px cropped detail). The considered acquisition setup is indicated within parentheses. The employed mask for Demosaic+Fusion is shown in Fig. 4.3a.

Simulated dataset

The performances of the proposed framework for RGB image bundles can be explained with any of the following three reasons:

- the degradation model of the PAN and the MS from the GT is not accurate;
- the amount of MS samples is insufficient to describe the color information in non-uniform regions;
- the regularization is not an accurate representation of the prior of certain structures of the image.

In this paragraph, we investigate the consideration on the degradation model, which is also backed up by a previous analysis from Duran et al. [64]. This work shows that the linear dependence between the PAN and spectral modalities of the GT may not be sufficient.

Additionally, as the spectral response of the PAN is wider than the combined one of the RGB components, 3 bands are probably not enough to properly characterize the structure of the PAN. This may also explain why fusion methods that use bands independently, such as the ones from the GLP class, and demosaicing methods which focus on modeling the dependencies locally, such as the RI classes, show better performances.

We initially tried to improve the model by adjusting the spectral degradation weights $\{w_{1k}\}_{k\in[1,\dots,N_b]}$, but only a slight improvement was obtained by decreasing the coefficient assigned to the blue band and increasing the one assigned to the green band.

To showcase the issue, we instead provide here an analysis with a perfect match between the simulation and reconstruction model. In this analysis, the reduced resolution PAN is not obtained through Wald's protocol, but instead as a linear combination of the bands of the GT; we additionally distorted the simulated acquisition with an additive white Gaussian noise (AWGN), such that the SNR is equal to 25dB.

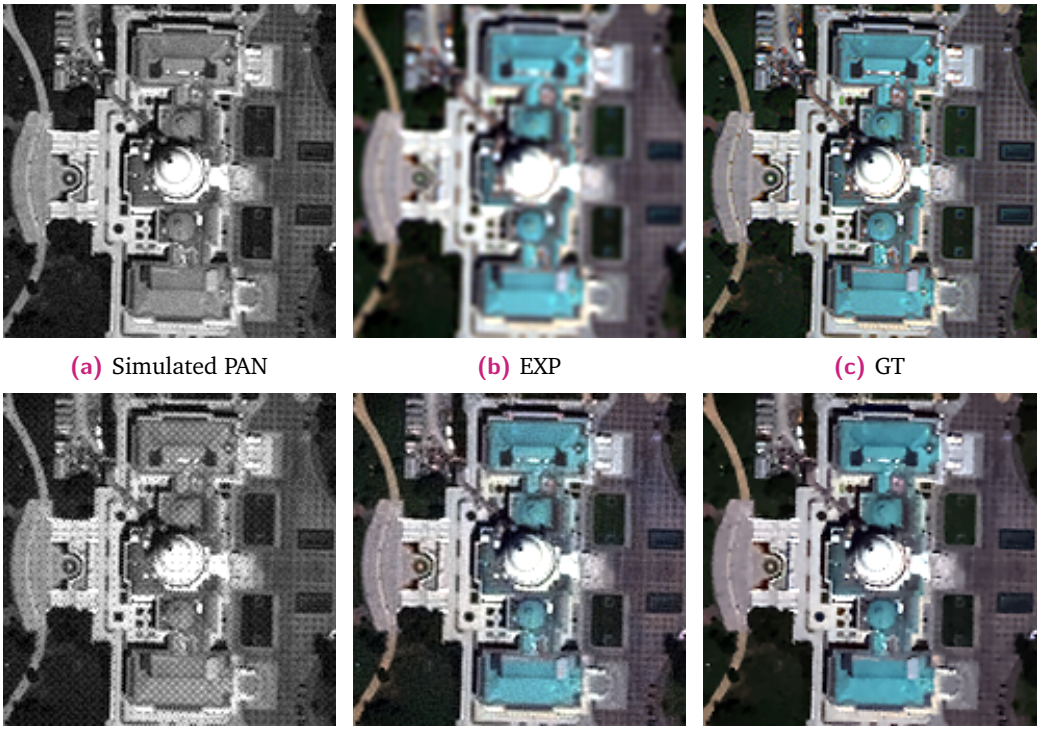
This analysis was performed for the "Washington" dataset, for which the quantitative and visual analysis are shown in Table 4.5 and products are shown in Fig. 4.9.

In this table, the "Demosaic Only" section is exactly the same as the one of Table 4.5, as the PAN is not involved. For our proposed framework in the "Demosaic+Fusion" problem, the obtained products not only recover their natural colors but also accurately reconstruct the homogeneous zones, such as the one of the roof in Fig. 4.9f.

This is also confirmed by indices, such as the structural similarity (SSIM) and $Q2^n$ index (Q^2n), which jointly evaluate the quality of spectral and spatial components.

Table 4.5. Results for the test scenarios described in Section 4.6.4 for the 3-band Washington dataset with simulated PAN with additive Gaussian noise such that SNR $= 25$dB and scale ratio $\rho = 2$. The PAN for the classic demosaic is interpolated via TPS RBF. The proposed method uses a (normalized) regularization parameter $\check{\lambda} = 1.8E-03$ and a blurring diameter $d_b = 1.5$.

	Sim. Wa.ton (RGB)	SSIM	PSNR	Q^2n	ERGAS	SAM	sCC
	Ideal (GT)	1	∞	1	0	0	1
Demosaic+Fusion	**MRCA+UTV+S$_1\ell_1$**	**0.9436**	35.07	**0.9336**	5.640	3.015	**0.8402**
	WB/EXP	0.7699	28.29	0.7720	12.22	3.929	0.2234
	ID/EXP	0.7923	28.81	0.8085	11.50	3.707	0.2560
	ItID/EXP	0.8298	29.67	0.8565	10.34	4.042	0.3669
	ARI/EXP	0.8480	30.48	0.8584	9.371	2.922	0.4423
	MLRI/EXP	0.8430	30.23	0.8573	9.647	3.025	0.4175
	AP/EXP	0.8316	29.67	0.8469	10.35	4.442	0.3675
	MSG/EXP	0.8407	30.02	0.8578	9.905	3.722	0.4016
	WB/GSA	0.9111	34.29	0.9126	6.175	3.932	0.8056
	ARI/GSA	0.9163	35.41	0.9198	5.359	2.998	0.8125
	WB/BDSD	0.8829	31.22	0.8864	8.880	4.247	0.7482
	ARI/BDSD	0.9327	35.82	0.9327	5.111	2.844	0.8171
	WB/ATWT	0.9170	33.85	0.9141	6.532	3.854	0.8003
	ARI/ATWT	0.9302	35.75	0.9317	5.171	2.918	0.8187
	WB/MTF-GLP-HPM	0.9059	32.62	0.8933	7.523	3.665	0.7815
	ID/MTF-GLP-HPM	0.9184	33.66	0.9125	6.665	3.394	0.7903
	ItID/MTF-GLP-HPM	0.9203	34.19	0.9166	6.347	3.760	0.8064
	ARI/MTF-GLP-HPM	0.9372	**36.22**	0.9315	**4.896**	**2.721**	**0.8211**
	MLRI/MTF-GLP-HPM	0.9359	35.89	0.9318	5.084	2.784	0.8167
	AP/MTF-GLP-HPM	0.9162	33.77	0.9041	6.683	4.153	0.8040
	MSG/MTF-GLP-HPM	0.9257	34.79	0.9187	5.909	3.443	0.8123
	WB/MTF-GLP-CBD	0.9018	32.71	0.8929	7.451	3.899	0.7849
	ARI/MTF-GLP-CBD	0.9299	36.01	0.9252	5.021	2.870	0.8137
	WB/BayesNaive	0.8684	31.17	0.8598	8.869	5.283	0.7276
	ARI/BayesNaive	0.9236	34.50	0.9115	5.899	3.292	0.8029
Demosaic Only	**MRCA+UTV+S$_1\ell_1$**	0.9603	36.57	0.9615	9.587	2.612	0.8386
	WB	0.9204	32.86	0.9221	14.80	3.505	0.6268
	ID	0.9464	34.58	0.9470	12.16	3.112	0.7374
	ItID	0.9546	36.15	0.9490	10.44	3.349	0.8529
	ARI	**0.9756**	**40.21**	**0.9727**	**6.311**	**2.181**	**0.9287**
	MLRI	0.9746	39.54	0.9715	6.814	2.248	0.9162
	AP	0.9458	35.29	0.9267	11.52	3.694	0.8241
	MSG	0.9578	37.15	0.9466	9.361	3.037	0.8762
Fusion Only	**MRCA+STV+S$_1\ell_1$**	**0.9639**	**38.91**	**0.9628**	**3.571**	**2.079**	**0.9121**
	EXP	0.8629	30.89	0.8767	8.923	2.435	0.4914
	GSA	0.9243	37.04	0.9297	4.434	2.537	0.8626
	BDSD	0.9449	38.04	0.9459	3.949	2.348	0.8684
	ATWT	0.9394	37.63	0.9425	4.146	2.394	0.8703
	MTF-GLP-HPM	0.9478	38.63	0.9443	3.691	2.207	0.8742
	MTF-GLP-CBD	0.9367	37.99	0.9351	3.976	2.378	0.8636
	BayesNaive	0.9371	36.10	0.9288	4.853	2.534	0.8574



(a) Simulated PAN (b) EXP (c) GT

(d) Acquisition (Demosaic+Fusion) Simulated PAN (e) ARI/MTF-GLP-HPM (Demosaic+Fusion) Simulated PAN - Classic (f) MRCA+UTV+$S_1\ell_1$ $d_b = 1.5$ (Demosaic+Fusion) Simulated PAN - Proposed

Fig. 4.9. Results for the image reconstruction in the Demosaic+Fusion experiment, in the case of a simulated PAN (160×160 px cropped detail). The first row shows the reference and uncompressed sources. The second row is associated with the "Demosaic+Fusion" scenario. The employed mask is shown in Fig. 4.3a.

4.6.5 Extension to multiple channels

In this section, we investigate the performances of the proposed framework, when addressing compressed acquisitions which involve an amount of channels superior to 3. There are multiple motivations for this analysis:

- To provide a preliminary analysis for the behaviour of the MRCA with hyperspectral (HS) images;
- To use a MS bundle, whose spectral coverage completely includes that of the PAN, as the previous analysis lacked information on the NIR bandwidths;
- To prove the flexibility of the algorithm to a wide variety of mask patterns.

The scenario under test employs the masks in Fig. 4.3b and Fig. 4.3d, that is, the MS sub-mask is a periodic pattern, for which some of the previously considered classic demosaic algorithms do not apply.

The analysis here was performed on two 4-band image bundles, "Washington" and "Janeiro", and two 8-band image bundles, "Janeiro" and "Stockholm". The objective quality assessment is given in Table 4.6, 4.7, 4.8 and 4.9, and the associated image products in Fig. 4.10, 4.11, 4.12, and 4.13, respectively.

When more channels are available, the multiresolution analysis (MRA) fusion methods, which operate on each bundle independently, start to show reduced performances in comparison to the proposed interconnected variational model. A quick analysis of reduced resolution quality indices for 4 bands shows that the proposed joint model consistently outperforms the classical methods in all scenarios under test. The results are confirmed by the visual analysis; by comparing Fig. 4.7f and 4.10f or Fig. 4.8f and 4.11f, it is evident that the proposed algorithm achieves a much more accurate color reconstruction. Some margin of improvement is however still possible, especially for localized pixels which interrupt homogeneous zones; i.e. the proposed algorithm struggles to reconstruct the bathers in the swimming pool of Fig. 4.11f.

This trend is further confirmed by analyzing the results at 8 bands, for which the proposed framework is additionally able to correct the spatial distortions that are present for the ¨Demosaic Only" scenario (as in Fig. 4.12i or 4.13i) with the help of the spatial one (as in Fig. 4.12f or 4.13f).

For the "Demosaic Only" setup, to the best of my knowledge, the proposed method reaches the state of the art performance for the case of masks with no dominant band.

Table 4.6. Results for the test scenarios described in Section 4.6.4 for the 4-band Washington dataset with scale ratio $\rho = 2$. The PAN for the classic demosaic is interpolated via TPS RBF. The proposed method uses a (normalized) regularization parameter $\check{\lambda} = 5.62E - 04$ and a blurring diameter $d_b = 1.3$

	Wash.ton (4-band)	SSIM	PSNR	Q^2n	ERGAS	SAM	sCC
	Ideal (GT)	1	∞	1	0	0	1
Demosaic+Fusion	**MRCA+UTV+S$_1\ell_1$**	**0.8886**	**29.44**	**0.9179**	**6.886**	**3.848**	0.6043
	WB/EXP	0.6987	25.73	0.7480	12.06	5.271	0.1740
	ID/EXP	0.7163	26.05	0.7664	11.38	5.279	0.2050
	ItID/EXP	0.6816	25.78	0.7069	11.84	6.279	0.1938
	SD/EXP	0.3926	12.31	0.2267	78.35	25.59	0.1551
	WB/GSA	0.8164	27.36	0.8763	8.091	6.040	0.5226
	ID/GSA	0.8173	27.49	0.8713	8.035	6.144	0.5330
	ItID/GSA	0.7763	26.74	0.7949	9.733	7.482	0.5315
	WB/BDSD	0.7815	26.79	0.8247	10.16	5.628	0.5562
	ID/BDSD	0.8138	27.46	0.8537	8.875	5.467	0.6057
	ItID/BDSD	0.8007	27.39	0.8068	9.170	6.328	0.6031
	WB/ATWT	0.8540	28.23	0.8844	7.858	5.406	0.6285
	ID/ATWT	0.8474	28.20	0.8727	7.744	5.437	0.6301
	ItID/ATWT	0.7853	27.13	0.7863	9.550	6.504	0.6146
	SD/ATWT	0.5011	12.34	0.2418	78.28	25.94	0.5994
	WB/MTF-GLP-HPM	0.8405	27.92	0.8634	8.526	5.370	0.6257
	ID/MTF-GLP-HPM	0.8426	28.13	0.8628	8.071	5.413	**0.6321**
	ItID/MTF-GLP-HPM	0.7877	27.33	0.7830	9.401	6.459	0.6157
	SD/MTF-GLP-HPM	0.4944	12.37	0.2439	77.99	25.59	0.5594
	WB/MTF-GLP-CBD	0.8033	27.03	0.8484	8.942	5.601	0.5113
	ID/MTF-GLP-CBD	0.8101	27.33	0.8530	8.389	5.685	0.5245
	ItID/MTF-GLP-CBD	0.7584	26.74	0.7780	9.637	6.738	0.5119
	SD/MTF-GLP-CBD	0.4099	12.32	0.2300	78.28	25.59	0.2828
	WB/BayesNaive	0.7714	25.95	0.8034	11.76	7.517	0.5794
	ID/BayesNaive	0.7839	26.23	0.8038	11.10	7.208	0.5972
	ItID/BayesNaive	0.7274	25.57	0.6895	12.71	8.501	0.5556
Demo. Only	**MRCA+STV+S$_1\ell_1$**	**0.9167**	**31.12**	**0.9407**	**12.09**	**3.195**	**0.6759**
	WB	0.8794	29.62	0.9046	15.37	3.886	0.5552
	ID	0.8964	30.56	0.9094	13.11	3.769	0.6481
	ItID	0.8239	29.32	0.8133	16.73	5.032	0.5965
	SD	0.4772	12.12	0.2386	161.3	25.65	0.4929
Fusion Only	**MRCA+UTV+S$_1\ell_1$**	**0.9488**	**33.93**	**0.9624**	**4.405**	**2.413**	**0.7849**
	EXP	0.8359	28.77	0.8792	8.508	3.302	0.4967
	GSA	0.8981	30.44	0.9354	5.789	4.192	0.6351
	BDSD	0.9165	31.39	0.9408	5.480	3.387	0.7295
	ATWT	0.9127	30.98	0.9399	5.543	3.641	0.7180
	MTF-GLP-HPM	0.9236	31.88	0.9430	5.182	3.570	0.7349
	MTF-GLP-CBD	0.9038	30.72	0.9359	5.393	3.973	0.6476
	BayesNaive	0.8775	28.51	0.8842	9.113	5.206	0.7159

Table 4.7. Results for the test scenarios described in Section 4.6.4 for the 4-band Janeiro dataset with scale ratio $\rho = 2$. The PAN for the classic demosaic is interpolated via TPS RBF. The proposed method uses a (normalized) regularization parameter $\check{\lambda} = 7.20E-04$ and a blurring diameter $d_b = 1.3$.

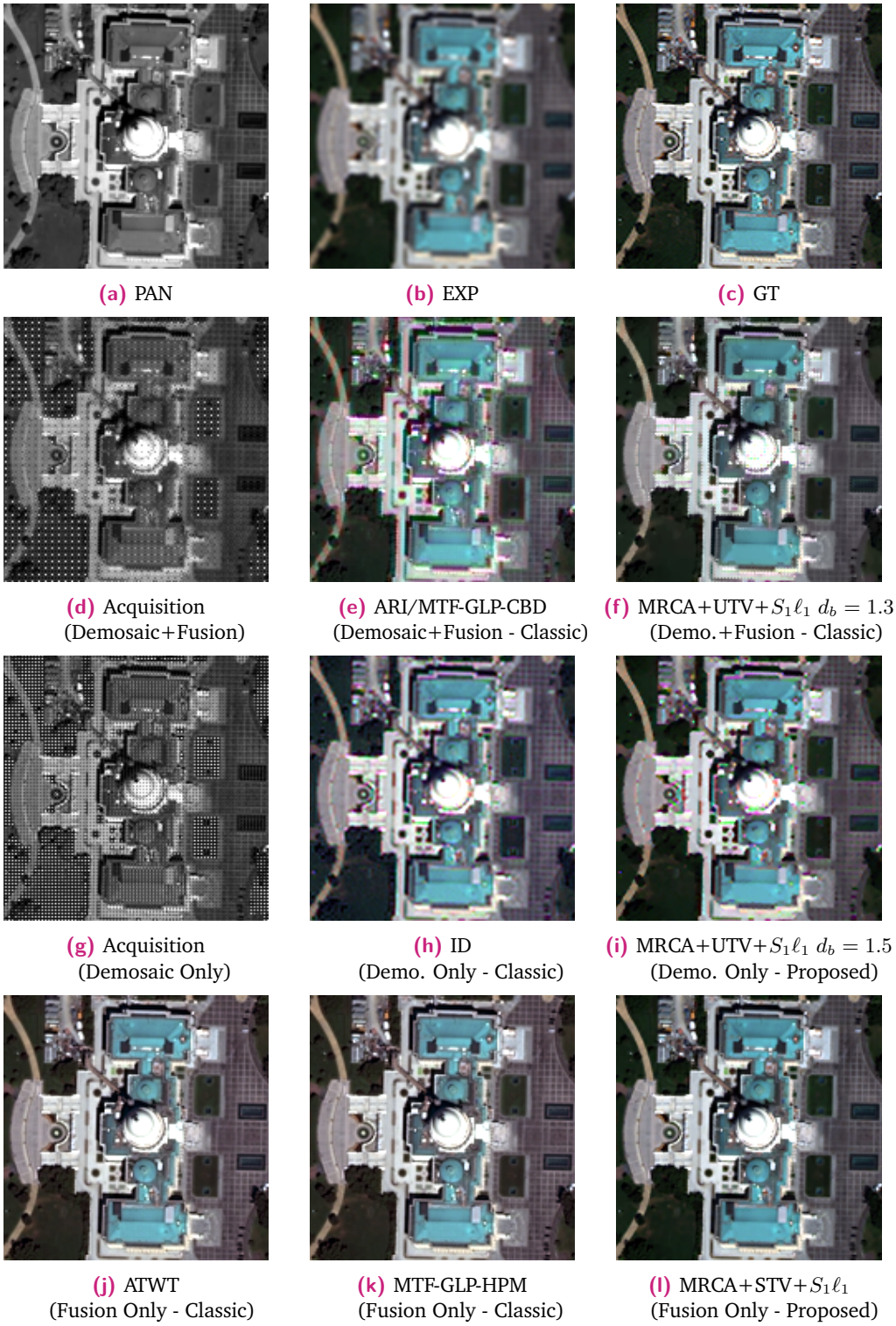
	Janeiro (4-band)	SSIM	PSNR	Q^2n	ERGAS	SAM	sCC
	Ideal (GT)	1	∞	1	0	0	1
Demosaic+Fusion	MRCA+UTV+$S_1\ell_1$	**0.8493**	28.67	**0.8823**	**10.95**	**5.559**	**0.5624**
	WB/EXP	0.6589	25.36	0.6985	16.32	7.546	0.1536
	ID/EXP	0.6823	25.76	0.7365	15.40	7.546	0.1978
	ItID/EXP	0.6513	25.66	0.7224	15.53	8.946	0.2116
	WB/GSA	0.8005	27.47	0.8469	12.40	7.877	0.5419
	ID/GSA	0.7999	27.56	0.8521	12.29	7.931	0.5453
	ItID/GSA	0.7559	26.85	0.8095	13.78	9.684	0.5426
	WB/BDSD	0.7517	26.51	0.7914	14.20	7.977	0.5073
	ID/BDSD	0.7879	27.27	0.8359	12.70	7.778	0.5481
	ItID/BDSD	0.7694	27.16	0.8180	12.92	9.138	0.5473
	WB/ATWT	0.8123	27.64	0.8495	12.25	7.459	0.5513
	ID/ATWT	0.8097	27.82	0.8567	11.89	7.490	0.5558
	ItID/ATWT	0.7472	26.98	0.8071	13.37	9.130	0.5449
	WB/MTF-GLP-HPM	0.7898	27.11	0.8181	13.17	7.288	0.5401
	ID/MTF-GLP-HPM	0.8005	27.52	0.8408	12.46	7.312	0.5330
	ItID/MTF-GLP-HPM	0.7486	27.03	0.8038	13.23	8.837	0.5404
	WB/MTF-GLP-CBD	0.7749	26.85	0.8128	13.41	7.634	0.5197
	ID/MTF-GLP-CBD	0.7876	27.29	0.8374	12.59	7.674	0.5283
	ItID/MTF-GLP-CBD	0.7378	26.82	0.8014	13.42	9.193	0.5205
	WB/BayesNaive	0.7891	27.04	0.8381	13.30	8.538	0.5271
	ID/BayesNaive	0.8043	27.42	0.8541	12.55	7.999	0.5394
	ItID/BayesNaive	0.7783	27.09	0.8323	12.98	9.101	0.5424
Demo. Only	MRCA+UTV+$S_1\ell_1$	0.8727	29.63	0.8962	19.27	**4.863**	0.6393
	WB	0.8546	28.88	0.8775	21.81	5.664	0.5108
	ID	**0.8845**	**30.13**	**0.9129**	**18.29**	5.429	0.6431
	ItID	0.8225	29.57	0.8713	20.12	6.928	**0.6643**
	SD	0.5938	18.77	0.4607	85.49	18.94	0.6357
Fusion Only	MRCA+UTV+$S_1\ell_1$	**0.9124**	**31.54**	**0.9357**	8.066	**3.701**	**0.6811**
	EXP	0.8180	28.34	0.8521	11.56	4.658	0.4867
	GSA	0.8695	29.56	0.9039	10.15	5.117	0.6267
	BDSD	0.8710	29.56	0.9028	10.03	4.948	0.6253
	ATWT	0.8915	30.43	0.9242	8.935	4.778	0.6544
	MTF-GLP-HPM	0.9051	30.92	0.9296	8.572	4.401	0.6636
	MTF-GLP-CBD	0.8946	30.60	0.9248	8.707	4.784	0.6501
	BayesNaive	0.8905	30.31	0.9113	9.347	4.293	0.6530

Table 4.8. Results for the test scenarios described in Section 4.6.4 for the 8-band Janeiro dataset with scale ratio $\rho = 2$. The PAN for the classic demosaic is interpolated via TPS RBF. The proposed method uses a (normalized) regularization parameter $\check{\lambda} = 9.21E - 04$ and a blurring diameter $d_b = 1.3$.

	Janeiro (8-band)	SSIM	PSNR	Q^2n	ERGAS	SAM	sCC
	Ideal (GT)	1	∞	1	0	0	1
Demosaic+Fusion	**MRCA+UTV+S$_1\ell_1$**	**0.8376**	**28.94**	**11.46**	**0.8582**	**8.110**	**0.5879**
	WB/EXP	0.6075	25.12	0.5973	18.58	11.03	0.06880
	ID/EXP	0.6070	25.22	0.6100	18.36	11.02	0.06054
	ItID/EXP	0.6134	25.62	0.6227	17.58	11.26	0.08454
	SD/EXP	0.4996	21.62	0.4384	28.57	20.40	0.1348
	WB/GSA	0.7931	27.52	0.8171	13.57	11.12	0.5739
	ID/GSA	0.7994	27.70	0.8247	13.27	11.13	0.5780
	ItID/GSA	0.7884	27.78	0.7806	13.38	11.33	0.5793
	SD/GSA	0.5737	20.28	0.4780	34.13	32.38	0.5688
	WB/BDSD	0.6570	25.40	0.6586	17.92	12.00	0.3883
	ID/BDSD	0.6599	25.53	0.6698	17.63	11.94	0.4089
	ItID/BDSD	0.6968	26.36	0.7053	16.04	11.97	0.4743
	WB/ATWT	0.7805	27.20	0.7880	14.34	10.82	0.5782
	ID/ATWT	0.7842	27.38	0.7961	14.02	10.81	0.5797
	ItID/ATWT	0.7612	27.52	0.7585	13.94	11.18	0.5654
	WB/MTF-GLP-HPM	0.7448	26.58	0.7383	15.55	10.73	0.5584
	ID/MTF-GLP-HPM	0.7482	26.75	0.7494	15.23	10.71	0.5519
	ItID/MTF-GLP-HPM	0.7402	27.15	0.7361	14.61	11.00	0.5449
	SD/MTF-GLP-HPM	0.5922	22.00	0.4968	27.42	20.40	0.4881
	WB/MTF-GLP-CBD	0.7249	26.32	0.7268	15.93	11.01	0.5399
	ID/MTF-GLP-CBD	0.7304	26.51	0.7416	15.54	11.00	0.5419
	ItID/MTF-GLP-CBD	0.7276	26.95	0.7311	14.85	11.30	0.5282
	SD/MTF-GLP-CBD	0.5780	21.96	0.4905	27.49	20.54	0.4704
Demo. Only	**MRCA+UTV+S$_1\ell_1$**	**0.8028**	**28.26**	**0.8099**	**25.78**	**7.877**	**0.5038**
	WB	0.7469	26.81	0.7570	30.73	10.07	0.2647
	ID	0.7496	26.89	0.7597	30.42	10.19	0.2430
	ItID	0.7555	27.62	0.7444	28.26	10.42	0.3279
	SD	0.5633	21.48	0.4577	59.74	21.61	0.3395
Fusion Only	**MRCA+UTV+S$_1\ell_1$**	**0.9336**	**33.34**	**0.9467**	**7.094**	**4.618**	**0.7208**
	EXP	0.8306	29.43	0.8495	11.34	5.785	0.4867
	GSA	0.9061	31.68	0.9260	8.552	5.922	0.6789
	BDSD	0.8919	30.70	0.9092	9.651	6.003	0.6658
	ATWT	0.9132	32.03	0.9323	8.127	5.690	0.6921
	MTF-GLP-HPM	0.9228	32.49	0.9362	7.781	5.411	0.7045
	MTF-GLP-CBD	0.9134	32.12	0.9312	8.041	5.705	0.6889

Table 4.9. Results for the test scenarios described in Section 4.6.4 for the 8 bands Stockholm dataset with scale ratio $\rho = 2$. The PAN for the classic demosaic is interpolated via TPS RBF. The proposed method uses a (normalized) regularization parameter $\breve{\lambda} = 9.21E-04$ and a blurring diameter $d_b = 1.3$.

	Stockholm (8-band)	SSIM	PSNR	Q^2n	ERGAS	SAM	sCC
	Ideal (GT)	1	∞	1	0	0	1
Demosaic+Fusion	**MRCA+UTV+$S_1\ell_1$**	**0.8359**	**27.55**	**0.8572**	**14.40**	**8.449**	0.7126
	WB/EXP	0.5406	22.87	0.5724	24.38	11.80	0.07145
	ID/EXP	0.5390	22.88	0.5875	24.36	11.81	0.04657
	ItID/EXP	0.5689	23.43	0.6440	22.86	11.70	0.09530
	SD/EXP	0.4747	20.53	0.5314	31.62	22.83	0.1834
	WB/GSA	0.8156	26.63	0.8385	16.04	11.39	0.7335
	ID/GSA	0.8262	26.86	0.8511	15.63	11.37	0.7361
	ItID/GSA	0.8246	27.20	0.8528	15.16	11.33	**0.7374**
	SD/GSA	0.5262	20.56	0.6853	33.03	47.00	0.7123
	WB/BDSD	0.6082	23.31	0.6464	23.20	12.99	0.5025
	ID/BDSD	0.6027	23.27	0.6542	23.34	13.08	0.4977
	ItID/BDSD	0.6549	24.27	0.7263	20.97	12.84	0.5580
	WB/ATWT	0.7726	25.70	0.7909	17.73	11.29	0.7223
	ID/ATWT	0.7788	25.84	0.8041	17.44	11.32	0.7238
	ItID/ATWT	0.7748	26.32	0.8187	16.58	11.38	0.7104
	SD/ATWT	0.6259	21.48	0.6444	28.44	24.71	0.6803
	WB/MTF-GLP-HPM	0.7196	24.73	0.7302	19.74	11.25	0.6802
	ID/MTF-GLP-HPM	0.7246	24.86	0.7462	19.46	11.24	0.6788
	ItID/MTF-GLP-HPM	0.7413	25.65	0.7859	17.81	11.17	0.6792
	WB/MTF-GLP-CBD	0.7123	24.72	0.7298	19.81	11.55	0.7008
	ItID/MTF-GLP-CBD	0.7394	25.65	0.7869	17.84	11.50	0.6915
Demo. Only	**MRCA+UTV+$S_1\ell_1$**	**0.7877**	**26.04**	**0.8092**	**33.82**	**8.186**	**0.5569**
	WB	0.7094	24.42	0.7492	40.67	11.10	0.2968
	ID	0.7121	24.42	0.7542	40.69	11.33	0.2494
	ItID	0.7521	25.63	0.7977	35.65	11.33	0.4079
	SD	0.5674	20.50	0.5784	64.20	24.91	0.4302
Fusion Only	**MRCA+UTV+$S_1\ell_1$**	**0.9393**	**32.09**	**0.9452**	**8.460**	**5.033**	0.8261
	EXP	0.8053	26.83	0.8349	15.45	6.281	0.5000
	GSA	0.9249	31.45	0.9397	9.242	6.230	0.8266
	BDSD	0.9129	30.26	0.9285	10.57	6.395	0.8082
	ATWT	0.9304	31.51	0.9430	9.175	5.894	0.8206
	MTF-GLP-HPM	0.9361	31.96	0.9415	8.689	5.585	**0.8361**
	MTF-GLP-CBD	0.9341	31.75	0.9395	8.951	5.800	0.8278



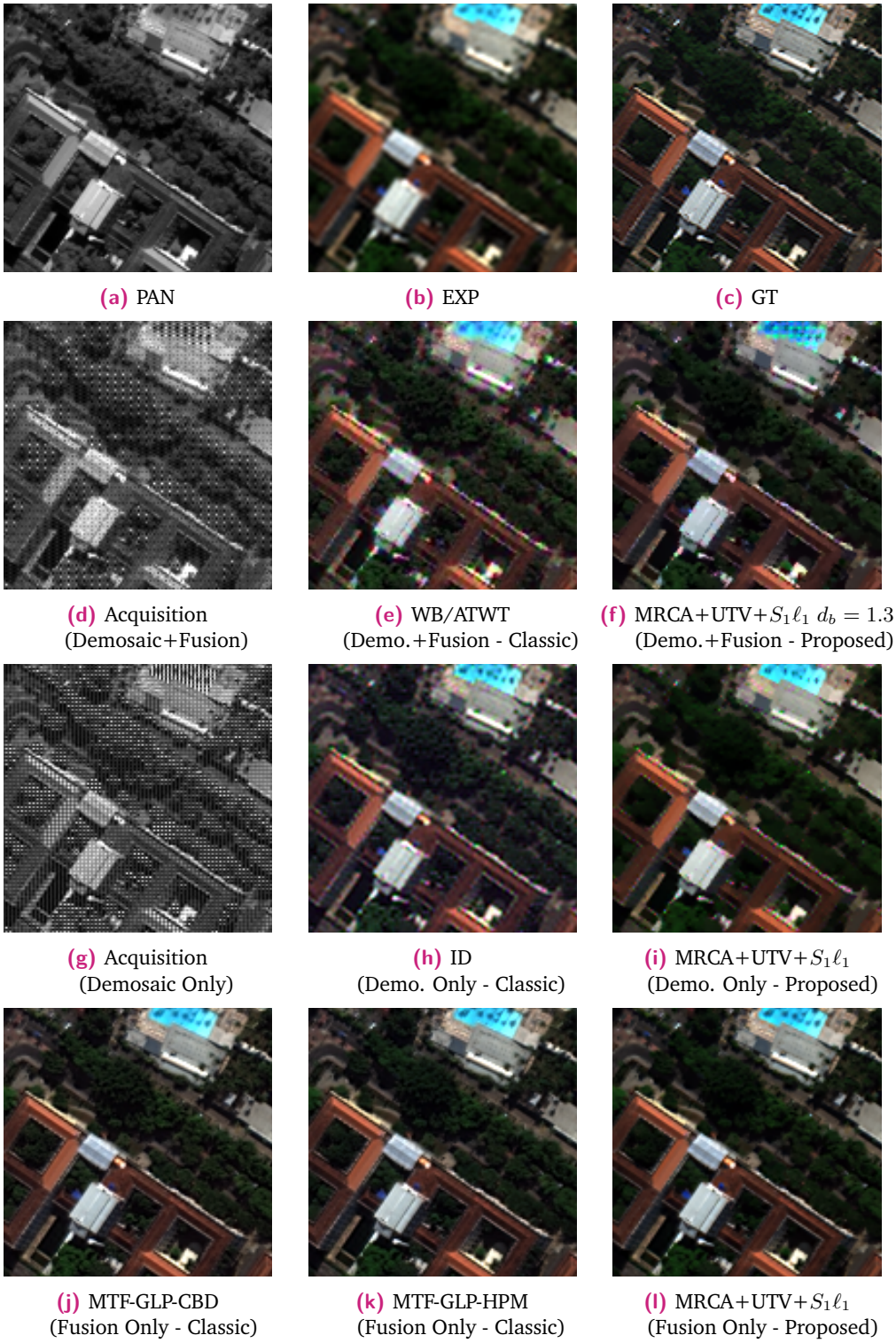
(a) PAN (b) EXP (c) GT

(d) Acquisition (Demosaic+Fusion) (e) ARI/MTF-GLP-CBD (Demosaic+Fusion - Classic) (f) MRCA+UTV+$S_1\ell_1$ $d_b = 1.3$ (Demo.+Fusion - Classic)

(g) Acquisition (Demosaic Only) (h) ID (Demo. Only - Classic) (i) MRCA+UTV+$S_1\ell_1$ $d_b = 1.5$ (Demo. Only - Proposed)

(j) ATWT (Fusion Only - Classic) (k) MTF-GLP-HPM (Fusion Only - Classic) (l) MRCA+STV+$S_1\ell_1$ (Fusion Only - Proposed)

Fig. 4.10. Comparison among joint and separate fusion and demosaic algorithms for the RGB+NIR bundle of the Washington dataset (160×160 px cropped detail). The considered acquisition setup is indicated within parentheses. The employed mask for Demosaic+Fusion is shown in Fig. 4.3b.



(a) PAN (b) EXP (c) GT

(d) Acquisition (Demosaic+Fusion) (e) WB/ATWT (Demo.+Fusion - Classic) (f) MRCA+UTV+$S_1\ell_1$ $d_b = 1.3$ (Demo.+Fusion - Proposed)

(g) Acquisition (Demosaic Only) (h) ID (Demo. Only - Classic) (i) MRCA+UTV+$S_1\ell_1$ (Demo. Only - Proposed)

(j) MTF-GLP-CBD (Fusion Only - Classic) (k) MTF-GLP-HPM (Fusion Only - Classic) (l) MRCA+UTV+$S_1\ell_1$ (Fusion Only - Proposed)

Fig. 4.11. Comparison among joint and separate fusion and demosaicing algorithms for the RGB/NIR bundle of the Janeiro dataset (128×128 px cropped detail). The considered acquisition setup is indicated within parentheses. The employed mask for Demosaic+Fusion is shown in Fig. 4.3b.

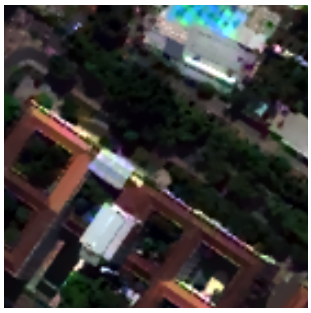
(a) PAN

(b) EXP

(c) GT

(d) Acquisition
(Demosaic+Fusion)

(e) ID/GSA
(Demosaic+Fusion - Classic)



(f) MRCA+UTV+$S_1\ell_1$ $d_b = 1.3$
(Demo.+Fusion - Proposed)

(g) Acquisition
(Demosaic Only)

(h) ItID
(Demo. Only - Classic)

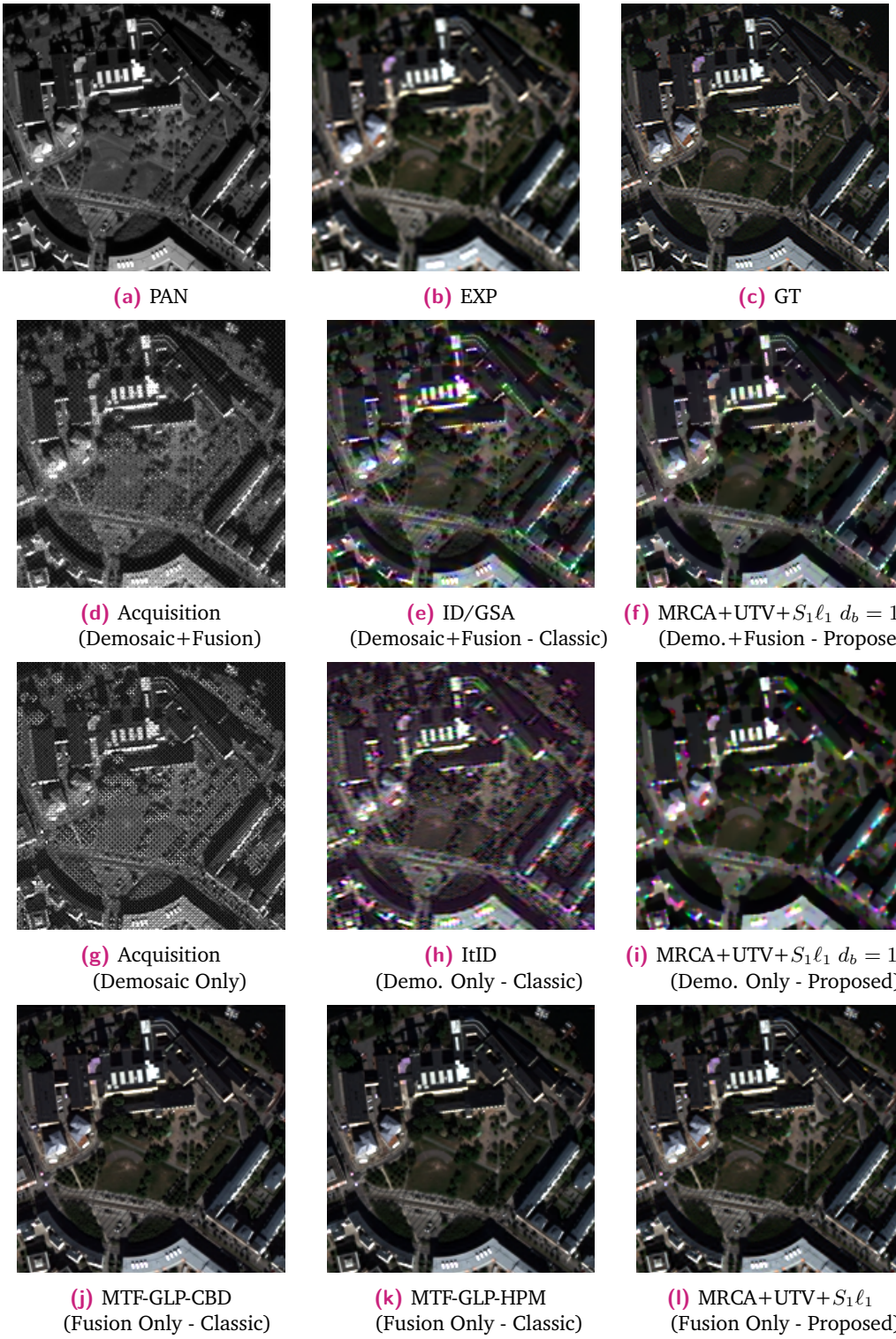
(i) MRCA+UTV+$S_1\ell_1$ $d_b = 1.4$
(Demo. Only - Proposed)

(j) MTF-GLP-CBD
(Fusion Only - Classic)

(k) MTF-GLP-HPM
(Fusion Only - Classic)

(l) MRCA+UTV+$S_1\ell_1$
(Fusion Only - Proposed)

Fig. 4.12. Comparison among joint and separate fusion and demosaic algorithms for the 8-band bundle of the Janeiro dataset (128×128 px cropped detail). The considered acquisition setup is indicated within parentheses. The employed mask for Demosaic+Fusion is shown in Fig. 4.3d.



(a) PAN (b) EXP (c) GT

(d) Acquisition (Demosaic+Fusion) (e) ID/GSA (Demosaic+Fusion - Classic) (f) MRCA+UTV+$S_1\ell_1$ $d_b = 1.3$ (Demo.+Fusion - Proposed)

(g) Acquisition (Demosaic Only) (h) ItID (Demo. Only - Classic) (i) MRCA+UTV+$S_1\ell_1$ $d_b = 1.4$ (Demo. Only - Proposed)

(j) MTF-GLP-CBD (Fusion Only - Classic) (k) MTF-GLP-HPM (Fusion Only - Classic) (l) MRCA+UTV+$S_1\ell_1$ (Fusion Only - Proposed)

Fig. 4.13. Comparison among joint and separate fusion and demosaic algorithms for the 8-band bundle of the Stockholm dataset (192×192 px cropped detail). The considered acquisition setup is indicated within parentheses. The employed mask for Demosaic+Fusion is shown in Fig. 4.3d.

4.6.6 Mask analysis

In this section we provide some preliminary tests on masks patterns based on the approaches described in Section 4.3. While the tests themselves are a quite exhaustive of a realistic set of patterns that could be implemented in practice, we still define them as preliminary in the sense that an accurate formalization of the problem requires a fully developed mathematical framework to strictly define the characteristics of the desired mask patterns.

Two mask design approaches are investigated:

- **Deterministic mask patterns**: which were described in Section 4.3.3. The test is performed with a variety of combination of PAN and MS masks. The former includes the COVE, VERT and DIAG patterns of Fig. 4.1c to 4.1b, while the latter includes the periodic (PERI) and MAXDIS patterns in Fig. 4.3b to 4.3c. The two masks are combined with the shifting procedure described in Section 4.3.3.
- **Random mask patterns**: which were described in Section 4.3.2. The investigated designs include the CASSI (Fig. 3.9), RAND and DIRI with a flat Dirichlet distribution.

The tests are performed on the 4-band bundles datasets "Hobart", "Janeiro" and "Washington", whose GT has sizes 512×512 px. The results of their objective analysis are given in Table 4.10, while the associated visual comparison is given in Fig. 4.14, 4.15, and 4.16.

The reconstruction with the proposed framework is performed with baseline settings, using a TV regularization embedded in an ℓ_{221}-norm, with the HRI blur operator $\mathbb{A}_b$ set to identity.

The reference GT channels are histogram matched to the PAN, before proceeding with the Wald's protocol, so that the dynamic range is consistent across all the samples of the acquisition. This choice was made for a fair comparison, as the procedure of eq. (4.6.1) is not applicable to any random mask, other than the RAND.

The main challenge for the reconstruction of acquisitions taken with a periodic mask is their texturing pattern, which can be corrected solely through the regularization. This effect is evident in products obtained with VERT pattern masked acquisitions, e.g. within the dome in Fig. 4.16d and 4.16g or along the road sides in Fig. 4.14g and 4.14g). The COVE masks, which were employed in the previous experiments

do not provide enough color information for the for small patches, such as in the case of the rooftops in Fig. 4.15h or the small habitations in Fig. 4.14h. The best performances are obtained with the combination DIAG+MAXDIS, as it has both the advantage of providing better localized channel information and PAN samples that are better distributed in the final acquisition.

It is worth noting that this last mask design, explicitly shown in Fig. 4.1f, is exactly equivalent to the design obtained by applying the dominant binary tree (DBT) [162], shown in Algorithm 1, by choosing the PAN as dominant band. If this dominant channel had the same spatial resolution of the MS channels, one could also think to adopt the demosaicing strategy proposed in [167], although adapting it to our case is not trivial.

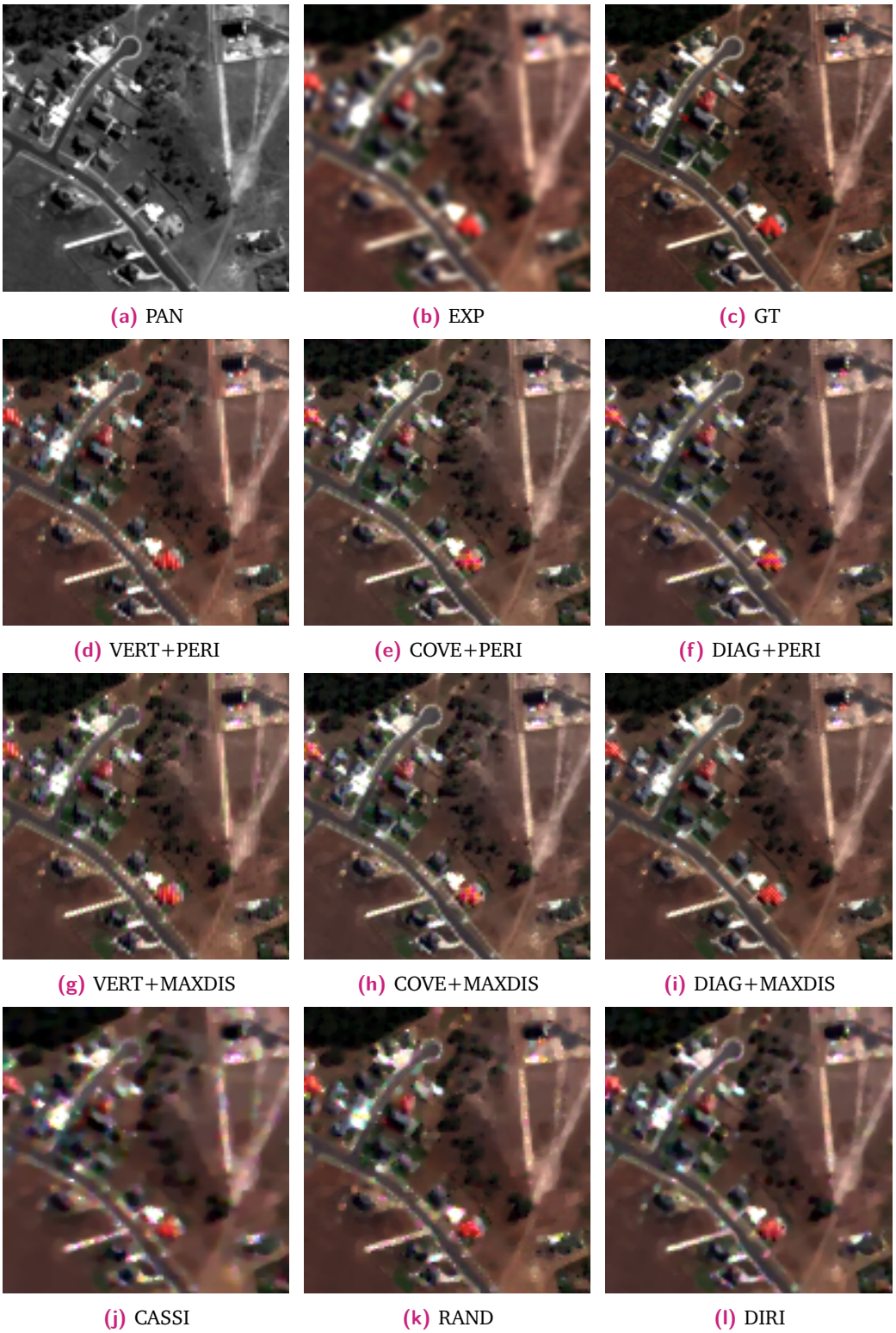
Random masks produce localized spectral distortions, which are more relevant in the CASSI-type masks, e.g. over the paved surfaces in Fig. 4.16j or across the roads in Fig. 4.14j. Among the tested randomized masks, the proposed DIRI pattern provides the best objective performances for the "Janeiro" and "Hobart" datasets, while the RAND design performs better in the "Washington" dataset. The first two datasets are in fact not characterized by which large uniform patches (Fig. 4.14l and 4.15l), where the reconstruction can be supported by the mixed pixel information of the DIRI, while the other case (Fig. 4.16k) does not require this level of detail, except in some spurious pixels.

To balance the accuracy of the spectral and spatial reconstruction, it is also possible to increase the occurrence ratio of the PAN samples in the mask pattern, i.e. by employing a generalized Dirichlet distribution in the DIRI design, or by employing non-uniform probabilities in the RAND design; in Table 4.10, we additionally test the case where half the samples in the pattern are from the PAN, but this led to no improvement.

Overall, the best periodic pattern (DIAG+MAXDIS) consistently outperforms random masks. Even if the DIRI design may be used to generate wider spectral responses, the same strategy can be employed even with deterministic masks, such as the cyan yellow magenta (CYM) arrangements consequently, relegating the employment of random masks to niche applications.

Table 4.10. Mask comparison results for the tests described in Section 4.6.6 for the 4-band 512×512 cut version of the Hobart, Janeiro (different cut) and Washington datasets. The dashed lines separate among the two considered classes: deterministic and random mask, respectively. Bold and underline represents the best and second best results for each dataset. The deterministic masks are combinations of masks shown in Fig. 4.3. All inversions are performed with the MRCA+TV+ℓ_{221} method, with the $\mathbb{A}_b$ operator set to identity. The percentage next to RAND denotes the amount of PAN samples. Only the best results for the (normalized) value $\check{\lambda}$ for the Q^2n parameter are given.

			SSIM	PSNR	Q^2n	ERGAS	SAM	sCC
		Ideal (GT)	1	∞	1	0	0	1
Hobart	**Deterministic**	**COVE+PERI**	0.9717	38.02	0.9204	5.401	3.619	0.6304
		DIAG+PERI	0.9788	38.57	0.9259	5.118	3.182	0.5496
		VERT+PERI	0.9734	38.10	0.9176	5.552	3.200	0.5545
		COVE+MAXDIS	0.9721	38.05	0.9207	5.353	3.610	**0.6309**
		DIAG+MAXDIS	**0.9847**	**39.35**	**0.9369**	**4.733**	**2.822**	0.5623
		VERT+MAXDIS	0.9734	38.06	0.9166	5.577	3.216	0.5484
	Aleatory	**CASSI**	0.9478	36.18	0.8585	6.924	4.313	0.4426
		RAND (20%)	0.9746	37.87	0.9076	5.789	3.070	0.4606
		RAND (50%)	0.9703	37.82	0.9091	5.630	3.478	0.5553
		DIRI	0.9703	38.07	0.9139	5.407	3.708	0.5914
Janeiro	**Deterministic**	**COVE+PERI**	0.9484	30.26	0.8606	9.143	4.634	0.6448
		DIAG+PERI	0.9594	30.11	0.8601	9.335	4.265	0.5630
		VERT+PERI	0.9494	30.09	0.8594	9.418	4.187	0.5894
		COVE+MAXDIS	0.9481	30.25	0.8601	9.162	4.643	**0.6449**
		DIAG+MAXDIS	**0.9653**	**30.42**	**0.8670**	**9.052**	**3.990**	0.5638
		VERT+MAXDIS	0.9489	30.00	0.8576	9.494	4.273	0.5779
	Aleatory	**CASSI**	0.9174	28.35	0.8162	11.47	5.209	0.4470
		RAND (20%)	0.9511	29.42	0.8507	10.18	4.249	0.4817
		RAND (50%)	0.9490	29.67	0.8480	9.865	4.571	0.5618
		DIRI	0.9594	30.32	0.8668	9.108	4.564	0.6031
Washington	**Deterministic**	**COVE+PERI**	0.9303	28.55	0.9071	7.790	5.154	**0.6145**
		DIAG+PERI	0.9506	29.49	0.9118	7.363	4.246	0.5856
		VERT+PERI	0.9448	28.88	0.9054	8.047	4.397	0.5726
		COVE+MAXDIS	0.9333	28.60	0.9070	7.779	5.176	0.6064
		DIAG+MAXDIS	**0.9623**	**29.99**	**0.9233**	**7.028**	**3.859**	0.5837
		VERT+MAXDIS	0.9440	28.78	0.9040	8.135	4.533	0.5595
	Aleatory	**CASSI**	0.8765	27.02	0.8278	10.28	5.465	0.4393
		RAND (20%)	0.9451	28.78	0.9010	8.330	4.104	0.4912
		RAND (50%)	0.9361	28.61	0.8991	8.285	4.661	0.5352
		DIRI	0.9306	28.83	0.8943	8.066	4.914	0.5637



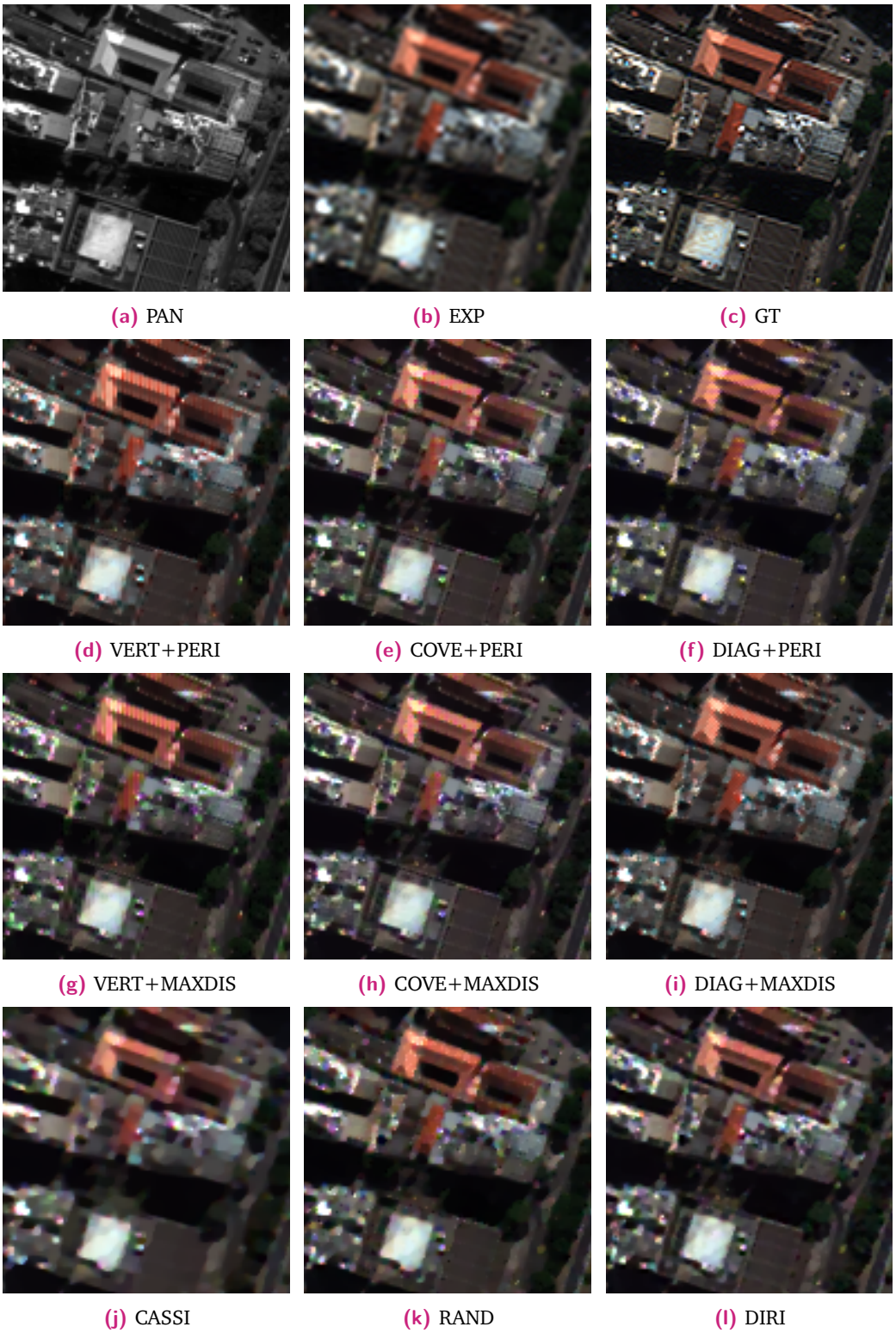
(a) PAN (b) EXP (c) GT

(d) VERT+PERI (e) COVE+PERI (f) DIAG+PERI

(g) VERT+MAXDIS (h) COVE+MAXDIS (i) DIAG+MAXDIS

(j) CASSI (k) RAND (l) DIRI

Fig. 4.14. Comparison of the reconstructed products with different mask designs for the 4-band Hobart dataset (128×128 px cropped detail). The combined MS/PAN masks are shown in Fig. 4.1. All inversions use the MRCA+TV+ℓ_{221} method.



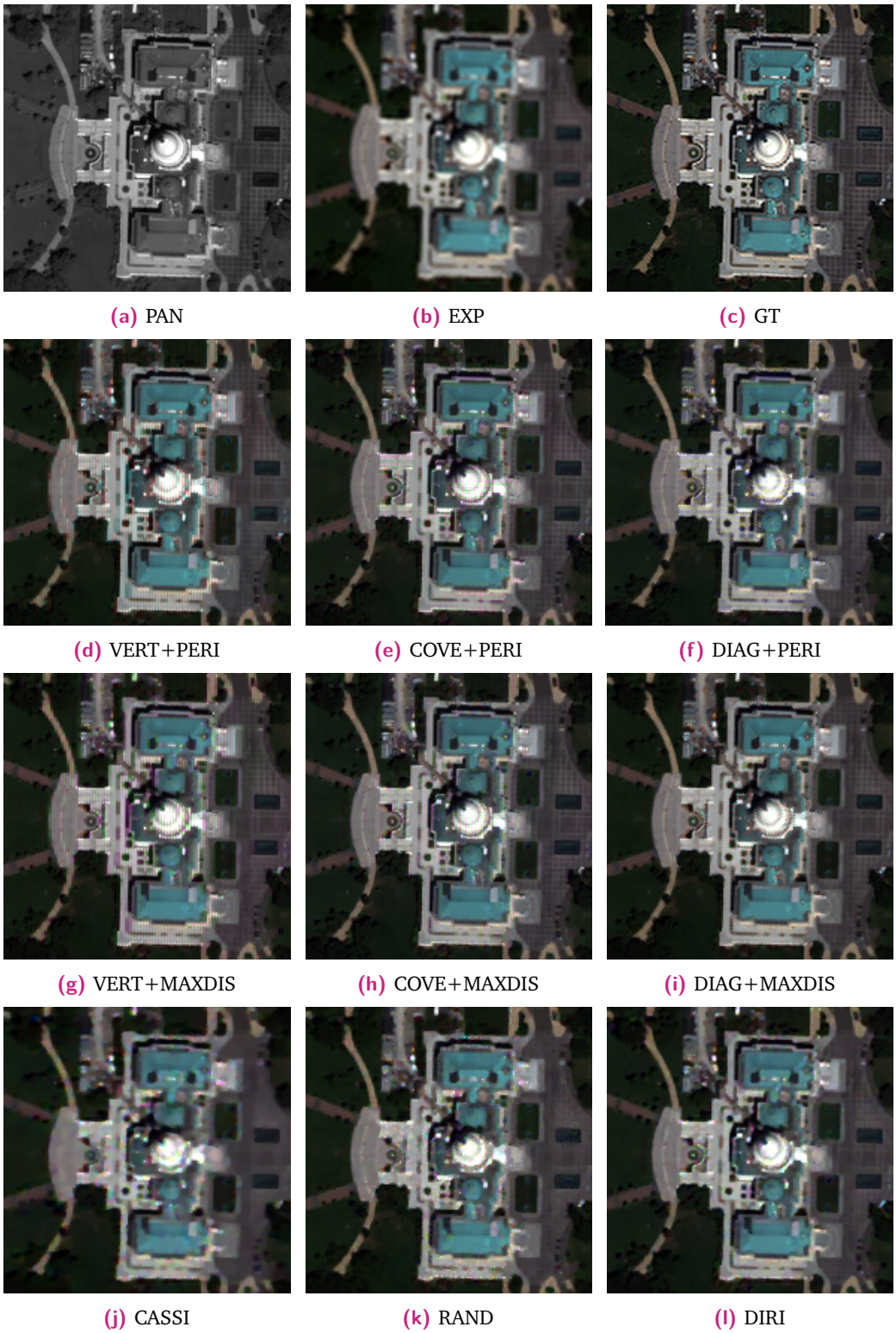
(a) PAN (b) EXP (c) GT

(d) VERT+PERI (e) COVE+PERI (f) DIAG+PERI

(g) VERT+MAXDIS (h) COVE+MAXDIS (i) DIAG+MAXDIS

(j) CASSI (k) RAND (l) DIRI

Fig. 4.15. Comparison of the reconstructed products with different mask designs for the 4-band Janeiro dataset (128×128 px cropped detail). The combined MS/PAN masks are shown in Fig. 4.1. All inversions use the MRCA+TV+ℓ_{221} method.



(a) PAN (b) EXP (c) GT

(d) VERT+PERI (e) COVE+PERI (f) DIAG+PERI

(g) VERT+MAXDIS (h) COVE+MAXDIS (i) DIAG+MAXDIS

(j) CASSI (k) RAND (l) DIRI

Fig. 4.16. Comparison of the reconstructed products with different mask designs for the 4-band Washington dataset (128×128 px cropped detail). The combined MS/PAN masks are shown in Fig. 4.1. All inversions use the MRCA+TV+ℓ_{221} method.

4.6.7 Setting the parameter values

We test here various possible parameters for the optimization of the proposed framework. We analyze here the acquisition system already employed for the "Demosaic+Fusion" setup in Section 4.6.4 and 4.6.5. The inversion algorithm MRCA+TV+ℓ_{221} with the blur operator $\mathbb{A}_b$ set to identity and a normalized regularization parameter $\check{\lambda} = 10^{-3}$, is set as baseline protocol.

We then perform a series of tests: in each of them, we allow the modification of a single parameter from that baseline setup, and evaluate the quality of the reconstructed product in terms of that parameter. We empirically experienced that setting up each parameter separately gives approximately the same result as if the parameters were set up jointly.

These tests are applied to the 4-band "Beijing" and "Washington" data and a summary of the results are given in Fig. 4.17 and 4.18, respectively. These results include the GT, the visual result of the baseline test and an indication of the best available setup of the proposed framework.

Specifically we perform the following tests:

- **Regularization Parameter** $\check{\lambda}$: The choice of the regularization parameter is the most important to ensure that the quality of the reconstructed image, as it provides a balance between the data fidelity term and the prior information (Section 2.1.2). Unfortunately, its choice is not trivial, as it depends on the SNR level of the acquisition, which in turn is a function of the characteristic of the sensors and the illumination n the scene. Some techniques for the automatic choice of the parameter, given the acquisition, such as generalized cross validation (GCV) [90], the ones based on Stein's unbiased risk estimate (SURE) [211] and the L-curve criterion [111, 113], will be explored in Chapter 6, as they are not easily adapted to the proposed framework. As a rule of thumb, we found that $\check{\lambda} = 10^{-3} \max \mathbf{y}$ is a good compromise in most scenarios; this is confirmed by the Q^2n results in Fig. 4.17d and 4.17d. In the visual comparison we provide the reconstructed products for implausible values of $\check{\lambda}$, to exaggerate its effects for the sake of presentation. If the regularization parameter is too low, there are relevant texture effects (Fig. 4.17e and 4.18e), as we impose no structure on the final image. If it is too high, the smoothing effect does not only apply on noisy regions, but to image features as well (Fig. 4.17f and 4.18f).

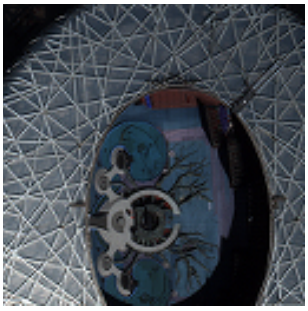
- **Blurring diameter** d_b: The blur operator $\mathbb{A}_b$ allows spreading the information of the PAN to adjacent samples at the cost of considering a degraded spatial resolution for the PAN. In our tests, the optimal value of the blur diameter was found to be in the range $1.3 - 1.5$ px for a scale ratio $\rho = 2$. For the "Beijing" dataset, the dependency is shown in Fig. 4.17g; a visual comparison of the reconstruction products for different values of d_b is shown in Fig. 4.17b, 4.17h and 4.17i. The optimal value is obtained as a trade-off between a more accurate recovery of the PAN sample and avoiding out-of-focus effects in the final product.

- **Regularizer norm**: In our proposed framework, the regularizer norm is an expression for the metric function $g(\cdot)$ which TV contributions in the domains of the spatial derivatives, channels and the pixels, according to the CTV [65, 66] formulation. The tested norms include the ℓ_{111}, ℓ_{211}, ℓ_{221}, $\ell_{\infty 11}$, $S_1\ell_1$ and $S_\infty\ell_1$ and the quantitative results are given in Fig. 4.17j for the "Beijing" dataset and in Fig. 4.18j for the "Washington" one. From this analysis, the ℓ_{221} norm (the baseline VTV) is the best compromise between quality of the reconstructed product and computational speed. Among the remaining choices, better performances are only achieved with the $S_1\ell_1$, which has the additional effect of whitening the noise contribution across different bands. For the visual analysis, one can immediately notice that some spectral spot-shaped spectral distortions in Fig. 4.17l when employing the ℓ_{111}. These distortions disappear in the other cases, such as in Fig. 4.17b) for the ℓ_{221} norm or, even better, in Fig. 4.17k for the $S_2\ell_1$ norm.

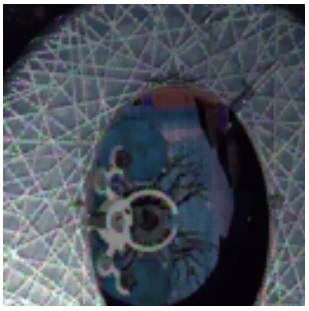
- **TV class**: For this test, we tested some different digital formulations for the discretization of the ROF model. In particular we tested the classic TV, the UTV [39] and the STV [1]; the STV was tested with an upscale by a factor of 2 and 3. For the "Washington" dataset, the qualitative comparison is given in Fig. 4.18g, which highlights that any of the considered choices shows no noticeable difference. The Q^2n index varies within a 0.002 difference margin. This is also confirmed by the visual inspection of Fig. 4.18i for the STV and Fig. 4.18h for the UTV, which feature no noticeable differences at the naked eye, in comparison to Fig. 4.18b.

To wrap the section, we would like to remind the reader that the accuracy of the direct model transfer function is of the utmost importance for an accurate inversion. The analysis of the accuracy of the direct model is outside of the scope of this work, but some techniques are available for the estimation of its parameters [225], in case of necessity. The histogram matching step of eq. (4.6.1), which was not studied

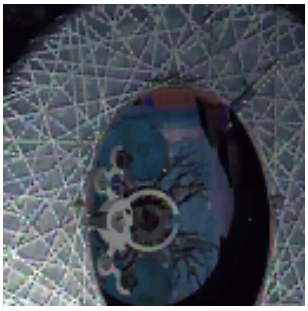
in detail in this context, allow for non-negligible increases in performance; this further confirms the requirement of an accurate description of the model for the multiresolution sensors.



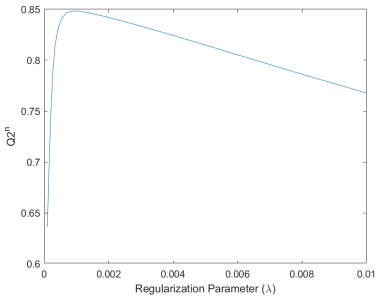
(a) GT



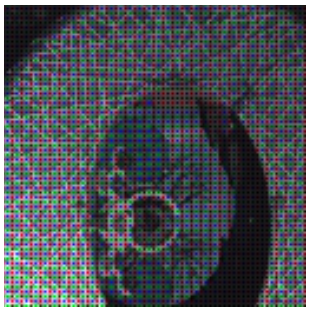
(b) MRCA+TV+ℓ_{221} $[\check{\lambda} = 10^{-3}]$ Baseline result



(c) MRCA+TV+$S_1\ell_1$ $[d_b = 1.5, \check{\lambda} = 9 \times 10^{-4}]$ Best result

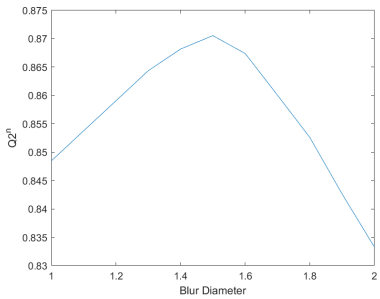


(d) Regular. Parameter Dependency

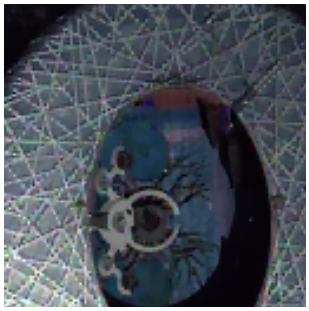


(e) Baseline $[\check{\lambda} = 10^{-8}]$

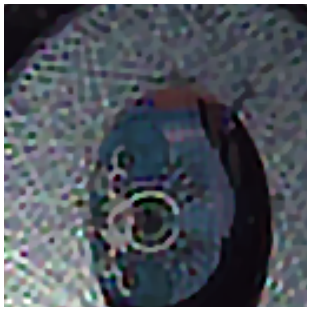
(f) Baseline $[\check{\lambda} = 10^{-1}]$



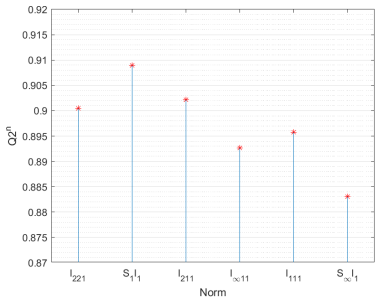
(g) Blur Diameter Dependency



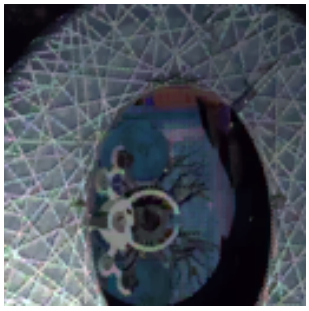
(h) Baseline $[d_b = 1.5]$



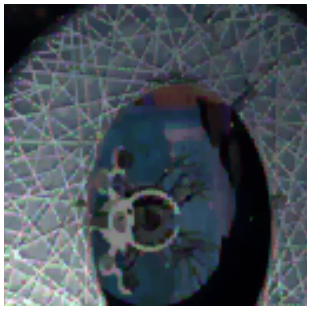
(i) Baseline $[d_b = 3]$



(j) Norm Dependency



(k) Baseline + Norm $S_1\ell_1$



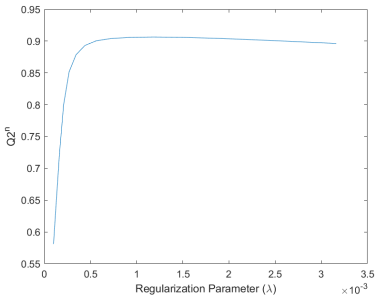
(l) Baseline + Norm ℓ_{111}

Fig. 4.17. Parameters settings for the 4-band Beijing dataset (128×128 px cropped detail). The first row shows the reference with the baseline and best performing methods. The following rows show the Q^2n and the reconstruction product by varying one parameter from the baseline setup.

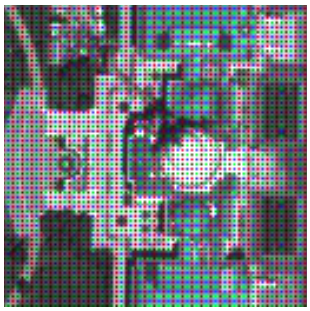
(a) GT

(b) MRCA+TV+ℓ_{221} [$\check{\lambda} = 10^{-3}$] Baseline result

(c) MRCA+UTV+$S_1\ell_1$ [$d_b = 1.3, \check{\lambda} = 6 \times 10^{-4}$] Best result

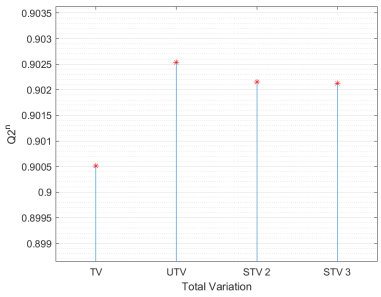


(d) Regular. parameter dependency

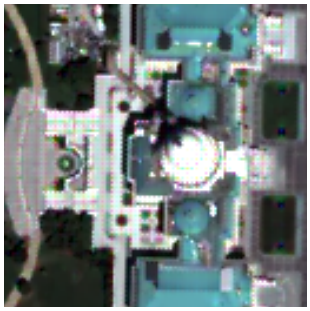


(e) Baseline [$\check{\lambda} = 10^{-8}$]

(f) Baseline [$\check{\lambda} = 10^{-1}$]

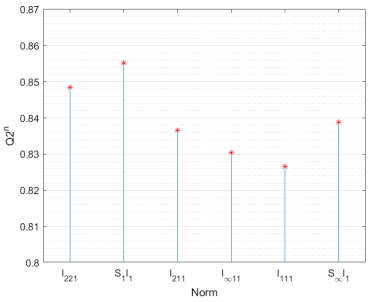


(g) TV Class Dependency

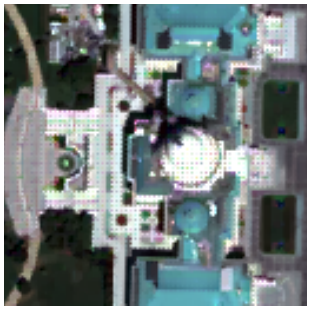


(h) Baseline+UTV

(i) Baseline+STV



(j) Norm Dependency



(k) Baseline + Norm $S_1\ell_1$

(l) Baseline + Norm ℓ_{111}

Fig. 4.18. Parameters settings for the 4-band Washington dataset (128×128 px cropped detail). The first row shows the reference with the baseline and best performing methods. The following rows show the Q^2n and the reconstruction product by varying one parameter from the baseline setup.

4.7 Conclusions and future perspectives

4.7.1 Conclusions

In this chapter, we proposed a model to jointly address the problem of image fusion and reconstruction of compressed data; a particular focus is given on ease of implementation with optical components on board of low-budget satellites, which could justify mass production of a constellation of the latter. The relaxed constraints on down-link resources is paid at a cost of a more complex and demanding processing on the ground segment.

We proposed the design of the MRCA, a multiresolution compressed acquisition system based on the CFA, both in its standard configuration and adapted to CASSI acquisitions, with a still theoretical, but feasible, optical implementation. We proposed a very flexible framework for the inversion of its acquisitions, cross checking results with a variety of widespread regularizers and good performances were achieved with an inversion based on the collaborative total variation. The promising results may justify mass-production of a constellation of very low-budget satellites, where the software reconstruction of the fused image is performed at the ground segment.

Additionally, we analyzed the effect of different masks for our proposed joint image fusion and reconstruction scheme. The results show that an ideal computational imaging-based design of the prototype should carefully balance the amount of provided information from the high and low resolution source, and avoid strongly localized patches with information pertaining to the same channel. The use of nonbinary random masks, which may allow the manufacture of a wide array of sensors with aleatory wider spectral responses with potentialities in overcoming the physical limitations of current MS platforms, is still quite limited compared to the case in which the design of deterministic masks is properly thought out.

When the spectral response of the PAN can be properly described by a combination of the spectral responses of the MS, our proposed framework reaches the state-of-the-art results. This is the case of 4 and 8-channel VIS/NIR bundles. In the case of RGB compressed acquisition, we propose to approach the inversion with a combination of classic demosaic and fusion methods, such as those based on the RI and GLP, respectively.

4.7.2 Future perspectives

Many different possible alternatives to extend this work are available. Among the possibilities, one interesting path is to expand the proposed framework to even more straightforward designs for commercial cameras. One of such examples could be inspired by the recent success of technologies such as Quad Bayer masks, which allow to combine together the energy acquired by multiple sensors to raise the SNR in condition of low illumination. In our case, our framework can formalize the setup of a single-chip acquisition system composed of wide-band pixels and 2×2 sized patches of MS sensors which can be combined together. Additionally, more advanced approaches for TV regularization could be employed, such the Elastica minimization [41] based of the work in [158] for inpainting problems, which is expected to have better reconstruction performances along curve borders. Some other models are also available, such as total generalized variation (TGV) [31], which also imposes a constraint on the Hessian operator, and the TV with more isometric properties by Condat [48]. Some more sophisticated alternatives are also available [128, 243]. The analysis could also focus on optimizing the computational time [79]. The expansion of this work can also focus on the design of the compression system itself, in similar vein of what is proposed in Section 4.3.1. The analysis may focus on optimizing the mask itself, expanding the considerations on compressed sensing of [131, 100]. Finally, in practical scenarios, the analysis may focus on the description of the direct model, making use of analysis such as Duran's work [64]. As the Bayesian framework is in general more robust to the case where more channels are involved, the most natural extension of this framework is to HS acquisition systems, which can be approached with techniques such as spectral unmixing [25], aimed at the reduction of the dimensionality of the channels.

5 Optics foundations for the ImSPOC acquisition system

This chapter is aimed at providing the foundational concepts necessary to model Fabry-Pérot (FP) interferometers and lens optics, which constitute the main building blocks for the description of the image spectrometer on chip (ImSPOC) concept. We describe the involved optical quantities of interest for the final user and link them with the observations on the detectors, characterizing the transformations operated by the involved optical components.

The chapter is divided into five parts: in Section 5.1 and 5.2 we introduce the optical variables at play, in Section 5.3 we perform a brief analysis of the lens optics, in Section 5.4 we describe the general principles of interferometry, and in particular the particular case of its realization through FP etalons. Finally, in Section 5.5, which is a novel contribution of this work, we provide a detailed mathematical model which describes the chain of optical operations of the ImSPOC concept.

5.1 Wave optics

The purpose of optical devices is to apply some transformation on the characteristic properties of light rays. Under particular propagation scenarios that are easily verifiable in practice, light rays are a particular instance of plane electro-magnetic (EM) waves. The aim of this section is to introduce the concept of EM radiation (Section 5.1.2), the properties of plane waves (Section 5.1.2), and the notable physical quantities associated with EM waves.

5.1.1 Electromagnetic radiations

EM fields describe a perturbation of the space due to the movement of electrical charges, whose dynamic, in their classical formulation, is described by Maxwell's equations. These equations define a set of coupled partial differential equations

in the space domain, described by a position vector $\mathbf{r}$ in a given system of tridimensional spatial coordinates (i.e. Cartesian), and in the time domain, described by a scalar value t. Their macroscopic form is given by:

$$\begin{cases} \nabla \cdot \mathbf{d} = \rho_f \,, & \text{(5.1.1a)} \\ \nabla \cdot \mathbf{b} = 0 \,, & \text{(5.1.1b)} \\ \nabla \times \mathbf{e} = -\dfrac{\partial \mathbf{B}}{\partial t} \,, & \text{(5.1.1c)} \\ \nabla \times \mathbf{h} = \mathbf{j}_f + \dfrac{\partial \mathbf{d}}{\partial t} \,, & \text{(5.1.1d)} \end{cases}$$

where $\mathbf{e}$, $\mathbf{b}$, $\mathbf{d}$, and $\mathbf{h}$ are known as the electric, magnetic, displacement and magnetizing field, respectively.[1] ρ_f and $\mathbf{j}_f$ denote the spatial density of charges and currents densities, respectively, and act as sources of energy. The dependency on $\mathbf{r}$ and t is made implicit in every term to simplify the notation. $\nabla\times$ and $\nabla\cdot$ denote the divergence and the curl operators in the spatial domain. $\mathbf{d}$ and $\mathbf{h}$ can be interpreted as auxiliary fields, with the former representing how $\mathbf{e}$ influences the distribution of electric charges and the latter how $\mathbf{b}$ influences the organization of magnetic dipoles in a given medium. This relationship across fields can be made explicit in dielectric materials [125], by defining their permittivity ε and permeability μ and obtaining:

$$\mathbf{D} = \varepsilon \mathbf{E} \,, \quad \text{(5.1.2a)}$$

$$\mathbf{B} = \mu \mathbf{H} \,. \quad \text{(5.1.2b)}$$

The variables ε and μ are in general second order tridimensional tensors, which are functions of both $\mathbf{r}$ and t. However, in the context of this work, we always consider the propagation in linear, non-lossy, homogeneous, and nondispersive isotropic media, for which ε and μ simplify to real scalar quantities, which do not vary with time or with their spatial position. Under these conditions and considering eq. (5.1.1) without charges or currents (in other words, with ρ_f and $\mathbf{j}_f$ being identically equal to zero), the only non-trivial solutions for $\mathbf{b}$ and $\mathbf{e}$ have to obey [125] the relationships:

$$\nabla^2 \mathbf{e} = \mu\varepsilon \frac{\partial^2 \mathbf{e}}{\partial t^2} \,, \quad \text{(5.1.3a)}$$

$$\nabla^2 \mathbf{b} = \mu\varepsilon \frac{\partial^2 \mathbf{b}}{\partial t^2} \,, \quad \text{(5.1.3b)}$$

[1] The names assigned to each field often vary across different authors, so it may be common to encounter $\mathbf{h}$ defined as magnetic field strength, H-field or simply as magnetic field, and similarly $\mathbf{b}$ as B-field or magnetic flux density.

where ∇^2 is the Laplacian operator in the spatial domain (i.e.: $\nabla^2 = \partial^2/\partial r_1^2 + \partial^2/\partial r_2^2 + \partial^2/\partial r_3^2$ in the Cartesian system with axis $[r_1; r_2; r_3]$) [2]. If expressed in Cartesian coordinates, each of the fields can be described by three components (e.g., so that $\mathbf{e} = [e_1; e_2; e_3]$) and each of the relations in (5.1.3) can be separated into three **Helmholtz equations**. The spatial evolution of each component can be then modeled as a wave, whose properties are described in the next section.

5.1.2 Helmholtz wave equation

In general terms, a wave defines any physical quantity $u(\mathbf{r}, t)$ whose evolution in time t and distribution in the space $\mathbf{r} \in \mathbb{R}^3$ is a solution of the so-called **Helmholtz equation**:

$$\nabla^2 u - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0\,, \tag{5.1.4}$$

where ∇^2 denotes the Laplacian operator in the spatial domain. The uniqueness of its solution is imposed by assigning boundary and initial conditions for u. It can be shown [202] that the **principle of superposition** applies. Assuming that u has limited energy, any particular solution u of (5.1.4) can be expressed with its Fourier decomposition:

$$u(\mathbf{r}, t) = \int_0^{+\infty} u_\nu(\mathbf{r}, t) d\nu\,, \tag{5.1.5}$$

In other terms, u is a combination of a set of **monochromatic wavefunctions** $u_\nu(\mathbf{r}, t)$ at a fixed frequency ν in the form:

$$u_\nu(\mathbf{r}, t) = |U(\mathbf{r}, \nu)| \cos\left(2\pi\nu t + \varphi(\mathbf{r})\right)\,, \tag{5.1.6}$$

where $|U(\mathbf{r}, \nu)|$ and $\arg\{(\mathbf{r}, \nu)\}$ denote the **amplitude** the **phase** of the monochromatic wave and are constant with time. For a given fixed frequency ν, it is common to adopt an alternative representation of the wavefunction $u_\nu(\mathbf{r}, t)$ in terms of the **complex function** $\check{U}_\nu(\mathbf{r}, t)$:

$$\begin{aligned} \check{U}_\nu(\mathbf{r}, t) &= |U(\mathbf{r}, \nu)| \exp\left(j\underline{/U(\mathbf{r}, \nu)}\right) \exp\left(j2\pi\nu t\right)) && \text{(5.1.7a)} \\ &= U(\mathbf{r}, \nu) \exp\left(j2\pi\nu t\right)\,, && \text{(5.1.7b)} \end{aligned}$$

where $U(\mathbf{r}, \nu) = |U(\mathbf{r}, \nu)| \exp\left(j\underline{/U(\mathbf{r}, \nu)}\right)$ is known in the literature as **complex amplitude**. This representation is fully reversible, so that it is possible to recover

[2] In the literature the tern $[r_1; r_2; r_3]$ is typically denoted with $[x, y, z]$.

the original wavefunction $u_\nu(\mathbf{r},t) = \mathrm{Re}\left\{\check{U}_\nu(\mathbf{r},t)\right\}$, where $\mathrm{Re}\{\cdot\}$ denotes the real part of its argument. As an additional advantage, the Helmholtz equation can map directly into the equivalent expression:

$$\nabla^2\check{U} - \frac{1}{c^2}\frac{\partial}{\partial t^2}\check{U} = 0\,, \tag{5.1.8a}$$

$$\nabla^2 U_\nu - \left(\frac{2\pi\nu}{c}\right)^2 \frac{\partial}{\partial t^2}U_\nu = 0\,, \tag{5.1.8b}$$

$$\nabla^2 U_\nu - k^2\frac{\partial}{\partial t^2}U_\nu = 0\,, \tag{5.1.8c}$$

and one can apply the same boundary/initial conditions to reach the same result in the new representation. The expression at a given frequency ν at the step (5.1.8b) is obtained by substituting eq. (5.1.7b) into (5.1.8a) and by defining the so-called **angular wavenumber** $k := 2\pi\nu/c$.

5.1.3 Plane waves

Plane waves are the solutions of Helmholtz equation with the condition of free charges propagation in a lossless, isotropic, and homogeneous medium [139]. These solutions are characterized by a complex amplitude in the form:

$$U(\mathbf{r},\nu) = U_0(\nu)\exp\left(-j\mathbf{k}\cdot\mathbf{r}\right)\,, \tag{5.1.9}$$

where $U_0(\nu) = |U_0(\nu)|\exp\left(j\underline{/U_0(\nu)}\right)$ is known as **complex envelope** and is defined by a complex scalar with amplitude $|U_0(\nu)|$ and phase $\underline{/U_0(\nu)}$. The **wavevector** $\mathbf{k}$ is an intrinsic characteristic of the plane wave; its magnitude (obtained by comparing eq. (5.1.9) with (5.1.8b)) is equal to the angular wavenumber and consequently strictly related to the frequency of the wave, while its unit vector defines its direction of propagation, regulated through the scalar product $\mathbf{k}\cdot\mathbf{r}$.

To intuitively illustrate the function of each parameter of a plane wave, let us suppose that, in a system of coordinates such as the Cartesian or cylindrical one, the propagation is in the direction of the z-coordinates, which we denote with r_3. With this assumption, eq. (5.1.9) simplifies to $U_\nu(r_3) = U_0(\nu)\exp\left(-jkr_3\right)$. According to eq. (5.1.7a), the associated wavefunction can be obtained by taking the real part of $U_\nu(r_3)$:

$$u_\nu(r_3,t) = |U_0(\nu)|\cos(\omega t - kr_3 + \underline{/U_0(\nu)}) \tag{5.1.10a}$$

$$= |U_0(\nu)|\cos(2\pi(\nu t - \frac{r_3}{\lambda}) + \underline{/U_0(\nu)})\,, \tag{5.1.10b}$$

where $k = 2\pi/\lambda$ and $\omega = 2\pi\nu$; λ is simply known as **wavelength**, while ω is frequently referred to as **angular frequency**. The resulting wavefunction $u_\nu(r_3, t)$ can be analyzed both in the time and space domains:

- If we fix a specific point in space r_3, the value $T = 1/\nu$, known as **period**, represents the time interval to wait for the wave to return the same state, or, equivalently in our context, to exhibit the same phase. Similarly, the frequency $\nu = \omega/(2\pi)$, when expressed in Hz, represents the average amount of times the wave returns to the same state within a 1 second interval.
- If we fix a specific instant of time t, the surfaces which exhibit the same phase, known as **wavefronts** are all planar and perpendicular to the direction of propagation, and the distance between two consecutive wavefronts is defined as the **wavelength** λ. Similarly, if we travel along the direction $\mathbf{k}$ for 1 meter, the (non-angular) **wavenumber** [3] $k/(2\pi)$, if expressed in m^{-1}, denotes the average amount of surfaces with the same phase for each meter along the direction of propagation.

It is worth noting that the variables λ and ν, which characterize the propagation of the plane wave with regard to space and time, respectively, are not independent, since they are related through the characteristic phase velocity c of the propagation medium itself, which imposes $c = \lambda\nu$. In the field of optics, it is common practice to define the phase velocity in relation to the speed of light in the vacuum, a universal physical constant $c_0 = 299792458\mathrm{m/s}$ with no dependency on the frequency ν. With this definition, the phase velocity is often defined in relative terms as $c = c_0/n$, where the so-called **refractive index** n acts as an intrinsic characteristic of the medium. The Fourier analysis of a **wave packet**, defined as a set of waves with different frequency, may be performed as a combination of different monochromatic waves ν. This frequency will stay fixed across different propagation media, while the wavelength λ may vary, as the phase velocity c varies as well when switching between different propagation media. To avoid this unnecessary complication, it is useful to introduce the wavelength and wavenumber in the vacuum, defined respectively as $\lambda_0 = c_0/\nu$ and $\sigma = 1/\lambda_0$, which will remain constant across different media. As a final remark, one could wonder how it is possible for a field to exist when we considered a charge-free expression of Maxwell's equations to derive eq. (5.1.3); this assumption was taken to consider a scenario in which the EM radiation is not perturbed by charges themselves. One common assumption for this condition to be verified, known as **far field**, is to consider the distance between any source and the position of the field itself to be longer than the **Fraunhofer's distance**

[3] The wavenumber is simply denoted with σ in most publications, but, not to generate confusion in the reader, in this thesis σ exclusively denotes the wavenumber in the vacuum.

$d_f = 2d_r^2/\lambda$, with d_r identifying the largest dimension of the source radiator. Under this approximation we can consider the radiator as a point source, and the solution of the Maxwell's equations is given by spherical wavefronts centered around the radiator, which can be locally approximated as planar wavefronts [18].

5.1.4 Transverse electro-magnetic waves

By comparing eq. (5.1.3) with (5.1.4), it can be promptly obtained that the phase velocity can be expressed in terms of the characteristic of the propagating medium as $c = 1/\sqrt{\varepsilon\mu}$. The solution of Helmotz equations (5.1.3) can thus be rewritten as:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_0 \exp\left(-j\mathbf{k}\cdot\mathbf{r}\right) , \tag{5.1.11a}$$

$$\mathbf{H}(\mathbf{r}) = \mathbf{H}_0 \exp\left(-j\mathbf{k}\cdot\mathbf{r}\right) , \tag{5.1.11b}$$

where $\mathbf{E}_0$ and $\mathbf{H}_0$ denote respectively the complex envelopes of $\mathbf{E}$ and $\mathbf{H}$, which are the complex functions associated with the fields $\mathbf{e}$ and $\mathbf{h}$. By substituting (5.1.11a) and (5.1.11b) into the system of Maxwell's equations (5.1.1) and applying it to complex functions, we obtain:

$$\mathbf{k}\times\mathbf{H}_0 = -\omega\varepsilon\mathbf{E}_0 , \tag{5.1.12a}$$

$$\mathbf{k}\times\mathbf{E}_0 = \omega\mu\mathbf{H}_0 . \tag{5.1.12b}$$

The resulting $\mathbf{E}$ and $\mathbf{H}$ are perpendicular both to the direction of propagation and to each other, generating a mode of propagation that is commonly known as **transverse electro-magnetic (TEM)**. The dependence between the fields is not simply limited to their direction, as it can be shown that their complex amplitudes are also related through the following expressions:

$$|\mathbf{H}_0| = \sqrt{\frac{\mu}{\varepsilon}}|\mathbf{E}_0| = \zeta|\mathbf{E}_0| , \tag{5.1.13}$$

where we have defined $\zeta = \sqrt{\mu/\varepsilon}$ as the **admittance** of the medium and we have substituted $|k| = \omega\sqrt{\varepsilon\mu}$ for simplicity.

5.2 Radiometry

In the field of Earth monitoring and remote sensing applications, which is the main target field of applications of this work, it is of the utmost importance to properly measure the transfer of energy associated with EM radiations, commonly known as **radiant energy**. The aim of this section it to briefly introduce and describe the measurable quantities associated with the radiant energy and their associated measurement techniques; those are the topic of the domain known as **radiometry**. Firstly, Section 5.2.1 introduces the Poynting vector, providing a bridge between the formalism of EM waves and radiometric quantities; the latter are then formally described in Section 5.2.3.

Finally, the description of the characteristics of the photodetector, a sensor devoted to their measurement of radiometric quantities, is the topic of Section 5.2.4. While the focus of this work is mainly on Earth remote sensing, the applicability of the concepts of this section goes far beyond the scope of Earth monitoring; for the interested reader, a more detailed treatment on the subject of radiometry is provided in the works of Grum and Wyatt [103, 235].

5.2.1 Poynting vector

Energy flux defines the energy transfer of a field per unit area and per unit time and is measured in $\mathrm{Wm}^{-2} = \mathrm{Js}^{-1}\mathrm{m}^{-2}$). Its directional value is described, for EM fields, by a mathematical entity known as **Poynting vector**, which allows to impose continuity conditions for the conservation of the energy through the **Poynting theorem** [195].

The expression of the Poynting vector $\mathbf{s}$, as it appears in the original work, is defined for linear nondispersive metals as:

$$\mathbf{s} = \frac{1}{2}\left(\mathbf{e} \times \mathbf{d} + \mathbf{b} \times \mathbf{h}\right) \tag{5.2.1a}$$

$$= \mathbf{e} \times \mathbf{h}\,. \tag{5.2.1b}$$

where the simplification to the second expression (5.2.1b) is only valid for dielectric media and obtained through eq. (5.1.2). As the Poynting vector varies with time, its instantaneous amplitude cannot be measured with any practical detector, as it is characterized by a certain response time. To avoid this inconvenience, it is more

practical to average the expression (5.2.1) over a time window Δt (i.e., the period $T = 1/\nu$ for plane waves) short enough compared to the other times of interest in the process under study. This average is commonly known as the **irradiance** and given by:

$$\mathcal{E}(\mathbf{r}, t) = \frac{1}{\Delta t} \int_{\Delta t} |\mathbf{s}| dt \,. \tag{5.2.2}$$

In terms of complex functions, the Poynting vector can be rewritten as:

$$\mathbf{s} = \mathrm{Re}\left\{\mathbf{E} \exp(j\omega t)\right\} \times \mathrm{Re}\{\mathbf{H} \exp(j\omega t)\} \tag{5.2.3a}$$

$$= \frac{1}{4} \left(\mathbf{E} \times \mathbf{H}^* + \mathbf{E} \times \mathbf{H} + \exp(j2\omega t)\mathbf{E} \times \mathbf{H} + \exp(-j2\omega t)\mathbf{E}^* \times \mathbf{H}^*\right) , \tag{5.2.3b}$$

where we have used the relation $\mathrm{Re}\{\mathbf{U} \exp(j\omega t)\} = \mathbf{U} \exp(j\omega t) + \mathbf{U}^* \exp(-j\omega t)$, which is valid for any complex envelope $\mathbf{U}$, as they have no dependence with t. By averaging eq. (5.2.3b) over an interval of time equal to any multiple of an oscillation cycle $1/\nu$, the third and fourth terms become null, which allows to define a **complex Poynting vector** $\mathbf{S}$:

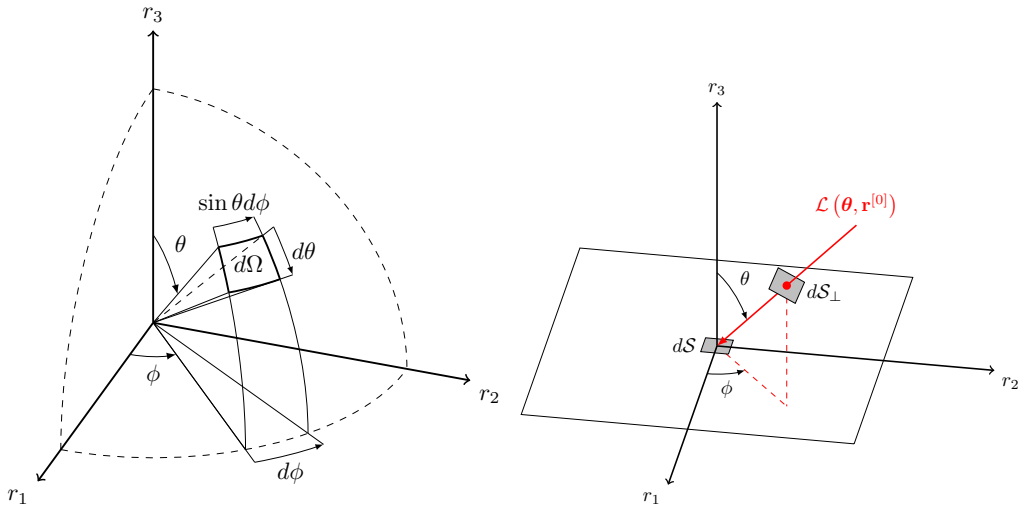
$$\mathbf{S} = \frac{1}{2} \mathbf{E} \times \mathbf{H}^* \,. \tag{5.2.4}$$

If the above relation is applied to TEM waves with complex envelopes $\mathbf{E}_0$, we obtain:

$$\mathcal{E} = |\mathbf{S}| = \frac{1}{2} \sqrt{\frac{\mu}{\varepsilon}} |\mathbf{E}_0|^2 \,. \tag{5.2.5}$$

5.2.2 Solid angle

A solid angle, denoted as Ω, is defined as the tridimensional angular volume that a certain surface in the space, projected over a unit sphere, subtends to its center point, known as **apex**. In the context of radiometry, it is used to identify the amount of field of view (FoV) from a particular point of observation covered by a certain object. The term solid angle sometimes also defines its measure in steradians (symbol: sr), which is equivalent to area of the surface $\mathcal{S}$ of the observed object projected over the unit sphere centered in the apex. It is useful to define the solid angle in terms of spanning angles in spherical coordinates with the origin in the apex. With reference to the geometry shown in Fig. 5.1a, each point of the unit surface cant be identified by a couple of spherical coordinates $\boldsymbol{\theta} = [\theta; \phi]$, whose elements are known as polar and azimuth angle, respectively. The area $\mathcal{S}$ projected over the unit sphere is composed by an infinite set of elementary contributions $d\Omega$, which can



(a) Solid angle in spherical coordinates **(b)** Radiance over a detector surface

Fig. 5.1. The figure on the left shows a representation of the elementary solid angle $d\Omega$ over an unitary sphere (bounded by the dashed lines). This visual representation shows how it's link with the variation $d\theta$ of the polar angle and $d\phi$ of the azimuth angle. On the right, a representation of the radiance incident vector over an elementary detection surface $d\mathcal{S}$ centered at $\mathbf{r}^{[0]}$.

be expressed in terms of the variation $d\theta$ of the polar angle and $d\phi$ of the azimuth angle as follows:

$$d\Omega = \sin\theta d\theta d\phi \,. \tag{5.2.6}$$

The total measure of the solid angle is thus given by $\int\int_{\mathcal{S}} \sin\theta d\theta d\phi$.

5.2.3 Radiometric measures

In the context of radiometry it is often required to quantify the transfer of radiant energy Q_e, which is the energy (measured in Joule) associated with the EM field, in terms of directional, temporal, or spectral parameters. This is useful to characterize the emission of a given radiator thanks to the detection of the incident radiation over a sensor. These quantities can be interpreted either in term of emitted or received energy, but this section is presented from the perspective of energy receptors, as our work is mostly focused on the detection process. The four fundamental quantities defined in the field of radiometry are presented below:

- **Radiant flux** Φ: defines the flux of the radiant energy Q_e per unit of time that flows across a given surface $\mathcal{S}$. In mathematical terms, Φ can be interpreted as the infinitesimal variation dQ_e/dt of Q_e with respect to an infinitesimal

variation dt of time. In other terms, the radiant flux Φ acts as a measure of EM power in the same way Q_e acts as a measure EM energy. The radiant flux is formally defined by:

$$\Phi = \frac{1}{\Delta t} \int_{\Delta t} \iint_{\mathcal{S}} |\mathbf{s}| \cdot \vec{\mathbf{i}_{\mathcal{S}}} \, d\mathcal{S} \, dt \,, \tag{5.2.7}$$

where $\vec{\mathbf{i}_{\mathcal{S}}}$ denotes the unit vector normal to the unit surface and Δt defines a time window multiple than the oscillation cycle, but shorter than the dynamics of all quantities of interest.

- **Irradiance** $\mathcal{E}$: defines the flow of radiant energy per unit area perpendicular to the direction of the flow itself. In mathematical terms, if $d\mathcal{S}_\perp$ is a surface element perpendicular to the direction of Φ, the irradiance can be expressed as the infinitesimal variation $\mathcal{E} = \frac{d\Phi}{d\mathcal{S}_\perp}$. As described in Section 5.2.1, the irradiance can also be expressed as the amplitude of the Poynting vector, averaged over an oscillation cycle. The full knowledge of $\mathcal{E}$ as a function of the position in space is a stronger information than the radiant flux, as the latter can be obtained integrating the irradiance over the surface under investigation.
- **Radiance** $\mathcal{L}$: In most practical scenarios, the detector element is described by a certain surface in the space $\mathcal{S}^{[t]}$, which can be decomposed in a series of elementary surfaces oriented in space. Any of those surfaces, denoted here with $d\mathcal{S}$, can be seen as a point detector in space, characterized by a given normal $\vec{\mathbf{i}_{\mathcal{S}}}$, acting as apex for its associated FoV. This FoV can be partitioned into a continuous set of elementary solid angles $d\Omega$; the FoV contained within each of these elements can be considered small enough within the volume of space identified by $d\Omega$, such that the angle of incidence of the rays contained in it is constant. From the perspective of the incident ray, the element area $d\mathcal{S}$ of the detector is seen with a relative angle with respect to its normal. Given the geometry of the problem shown in Fig. 5.1b, the projected area $d\mathcal{S}_\perp$ of $d\mathcal{S}$ in the direction orthogonal to the direction of incidence is given by $d\mathcal{S}_\perp = d\mathcal{S} \cos\theta$, where θ defines the angle between the direction of the incident ray and the normal to the detector area. The radiance is henceforth defined as the radiant flux flowing across a given surface, per unit solid angle $d\Omega$ with apex centered in $d\mathcal{S}$ per unit projected area $d\mathcal{S}_\perp$, or in mathematical terms:

$$\mathcal{L} = \frac{d^2\Phi}{d\Omega d\mathcal{S}_\perp} = \frac{d^2\Phi}{d\Omega d\mathcal{S} \cos\theta} \,. \tag{5.2.8}$$

The radiance fully characterizes a detection process, since the radiant flux flowing across any given detector can be obtained by integrating the radiance both over its surface and over the full extent of its FoV.

- **Radiant intensity** $\mathfrak{I}$: is defined as the flow of radiant energy per unit solid angle $\mathfrak{I} = \frac{\partial \Phi}{\partial \Omega}$ and acts as a simplification of the concept of radiance, but applied to a point detector. The radiant intensity is a directional quantity and it can be expressed in terms of spherical coordinates as $\mathfrak{I}(\boldsymbol{\theta})$, which roughly represents the flow of radiant energy received by a point detector in a certain angular direction $\boldsymbol{\theta} = [\theta; \phi]$.
- **Radiant Intensity** $\mathcal{D}$: is defined as the incident amount radiant flux incident to the point source per unit solid angle, that is $\mathcal{D} = d\Phi/d\Omega$. The radiant intensity is a simplification of the concept of radiance for point detectors, as we can imagine that, if the whole surface of a given detector is distant enough from all emitting sources, then the radiant flux can only be considered only as a function of the incident angle.

The radiometric variables are also often defined in terms of their spectral density. For example, for the radiance, it is possible to define a **spectral radiance** either in terms of the optical frequency optical frequency $\mathcal{L}_\nu = \frac{d\mathcal{L}}{d\nu}$, of the wavelengths $\mathcal{L}_\lambda = \frac{d\mathcal{L}}{d\lambda}$, or of the wavenumbers $\mathcal{L}_\sigma = \frac{d\mathcal{L}}{d\sigma}$. Each of these representation can be expressed in terms of one another with appropriate mathematical adjustments [58]; in this thesis, unless otherwise noted, we employ the expression in terms of wavenumbers σ. The passage from the spectral to non-spectral definition is obtained by integrating over σ, i.e. for the radiance:

$$\mathcal{L} = \int_0^{+\infty} \mathcal{L}_\sigma d\sigma \,. \tag{5.2.9}$$

The definitions of the radiometric quantities are summarized in Table 5.1.

Table 5.1. Summary of the radiometric quantities introduced in Section 5.2.3.

Name	Symbol	Unit	Definition
Radiant Energy	Q_e	J	
Radiant Flux	Φ	W	dQ_e/dt
Spectral Radiant flux	Φ_σ	$\mathrm{W}\,\mu\mathrm{m}$	$d\Phi/d\sigma$
Irradiance	$\mathcal{E}$	$\mathrm{W\,m^{-2}}$	$d\Phi/d\mathcal{S}_\perp$
Spectral Irradiance	$\mathcal{E}_\sigma$	$\mathrm{W\,m^{-2}}\,\mu\mathrm{m}$	$d\mathcal{E}/d\sigma$
Radiant Intensity	$\mathcal{D}$	$\mathrm{W\,sr^{-1}}$	$d\Phi/d\Omega$
Spectral Radiant intensity	$\mathcal{D}_\sigma$	$\mathrm{W\,sr^{-1}}\,\mu\mathrm{m}$	$d\mathcal{D}/d\sigma$
Radiance	$\mathcal{L}$	$\mathrm{W\,sr^{-1}\,m^{-2}}$	$d^2\Phi/(d\Omega\cos\theta\partial\mathcal{S})$
Spectral Radiance	$\mathcal{L}_\sigma$	$\mathrm{W\,sr^{-1}\,m^{-2}}\,\mu\mathrm{m}$	$d\mathcal{L}/\partial\sigma$

The quantities we discussed are often defined as densities (e.g.: spatial, directional or spectral), which are defined in a continuous space and thus only measurable with a finite precision, that is, by partitioning the continuous domains into sufficiently

small intervals. E.g., a representation of the spectral radiance can be given as set the radiant fluxes, whose generic element Φ_{li} is:

$$\Phi_{li} = \int_{\sigma_l - \frac{\Delta\sigma}{2}}^{\sigma_l + \frac{\Delta\sigma}{2}} \iint_{\Omega_i} \mathcal{L}_\sigma(\boldsymbol{\theta})\, d\Omega\, d\sigma \tag{5.2.10a}$$

$$\approx \Delta\sigma \mathcal{S}_{\Omega_i} \mathcal{L}_{\sigma_l}(\boldsymbol{\theta}_i)\,, \tag{5.2.10b}$$

where the spectrum is divided in intervals such that the l-th one is $[\sigma_l - \Delta\sigma/2, \sigma_l + \Delta\sigma/2]$ and the FoV is divided in solid angles, such that the i-th one Ω_i has a subtended area on the unit sphere equal to $\mathcal{S}_{\Omega_i}$. If the spectral radiance is reasonably uniform over the span of both intervals (and equal to $\mathcal{L}_{\sigma_l}(\boldsymbol{\theta}_i)$), this justifies the approximation 5.2.10b.

5.2.4 Photodetectors

Photodetectors define a class of sensors for the detection of the incident EM radiation. We provide in this paragraph just a brief introduction; a more in depth analysis, which includes more advanced concepts of photonics, is left to the specialized literature [202, 43].

The most common technology of modern photodetectors are the **photodiodes**, which are composed by a p-n junction (or a PIN structure), operating in zero or in reverse bias mode. Their operating principle is based on the **inner photoelectric effect**, which defines the generation of an electron-hole pair within the photodiode, because of the the energy generated by a photon striking on its surface. If the pair is generated in the depletion region (the region in the proximity of the doping discontinuities where mobile charges have been diffused by the electric field and is depleted of carriers), the electric field generates a drift current, so that electrons can be collected at the cathode and holes at the anode. Therefore, the resulting photo-current is approximately proportional to the incident irradiance. The vast majority of photodiodes employ:

- **a PIN structure**: to enlarge the depletion area by inserting an intrinsic region between the p-n junction;
- **a reverse bias mode of operation**: to increase the speed of the process of photon absorption at the cost of more intense dark currents (current that circulates in the device, even with no illumination).

Some devices, known as **avalanche photodiodes**, additionally make use of the avalanche effect to multiply the collected carriers.

When photons impact the photodiode, they are not necessarily absorbed, as the electron-hole pair may either not be formed or recombine before its collection over the electrical contacts. To model this nonideality effect, it is common to define an associated function $0 < \eta < 1$, called **quantum efficiency**, which defines the probability that an incident photon becomes part of the flux of generated electrons, which is typically a function of the wavelength of the incident ray.

The physical processes that are involved for the detection of the incident photons introduce sources of unwanted deviations for the desired transfer function, which can be categorized into two types:

- **Thermal noise**: is the main effect of **photon noise**, which are due to physical phenomena in the detector itself. Thermal noise is determined by the thermal fluctuation of circulation of carriers in the electronic circuit, which can be modeled with an additive white Gaussian noise (AWGN) whose standard deviation is proportional to the temperature and the resistance of the circuit conductor.
- **Shot noise**: is the most relevant effect of **circuit noise**, which is due to the electrical circuits associated the receiver. Shot noise is due to the fluctuation of the bias current through depleted regions. This can be modeled by a Poisson process with zero mean and power spectral density (PSD) proportional to the dark current in the device. For a sufficiently large number of incident photons, shot noise approaches a normal distribution.

5.3 Optics of lenses

As radiometric measures involve the directional measure of the radiance, one of the main requirements is to be able to focus over a single detector a set of parallel rays incoming from a certain direction; this focusing effect is generally accomplished by the use of lenses. In this section we describe the optical phenomena related to the transfer of rays across media characterized by stepwise uniform refraction indices; we also introduce some basic tools to model the behaviour of focusing lenses.

5.3.1 Fresnel equations

Let a monochromatic plane wave be incident to a planar boundary between two linear, homogeneous, isotropic material with refractive indices n_1 and n_2 respectively. When a light ray is incident to this boundary from the first medium, it is split into two components, one which is reflected back into the medium itself and the other one which is transmitted into the second medium. We denote the incident, reflected and transmitted rays with the superscripts $[in]$, $[r]$ and $[t]$, respectively. If the boundary itself is lossless, the principle of the conservation of energy implies that the overall amount of EM energy balance is preserved, or, in other words, that the incident energy is equal to the sum of the reflected and transmitted energy. The reflected and transmitted rays can be characterized by their direction of propagation and their associated quota of energy; such characterization can be obtained as solution of Maxwell's equations, assuming as boundary conditions that the components tangent to the discontinuity of both the electric and magnetization field are the same across the two media. With regards to the direction, let $\theta^{[in]}$, $\theta^{[r]}$, and $\theta^{[t]}$ denote the **incidence angle**, the **reflection angle**, and the **transmission angle**, respectively. Those angles obey the following relationships:

$$\theta^{[r]} = \theta^{[in]}\,, \tag{5.3.1a}$$

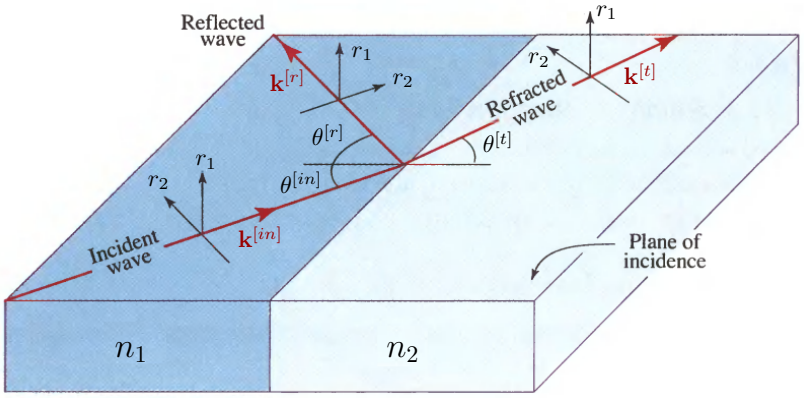
$$n_1 \sin\theta^{[in]} = n_2 \sin\theta^{[t]}\,, \tag{5.3.1b}$$

respectively known as **law of reflection** and **Snell's law**. With regard to the energy balance, a common practice to decouple the field in terms of its polarization. Let $\mathbf{k}^{[in]}$, $\mathbf{k}^{[r]}$, and $\mathbf{k}^{[t]}$ denote the wavevectors associated with the incident, reflected and transmitted ray, respectively. The polarization determines the direction of the electric fields with respect to the the plane formed by the wavevector $\mathbf{k}^{[in]}$ of the incoming wave and the normal to the boundary, commonly known as **incidence plane**.

The reference coordinate system is shown in Fig. 5.2 and is assumed to be taken according to the direction of the propagation, with r_3 defining the direction of $\mathbf{k}^{[in]}$ and r_2 the direction parallel to the incidence plane and perpendicular to $\mathbf{k}^{[in]}$.

We can decompose the plane wave into:

- a **transverse electric (TE) mode**, where the electric field complex envelope $E_\perp$ is directed in the r_1 direction, perpendicular to the incidence plane;
- a **transverse magnetic (TM) mode** where the electric field complex envelope $E_\parallel$ is directed in the r_2 direction, parallel to the incidence plane.



Source: Adapted from [202]

Fig. 5.2. Coordinate system for the reflected and transmitted ray generated by a ray incident on a discontinuity slab.

In the following, we will denote the complex envelopes in the assigned directions for the electric and magnetization fields as follows:

$$E^{[in]} = E_0^{[in]} \exp(-j\mathbf{k}^{[in]} \cdot \mathbf{r}), \tag{5.3.2a}$$

$$E^{[r]} = \tilde{r}_{\perp,\parallel} E_0^{[in]} \exp(-j\mathbf{k}^{[r]} \cdot \mathbf{r}), \tag{5.3.2b}$$

$$E^{[t]} = \tilde{t}_{\perp,\parallel} E_0^{[in]} \exp(-j\mathbf{k}^{[t]} \cdot \mathbf{r}), \tag{5.3.2c}$$

$$H^{[in]} = \zeta_1 E_0^{[in]} \exp(-j\mathbf{k}^{[in]} \cdot \mathbf{r}), \tag{5.3.2d}$$

$$H^{[r]} = \zeta_1 \tilde{r}_{\perp,\parallel} E_0^{[in]} \exp(-j\mathbf{k}^{[r]} \cdot \mathbf{r}), \tag{5.3.2e}$$

$$H^{[t]} = \zeta_2 \tilde{t}_{\perp,\parallel} E_0^{[in]} \exp(-j\mathbf{k}^{[t]} \cdot \mathbf{r}), \tag{5.3.2f}$$

where we denoted with $E^{[in]}$, $E^{[r]}$, $E^{[t]}$ the complex functions of the component of electric field, with $H^{[in]}$, $H^{[r]}$, $H^{[t]}$ the complex functions of the component of the magnetization field, with $E_0^{[in]}$ the complex envelope of the incident electric field, with the superscript denoting if they are assigned the incident, reflected or transmitted ray. The values $\tilde{r}_{\perp,\parallel}$ and $\tilde{t}_{\perp,\parallel}$ denote the **reflection coefficient** and the **transmission coefficient**, respectively, with the subscript $\parallel$ corresponding to the TM mode and $\perp$ to the TE mode. Finally ζ_1 and ζ_2 are the admittances of the two media, which act as a link between the complex envelope of the magnetization and electric field through eq. (5.1.13).

We derive the expressions of $\tilde{r}_\perp$ and $\tilde{t}_\perp$ by imposing the boundary conditions for the TE mode, and $\tilde{r}_\parallel$ and $\tilde{t}_\parallel$ by imposing them for the TM mode:

- **Transverse electric (TE)**: For the TE mode, the boundary conditions assume the following form in correspondence to the discontinuity:

$$E^{[in]} + E^{[r]} = E^{[t]}\,, \tag{5.3.3a}$$
$$H^{[in]}\cos\theta^{[in]} - H^{[r]}\cos\theta^{[in]} = H^{[t]}\cos\theta^{[t]}\,, \tag{5.3.3b}$$

  and, by substituting them in (5.3.2), we obtain:

$$\tilde{r}_{\perp} = \frac{\zeta_1\cos\theta^{[in]} - \zeta_2\cos\theta^{[t]}}{\zeta_1\cos\theta^{[in]} + \zeta_2\cos\theta^{[t]}}\,, \tag{5.3.4a}$$
$$\tilde{t}_{\perp} = \frac{2\zeta_1\cos\theta^{[in]}}{\zeta_1\cos\theta^{[in]} + \zeta_2\cos\theta^{[t]}}\,. \tag{5.3.4b}$$

- **Transverse magnetic (TM)**: For the TM mode, the boundary conditions at the discontinuity assume instead the following form:

$$E^{[in]}\cos\theta^{[in]} - E^{[r]}\cos\theta^{[in]} = E^{[t]}\cos\theta^{[t]}\,, \tag{5.3.5a}$$
$$H^{[in]} + H^{[r]} = H^{[t]}\,, \tag{5.3.5b}$$

  and by substituting them in (5.3.2), we obtain:

$$\tilde{r}_{\parallel} = \frac{\zeta_2\cos\theta^{[in]} - \zeta_1\cos\theta^{[t]}}{\zeta_2\cos\theta^{[in]} + \zeta_1\cos\theta^{[t]}}\,, \tag{5.3.6a}$$
$$\tilde{t}_{\parallel} = \frac{2\zeta_1\cos\theta^{[in]}}{\zeta_2\cos\theta^{[in]} + \zeta_1\cos\theta^{[t]}}\,. \tag{5.3.6b}$$

In terms of irradiances $\mathcal{E}^{[in]}$, $\mathcal{E}^{[r]}$ and $\mathcal{E}^{[t]}$ of the incident, reflected and transmitted wave, respectively, we can define the **reflectivity** $\mathcal{R}$ and the **transmissivity** $\mathcal{T}$ as:

$$\mathcal{E}^{[r]} = \mathcal{R}\mathcal{E}^{[in]}\,, \tag{5.3.7a}$$
$$\mathcal{E}^{[t]} = \mathcal{T}\mathcal{E}^{[in]}\,. \tag{5.3.7b}$$

Once again, the reflectivity $\mathcal{R}_{\perp,\parallel}$ and transmissivity $\mathcal{T}_{\perp,\parallel}$ are different for the TE and TM mode. Since the irradiance is derived in the expression (5.2.5) as the amplitude of the complex Poynting vector, we obtain:

$$\mathcal{R}_{\perp,\parallel} = \left|\tilde{r}_{\perp,\parallel}\right|^2 , \tag{5.3.8a}$$

$$\mathcal{T}_{\perp,\parallel} = 1 - \mathcal{R}_{\perp,\parallel} , \tag{5.3.8b}$$

where eq. (5.3.8b) was derived as a result of the principle of conservation of energy.

By substituting eq. (5.3.4a) and (5.3.4b) into (5.3.8a), we finally obtain:

$$\mathcal{R}_{\perp} = \left|\frac{\zeta_1 \cos\theta^{[in]} - \zeta_2 \cos\theta^{[t]}}{\zeta_1 \cos\theta^{[in]} + \zeta_2 \cos\theta^{[t]}}\right|^2 = \left|\frac{n_1 \cos\theta^{[in]} - n_2 \cos\theta^{[t]}}{n_1 \cos\theta^{[in]} + n_2 \cos\theta^{[t]}}\right|^2 , \tag{5.3.9a}$$

$$\mathcal{R}_{\parallel} = \left|\frac{\zeta_2 \cos\theta^{[in]} - \zeta_1 \cos\theta^{[t]}}{\zeta_2 \cos\theta^{[in]} + \zeta_1 \cos\theta^{[t]}}\right|^2 = \left|\frac{n_2 \cos\theta^{[in]} - n_1 \cos\theta^{[t]}}{n_2 \cos\theta^{[in]} + n_1 \cos\theta^{[t]}}\right|^2 . \tag{5.3.9b}$$

The rightmost side of the equation is valid only for non-magnetic media, or in other words such that their permeability is equal to that of the void, which is a commonly verified hypothesis for the range of optical frequencies. In practical application, where light is not artificially polarized, the amount of TE and TM modes are present in equal amount, so that the effective reflectivity of the discontinuity can be assumed to be an average of the two contributions:

$$\mathcal{R} \approx \frac{1}{2}\left(\mathcal{R}_{\perp} + \mathcal{R}_{\parallel}\right) . \tag{5.3.10}$$

In the special case of normal incidence ($\theta^{[in]} = \theta^{[t]} = 0$), the reflective coefficients are the same for both polarizations and the reflectivity can be simplified to:

$$\mathcal{R} = \left|\frac{n_1 - n_2}{n_1 + n_2}\right|^2 . \tag{5.3.11}$$

5.3.2 Ray transfer matrix analysis

The **ray transfer matrix analysis** defines a series of tools to describe the direction and position of rays in the space travelling across optical systems, such as lenses.

From a physical point of view, if we consider light rays that travel across uniform media in space, their path is straight everywhere except across their boundaries, as the light beam changes direction when transmitted across adjacent media. The ray transmission is regulated by Snell's law (5.3.1b), where the angle of incidence and transmission are intended here with respect to the normal to a generally curve discontinuity.

Mathematically, lenses can be defined as a closed region of a transparent medium, typically with rotational symmetry around a certain axis known as **optical axis**. In the framework of the **ray transfer matrix analysis**, the incident and transmitted rays can be defined as the light rays travelling across an input and output plane, respectively, perpendicular to the optical axis.

Given a certain surface (either input or output), let $\mathbf{r}^{[0]}$ its intersection with the optical axis and let $[r_1, r_2, r_3]$ be a system of Cartesian coordinates centered in $\mathbf{r}^{[0]}$, with r_3 oriented in the direction of the optical axis. Under this system, each ray direction can be uniquely represented by two sets of parameters [88]:

- The **position in the space**, defined by the intersection of the ray with the considered surface, which is identified by the vector $\mathbf{r} = [r_1; r_2]$;
- The **direction of incidence** defines the orientation of the ray with respect to the surface. If the axis origin is shifted by $\mathbf{r}^{[0]}$, this direction is fully described by the polar angle θ between the direction of the ray and r_3 and by the azimuth angle ϕ that the the projection of $\mathbf{k}^{[in]}$ on the discontinuity makes with the r_1 axis. To simplify the notation, the couple is denoted with the vector form $\boldsymbol{\theta} = [\theta; \phi]$.

The described set of variables is also shown (with a slightly different notation) in Fig. 5.8. If we consider $[\mathbf{r}; \boldsymbol{\theta}]$ as a vector of 4 elements, the effect of a certain lens of changing the direction and shifting the position can be modeled with a certain transfer function $\Theta : \mathbb{R}^4 \to \mathbb{R}^4$:

$$\begin{bmatrix} \boldsymbol{\theta}^{[t]} \\ \mathbf{r}^{[t]} \end{bmatrix} = \Theta \left(\begin{bmatrix} \boldsymbol{\theta}^{[in]} \\ \mathbf{r}^{[in]} \end{bmatrix} \right) , \tag{5.3.12}$$

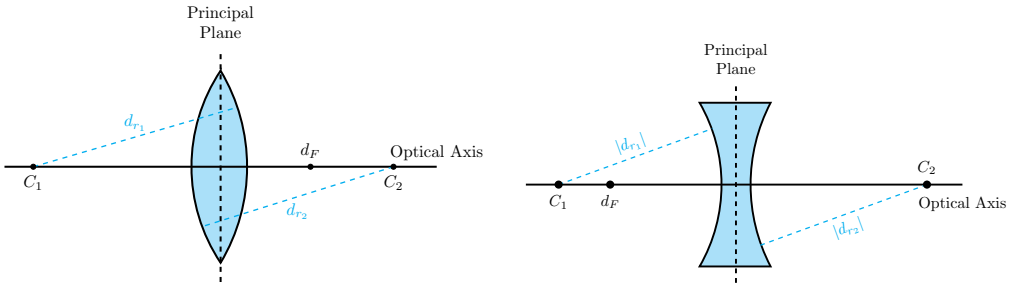
that links the incident vector $\left[\boldsymbol{\theta}^{[in]}; \mathbf{r}^{[in]}\right]$ to the transmitted one $\left[\boldsymbol{\theta}^{[t]}; \mathbf{r}^{[t]}\right]$. If the conditions of linearity between the input and output variables are satisfied, this relationship can be modeled as multiplication by a 4×4 matrix.

A great variety of optical devices is included in this model [120], but we focus our attention to the case of lenses with perfect rotational symmetry. In this case, the incident ray may be specified just with the two scalar variables $r = \|\mathbf{r}\|_2 = \sqrt{r_1^2 + r_2^2}$

and θ, and the transfer function Θ in (5.3.12) is a multiplication by a matrix of sizes 2×2.

5.3.3 Spherical lenses

Spherical lenses are a special case of **simple lenses**, which define a transmissive optical devices manufactured with a single transparent medium (in other words a medium which is homogeneous, nondispersive, isotropic media, and uniform with the wavelength). For spherical lenses, the boundaries of the simple lens are two surfaces with a constant curvature, that is surfaces of spheres with fixed radii d_{r_1} and d_{r_2}. By convention, these radii are given with a positive sign for convex lenses or negative for concave lenses (e.g. if the bounding surfaces bulges inward); an example of a convex and concave lens is show in Fig. 5.3.



(a) Biconvex lens **(b)** Biconcave lens

Fig. 5.3. Visual representation of convex and concave lenses. The focus d_F represents the point of convergence of incident rays parallel to the optical axis coming from the left side. Thin lenses are characterized by only one principal plane.

Given the rotational symmetry of the structure, the optical axis corresponds to the straight path that links the centers of curvature of the two boundaries. Let n be the refraction index within the lens and let n_{0_1} and n_{0_2} the refraction indices on the two external sides of the lens (typically $n_{0_1} = n_{0_2} \approx 1$ for lenses surrounded by air). The ray transfer matrix $\boldsymbol{\Theta} = \boldsymbol{\Theta}^{[3]}\boldsymbol{\Theta}^{[2]}\boldsymbol{\Theta}^{[1]}$ of the overall system can be obtained as cascade of three components: two matrices $\boldsymbol{\Theta}^{[1]}$ and $\boldsymbol{\Theta}^{[3]}$ for the transmission over the discontinuities and a matrix $\boldsymbol{\Theta}^{[2]}$ for the propagation within the lens itself.

To derive the expression of each contribution, we assume that the angle of incidence $\theta^{[in]}$ is small enough so that we can approximate the associated sinusoidal functions with the first term Fourier decomposition: $\sin\theta^{[in]} \approx \theta^{[in]}$. This condition is known as **paraxial approximation**. Implicitly, this assumption implies that the radii of curvature are big enough that each discontinuity surface is approximately planar for

every intersection point $r^{[in]}$; this allows to assume that all the refraction phenomena happen over planes perpendicular to the optical axis, known as **principal planes**. Two such planes can be identified situated at opposite sides of any thick lens, separated by a distance d_L.

Due to the geometry of the ray displacement, the intersection on the back plane is obtained as $r^{[t]} = r^{[in]} + d_L \sin\theta^{[in]}$, which under the paraxial approximation becomes:

$$\boldsymbol{\Theta}^{[2]} = \begin{bmatrix} 1 & d_L \\ 0 & 1 \end{bmatrix} . \tag{5.3.13}$$

With regard to the transfer across a spherical surface discontinuity, one could simply apply Snell's law from eq. (5.3.1b), with the adjustment that angles of incidence are to be considered with respect to the normal to the surface itself, which is directed on the outside of the bounded region for $\boldsymbol{\Theta}^{[1]}$ and on the inside for $\boldsymbol{\Theta}^{[3]}$:

$$\boldsymbol{\Theta}^{[1]} = \begin{bmatrix} 1 & 0 \\ \frac{n_1 - n_0}{R_1 n_0} & \frac{n_0}{n_1} \end{bmatrix} , \tag{5.3.14a}$$

$$\boldsymbol{\Theta}^{[3]} = \begin{bmatrix} 1 & 0 \\ \frac{n_0 - n_1}{R_2 n_0} & \frac{n_1}{n_0} \end{bmatrix} . \tag{5.3.14b}$$

The expression of $\boldsymbol{\Theta}$ can be further simplified in the case in which **thin lens approximation** holds, that is if the thickness d_L of the lens is negligible compared to the radii of curvature d_{r_1} and d_{r_2}. Substituting $d_L = 0$ in eq. (5.3.13) for Θ_2 produces an identity matrix, so the overall transfer function becomes:

$$\begin{bmatrix} \rho^{[t]} \\ \theta^{[t]} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{d_F} & 1 \end{bmatrix} \begin{bmatrix} \rho^{[in]} \\ \theta^{[in]} \end{bmatrix} . \tag{5.3.15}$$

A quick analysis of (5.3.15) shows that $r^{[t]} = r^{[in]}$, which implies that the two principal planes overlaps for thin lenses, as we expected. For the case $n_{0_1} = n_{0_2} = n_0$, eq. (5.3.15) reduces to the **lensmaker equation**:

$$\frac{1}{d_F} = \frac{n - n_0}{n_0} \left(\frac{1}{d_{r_1}} - \frac{1}{d_{r_2}} \right) . \tag{5.3.16}$$

for which we made the value of the **focus** d_F explicit. The obtained result $\theta^{[t]} = \theta^{[in]} - r^{[in]}/d_F$ proves the focusing principle of the lens. In fact, if we consider a series of incident ray beams perpendicular to the principal plane $\theta^{[in]}$, they will

appear deflected on the other side, such that all rays focus on a single point on the optical axis at a distance d_F from the principal plane.

5.4 Principles of interferometry

In this section, we provide a general introduction to the concept of interferometry in Section 5.4.1 and specialize it to the case of two coherent EM waves in Section 5.4.2. The Fourier transform spectrometers (FTSs), a special class of instruments of which ImSPOC is part of, are the topic of Section 5.4.3. A special instance of interferometry, known as Fabry-Pérot (FP), is then detailed in Section 5.4.4 and a mathematical model for the transfer of energy is the topic of Section 5.4.5.

5.4.1 Introduction

Interferometry refers to the set of techniques for the measurement of superimposed EM waves causing interference with each other. The term **interference**, in general terms, refers to the interaction between EM waves, which allows to extract information on the nature of the EM radiation itself.

In the simplest scenario, two monochromatic plane waves with a given frequency ν share the same direction of propagation and, according to the principle of superposition, their combination generates irradiance patterns whose expression is determined by the phase difference between the two waves.

When they are in phase, they generate a **constructive interference**, and their combined amplitude increases, while when they are in opposition of phase, they generate a **destructive interference** and their combined amplitude decreases. Although it is also possible, in very controlled environments, to extract information from the interference of incoherent sources [129], the typical required condition for the exploitation of interferometry is for the waves to exhibit **coherence**, that is to have the same frequency waveform and a constant phase difference. These conditions can be partially relaxed, up to a certain degree which is set by the required amount of correlation between the packet of waves under test [234].

The most common technique to allow wave packets to interfere is to split (e.g with a beam splitter) a single wave into two or more replicas that travel across different paths before recombining. The resulting combined wave, called **interferogram**,

features different **interference fringes**, a series of intensity peaks of the interferogram, which provide information about the different **optical path lengths (OPLs)** travelled by each of the originating waves.

5.4.2 Interference of two waves

Let us assume that two TEM waves are propagating in the space with the same frequency, direction of propagation, and polarization (or, in other terms, that the electric fields of the two waves share the same direction). According to the Fourier analysis, as seen in Section 5.1.1, one could approach the problem by decomposing the waves into a packet of monochromatic waves, each with a given optical frequency ν. The complex functions $\mathbf{E}_1(\nu)$ and $\mathbf{E}_2(\nu)$ of the electic fields of the two waves at ν are in the form:

$$\mathbf{E}_1(\nu) = |\mathbf{E}_1(\nu)| \exp j\varphi_1(\nu)\,, \tag{5.4.1a}$$

$$\mathbf{E}_2(\nu) = |\mathbf{E}_2(\nu)| \exp j\varphi_2(\nu)\,, \tag{5.4.1b}$$

where $|\mathbf{E}_1(\nu)|$ and $|\mathbf{E}_2(\nu)|$ are the amplitudes and $\varphi_1(\nu) = \underline{/\mathbf{E}_1(\nu)}$ and $\varphi_2(\nu) = \underline{/\mathbf{E}_2(\nu)}$ are the phases of each of the two waves, respectively. The complex envelope $\mathbf{E}$ of the combined wave, resulting from the interference of the two, can be simply expressed, because of the superposition principle, as $\mathbf{E}(\nu) = \mathbf{E}_1(\nu) + \mathbf{E}_2(\nu)$. Its associated spectral irradiance $\mathcal{E}_\nu = \frac{1}{2\zeta}|\mathbf{E}(\nu)|$ (with ζ defining the impedance of the medium), is obtained by the expression (5.2.5) of the Poynting vector, and assumes the following form, known as **interference equation**:

$$\mathcal{E}_\nu = \frac{1}{2\zeta}|\mathbf{E}_{\nu,1} + \mathbf{E}_{\nu,2}|^2 \tag{5.4.2a}$$

$$= \frac{1}{2\zeta}\left(|\mathbf{E}_1|^2 + |\mathbf{E}_2|^2 + \mathbf{E}_1^*\mathbf{E}_2 + \mathbf{E}_1\mathbf{E}_2^*\right) \tag{5.4.2b}$$

$$= \mathcal{E}_{\nu,1} + \mathcal{E}_{\nu,2} + 2\sqrt{\mathcal{E}_{\nu,1}\mathcal{E}_{\nu,2}}\cos\varphi\,, \tag{5.4.2c}$$

which allows to express the total irradiance as function of the irradiance associated with each of the two waves $\mathcal{E}_{\nu,1}$ and $\mathcal{E}_{\nu,2}$. The **phase difference** $\varphi = \varphi_2 - \varphi_1$ is the most important parameter for the process of interference, as it decides the amount of constructive inference between the two waves.

As described in Section 5.4.1, the phase difference between coherent waves can be induced artificially by having the two waves travelling across different paths.

In media that are non-dispersive, lossless, isotropic, and homogeneous except for a countable amount of discontinuities, it is possible to define the refractive index $n(\mathbf{r})$ as a function of the space coordinates $\mathbf{r}$. Let the splitting point be defined as the last point where the waves are completely in phase, with a certain initial phase which we can assume equal to zero without loss of generality. The subsequent phase contributions are introduced by the different paths travelled by the two waves. Those two paths $\{L_m\}_{m=1,2}$ can be partitioned into a set of infinitesimal contributions $d\mathbf{r}$, which allow to determine the phase contributions $\{\varphi_m\}_{m=1,2}$ through the line integral:

$$\varphi_m = \int_{L_m} \mathbf{k}(\mathbf{r}) \cdot d\mathbf{r} = \int_{L_m} (2\pi\sigma n(\mathbf{r}))dr \tag{5.4.3a}$$

$$= 2\pi\sigma \int_{L_m} n(\mathbf{r})dr = 2\pi\sigma\delta_m \,, \tag{5.4.3b}$$

where in eq. (5.4.3a) we make use of the fact that the direction of propagation is always parallel to the travelled path, and in eq. 5.4.3b we have substituted $|\mathbf{k}(\mathbf{r})| = 2\pi c(\mathbf{r})/\nu = 2\pi c_0 n(\mathbf{r})$, where σ is the wavenumber in the vacuum and $n(\mathbf{r})$ is the refraction index in the generally inhomogeneous medium. The quantity $\delta_m := \int_{L_m} n(\mathbf{r})dr$ is known as **optical path length (OPL)** and, in homogeneous media, as light propagates in straight lines, it is equal to the product between the geometric path length and the refraction index of the media.

The irradiances associated with each of the two rays at the recombination point may be unequal, as the paths themselves may feature different absorption coefficients. This behaviour can be modeled with a quadratic coefficient $0 < \mathcal{R}^2 < 1$, which defines the ratio between the two irradiances at the recombination spot, so that $\mathcal{E}_2 = \mathcal{R}^2\mathcal{E}_1$. This allows to rewrite eq. (5.4.2) as:

$$\mathcal{E}_\sigma = \left(1 + \mathcal{R}^2(\sigma) + 2\mathcal{R}(\sigma)\cos(2\pi\sigma\delta(\sigma))\right)\mathcal{E}_{\sigma,1} \,, \tag{5.4.4}$$

where we have expressed the spectral irradiance in terms of σ instead of ν and we have defined the **optical path difference (OPD)** as $\delta := \delta_2 - \delta_1$. The explicit dependence by σ in δ is due to the spectral component of the refractive index n.

5.4.3 Fourier transform spectrometers

By analyzing eq. (5.4.4) over all frequencies, as previously shown in Section 5.4.2, we can obtain the irradiance $\mathcal{E}$ of the combined rays, by integrating $\mathcal{E}_\sigma$ over the

domain of the wavenumbers. One implicit assumption is that the complex amplitude of both waves for any given wavenumber is either constant with time (for fully coherent waves) or at least changing slowly with respect to the period of oscillation (for which the waves are often called quasi-monochromatic). In particular, if the refraction index n and the attenuation $\mathcal{R}$ are constant with the frequency ν (hence with σ), we obtain:

$$\mathcal{E}(\delta) = \int_0^{+\infty} \mathcal{E}_\sigma d\sigma \tag{5.4.5a}$$

$$= \int_0^{+\infty} \left(1 + \mathcal{R}^2 + 2\mathcal{R}\cos(2\pi\sigma\delta)\right) \mathcal{E}_{\sigma,1}\, d\sigma \tag{5.4.5b}$$

$$= (1 + \mathcal{R}^2) \int_0^{+\infty} \mathcal{E}_{\sigma,1}\, d\sigma + 2\mathcal{R} \int_0^{+\infty} \cos(2\pi\delta\sigma)\mathcal{E}_{\sigma,1}\, d\sigma \tag{5.4.5c}$$

$$= \frac{1 + \mathcal{R}^2}{(1 + \mathcal{R})^2}\mathcal{E}(\delta = 0) + 2\mathcal{R} \int_0^{+\infty} \cos(2\pi\delta\sigma)\mathcal{E}_{\sigma,1}\, d\sigma \tag{5.4.5d}$$

where we have substituted $\mathcal{E}(\delta = 0) = \int_0^{+\infty} (1 + \mathcal{R})^2\, \mathcal{E}_{\sigma,1} d\sigma$ as the value of the irradiance $\mathcal{E}(\delta)$ in $\delta = 0$, which is commonly known as **OPD-zero** irradiance.

If we factorize eq. (5.4.5) as follows:

$$\frac{1}{2\mathcal{R}}\mathcal{E}(\delta) - \frac{1 + \mathcal{R}^2}{2\mathcal{R}(1 + \mathcal{R})^2}\mathcal{E}(\delta = 0) = \int_0^{+\infty} \cos(2\pi\sigma\delta)\mathcal{E}_{\sigma,1}\, d\sigma \tag{5.4.6a}$$

$$= \int_{-\infty}^{+\infty} \cos(2\pi\sigma\delta)\mathcal{E}_{|\sigma|,1}\, d\sigma \tag{5.4.6b}$$

$$= \int_{-\infty}^{+\infty} \frac{e^{-j2\pi\sigma\delta} + e^{+j2\pi\sigma\delta}}{2}\mathcal{E}_{|\sigma|,1}\, d\sigma \tag{5.4.6c}$$

$$= \int_{-\infty}^{+\infty} e^{-j2\pi\sigma\delta}\mathcal{E}_{|\sigma|,1}\, d\sigma = \mathcal{F}\left[\mathcal{E}_{|\sigma|,1}\right], \tag{5.4.6d}$$

then we can immediately recognize on the right side the Fourier transform of $\mathcal{E}_{|\sigma|,1}$, the even extension of $\mathcal{E}_{\sigma,1}$, that we have denoted in eq. (5.4.6d) with $\mathcal{F}[\mathcal{E}_{|\sigma|,1}]$.[4]

Eq. (5.4.6d) can be seen as a transformation of $\mathcal{E}_{\sigma,1}$ from the domain of the wavenumber σ to the domain of the OPD δ. The user can obtain information on the spectral signature of the input irradiance $\mathcal{E}_{\sigma,1}$ by analyzing the detected irradiance at the output of the interferometer for a different (ideally continuous) set of OPDs. The obtained profile $\mathcal{E}$ as function of the OPD δ is known in the literature as **interferogram** and it typically (but not necessarily) features oscillations, known as **fringes**, around a central value.

[4] In the literature, the result of a Fourier transform is often referred as spectrum, but we choose not to employ this convention to avoid confusion, as in the domain of the interferometry the input of the Fourier transform itself is a description of a spectrum.

The devices that are able to record an interferogram, by varying the OPD are known as **Fourier transform spectrometers (FTSs)**, and, broadly speaking, they are aimed at the acquisition of a Fourier transformation of the input spectrum.

A multitude of techniques are available to generate different OPLs [117], but the most straightforward example is the **Michelson interferometer**. In this device, the input signal is divided into two arms through a **beam splitter**, usually manufactured with a half-silvered mirror to make its faces partially reflective. The split rays are then reflected by a series of mirrors in order to recombine them along the path that leads to the detector. By adjusting the distance of one (or more) mirrors with respect to the detector, one can adjust the OPD between the two rays, and perform multiple shot acquisitions to obtain the expression of $\mathcal{E}(\delta)$ for different values of δ.

5.4.4 Fabry-Pérot interferometry

The **FP interferometer**, also known as **FP etalon**, is an optical cavity made of two parallel reflecting surfaces, that allows optical rays to interfere by making them resonate within. The first prototype was developed by Charly Fabry and Alfred Pérot in 1899 [185].

To mathematically characterize a FP interferometer, let us assume we have a slab of transparent material bounded by two parallel reflective faces, which ideally extend infinitely in the space and are separated by a distance L. The faces act as a boundary for a certain medium, known as **optical cavity**, with refraction index n, sandwiched between two media with refraction index n_{0_1} and n_{0_2}.

Let us consider a monochromatic plane wave incident to the surface that forms an angle $\theta^{[in]}$ with the normal to the surface itself, characterized by a certain input irradiance $\mathcal{E}^{[in]}$; the associated incident plane can be defined as the one containing both the wavevector of the incident wave and the normal to the surface at the incidence point.

The main principle of operation of the device consists in letting the ray bounce multiple times within the etalon, so that a set of interfering beam emerges on the other side. To model this physical behaviour, we use the convention of the wave optics, by analyzing the intersection of the ray paths with all the discontinuity surfaces and by applying there the energy balance principle to obtain the resulting reflected and transferred rays.

With regard to the ray paths and with respect to the geometry of Fig. 5.4, we use the Snell's law (5.3.1b) to obtain the transmitted angle θ within the cavity and $\theta^{[t]}$ of the rays emerging on the other side, which yields:

$$n_{0_1} \sin\theta^{[in]} = n \sin\theta = n_{0_2} \sin\theta^{[t]} \,, \tag{5.4.7a}$$

$$\theta = \arcsin\left(\frac{n_{0_1}}{n} \sin\theta^{[in]}\right) \,, \tag{5.4.7b}$$

$$\theta^{[t]} = \arcsin\left(\frac{n_{0_1}}{n_{0_2}} \sin\theta^{[in]}\right) \,. \tag{5.4.7c}$$

We assume that the the internal sides of the slab are characterized by reflective coefficients $\tilde{r}_1$ and $\tilde{r}_2$ and transmission coefficients $\tilde{t}_1$ and $\tilde{t}_2$, while the transmission coefficient at the incidence point is $\tilde{t}_0$. As usual, one can characterize any component of the plane wave (either the electric or the magnetization field) by the means of the associated complex functions. If we assume that the complex function associated with the incident ray is $E^{[in]}$, we can define a series of emerging rays with complex function $\{E_m\}_{m\in\mathbb{N}}$, such that the k-th element of the set can be obtained, multiplying it by a certain factor α_m, which take into account all the attenuations that the m-th ray finds on its path, and a phase difference $2\pi\sigma\delta_m$, due to the ray travelling within the slab an OPL equal to δ_m, then:

$$E_m = \alpha_m E^{[in]} \exp(-j2\pi\sigma\delta_m) \,. \tag{5.4.8}$$

As multiple rays are involved in the interference, it is common practice to define the OPD of FP as the differences between the OPLs of two consecutive emerging rays, so that $\delta = \delta_{m+1} - \delta_m$. To evaluate its expression, let us consider a plane perpendicular to the direction of propagation of the emerging rays.

With reference to Fig. 5.4, the OPD δ is given by the difference between two contributions: the round trip OPL $\delta^{[rt]}$, depicted in green, travelled by a ray between two consecutive impacts over the same face and the projection $\delta^{[t]}$ of the distance between said impact points over the considered plane, depicted in red:

$$\delta = \delta^{[rt]} - \delta^{[t]} \tag{5.4.9a}$$

$$= n\frac{2L}{\cos\theta} - n_{0_2}(2L\tan\theta\sin\theta^{[t]}) \tag{5.4.9b}$$

$$= n\,(2L/\cos\theta - 2L\tan\theta\sin\theta) = 2nL\cos\theta \,, \tag{5.4.9c}$$

where Snell's law (5.4.7a) is used in eq. (5.4.9a).

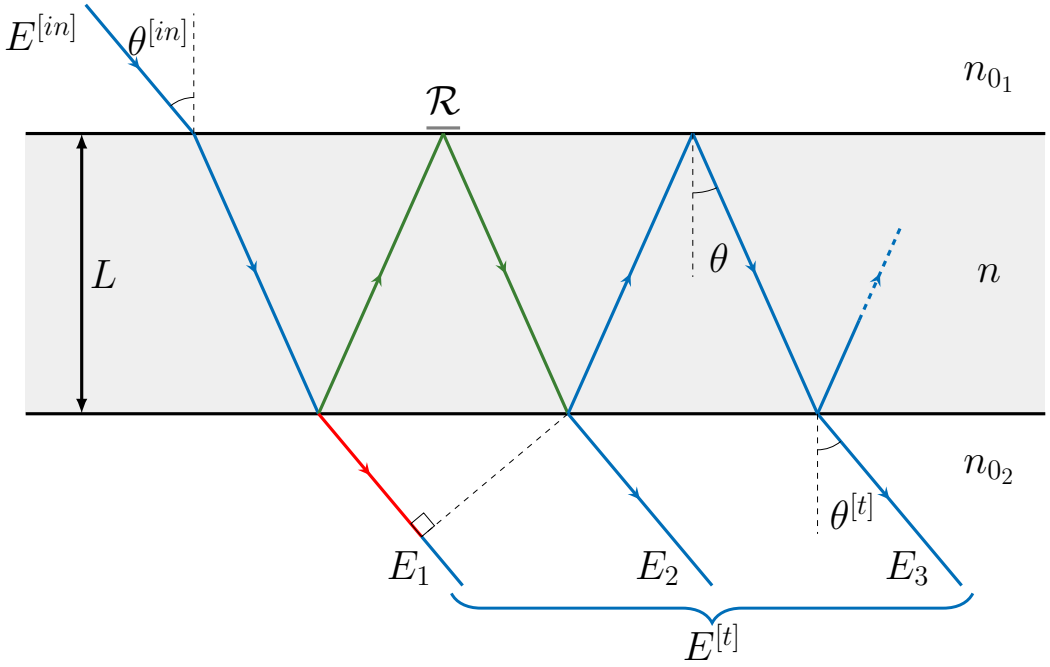


Fig. 5.4. Ray path tracing in a FP interferometer. Its OPD is obtained as a difference between the reflected path in green and the transmitted path in red.

With regard to the attenuation coefficients α_m, one can consider that the first emerging ray E_1 is obtained by two consecutive transmissions through both surfaces of the cavity. All subsequent waves have to be additionally reflected $2(m+1)$ times over the internal surfaces of the cavity, consequently, as $\alpha_{m+1} = \tilde{t}_0\tilde{t}_1(\tilde{r}_1\tilde{r}_2)^m$, the expression of each emerging wavefunction can be rewritten as:

$$E_{m+1} = \tilde{t}_0\tilde{t}_1(\tilde{r}_1\tilde{r}_2)^m E^{[in]} e^{-j\pi\delta_{rt}\sigma} e^{-j2m\pi\delta\sigma} \tag{5.4.10a}$$

where we have imposed for simplicity $\varphi = 2\pi\sigma\delta$, which represents the phase difference (as it has the same purpose of the homonym quantity presented in Section 5.4.2).

5.4.5 Wave transfer models

As shown in the previous section, one could define a series of subsequently attenuated rays emerging from a cavity generated from a single incident ray with incident angle $\theta^{[in]}$ and irradiance $\mathcal{E}_\sigma^{[in]} = |E^{[in]}|^2$.

According to the principle of superposition, the irradiance $\mathcal{E}_\sigma^{[t]}$ associated with the monochromatic beam of a given number N_m of emerging waves can be obtained as the irradiance associated with the sum of the complex functions of each emerging ray; in mathematical terms:

$$\mathcal{E}_\sigma^{[t]} = \left| \sum_{m=0}^{N_m-1} E_{m+1} \right|^2 = \left| \sum_{m=0}^{N_m-1} \alpha_{m+1} E^{[in]} e^{-j\pi\delta_{rt}\sigma} e^{-jm\varphi} \right|^2 \tag{5.4.11a}$$

$$= \left| \sum_{m=0}^{N_m-1} \left(\tilde{r}_1\tilde{r}_2 e^{-j\varphi}\right)^m \right|^2 |\tilde{t}_0\tilde{t}_1|^2 \left|E^{[in]}\right|^2 \tag{5.4.11b}$$

$$= \left| \sum_{m=0}^{N_m-1} \left(\mathcal{R}(\sigma) e^{-j\varphi}\right)^m \right|^2 \mathcal{T}^2(\sigma)\mathcal{E}_\sigma^{[in]} \tag{5.4.11c}$$

In the above expression, we defined $\mathcal{T}(\sigma) := \tilde{t}_0\tilde{t}_1$ and implicitly supposed here that the faces are lossless so one can impose $\mathcal{R}(\sigma) = |\tilde{r}_1\tilde{r}_2|$; indeed if the structure is symmetric from both directions, $\mathcal{R}(\sigma)$ represents the reflectivity of any of the two surfaces, as defined in Section 5.3.1. It is worth stressing here that, other than the explicit dependency by the wavenumber in the void σ, the terms $\mathcal{R}$, $\mathcal{T}$ and φ are also function of the incident angle $\theta^{[in]}$ through δ (according to eq. (5.3.4), (5.3.6) and (5.4.9)).

We aim now to define the **irradiance ratio** $\mathfrak{T}(\sigma) = \mathcal{E}_\sigma^{[t]}/\mathcal{E}_\sigma^{[in]}$ for various values of the amount of waves N_m.

- **2-wave model**: The FP interferometer is a specific instance of a FTS, as long as we consider the amplitude of the first emerging ray is much larger than the rest ($E_m \approx 0$ for $m \geq 2$). For this case, the expression of the irradiance ratio $\mathfrak{T}_2$ has the same structure of the 2-wave interferometry described in Section 5.4.2:

$$\mathfrak{T}_2(\sigma) = \left.\frac{\mathcal{E}_\sigma^{[t]}}{\mathcal{E}_\sigma^{[in]}}\right|_{N_m=2} = \left|1 + \mathcal{R}(\sigma)e^{-j\varphi}\right|^2 \mathcal{T}^2(\sigma) \tag{5.4.12a}$$

$$= \left|1 + \mathcal{R}(\sigma)\cos\varphi - j\mathcal{R}(\sigma)\sin\varphi\right|^2 \mathcal{T}^2(\sigma) \tag{5.4.12b}$$

$$= \left((1 + \mathcal{R}(\sigma)\cos\varphi)^2 + \mathcal{R}^2(\sigma)\sin^2\varphi\right)\mathcal{T}^2(\sigma) \tag{5.4.12c}$$

$$= \left(1 + \mathcal{R}^2(\sigma) + 2\mathcal{R}(\sigma)\cos\varphi\right)\mathcal{T}^2(\sigma)\,. \tag{5.4.12d}$$

- N_m**-wave model**: If the 2-wave condition is not verified, let N_m be a fixed amount of emerging rays with non negligible energy, then the irradiance ratio $\mathfrak{T}_{N_m}$ can be expressed in closed form, as eq. (5.4.11c) is a geometric series:

$$\mathfrak{T}_{N_m}(\sigma) = \frac{\mathcal{E}_\sigma^{[t]}}{\mathcal{E}_\sigma^{[in]}} = \left| \frac{1 - \mathcal{R}^{N_m}(\sigma) e^{-jN_m\varphi}}{1 - \mathcal{R}(\sigma) e^{-j\varphi}} \right|^2 \mathcal{T}^2(\sigma) \tag{5.4.13a}$$

$$= \frac{1 + \mathcal{R}^{2N_m}(\sigma) - 2\mathcal{R}^{N_m}(\sigma)\cos(N_m\varphi)}{1 + \mathcal{R}^2(\sigma) - 2\mathcal{R}(\sigma)\cos\varphi} \mathcal{T}^2(\sigma) \tag{5.4.13b}$$

- ∞**-wave model**: In the ∞-wave case ($N_m \to \infty$), results in the following transfer function, known as **Airy's distribution**:

$$\mathfrak{T}_\infty(\sigma) = \left. \frac{\mathcal{E}_\sigma^{[t]}}{\mathcal{E}_\sigma^{[in]}} \right|_{N_m \to \infty} = \frac{1}{1 + \mathcal{R}^2(\sigma) - 2\mathcal{R}(\sigma)\cos\varphi} \mathcal{T}^2(\sigma) \tag{5.4.14a}$$

$$= \frac{1}{(1 - 2\mathcal{R}(\sigma) + \mathcal{R}^2(\sigma)) - 2\mathcal{R}(\sigma)(\cos\varphi - 1)} \mathcal{T}^2(\sigma) \tag{5.4.14b}$$

$$= \frac{1}{(1 - \mathcal{R}(\sigma))^2 - 4\mathcal{R}(\sigma)\sin^2(\varphi/2)} \mathcal{T}^2(\sigma) \tag{5.4.14c}$$

$$= \frac{1}{1 + \frac{4\mathcal{R}(\sigma)}{(1-\mathcal{R}(\sigma))^2}\sin^2(\varphi/2)} \left(\frac{\mathcal{T}^2(\sigma)}{(1 - \mathcal{R}(\sigma))^2} \right) \tag{5.4.14d}$$

$$= \frac{\mathfrak{T}_\infty(\sigma)|_{\delta=0}}{1 + \frac{4\mathcal{R}(\sigma)}{(1-\mathcal{R}(\sigma))^2}\sin^2(\varphi/2)}\,, \tag{5.4.14e}$$

  where we have denoted with $\mathfrak{T}_\infty(\sigma)|_{\delta=0}$ the value of the transfer function of an equivalent cavity with OPD equal to zero.

If we suppose that the incident wave is nonmonochromatic, but still coherent, one can obtain the total output irradiance $\mathcal{E}^{[t]}$ by integrating the spectral irradiance $\mathcal{E}_\sigma^{[t]}$ over the wavenumber in the vacuum σ:

$$\mathcal{E}^{[t]} = \int_0^{+\infty} \mathcal{E}_\sigma^{[t]} d\sigma = \int_0^{+\infty} \mathcal{E}_\sigma^{[t]} d\sigma = \int_0^{+\infty} \mathfrak{T}(\sigma)\mathcal{E}_\sigma^{[in]} d\sigma\,, \tag{5.4.15}$$

where $\mathfrak{T}(\sigma)$ can be any of the expressions from eq. (5.4.12), (5.4.13) or (5.4.14).

As described in Section 5.4.2, if $\mathcal{R}$ and $\mathcal{T}$ are constant with σ, the choice of $\mathfrak{T}_2(\sigma)$ describes, if we exclude the bias and the multiplication coefficient, a cosine Fourier transform from the domain of the wavenumber in vacuum σ to the domain of the OPD δ.

5.4.6 Filtering effect of Fabry-Pérot interferometers

The irradiance ratio $\mathfrak{T}_\infty(\sigma)$ in 5.4.14 shows that FP interferometers act as filter for the incident spectral irradiance, with a periodic response in σ. It is useful to describe some characteristics of this filtering behaviour, in terms of some intrinsic parameters of the FP etalon.

Two particular quantities are of interest [5]:

- **Free spectral range (FSR)** $\Delta\sigma_{FSR}$: it defines the spacing in wavenumbers between two successive transmitted optical intensity maxima or minima of the interferometer. A quick analysis of eq. (5.4.14) shows that $\mathfrak{T}_\infty(\sigma)$ exhibits a periodicity of π in terms of $\varphi = 2\pi\delta\sigma$. If we rewrite it as:

$$\mathfrak{T}_\infty(\sigma) = \frac{\mathfrak{T}_\infty(\sigma)|_{\delta=0}}{1 + \frac{4\mathcal{R}(\sigma)}{(1-\mathcal{R}(\sigma))^2}\sin^2\left(\pi\frac{\sigma}{\Delta\sigma_{FSR}}\right)}\,, \tag{5.4.16}$$

  we can immediately obtain that in terms of σ, the periodicity $\Delta\sigma_{FSR} = 1/\delta$ is the reciprocal of the OPD.
- **Full width at half maximum (FWHM)** $\Delta\sigma_{FWHM}$: it defines the difference between the two extreme values of the wavenumber at which the intensity of the superposed wave is equal to half of its maximum value. This condition is equivalent to find the difference between two consecutive values of the frequency $\sigma_{1,2}$ around a peak for which $\mathfrak{T}_\infty(\sigma) = \mathfrak{T}(\sigma)|_{\delta=0}/2$. This is possible only for sufficiently big values of the reflectivity ($\mathcal{R} > 0.172$) and from eq. (5.4.16) we obtain:

$$\frac{4\mathcal{R}(\sigma)}{(1-\mathcal{R}(\sigma))^2}\sin^2\left(\pi\frac{\sigma_{1,2}}{\Delta\sigma_{FSR}}\right) = 1\,, \tag{5.4.17a}$$

$$\Delta\sigma_{FWHM} := |\sigma_2 - \sigma_1| = \frac{2}{\pi}\Delta\sigma_{FSR}\arcsin\left(\frac{1-\mathcal{R}}{2\sqrt{\mathcal{R}}}\right)\,. \tag{5.4.17b}$$

[5] Those quantity are typically defined in terms of optical frequency ν, although to simplify the exposition, they are equivalently defined in terms of wavenumbers $\sigma = \nu/c_0$.

In most works, it is customary to characterize the etalon with a single parameter known as **finesse** $\mathscr{F}$, defined as the ratio between $\Delta\sigma_{FSR}$ and $\Delta\sigma_{FWHM}$:

$$\mathscr{F} := \frac{\Delta\sigma_{FSR}}{\Delta\sigma_{FWHM}} = \frac{\pi}{2\arcsin\left(\frac{1-\mathcal{R}}{2\sqrt{\mathcal{R}}}\right)} \tag{5.4.18a}$$

$$\approx \frac{\pi\sqrt{\mathcal{R}}}{1-\mathcal{R}}\,, \tag{5.4.18b}$$

where the approximation (5.4.18b) is just valid for high values of the parameter $\mathcal{R}$ (typically $\mathcal{R} > 0.5$). The value of the finesse influences the shape and the filtering power of the transfer function, as shown in Fig. 5.5a. For devices characterized by:

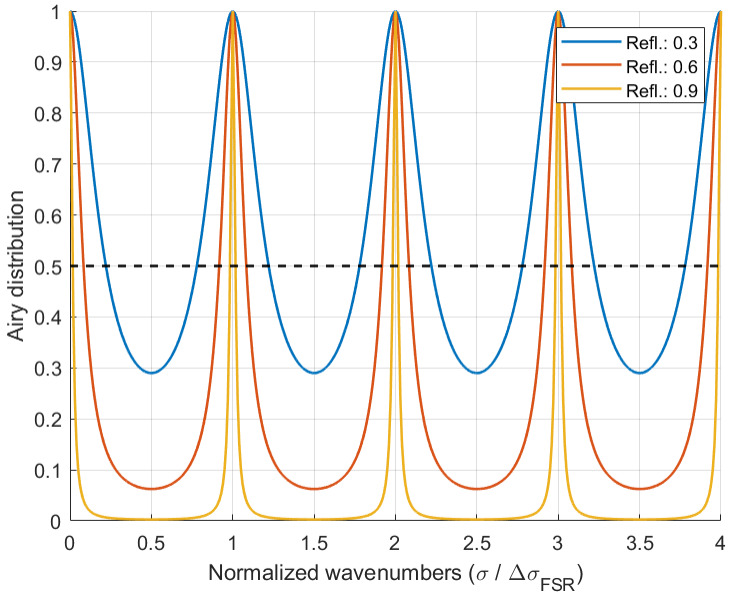
- **high finesse**: or in other words high reflectivity coefficients, the transfer function shows a very narrow interferometric fringes, and they act as bandpass filters;
- **low finesse**: the transfer function is approximately sinusoidal in nature, and its behaviour is closer to a standard FTS with more energy is available for the photodetectors to collect.

The effect of the finesse can also be seen in Fig. 5.5b, which shows the spatial response of an interferometer with diffused incident light. The center of each subimage shows the response for null polar incidence angle (perpendicular incidence to the input plane), which linearly increases proportionally to the distance from the center of each subimage. A quick visual analysis shows that the acquisitions with interferometers characterized by high finesse appear noticeably sharper than their low finesse counterparts and the density of fringes within the same FoV increases for thicker interferometers, as $\Delta\sigma_{FSR}$ is inversely proportional with the OPD.

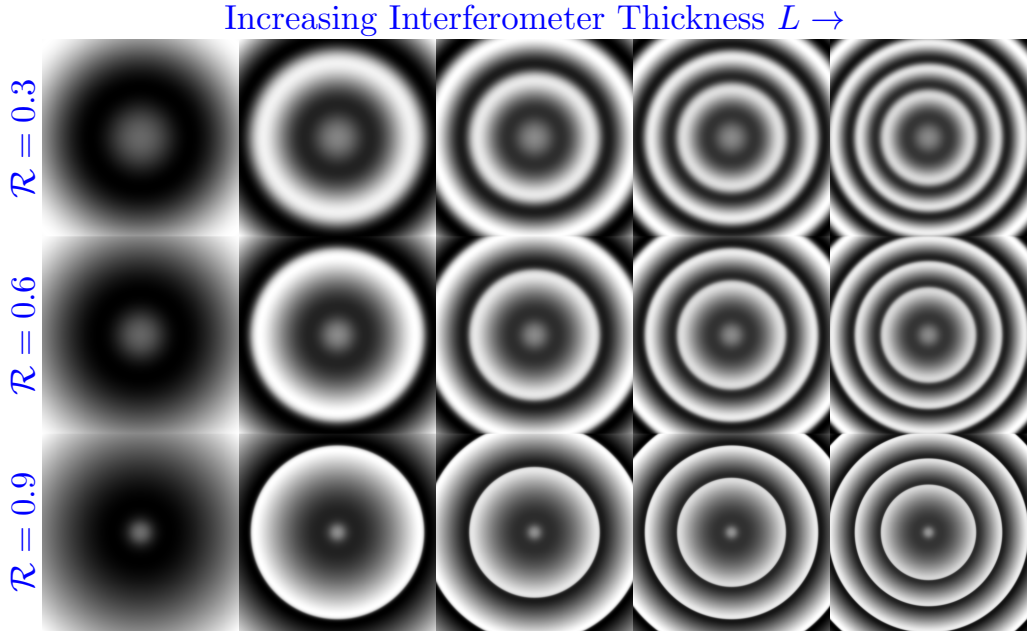
The value of $\mathcal{R}$ is not fixed by the nature of the media which surround the discontinuity, as the expressions derived in Section 5.3.1 seem to imply, as we have a degree of control over its expression by adding a coating over the surfaces.

5.5 The ImSPOC acquisition model

The main target of this section is to describe, mathematically characterize, and quantify the physical transformations of the incident radiance, that are involved in the generation of the acquisition with an ImSPOC device.



(a) Finesse



(b) Position of the fringes

Fig. 5.5. Representation of the interferometric fringes. In Fig. 5.5a, a representation of the transfer function for different values of reflectivity. The dashed line marks the half width bandwidth of the filter. In Fig. 5.5b, each subimage represents the interferogram obtained by illuminating an interferometer to a diffused white source. The center of each subimage is the answer for incident rays perpendicular to the interferometer, with progressively increasing incident angles moving towards the edges. Each row of the matrix is for a given finesse and each column for a given interferometer thickness.

5.5.1 Context

In recent decades, the field of remote sensing has become of the utmost importance for monitoring the Earth's surface (its lands, urbanisation, oceans, agriculture, etc.) and its atmosphere [220, 23, 2, 135]. In particular, the need for accurate measurements of gases in the atmosphere is ever increasing for tasks as monitoring climate change and air quality study and regulation issues. These tasks require data acquisitions with ever higher spatial, spectral and temporal resolutions, motivating the development of new sensors and their corresponding signal processing methods. In modern days, different types of hyperspectral (HS) imaging systems, based on different techniques (e.g., spatial, snapshot or spectral scanning [28]), are available and are tailored to specific applications.

Monitoring atmospheric gases in the context of climate change requires increasingly accurate sensors and less spaced acquisitions, and HS imaging acquisitions specifically offer the necessary diversity of spectral information for applications such as atmospheric gas monitoring [135]. Most current HS imaging systems have to balance a trade-off between spectral, spatial and temporal resolution and new technologies aim to either overcome those limitations or make it more efficient in terms of price. An increasing interest is also shifting on the miniaturisation, especially with the purpose of installation over airborne vehicles or nano-satellites [194, 208]. The implication on costs' reduction may involve various fields: both the production, maintenance and usage costs will be amortized, and more load would be available for the platform to be mounted on.

This section focuses on the groundbreaking HS image acquisition technology based on a miniaturised static (snapshot) interferometric imaging spectrometer, called **ImSPOC** [104]. The main targeted applications of this device are in the Earth observation field, in particular to monitor gases in the atmosphere, but requires accurate signal processing developments to provide intelligible and well-calibrated acquisitions for the final users.

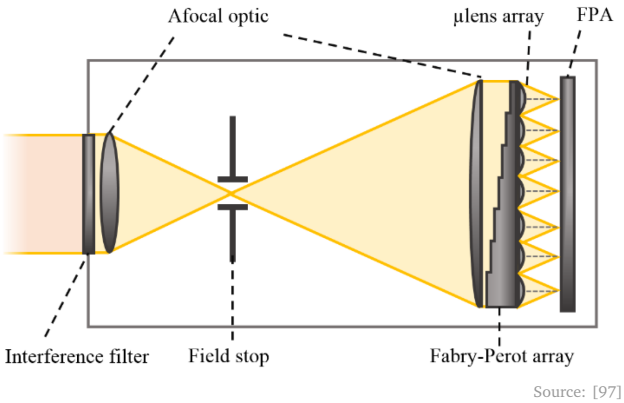
This design is aimed to the manufacture a competitive low-cost snapshot spectrometer with spectral resolution comparable to the conventional remote sensing systems for measuring gases with similar performance, while also providing directional information on the incoming radiance. The ImSPOC sensor poses a series of novel challenges; on the manufacturing side, which is outside the scope of this work, some of these issues involve the choice of the appropriate design, materials, layout, dimensioning, and packaging [96, 97, 98, 76]. We aim instead to describe here the

optical transformations that characterize the device, and investigate the task of the data processing of raw data to Chapter 6.

5.5.2 Physical structure

ImSPOC describes a concept for an innovating spectro-imaging system, whose patent was jointly presented by the Institut de Planétologie et d'Astrophysique de Grenoble (IPAG) and Office National d'Etudes et de Recherches Aérospatiales (ONERA) in 2016 and deposited in 2018 [104]. It is a miniaturized snapshot acquisition system for HS imagery, whose principle of operation is based on the interaction between a matrix of micro-lenses and a staircase-shaped optical plate, superposed to a focal plane. Each element (a step of the staircase pattern) in the optical plate forms a FP etalon and is associated with a specific micro-lens; this block is in charge of the acquisition of a single subimage over an assigned area of the focal plane.

In terms of its physical behaviour, the ImSPOC family of instruments can be considered as a spectrometer aimed at the measure of the incoming spectral radiance through simultaneous acquisitions, each of which is performed through interferometric measurements with FP etalons with different thicknesses. A visual representation of the device, complete with its leading optic, is shown in Fig. 5.6.



Source: [97]

Fig. 5.6. Optical concept of the ImSPOC device.

This allows the device to operate as a static snapshot detector, so that no complicated methodologies, such as a temporal scanning of the scene (e.g., with a classical spectrometer) or sequential variation of the OPD (e.g., in the case of standard Michelson interferometers), are needed.

In this chapter a fully conceptual approach, partially divorced from its technological realization, aimed at providing the mathematical tools to characterize the operations

of the device. To this end, the ImSPOC concept will be simply intended as an array of N_k adjacent elements, each composed by a stacked FP interferometer with an associated lens, superimposed to a matrix of photodetectors.

The raw product of ImSPOC is a monochromatic acquisition with a series of N_k small images next to one another, which we call **subimages** in the following, each associated with the response of one of the interferometers, as shown in Fig. 6.4d.

5.5.3 Description of the ray transfer function

In this thesis we employ the approach to characterize the optical system through the formalism of the **ray transfer matrix** described in Section 5.3.2; the full description of all aberration introduced the device may require a full **ray tracing** procedure [91], which is outside the scope of this work.

The ray transfer method describes each ray by its direction in terms of spherical coordinates, and its intersection with a given plane in terms Cartesian coordinates. So, i.e., the incident ray is represented by the quadruplets of coordinates $\left[\boldsymbol{\theta}^{[in]}; \mathbf{r}^{[in]}\right]$, where $\boldsymbol{\theta}^{[in]} = \left[\theta^{[in]}; \phi^{[in]}\right]$ and $\mathbf{r}^{[in]} = \left[r_1^{[in]}; r_2^{[in]}\right]$ denote said direction and intersection, respectively.

The ray path is traced across different parallel planes, and its evolution between two consecutive planes is described with an assigned transfer function. The planes that are considered in our context are:

- **Input plane**: The plane of the incoming radiance to measure, which in our case is at the back plane of the interferometers;
- **Principal plane(s)**: The hypothetical plane(s) of the lens where all refraction are supposed to happen;
- **Focal plane**: The hypothetical plane where all parallel incident on the lens converge, which in our context is at the plane of the photodetectors.

An example of the placement of the planes is shown in Fig. 5.7.

The framework of Section 5.3.2 defines a ray transfer matrix which maps each incident ray to an associated transmitted ray. However, in the ImSPOC concept, multiple replicas of a single ray are generated by the internal reflections within the FP etalon. To model this behaviour, we propose to define a series of ray transfer functions $\{\Theta_{km}\}_{\substack{k\in[1,\dots,N_k] \\ m\in[1,\dots,N_m]}}$, so that Θ_{km} is associated with the m-th replica generated within the k-th interferometer.

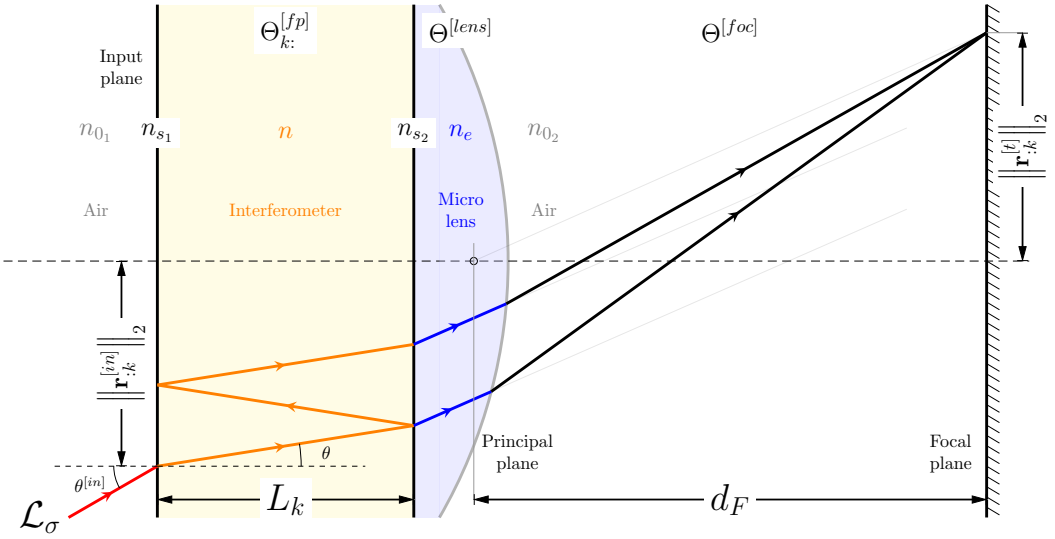


Fig. 5.7. Cross section of the ray path in an element of the ImSPOC taken along the plane of incidence. The figure shows how the directly transmitted ray and its replica generated by an internal reflection within the interferometer (in orange) are both focused on the same spot on the focal plane, under idealized operating conditions.

The ray transfer function Θ_{km} is modeled as a cascade of three transformations:

- $\Theta^{[fp]}_{km}$, which models to the m-th emerging ray within the k-th interferometer;
- $\Theta^{[lens]}_{k}$, which models the deflection of the ray due to the k-th microlens;
- $\Theta^{[foc]}$, which models the free propagation in the space between the principal plane the and the focal plane, whose distance is denoted by d_F.

The description $\left[\boldsymbol{\theta}^{[t]}_{km}, \mathbf{r}^{[t]}_{:km}\right]$ of the m-th replica associated with the k-th interferometer is then given by:

$$\begin{bmatrix}\boldsymbol{\theta}^{[t]}_{km}\\ \mathbf{r}^{[t]}_{:km}\end{bmatrix} = \Theta_{km}\left(\begin{bmatrix}\boldsymbol{\theta}^{[in]}\\ \mathbf{r}^{[in]}_{:k}\end{bmatrix}\right) = \Theta^{[foc]}\left(\Theta^{[lens]}_{k}\left(\Theta^{[fp]}_{km}\left(\begin{bmatrix}\boldsymbol{\theta}^{[in]}\\ \mathbf{r}^{[in]}_{:k}\end{bmatrix}\right)\right)\right), \tag{5.5.1}$$

where we have denoted:

$$\mathbf{r}^{[in]}_{:k} = \mathbf{r}^{[in]} - \mathbf{r}^{[0]}_{:k} \tag{5.5.2}$$

the vector difference $\mathbf{r}^{[in]}_{:k}$ between $\mathbf{r}^{[in]}$ and the position $\mathbf{r}^{[0]}_{:k}$ of the optical axis of the k-th lens with the input plane. A visual representation of the system of coordinates is shown in Fig. 5.8.

The above equation intrinsically assumes that if one ray is transmitted through the FP cavity, it is also fully transmitted through the lens surface, so that no extra replicas are generated within the system. Under idealized conditions that the k-th interferometer is a homogeneous medium with refraction index n and with thickness

L_k bounded by parallel faces between two homogeneous media with refraction indices n_{0_1} and n_{0_2}, and supposing that the paraxial approximation holds, the transfer function $\Theta_{mk}^{[fp]}$ can be expressed as:

$$\Theta_{km}^{[fp]}\left(\begin{bmatrix}\theta\\ \phi\\ r_1\\ r_2\end{bmatrix}\right) \approx \begin{bmatrix}\frac{n_{0_1}}{n_{0_2}}\theta\\ \phi\\ r_1+\frac{n_{0_1}}{n}\theta(2m-1)L_k\cos\phi\\ r_2+\frac{n_{0_1}}{n}\theta(2m-1)L_k\sin\phi\end{bmatrix}, \tag{5.5.3}$$

where the expression for the polar angles a consequence of Snell's law, and the new position is updated considering that the the m-th ray travels a total of $2m+1$ half trips within the cavity before emerging on the other side.

A similar analysis can be applied for the ray transfer function describing the last leg of the transformation, which is equivalent to a direct transmission across a single homogeneous layer with thickness d_F, so that:

$$\Theta^{[foc]}\left(\begin{bmatrix}\theta\\ \phi\\ r_1\\ r_2\end{bmatrix}\right) = \begin{bmatrix}\theta\\ \phi\\ r_1+(d_F\cos\phi)\,\theta\\ r_2+(d_F\sin\phi)\,\theta\end{bmatrix}. \tag{5.5.4}$$

The expression for the lens matrix depends on its peculiar geometry and can be characterized in the calibration phase, however, in the interest of providing at least an idealized description for a thin spherical lens with focal distance d'_F, typically chosen to be equal to d_F, the ray transfer function is equal to:

$$\Theta^{[lens]}\left(\begin{bmatrix}\theta\\ \phi\\ r_1\\ r_2\end{bmatrix}\right) \approx \begin{bmatrix}\theta-\frac{\sqrt{r_1^2+r_2^2}}{d'_F}\\ \phi\\ r_1\\ r_2\end{bmatrix}. \tag{5.5.5}$$

5.5.4 Acquisition model

The target of this section is to mathematically describe the optical processes that model the raw acquisition, described by the electrical charges collected by the readout circuitry, characterizing the sensor matrix.

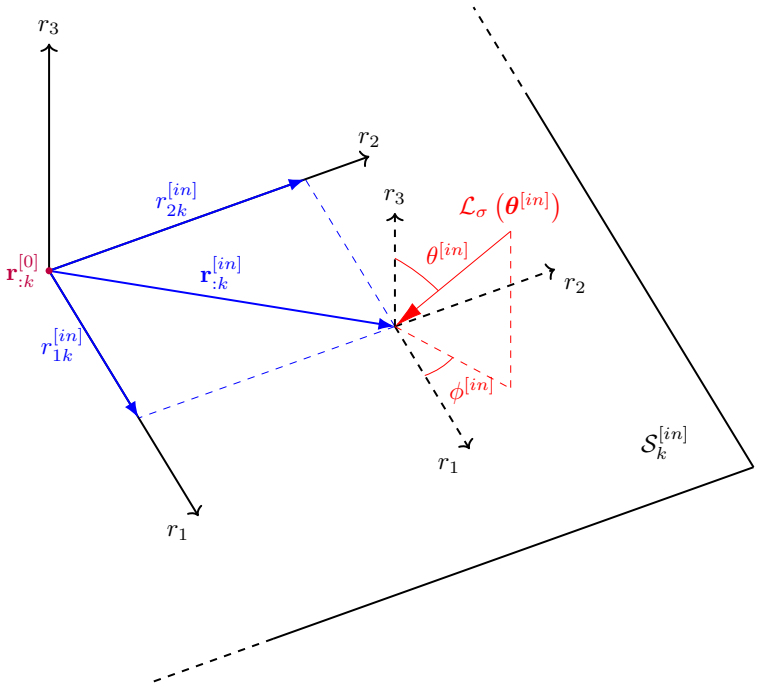


Fig. 5.8. Geometrical system of coordinate for the an incident ray over the input plane surface of the k-th interferometer. The coordinates relative to the position of the intersection are highlighted in blue, while those relative to the angle of incidence are highlighted in red.

This acquisition is a stochastic process in nature, since various sources of uncertainty are introduced by the technical characteristics of the photodetectors, whose noise characterization was introduced in Section 5.2.4. Photodetector noise is notoriously difficult to separate from the relevant information, but it may be possible to compensate some other sources of uncertainty, such as the **speckle noise**. The latter describes the effects of oscillations of wavefronts with different phases in the path between the detector and the receiver, which combine to generate a wave whose intensity varies randomly. This effect is deterministic nature in nature for a fixed configuration, and can be compensated with either multi-look analysis or reconstructions with wavelet regularization [81].

The analysis of this work assumes that all noise contributions are zero-mean, additive and concentrated to the last node of the transfer model, making it possible to separate the stochastic process into two components: a deterministic model, which interpreted as the expected value of the stochastic process, and a random component. If the signal to noise ratio (SNR) of the captured energy is above a certain threshold, shot noise can be reasonably assumed to be a Gaussian process, justifying this assumption.

Under these premises, the main target becomes to find a mathematical relation, quantifying the contribution of the incident spectral radiance $\mathcal{L}_\sigma\left(\left[\boldsymbol{\theta}^{[in]};\mathbf{r}^{[in]}_{:k}\right]\right)$ incident over the k-th interferometer for the acquisition y_o over the generic o-th detector.

According to the definition of radiometric measures of Section 5.2.3, the intensity level of the acquisition may be simply obtained by integrating the radiance at the focal plane $\mathcal{L}^{[t]}_\sigma([\boldsymbol{\theta}^{[t]};\mathbf{r}^{[t]}])$. The is a function of three variables:

- the incidence angle $\boldsymbol{\theta}^{[t]}$, spanning over an emisphere Ω;
- the coordinates $\mathbf{r}^{[t]}$ of the intersection with the focal plane, which span over the surface $\mathcal{S}^{[t]}_o$ of the o-th sensor that is sensitive to the light stimulus;
- the wavenumber σ, spanning over the range of the instrument $[\sigma_{min},\sigma_{max}]$.

We model the detected intensity level y_o of the o-th detector with the integration:

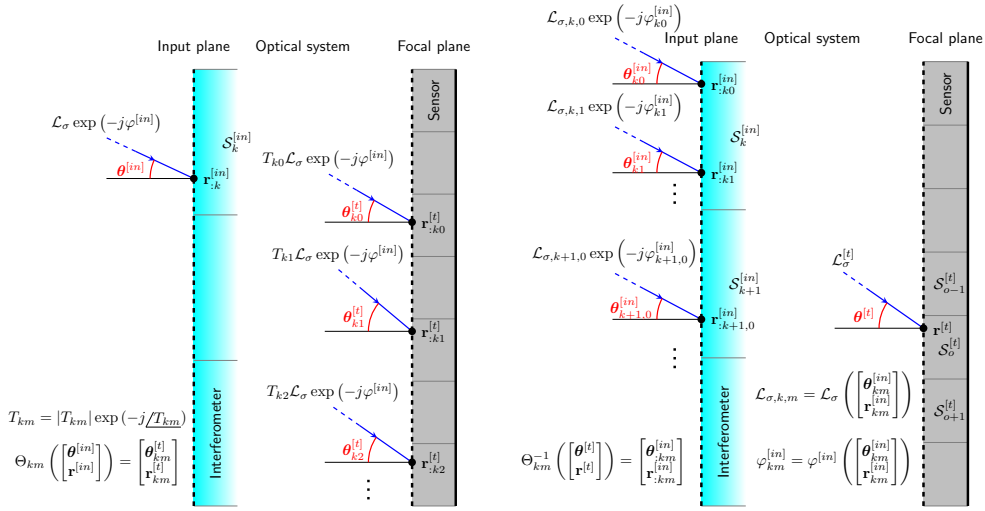
$$y_o = \iint\limits_{\mathcal{S}^{[t]}_o} \iint\limits_{\Omega} \int\limits_{\sigma_{min}}^{\sigma_{max}} \mathcal{L}^{[t]}_\sigma\left(\left[\boldsymbol{\theta}^{[t]};\mathbf{r}^{[t]}\right]\right)\cos\theta^{[t]}\eta_o(\sigma)\,d\sigma\,d\mathbf{r}^{[t]}\,d\boldsymbol{\theta}^{[t]}\,, \tag{5.5.6}$$

where we also defined $\eta_o(\sigma)$ as the quantum efficiency of the o-th sensor to transform the illuminating photons into a transfer of electrons. The term $\cos\theta^{[t]}$ is necessary to take into account that the flux of photons is not necessarily perpendicular to the surface of the photo-detector.

Unfortunately, eq. (5.5.6), the radiant energy collected on the sensor may bear no resemblance to the incident radiance $\mathcal{L}_\sigma$, as their relationship is given by the characteristics of the optical system. In the case of classical cameras, the ray transfer matrix analysis (Section 5.3.2) allows a one-to-one association between the direction of an incident ray at the level of the input and the focal plane. In our case, however, as shown in fig. 5.9a, a single incident ray generates a set of emerging rays. This set is countable, as each of them is associated to a certain amount m of round trips within the interferometer. Consequently, the relationship between the direction and position $[\mathbf{r}^{[in]}_{:k},\boldsymbol{\theta}^{[in]}]$ of the incident radiance and the corresponding ones $[\mathbf{r}^{[t]}_{km},\boldsymbol{\theta}^{[t]}]$ of the m-th replica on the focal plane can be described analytically by a set of transfer functions Θ_{km}, defined as follows:

$$\Theta_{km}\left(\begin{bmatrix}\boldsymbol{\theta}^{[in]}\\ \mathbf{r}^{[in]}_{:k}\end{bmatrix}\right) = \begin{bmatrix}\boldsymbol{\theta}^{[t]}_{km}\\ \mathbf{r}^{[t]}_{km}\end{bmatrix}\,, \tag{5.5.7}$$

which may additionally be expressed as a function of the wavenumber σ to take into account the effects of spectral aberrations. For our purposes, however, we are mostly



(a) Incident ray perspective **(b)** Receiver perspective

Fig. 5.9. Ray tracing seen from the perspective of an incident ray and from the perspective of the receiver. In the left figure, the ray crosses the surface of he k-th interferometer and it is split into a set of rays that impact the focal plane in different positions; the m-th emerging ray is attenuated by a complex factor T_{km}. In the right figure, the ray that is absorbed by the sensor is traced backwards, identifying a countable set of incident rays that combine together to generate the collected radiance $\mathcal{L}^{[t]}$.

interested in the so-called **backward ray tracing**, that is expressing a generic couple position/direction $[\mathbf{r}^{[t]}, \boldsymbol{\theta}^{[t]}$ in terms of the corresponding set of the generating rays incident to the input plane, as shown in fig. 5.9b. It is hence convenient to define an inverse relationship:

$$\Theta_{km}^{-1}\left(\begin{bmatrix}\boldsymbol{\theta}^{[t]}\\ \mathbf{r}^{[t]}\end{bmatrix}\right) = \begin{bmatrix}\boldsymbol{\theta}^{[t]}_{:km}\\ \mathbf{r}^{[t]}_{:km}\end{bmatrix} \Leftrightarrow \Theta_{km}\left(\begin{bmatrix}\boldsymbol{\theta}^{[in]}_{:km}\\ \mathbf{r}^{[in]}_{:km}\end{bmatrix}\right) = \begin{bmatrix}\boldsymbol{\theta}^{[t]}\\ \mathbf{r}^{[t]}\end{bmatrix}. \tag{5.5.8}$$

It is worth noting that not all values necessarily exist for all available indices k and m, but this can be easily fixed by imposing that the transfer function is zero for

The set of rays we identify with this process is generally incoherent under the current hypothesis, so to completely characterize the input we need to define the following two component:

- **Incident radiance**, which defines our quantity of interest in terms of energy to detect:

$$\mathcal{L}_{\sigma,k,m} := \mathcal{L}_{\sigma}\left(\left[\boldsymbol{\theta}^{[in]}_{:km}; \mathbf{r}^{[in]}_{:km}\right]\right); \tag{5.5.9}$$

- **Phase component**, which in conjunction with the radiance, allows to fully define the complex amplitude of the incoming wave:

$$\varphi_{\sigma,k,m} := \varphi^{[in]}\left(\left[\boldsymbol{\theta}^{[in]}_{:km}; \mathbf{r}^{[in]}_{:km}\right], \sigma\right) . \tag{5.5.10}$$

The term $\mathcal{L}^{[t]}_{\sigma}$ from eq. (5.5.6) can be finally expressed as:

$$\mathcal{L}^{[t]}_{\sigma}\left(\left[\mathbf{r}^{[t]}; \boldsymbol{\theta}^{[t]}\right]\right) = \left| \sum_{k=0}^{N_k} \sum_{m=0}^{+\infty} T_{km}\left(\left[\mathbf{r}^{[in]}_{:km}, \boldsymbol{\theta}^{[in]}_{:km}\right], \sigma\right) \mathcal{L}_{\sigma,k,m} \exp\left(-j\varphi^{[in]}_{\sigma,k,m}\right) \right| . \tag{5.5.11}$$

In the above equation the term $T_{km}\left(\left[\boldsymbol{\theta}^{[in]}; \mathbf{r}^{[in]}_{:k}\right], \sigma\right)$ is the **attenuation factor** due to the path traveled by m-th replica of the incident ray characterized by the pair $\left[\boldsymbol{\theta}^{[in]}; \mathbf{r}^{[in]}_{:k}\right]$; the attenuation factor is generally complex (in other words in the form $T_{km} = |T_{km}| \exp\left(-j\underline{/T_{km}}\right)$), to take into account for the path crossed by the ray within the optical system, as well as some effects of phase shift due to bouncing elements within the device. The expression (5.5.11) assumes that, for every interferometer, it is always possible to identify a certain ray path that, after any number m of round trips, emerges in the direction under study; this condition is not realistic in practice, as typically a certain area of the focal plane is just assigned to a single interferometer, but this behaviour can be treated mathematically by setting T_{km} equal to zero for all the indices k which are not concerned with the subimage area on the focal plane.

Eq. (5.5.6) is able to justify the behaviour of many nonidealities of the device, which include, but are not limited to:

- non parallel faces manufacturing of the interferometers;
- not perfectly spherical lenses;
- unverified hypothesis of long range source;
- spectral aberration effects;
- roughness of the surfaces;
- aberration effects of the structure;
- inhomogeneous spectral response of the reflective surfaces of the interferometer;
- cross talk between subimages belonging to different interferometers.

However, some effects are still not included:

- parasite effect of an interferometer with thickness d_F in the free space between the lens and the focal plane array (FPA);

- exposure time, which is influenced by the photonic phenomena within the photodetector.

5.5.5 Far field approximation

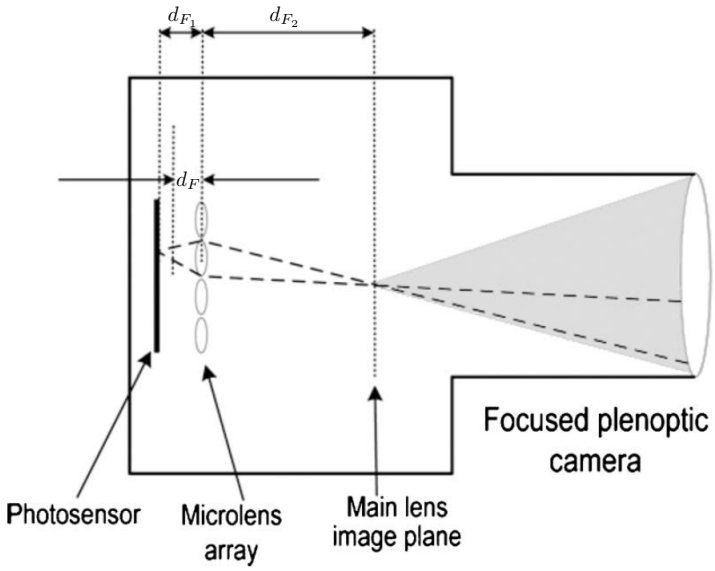
In this section, we want to define some assumption that allow to simplify the transfer model (5.5.6). We specifically assume:

- **Far field approximation**: if the target is sufficiently distant (technically more distant than the Fraunhofer distance), the whole spectrometer can be seen as a point-target with regard to the scene, so any displacement of the intersection spot on the device can be ignored, with respect to the distance of the target. Under this assumption, the incident radiance $\mathcal{L}_\sigma$ can be seen as independent from the intersection point and a given radiator on the scene would have the same incident angle across the whole input plane of the instrument.

  In terms of the description of the incident field of Section 5.5.4, this is equivalent to impose that:

$$\begin{cases} \mathcal{L}_{\sigma,k,m} &= \mathcal{L}_\sigma\left(\left[\boldsymbol{\theta}^{[in]}_{:km}, \mathbf{r}^{[in]}_{:km}\right]\right) = \mathcal{L}_\sigma\left(\boldsymbol{\theta}^{[in]}_{:km}\right), \\ \varphi^{[in]}_{\sigma,k,m} &= \varphi^{[in]}\left(\left[\boldsymbol{\theta}^{[in]}_{:km}, \mathbf{r}^{[in]}_{:km}\right], \sigma\right) = \varphi^{[in]}\left(\boldsymbol{\theta}^{[in]}_{:km}, \sigma\right), \end{cases} \quad \forall \begin{cases} \mathbf{r}^{[in]}_{:km} \in \mathcal{S}^{[in]}_k, \\ k \in [1, \dots, N_k]. \end{cases} \tag{5.5.12}$$

  This hypothesis, when not verified, can be sometimes exploited to achieve certain desired purposes; i.e., the **plenoptic camera** [87, 153], whose conceptual design is shown in Fig. 5.10, shows many similarities with the ImSPOC concept, although the former does not feature any interposed array of interferometers. A plenoptic campera can be used to estimate the **depth map** of the scene, by analyzing the parallax effect of the scene's geometry. If the ImSPOC employs this mode of operation, the subimages would be affected by an interferometric modulations, so one possible strategy could be to decompose the reconstructed hyper-cube into a panchromatic (PAN) component, which is common to all subimages, and a specific modulation, characteristic of each subimage [210]. In this case the inversion problem could be considered as a joint problem of spatial alignment and superresolution [73], which can be solved by dedicated optimization protocols [177].



Source: [153]

Fig. 5.10. Design of the plenoptic camera 2.0.

- **No cross-talk effect**: For this assumption a subset of detectors is assigned exclusively for the detection of rays emerging by a single interferometer. In other terms, we define a subset of detectors $\mathcal{Q}_k$ with cardinality $|\mathcal{Q}_k|$ associated with the k-th element of the array, such that:

$$o \in \mathcal{Q}_k \Leftrightarrow \exists\, \boldsymbol{\theta}^{[in]} \in \Omega : \mathbf{r}^{[t]}_{:km}\left(\boldsymbol{\theta}^{[in]}\right) \in \mathcal{S}^{[t]}_o\,. \tag{5.5.13}$$

  The o-th sensor can then be uniquely associated to a pair of indices $[k, i]$, such that the i-th element of the set $\mathcal{Q}_k$ of detectors is solely associated to the k-th subimage. From now on we hence redefine the variables which are intrinsically associated to a detector with the new convention:

$$y_{ki} := y_o\,, \quad \mathcal{S}^{[t]}_{ki} := \mathcal{S}^{[t]}_o\,, \quad \eta_{ki} := \eta_o\,. \tag{5.5.14}$$

- **Common focal spot**: which defines the condition that all the parallel emerging rays for a given interferometer focus on the same spot on the focal plane. We previously expressed the focal spot $\mathbf{r}^{[t]}_{:km}$ associated with the m-th replica generated by the k-th element as the second component of the ray transfer relationship $\Theta_{km}\left(\left[\boldsymbol{\theta}^{[in]}, \mathbf{r}_{:k}\right]\right)$, which is a function of the direction $\boldsymbol{\theta}^{[in]}$ and position $\mathbf{r}_{:k}$ of a given input ray. The condition is then mathematically equiva-

lent to suppose that this second component shows no dependency with the position of the input ray $\mathbf{r}_{:k}$. In other words:

$$\mathbf{r}^{[t]}_{:km}\left(\begin{bmatrix}\boldsymbol{\theta}^{[in]}\\ \mathbf{r}^{[in]}_{:k}\end{bmatrix},\sigma\right)=\mathbf{r}^{[t]}_{:k0}\left(\boldsymbol{\theta}^{[in]}\right),\qquad \forall\begin{cases}\mathbf{r}^{[in]}_{:k}\in\mathcal{S}^{[in]}_{k}\,, & k\in[1,\dots,N_k]\,,\\ m\in\mathbb{N}\,, & \sigma\in[\sigma_{min},\sigma_{max}]\,,\end{cases} \tag{5.5.15}$$

where we dropped the dependency from m because all replicas at the principal plane are parallel and from σ because we supposed that spectral aberrations are negligible. A visual representation of this effect is shown in Fig. 5.7. The above relation can also be seen as a 1-to-1 mapping between the incident angle $\boldsymbol{\theta}^{[in]}$ and its associated illuminated spot on the focal plane due to the k-th interferometer/lenslet couple. In physical terms, this condition can be obtained with a matrix of perfectly spherical thin lenses under the condition of paraxial approximation, as shown in Section 5.3.

When the three conditions are considered jointly, the FoV Ω can thus partitioned in a set of solid angles $\{\Omega_{ki}\}_{\substack{k\in[1,\dots,N_k]\\ i\in[1,\dots,|\mathcal{Q}_k|]}}$ such that:

$$\boldsymbol{\theta}^{[in]}\in\Omega_{ki}\Leftrightarrow\mathbf{r}_{:km}\left(\boldsymbol{\theta}^{[in]}\right)\in\mathcal{S}^{[t]}_{ki}\,. \tag{5.5.16}$$

In other terms, Ω_{ki} denotes a continuous set of angles of incidence, such that, if it intersect the input plane belonging to the k-th interferometer, it is focused on the i-th sensor of the set $\mathcal{Q}_k$.

We now want to rewrite the integration of the expression (5.5.11) for the transmitted radiance $\mathcal{L}_\sigma^{[t]}$ with respect to the angle of incidence, supposing the three conditions are jointly verified. The expression becomes as follows:

$$\iint_{\Omega} \mathcal{L}_\sigma^{[t]} \left(\left[\boldsymbol{\theta}^{[t]}; \mathbf{r}_{:k0}^{[t]} \right] \right) \cos\theta^{[t]} \, d\boldsymbol{\theta}^{[t]} \tag{5.5.17a}$$

$$= \iint_{\Omega} \left| \sum_{m=0}^{+\infty} T_{km} \left(\left[\boldsymbol{\theta}_{:km}^{[in]}; \mathbf{r}_{:km}^{[in]} \right], \sigma \right) \mathcal{L}_\sigma \left(\boldsymbol{\theta}_{:km}^{[in]} \right) e^{-j\varphi^{[in]}\left(\boldsymbol{\theta}_{km}^{[in]}\right)} \right| \cos\theta^{[t]} \, d\boldsymbol{\theta}^{[t]} \tag{5.5.17b}$$

$$= \left\{ \iint_{\Omega} \left| \sum_{m=0}^{+\infty} T_{km} \left(\left[\boldsymbol{\theta}^{[in]}; \mathbf{r}_{km}^{[in]} \right], \sigma \right) \right| \cos\theta^{[t]} \, d\boldsymbol{\theta}^{[t]} \right\} \mathcal{L}_\sigma \left(\boldsymbol{\theta}^{[in]} \right) \tag{5.5.17c}$$

$$= T_k'' \left(\boldsymbol{\theta}^{[in]}, \sigma \right) \mathcal{L}_\sigma \left(\boldsymbol{\theta}^{[in]} \right), \tag{5.5.17d}$$

where we have exploited the far field approximation in eq. (5.5.17b), which allows to remove the dependency from the position of incidence in the incident complex amplitude and the no cross-talking to get rid of the summation over all interferometers. Additionally, the common focal spot condition allows to uniquely identify that all the rays that focus on the coordinate $\mathbf{r}_{k0}^{[t]}$ have a specific incident angle $\boldsymbol{\theta}^{[in]}$, which is exploited in eq. (5.5.17c) to factor out of the integral the expression of the input radiance $\mathcal{L}_\sigma \left(\boldsymbol{\theta}^{[in]} \right)$, and its phase term disappears thanks by taking its module. Finally we defined the term:

$$T_k'' \left(\boldsymbol{\theta}^{[in]}, \sigma \right) := \iint_{\Omega} \left| \sum_{m=0}^{+\infty} T_{km} \left(\left[\boldsymbol{\theta}^{[in]}; \mathbf{r}_{:km}^{[in]} \right], \sigma \right) \right| \cos\theta^{[t]} \, d\boldsymbol{\theta}^{[t]} . \tag{5.5.18}$$

which is just a function of σ and $\boldsymbol{\theta}^{[in]}$, as the term $\mathbf{r}_{:km}^{[in]}$ can be expressed just as a function of the integration variable $\boldsymbol{\theta}^{[t]}$ and of $\mathbf{r}_{:k0}^{[t]}$, which has a one-to-one relationship with $\boldsymbol{\theta}^{[in]}$ according to eq. (5.5.12).

With this substitution, the acquisition of eq. (5.5.6) can be rewritten as follows:

$$y_{ki} = \iint_{\mathcal{S}_{ki}^{[t]}} \int_{\sigma_{min}}^{\sigma_{max}} T_k''(\boldsymbol{\theta}^{[in]}, \sigma) \mathcal{L}_\sigma \left(\boldsymbol{\theta}^{[in]} \right) \cos\theta^{[t]} \eta_{ki}(\sigma) \, d\sigma \, d\mathbf{r}^{[t]} \tag{5.5.19a}$$

$$= \iint_{\Omega_{ki}} \int_{\sigma_{min}}^{\sigma_{max}} T_k''(\boldsymbol{\theta}^{[in]}, \sigma) \mathcal{L}_\sigma \left(\boldsymbol{\theta}^{[in]} \right) \cos\theta^{[t]} \eta_{ki}(\sigma) \left| \frac{d\mathbf{r}^{[t]}}{d\boldsymbol{\theta}^{[in]}} \right| d\sigma \, d\boldsymbol{\theta}^{[in]} , \tag{5.5.19b}$$

where the Jacobian $\left|d\mathbf{r}^{[t]}/d\boldsymbol{\theta}^{[in]}\right|$, which is necessary for the change of variable. We have additionally defined for convenience the solid angle $\{\Omega_{ki}\}_{\substack{k\in[1,\dots,N_k]\\ i\in[1,\dots,|\mathcal{Q}_k|]}}$ as a continuous collection of all the incident directions $\boldsymbol{\theta}^{[in]}$ that converge on the surface $\mathcal{S}_{ki}^{[t]}$ of the detector:

$$\boldsymbol{\theta}^{[in]} \in \Omega_{ki} \Leftrightarrow \mathbf{r}_{:km}\left(\boldsymbol{\theta}^{[in]}\right) \in \mathcal{S}_{ki}^{[t]}\,. \tag{5.5.20}$$

5.5.6 Discretization of the acquisition model

With the analysis of the previous section, we assumed that it is possible to partition the FoV of an interferometer in a set of solid angles $\{\Omega_{ki}\}_{i\in\mathcal{Q}_k}$, which map the incident radiance to an associated set of acquisitions $\{y_{ki}\}_{i\in\mathcal{Q}_k}$. When different interferometers are considered together, i.e. for the purpose of gathering information from a snapshot acquisition, the fact that each interferometer partitions the FoV of the instrument in a different way may pose an issue. For the following analysis, we hence assume that this partition is exactly the same for every $k \in [1,\dots,N_k]$; in other words, we can define a certain disjoint set of solid angles of incidence $\{\Omega_i\}_{i\in[1,\dots,N_i]}$, such that all the incident rays within Ω_i are mapped on a given detector area, specific for each subimage:

$$\exists\Omega_i : \forall\boldsymbol{\theta}^{[in]} \in \Omega_i,\ k \in [1,\dots,N_k]\,, \mathbf{r}_{:k}^{[t]}\left(\boldsymbol{\theta}^{[in]}\right) \in \mathcal{S}_{ki}^{[t]}\,. \tag{5.5.21}$$

This condition 5.5.21 is not realistic in practice, but this misalignment effect can be corrected with either:

- a hardware implementation that compensates the parallax effect with a certain leading optical system;
- a software post-processing of the raw acquired data for the co-registration of the subimages, e.g. with the approaches which we present in Section 6.3.

Eq. (5.5.19) can be rewritten for convenience to:

$$y_{ki} = \iint_{\Omega_i}\int_{\sigma_{min}}^{\sigma_{max}} T_k''(\boldsymbol{\theta}^{[in]},\sigma)\mathcal{L}_\sigma\left(\boldsymbol{\theta}^{[in]}\right)\,d\sigma\,d\boldsymbol{\theta}^{[in]}\,, \tag{5.5.22}$$

where we have defined for simplicity $T_k'(\boldsymbol{\theta}^{[in]},\sigma) := T_k''(\boldsymbol{\theta}^{[in]},\sigma)\eta(\sigma)\left|\frac{\mathbf{r}^{[t]}}{d\boldsymbol{\theta}^{[in]}}\right|$, assuming that the quantum efficiency of all detectors is the same. We will assume here

that this term includes also the effects of the leading optic, which we did not include to simplify the exposition.

A computationally tractable version of eq. (5.5.22) requires a discrete approximation of the above expression, which can be obtained by an appropriate partition of the spaces of the involved variables. In particular, the space of the wavenumbers can be divided into N_b equally spaced intervals, whose midpoints are denoted with $\{\sigma_l\}_{l\in[1,\dots,N_b]}$. Additionally, $T'_k(\boldsymbol{\theta}^{[in]},\sigma)$ is considered approximately uniform both within a given wavenumbers' interval $[\sigma_l-\Delta\sigma/2,\sigma_l+\Delta\sigma/2]$ and within the solid angle Ω_i. This constant value is denoted for short as $T'_k(\boldsymbol{\theta}_i^{[in]},\sigma_l)$, where $\boldsymbol{\theta}_i^{[in]}$ is the centroid of the solid angle Ω_i.

Eq. (5.5.22) can then be rewritten in any of the following forms:

$$y_{ki}=\sum_{l=1}^{N_b}a_{kli}x_{li}\,, \tag{5.5.23a}$$

$$\mathbf{y}_{:i}=\mathbf{A}_{::i}\mathbf{x}_{:i}\,, \tag{5.5.23b}$$

where $\mathbf{y}_{:i}=\{y_{ki}\}_{k\in[1,\dots,N_b]}$ is a column vector representing the acquisitions relative to the solid angle of incidence Ω_i. The elements of $\mathbf{x}_{:i}=\{x_{li}\}_{l\in[1,\dots,N_b]}$ and $\mathbf{A}_{::i}=\{a_{kli}\}_{k\in[1,\dots,N_k],l\in[1,\dots,N_b]}$ are defined as:

$$a_{kli}:=T'_k\left(\boldsymbol{\theta}_i^{[in]},\sigma_l\right)\,, \tag{5.5.24a}$$

$$x_{li}:=\iint_{\Omega_i}\int_{\sigma_l-\frac{\Delta\sigma}{2}}^{\sigma_l+\frac{\Delta\sigma}{2}}\mathcal{L}_\sigma\left(\boldsymbol{\theta}^{[in]}\right)\,d\sigma\,d\boldsymbol{\theta}^{[in]}\,. \tag{5.5.24b}$$

With this formalism we obtain that:

- $\{x_{li}\}_{i\in[1,\dots,N_i],l\in[1,\dots,N_b]}$ are a discrete characterization of the incoming spectral radiance $\mathcal{L}_\sigma\left(\boldsymbol{\theta}^{[in]}\right)$ obtained through the measure of incident radiant flux within the set of solid angle $\{\Omega_i\}_{i\in[1,\dots,N_i]}$ and in the wavelength ranges $\{[\sigma_l-\Delta\sigma/2,\sigma_l-\Delta\sigma/2]\}_{i\in[1,\dots,N_i]}$.
- $\mathbf{y}_{:i}=\{y_{ki}\}_{k\in[1,\dots,N_b]}$ is a sampled version of the interferogram associated with a specific portion of the scene, as we have shown in the introduction in Fig. 1.6.

5.5.7 Definition of the transfer matrix

This section introduces a series of simplification of the model of the transfer function of eq. (5.5.24a) to be able to link the acquisition to the ideal behaviour of a FTS. The main assumption is that $T^{[opt]}$ and the quantum efficiency η exhibit no variation with respect to their spatial or spectral dependency, so that they can be set equal to 1 without loss of generality. The coefficient a_{kli} of the transfer matrix can subsequently be expressed as a sampled version $T_k\left(\boldsymbol{\theta}_i^{[in]}, \sigma_l\right)$ of the response of the interferometer. According to the analysis of Section 5.4.5 for the 2, N_m and ∞-wave model, in eq. (5.4.12), (5.4.13) and (5.4.14), its explicit expression becomes:

$$a_{kli} = \begin{cases} \left(1 + \mathcal{R}_{kli}^2 + 2\mathcal{R}_{kli}\cos\varphi_{kli}\right)\mathcal{T}_{kli}^2 & \text{2-wave model} \quad (5.5.25a) \\ \dfrac{1 + \mathcal{R}_{kli}^{2N_m} - 2\mathcal{R}_{kli}^{N_m}\cos(N_m\varphi_{kli})}{1 + \mathcal{R}_{kli}^2 - 2\mathcal{R}_{kli}\cos\varphi_{kli}}\mathcal{T}_{kli}^2 & N_m\text{- wave model} \quad (5.5.25b) \\ \dfrac{1}{(1-\mathcal{R}_{kli})^2 + 4\mathcal{R}_{kli}\sin^2(\varphi_{kli}/2)}\mathcal{T}_{kli}^2 & \infty\text{-wave model} \quad (5.5.25c) \end{cases}$$

where $\mathcal{T}_{ikl}$ and $\mathcal{R}_{ikl}$ are the geometric mean of the reflectivities and the transmissivities of the two surfaces of the interferometer, respectively, while φ_{ikl} denotes the round trip phase difference.

If the interferometer is bounded by two layers on its surface with the same refractive index $n_s = n_{s_1} = n_{s_2}$ and immersed in a medium with refractive index $n_0 = n_{0_1} =$

n_{0_2} (typically air), according to the analysis of Section 5.3.1, the relevant variables of eq. 5.5.25 can be obtained as follows:

$$\theta_{li} = \arcsin\left(\frac{n_0}{n(\sigma_l)}\sin\theta_i^{[in]}\right) \approx \frac{n_0}{n(\sigma_l)}\theta_i^{[in]} \tag{5.5.26a}$$

$$\varphi_{kli} = 2\pi\sigma_l\delta_{kli} - \varphi_{ki}^{[0]} = 2\pi\sigma_l n(\sigma_l) L_k \cos\theta_{li} - \varphi_{ki}^{[0]} \tag{5.5.26b}$$

$$\mathcal{R}_{kli,\perp} = \left|\frac{n(\sigma_l)\cos\theta_{li} - n_s(\sigma_l)\sqrt{1-\left(\frac{n(\sigma_l)}{n_s(\sigma_l)}\sin\theta_{li}\right)^2}}{n(\sigma_l)\cos\theta_{li} - n_s(\sigma_l)\sqrt{1-\left(\frac{n(\sigma_l)}{n_s(\sigma_l)}\sin\theta_{li}\right)^2}}\right|^2 \tag{5.5.26c}$$

$$\mathcal{R}_{kli,\parallel} = \left|\frac{n(\sigma_l)\sqrt{1-\left(\frac{n(\sigma_l)}{n_s(\sigma_l)}\sin\theta_{li}\right)^2} - n_s(\sigma_l)\cos\theta_{li}}{n(\sigma_l)\sqrt{1-\left(\frac{n(\sigma_l)}{n_s(\sigma_l)}\sin\theta_{li}\right)^2} + n_s(\sigma_l)\cos\theta_{li}}\right|^2 \tag{5.5.26d}$$

$$\mathcal{R}_{kli} \approx \frac{1}{2}\left(\mathcal{R}_{kli,\perp} + \mathcal{R}_{kli,\parallel}\right) \approx \left|\frac{n(\sigma_l) - n_s(\sigma_l)}{n(\sigma_l) + n_s(\sigma_l)}\right|^2 \tag{5.5.26e}$$

$$\mathcal{T}_{kli} = 1 - \mathcal{R}_{kli} \tag{5.5.26f}$$

where θ_{li} denotes the internal reflection angle, δ_{kli} represents the round trip OPD of the consecutively reflecting rays within the interferometer, L_k denotes the thickness of the k-th interferometer, while $\mathcal{R}_{kli,\perp}$ and $\mathcal{R}_{kli,\parallel}$ are the internal reflectivities in case the incident rays have perpendicular or parallel polarization, respectively. $\varphi_{ki}^{[0]}$ defines a phase shift term to take into account all additional effect of phase differences between consecutively transmitted rays which are not due to the difference in the optical path.

In the above equations, we have made explicit the dependency from the wavenumber σ_l both for the internal refraction index n and for the one at the surface n_s. It was instead made implicit for the term n_0, as it is usually air, as it is approximately constant with variation of the wavelengths. In case no layer is present on the surface on the interferometer, then $n_s = n_0$.

The approximation in eq. (5.5.26a) is a result of eq. (5.4.3b), where the internal reflection angle is expressed as function of the polar component $\theta_i^{[in]}$ of the incident angle via Snell's law in the paraxial approximation. The other equations are an instance of the Fresnel equations, which were derived from the analysis of the surface reflectivity and transmissivity in Section 5.3.1; for a fixed incidence angle and wavenumber, the expression of $\mathcal{R}_{kli}$ and $\mathcal{T}_{kli}$ is the same for every interferometer, as consequence of the assumption of even coating over all the surfaces. The dependency from the index k is hence just kept to take into account possible manufacturing differences among different interferometers.

5.5.8 Link with the Fourier transform spectrometer

The typical behaviour of a FTS can be obtained by assuming a 2-wave model on the simplified expression of the transfer function that was derived in the previous section, for which the transfer function becomes:

$$T'_k(\boldsymbol{\theta}_i^{[in]}, \sigma) = \mathcal{T}_{ki}^2 \left(1 + \mathcal{R}_i^2 + \mathcal{R}_i \cos(2\pi\sigma\delta_{ki} - \varphi_k^{[0]})\right) \tag{5.5.27}$$

and the associated acquisition y_{ki}:

$$\begin{aligned}
y_{ki} &= \int_0^\infty T_k\left(\boldsymbol{\theta}_i^{[in]}, \sigma\right) \mathcal{E}_\sigma(\Omega_i)\, d\sigma && \text{(5.5.28a)}\\
&= \mathcal{T}_{ki}^2 \int_0^\infty \left(1 + \mathcal{R}_{ki}^2 + 2\mathcal{R}_{ki}\cos\left(2\pi\delta_{ki}\sigma - \varphi_{ki}^{[0]}\right)\right) \mathcal{E}_\sigma(\Omega_i)\, d\sigma && \text{(5.5.28b)}\\
&= \mathcal{T}_{ki}^2(1 + \mathcal{R}_{ki}^2) \int_0^\infty \mathcal{E}_\sigma(\Omega_i)\, d\sigma \\
&\qquad + 2\mathcal{R}_{ki} \int_0^{+\infty} \mathcal{E}_\sigma(\Omega_i) \cos\left(2\pi\delta_{ik}\sigma - \varphi_{ki}^{[0]}\right)\, d\sigma && \text{(5.5.28c)}\\
&= \mathcal{T}_{ki}^2(1 + \mathcal{R}_{ki}^2)\overline{\mathcal{E}}_\sigma(\Omega_i) + 2\mathcal{R}_{ki}\mathcal{T}_{ki}^2 \cos\varphi_{ki}^{[0]} \int_{-\infty}^{+\infty} \mathcal{E}_{|\sigma|}(\Omega_i) e^{-j2\pi\delta_{ki}\sigma}\, d\sigma\,, && \text{(5.5.28d)}
\end{aligned}$$

where $\mathcal{E}_\sigma(\Omega_i) = \iint_{\Omega_i} \mathcal{L}_\sigma\left(\boldsymbol{\theta}^{[in]}\right) d\boldsymbol{\theta}^{[in]}$ is the spectral irradiance associated with the solid angle of incidence Ω_i, $\overline{\mathcal{E}}_{|\sigma|}(\Omega_i)$ is its average value with respect to the wavenumbers range and $\mathcal{E}_{|\sigma|}(\Omega_i)$ is its even symmetrical extension to negative wavenumbers.

If $\overline{\mathcal{E}}_{|\sigma|}(\Omega_i)$ is known, eq. (5.5.28d) allows to express the acquisition y_{ki} as a linear transformation of the Fourier transform of $\mathcal{E}_{|\sigma|}(\Omega_i)$; this constitutes the expected ideal behaviour of ImSPOC as a FTS, described in general terms in Section 5.4.3.

If no prior is considered, according to Shannon-Nyquist theorem, the absolute limit for a perfect reconstruction of a spectral irradiance $\mathcal{E}_\sigma\left(\Omega_i^{[in]}\right)$ with monolateral bandwidth $B_\sigma = \sigma_{max} - \sigma_{min}$ from $\mathbf{y}_i$, which avoids aliasing, the average OPD's step size $\Delta\delta_i = \sum_{k=1}^{N_k-1}(\delta_{i,k+1} - \delta_{ik})$, must be such that:

$$\Delta\delta_i < 2/B_\sigma\,, \qquad \forall i \in [1, \ldots, N_i]\,. \tag{5.5.29}$$

One of the first designs of ImSPOC, known as **microSPOC**, was conceived to use a triangular slab, so that one could obtain a situation of interferometer thicknesses

varying continuously [89], which would imply $\Delta\delta_i \to 0$. Technical issues, especially related to cross-talk, eventually led to the current staircase design for the interferometers' matrices, which implies that their thicknesses can only be chosen within a discrete set $\{L_k\}_{k\in[1,\dots,N_k]}$, which we consider arranged in increasing order.

The condition of eq. (5.5.29) must be valid for every angle of incidence, the average difference $\Delta L = \sum_{k=1}^{N_k-1} L_k$ between consecutive thicknesses of the interferometer must abide the following condition:

$$2n\ \Delta L \cos\left(\frac{n_0}{n}\theta_{max}^{[in]}\right) < \frac{2}{B_\sigma}\,, \tag{5.5.30a}$$

$$\Delta L < \frac{1}{nB_\sigma \cos\left(\frac{n_0}{n}\theta_{max}^{[in]}\right)}\,, \tag{5.5.30b}$$

where $\theta_{max}^{[in]}$ is the maximum polar angle of incidence allowed by the device; in the case of paraxial approximation, the cosine term is close to 1, so the condition simplifies to $\Delta L < 1/(nB_\sigma)$.

In practice, however, the sampling period $\Delta\delta_i$ of the OPD space must be much smaller than this upper limit, both because of manufacturing imperfections, which does not allow to limit the spectrum bandwidth to its nominal values. Additionally, extra mathematical conditions have to be set up if the spectrum is not in its base-band form, which opens up to the possibility for replicas of the reconstructed spectrum to overlap in case the sampling frequency $\Delta\delta_i\sigma_{min}$ is not a multiple of σ_{min}. To deal with this case, one can make use of the **bandpass sampling theorem** [223], to set up extra conditions to avoid aliasing. In particular for a uniform sampling ($\Delta\delta = \delta_{k+1} - \delta_k,\ \forall k = 2, ..., N_k$), the condition has to be extended as follows:

$$\frac{q-1}{2\sigma_{min}} \le \Delta\delta \le \frac{q}{2\sigma_{max}}\,, \qquad \forall q \in \mathbb{N} \text{ and } 1 \le q \le \left\lfloor \frac{\sigma_{max}}{B_\sigma} \right\rfloor\,, \tag{5.5.31}$$

where $\lfloor\cdot\rfloor$ denotes the highest integer smaller than its argument. The allowed ranges of sample frequencies $1/\Delta\delta$ are shown in Fig. 5.11.

Additional restrictions come into play when we consider different models for the transfer function; under the constraints that we are considering in this chapter, the transfer function $T_k(\boldsymbol{\theta}_i^{[in]}, \sigma)$, in any of the models proposed in Section 5.4.5 is a function of σ only through the term $2\pi\delta_{ki}\sigma$.

In particular, for the general N_m-wave case (which includes as limit the ∞-wave case), we obtain the following Fourier series decomposition:

$$T_k\left(\boldsymbol{\theta}_i^{[in]},\sigma\right) = \mathcal{T}_{ki}^2 \frac{1+\mathcal{R}_{ki}^{2N_m}-2\mathcal{R}_{ki}^{N_m}\cos\left(2N_m\pi\delta_{ki}\sigma\right)}{1+\mathcal{R}_{ki}^2-2\mathcal{R}_{ki}\cos\left(2\pi\delta_{ki}\sigma\right)} \tag{5.5.32a}$$

$$= \mathcal{T}_{ki}^2 \frac{1-\mathcal{R}_{ki}^{2N_m}}{1-\mathcal{R}_{ki}^2}\left(1 + 2\sum_{m=1}^{N_m-1}\mathcal{R}_{ki}^m \frac{1-\mathcal{R}_{ki}^{2(N_m-m)}}{1-\mathcal{R}_{ki}^{2N_m}}\cos\left(2m\pi\delta_{ki}\sigma\right)\right). \tag{5.5.32b}$$

For a perfect reconstruction, the Nyquist condition becomes $\Delta\delta_i < 2/((N_m-1)B_\sigma)$ as the terms $\cos\left(2m\pi\delta_{ki}\sigma\right)$ are equivalent to scaling the wavenumber range by a factor of m. This condition usually imposes a strong constraint on the design of the interferometers' array matrix, hence the ImSPOC prototypes are usually designed with low reflectivities $\mathcal{R}_{ki}$ to make negligible all higher order terms of the Fourier series.

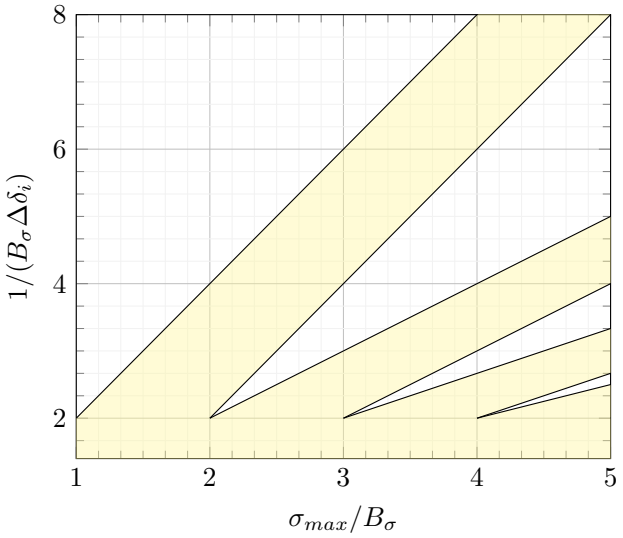


Fig. 5.11. Plot of the forbidden area (in yellow) for the perfect reconstruction of a Fourier transform of a spectrum with bandwidth B_σ and maximum wavenumber σ_{max}, through an interferometer sampled with step size $\Delta\delta_i$.

6 Data processing pipeline of ImSPOC acquisitions

This chapter presents the image spectrometer on chip (ImSPOC) concept from the perspective of signal processing, with the aim to present the necessary operations to transform its raw acquisition into intelligible products, a spectral representation of the scene and its associated image.

The discussion is intended as an accessible tutorial for a data processing engineer, and acts as a detailed description of some preliminary approaches to address each stage of the processing pipeline.

In contrast with the more theoretical approach of Section 5.5, we also address some challenges associated with the nonidealities of real devices. Nevertheless, the chosen methodology is still based on the physical model of the optical transformation, so that the proposed procedures are available even with a limited amount of available prototypes and acquisitions. Additionally, we favor here fast algorithms for the recovery of the quantities of interest, as the final goal is to embark the device on board of embedded platforms.

6.1 Introduction

In this chapter, we present a comprehensive discussion of the pipeline of operations that transform the raw acquisitions of an ImSPOC prototype to user intelligible products, which can be exploited to extract information on the spectral characteristics of the scene.

The ImSPOC concept was described in detail in Section 5.5.2, so we just recall here that its concept defines an imaging spectrometer based on the interferometry of Fabry-Pérot (FP). Its optical components, shown in Fig. 6.1, are composed of an array of low finesse FP etalons of different thicknesses, disposed in a staircase pattern. The array is overlaid over a matrix of microlenses, which focuses a packet of parallel rays incident to the input plane to an array of spots on the focal plane.

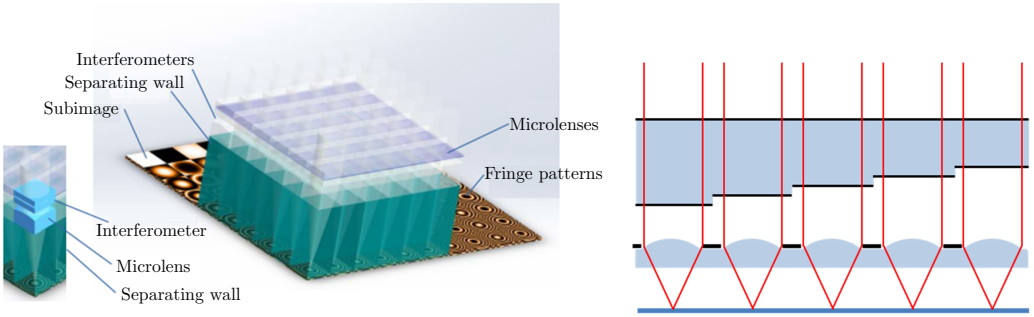
The raw acquisition is composed of a series of subimages, which can be ideally stacked to identify the interferometers associated with the spectral radiance in a given direction.

The main envisioned application for the ImSPOC device is to estimate the concentration of a given component (e.g. a gas) in the atmosphere. However, in comparison with other approaches [137], we assume that the instrument mode of operation is that of a spectro-imaging system. In this configuration, the goal is to estimate a discrete representation of the spectral radiance, and an associated image, for which each pixel is associated a given partition of the field of view (FoV).

The proposed pipeline of data processing is aimed at addressing the following issues, which are discussed in their dedicated sections:

- Segmentation of the subimages (Section 6.2);
- Spatial alignment of the datacube of subimages (Section 6.3);
- Spectral calibration of the optical transfer function, which characterizes the device (Section 6.4);
- Reconstruction of the input spectra relative to the incident spectral radiance (Section 6.5).

In this introduction, after we describe how the ImSPOC images are represented, the pipeline of operations is laid out in terms of the desired products, challenges, and protocols to implement.



(a) Tridimensional view **(b)** Side view

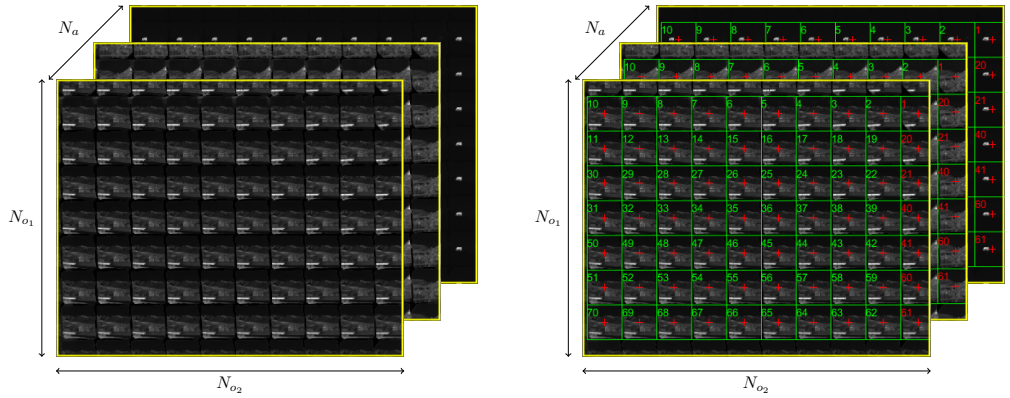
Fig. 6.1. Concept of the ImSPOC acquisition device. In the side view, from the top, the device is composed by a staircase shaped array of interferometers, a matrix of microlenses and a set of photodetectors on the focal plane.

6.1.1 Image representation

Given a focal plane made up of $N_{o_1} \times N_{o_2}$ photodetectors, distributed over a focal plane, the natural representation of a bundle of N_a raw acquisition is given by a 3-dimensional array $\boldsymbol{\mathcal{U}}^{[o]} \in \mathbb{R}^{N_{o_1} \times N_{o_2} \times N_a}$. An example of a set of acquisitions is shown in Fig. 6.2a.

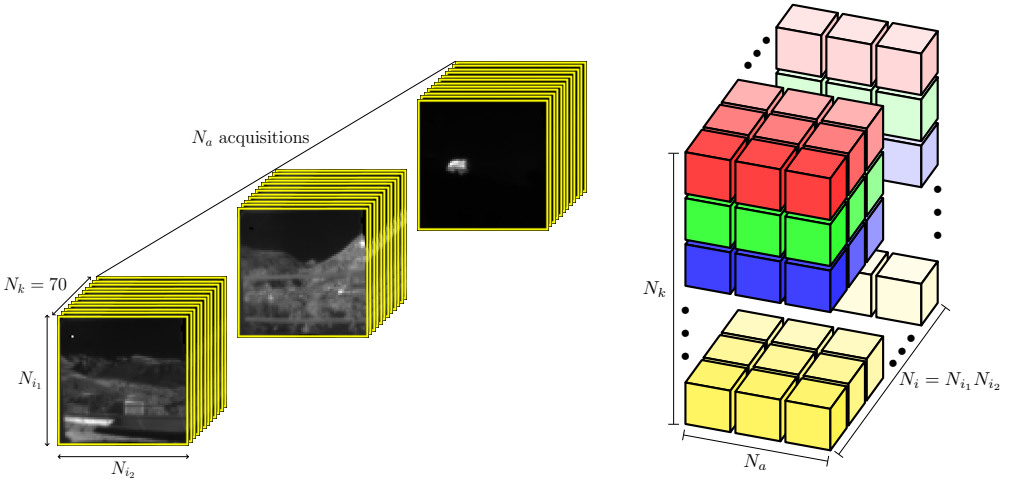
Given the particular physical design of the ImSPOC acquisition system, it is however convenient to generate a datacube $\boldsymbol{\mathcal{U}}^{[s]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_k \times N_a}$ by cropping and stacking N_k subimages of sizes $N_{i_1} \times N_{i_2}$. This operation is detailed in Section 6.2.1 and shown in Fig. 6.2c.

In this chapter, we often employ a different representation of the image, where elements are arranged differently. Specifically, the unfolding operation $\boldsymbol{\mathcal{S}} = \mathrm{reshape}(\boldsymbol{\mathcal{U}}^{[s]})$ reshapes the tensor $\boldsymbol{\mathcal{U}}^{[s]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_k \times N_a}$ into a new tensor $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{N_k \times N_a \times N_i}$, where $N_i = N_{i_1} N_{i_2}$. This operation is still a representation in lexicographic order, but the dimensions are permuted to make the interferogram samples appear along columns. This arrangement is shown in Fig. 6.2d.



(a) Raw acquisition $\mathcal{U}^{[o]}$

(b) Estimation of the centers of the subimages



(c) Stacked interferogram - $\mathcal{U}^{[s]} = \text{stack}(\mathcal{U}^{[o]}, \mathbf{R}^{[0]}, [N_{i_1}, N_{i_2}])$

(d) Permuted lexicographic order - $\mathcal{S} = \text{reshape}(\mathcal{U}^{[s]})$

Fig. 6.2. Various image representations of ImSPOC acquisition. On the first row, a set of acquisitions in their natural representation, where $[N_{o_1}, N_{o_2}]$ are the column and row pixels of the focal plane. In the second row the subimages are stacked to obtain the datacube $\mathcal{U}^{[y]}$, whose dimensions are rearranged in Fig. 6.2d to obtain the permuted lexicographic order.

6.1.2 Notation

In this chapter we denote the following variables in their natural representation, whose first two dimensions denote the column and row pixel:

- The raw acquisitions: $\boldsymbol{\mathcal{U}}^{[o]} \in \mathbb{R}^{N_{o_1} \times N_{o_2} \times N_a}$,
- The raw datacube of subimages: $\boldsymbol{\mathcal{U}}^{[s]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_k \times N_a}$;
- The co-registered datacube of subimages: $\boldsymbol{\mathcal{U}}^{[y]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_k \times N_a}$.

The following images are instead described in their permuted lexicographic order described in Section 6.1.1:

- the co-registered datacube of subimages: $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{N_k \times N_a \times N_i}$;
- the ideal product to reconstruct: $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{N_b \times N_a \times N_i}$;
- the estimated product: $\widehat{\boldsymbol{\mathcal{X}}} \in \mathbb{R}^{N_b \times N_a \times N_i}$;
- the transfer model: $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{N_k \times N_b \times N_i}$.

The cardinality of each domain is given below:

- N_{o_1} and N_{o_2}: the number of photodetectors on the column and on the row of the focal plane array (FPA), whose product $N_o = N_{o_1} N_{o_2}$ is the total number of photodetectors;
- N_{i_1} and N_{i_2}: the column and row pixel size of each subimage, whose product is $N_i = N_{i_1} N_{i_2}$;
- N_a: the number of acquisitions;
- N_k: the number of active interferometers, or equivalently, the amount of samples in the interferogram;
- N_b: the number of spectrum samples for the reconstruction;
- N_b': the number of spectrum samples employed in the direct model.

For a single acquisition ($N_a = 1$) the tensor variables are represented with a matrix notation, i.e. $\mathbf{U}^{[o]}$ denotes a single raw acquisition and $u^{[o]}_{i_1,i_2}$ is a generic pixel on the rectangular grid of photodetectors, which forms the FPA.

This notation is chosen to keep the standard of matrices being denoted with an uppercase bold letter, vectors with a lowercase bold letter, scalars with a lowercase non-bold letter. The arrays with dimensions superior to 3 are shown in their tensor notation (e.g.: $\boldsymbol{\mathcal{U}}$).

The variables can be sliced in the following ways:

- **Frontal Slicing**: The i-th frontal slice of any 3-way tensor is denoted with their matrix notation, so i.e. $\mathbf{Y}_{::i}$ is a frontal slice of $\boldsymbol{\mathcal{Y}}$, which defines a collection of N_k interferograms over N_a acquisitions, all relative to a given angle of incidence Ω_i. To simplify the notation, if Ω_i is fixed, the listed variables can be simply rewritten as the matrices $\mathbf{Y} \in \mathbb{R}^{N_k \times N_a}$, $\mathbf{X} \in \mathbb{R}^{N_b \times N_a}$ and $\mathbf{A} \in \mathbb{R}^{N_k \times N_b}$.
- **Vector Slicing**: a specific column or row of a matrix is represented with a vector notation, e.g. $\mathbf{y}_{:k}$ and $\mathbf{y}_{l:}$ are the k-th column and l-th row of $\mathbf{Y}$ (with Ω_i fixed), respectively.
- **Element Slicing**: The generic element of any matrix is denoted with an indexed scalar notation, e.g. a_{kli} denotes the element of $\boldsymbol{\mathcal{A}}$ associated with the k-th interferometer, to the l-th spectrum sample and to the i-th solid angle of incidence Ω_i. Once again, if Ω_i is fixed, its generic element is simply denoted with a_{kl} (or sometimes with $a_{k,l}$).

Given a generic bidimensional matrix $\mathbf{C}$, we denote with $\mathbf{C}^{\mathsf{T}}$, $\mathbf{C}^{\dagger}$, $\bar{\mathbf{c}}_{k:}$, and $\bar{\mathbf{c}}_{:l}$ its transpose, its (Moore-Penrose) pseudo-inverse, and the mean value of its k-th row and l-th column, respectively.

Whenever any operation is performed between a matrix and a scalar, it is implicitly assumed that the operation is broadcast to all elements of the matrix, e.g. a difference $\mathbf{C} - \bar{\mathbf{c}}_{k:}$ implies that the mean of the k-th row of $\mathbf{C}$ is subtracted from each element of $\mathbf{C}$.

Additionally, $[\cdot, \cdot]$ and $[\cdot; \cdot]$ stand for row and column concatenation, while $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_F$ and $< \cdot, \cdot >$ denote the ℓ_1 norm, ℓ_2 norm, the Frobenius norm, and the scalar product, respectively.

6.1.3 Desired products

In the interest of describing the necessary operations aimed at providing intelligible results for the final users, it is useful to formally identify a stack of data processing operations, aimed at transforming raw acquisitions to useful data for the final applications; as no unique standard is common practice to label such operations, we propose in this work to employ a nomenclature inspired by the terminology of National Aeronautics and Space Administration (NASA)'s **Data Processing Levels**.

This formalism consists in identifying a sequence of products at different stages of the processing procedure, which we show in Table 6.1 for the reader convenience. They are specifically defined as:

- **Level 0 (L0)**: The matrix of photo-detectors' raw acquisitions.
- **Level 1 (L1)**: A datacube of tractable interferogram; in general terms, the operation to obtain L1 products include optical corrections and spatial alignments, aimed at relating the spatial coordinates of each slice to the same unique set of incident angles.
- **Level 2 (L2)**: The spectral characterization of the incident radiance; in our context, it requires an inversion of the L1 products, given a particular knowledge of the transfer model of the optical device.
- **Level 3 (L3)**: User ready products, such as gas concentrations or composition in the atmosphere. Those are obtained through post processing of the reconstructed spectra, such as spectral unmixing.

This separation between levels is just for the sake of exposition; they will be kept separate in this thesis, but nothing prevents to merge multiple processing steps in future works if particular necessities arise.

Table 6.1. Data processing levels, related to the ImSPOC acquisitions pipeline of operations.

Level	Data Description	ImSPOC Format	Variable
L0	Raw Acquisition	Photodetectors' intensity levels	Raw: $\boldsymbol{\mathcal{U}}^{[o]}$ Stacked: $\boldsymbol{\mathcal{U}}^{[s]}$
L1	Calibrated Output	Adjusted interferogram	Datacube: $\boldsymbol{\mathcal{U}}^{[y]}$ Lexicographic: $\boldsymbol{\mathcal{Y}}$
L2	Reconstructed Input	Directional Spectrum	Ideal: $\boldsymbol{\mathcal{X}}$ Reconstructed: $\widehat{\boldsymbol{\mathcal{X}}}$
L3	Application Ready Products	Gas Concentration, composition, etc.	Application dependent

We detail here which variables we associate to each of the processing levels.

- **Raw Input (L0)**: At this stage the available information is just made up of the intensity levels acquired by the photodetectors. Since the FPA is composed by a set of photodetectors arranged in a rectangular matrix of size $N_{o_1} \times N_{o_2}$, the natural representation of the raw input is a 3-dimensional array $\boldsymbol{\mathcal{U}}^{[o]} \in \mathbf{R}^{N_{o_1} \times N_{o_2} \times N_a}$. According to the discussion of Section 6.1.1, the relevant intensity levels are typically rearranged in a datacube $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{N_k \times N_a \times N_i}$ made up of N_k subimages of N_i pixels.
- **Co-registered datacube (L1)**: At this stage, we require, for each vector slice $\mathbf{s}_{:li}$ of the datacube $\boldsymbol{\mathcal{S}}$, to describe an intereferogram associated with the solid angle of incidence Ω_i. Unfortunately, the different optical properties of each of the N_k microlenses introduce slightly different optical distortions, that map to slight misalignments across different subimages. These have to be

corrected with a co-registration procedure to generate an aligned datacube $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{N_k \times N_a \times N_i}$.

- **Reconstructed spectra (L2)**: This is a discretized representation of the incident spectral radiance. Specifically, the FoV is segmented into a set of solid angle of incidence $\{\Omega_i\}_{i\in[1,\dots,N_i]}$, and the wavenumber range into a set of intervals $\{[\sigma_l - \Delta\sigma/2, \sigma_l + \Delta\sigma/2]\}_{l\in[1,\dots,N_b]}$. With this representation the 3-dimensional array $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{N_b \times N_a \times N_i}$ made up of the elements:

$$x_{lmi} = \int_{\sigma_l - \frac{\Delta\sigma}{2}}^{\sigma_l + \frac{\Delta\sigma}{2}} \iint_{\Omega_i} \mathcal{L}_{\sigma}^{[m]}\left(\boldsymbol{\theta}^{[in]}\right) \, d\boldsymbol{\theta}^{[in]} \, d\sigma \,. \tag{6.1.1}$$

  Here, $\mathcal{L}_{\sigma}^{[m]}\left(\boldsymbol{\theta}^{[in]}\right)$ denotes the spectral radiance of the m-th acquisition in terms of the wavenumber σ (or, in other words, the reciprocal of the wavelength in vacuum) and of the incident angle $\boldsymbol{\theta}^{[in]} = \left[\theta^{[in]}, \varphi^{[in]}\right]$. The latter is expressed in spherical coordinates in terms of the incident polar and azimuthal angle, respectively.
- **Information on the scene (L3)**: The final target of the instrument is to provide spectral information of the scene under target, as the required radiometric measure involves a certain detection of its direct, reflected, and diffracted electro-magnetic (EM) radiation. This in turn depends not only on the characteristic of the scene itself, but also on the atmospheric path travelled by the EM radiations; some post-processing operations is thus required to link the EM radiations incident to the device to the characteristics of the scene. As the scope of this work does not involve any discussion on the problem of **atmospheric correction**, we redirect the reader to the relevant literature for more details [193, 201, 204].

The scope of this thesis is limited to the analysis on the reconstruction up to the L2 of the pipeline of operation, which is equivalent to recovering the spectral characteristics of the incident light beams.

The spectral radiance $\mathcal{L}_{\sigma}^{[m]}(\boldsymbol{\theta}^{[in]})$ is a function of continuous variables, so we need some criterion to classify the quality of the reconstructed product $\widehat{\boldsymbol{\mathcal{X}}}$, which is discrete in nature. In simple terms, we say that $\widehat{\boldsymbol{\mathcal{X}}}$ has a:

- a high **spectral resolution** if short step sizes are chosen for the wavenumber σ (i.e. if the amount of samples N_b is large);
- a high **spatial resolution** if the FoV is partitioned into small solid angles (i.e. if N_i is large).

6.1.4 Challenges

The ImSPOC acquisition can be seen as a problem of **computational imaging**, as the raw acquisition actively lies in a domain other than the desired one. As common in this scenario, this situation demands an inversion of the measured quantities in order to recover the input spectra [117, 127, 172].

The main problem to overcome is related to the uncertainty of the measurements. This is due to the noisy phenomena which characterize the sensors (Section 5.2.4). Additionally, the quantities to infer are intrinsically in a continuous domain, while we seek here for a discrete representation of that same input spectrum.

These conditions describe an ill-posed problem in the Hadamard sense [106], as it was described in Section 2.1.1. A simple inversion of the transfer matrix which describes the optical transformation would also enhance the noise, so we need to introduce here some penalization term to impose the reconstruction to be well-behaving.

Additionally, in comparison with the discussion of Section 5.5, the real acquisition system is characterized by certain nonidealities, which we need to compensate.

These effects include:

- a not perfectly symmetrical behaviour of the optical system associated with each interferometry, which causes spatial distortions on each subimage;
- a not perfect knowledge of the geometry of each FP etalon, which causes a mismatch between the expected and the real behaviour of the system.

6.1.5 Protocol of operations

The data processing procedures related to the ImSPOC device can be classified into:

- **Calibration**: where the target is to characterize the device under test. This one-time operation is performed in a controlled environment and setup of the input, and the obtained measurement is employed to improve the accuracy of the data processing;
- **Operation**: where the raw acquisition that describes an unknown scene is processed to recover the information on the incident spectral radiance.

The pipeline of the data processing operations of the ImSPOC acquisitions proposed in this chapter follows a layered approach and is described below:

- **Stacking**: whose aim is to segment the regions of interest (RoIs) associated with each subimage in order to arrange them in a datacube. This is equivalent to estimating the centers $\widehat{\mathbf{R}}^{[0]} \in \mathbb{R}^{2\times N_k}$ of each subimage, as the operation of piling subimages of size $N_{i_1} \times N_{i_2}$ is described by a function $\boldsymbol{\mathcal{U}}^{[s]} = \mathrm{stack}(\boldsymbol{\mathcal{U}}^{[o]}, \mathbf{R}^{[0]}, [N_{i_1}, N_{i_2}])$;
- **Co-registration**: whose aim is to generate a coherent datacube of interferograms. This phase (also known as **spatial calibration**) involves the compensation of spatially incoherent behaviours between different subimages, typically caused by a not fully verified parallax approximation of the view of the scene or small differences in the manufacture of the micro-lenses or optical cavities. The main procedure we propose here is to infer a geometry transformation function that describes an coordinate shifting procedure from the original stack $\boldsymbol{\mathcal{U}}^{[s]}$ to a novel stack $\boldsymbol{\mathcal{U}}^{[y]}$. The i-th frontal slice $\mathbf{Y}_{::i}$ of its permuted lexicographic representation describes then an interferogram that is exclusively related to a fixed FoV partition Ω_i.
- **Model characterization**: At this point the direct model of the optical system is described by the relation $\mathbf{X}_{::i} = \mathbf{A}_{::i}\mathbf{Y}_{::i}$; the aim of the spectral calibration is to match a set of known inputs and outputs, acquired in a controlled environment, from which we infer a description $\mathbf{A}^{[\widehat{\boldsymbol{\mathcal{B}}}]}$ of the direct transfer matrix, in terms of the set of parameters $\widehat{\boldsymbol{\mathcal{B}}}$. The proposed formulation assumes that the coefficient of the transfer matrix are samples of the models described in Section 5.5.7; e.g., these models includes the 2-wave, which defines the Fourier transform spectrometer (FTS) behaviour, and the ∞-wave model, also known as Airy's distribution. Among the parameters to estimate, the most relevant is the thickness of the array of interferometers, which generally differs from it nominal value; in fact, as manufacturing techniques can cause imperfections, the hypothesis of parallel face interfaces for the FP structures may not fully hold.
- **Inversion**: whose target is an estimation $\widehat{\mathbf{X}}_{::i}$ of the spectra related to the incident radiance, subject to the condition that $\widehat{\mathbf{X}}_{::i} \approx \mathbf{A}_{::i}\mathbf{Y}_{::i}$, where the matrix $\mathbf{A}_{::i}$ is given by the previous estimation. This is a classic inversion problem, which requires a regularization to counter the ill-posedness of the problem.

A summary of these procedures is shown in Fig. 6.3. The flowchart highlights the separation between calibration and operation of the device, illustrating the products that we expect to obtain at each stage of the processing pipeline.

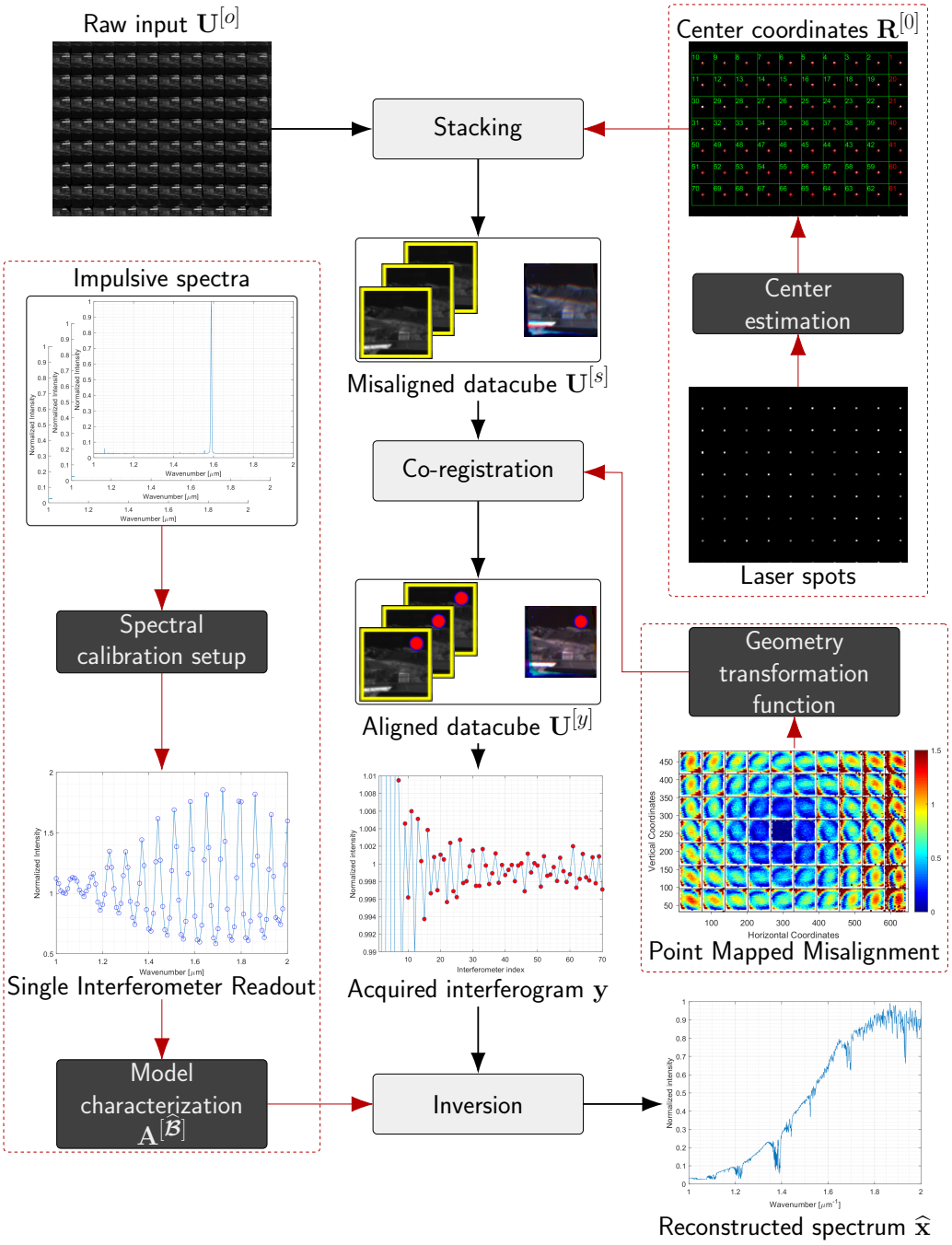


Fig. 6.3. Visual representation of the pipeline of operations that we employ for the spectrum reconstruction from an ImSPOC acquisition. The dashed rectangles denotes calibration operations (which are one-time only), while the flowchart linked by black arrows is necessary for each acquisition.

6.1.6 Novel contributions

The novel contributions presented in this chapter are listed below:

- Three novel approaches to identify the positions of the center positions of the subimages, based on the fitting of a Gaussian function, on the centroids of regions processed with mathematical morphology, and with a scanline approach;
- An analysis of the co-registration procedures for subimages, based on a point mapping approach and employing polynomial geometry transformation functions;
- Three novel methods for the estimation of the parameters (e.g. the interferometers' thicknesses) of the transfer function exploiting the description given by the Airy's distribution; the methods are based on its maximum likelihood (ML) formulation, on an exhaustive search (ES) and on nonlinear regression;
- A comparison between Bayesian frameworks for the inversion of the interferograms captured over a single pixel and a discussion of regularization both with a penalized matrix decomposition (PMD) and a least absolute shrinkage and selection operator (LASSO) approach;
- An analysis of the effect of the distortions in the reconstruction in the case of a mismatch between the acquisition and the inversion model.

6.1.7 Available prototypes

Each of the datasets employed in this work were captured with one of four available prototypes, manufactured in the context of the projects described in Appendix A.2 and whose characteristics are described in Table 6.2.

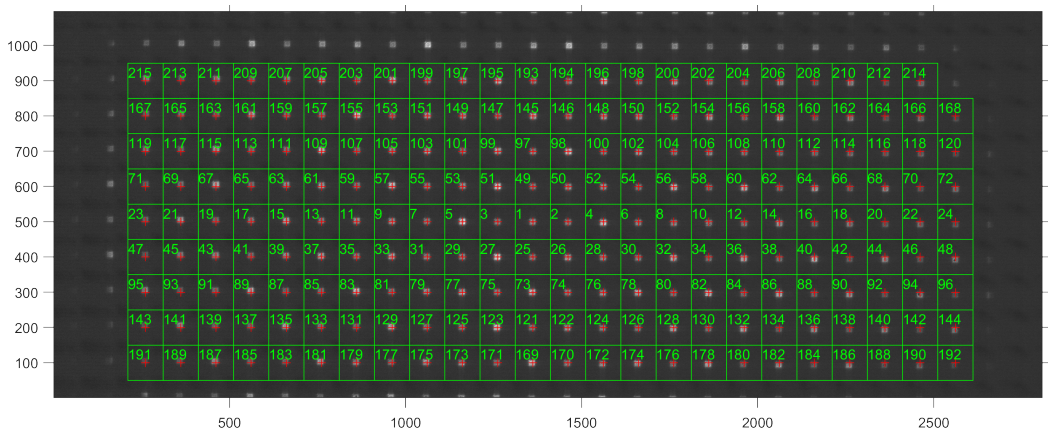
They will be generically labeled as **PROTO** with a progressive number to keep an internal consistency within the document. The **prototype ImSPOC-UV/VIS (PROTO-1)** and its evolution the **prototype ImSPOC-UV-drone (PROTO-2)** were developed in the context of the ImSPOC-UV project and they were originally designed for acquisitions in the ultraviolet (UV) range, but they were successively repurposed for operations in the visible (VIS) range. The **prototype ImaGAZ-1 (PROTO-3)** is also realized with the same technology (glass etalon with an ad-hoc reflective layer grown on the surfaces), but with a specific focus to cover the near infrared (NIR) wavelengths, according to the aims of the ImaGAZ (ImaGAZ) project.

All prototypes contain a matrix of interferometers disposed over a bidimensional matrix in a staircase pattern with linearly increasing thicknesses. An index is assigned to each interferometer, starting from 1, which corresponds to the interferometer at the optical contact (i.e., with its reflecting surfaces directly touching each other); the indices then increase sequentially following the increasing order of their nominal thickness.

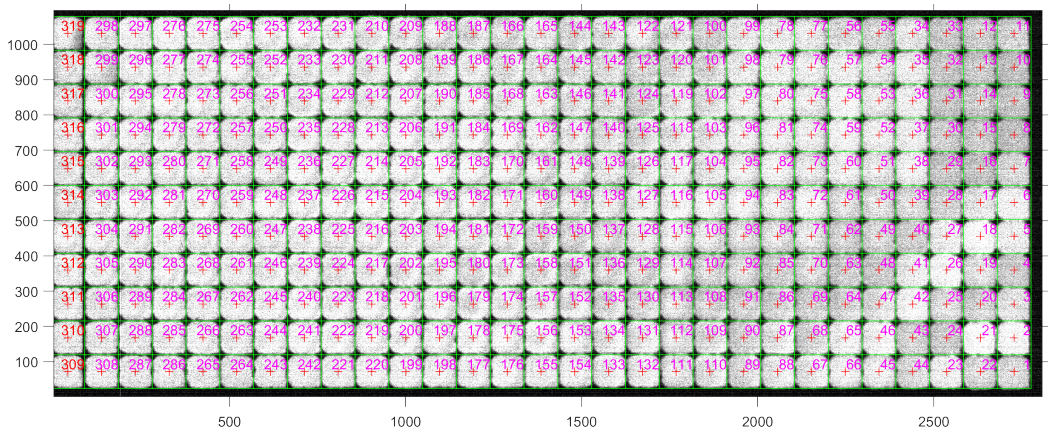
The PROTO-3 is characterized by a slightly different nominal surface reflectivity compared to the ImSPOC-UV prototypes, due to a less thick surface layer. Additionally, the optical behaviour of its reflective coating is not spectrally uniform, as its reflectivity changes significantly in the range $800 - 1100\ \mathrm{nm}$.

The **prototype NanoCarb-1 (PROTO-4)** has more notable manufacturing differences, as it follows the design principles of the NanoCarb technology (silicon etalons with no additional surface layer) and features four separated sets of interferometers, each dedicated to an estimation of gas concentration in different regions of the spectrum [97].

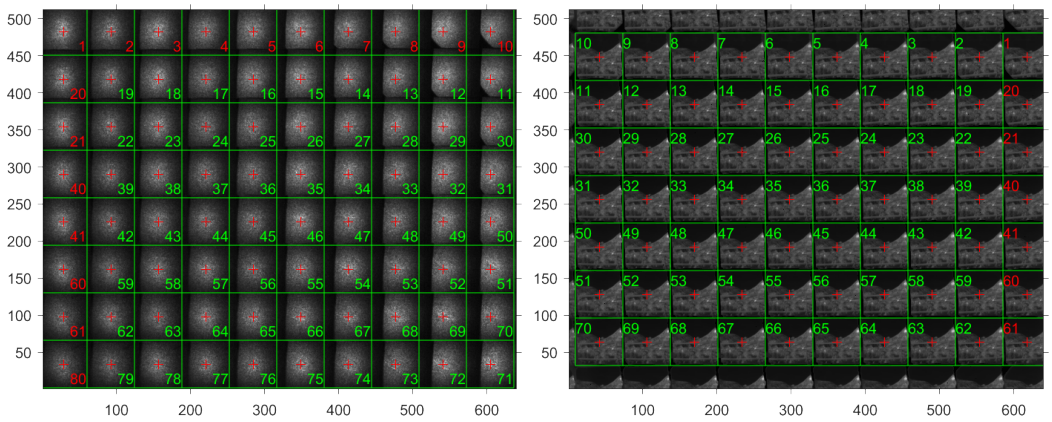
Each prototype has a slightly different arrangement of the staircase pattern. For each prototype, the indices shown in Fig. 6.4 are given in increasing order of interferometers' thicknesses. The PROTO-3 is the only notable exception, as its design also includes a second optical contact-type interferometer in position 80 to double check the accuracy of the manufacturing process. The staircase design of the PROTO-1 was designed to allow for the interferometers' thicknesses to increase linearly on both sides of a central vertical line. Its subimage in the bottom right corner is not operational as it is insufficiently illuminated by the incident field and was ignored in our analysis.



(a) PROTO-1



(b) PROTO-2



(c) PROTO-3 (d) PROTO-4

Fig. 6.4. Spatial arrangement of the subimages for different prototypes with numbers denoting the nominal thicknesses in increasing order. The subimage area is marked by a green outline and its center by a red cross. The background picture is a generic acquisition of the specified prototype. Some subimages extending outside the focal plane are marked by a red colored index.

Table 6.2. Characteristics of the available ImSPOC prototypes concept studied in this work. Sizes are in length by height.

Prototype	**PROTO-1**	**PROTO-2**	**PROTO-3**	**PROTO-4**
Project	ImSPOC-UV	ImSPOC-UV	ImaGAZ	NanoCarb
Wavelength range [nm]	$500 - 1000$	$380 - 1000$	$800 - 1700$	Reg. 1: ~ 780, Reg. 2: ~ 1600, Reg. 3: ~ 1660, Reg. 4: ~ 2060.
Acquisition's sizes [px]	2808×1096	2808×1096	640×512	640×512
Subimages' sizes [px]	100×100	96×96	64×64	64×64
Operative interferometers	215	319	$79(+1)$	70
Subimage arrangement	24×9	29×11	10×8	10×8
Subimages' step size [nm]	100	87.5	200	Variable
Nominal reflectivity	0.13	0.13	0.12	0.32

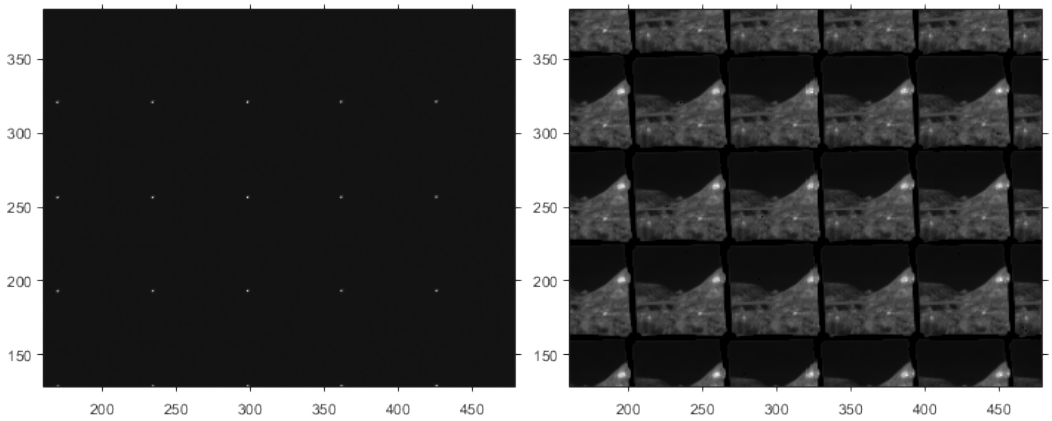
6.2 Center estimation

The particular concept design of ImSPOC introduces the particular challenge that multiple subimages are registered on the same focal plane. We are however interested in stacking those images in datacubes, so that each of the frontal slices represents a specific subimage. The determination of the regions associated with each subimage requires the determination of their centers.

We consider two possible scenarios for the determination of the centers:

- **Extended objects**: whose sources have a finite extension over the focal plane, such as in the case of a flat field illumination or a standard in situ acquisition. We propose to address this problem by pre-processing the image with a set of morphological transformations to identify the RoIs associated with the sources (Section 6.2.2). Then, we propose the centroid center estimation (CCE) and the scanline center estimation (SCE) methods to estimate their centers (Section 6.2.2).
- **Point objects**: such as a laser beam incident to the input surface, which maps to multiple replica on the captured image, as shown in Fig.6.5a. We propose to address this case with the Gaussian-fit center estimation (GCE) method, described in Section 6.2.3, for which we assume to already have an available datacube of stacked subimages, and only one laser spot is detectable on each frontal slice of the datacube.

A final discussion on the results is given in Section 6.2.4.



(a) Point source **(b)** Extended source

Fig. 6.5. Two different acquisitions taken with the PROTO-4, in response of a point and an extended source.

6.2.1 Coordinate system

In this chapter, the spatial coordinates are given in a list of generically N_r pairs, arranged over a generic matrix $\mathbf{R} \in \mathbb{R}^{2\times N_r}$. With this description, the k-th column $r_{:k}$ is a column vector whose elements define the vertical and horizontal coordinate of the k-th pixel.

For our purposes, we are interested in defining two different coordinate systems, one for the image represented in absolute coordinates, which describes the full FPA, and one in relative coordinates, which describes the positions of its stacked version.

Absolute coordinates

The FPA is an array of N_o photo-detectors, which we consider regularly spaced and arranged over a rectangular grid of sizes $N_{o_1} \times N_{o_2}$.

Let $\mathbf{R}^{[t]} \in \mathbb{R}^{2\times N_o}$ denote the centers of these photodetectors. If we assume the distance between adjacent photodetectors is a single unit, the elements of $\mathbf{R}^{[t]}$ are given by:

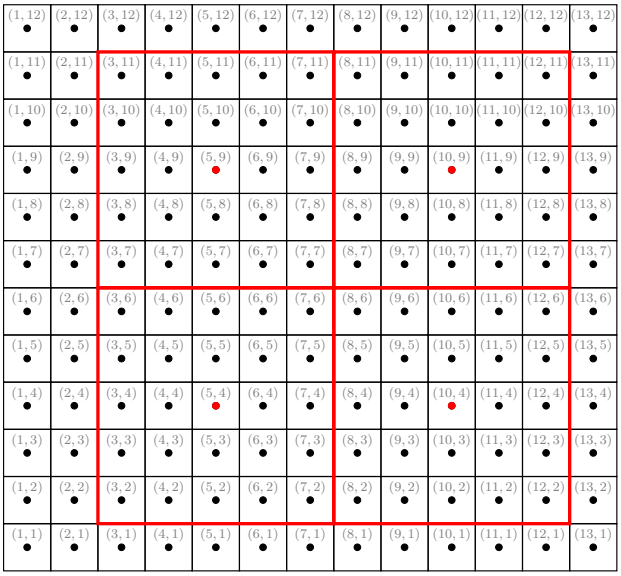
$$\begin{cases} r^{[t]}_{1,(i_2-1)N_{o_1}+i_1} & = i_1\,, \\ r^{[t]}_{2,(i_2-1)N_{o_1}+i_1} & = i_2\,, \end{cases} \qquad \forall i_1 \in [1,\dots,N_{o_1}]\,,\ i_2 \in [1,\dots,N_{o_2}]\,, \tag{6.2.1}$$

so that all coordinates are integer values, as it is shown in Fig. 6.6a.

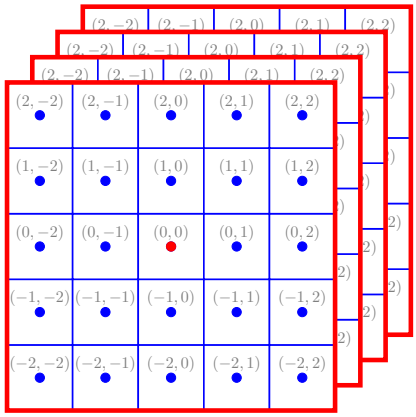
The raw acquisition is made up of a given amount N_k of subimages, whose centers $\mathbf{R}^{[0]} \in \mathbb{R}^{2\times N_k}$ can be expressed in the same coordinate system. These centers are in general not integer values (i.e., in Fig. 6.6c), but it is always possible (excluding border effects) to crop rectangles of given sizes $N_{i_1} \times N_{i_2}$ around the set of coordinates $\mathbf{R}^{[0]}$.

Specifically, the i-th detector is associated the set of pixels $\mathcal{Q}_k$ assigned to the k-th subimage if the following condition are verified simultaneously:

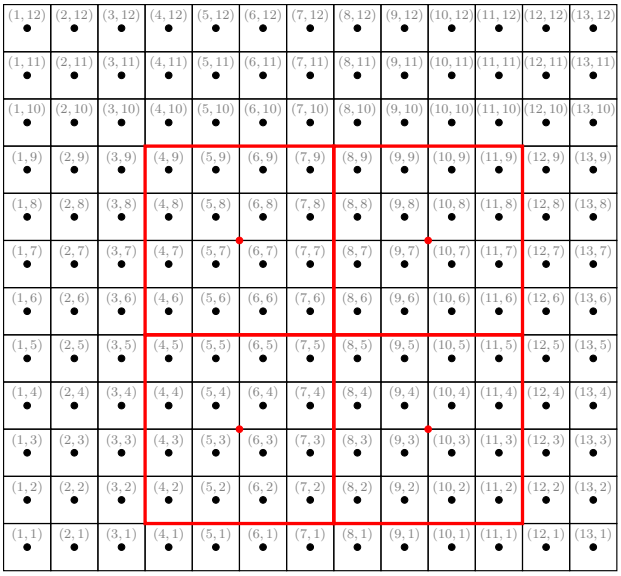
$$i \in \mathcal{Q}_k \Leftrightarrow \begin{cases} \left| r^{[t]}_{1i} - r^{[0]}_{1k} \right| & \leq \frac{N_{i_1}}{2}\,, \\ \left| r^{[t]}_{2i} - r^{[0]}_{2k} \right| & \leq \frac{N_{i_2}}{2}\,, \end{cases} \qquad \forall i \in [1,\dots,N_o]\,,\ k \in [1,\dots,N_k]\,. \tag{6.2.2}$$



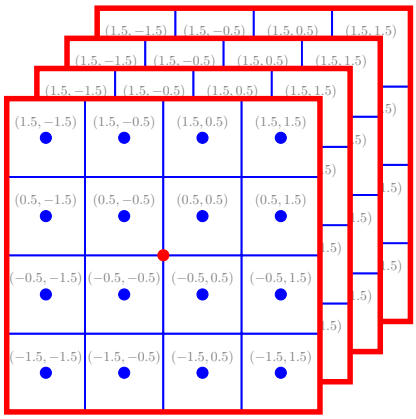
(a) Absolute coordinate system (subimages of odd sizes)



(b) Relative coordinate system (subimages of odd sizes)



(c) Absolute coordinate system (subimages of even sizes)



(d) Relative coordinate system (subimages of even sizes)

Fig. 6.6. Comparison between the absolute and relative coordinate system. On the left column, the red dots $\mathbf{R}^{[0]}$ define the centers of the subimages, which corresponds to the origins in the relative coordinate systems. These origins are in half pixel positions in the case of even length subimages. On the left column the black dots $\mathbf{R}^{[t]}$ define the centers of the photodetectors, which correspond to the blue dots $\check{\boldsymbol{\mathcal{R}}}^{[t]}$ in the second coordinate system.

Given a certain raw acquisition $\mathbf{U}^{[o]} \in \mathbb{R}^{N_{o_1} \times N_{o_2}}$, eq. (6.2.2) can guide the creation of a stacked datacube $\boldsymbol{\mathcal{U}}^{[s]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_k}$, by piling up each cropped subimage. This operation is denoted with $\boldsymbol{\mathcal{U}}^{[s]} = \text{stack}\left(\mathbf{U}^{[o]}, \mathbf{R}^{[0]}, [N_{i_1}, N_{i_2}]\right)$ and an example is shown in Fig. 6.6b.

Relative coordinates

The datacube $\boldsymbol{\mathcal{U}}^{[s]}$ of stacked subimages can be interpreted as a spatial translation operation of all subimages to a common origin. Consequently, we can define a coordinate system shared by all frontal slices $\left\{\mathbf{U}^{[s]}_{::k}\right\}_{k\in[1,\dots,N_k]}$ of $\boldsymbol{\mathcal{U}}^{[s]}$.

All centers defined in this relative coordinate system are denoted with a specific symbol $\check{\mathbf{R}}$, to differentiate them with respect to the ones defined in the previous section.

If the origin is chosen to be at the center of the subimages, then the coordinates $\check{\mathbf{R}}^{[t]} \in \mathbb{R}^{2\times N_i}$ associated with the intensity values of the subimages are composed by the following elements:

$$\begin{cases} \check{\mathbf{r}}^{[t]}_{1,((i_2-1)N_{i_1}+i_1)} &= -\frac{N_{i_1}-1}{2} + i_1\,, \\ \check{\mathbf{r}}^{[t]}_{2,((i_2-1)N_{i_1}+i_1)} &= -\frac{N_{i_2}-1}{2} + i_2\,, \end{cases} \qquad \forall i_1 \in [1,\dots,N_{i_1}]\,,\ i_2 \in [1,\dots,N_{i_2}]\ . \tag{6.2.3}$$

A typical application is to estimate a matrix of coordinates $\check{\boldsymbol{\mathcal{R}}}^{[f]} \in \mathbb{R}^{2\times N_k \times N_a}$ of N_a features (e.g., a contour) that are shared across all N_k subimages. The main assumption here is that each feature corresponds to a single replica in each subimage.

The obtained features' positions can be shifted to the original coordinate system, obtaining:

$$\boldsymbol{\mathcal{R}}^{[f]} = \check{\boldsymbol{\mathcal{R}}}^{[f]} + \mathbf{R}^{[0]}\,, \tag{6.2.4}$$

for which we remind here that $\mathbf{R}^{[0]}$ is broadcast to allow for a sum over consistent dimensions.

6.2.2 Extended sources

In a general acquisition, the scene is typically an extended sources as it is composed by multiple radiators, which for the raw ImSPOC product map to a set of replicas that cover a certain finite surface on each subimage. From these raw products, we

aim in this section to estimate a matrix $\mathbf{R}^{[0]} \in \mathbb{R}^{2\times N_k}$ of centers and, eventually, the sizes $[N_{i_1}, N_{i_2}]$ of the subimages.

We propose here a processing scheme to isolate the surfaces of the regions associated with extended areas, which we denote as RoIs. The scheme employs an approach based on mathematical morphology; we also propose two methods for the estimation of the centers of the obtained RoIs.

Definition of the regions of interest

In the case of extended sources, the definition of the center is a challenging task, as the object covers a large area on the focal plane.

In these instances, it is often possible to distinguish between **regions of interest (RoIs)**, which contain relevant information, and **background**, which is purely noisy. A naive approach to separate between these two classes involves noticing that the two classes are characterized by different intensities levels, with the RoIs associated with brighter intensities.

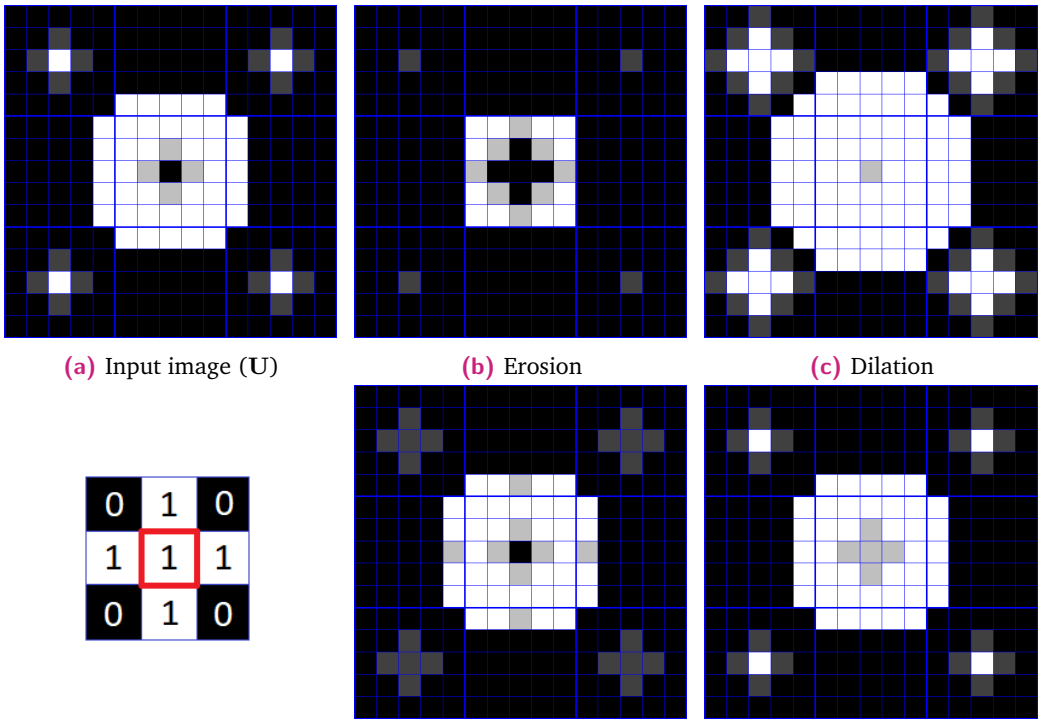
We propose here to address the determination of the RoIs with an approach based on **mathematical morphology** [115, 175, 206]. The most basic operations related to mathematical morphology consist in probing the image with a simple shape, known as **structuring element**, which defines an entity composed by a series of pixels on a grid with an associated origin.

To simplify the exposition, let us assume the structuring element $\mathbf{B} \in \mathbb{R}^{N_{d_1}\times N_{d_2}}$ is a odd length bidimensional binary mask, with its origin in the center. The generic element b_{d_1,d_2} of $\mathbf{B}$ is either a zero or a one, and its indices span the ranges $d_1 \in [-(N_{d_1}-1)/2, \dots, (N_{d_1}-1)/2]$ and $d_2 \in [-(N_{d_2}-1)/2, \dots, (N_{d_2}-1)/2]$. An example of such mask is shown in Fig. 6.7d.

The mask shifts along all position of an input image $\mathbf{U}$ and, at each overlap, the ones in the mask select a subset of the pixels of $\mathbf{U}$, according to the current position of the ones in the mask. Then, a a nonlinear operation is computed which involves the selected pixels and the result is stored at the current position of the origin. The following four basic nonlinear operations are the most common:

- **Dilation**: computes the maximum value;
- **Erosion**: computes the minimum value;
- **Opening**: an erosion followed by a dilation;
- **Closing**: a dilation followed by an erosion.

An examples of the results of each of these operations is shown in Fig. 6.7.



(a) Input image ($\mathbf{U}$) **(b)** Erosion **(c)** Dilation

(d) Structuring element ($\mathbf{B}$) **(e)** Opening (Erosion→Dilat.) **(f)** Closing (Dilation→Erosion)

Fig. 6.7. Results of mathematical morphology operation with a 3×3 cross-shaped structuring element (Fig. 6.7d), whose origin is in the center pixel (red border). The operations are applied over an image with four intensity levels (Fig. 6.7a). The products of the morphological operators show how the opening has the effect of suppressing the spurious pixels in the corners (Fig. 6.7e) and the closing to flatten the hole in the center of the image (Fig. 6.7f).

In mathematical terms, applying an erosion or a dilation on an image $\mathbf{U} \in \mathbb{R}^{N_{o_1} \times N_{o_2}}$, whose generic element is denoted with u_{i_1,i_2}, generates another image $\mathbf{U}' \in \mathbb{R}^{N_{o_1} \times N_{o_2}}$ whose elements are given by:

$$u'_{i_1,i_2} = \begin{cases} \min\limits_{[d_1,d_2]:s_{d_1,d_2}=1} u_{i_1-d_1,i_2-d_2} & \text{for an erosion}, \quad (6.2.5a) \\ \max\limits_{[d_1,d_2]:s_{d_1,d_2}=1} u_{i_1-d_1,i_2-d_2} & \text{for a dilation}. \quad (6.2.5b) \end{cases}$$

In the equation above, the elements outside of the boundaries of the image are ignored when evaluating the $\max$ or $\min$.

A quick analysis of Fig. 6.7e and 6.7f shows how the opening has the effect to suppress the spurious bright pixels, while the closing flattens darker pixels. This is the desired effect we were looking for to identify the RoIs.

More in detail, we propose the following sequence of operations to label the RoI of $\mathbf{U}$, which we denote as `GenerateROI`$(\mathbf{U})$ in Algorithm 4:

- An opening with a symmetric structuring element (either a disk or a square);
- A closing with a structuring element, with the same shape, but generally a different size;
- A thresholding operation with a given threshold t, to transform each pixel u_{i_1,i_2} in a binary value:

$$\mathrm{thres}_t\left(u_{i_1,i_2}\right) = \begin{cases} 1\,, & \text{if } u_{i_1,i_2} \geq t\,, \\ 0\,, & \text{otherwise}\,, \end{cases} \tag{6.2.6}$$

  such that its intensity levels 0 and 1 label the background and the RoIs, respectively;

This method is versatile, but requires the setup of a relatively large amount of parameters: the shape and size of the structuring elements, and the threshold level. As a rule of thumb, we suggest to choose the shape of the structuring element to approximately follow the shape of the regions, i.e. to employ a square if the subimages are approximately rectangular. Their sizes should be set up with a visual analysis: the size of the structuring element for the opening should be increased until the spurious pixels are eliminated, while that for the closing should be increased until the RoIs are fully connected. With respect to the threshold, an initial guess can be obtained by the procedure known as Otsu's method [180]. An example of such product is shown in Fig. 6.8c.

Centroid and scanline center estimation methods

We propose here two different methods for the determination of the centers of the RoI, starting from the image processed with the mathematical morphology techniques, described in previous section.

In our context, we are faced with two choices, we can either estimate the center of the the regions regardless of their position on the focal plane, for which we propose the CCE method. If this is not sufficient, and assume that the RoIs associated with each interferometer are reasonably well aligned with the geometry of the focal plane, so that the regions are approximately regularly spaced on the horizontal and vertical direction. This condition may be exploited, in the proposed SCE method, to estimate the horizontal and vertical coordinates of their centers separately.

- **Centroid center estimation (CCE)**: For this method each of the RoI is labeled according to a connectivity-8 criterion, that is if any of the pixels surrounding a white pixel is also white, they are labeled as part of the same region. The final estimation of the center of each RoI obtained by evaluating the centroid of each of the obtained regions, that is, by evaluating the average coordinate of the pixel associated with each region.

- **Scanline center estimation (SCE)** [1]: The algorithm is based on the detection of trigger conditions that define the boundaries of each set of RoIs. Let us imagine a vertical line travelling bottom to top across the image. Given the binary nature of the image, the line will eventually intersect any of the bright pixels; this is the first trigger condition. The line then continues scanning the image, until it eventually its whole length exclusively intersects a background area (black pixels); this is the second trigger condition. The coordinates of these two trigger conditions, stored as a pair, represent the boundaries of a given horizontally aligned set of RoIs. The process is repeated until the whole image is scanned and the midpoints of all pairs are stored in a vector $\mathbf{r}^{[1]} \in \mathbb{R}^{N_{k_1}}$. An equivalent procedure is repeated, scanning the image from left to right, returning a vector $\mathbf{r}^{[2]} \in \mathbb{R}^{N_{k_2}}$ of midpoints of vertically aligned RoIs. The desired matrix of center estimation $\widehat{\mathbf{R}}^{[0]}$ is finally obtained by considering all combinations of the sets $\mathbf{r}^{[1]}$ and $\mathbf{r}^{[2]}$.

Each of the two methods is described in detail in the Algorithm 4. The estimated centers $\widehat{\mathbf{R}}^{[0]}$ can also be used to find an estimation $\left[\widehat{N}_{i_1}, \widehat{N}_{i_2}\right]$ of the sizes of each subimage, which are given by:

$$\widehat{N}_{i_1} = \sum_{k=1}^{N_{k_1}-1} \left| r_{k+1}^{[1]} - r_k^{[1]} \right| , \tag{6.2.7a}$$

$$\widehat{N}_{i_2} = \sum_{k=1}^{N_{k_2}-1} \left| r_{k+1}^{[2]} - r_k^{[2]} \right| . \tag{6.2.7b}$$

As an example, both the proposed approaches were applied to an acquisition taken by the PROTO-4, in the case the device is illuminated with an extended monochromatic source. The results are shown in Fig. 6.8.

[1] The idea of the SCE method was originally proposed in the context of the "projet Intégrateur: Analyse de données hyperspectrales pour la quantification de gaz dans l'atmosphère". The project, which lasted from September 2019 to January 2020, was led by Aneline Dolet (ASI, ENSE3, Grenoble-INP) and developed by Sandrine Gayrard, Franklin Mathiot, and Matty Battista. The algorithm presented here is an expanded version based on that original code.

Algorithm 4: Centroid center estimation (CCE) and scanline center estimation (SCE).

Result:

- List of centers $\widehat{\mathbf{R}}^{[0]} \in \mathbb{R}^{2\times N_k}$ (and, consequently, the amount of RoIs N_k);
- Thumbnail Sizes: $[\widehat{N}_{i_1}, \widehat{N}_{i_2}]$ (only for the SCE).

Data:

- Acquisitions matrix $\mathbf{U} \in \mathbb{R}^{N_{o_1}\times N_{o_2}}$ with elements u_{i_1,i_2};
- Structural element $\mathbf{B}^{[o]}$ for the opening and $\mathbf{B}^{[c]}$ for the closing;
- Threshold value t.

Function `GenerateRoI`$(\mathbf{U}, \mathbf{B}^{[o]}, \mathbf{B}^{[c]}, t)$ `// Morphological operations`

$\mathbf{U} \leftarrow$ Opening (6.2.5a)+(6.2.5b) on $\mathbf{U}$ with structuring element $\mathbf{B}^{[o]}$

$\mathbf{U} \leftarrow$ Closing (6.2.5b)+(6.2.5a) on $\mathbf{U}$ with structuring element $\mathbf{B}^{[c]}$

$\mathbf{U} \leftarrow$ Thresholding on $\mathbf{U}$ with threshold t through eq. (6.2.6)

return $\mathbf{U}$

Function `CCE`$(\mathbf{U}, \mathbf{B}^{[o]}, \mathbf{B}^{[c]}, t)$ `// Centroid center estimation`

$\mathbf{U} \leftarrow$`GenerateRoI`$(\mathbf{U}, \mathbf{B}^{[o]}, , \mathbf{B}^{[c]}, t)$

Generate sets $\{\mathcal{Q}_k\}_{k\in[1,\dots,N_k]}$ of connect-8 regions of $\mathbf{U}$ with cardinality $|\mathcal{Q}_k|$

$\widehat{\mathbf{r}}^{[0]}_{:k} \leftarrow \frac{1}{|\mathcal{Q}_k|}\sum_{\mathbf{r}\in\mathcal{Q}_k}\mathbf{r}$ `// Evaluate Centroids`

$\widehat{\mathbf{R}}^{[0]} \leftarrow \{\widehat{\mathbf{r}}^{[0]}_{:k}\}_{k\in[1,\dots,N_k]}$

return $[\widehat{\mathbf{R}}^{[0]}, N_k]$

Function `SCE1D`$(\mathbf{U})$ `// SCE on a single direction`

$i' \leftarrow$ NaN `// Tracks the beginning of a RoI`

$N_{k_1} \leftarrow 0$ `// Counts the amounts of RoI per row`

Assign $[N_{o_1}, N_{o_2}]$ as the sizes of $\mathbf{U}$

for $i_1 = 1, \dots, N_{o_1}$ **do**

  if $i' =$ *NaN* ***and*** $\sum_{i_1=1}^{N_{o_1}} u_{i_1,i_2} > 0$ **then**

    $i' \leftarrow i_1$

  else if $\sum_{i_1=1}^{N_{o_1}} u_{i_1,i_2} = 0$ **then**

    $i' \leftarrow$ NaN

    $N_{k_1} \leftarrow N_{k_1} + 1$ `// Updates the count`

    $r^{[1]}_{N_{k_1}} \leftarrow (i' + i_1)/2$ `// Average between boundaries of RoI`

end

$\widehat{N}_{i_1} \leftarrow \frac{1}{N_{k_1}-1}\sum_{k=1}^{N_{k_1}-1}\left(r^{[1]}_{k+1} - r^{[1]}_{k}\right)$

$\mathbf{r}^{[1]} \leftarrow \{r^{[1]}_k\}_{k\in[1,\dots,N_{k_1}]}$

return $[\mathbf{r}^{[1]}, \widehat{N}_{i_1}, N_{k_1}]$

Function `SCE`$(\mathbf{U}, \mathbf{B}^{[o]}, , \mathbf{B}^{[c]}, t)$ `// Scanline center estimation`

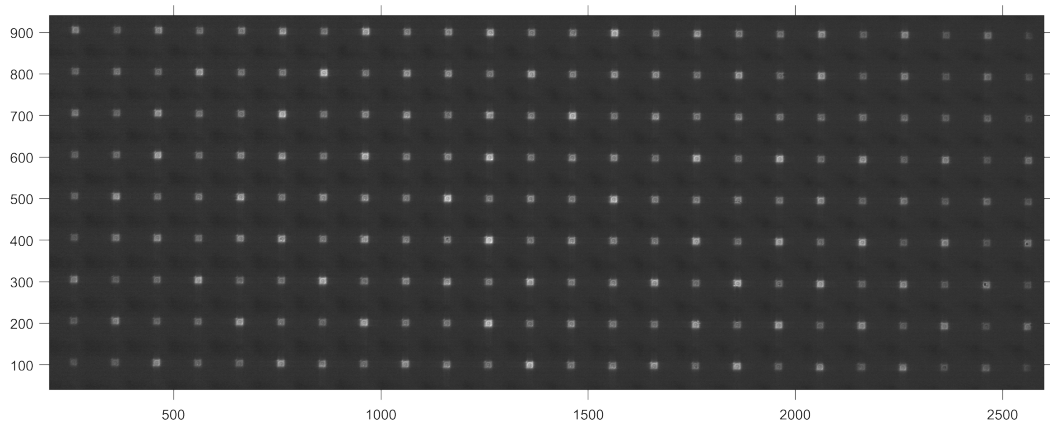
$\mathbf{U} \leftarrow$`GenerateRoI`$(\mathbf{U}, \mathbf{B}^{[o]}, \mathbf{B}^{[c]}, t)$

$[\mathbf{r}^{[1]}, \widehat{N}_{i_1}, N_{k_1}] \leftarrow$`SCE1D`$(\mathbf{U})$

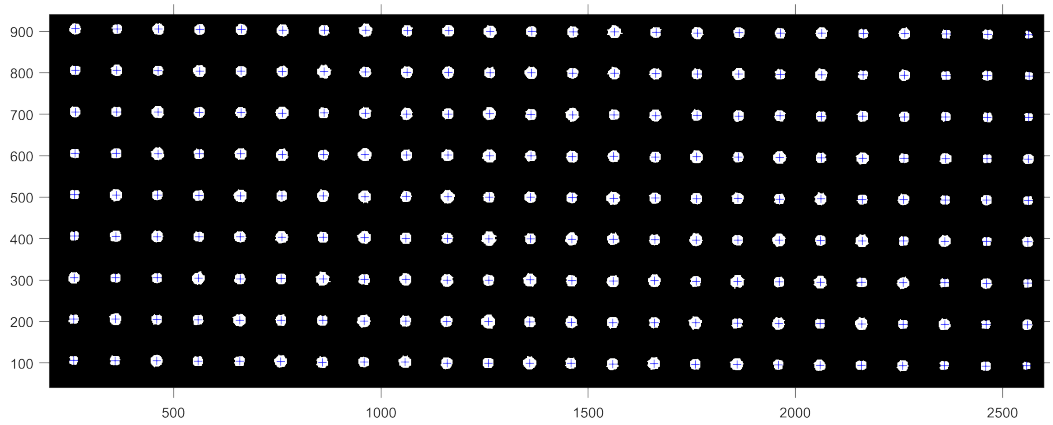
$[\mathbf{r}^{[2]}, \widehat{N}_{i_2}, N_{k_2}] \leftarrow$`SCE1D`$(\mathbf{U}^T)$ `// Repeat for the other direction`

$\widehat{\mathbf{R}}^{[0]} \leftarrow \left\{[r^{[1]}_{k_1}, r^{[2]}_{k_2}]\right\}_{k_1\in[1,\dots,N_{k_1}],\, k_2\in[1,\dots,N_{k_2}]}$ `// Assign all combinations`

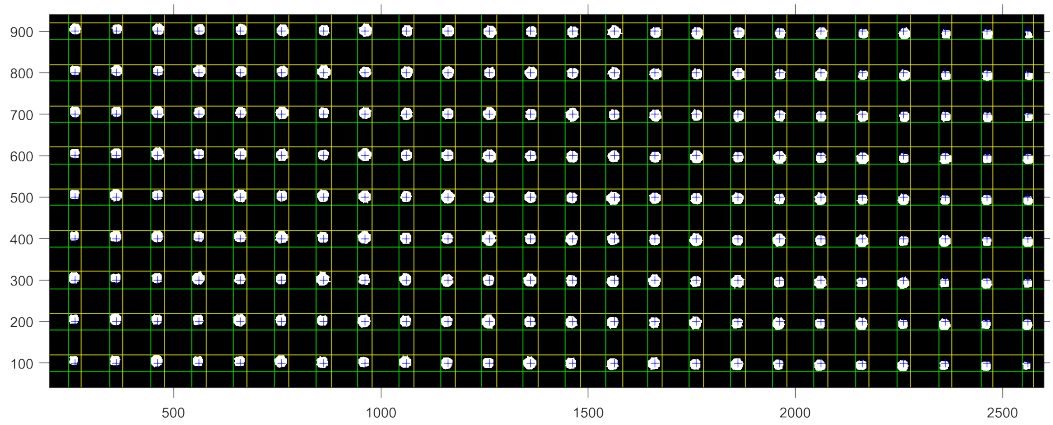
return $[\widehat{\mathbf{R}}^{[0]}, \widehat{N}_{i_1}, \widehat{N}_{i_2}]$



(a) Input image (PROTO-1)



(b) CCE



(c) SCE

Fig. 6.8. Estimation of the centers for a PROTO-1 acquisition for an illumination with a monochromatic extended source. The raw image (Fig. 6.8a) is first processed with morphological operations described in Section 6.2.2; the opening and closing use a disk-shaped structuring element with a diameter of 2 and 12 px, respectively, while the thresholding employs a threshold level equal to 10% of the maximum recorded intensity. The centers obtained determined with the CCE (Fig. 6.8b) and the SCE (Fig. 6.8c) and shown with blue crosses. For the SCE, the green and yellow lines denote opposing boundaries of the estimated sets of aligned RoIs.

6.2.3 Point sources

In this section we consider the situation of a laser beam incident to an ImSPOC device, acting as a point source, with the aim of estimating the position of each replica of the laser spot illuminating the focal plane. This is a special case of finding a set of spatial features, which are common across the N_k subimages. Each of these features are measured separately with N_a different acquisitions.

Let $\boldsymbol{\mathcal{U}}^{[s]} \in \mathbb{R}^{N_{i_1} \times N_{i_2} \times N_k \times N_a}$ be a set of N_k stacked subimages from of a set of N_a acquisitions and $\check{\boldsymbol{\mathcal{R}}}^{[f]} \in \mathbb{R}^{2 \times N_k \times N_a}$ be the relative positions of the center spots, as described in Section 6.2.1.

We propose here to estimate $\check{\boldsymbol{\mathcal{R}}}^{[f]}$ by solving a problem of nonlinear regression, through the fit of a Gaussian function.

Gaussian-fit center estimation method

The proposed **Gaussian-fit center estimation (GCE)** defines a method to find the position of a laser spots $\check{\boldsymbol{\mathcal{R}}}^{[f]}$ on different subimages by fitting a bidimensional Gaussian function to the available intensity levels. In particular, we aim to estimate of a matrix of parameters $\widehat{\boldsymbol{\mathcal{B}}} \in \mathbb{R}^{4 \times N_k \times N_a}$. For the k-th subimage of the l-th acquisition, a vertical slice $\widehat{\boldsymbol{\beta}}_{:kl}$ is composed by four elements, which are composed by the following four characteristics of the Gaussian function:

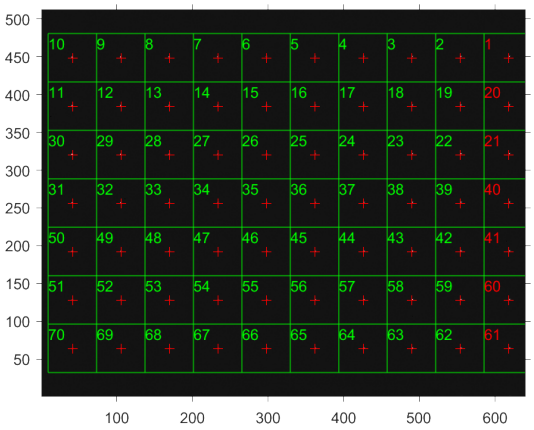
- $\beta_1 \in [-N_{i_1}/2, N_{i_1}/2] \subseteq \mathbf{R}$ is the vertical coordinate of its center;
- $\beta_2 \in [-N_{i_2}/2, N_{i_2}/2] \subseteq \mathbf{R}$ is the horizontal coordinate of its center;
- $\beta_3 \in \mathbf{R}^+$ is its peak intensity;
- $\beta_4 \in \mathbf{R}^+$ is its standard deviation (STD), which we assume for simplicity equal in the vertical and horizontal direction.

The solution is then given by the following minimization:

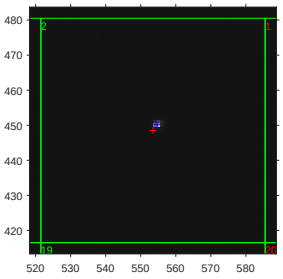
$$\widehat{\boldsymbol{\beta}}_{:kl} = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{N_i} \left(\beta_3 \exp\left(-\frac{\|\mathbf{r}_{:k}^{[s]} - [\beta_1; \beta_2]\|_2^2}{2\beta_4^2} \right) - s_{kli} \right)^2 . \tag{6.2.8}$$

where s_{k1i} is the generic element of the permuted lexicographic representation $\boldsymbol{\mathcal{S}} = \text{reshape}(\mathbf{U}^{[s]}) \in \mathbb{R}^{N_k \times N_a \times N_i}$ of the stacked image.

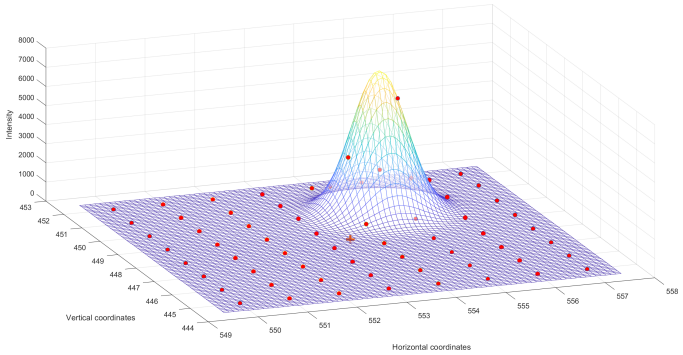
The above functional has a nonlinear dependency with the set of parameters; this is a problem of nonlinear regression, that can be solved via the Gauss-Newton algorithm (GNA) [78], described in Section2.3.3. With regard to the initialization, β_1 and



(a) GCE for PROTO-4



(b) Detail for thumbnail 2



(c) Gaussian Fit for thumbnail 2

Fig. 6.9. Results obtained by applying the GCE protocol to a laser spot acquisition taken with the PROTO-4. The estimated centers and their initial guesses are marked with blue and red crosses, respectively. A detail for the interferometer 2 (Fig. 6.9b) is used to show how the associated Gaussian bell shape can fit the acquisition samples (red dots in Fig. 6.9c).

β_2 can be initialized as the coordinates of the brightest spot, β_3 as the maximum intensity in the subimage and β_4 can be set to 1 for simplicity.

Finally, the estimation center coordinates are given by: $\mathbf{r}_{:kl}^{[f]} = [\widehat{\boldsymbol{\beta}}_{1kl};\ \widehat{\boldsymbol{\beta}}_{2kl}]$. Additionally, the estimated peak intensity $\widehat{\beta}_{3kl}$ can be compared to a user-defined threshold to assess if the laser spot was actually present in the subimage.

6.2.4 Discussion of the proposed methods

This section is not intended as a systematic discussion to the problem of center estimation, as no experimental result is provided. However, we still believe that the discussed techniques are of interest to the end user for the automatization of the processing pipeline of ImSPOC and to provide some feedback based on our practical experience.

As a rule of thumb, for extended sources, we suggest to firstly employ the SCE method to have a reasonable estimation of the subimage sizes and then to refine the center position of the sources with the CCE method. The refinement can be obtained by picking the closest centers to the ones determined with the SCE method, which also allows to arrange them in a systematic order.

The point source case can also be tackled with the methods developed for extended sources, but we suggest to skip the morphological opening to avoid suppressing the relevant information. For the SCE method in particular, it may also be useful to perform a dilation on the input image, to aid the vertical alignment of the RoIs. This may be useful to estimate the initial partition of the subimages, which can then be used as pre-processing step for the GCE method.

6.3 Co-registration of subimages

In this section, we describe the processing operations to transform a datacube of raw acquisitions $\mathcal{U}^{[s]}$ (L0 product) into a co-registered datacube $\mathcal{U}^{[y]}$ (L1 product).

The main goal is to correct the spatial distortions introduced by eventual asymmetries in the optical path travelled by the incident rays for each element of the structure, by aligning each of the subimages collected in the datacube.

We propose to address this problem by estimating a polynomial geometry transformation function and resampling over a regular grid in Section 6.3 and we test its effectiveness with real acquisitions in Section 6.3.3.

6.3.1 Problem statement

The operating principle of ImSPOC relies on an array of optical elements with different characteristics. The ideal behaviour of the instrument requires that, for a given solid angle of incidence (or in turn for a given portion of the targeted scene),

it is possible to uniquely identify an associated sample on the focal plane for each of the N_k interferometers.

We employ here the relative coordinate system described in Section 6.2.1, obtained by shifting the k-th subimage by $\mathbf{r}_{:k}^{[0]}$, which here denotes the intersection between the optical axis of the k-th lenslet and the focal plane.

Let us partition the FoV into N_a sufficiently small solid angles $\{\Omega_i\}_{i\in[1,\dots,N_a]}$ and let $\check{\boldsymbol{\mathcal{R}}}^{[f]} \in \mathbb{R}^{2\times N_k\times N_a}$, be the focal spots of the incident rays associated with the set of solid angles [2]. The alignment condition is then given by the following property:

$$\check{\mathbf{r}}_{:kl}^{[f]} = \mathbf{r}_{:k'l}^{[f]} \qquad \forall k \in [1,\dots,N_k]\,,\; l \in [1,\dots,N_a]\;, \tag{6.3.1}$$

where k' is the index of a subimage of reference.

The position of each of these features can be rewritten as a relative shift $\boldsymbol{\mathcal{R}}^{[d]}$ with respect to the position of the corresponding focal spot in the reference subimage. This description is given by:

$$\check{\boldsymbol{\mathcal{R}}}^{[d]} = \check{\boldsymbol{\mathcal{R}}}^{[f]} - \check{\mathbf{R}}_{:k':}^{[f]}\,, \tag{6.3.2}$$

where k' is the index of the reference subimage. With this description, eq.(6.3.1) is equivalent to impose all elements of $\check{\boldsymbol{\mathcal{R}}}^{[d]}$ to be equal to zero.

As seen in Section 5.5.6, verifying this condition allows to discretize the direct model which describes the optical transformation performed by the ImSPOC device.

Unfortunately, this condition is unrealistic to be verified in practice for real devices, as differences between the geometry of the lenslets and parallax effects introduce misalignments across different subimages.

Misalignments due to parallax are exploited by optical devices such as plenoptic cameras [87] to extract depth information of a close field acquisition of the scene and the recognition of common features across different subimages is known as **correspondence problem** [95]. The final target of our context is quite different, as we are interested here to identify samples of the interferogram which are related to the same portion of the scene, but the necessity to identify common features is shared by both scenarios. Conversely, we are not interested in the information provided by the amount of misalignment itself, which we aim to fully compensate with a co-registration procedure.

[2] The case of spectral aberrations, which are due to a different directional response of the optical system at different wavelengths is considered outside the the scope of this work, but it could be possible characterize them through Zernike's polynomials [71].

The main target is then to define a transformation $\mathbb{T}_s(\cdot)$ to generate a simulated acquisition $\boldsymbol{\mathcal{U}}^{[y]} = \mathbb{T}_s(\boldsymbol{\mathcal{U}}^{[s]})$, where the the new measure $\boldsymbol{\mathcal{R}}^{[d]}$ is approximately identically equal to zero. We will define the proposed procedure in the next section.

6.3.2 Point mapping calibration

The scope of this work is limited to **point mapping** alignment procedures, for which the common features that are shared across different subimages are made up of point objects. An example of such calibration procedure, that is specifically used in this thesis, consists in a setup with a laser beam that is incident to the input plane of the ImSPOC prototype under test. The laser beam is placed over a moving platform so that it is possible to vary its angle of incidence; this allows to record N_a different acquisitions, arranged in a tensor $\boldsymbol{\mathcal{U}}^{[s]} \in \mathbb{R}^{2\times N_k \times N_a}$, each with a different orientation of the laser beam.

For each acquisition, a laser spot illuminates each of the subimages, which allows to estimate the vertical and horizontal misalignments introduced by the not perfectly matching geometry of the microlenses and by other effects.

The full procedure involves the following steps [95], as it was already described in Section 3.4.2:

- **Center Estimation**: which involes estimating the position $\boldsymbol{\mathcal{R}}^{[d]}$ of the illuminated spots on each subimage, with respect to a reference subimage. We propose to perform this step here with the GCE method described in 6.2.3.
- **Determination of a transformation function**: which involves the inference of parameters of an analytic function that characterizes the geometric transformation of the coordinate system to be aligned with the reference one. More specifically, we propose here to define a polynomial bidimensional function $\mathbf{p}_k : \mathbb{R}^2 \to \mathbb{R}^2$ of order N_m:

$$\mathbf{p}_k\left(\check{\mathbf{r}}^{[f]}_{:kl}, \boldsymbol{\beta}_{::k}\right) = \sum_{m=0}^{N_m}\sum_{n=0}^{N_m-m} \beta_{mnk}\ \left(\check{r}^{[f]}_{1kl}\right)^m \left(\check{r}^{[f]}_{2kl}\right)^n , \tag{6.3.3}$$

  such that the each function of the set $\{\mathbf{p}_k\}_{k\in[1,\dots,N_k]}$ is used to fit a transformation from a generic position $\check{\mathbf{r}}^{[f]}_{:kl}$ of the k-th subimage to the displacement

$\check{\mathbf{r}}_{:kl}^{[d]}$ with respect to the reference image. Consequently the estimation $\widehat{\boldsymbol{\beta}}$ of the coefficients $\boldsymbol{\beta}_{::k} = \{\beta_{mnk}\}_{(m,n)\in\mathbb{N}^2:m+n\leq N_m}$ of $\mathbf{p}_k$ may be inferred as:

$$\widehat{\boldsymbol{\beta}}_{::k} = \arg\min_{\boldsymbol{\beta}_{::k}} \sum_{l=1}^{N_a} \left\| \mathbf{p}_k \left(\check{\mathbf{r}}_{:kl}^{[f]}, \boldsymbol{\beta}_{::k} \right) - \check{\mathbf{r}}_{:k'l}^{[d]} \right\|_2^2 . \tag{6.3.4}$$

This is once again a problem of nonlinear regression, as which is suitable to be solved with the Levenberg-Marquardt algorithm [168].

- **Resampling**: Once the geometry of the transformation is known, one can determine the coordinate positions $\check{\mathbf{R}}^{[s]} \in \mathbb{R}^{2\times N_k \times N_i}$ of the samples to interpolate in the target subimage, corresponding to the regular grid positions $\check{\mathbf{R}}^{[t]}$ of the reference subimage.
  Specifically, the elements of $\check{\mathbf{R}}^{[s]}$ are given by:

$$\check{\mathbf{r}}_{:kl}^{[s]} = \mathbf{p}_k \left(\check{\mathbf{r}}_{:kl}^{[t]}, \widehat{\boldsymbol{\beta}}_{::k} \right) + \check{\mathbf{r}}_{:kl}^{[t]} , \qquad \forall k \in [1, \dots, N_k] , \, l \in [1, \dots, N_i] . \tag{6.3.5}$$

  As the the samples of the target subimage are arranged over a regular grid, the interpolation can be performed with one of the techniques described in Section 3.4.2 for grid interpolation, e.g., with a cubic spline, bicubic kernel, and other similar approaches. [3].

6.3.3 Experimental results

Experimental setup

The pipeline of spatial calibration is applied in this section for the registration of subimages from acquisitions taken with the PROTO-4, which is the only available prototype which is accompanied with a full characterization of its spatial response.

The point mapping setup is made up of 648 raw images, obtained by illuminating the input plane with a laser beam with different orientations. The orientations were chosen to span a relatively large set of incidence angles for a reasonable coverage of the complete FoV of the instrument. From the perspective of the detectors, the laser beam is indistinguishable from a radiation generated by a point source at infinite distance, which is a reasonable simulation of a far field acquisition.

[3] The function $\mathbf{p}_k$ defines a transformation from the coordinate system of the k-th subimage to that of the reference subimage. If we performed the reverse transformation, this would not allow to perform a resampling over a regular grid.

The center positions $\mathbf{R}^{[0]}$ of the subimages were estimated from a preliminary acquisition where the laser beam was pointed perpendicularly to the input plane. An area around the reference subimage was cropped to determine the position $\mathbf{r}^{[0]}_{:k'}$ of laser spot with the GCE method. The remaining center positions are set to be regularly spaced over the FPA, so that it can be segmented into 70 non-overlapping rectangular subimages with no gaps in between. The reference subimage corresponds to the index $k' = 35$, which is located around the middle position of the FPA.

The full set of 648 acquisitions is then stacked into a datacube $\boldsymbol{\mathcal{U}}^{[s]} \in \mathbb{R}^{64\times64\times70\times648}$, and the position of focal spots of the laser beams was estimated with the GCE method for each subimage. The results were arranged in a tridimensional array $\check{\boldsymbol{\mathcal{R}}}^{[f]} \in \mathbb{R}^{2\times70\times648}$.

Among the elements of $\check{\boldsymbol{\mathcal{R}}}^{[f]}$, a given pair of coordinate $\check{\mathbf{r}}^{[f]}_{:kl}$ is labeled as **outlier** if it does not satisfy both of the following conditions:

- the Euclidean distance from the reference is below 4 units, or in other terms if:

$$\left\| \check{\mathbf{r}}^{[f]}_{:kl} - \check{\mathbf{r}}^{[f]}_{:k'l} \right\|_2 = \left\| \check{\mathbf{r}}^{[d]}_{:kl} \right\|_2 \leq 4\,; \tag{6.3.6}$$

- the peak intensity estimated with the GCE is less than 2% of the maximum recorded peak intensity.

The set of polynomial geometry transformation functions $\{\mathbf{p}_k\}_{k\in[1,\dots,N_k]}$ is then obtained by solving eq. 6.3.4, ignoring the outliers. In our experiments, we tested the estimation of $\mathbf{p}_k$ with polynomial degree N_m ranging from 0 up to 5, in order to find the best compromise between accuracy of the geometry transformation and avoiding the overfitting problem.

A visual representation of the estimated geometry transformation function is given in Fig. 6.11, expressed in the components of vertical and horizontal shift. Fig. 6.11c to 6.11d showcase the limitations of a linear fitting function, which is unable to analytically simulate the behaviour of the registered samples (Fig. 6.11a and 6.11b), in comparison to the case of $N_m = 3$ (Fig. 6.11e to 6.11f). These considerations will be justified analytically in the next sections.

We now have all the ingredients necessary to setup the calibration procedure $\mathtt{T}_s(\cdot)$ described in Section 6.3.2 on any raw image, which will be tested on two different scenarios in the following sections. In those experiments, we test the resampling with the following list of kernels:

- nearest neighbour,
- linear,

- bicubic,
- cubic spline.

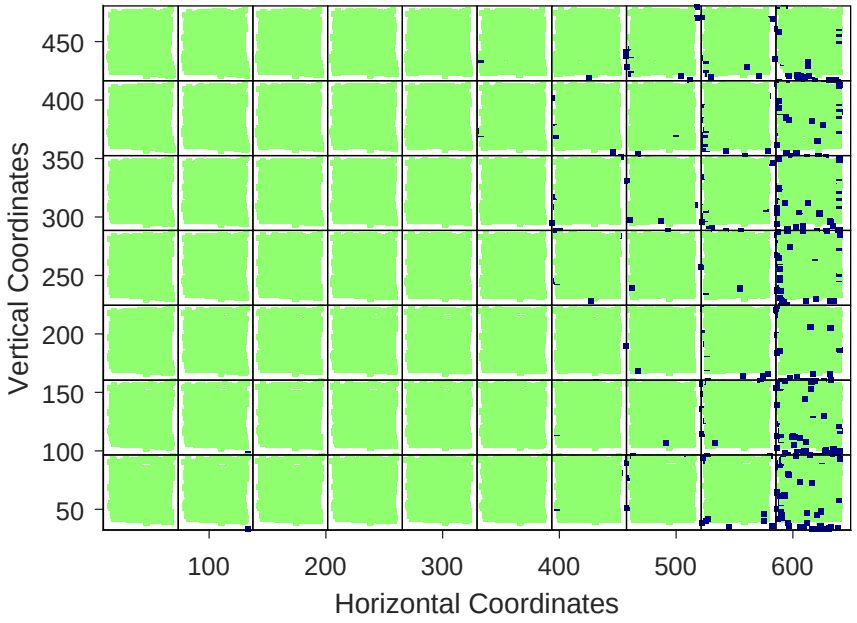
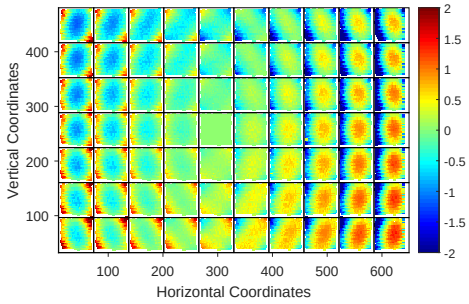
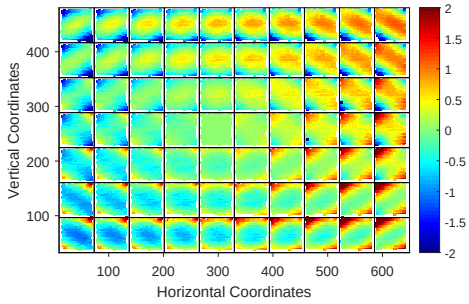


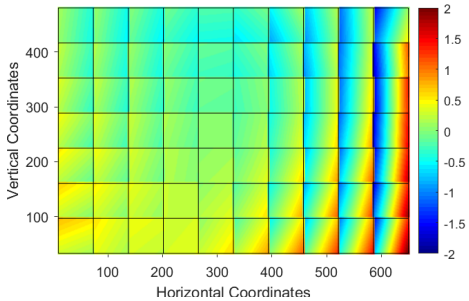
Fig. 6.10. Position of the laser spots of the 648 acquisitions estimated with the GCE method. The centers labeled in blue were labeled as outliers.



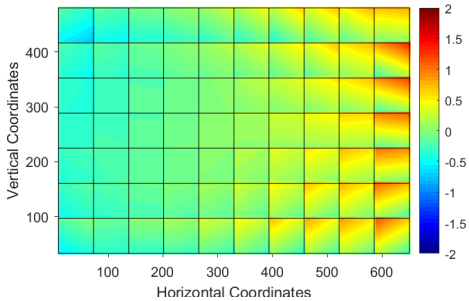
(a) Horizontal displacement (not calibrated)



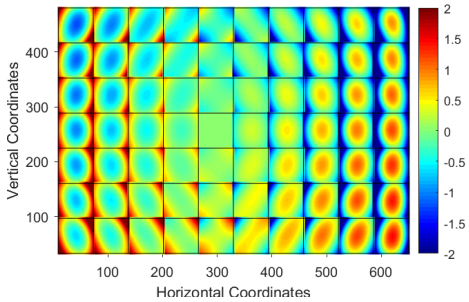
(b) Vertical displacement (not calibrated)



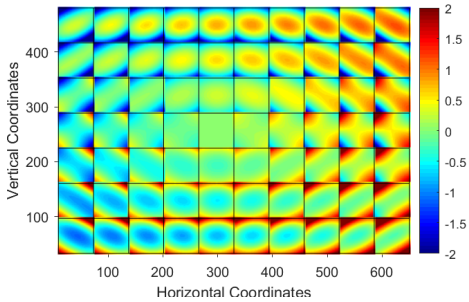
(c) Horizontal displacement fit function (polynomial degree: 1)



(d) Vertical displacement fit function (polynomial degree: 1)



(e) Horizontal displacement fit function (polynomial degree: 3)



(f) Vertical displacement fit function (polynomial degree: 3)

Fig. 6.11. In Fig. 6.11a and 6.11a, the measured amount of horizontal and vertical misalignment, expressed in unit values, with respect to the central thumbnail, before the calibration procedure. The remaining figures show the two components of the inferred geometry transformation function, for different polynomial degrees. Green dots denote close to no misalignment, while red and blue dots denote a strong misalignment in a given direction.

Calibration dataset

The co-registration procedure is firstly applied to the datacube of calibration acquisitions themselves. This is the training set, so we intend this experiment just as a double check that the laser spots are shifted properly over the desired positions in the processed datacube. These new center positions are calculated once again with the GCE method.

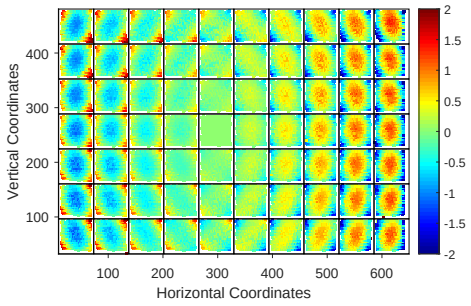
The quantitative validation is performed by evaluating the mean Euclidean distance (MED) index, which we define as:

$$\text{MED} = \sum_{l=1}^{N_a} \sum_{k=1}^{N_k} \left\| \check{\mathbf{r}}_{:kl}^{[d]} \right\|_2 , \tag{6.3.7}$$

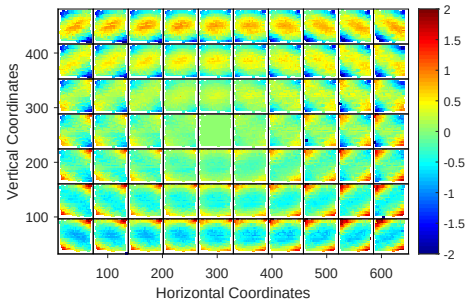
where $\check{\mathbf{r}}_{:kl}^{[d]}$ defines the estimated displacement of the laser spot of k-th subimage with respect to the reference in the l-th processed acquisition. The results are given in Table 6.3 for all the parameters under test; the GCE allows to also define a new set of outliers, which are also given in the table. For a fair comparison, the outliers of the raw image are excluded from the computation of the MED in eq. (6.3.7), but the new outliers generated by the processing are included.

The analysis of the MED shows that an increasing degree of the fitting polynomial yields a more accurate approximation of the training set, but this eventually comes at a cost of overfitting the data, as shown in the next section. The qualitative results also seem to imply that the linear resampling has the best performances, but this result must be taken with a grain of salt. This effect is most likely due to the fact that a linear transformation introduces less spatial distortions in the Gaussian bell that characterizes the laser spot, which simplifies the GCE procedure to assess the center positions.

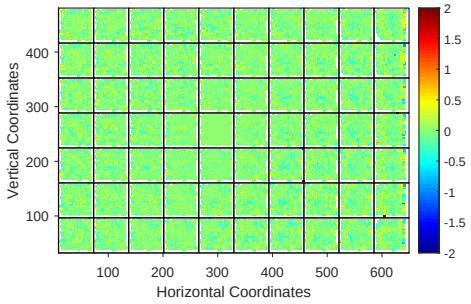
A visual comparison of the results is given in Fig. 6.12, which also shows the beneficial effect of polynomials of degree superior to 1, especially in comparison to the unprocessed data of Fig. 6.11.



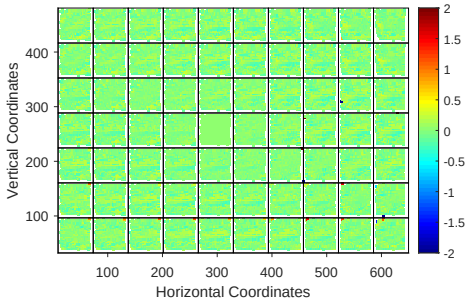
(a) Co-registered horizontal displacement. Polynomial degree: 1, resampling: spline



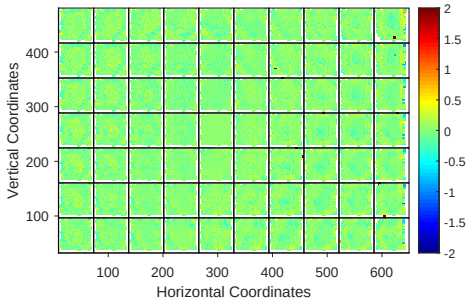
(b) Co-registered vertical displacement. Polynomial degree: 1, resampling: spline



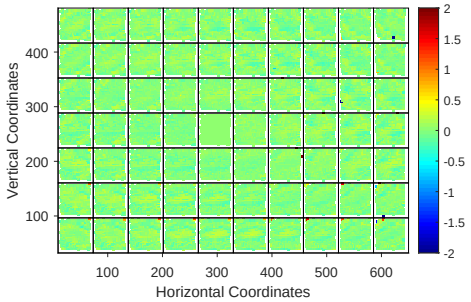
(c) Co-registered horizontal displacement. Polynomial degree: 3, resampling: spline



(d) Co-registered vertical displacement. Polynomial degree: 3, resampling: spline



(e) Co-registered vertical displacement. Polynomial degree: 3, resampling: linear



(f) Co-registered vertical displacement. Polynomial degree: 3, resampling: linear

Fig. 6.12. Position of the centers of the laser spots, after the images were spatially aligned with the parameters indicated in the small captions. The left and right column refer to horizontal and vertical misalignment with respect to the reference thumbnail, respectively. The intensity scale is the same as Fig. 6.11a and 6.11b.

Table 6.3. Quantitative results for the 648 laser beam acquisitions co-registration procedure. The given values are the average and STD with respect to the subimages. The new outliers are included in the computation of the MED. Best results are in bold, second best are underlined.

Poly. degree N_m	**Resampling**	**MED** (Ideal value: 0)	**New outliers**
0	Spline	0.6173 ± 0.4899	8
1	Spline	0.5633 ± 0.4214	10
2	Spline	0.1555 ± 0.0967	6
3	Spline	0.1476 ± 0.0902	6
4	Spline	0.1428 ± 0.0872	5
5	Spline	0.1406 ± 0.0858	8
3	Linear	$\mathbf{0.1249 \pm 0.0851}$	11
3	Cubic	$\underline{0.1353 \pm 0.0885}$	8
3	Near. Neig.	0.3382 ± 0.1689	18
Raw unprocessed data		0.6703 ± 0.5325	0

In situ datasets

The same co-registration procedure is then applied on 91 in situ acquisitions, made up of 26 shots of a "Landscape" scene, 47 shots of a "Mountain" scene; and 18 shots of a "Sunny sky" scene.

The quality of the alignment is assessed by evaluating the structural similarity (SSIM) [229] between the reference and the co-registered subimage. This validation was limited to a subset area of 56×56 pixels, as the spatial domain of the processed subimages is slightly restricted after the stretching needed to align them with the reference, especially in the corner areas shown in Fig. 6.13. The cut was chosen arbitrarily to reasonably encompass a common overlapping area. While the SSIM provides a great metric to evaluate the alignment of common features across different replicas of the image, it is worth noting that we do not target its ideal value of 1 as each slice of the calibrated datacube has to provide different information on the input spectrum.

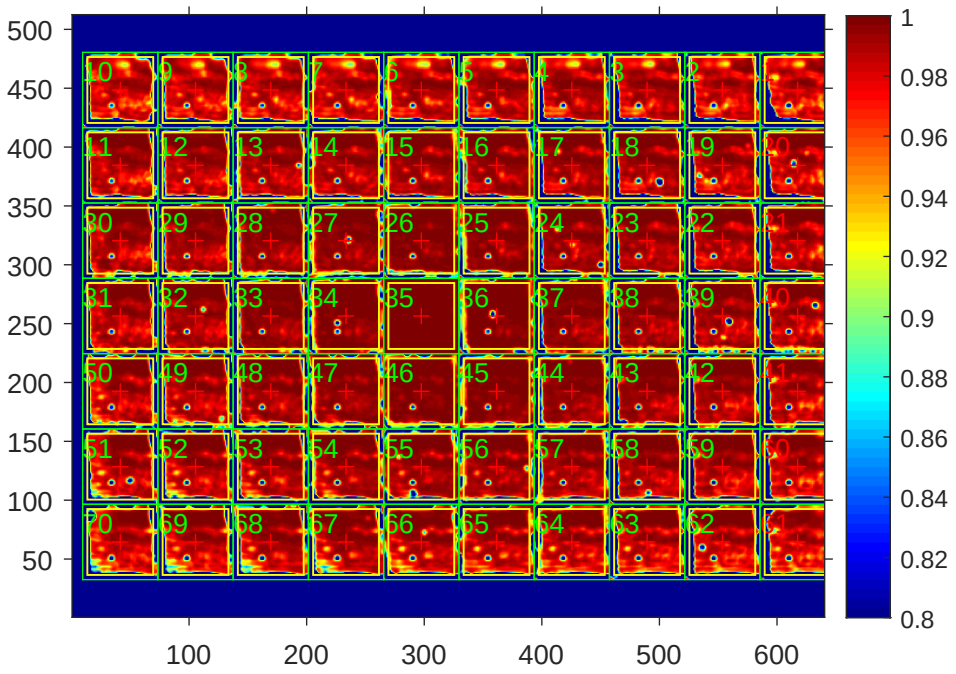


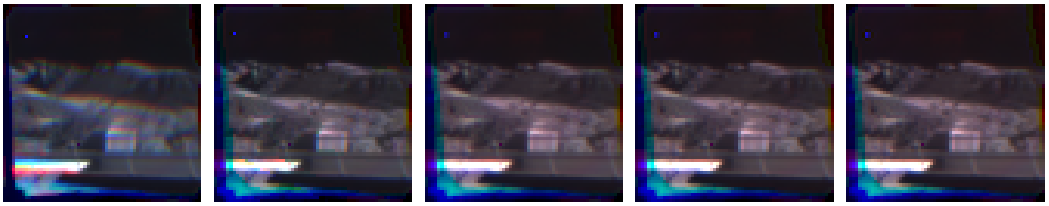
Fig. 6.13. SSIM map associated with the first in situ acquisition after the proposed subimage co-registration procedure. The red and blue zones are associated with higher and low structural similarities, respectively, with regards to the central thumbnail (index 35). The yellow frame identifies the boundaries of the regions used for the calculation of the global indices in Table 6.4. Subimage pixels falling outside of the focal plane are assumed to be equal to 1.

The average value of the computed SSIM with respect to all the subimages and all the acquisitions are given in Table 6.3. The analysis of the SSIM shows that the quadratic degree polynomial is the minimum requirement to appropriately describe the geometry of the transformation, with slowly degrading performances for degrees above 4. The cubic spline also proves to be the leading method of resampling for in situ acquisitions, although all kernels, with the exclusion the nearest neighbour, perform reasonably well.

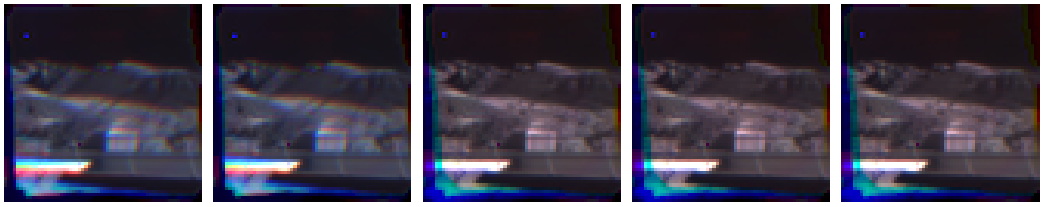
A visual comparison of the co-registered subimages for a landscape acquisition is shown in Fig. 6.14, 6.15, and 6.16 for the "Landscape", "Mountain" and "Sunny sky" dataset. The aligned subimages are represented as red green blue (RGB) channels to provide a visual feedback on their alignment. For all the considered datasets, it can be immediately verified how the aberrations across the borders of the image tend disappear with a sufficiently large degree polynomial geometric transformation and with at least a linear resampling.

Table 6.4. Quantitative results on the 91 in situ acquisitions for the subimage registration experiments described in Section 6.3.3. The given values are the mean and STD with respect to both the acquisitions and the subimages for each of the three datasets. The value N_m defines the degree of the geometry transformation polynomial function. Best results are in bold, second best are underlined.

Dataset:		**Landscape**	**Mountain**	**Sunny sky**
Number of acquisitions:		26	47	18
N_m	Resampling	SSIM (Ideal value: 1)		
0	Spline	0.9324 ± 0.0359	0.9382 ± 0.0284	0.9891 ± 0.0057
1	Spline	0.9350 ± 0.0338	0.9371 ± 0.0310	0.9871 ± 0.0070
2	Spline	$\mathbf{0.9633 \pm 0.0178}$	$\underline{0.9586 \pm 0.0180}$	0.9969 ± 0.0021
3	Spline	0.9632 ± 0.0180	$\mathbf{0.9587 \pm 0.0179}$	$\mathbf{0.9970 \pm 0.0020}$
4	Spline	0.9623 ± 0.0184	0.9579 ± 0.0183	0.9969 ± 0.0020
5	Spline	0.9618 ± 0.0184	0.9575 ± 0.0181	0.9969 ± 0.0018
3	Bicubic	$\underline{0.9633 \pm 0.0178}$	0.9585 ± 0.0179	$\underline{0.9969 \pm 0.0020}$
3	Linear	0.9623 ± 0.0174	0.9567 ± 0.0176	0.9968 ± 0.0018
3	Near. Neig.	0.9479 ± 0.0192	0.9397 ± 0.0209	0.9939 ± 0.0030
Raw unprocessed data		0.9315 ± 0.0376	0.9286 ± 0.0332	0.9890 ± 0.0058

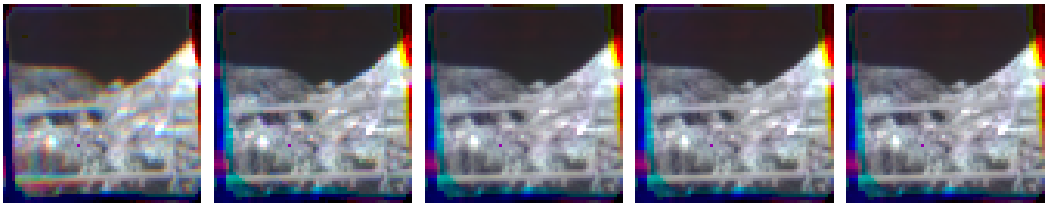


(a) Raw unprocessed data | **(b)** Poly. degree: 3, resampling: near. neig. | **(c)** Poly. degree: 3, resampling: linear | **(d)** Poly. degree: 3, resampling: bicubic | **(e)** Poly. degree: 3, resampling: spline

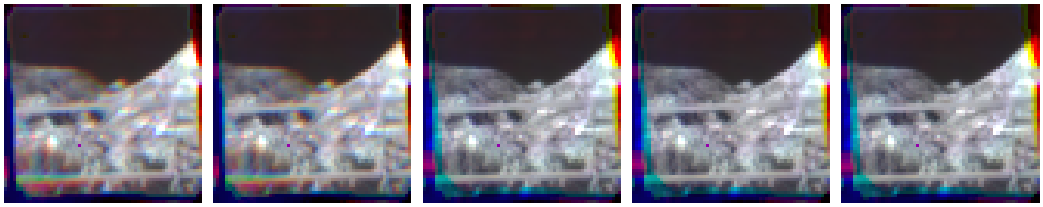


(f) Poly. degree: 0, resampling: spline | **(g)** Poly. degree: 1, resampling: spline | **(h)** Poly. degree: 2, resampling: spline | **(i)** Poly. degree: 4, resampling: spline | **(j)** Poly. degree: 5, resampling: spline

Fig. 6.14. Visual representation of the aligned thumbnails, with different calibration methods, for the first acquisition in the "Landscape" dataset. The subimages identified by the indices 2, 35, and 70 are visualized as RGB bands.



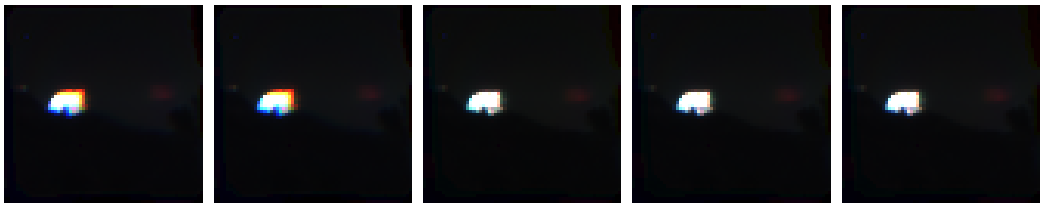
(a) Raw unprocessed data | **(b)** Poly. degree: 3, resampling: near. neig. | **(c)** Poly. degree: 3, resampling: linear | **(d)** Poly. degree: 3, resampling: bicubic | **(e)** Poly. degree: 3, resampling: spline



(f) Poly. degree: 0, resampling: spline | **(g)** Poly. degree: 1, resampling: spline | **(h)** Poly. degree: 2, resampling: spline | **(i)** Poly. degree: 4, resampling: spline | **(j)** Poly. degree: 5, resampling: spline

Fig. 6.15. Visual representation of the aligned thumbnails, with different calibration methods, for the first acquisition in the "Mountain" dataset. The subimages identified by the indices 2, 35, and 70 are visualized as RGB bands.

(a) Raw unprocessed data | **(b)** Poly. degree: 3, resampling: near. neig. | **(c)** Poly. degree: 3, resampling: linear | **(d)** Poly. degree: 3, resampling: bicubic | **(e)** Poly. degree: 3, resampling: spline



(f) Poly. degree: 0, resampling: spline | **(g)** Poly. degree: 1, resampling: spline | **(h)** Poly. degree: 2, resampling: spline | **(i)** Poly. degree: 4, resampling: spline | **(j)** Poly. degree: 5, resampling: spline

Fig. 6.16. Visual representation of the aligned thumbnails, with different calibration methods, for the first acquisition in the "Sunny sky" dataset. The subimages identified by the indices 2, 35, and 70 are visualized as RGB bands.

6.4 Model Characterization

This section's goal is to present the algorithms aimed at the inference of the parameters of the transfer matrix which characterizes the optical transformation performed by an ImSPOC device. The problem is introduced in Section 6.4.1, the algorithms for the parametric inference are the topic of Section 6.4.3, and the results of the related experiments are given in Section 6.4.4.

6.4.1 Problem statement

The manufacture of an ImSPOC device follows a series of design choices, which aim to satisfy the requirements of the user application. It is however necessary, especially for initial prototypes, to verify that the optical transformations performed by the real device follow, within a certain degree of accuracy, the desired behaviour.

The procedure of model characterization aims at estimating the characteristics of the direct model that describes the optical transformations of the device. The characterization is not only useful to check that the device operates as intended. In fact, as it will be discussed in Section 6.5, an accurate knowledge of the direct model can dramatically improve the performances for the interferogram inversion algorithms.

A typical characterization procedure consists in illuminating the device with a set of N_a known spectra and measure the intensity levels associated with each interferometer over the illuminated areas. Formally, we define the following matrices:

- $\mathbf{X}_{::i} \in \mathbb{R}^{N_b \times N_a}$ is a matrix of input radiant fluxes associated with a given incident solid angle Ω_i, for which the l-th column $\mathbf{x}_{:\mathbf{li}}$ is a given input spectrum, made up of N_b samples.
- $\mathbf{Y}_{::i} \in \mathbb{R}^{N_k \times N_a}$ is a matrix of spatially co-registered acquisitions, such that the measurements are all due to a specific solid angle of incidence Ω_i. The l-th column $\mathbf{y}_{:li}$ defines the interferogram measurement for the l-th acquisition; its k-th sample y_{kli} is due to the optical path difference (OPD) introduced by the k-th interferometer [4].

[4] $\mathbf{Y}_{::i}$ is the i-th frontal slice of the co-registered datacube $\boldsymbol{\mathcal{Y}} = \text{reshape}\left(\boldsymbol{\mathcal{U}}^{[y]}\right)$ obtained in Section 6.3.

We propose to approach this problem with a linear model, where the acquisition is described with the following stochastic process:

$$\mathbf{Y}_{::i} = \mathbf{A}_{::i}\mathbf{X}_{::i} + \mathbf{E}_{::i}\,, \tag{6.4.1}$$

where $\mathbf{A}_{::i} \in \mathbb{R}^{N_k \times N_b}$ is a transfer matrix that we want to estimate and $\mathbf{E}_{::i}$ is a realization of a certain additive noise. Ideally, the estimation of $\mathbf{A}_{::i}$ should provide the best available spectral resolution, or, in other terms, the cardinality N_b of the wavenumber space should be as large as possible.

We identify here two possible approaches for the estimation $\widehat{\mathbf{A}}_{::i}$ of the direct model:

- **Coefficient estimation**: for which the goal is a separate estimation $\widehat{a}_{kli}$ of all the coefficients $\{a_{kli}\}_{k\in[1,\dots,N_k],\, l\in[1,\dots,N_b],\, i\in[1,\dots,N_i]}$ of $\boldsymbol{\mathcal{A}}$. If we consider the model (6.4.1) as noiseless, then, for the m-th acquisition, a naive estimation is given by:
$$\widehat{a}_{kli} := \left.\frac{y_{kmi}}{x_{lmi}}\right|_{x_{l'mi}=0,\,\forall l'\neq l}\,, \tag{6.4.2}$$
or in other terms, $\widehat{a}_{kli}$ is equivalent to the stimulus due to a radiation characterized by an impulsive spectrum centered at the wavenumber σ_l, on the area of the subimage associated with the k-th interferometer and assigned to the solid angle Ω_i. This consideration suggests that, for the most efficient calibration setup, it is advisable to work with input spectra $\mathbf{X}_{::i}$ that closely approximate an identity matrix. The spacing $\Delta\sigma$ between the set of central wavelengths $\{\sigma_m\}_{m\in[1,\dots,N_a]}$ determines the spectral resolution of the estimated transfer matrix.
- **Parametric estimation**: This approach consists in assuming that the elements $\left\{a^{[\boldsymbol{\beta}]}_{kli}\right\}_{k\in[1,\dots,N_k],\, l\in[1,\dots,N_b],\, i\in[1,\dots,N_i]}$ of the direct model, which we denote here with $\boldsymbol{\mathcal{A}}^{[\boldsymbol{\beta}]}$, are samples of an analytical function in the set of parameters $\boldsymbol{\mathcal{B}}$, for which we aim to find the estimation $\widehat{\boldsymbol{\mathcal{B}}}$. If the direct model accurately describes the actual optical behaviour of the instrument and the parameters $\boldsymbol{\mathcal{B}}$ are estimated correctly, then the spectral resolution of $\boldsymbol{\mathcal{A}}^{[\boldsymbol{\mathcal{B}}]}$ (i.e., the number of bands N_b) can be chosen as high as desired.

In this thesis, we focus on the parametric approach, although the calibration setup for the characterization matches the one described in the coefficient estimation approach.

6.4.2 Model description

The most natural choice for the analytical function to assign to the direct model for the parametric estimation is given by the interferometer optical transfer function, which we presented in Section 5.5. In that discussion, the measurement was interpreted as a combination of N_m interfering light rays, generated by multiple reflections inside the FP etalons. This interpretation led to the definition of the "N_m-wave model", whose special cases are the "2-wave model", which can be interpreted as a discrete cosine transform (DCT), and the "∞-wave model", also known as Airy's distribution. For the reader's convenience, the expressions for each of these models are given in Table 6.5.

Table 6.5. Summary of the model of the coefficients a_{kli} of the interferometry transfer matrix models. The pedex kli refers to the dependency from the incident angle, the OPD and the wavenumbers, respectively. The shorthand values of the phase φ_{kli} is defined in eq. 6.4.3. The expected average value in a full phase period and the zero phase value are also provided for completeness.

Model	Transfer function	Mean value	Zero phase value
2-wave	$\mathcal{T}_{kli}^2\left(1+\mathcal{R}_{kli}^2+2\mathcal{R}_{kli}\cos\varphi_{kli}\right)$	$\mathcal{T}_{kli}^2(1+\mathcal{R}_{kli}^2)$	$\mathcal{T}_{kli}^2(1+\mathcal{R}_{kli})^2$
N_m**-wave**	$\mathcal{T}_{kli}^2\frac{1+\mathcal{R}_{kli}^{2N_m}-2\mathcal{R}_{kli}^{N_m}\cos(N_m\varphi_{kli})}{1+\mathcal{R}_{kli}^2-2\mathcal{R}_{kli}\cos\varphi_{kli}}$	$\mathcal{T}_{kli}^2\frac{1-\mathcal{R}_{kli}^{2N_m}}{1-\mathcal{R}_{kli}^2}$	$\mathcal{T}_{kli}^2\frac{\left(1-\mathcal{R}_{kli}^{N_m}\right)^2}{(1-\mathcal{R}_{kli})^2}$
∞**-wave**	$\frac{\mathcal{T}_{kli}^2}{(1-\mathcal{R}_{kli})^2+4\mathcal{R}_{kli}\sin^2(\varphi_{kli}/2)}$	$\frac{\mathcal{T}_{kli}^2}{1-\mathcal{R}_{kli}^2}$	$\frac{\mathcal{T}_{kli}^2}{(1-\mathcal{R}_{kli})^2}$

With regard to the formalism, $\mathcal{R}_{kli}$ and $\mathcal{T}_{kli}^2$ respectively denote the surface reflectivity and the gain of the system, while φ_{kli} is defined as:

$$\varphi_{kli} := 2\pi\delta_{kli}\sigma_l - \varphi_k^{[0]}\,, \tag{6.4.3a}$$

$$\delta_{kli} := 2n_{kl}L_k\cos\left(\frac{n_0}{n_{kl}}\theta_i^{[in]}\right)\,, \tag{6.4.3b}$$

where σ_l denotes the l-th wavenumber sample, $\theta_i^{[in]}$ is the i-th incidence angle, and δ_{kli} is the associated OPD related to the k-th interferometer with thickness L_k. We keep here the dependence of the OPD on the wavenumber through the refractive index n_{kl} of the medium inside the k-th interferometer at σ_l. Compared to the previous analysis, a new term $\varphi_k^{[0]}$ was introduced, to consider the effect of phase shift between consecutively reflected waves that are not taken into account by the OPD. This additional phase shift may be introduced at the reflection over the surfaces.

A major assumption of this work is that both the reflectivity $\mathcal{R}_{kli}$, the gain $\mathcal{T}^2_{kli}$, and the OPD δ_{kli} are:

- constant over the whole instrument's bandwidth, which allows to remove the dependency on the wavenumber; an analysis of the validity of such assumption is provided in the experiments of Section 6.4.4;
- relative to a specific angle of incidence $\boldsymbol{\theta}_i^{[in]}$. In particular, it is common practice to perform calibrations where the illumination sources are perpendicular to the incidence plane, for which the incident polar angle $\theta_i^{[in]}$ is equal to zero.

This allows to simplify the notation for this section, by rewriting the relevant variables as:

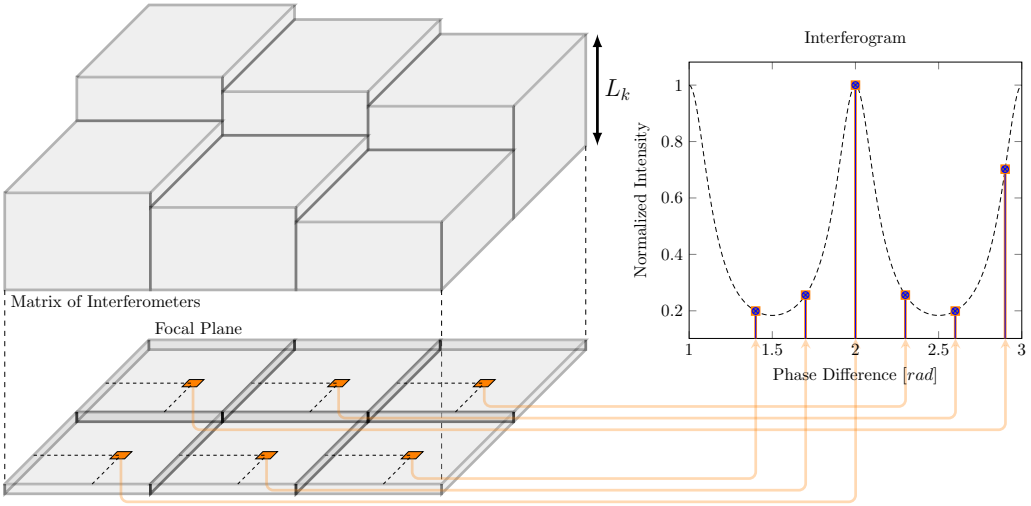
$$\begin{aligned} \mathcal{R}_k &:= \mathcal{R}_{kli}\,, & \mathbf{X} &:= \mathbf{X}_{::i}\,, \\ \mathcal{T}^2_k &:= \mathcal{T}^2_{kli}\,, & \mathbf{Y} &:= \mathbf{Y}_{::i}\,, \\ \delta_k &:= \delta_{kli}\,, & \mathbf{A} &:= \mathbf{A}_{::i}\,. \end{aligned} \tag{6.4.4}$$

We finally want to provide an example of how an accurate knowledge of the transformation matrix is useful for the inversion process. As the ImSPOC concept is a particular case of a FTS, the OPDs are typically designed to be equally spaced, as shown in Fig. 6.17a. This would allow to employ the inversion technique based on the Fourier transform to be discussed in Section 6.5.2; however, the calibration procedure may show that the real OPDs are spaced irregularly. If this information was not considered, the interferogram samples are placed incorrectly in the OPD domain, as shown in Fig. 6.17b, which strongly degrades the quality of the inversion.

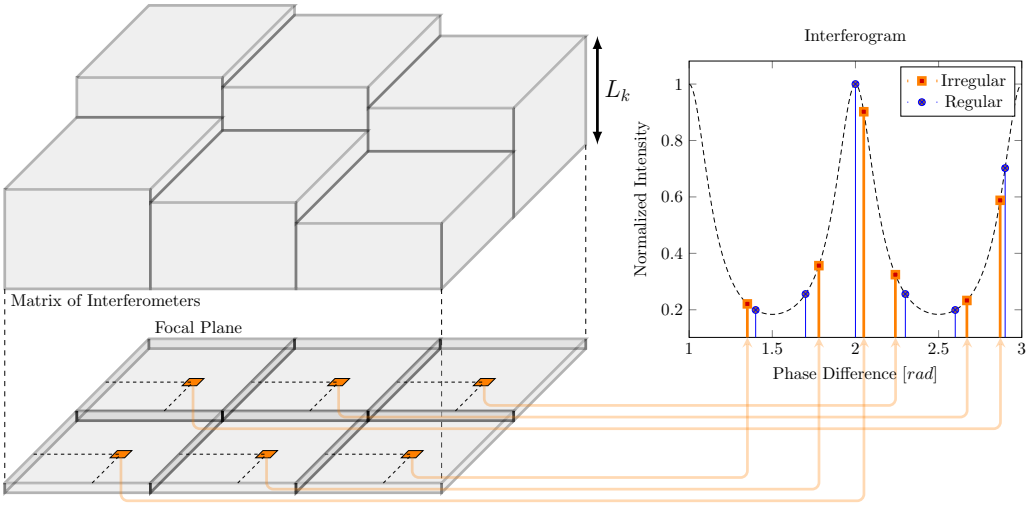
6.4.3 Parameter estimations algorithms

The analysis of this section will be limited to techniques aimed at the estimation of the quadruplets of parameters for the analytical functions shown in table 6.5. For the generic k-th interferometer, this quadruplet is identified by a column vector $\boldsymbol{\beta}_{:k} = [\delta_k;\, \mathcal{R}_k;\, \varphi_k^{[0]};\, \mathcal{T}_k^2]$, spanning the space $\mathtt{E}_b \subseteq \mathbb{R}^4$ and made up of the following four parameters:

- the **OPD** $\delta_k \in [0, 1/(2\Delta\sigma)]$;
- the **surface reflectivity** $\mathcal{R}_k \in (0, 1)$;
- the **phase shift** $\varphi_k^{[0]} \in [0, 2\pi)$;
- the **system gain** $\mathcal{T}_k^2 \in \mathbb{R}^+$.



(a) Nominal OPD values



(b) Exact OPD values

Fig. 6.17. Comparison between interferograms' samples with nominal and exact values for the OPDs.

Here, $\Delta\sigma = \sum_{l=1}^{N_a-1}(\sigma_{l+1} - \sigma_l)$ is the step size of the reconstruction space of the wavenumbers. The limitation on the domain of the OPD is to avoid aliasing, according to the Shannon-Nyquist theorem, similarly to what was discussed in Section 5.5.8. In fact, a certain estimation $\widehat{\delta}_k$ of the OPD and its alias $1/2\Delta\sigma - \widehat{\delta}_k$ lead to the same expression of the transfer matrix. The ambiguity arises due to its dependence on δ_k through the phase term $\cos(2\pi\sigma_l\delta_k)$, which for the alias becomes:

$$\cos\left(2\pi\left(\frac{1}{2\Delta\sigma} - \widehat{\delta}_k\right)\sigma_l\right) = \cos\left((l-1)\pi - 2\pi\widehat{\delta}_k\sigma_l\right) \tag{6.4.5a}$$

$$= \cos(2\pi\widehat{\delta}_k\sigma_l)\,, \qquad \forall l \in [1, \dots, N_b]\,, \tag{6.4.5b}$$

where we described $\sigma_l = (l-1)\Delta\sigma$ as an arithmetic sequence.

In the physical model of Section 5.5.7, a direct relationship between the terms $\mathcal{T}_k$ and $\mathcal{R}_k$ was given based on the conservation of energy. However, we consider them here as independent entities to include in $\mathcal{T}_k^2$ all the attenuation effects in the optical path that are not due to the interferometer itself. Furthermore, we estimate the reflectivity $\mathcal{R}_k$ separately for each interferometer; while it is a reasonable assumption for the reflectivity to be the same across all interferometers, this choice allows to simplify the problem and to take into account manufacturing differences among different reflective coatings.

While the core strategies of each algorithm may differ, their shared feature is to consider for simplicity the noise from eq. (6.4.1) as instances of additive white Gaussian noise, which is a realistic assumption for moderately high signal to noise ratio (SNR) (Section 5.2.4). The inference problem, as demonstrated in Section 2.1.2, becomes thus equivalent to minimizing the following cost function:

$$\widehat{\boldsymbol{\mathcal{B}}} = \arg\min_{\boldsymbol{\mathcal{B}}} \left\| \mathbf{A}^{[\boldsymbol{\mathcal{B}}]}\mathbf{X} - \mathbf{Y} \right\|_F^2 \tag{6.4.6}$$

where $\mathbf{A}^{[\boldsymbol{\mathcal{B}}]}$ denotes the transfer matrix evaluated in the matrix of parameters $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{4\times N_k}$, obtained by row concatenation of the vectors $\{\boldsymbol{\beta}_{:k}\}_{k\in[1,\dots,N_k]}$.

Since the parameters $\boldsymbol{\beta}$ were assumed to independent on the wavelength, the k-th row $\mathbf{a}_{k:}$ of $\mathbf{A}^{[\boldsymbol{\mathcal{B}}]}$ is just a function of $\boldsymbol{\beta}_{:k}$. This allows to separate eq. (6.4.6) in a set of N_k estimations $\left\{\widehat{\boldsymbol{\beta}}_{:k}\right\}_{k\in[1,\dots,N_k]}$ in the form:

$$\widehat{\boldsymbol{\beta}}_{:k} = \arg\min_{\boldsymbol{\beta}\in\mathbb{E}_b} \left\| \mathbf{a}_{k:}^{[\boldsymbol{\beta}]}\mathbf{X} - \mathbf{y}_{k:} \right\|_2^2\,, \tag{6.4.7}$$

where $\mathbf{a}_{k:}^{[\beta]}$ is a row vector, expressed as a function of an array of parameters $\beta \in \mathrm{E}_b \subseteq \mathbb{R}^4$ $\mathbf{y}_{k:}$ is the k-th row of the matrix $\mathbf{Y}$.

The model (6.4.7) is structurally similar to an inversion problem, where the prior is assigned intrinsically by imposing its parametric relationship. This model may be extended to consider additional priors, e.g. by imposing that the thicknesses of the interferometers increase in a given order. Such extensions come however at the cost of introducing a correlation between the information of different interferometers, which is not approachable with the formalism of eq. (6.4.7) and requires longer computation times.

Maximum likelihood approach

In general terms, the parametric estimation can be considered as a problem of data fitting. We employ in this section an approach based on maximum likelihood estimation (MLE), which is an extension and mathematical justification of the work we presented in [59].

The **maximum likelihood (ML)** method itself is based on two main assumptions:

- the input spectra are **strictly impulsive**, so that $\mathbf{X}$ is an identity matrix of sizes $N_a \times N_a$, with $N_a = N_b$; this allows to rewrite the simulated acquisition as:

$$\mathbf{Y}^{[sim]} = \mathbf{A}^{[\boldsymbol{\mathcal{B}}]}\mathbf{X} = \mathbf{A}^{[\boldsymbol{\mathcal{B}}]}\,; \tag{6.4.8}$$

- the elements of the transfer matrix $\mathbf{A}$ are samples of a **2-wave model** function, so that its coefficients are given by:

$$a_{kl} = \mathcal{T}_k^2 \left(1 + \mathcal{R}_k^2 + 2\mathcal{R}_k \cos\left(2\pi\delta_k\sigma_l - \varphi_k^{[0]}\right)\right)\,. \tag{6.4.9}$$

We want to rewrite eq. (6.4.7) by normalizing the acquisitions as follows:

$$\widetilde{\mathbf{y}}_{k:} := \frac{\mathbf{y}_{k:} - \overline{\mathbf{y}}_{k:}}{\overline{\mathbf{y}}_{k:}}\,, \tag{6.4.10}$$

for which we obtain:

$$\widehat{\boldsymbol{\beta}}_{:k} = \arg\min_{\boldsymbol{\beta}\in\mathbf{E}_b} \left\| \widetilde{\mathbf{y}}_{k:}^{[sim]} - \widetilde{\mathbf{y}}_{k:} \right\|_2^2 \tag{6.4.11a}$$

$$= \arg\min_{\boldsymbol{\beta}\in\mathbf{E}_b} \left\| \frac{\mathbf{a}_{k:}^{[\beta]} - \overline{\mathbf{a}}_{k:}^{[\beta]}}{\overline{\mathbf{a}}_{k:}^{[\beta]}} - \widetilde{\mathbf{y}}_{k:} \right\|_2^2 \tag{6.4.11b}$$

$$= \arg\min_{\boldsymbol{\beta}\in\mathbf{E}_b} \sum_{l=1}^{N_a} \left(\frac{2\mathcal{R}_k}{(1+\mathcal{R}_k)^2} \cos\left(2\pi\delta_k\sigma_l - \varphi_k^{[0]}\right) - \widetilde{y}_{kl} \right)^2 \tag{6.4.11c}$$

$$= \arg\min_{\boldsymbol{\beta}\in\mathbf{E}_b} \sum_{l=1}^{N_a} \left(\alpha_k \cos\left(2\pi\delta_k\sigma_l - \varphi_k^{[0]}\right) - \widetilde{y}_{kl} \right)^2 , \tag{6.4.11d}$$

where we made use of both the conditions (6.4.8) and (6.4.9) and we have defined:

$$\alpha_k = \frac{2\mathcal{R}_k}{(1+\mathcal{R}_k)^2} . \tag{6.4.12}$$

Eq 6.4.11 is a classical problem of estimation of the parameters of a sinusoid affected by Gaussian noise, whose MLE solution is a well known result in the literature (e.g. the reader can refer to example 7.16 in [130]). Specifically, the inference of the parameters is done as follows. First, we obtain the estimation $\widehat{\delta}_k$ of the OPD:

$$\widehat{\delta}_k = \arg\max_{\delta_k\in\left[0,\frac{1}{2\Delta\sigma}\right]} \begin{bmatrix} \mathbf{y}_{k:}\mathbf{c}_k \\ \mathbf{y}_{k:}\mathbf{s}_k \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \mathbf{c}_k^{\mathsf{T}}\mathbf{c}_k & \mathbf{c}_k^{\mathsf{T}}\mathbf{s}_k \\ \mathbf{s}_k^{\mathsf{T}}\mathbf{c}_k & \mathbf{s}_k^{\mathsf{T}}\mathbf{s}_k \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{y}_{k:}\mathbf{c}_k \\ \mathbf{y}_{k:}\mathbf{s}_k \end{bmatrix} , \tag{6.4.13}$$

where $\mathbf{c}_k^{\mathsf{T}} = \frac{1}{N_a}[\cos(2\pi\delta_k\sigma_1), ..., \cos(2\pi\delta_k\sigma_{N_a})]$, $\mathbf{s}_k^{\mathsf{T}} = \frac{1}{N_b}[\sin(2\pi\delta_k\sigma_1), ..., \sin(2\pi\delta_k\sigma_{N_a})]$. If δ_k is not close to either zero or to $1/(2\Delta\sigma)$, then:

$$\mathbf{s}_k^{\mathsf{T}}\mathbf{c}_k = \mathbf{c}_k^{\mathsf{T}}\mathbf{s}_k = \frac{1}{2N_b}\sum_{l=1}^{N_a} \sin(4\pi\delta_k\sigma_l) \approx 0 , \tag{6.4.14}$$

for which eq. (6.4.13) can be simplified to:

$$\widehat{\delta}_k \approx \arg\max_{\delta\in\left[0,\frac{1}{2\Delta\sigma}\right]} \left| \sum_{l=1}^{N_b} \widetilde{y}_{kl} e^{-j2\pi\delta\sigma_l} \right| . \tag{6.4.15}$$

The MLE result $\widehat{\delta}_k$ is hence equivalent to choosing the OPD which maximizes the **periodogram**, that is, the generalized discrete Fourier transform (DFT) of $\widetilde{\mathbf{y}}_{k:}$. Given $\widehat{\delta}_k$, the estimations $\widehat{\alpha}_k$ of α_k and $\widehat{\varphi}_k^{[0]}$ of $\varphi_k^{[0]}$ are given by:

$$\widehat{\alpha}_k = \frac{2}{N_a}\left|\sum_{l=1}^{N_a} \widetilde{y}_{kl} e^{-j2\pi\widehat{\delta}_k\sigma_l}\right| \tag{6.4.16a}$$

$$\widehat{\varphi}_k^{[0]} = \arctan\frac{\sum_{l=1}^{N_a} \widetilde{y}_{kl}\sin(2\pi\widehat{\delta}_k\sigma_l)}{\sum_{l=1}^{N_a} \widetilde{y}_{kl}\cos(2\pi\widehat{\delta}_k\sigma_l)} \tag{6.4.16b}$$

where $\arctan$ denotes a four quadrant inverse tangent of a point whose coordinate are the denominator and the numerator of the argument; $\widehat{\varphi}_k^{[0]}$ can hence assume any value in the range $[0, 2\pi)$. The procedure is fully described in Algorithm 5.

Algorithm 5: Maximum likelihood (ML) parameter estimation.

Result: Parameters' Estimation:

- $\widehat{\boldsymbol{\beta}}_{:k} = \left[\widehat{\delta}_k, \widehat{\mathcal{R}}_k; \widehat{\varphi}_k^{[0]}; \widehat{\mathcal{T}}_k^2\right], \forall k \in [1, \dots, N_k]$

Input:

- Input spectra's central wavenumbers: $\{\sigma_l\}_{l\in[\![1..N_a]\!]}$
- Matrix of acquisitions: $\mathbf{Y} \in \mathbb{R}^{N_k\times N_a}$ whose rows are $\{\mathbf{y}_{k:}\}_{k\in[1,\dots,N_k]}$
- OPD Sample Space: $\{\delta'_s\}_{s=1,\dots,N_s} \in \left[0, \frac{1}{2\Delta\sigma}\right]$ where $\delta_s = \frac{s-1}{2\Delta\sigma}$ and $\Delta\sigma = \frac{1}{N_b-1}\sum_{l=1}^{N_b-1}(\sigma_{l+1} - \sigma_l)$

Procedure:

Construct $\mathbf{W} \in \mathbf{R}^{N_s\times N_a}$ with elements $w_{lk} = e^{-j2\pi\delta'_k\sigma_l}$
for $k = 1, \dots, N_k$ **do**
  $\widetilde{\mathbf{y}}_{k:} = (\mathbf{y}_{k:} - \overline{\mathbf{y}}_{k:}) \,/\, \overline{\mathbf{y}}_{k:}$
  $\mathbf{r} \leftarrow \left|\mathbf{W}\widetilde{\mathbf{y}}_{k:}^{\mathsf{T}}\right| \in \mathbb{R}^{N_s}$ and denote the s-th element with J_{sk}
  $\widehat{s} = \arg\max_{s\in[1,\dots,N_s]} r_s$
  $\widehat{\delta}_k = \delta'_{\widehat{s}}$
  Calculate $\widehat{\alpha}_k$ and $\widehat{\varphi}_k^{[0]}$ analytically from eq. 6.4.16
  $\widehat{\mathcal{R}}_k = \left(1 - \sqrt{1-\widehat{\alpha}_k^2}\right)/\widehat{\alpha}_k$
  $\widehat{\mathcal{T}}_k^2 = \frac{\overline{\mathbf{y}}_{k:}}{1+\widehat{\alpha}_k\sum_{l=1}^{N_b}\cos(2\pi\delta_k\sigma_l)}$
end

The ML method requires very low computational power, but its applicability is limited by validity of its assumptions. Specifically, the transfer function model can be assumed as 2-wave only for devices for reasonably low reflectivities (typically below 0.3), which is the case for the low finesse interferometers which are used in the current ImSPOC designs.

Exhaustive search approach

The **exhaustive search (ES)** approach, which we originally proposed in [190], consists in sampling the parameters space $\mathtt{E}_b$ in order to define a set of N_s candidate vectors $\{\boldsymbol{\beta}'_{:s}\}_{s\in[1,\dots,N_s]}$ that can represent, up to a given desired accuracy, the whole space. The objective function (6.4.7) is then computed for each of those candidates and the one that corresponds to the minimum value is chosen as the estimation.

Formally, the estimation can be expressed as:

$$\widehat{\boldsymbol{\beta}}_{:k} = \arg \min_{\{\boldsymbol{\beta}'_{:s}\}_{s\in[1,\dots,N_s]}} \left\| \mathbf{a}_{k:}^{[\boldsymbol{\beta}'_{:s}]}\mathbf{X} - \mathbf{y}_{k:} \right\|_2^2 = \arg \min_{\{\boldsymbol{\beta}'_{:s}\}_{s\in[1,\dots,N_s]}} \left\| \mathbf{y}_{k:}^{[sim]}(\boldsymbol{\beta}'_{:s}) - \mathbf{y}_{k:} \right\|_2^2 , \tag{6.4.17}$$

which is equivalent to comparing a simulated acquisition $\mathbf{y}_{k:}^{[sim]}(\boldsymbol{\beta}'_{:s}) = \mathbf{a}_{k:}^{[\boldsymbol{\beta}'_{:s}]}\mathbf{X}$ with the measured one $\mathbf{y}_{k:}$.

The main challenge of this approach is to choose the segmentation of the parameter space $\mathtt{E}_b$ that provides the best compromise between the amount of checks and the accuracy of the estimation. There is a vast literature dedicated to this problem, especially in the domain of machine learning where this field is known as **hyper-parameter search** [24]; some possible approaches include grid searches, random searches, Bayesian optimization, and more. Determining the most efficient solution is however outside the scope of this work, as we simply investigate a regular grid sampling of the parameter space $\mathtt{E}_b$.

As the scanning of a three dimensional parameter space requires too much computational time, it is useful to operate in domains that are independent to the variations of a given variable. For impulsive inputs $\mathbf{X}$, we propose to approach this issue with the following two adjustments:

- **Gain**: To get rid of the dependency from the gain $\mathcal{T}_k^2$, one simple adjustment is to normalize the acquisitions for their mean to be equal to 1, so that eq. (6.4.17) can be rewritten as:

$$\widehat{\boldsymbol{\beta}}_{:k} = \arg \min_{\{\boldsymbol{\beta}'_{:s}\}_{s\in[1,\dots,N_s]}} \left\| \frac{\mathbf{y}_{k:}^{[sim]}(\boldsymbol{\beta}'_{:s})}{\overline{\mathbf{y}}_{k:}^{[sim]}(\boldsymbol{\beta}'_{:s})} - \frac{\mathbf{y}_{k:}}{\overline{\mathbf{y}}_{k:}} \right\|_2^2 , \tag{6.4.18}$$

  so that just check for sample vectors $\boldsymbol{\beta}'_{:s}$ with a fixed $\mathcal{T}_k^2 = 1$.

- **Phase shift**: To make our estimation independent of the phase shift, we propose to scan the parameter space in the Fourier amplitude domain. Specifically, let $\mathbf{W} \in \mathbb{R}^{N_a \times N_a}$ define a DFT transformation matrix whose elements

$w_{kl} = \exp\left(-j2\pi\check{\delta}_k\sigma_l\right)$, where $\{\sigma_l\}_{l\in[1,\dots,N_a]}$ are the central wavenumbers of the input spectra, whose mean step size is $\Delta\sigma$ and $\check{\delta}_k^{[ref]} = (k-1)/(2\Delta\sigma)$. Then, we propose to estimate the OPD and the reflectivity with the functional:

$$\left[\widehat{\delta}_k, \widehat{\mathcal{R}}_k\right] = \arg \min_{\substack{[\delta'_s,\mathcal{R}'_s]:\\ s\in[1,\dots,N_s]}} \left\| \frac{\left|\mathbf{W}\left(\mathbf{y}_{k:}^{[sim]}([\delta'_s,\mathcal{R}'_s])\right)^{\mathsf{T}}\right|}{\overline{\mathbf{y}}_{k:}^{[sim]}([\delta'_s,\mathcal{R}'_s])} - \frac{|\mathbf{W}\mathbf{y}_{k:}^{\mathsf{T}}|}{\overline{\mathbf{y}}_{k:}} \right\|_2^2, \quad (6.4.19)$$

The main motivation behind this approach is due to the DFT shifting theorem, which states that the magnitude of the DFT is invariant to any circular shift of $\delta_k^{[ref]}\Delta\sigma$, where $\delta_k^{[ref]}$ is the real OPD to estimate. This is the maximum error we allow on the estimation of the phase shift, which can be made small enough if we raise the amount N_a of acquisitions.

Once the OPD and the reflectivity are estimated, those parameters can be assumed as fixed for the estimation of the phase shift. The full procedure is described in the Algorithm 6. The ES method allows to extend the interferometer parameter estimation to the cases in which the 2-wave model is not a sufficiently accurate representation of the optical transformation, but its accuracy is limited by the discrete segmentation of the sample space, which has to be very dense to return precise results.

Algorithm 6: Exhaustive search (ES) parameter estimation.

Result:

- Parameters' Estimation: $\widehat{\boldsymbol{\beta}}_{k:} = \left[\widehat{\delta}_k;\, \widehat{\mathcal{R}}_k;\, \widehat{\varphi}_k^{[0]};\, \widehat{\mathcal{T}}_k^2\right],\ \forall k \in [1, \dots, N_k]$

Input:

- Matrix of acquisitions: $\mathbf{Y} \in \mathbb{R}^{N_k \times N_a}$ whose rows are $\{\mathbf{y}_{k:}\}_{k\in[1,\dots,N_k]}$
- Amount of samples for parameters' space tessellation: $[N_{s_1}, N_{s_2}, N_{s_3}]$
- Wave model label (e.g. 2, N_m or ∞ waves)
- Input spectra $\mathbf{X} \in \mathbb{R}^{N_b \times N_a}$ and their central wavenumbers $\{\sigma_l\}_{l\in[1,\dots,N_a]}$

Procedure:

Define the sample spaces: $\delta'_{s_1} = \frac{s_1-1}{N_{s_1}}\frac{1}{2\Delta\sigma}$, $\mathcal{R}'_{s_2} = \frac{s_2-1}{N_{s_2}}$, $\varphi_{s_3}^{[0]\prime} = \frac{s_3-1}{N_{s_3}}2\pi$,
$\forall s_1 \in [1, \dots, N_{s_1}],\ s_2 \in [1, \dots, N_{s_2}],\ s_3 \in [1, \dots, N_{s_3}]$ with
$\Delta\sigma = \frac{1}{N_b-1}\sum_{l=1}^{N_b-1}(\sigma_{l+1} - \sigma_l)$
Construct $\mathbf{W} \in \mathbf{R}^{N_s \times N_a}$ with elements $w_{lk} = e^{-j2\pi\delta'_k\sigma_l}$
for $k = 1, ..., N_k$ **do**
  Evaluate $\overline{\mathbf{y}}_{k:} = \sum_{l=1}^{N_b} y_{kl}$
  $J_{min} \leftarrow \infty$
  for $s_1 = 1, ..., N_{s_1},\ s_2 = 1, ..., N_{s_2}$ **do**
    $\boldsymbol{\beta}' \leftarrow \left[\delta'_{s_1};\, \mathcal{R}'_{s_2};\, 0;\, 1\right]$
    Evaluate $\mathbf{a}_{k:}^{[\beta']}$ with the selected model from Table 6.5
    Evaluate $\mathbf{y}^{[sim]} = \mathbf{a}_{k:}^{[\beta']}\mathbf{X}$ and its mean $\overline{\mathbf{y}}^{[sim]}$
    Evaluate $J' = \left\| \frac{|\mathbf{W}\mathbf{y}^{[sim]\intercal}|}{\overline{\mathbf{y}}^{[sim]}} - \frac{|\mathbf{W}\mathbf{y}_{k:}^{\intercal}|}{\overline{\mathbf{y}}_{k:}} \right\|_2$
    if $J' < J_{min}$ **then**
      $J_{min} \leftarrow J',\ \widehat{\delta}_k \leftarrow \delta'_{s_1},\ \widehat{\mathcal{R}}_k \leftarrow \mathcal{R}'_{s_2}$
    end
  end
  $J_{min} \leftarrow \infty$
  for $s_3 = 1, ..., N_{s_3}$ **do**
    $\boldsymbol{\beta}' \leftarrow [\widehat{\delta}_k;\, \widehat{\mathcal{R}}_k;\, \varphi_{s_3}^{[0]\prime};\, 1]$
    Evaluate $\mathbf{a}_{k:}^{[\beta']}$ with the selected model from Table 6.5
    Evaluate $\mathbf{y}^{[sim]} = \mathbf{a}_{k:}^{[\beta']}\mathbf{X}$ and its mean $\overline{\mathbf{y}}^{[sim]}$
    Evaluate $J' = \left\| \frac{\mathbf{y}^{[sim]}}{\overline{\mathbf{y}}^{[sim]}} - \frac{\mathbf{y}_{k:}}{\overline{\mathbf{y}}_{k:}} \right\|_2$
    if $J' < J_{min}$ **then**
      $J_{min} \leftarrow J',\ \widehat{\varphi}_k^{[0]} \leftarrow \varphi_{s_3}^{[0]\prime}$
    end
  end
  Evaluate $\mathbf{y}^{[sim]} = \mathbf{a}_{k:}^{[\beta']}\mathbf{X}$ and its mean $\overline{\mathbf{y}}^{[sim]}$
  $\widehat{\mathcal{T}}_k^2 \leftarrow \frac{\overline{\mathbf{y}}_{k:}}{\overline{\mathbf{y}}^{[sim]}}$
end

Gauss-Newton algorithm approach

In this section, we present the **Gauss-Newton algorithm (GNA)** method, which is an alternative procedure for the estimation of the parameters of an interferometer based on the solution of a nonlinear regression problem. As opposite to ES method, the GNA method allows to estimate the parameters with a user-defined degree of precision. which can be set by choosing the amount of iterations of the solving algorithm.

Eq. (6.4.17) can be interpreted as a problem of fitting a nonlinear function to each of the acquisitions; this is a problem of nonlinear regression, which can be solved with any of the techniques discussed in Section 2.3.3.

The GNA method iterates a two step procedure; if we suppose $\boldsymbol{\beta}_{:k}^{(q)}$ is the estimation of the parameters at the q-th iteration, the update step consists in:

- **Linearization of the model**: where the transfer matrix $\mathbf{a}_{k:}^{[\boldsymbol{\beta}_{:k}^{(q)}]}$ is represented by a Taylor decomposition at $\boldsymbol{\beta}_{:k}^{(q)}$, truncated to its first derivatives.

  Formally, for a the k-th interferometer, we define **residual** $\mathbf{r}^{(q)}$ at the q-th iteration the quantity:

  $$\mathbf{r}^{(q)} := \mathbf{a}_{k:}(\boldsymbol{\beta}_{:k}^{(q)})\mathbf{X} - \mathbf{y}_{k:}\,. \tag{6.4.20}$$

  Its gradient $\boldsymbol{\nabla}_r^{(q)} \in \mathbb{R}^{4\times N_a}$ is expressed in terms of the gradient $\boldsymbol{\nabla}_a^{(q)}$ of $\mathbf{a}_{k:}$ as:

  $$\boldsymbol{\nabla}_r^{(q)} = \boldsymbol{\nabla}_a^{(q)}\mathbf{X}\,, \tag{6.4.21}$$

  where

  $$\boldsymbol{\nabla}_a^{(q)} = \left[\frac{\partial \mathbf{a}_{k:}}{\partial \boldsymbol{\beta}_{:k}^{(q)}}\right] = \left[\frac{\partial \mathbf{a}_{k:}}{\partial \delta_k}; \frac{\partial \mathbf{a}_{k:}}{\partial \mathcal{R}_k}; \frac{\partial \mathbf{a}_{k:}}{\partial \varphi_k^{[0]}}; \frac{\partial \mathbf{a}_{k:}}{\partial \mathcal{T}_k^2}\right]\,. \tag{6.4.22}$$

  The gradient $\boldsymbol{\nabla}_a^{(q)} \in \mathbb{R}^{4\times N_b}$ is a concatenation of partial derivatives in terms of each of the four components of $\boldsymbol{\beta}^{(q)}$. For the 2-wave and ∞-wave model, those are given in Table 6.6.

- **Linear Regression**: where the objective function is solved as a classic regression, assuming that the direct model is linear.

  The update for the parameters is then given by:

  $$\boldsymbol{\beta}_{:k}^{(q+1)} = \boldsymbol{\beta}_{:k}^{(q)} - \left(\boldsymbol{\nabla}_r^{(q)}\right)^{\dagger}\mathbf{r}^{(q)} \tag{6.4.23}$$

where $\left(\boldsymbol{\nabla}_r^{(q)}\right)^{\dagger} = \left(\boldsymbol{\nabla}_r^{(q)\mathsf{T}}\boldsymbol{\nabla}_r^{(q)}\right)^{-1}\boldsymbol{\nabla}_r^{(q)\mathsf{T}}$ is the pseudo-inverse of $\boldsymbol{\nabla}_r^{(q)}$.

Table 6.6. Summary of the coefficients of the Jacobian matrix associated with the transfer matrix for the 2 and ∞-wave model.

	2-wave model	$\infty-$**wave model**
φ_{kl}	$2\pi\delta_k\sigma_l - \varphi_k^0$	
a_{kl}	$\mathcal{T}_k^2\left(1+\mathcal{R}_k^2+2\mathcal{R}_k\cos\varphi_{kl}\right)$	$\frac{\mathcal{T}_k^2}{(1-\mathcal{R}_k)^2+4\mathcal{R}_k\sin^2(\varphi_{kl}/2)}$
$\frac{\partial a_{kl}}{\partial\delta_k}$	$-4\pi\sigma_l\mathcal{T}_k^2\mathcal{R}_k\sin\varphi_{kl}$	$-4\pi\sigma_l\mathcal{R}_k\frac{a_{kl}^2}{\mathcal{T}_k^2}\sin\varphi_{kl}$
$\frac{\partial a_{kl}}{\partial\mathcal{R}_k}$	$2\mathcal{T}_k^2(\mathcal{R}_k+\cos\varphi_{kl})$	$-2\frac{a_{kl}^2}{\mathcal{T}_k^2}\left(2\sin^2\left(\frac{\varphi_{kl}}{2}\right)+\mathcal{R}_k-1\right)$
$\frac{\partial a_{kl}}{\partial\varphi_k^0}$	$2\mathcal{T}_k^2\mathcal{R}_k\sin\varphi_{kl}$	$2\mathcal{R}_k\frac{a_{kl}^2}{\mathcal{T}_k^2}\sin\varphi_{kl}$
$\frac{\partial a_{kl}}{\partial\mathcal{T}_k^2}$	$\frac{a_{kl}}{\mathcal{T}_k^2}$	

The iterations are then repeated for either a fixed number of times or until the decrease of the objective function $J^{(q)} = \|\mathbf{r}_k^{(q)}\|_2^2$ is below a certain user defined threshold. This method exhibits desirable features:

- the solution can converge to any point parameter space E_b;
- it allows for more sophisticated expressions for the transfer model, such as the ∞-wave model;
- it may use any set of input spectra $\mathbf{X}$.

Its main limitation is the requirement for an accurate initialization $\boldsymbol{\beta}^{[0]}$ for the process to converge to the desired local minimum; a reasonable initialization, when available, can be provided by the result of the ML Algorithm 5.

The whole procedure is described in Algorithm 7.

Algorithm 7: Gauss-Newton algorithm (GNA) for parameter estimation.

Result:

- Parameters' Estimation: $\widehat{\boldsymbol{\beta}}_{:k} = \left[\widehat{\delta}_k;\, \widehat{\mathcal{R}}_k;\, \widehat{\varphi}_k^{[0]};\, \widehat{\mathcal{T}}_k^2\right],\ \forall k \in [1, \dots, N_k]$

Input:

- Matrix of acquisitions: $\mathbf{Y} \in \mathbb{R}^{N_k \times N_a}$ whose rows are $\{\mathbf{y}_{k:}\}_{k\in[1,\dots,N_k]}$
- Wave model label (e.g., 2 or ∞ waves)
- Input spectra $\mathbf{X} \in \mathbb{R}^{N_b \times N_a}$
- Maximum number of iterations N_q
- Tolerance: J_{tol}

Procedure:

Initialize $\left\{\boldsymbol{\beta}_k^{(0)}\right\}_{k\in[\![1..N_k]\!]}$ with Algorithm 5 (ML method)

for $k = 1, ..., N_k$ **do**

$J^{(-1)} \leftarrow \infty$ `// Initialize the objective function value`

$q \leftarrow 0$ `// Initialize the iterations`

$\mathbf{a}_{k:} \leftarrow \left\{ a_{kl} \middle|_{\left[\delta_k;\mathcal{R}_k;\varphi_k^{[0]};\mathcal{T}_k^2\right]=\boldsymbol{\beta}_k^{(0)}} \right\}_{l\in[1,\dots,N_b]}$ with a_{kl} from Table 6.6

$J^{(0)} \leftarrow \|\mathbf{a}_{k:}\mathbf{X} - \mathbf{y}_{k:}\|_2$

while $q < N_q$ ***and*** $\left|J^{(q)} - J^{(q-1)}\right| > J_{tol}$ **do**

$q \leftarrow q + 1$

$\mathbf{a}_{k:} \leftarrow \left\{ a_{kl} \middle|_{\left[\delta_k;\mathcal{R}_k;\varphi_k^{[0]};\mathcal{T}_k^2\right]=\boldsymbol{\beta}_k^{(q)}} \right\}_{l\in[1,\dots,N_b]}$ with a_{kl} from Table 6.6

$\boldsymbol{\nabla}_A \leftarrow \left\{ \left[\frac{\partial a_{kl}}{\partial\delta_k};\, \frac{\partial a_{kl}}{\partial\mathcal{R}_k};\, \frac{\partial a_{kl}}{\partial\varphi_k^{[0]}};\, \frac{\partial a_{kl}}{\partial\mathcal{T}_k^2}\right] \middle|_{\left[\delta_k;\mathcal{R}_k;\varphi_k^{[0]};\mathcal{T}_k^2\right]=\boldsymbol{\beta}_k^{(q)}} \right\}_{l\in[1,\dots,N_b]}$ with partial derivatives from Table 6.6

$\mathbf{r} \leftarrow \mathbf{a}_{k:}\mathbf{X} - \mathbf{y}_{k:}$ `// Evaluate the residual`

$\boldsymbol{\nabla}_r \leftarrow \boldsymbol{\nabla}_A\mathbf{X}$ `// Compute the Jacobian`

$\boldsymbol{\beta}_{:k}^{(q)} = \boldsymbol{\beta}_{:k}^{(q-1)} - \left(\boldsymbol{\nabla}_r\boldsymbol{\nabla}_r^{\mathsf{T}}\right)^{-1}\boldsymbol{\nabla}_r\,\mathbf{r}^{\mathsf{T}}$ `// Update the parameters`

$J^{(q)} = \left\|\mathbf{r}^{\mathsf{T}}\right\|_2^2$ `// Compute the current error`

end

$\widehat{\boldsymbol{\beta}}_{:k} \leftarrow \boldsymbol{\beta}_{:k}^{(q)}$

end

6.4.4 Calibration experimental results

Experimental setup

In this section we compare the performances of the model charaterization methods we proposed for both real and simulated acquisitions.

The experiment is set up as follows:

- The device under test is illuminated with a series of monochromatic acquisitions with central wavelengths $\{\sigma_l\}_{l\in[1,\dots,N_a]}$;
- The average values of the illuminated region of each subimage is arranged in a matrix $\mathbf{Y} \in \mathbb{R}^{N_k \times N_a}$;
- The space of parameters $\widehat{\boldsymbol{\mathcal{B}}} \in \mathbb{R}^{4\times N_k}$ is estimated with the investigated method;
- A simulated acquisition $\mathbf{Y}^{[sim]} \in \mathbb{R}^{N_k \times N_a}$ is evaluated as $\mathbf{A}^{[\widehat{\boldsymbol{\mathcal{B}}}]}\mathbf{X}$;
- The simulated acquisition $\mathbf{Y}^{[sim]}$ is compared with $\mathbf{Y}$ with a series of quality indices.
- If the nominal set of OPDs is known (i.e. from the design sheets provided to the manufacturer), their values are compared with the estimated OPDs.

The reconstructed interferogram $\mathbf{Y}^{[sim]}$ is compared with the real acquisition $\mathbf{Y}$, by evaluating the root mean square error (RMSE) directly in the domain of the OPD:

$$\text{RMSE} = \left\| \mathbf{Y}^{[sim]} - \mathbf{Y} \right\|_F = \left\| \mathbf{A}^{[\widehat{\boldsymbol{\mathcal{B}}}]}\mathbf{X} - \mathbf{Y} \right\|_F . \tag{6.4.24}$$

We also evaluate the average Fourier norm (AFN) to compare their amplitudes in the Fourier domain; this index is defined as follows:

$$\text{AFN} = \left\| \left| \left(\mathbf{A}^{[\widehat{\boldsymbol{\mathcal{B}}}]}\mathbf{X} - \mathbf{Y} \right) \mathbf{W}^{\mathsf{T}} \right| \right\|_F , \tag{6.4.25}$$

where $\mathbf{W} \in \mathbb{R}^{N_k \times N_k}$ is a standard DFT transformation matrix. The RMSE and AFN are expected to yield similar results in terms of hierarchies of the methods' performances, but the latter is reasonably exempt from the effect of phase mismatches. If a set of nominal OPDs $\{\delta_k^{[ref]}\}_{k\in[1,\dots,N_k]}$ is available, this can be compared with the array of estimated OPDs $\{\widehat{\delta}_k\}_{k\in[1,\dots,N_k]}$. For this purpose, the most natural quality index is the mean absolute error (MAE):

$$\text{MAE} = \frac{1}{N_k} \sum_{k=1}^{N_k} |\delta_k^{[ref]} - \widehat{\delta}_k| . \tag{6.4.26}$$

However, the current technology for manufacturing an array of interferometer is able to control the variation of the thickness between adjacent interferometers with

a much larger degree of accuracy than that of absolute thicknesses. To address this issue, we define an alternative quality index, the mean absolute unbiased error (MAUE), as:

$$\text{MAUE} = \frac{1}{N_k} \sum_{k=1}^{N_k} \left| \left(\delta_k^{[ref]} - \overline{\boldsymbol{\delta}}^{[ref]} \right) - \left(\widehat{\delta}_k - \overline{\widehat{\boldsymbol{\delta}}} \right) \right| , \tag{6.4.27}$$

where $\overline{\boldsymbol{\delta}}$ denotes the average of the associated OPD over all interferometers.

With respect to the tested methods, we compare the performances of the three approaches proposed in Section 6.4, assuming different transfer matrix analytical models. Specifically, the ML method always assumes that the optical transformation is described by a 2-wave model, but, once the parameters are estimated, the simulated acquisition $\mathbf{Y}^{[sim]}$ may be generated with any model. For the remaining methods, we impose instead that the estimation and simulation model are the same. In summary, the following method are tested:

- the ML method, with a 2-wave, 3-wave, and ∞-wave simulation model;
- the ES method, with a 2-wave, 3-wave, and ∞-wave estimation/simulation model;
- the GNA method, with a 2-wave, and an ∞-wave estimation/simulation model.

Real data results

For the real dataset, three different spectral spetral calibration procedures were performed at the Institut de Planétologie et d'Astrophysique de Grenoble (IPAG), for the PROTO-1, PROTO-2, and PROTO-3, whose general characteristics were described in Section 6.1.7.

For each spectral calibration procedure, an extended source (i.e., a lamp) generates a wideband incident field; its incident light is filtered by a diffraction grating, which acts as a bandpass filter whose bandwidth is centered around a user-selected center wavelength. The filtered light illuminates the input plane in a direction roughly perpendicular to the device under test, and the readout of the instrument is recorded. The procedure is repeated for N_a acquisitions, selecting a different central wavelength on the diffraction grating every time, to simulate the response of the device to N_a monochromators.

For each acquisition, the central area of each subimage is cropped into 12×12 px squares and the average values are arranged into the matrix of acquisitions $\mathbf{Y} \in \mathbb{R}^{N_k \times N_a}$.

The method for the determination of the centers of the illuminated spots are slightly different for each prototype:

- For the PROTO-1, we perform the CCE on all acquisitions, and then we pick the median coordinates across all acquisitions;
- For the PROTO-2, the center coordinates were given by the design datasheets;
- For the PROTO-3, the center coordinate of the central subimage was obtained with the CCE on a sample image, and the remaining coordinates are chosen to be regularly spaced on a grid.

These informations are summarized in Table 6.7.

Table 6.7. Characteristics of the spectral calibration procedures for the characterization of the transfer matrix of an ImSPOC.

Prototype	**PROTO-1**	**PROTO-2**	**PROTO-3**
Acquisitions (N_a)	101	721	343
Central wavenumber range $[\mathrm{mm}^{-1}]$	[1000,2000]	[1000,2800]	[625,1000]
Central wavenumber step size $[\mathrm{mm}^{-1}]$	10	2.5	Variable (1.1 ± 1.2)
Center estimation of the illuminated area	Median of CCE	Design datasheet	CCE+ regular grid
Illuminated region (px)	12×12	12×12	12×12

The results of the experiments on the three prototypes are given in tables 6.8, 6.9, and 6.10; an analysis of these tables highlights the advantages of operating in continuous domains, which yields a particular hierarchy of performances: GNA, followed by the ML and the ES. The three methods explore a increasingly less dense domain of parameters, which respectively are fully continuous, continuous everywhere except for the OPD, and fully discrete. The results do not contradict the previous literature [190], as the presented implementation of the ML is a more refined version of the one presented by Dolet et al. [59], as the presented method provides an estimation for both the phase shift and the reflectivity, other than just the OPD.

For each of the three prototypes, Fig. 6.18 provides a visual representation of the estimated OPD. They are expected to be roughly equivalent to two times the interferometers' thicknesses and are sorted with respect to their identifying indices of Fig. 6.4. Those results closely match the nominal increase in thickness of Table 6.2. With regard to some design specific considerations:

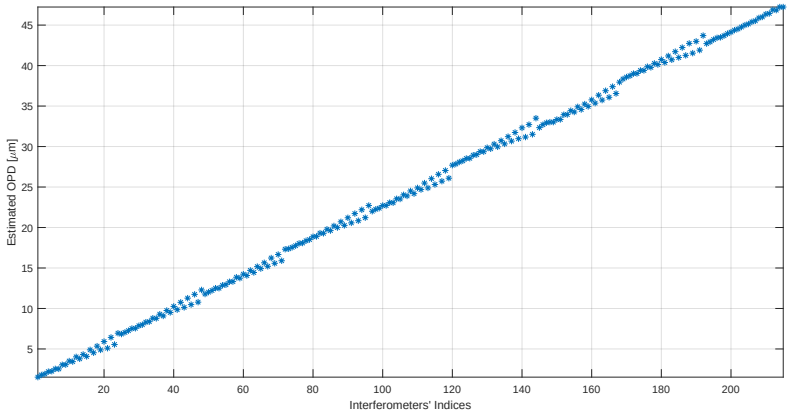
- The PROTO-1 is manufactured with a particular staircase design pattern, which increases linearly from a central vertical axis both on the left and right direction

(according the order shown in Fig. 6.4a). The step size of the thicknesses is supposed to be the same on both sizes, but our results show that this is not an accurate statement, probably because faces of the interferometers are not exactly parallel, but instead bent on their top side. The effect of bending causes a different behaviour on each side of the staircase, which causes the alternating pattern of Fig. 6.18a;

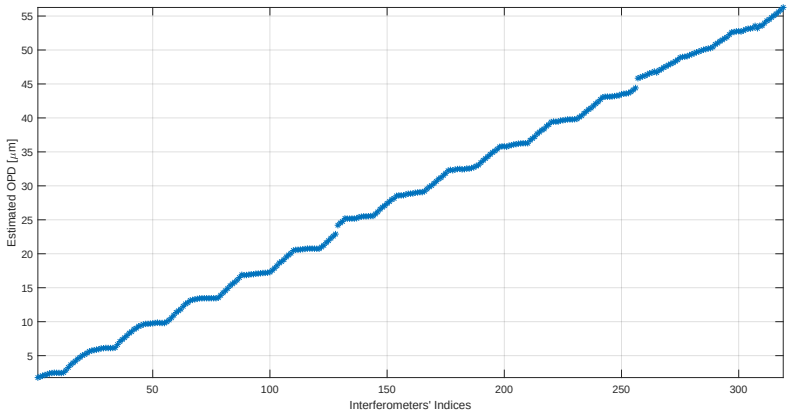
- The PROTO-3 is characterized by two elements at the optical contact (the opposing reflective faces touch each other directly), which can be easily identified at the two ends of Fig. 6.18c.

The analysis of the MAUE seems to indicate that the estimation of the OPD with the ML is more accurate, but it is difficult to assess if this is due to the performances of the inference method or because of imperfections introduced during the process of manufacturing the device, especially considering that the GNA provides better performances for simulated datasets. Another visual comparison between the observed and simulated acquisition is available in Fig. 6.19 for PROTO-1 and in Fig. 6.20 for PROTO-3, showcasing the fitting of the analytical function for a relatively thin and relatively thick interferometer. This visualization can be used both to show the benefits of a proper choice of the options of the model characterization method and to highlight some limitations of the proposed models; in particular a comparison between Fig. 6.20a and 6.20e (or similarly Fig. 6.19a and 6.19e) shows that the ∞-wave model provides a better approximation of the acquisition by generating oscillations that are more rounded towards the bottom. Similarly, a visual analysis of Fig. 6.20b and 6.20h (and similarly, Fig. 6.19b and 6.19h) shows how the inference of the OPD over a finer domain allows to follow more closely the oscillations of the acquisition.

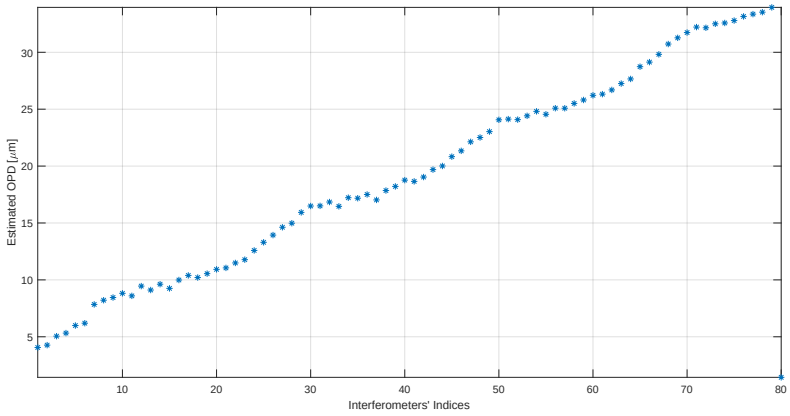
With respect to the limitations of the model, this work has assumed that the reflectivity is constant with the wavelength, so the proposed model is not capable to take into account variations in the amplitude of the oscillation, which are instead present on the real data across all prototypes, especially for PROTO-3 where the coating exhibits a decrease in its efficiency for low wavelengths. PROTO-3 also shows an anomalous effect in Fig. 6.20 of increase in frequency of oscillations for high wavenumbers; to the best of my knowledge, this effect is not mainly due to a variation of the OPD with the wavenumber, but instead because of uncompensated spurious impulses generated by the monochromator.



(a) PROTO-1

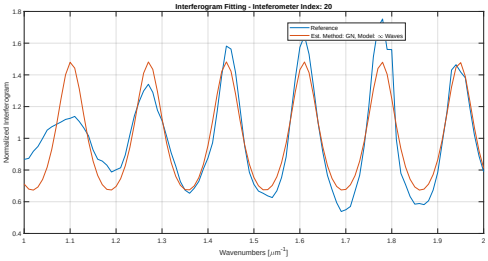


(b) PROTO-2

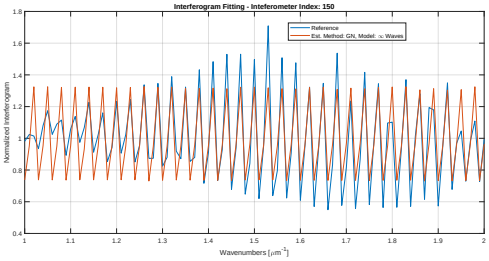


(c) PROTO-3

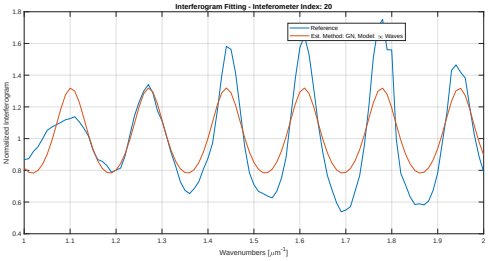
Fig. 6.18. Visual representation of the estimated values of the OPD for different prototypes, performed with the GNA method and an ∞-wave model.



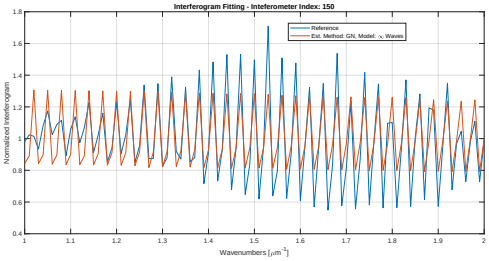
(a) Inference method: GNA, model: ∞-wave, estimated $\mathcal{R}$, interferometer index: 20



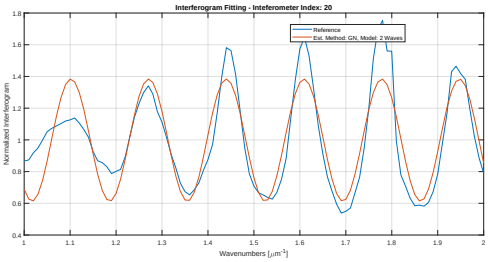
(b) Inference method: GNA, model: ∞-wave, estimated $\mathcal{R}$, interferometer index: 150



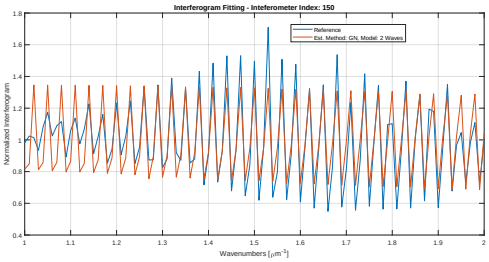
(c) Inference method: GNA, model: ∞-wave, fixed $\mathcal{R}$, interferometer index: 20



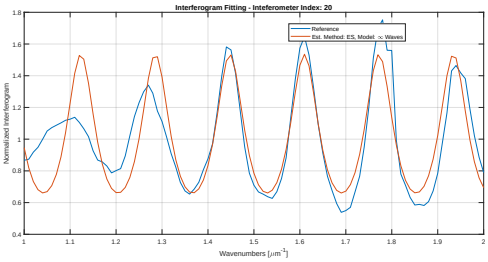
(d) Inference method: GNA, model: ∞-wave, fixed $\mathcal{R}$, interferometer index: 150



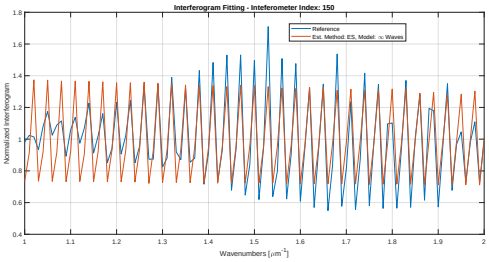
(e) Inference method: GNA, model: 2-wave, estimated $\mathcal{R}$, interferometer index: 20



(f) Inference method: GNA, model: 2-wave, estimated $\mathcal{R}$, interferometer index: 150

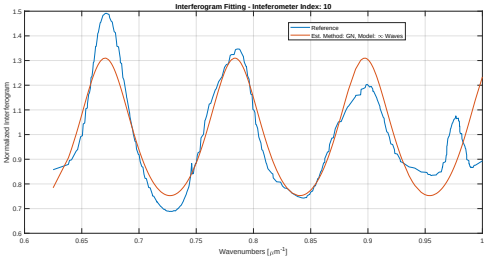


(g) Inference method: GNA, model: ∞-wave, estimated $\mathcal{R}$, interferometer index: 20

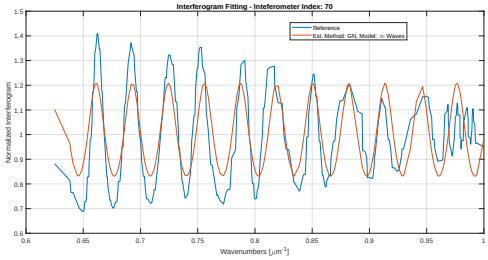


(h) Inference method: ES, model: ∞-wave, estimated $\mathcal{R}$, interferometer index: 150

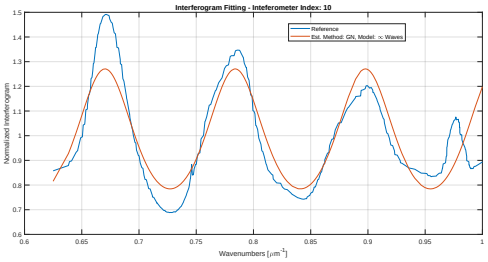
Fig. 6.19. Comparison between the observed interferogram (in blue) and its parametric reconstruction (in red) with various configurations for the PROTO-1 prototype.



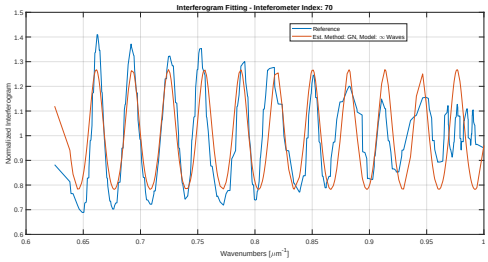
(a) Inference method: GNA, model: ∞-wave, estimated $\mathcal{R}$, interferometer index: 10



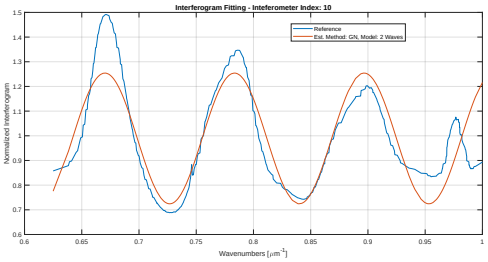
(b) Inference method: GNA, model: 2-wave, estimated $\mathcal{R}$, interferometer index: 70



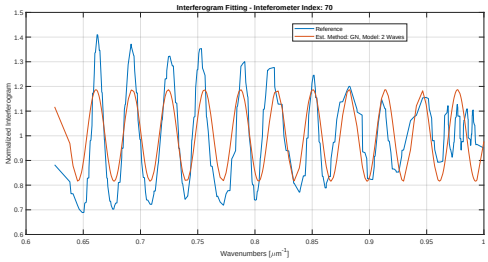
(c) Inference method: GNA, model: 2-wave, fixed $\mathcal{R}$, interferometer index: 10



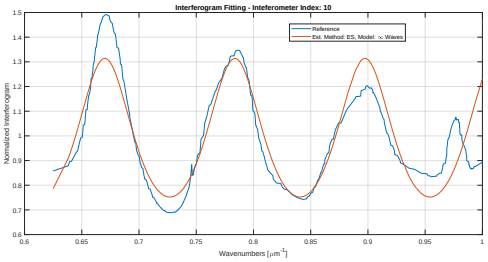
(d) Inference method: GNA, model: ∞-wave fixed $\mathcal{R}$, interferometer index: 70



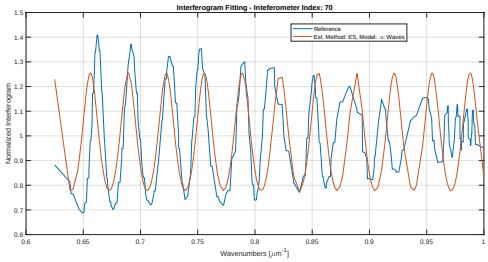
(e) Inference method: GNA, model: 2-wave, estimated $\mathcal{R}$, interferometer index: 10



(f) Inference method: GNA, model: 2-wave, estimated $\mathcal{R}$, interferometer index: 70



(g) Inference method: ES, model: ∞-wave, estimated $\mathcal{R}$, interferometer index: 10



(h) Inference method: ES, model: ∞-wave, estimated $\mathcal{R}$, interferometer index: 70

Fig. 6.20. Comparison between the observed interferogram (in blue) and its parametric reconstruction (in red) with various configurations for the PROTO-3.

Table 6.8. Results for the model characterization of the PROTO-1. The cardinality of the sample space for methods operating in a discretized space is $[10000, 100, 10000]$ and the tolerance for GNA is set to 10^{-15}. Best results are in bold and second best are underlined.

	Method	Model	RMSE $\times 10^{-1}$	AFN	MAUE $[\mu m]$
Estimated $\mathcal{R}$	ML	2	0.2845 ± 0.1211	4.210 ± 1.657	$\mathbf{0.3496 \pm 0.2309}$
	ML	∞	$\underline{0.2489 \pm 0.1080}$	$\underline{3.848 \pm 1.621}$	$\mathbf{0.3496 \pm 0.2309}$
	ES	2	0.3397 ± 0.1653	4.534 ± 1.370	0.3714 ± 0.2596
	ES	3	0.2953 ± 0.1412	4.152 ± 1.352	0.3701 ± 0.2637
	ES	∞	0.2906 ± 0.1368	4.133 ± 1.345	0.3683 ± 0.2629
	GNA	2	0.2833 ± 0.1202	4.194 ± 1.671	$\underline{0.3499 \pm 0.2323}$
	GNA	∞	$\mathbf{0.2438 \pm 0.1048}$	$\mathbf{3.776 \pm 1.747}$	0.3511 ± 0.2355
Fixed $\mathcal{R}$	ML	2	0.3564 ± 0.1532	6.092 ± 5.320	$\mathbf{0.3496 \pm 0.2309}$
	ML	∞	0.3268 ± 0.1394	5.748 ± 5.212	$\mathbf{0.3496 \pm 0.2309}$
	ES	2	0.3842 ± 0.1745	6.333 ± 5.576	0.3725 ± 0.2592
	ES	3	0.3571 ± 0.1697	6.011 ± 5.428	0.3758 ± 0.2697
	ES	∞	0.3567 ± 0.1705	6.007 ± 5.426	0.3709 ± 0.2645
	GNA	2	0.3537 ± 0.1517	6.018 ± 5.163	0.3500 ± 0.2320
	GNA	∞	0.3236 ± 0.1376	5.665 ± 5.062	0.3508 ± 0.2329

Table 6.9. Results for the model characterization of the PROTO-2. The cardinality of the sample space for methods operating in a discretized space is $[2500, 100, 2500]$ and the tolerance for GNA is set to 10^{-15}. Best results are in bold and second best are underlined.

	Method	Model	RMSE	AFN	MAUE $[\mu m]$
Estimated $\mathcal{R}$	ML	2	0.8879 ± 0.7539	258.706 ± 284.550	0.4758 ± 0.3327
	ML	∞	0.8892 ± 0.7649	$\mathbf{255.162 \pm 281.721}$	0.4758 ± 0.3327
	ES	2	0.9285 ± 0.7837	259.462 ± 278.893	2.2738 ± 7.7980
	ES	3	0.9286 ± 0.8115	259.260 ± 279.750	2.3636 ± 7.9696
	ES	∞	0.9523 ± 0.9092	261.687 ± 277.691	2.7474 ± 8.5237
	GNA	2	$\underline{0.8872 \pm 0.7537}$	258.528 ± 284.455	0.4747 ± 0.3312
	GNA	∞	$\mathbf{0.8804 \pm 0.7539}$	$\underline{256.886 \pm 285.042}$	0.4750 ± 0.3305
Fixed $\mathcal{R}$	ML	2	0.9293 ± 0.7688	282.694 ± 300.610	0.4758 ± 0.3327
	ML	∞	0.9252 ± 0.7684	281.521 ± 300.362	0.4758 ± 0.3327
	ES	2	0.9382 ± 0.7749	284.017 ± 301.077	2.2729 ± 7.7984
	ES	3	0.9341 ± 0.7752	282.959 ± 301.001	2.6235 ± 8.3448
	ES	∞	0.9344 ± 0.7750	282.942 ± 301.041	2.6237 ± 8.3450
	GNA	2	0.9277 ± 0.7683	281.977 ± 299.885	$\mathbf{0.4746 \pm 0.3311}$
	GNA	∞	0.9233 ± 0.7679	280.719 ± 299.618	$\underline{0.4746 \pm 0.3304}$

Table 6.10. Results for the model characterization of the PROTO-3. The cardinality of the sample space for methods operating in a discretized space is $[10000, 100, 10000]$ and the tolerance for GNA is set to 10^{-15}. Best results are in bold and second best are underlined.

	Method	Model	RMSE $\times 10^{-1}$	AFN	MAUE $[\mu\text{m}]$
Estimated $\mathcal{R}$	ML	2	0.1507 ± 0.0378	0.6640 ± 0.4765	$\mathbf{0.6016 \pm 0.3739}$
	ML	∞	0.1458 ± 0.0396	0.6342 ± 0.4690	$\mathbf{0.6016 \pm 0.3739}$
	ES	2	0.2222 ± 0.0861	0.8693 ± 0.5052	0.8299 ± 0.7679
	ES	3	0.2115 ± 0.1065	0.8102 ± 0.4389	0.8110 ± 0.6213
	ES	∞	0.2088 ± 0.1036	0.8065 ± 0.4403	0.7785 ± 0.5817
	GNA	2	0.1449 ± 0.0404	$\underline{0.6160 \pm 0.4329}$	0.6620 ± 0.5004
	GNA	∞	$\mathbf{0.1394 \pm 0.0418}$	$\mathbf{0.5902 \pm 0.4096}$	0.6421 ± 0.4646
Fixed $\mathcal{R}$	ML	2	0.1526 ± 0.0371	0.6729 ± 0.4824	$\mathbf{0.6016 \pm 0.3739}$
	ML	∞	0.1478 ± 0.0390	0.6439 ± 0.4723	$\mathbf{0.6016 \pm 0.3739}$
	ES	2	0.1979 ± 0.0732	0.8743 ± 0.6010	0.8307 ± 0.7345
	ES	3	0.1925 ± 0.0677	0.8297 ± 0.5718	0.7704 ± 0.5386
	ES	∞	0.1918 ± 0.0686	0.8250 ± 0.5660	0.7680 ± 0.5423
	GNA	2	0.1488 ± 0.0381	0.6581 ± 0.4587	0.6435 ± 0.4387
	GNA	∞	$\underline{0.1437 \pm 0.0390}$	0.6275 ± 0.4326	$\underline{0.6313 \pm 0.4397}$

Simulated dataset

In this section the same experiments are carried out over a set of simulated acquisitions, generated with our analytical model both under ideal conditions, which we denote as **baseline experiment**, and with some elements of nonideality, to assess the validity of the proposed algorithms in different testbeds.

In the baseline experiment, the matrix of acquisitions $\mathbf{Y}$ is assumed to be generated as a matrix multiplication $\mathbf{AX}$, such that:

- $\mathbf{A}$ is a matrix obtained by sampling an ∞-wave model analytical function, whose parameters are chosen to match the nominal ones of the PROTO-1 from Table 6.2. Specifically:
  - the nominal OPDs increase with a step of 200 nm from a minimum value of 1000 nm;
  - the reflectivity is set to 0.13;
  - the phase shift is set to zero;
  - the system gain is set to 1;
- $\mathbf{X} \in \mathbb{R}^{101\times 101}$ is made up of 101 perfectly impulsive spectra, whose central wavenumbers are evenly spaced over the range $[1000, 2000]$ mm^{-1}.

The following nonidealities are considered as variations on the baseline experiment:

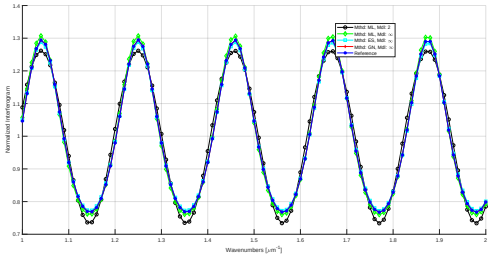
- **Noisy acquisition**: in which we add additive white Gaussian noise (AWGN) to each element of the acquisition $\mathbf{Y}$, such that the SNR is equal to 20 dB;
- **Non-impulsive inputs**: For which the input is a matrix $\mathbf{X} \in \mathbb{R}^{10001\times 101}$, for which each input is a Gaussian curve with STD of 5 mm^{-1} (the resolution of the spectra is thus 0.1 mm^{-1});
- **Uncertain OPDs**: where the nominal OPDs are perturbed by adding a zero mean additive noise with STD of 20 nm.

The STD of the Gaussian is chosen to be below 10 mm^{-1}, as the Gaussian can be seen as a low pass operation in cascade of the impulsive input. In fact, to be able to resolve a maximum OPD $\delta^{[max]} = 47$ nm, the Shannon-Nyquist theorem requires an STD not greater than $1/(2\delta_{max}) \approx 10$ $\mu\mathrm{m}^{-1}$.

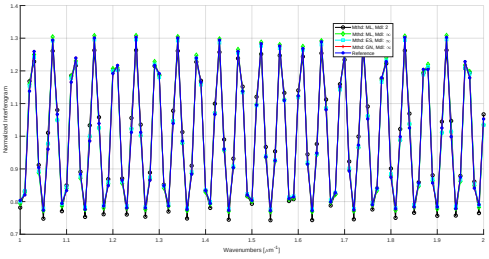
For all combinations, the complete set of parameters (OPD, reflectivity, phase shift and gain) was inferred with the same methods investigated over real datasets, and validated graphically in Fig. 6.21 and quantitatively in Table 6.11. A quick analysis of the obtained results shows that the GNA method consistently outperforms the

other approaches, especially when there is a match between the simulation and the reconstruction models. In the case of noiseless acquisitions and impulsive inputs, the spectra are perfectly reconstructed, regardless of the value of the nominal OPDs.

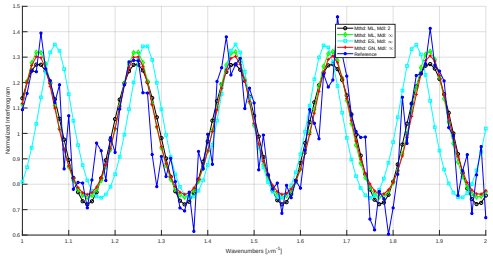
The GNA method shows impressive results even when the input spectra are not impulsive and the acquisitions are noisy, as the OPD are consistently inferred accurately. As opposite to the previous tests on real data, we have here perfect knowledge of the reference OPD and the results justify the advantage of exploring continuous parameters' spaces.



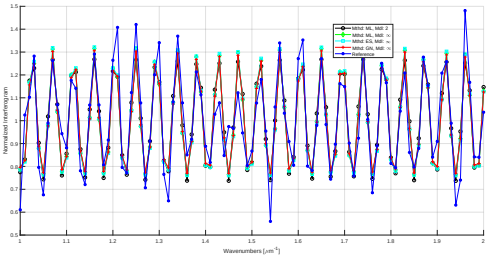
(a) SNR: ∞ dB, impulsive input interferometer index: 20



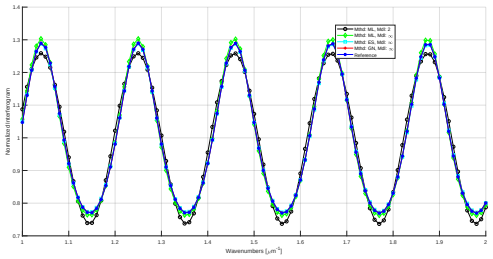
(b) SNR: ∞ dB, impulsive input interferometer index: 20



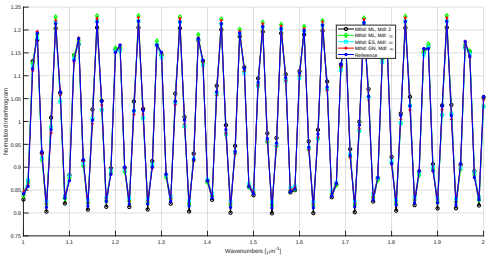
(c) SNR: 20 dB, impulsive input interferometer index: 20



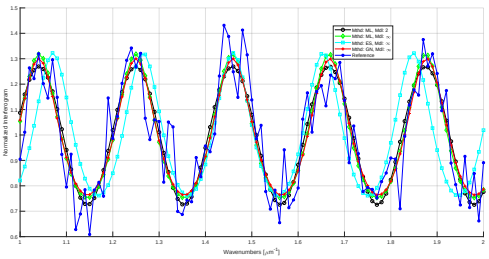
(d) SNR: 20 dB, impulsive input interferometer index: 100



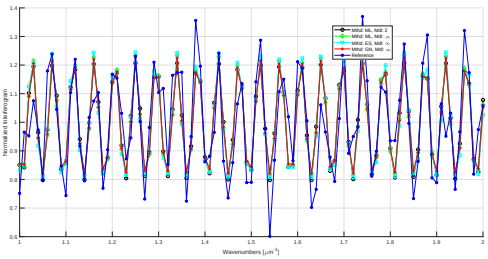
(e) SNR: ∞ dB, Gaussian input STD: 5 mm^{-1} interferometer index: 20



(f) SNR: ∞ dB, Gaussian input STD: 5 mm^{-1} interferometer index: 100



(g) SNR: 20 dB, Gaussian input STD: 5 mm^{-1} interferometer index: 20



(h) SNR: 20 dB, Gaussian input STD: 5 mm^{-1} interferometer index: 100

Fig. 6.21. Comparison between the simulated interferogram (in blue) and its parametric reconstruction with various methods. Each row is relative to a specific configuration. The simulation assumes an ideally constructed transfer matrix with the nominal parameters of PROTO-1, with a bias on the nominal OPD of 1 μm.

Table 6.11. Quantitative results for the model characterization for a set of 101 simulated impulsive acquisitions with different parameters indicated in the sideline. The cardinality of the search space is set as $[1001, 101, 1000]$ and the tolerance for the GNA method is 10^{-15}. Groups separated by double horizontal lines have a different reference. Best results are in bold, second best ones are underlined.

			Mthd	**Mdl**	**RMSE** $\times 10^{-3}$	**AFN** $\times 10^{-1}$	**MAE** $\times 10^{-1}$ $[\mu m]$
			Ideal	∞	0	0	0
Impulse spectrum Gaussian STD: 0 mm^{-1}	OPD STD: 0 nm	SNR: ∞ dB	ML	2	0.6607 ± 0.1276	0.9030 ± 0.8060	0.1666 ± 0.1169
			ML	3	0.0883 ± 0.0952	0.1158 ± 0.0492	0.1666 ± 0.1169
			ML	∞	$\underline{0.0790 \pm 0.0925}$	$\underline{0.1074 \pm 0.0506}$	0.1666 ± 0.1169
			ES	2	0.7757 ± 0.6540	1.0215 ± 0.8057	0.2020 ± 0.2770
			ES	3	0.1411 ± 0.4829	0.1707 ± 0.0763	0.1806 ± 0.2488
			ES	∞	0.1328 ± 0.4860	0.1598 ± 0.0772	0.1828 ± 0.2518
			GNA	2	0.5970 ± 0.0755	0.8478 ± 0.8146	$\underline{0.0314 \pm 0.1372}$
			GNA	∞	$\mathbf{0.0000 \pm 0.0000}$	$\mathbf{0.0000 \pm 0.0000}$	$\mathbf{0.0000 \pm 0.0001}$
		SNR: 20 dB	ML	2	1.0391 ± 0.3287	1.2241 ± 0.7440	0.2768 ± 0.2036
			ML	3	0.5116 ± 0.3904	$\underline{0.5583 \pm 0.1243}$	0.2768 ± 0.2036
			ML	∞	$\underline{0.5045 \pm 0.3944}$	0.5589 ± 0.1239	0.2768 ± 0.2036
			ES	2	3.6414 ± 5.7534	4.2748 ± 0.7594	1.0721 ± 1.4333
			ES	3	3.0694 ± 5.3210	3.7845 ± 1.0798	1.0476 ± 1.2776
			ES	∞	3.3082 ± 5.8621	4.0732 ± 1.0018	1.0796 ± 1.3566
			GNA	2	0.9957 ± 0.2847	1.1992 ± 0.7537	$\underline{0.2473 \pm 0.2174}$
			GNA	∞	$\mathbf{0.4064 \pm 0.2782}$	$\mathbf{0.4338 \pm 0.1189}$	$\mathbf{0.2260 \pm 0.1796}$
	OPD STD: 20 nm	SNR: ∞ dB	ML	2	0.6620 ± 0.1398	0.1614 ± 0.1255	0.1614 ± 0.1255
			ML	3	0.0889 ± 0.1021	0.1614 ± 0.1255	0.1614 ± 0.1255
			ML	∞	$\underline{0.0794 \pm 0.0994}$	0.0961 ± 0.0278	0.1614 ± 0.1255
			ES	2	0.8410 ± 0.8901	0.2220 ± 0.3563	0.2222 ± 0.3564
			ES	3	0.1422 ± 0.4345	0.1836 ± 0.2454	0.1839 ± 0.2477
			ES	∞	0.1268 ± 0.4244	0.1538 ± 0.0468	0.1823 ± 0.2419
			GNA	2	0.5962 ± 0.0786	$\underline{0.0360 \pm 0.1364}$	$\underline{0.0322 \pm 0.1378}$
			GNA	∞	$\mathbf{0.0000 \pm 0.0000}$	$\mathbf{0.0000 \pm 0.0000}$	$\mathbf{0.0000 \pm 0.0001}$
Impulse spectrum Gaussian STD: 5 mm^{-1}	OPD STD: 20 nm	SNR: ∞ dB	ML	2	0.2170 ± 0.2801	0.2405 ± 0.1798	0.1607 ± 0.1250
			ML	3	0.0682 ± 0.0939	$\underline{0.0793 \pm 0.0337}$	0.1607 ± 0.1250
			ML	∞	$\underline{0.0676 \pm 0.0888}$	0.0801 ± 0.0354	0.1607 ± 0.1250
			ES	2	0.2498 ± 0.2993	0.2640 ± 0.1650	0.1933 ± 0.2513
			ES	3	0.0783 ± 0.1888	0.0855 ± 0.0231	0.1647 ± 0.1964
			ES	∞	0.0779 ± 0.1897	0.0863 ± 0.0249	0.1644 ± 0.1963
			GNA	2	0.1688 ± 0.2005	0.1972 ± 0.1737	$\underline{0.0210 \pm 0.1331}$
			GNA	∞	$\mathbf{0.0180 \pm 0.0132}$	$\mathbf{0.0230 \pm 0.0184}$	$\mathbf{0.0065 \pm 0.0167}$
		SNR: 20 dB	ML	2	0.6453 ± 0.4249	0.6542 ± 0.1938	0.3923 ± 0.3138
			ML	3	$\underline{0.5193 \pm 0.3553}$	$\underline{0.5604 \pm 0.1204}$	0.3923 ± 0.3138
			ML	∞	0.5193 ± 0.3551	0.5632 ± 0.1210	0.3923 ± 0.3138
			ES	2	3.2669 ± 4.7574	3.6522 ± 0.6656	2.7486 ± 17.2989
			ES	3	3.2256 ± 4.3751	3.8310 ± 0.7686	2.7697 ± 17.2658
			ES	∞	3.2445 ± 4.3662	3.8557 ± 0.6673	2.8022 ± 17.2696
			GNA	2	0.5948 ± 0.3618	0.6046 ± 0.1904	$\underline{0.3609 \pm 0.3189}$
			GNA	∞	$\mathbf{0.4466 \pm 0.3091}$	$\mathbf{0.4725 \pm 0.1205}$	$\mathbf{0.3570 \pm 0.3063}$

6.5 Inversion of interferograms

6.5.1 Problem statement

For the devices based on the ImSPOC concept, the **inversion problem** defines the series of procedures aimed at recovering the incident spectral radiance, given the measured intensity values associated with the matrix of interferometers.

From the previous analysis, the acquisition is defined as a datacube of co-registered subimages $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{N_k \times N_a \times N_i}$, so that the generic element y_{kmi} is the m acquisition due to the k-th interferometer and associated with the solid angle Ω_i.

With this description, the acquisition is then modeled as:

$$\mathbf{Y}_{::i} = \mathbf{A}_{::i}\mathbf{X}_{::i} + \mathbf{E}_{::i}\,, \tag{6.5.1}$$

where $\mathbf{E}_{::i} \in \mathbb{R}^{N_k}$ is once again assumed to be realization of AWGN, with the same caveats of Section 6.4.1 and $\mathbf{A}_{::i}$ is the calibrated matrix that results from the model characterization (Section 6.4). This is an ill-conditioned problem, which we aim to address with the methods described in the following section, in order to recover the reconstructed product $\widehat{\mathbf{X}}_{::i}$.

6.5.2 Inversion protocols

Three different approaches are investigated for the resolution of the inversion problem (6.5.1) in the following section.

We employ in this context three different approaches:

- an inversion in the Fourier domain;
- a PMD for the inversion of the direct model;
- a regularization with a LASSO framework.

They are roughly ordered in increasing order of computational complexity, and in decreasing order of versatility.

Fourier transform based methods

As the ImSPOC device can be seen as a particular instance of a FTS, it is pedagogical to describe the envisioned inversion method that guides the design of the device, based on the description of the interferograms in the Fourier domain. The main drawback of such theoretical technique is that its scope is limited to a very restrictive set of assumptions, specifically:

- the transfer matrix $\mathbf{A}_{::i}$ is perfectly described as samples of a 2-wave model with zero phase shift;
- the OPDs associated with said matrix are described by a regularly spaced arithmetic sequence of increasing values starting from zero;
- the reflectivity and gain parameter, denoted with $\mathcal{R}_i$ and $\mathcal{T}_i^2$ respectively, are equal for every interferometer and are constant in the whole wavelength range of operation of the device;
- the wavenumber spectrum of the incident radiance is strictly limited to a baseband interval $[0, B_\sigma]$.

According to the analysis of Section 5.5.8, the condition for an alias-free reconstruction of a spectrum with bandwidth B_σ is $\Delta\delta_i \leq 1/(2B_\sigma)$. For the sake of exposition, we target a spectrum reconstruction over $N_b = N_k$ samples, which is the maximum spectral resolution allowed by the Nyquist-Shannon theorem. In summary, the set of conditions is equivalent to having coefficients a_{kli} of the transfer matrix $\boldsymbol{\mathcal{A}}$ in the form:

$$a_{kli} = \mathcal{T}_i^2 \left(1 + \mathcal{R}_i^2 + 2\mathcal{R}_i \cos(2\pi\sigma_l \delta_{ki})\right) \tag{6.5.2a}$$

$$\delta_{ki} = (k-1)\Delta\delta_i = (k-1)\frac{1}{2B_\sigma} \tag{6.5.2b}$$

$$\sigma_l = \frac{l-1}{N_k - 1} B_\sigma \tag{6.5.2c}$$

which yields the following deterministic model for the acquisition:

$$y_{kmi} = \sum_{l=1}^{N_k} a_{kli} x_{lmi} \tag{6.5.3a}$$

$$= \mathcal{T}_i^2 \left(1 + \mathcal{R}_i^2\right) \overline{\mathbf{x}}_{:mi} + 2\mathcal{T}_i^2 \mathcal{R}_i \sum_{k=1}^{N_k} \cos\left(\frac{\pi}{N_k - 1}(k-1)(l-1)\right) x_{lmi}\,. \tag{6.5.3b}$$

With some algebraic modification, eq. (6.5.3b) can be rewritten in matrix form as:

$$\widetilde{\mathbf{y}}_{:mi} = \mathbf{W}\widetilde{\mathbf{x}}_{:mi}\,. \tag{6.5.4}$$

The coefficients of the matrices $\widetilde{\mathbf{y}}_{:mi} = \{\widetilde{y}_{kmi}\}_{k\in[1,\dots,N_k]}$, $\mathbf{W} = \{w_{kl}\}_{\substack{l\in[1,\dots,N_b]\\ k\in[1,\dots,N_k]}}$, and $\widetilde{\mathbf{x}}_{:mi} = \{\widetilde{x}_{lmi}\}_{k\in[1,\dots,N_k]}$ are described by the equations:

$$\widetilde{y}_{kmi} := \frac{y_{kmi} - \mathcal{T}_i^2\left(1+\mathcal{R}_i^2\right)\overline{\mathbf{x}}_{:mi}}{2\mathcal{T}_i^2\mathcal{R}_i}\sqrt{1+\mathcal{I}_{l,\{1,N_k\}}}\sqrt{\frac{N_k-1}{2}}\,, \tag{6.5.5a}$$

$$\widetilde{x}_{lmi} := x_{lmi}\sqrt{1+\mathcal{I}_{k,\{1,N_k\}}}\,, \tag{6.5.5b}$$

$$w_{kl} := \sqrt{\frac{2}{N_k-1}}\sqrt{\frac{1}{\left(1+\mathcal{I}_{l,\{1,N_k\}}\right)\left(1+\mathcal{I}_{k,\{1,N_k\}}\right)}}\cos\left(\frac{\pi}{N_k-1}(k-1)(l-1)\right)\,, \tag{6.5.5c}$$

where $\mathcal{I}_{k,\{1,N_k\}}$ is a Kronecker delta, defined as:

$$\mathcal{I}_{k,\{1,N_k\}} = \begin{cases} 1 & \text{if } k=1 \text{ or } k=N_k\,, \quad (6.5.6a)\\ 0 & \text{otherwise}\,. \quad (6.5.6b)\end{cases}$$

The matrix $\mathbf{W}$ may be easily recognized as a **DCT-I** matrix, an orthogonal transformation (i.e. $\mathbf{W}^\mathsf{T}\mathbf{W}$ is an identity matrix), which is widely used in signal processing and can be implemented very efficiently by the fast Fourier transform (FFT) [3, 198]. The orthogonality of $\mathbf{W}$ allows to estimate $\widehat{\widetilde{\mathbf{x}}}_{:mi}$ in eq. (6.5.4) through the matrix multiplication:

$$\widehat{\widetilde{\mathbf{x}}}_{:mi} = \mathbf{W}^\mathsf{T}\widetilde{\mathbf{y}}_{:mi}\,. \tag{6.5.7}$$

The vector $\mathbf{x}_{:mi}$ is in turn obtained after scaling the first and last element of $\widehat{\widetilde{\mathbf{x}}}_{:mi}$ by a factor $\sqrt{2}$. The formula of eq. (6.5.7) has the advantage to avoid any form of matrix inversion, which generally leads to relatively stable results [123]. According to the bandpass sampling theorem, briefly introduced in Section 5.5.8, the constraint on the instrument spectral range of operation and on the OPD thickness can be partially relaxed. In fact, if the device wavenumber range is in the form $[qB_\sigma, (q+1)B_\sigma]$ with $q \in \mathbb{Z}$ and we accordingly shift the reconstruction samples $\{\sigma_l\}_{l\in[1,\dots,N_k]}$ by qB_σ from (6.5.2c), the cosine term in a_{kli} is equal to:

$$\cos\left(2\pi(\sigma_l + qB_\sigma)\delta_{ki}\right) = (-1)^{q(l-1)}\cos(2\pi\sigma_l\delta_{ki})\,. \tag{6.5.8}$$

Perfect reconstruction can thus still be obtained from the inversion procedure (6.5.7), with no modification other than by swapping the sign of the even position samples in $\widetilde{\mathbf{y}}_{:mi}$ if q is odd. A similar remark can be found if the set of OPDs $\{\delta_{ik}\}_{k\in[1,\dots,N_k]}$ are shifted by $q/(2B_\sigma)$ for which the same cosine term becomes:

$$\cos\left(2\pi\sigma_l\left(\delta_{ki} + \frac{q}{2B_\sigma}\right)\right) = (-1)^{q(k-1)}\cos(2\pi\sigma_l\delta_{ik})\,, \tag{6.5.9}$$

where the adjustments apply here to $\widehat{\widehat{\mathbf{x}}}_{:mi}$ instead.

Penalized matrix decomposition based methods

In practical situations, it is necessary to develop quick inversion procedures that are able to invert any known stochastic process in the form (6.5.1), regardless of the form of the transfer matrix $\boldsymbol{\mathcal{A}}$. I.e., its expression can be obtained by either a full characterization procedure aimed at evaluating each coefficient of $\boldsymbol{\mathcal{A}}$ or through a Bayesian inference of its parameters, as described in Section 6.4.

The ML solution $\widehat{\mathbf{X}} \in \mathbb{R}^{N'b \times N_a}$ of the stochastic process (6.5.1), if the noise is assumed to be AWGN, is given by:

$$\widehat{\mathbf{X}} = \arg \min_{\mathbf{X} \in \mathbb{R}^{N_b \times N_a}} \frac{1}{2} \|\mathbf{A}\mathbf{X} - \mathbf{Y}\|_F^2 , \tag{6.5.10}$$

where we have dropped the dependence from the third dimension, as we assume to work at a fixed incident solid angle Ω_i. Its closed form solution is $\widehat{\mathbf{X}} = \mathbf{A}^{\dagger}\mathbf{Y}$ makes use of the Moore-Penrose pseudo-inverse $\mathbf{A}^{\dagger} = (\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{T}}$, an operation that can be implemented efficiently through singular value decomposition (SVD). In its economic form, the decomposition of $\mathbf{A}$ is given by:

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Sigma}\mathbf{U}^{\mathsf{T}} , \tag{6.5.11}$$

where $\mathbf{V} \in \mathbf{R}^{N_k \times N_r}$ and $\mathbf{U} \in \mathbf{R}^{N_b \times N_r}$ are unitary matrices (i.e. $\mathbf{V}^{\mathsf{T}}\mathbf{V}$ and $\mathbf{U}^{\mathsf{T}}\mathbf{U}$ are identity matrices) and $\boldsymbol{\Sigma} \in \mathbf{R}^{N_r \times N_r}$ is a diagonal matrix, whose elements on the main diagonal $\{\check{\zeta}_r\}_{r \in [1,\dots,N_r]}$ are defined as singular values of $\mathbf{A}$ and assumed to be ordered in decreasing order. N_r denotes the rank of $\mathbf{A}$, which, in the vast majority of the practical applications, is $N_r = \min(N_k, N_b)$. The straightforward inversion method, denoted from now on as **pseudo-inversion (PINV)**, consists in computing:

$$\mathbf{A}^{\dagger} = \mathbf{U}\boldsymbol{\Sigma}^{-1}\mathbf{V}^{\mathsf{T}} , \tag{6.5.12}$$

where $\boldsymbol{\Sigma}^{-1}$ is a diagonal matrix whose elements on the main diagonal are $\{1/\check{\zeta}_r\}_{r \in [1,\dots,N_r]}$. As formulated, the problem is heavily ill-conditioned, so the PINV solution is accurate only in a completely noise-free environment, or in other words if we consider $\mathbf{x}_{:m}$ and $\mathbf{y}_{:m}$ as purely mathematical entities obtained via matrix multiplication, with no errors introduced either by the system or by the finite precision of the calculator.

The penalized matrix decomposition (PMD) methods operate with modified version of the matrix $\mathbf{\Sigma}^{-1}$, whose singular values $\{\check{\zeta}\}_{r\in[1,\dots,N_r]}$ have a well-defined upper bound. Two such methods are the most widespread:

- **Truncated singular value decomposition (TSVD)**: In the TSVD approach [114], this modified matrix is obtained through a low rank decomposition of $\mathbf{A}^{\dagger}$, which preserves the N_e smallest singular values and equates the rest to zero. In other words:

$$\check{\zeta}'_r = \begin{cases} 1/\check{\zeta}_r & \text{if } r \in [1,\dots,N_e]\ , \quad (6.5.13a) \\ 0 & \text{otherwise}\,. \quad (6.5.13b) \end{cases}$$

- **Ridge regression (RR)**: For the RR method [92], the objective is to penalize each singular value with a factor $\check{\lambda}$, so that:

$$\check{\zeta}'_r = \frac{\check{\zeta}_r}{\check{\zeta}_r^2 + \check{\lambda}^2} \quad r \in [1,\dots,N_r]\ . \quad (6.5.14)$$

  This can be shown to be equivalent to the following estimation:

$$\widehat{\mathbf{x}}_{TIK} = \arg\min_{\mathbf{x}\in\mathbb{R}^{N_b}} \frac{1}{2}\left\|\mathbf{A}\mathbf{x}-\mathbf{y}\right\|_2^2 + \check{\lambda}^2\left\|\mathbf{x}\right\|^2\ . \quad (6.5.15)$$

The optimal choice of the parameters N_e for the TSVD or $\check{\lambda}$ for RR is a critical open problem in the literature. Some techniques for its estimation involve the **generalized cross validation (GCV)** [93] and the **L-curve criterion** [113] (which were described in Section 2.2.1).

LASSO regularizer methods

While the methods proposed in the previous section have the advantage to be evaluated with a one-shot matrix multiplication, it is sometimes possible to sacrifice some computational time for a better accuracy of the results. This section focuses on iterative algorithms aimed at the resolution of problems with the framework known in the literature as LASSO [218]. In this formalism, the estimation $\widehat{\mathbf{x}}_{:m}$ of the spectrum of the m-th acquisition is given by:

$$\widehat{\mathbf{x}}_{:m} = \arg\min_{\mathbf{x}\in\mathbb{R}^{N_b}} \frac{1}{2}\left\|\mathbf{A}\mathbf{x}_{:m}-\mathbf{y}_{:m}\right\|_2^2 + \check{\lambda}\|\,\mathbb{L}(\mathbf{x}_{:m})\|_1\,. \quad (6.5.16)$$

In the expression above, the regularizer is given by the minimization of an ℓ_1 norm, which in turn tends to minimize the amount of the non-zero coefficients of the estimation $\mathbb{L}(\widehat{\mathbf{x}}_{:m})$ expressed in the transformed domain identified by a certain linear operator $\mathbb{L}$. The aforementioned transformation should be ideally realized to highlight a particular effect of sparsity of the spectrum, so that one can classify a particular approach based on the choice of $\mathbb{L}$. Three possible choices will be investigated here:

- **Discrete cosine transform (DCT)**: The operator assumes in this case the form of a multiplication by an orthogonal matrix, e.g. whose coefficients are in the form (6.5.5c). If the spectrum contains a relatively limited amount of fast oscillations, the coefficients associated with high frequencies tend to be smaller, justifying the sparsity.
- **Discrete wavelet transform (DWT)**: The implementation of $\mathbb{L}$ consists in successive decomposition of the input in its low-pass and high-pass component through quadrature mirror filter (QMF) (a cascade of fully reversible filtering and decimation). The choice of the filters determines the nature of the DWT. This work assumes 3 stages of decompositions with **Daubechies 8** analysis filters. The sparsity is here imposed as the amplitude of the coefficients associated with low-pass components is typically much higher than their high-pass counterparts.
- **Total variation (TV)**: It imposes a minimization of the difference across consecutive samples in the spectrum. The operation $\mathbb{L}(\mathbf{x})$ is defined as $\{x_{l+1} - x_l\}_{l\in[1,\dots,N_b-1]}$.

To avoid ambiguities, from now on we will label these three techniques as **LDCT**, **LDWT** and **LTV**, respectively.

Many methods are available to find a solution for the LASSO regression, but the choice here is to employ proximal methods [181]. In particular, it is possible to define a proximal operator for the Fenchel conjugate [20] $g^\star$ of the ℓ_1 norm operator $g(\mathbf{x}) = \|\mathbf{x}\|_1$ as a column vector of the same size of $\mathbf{x}$, whose l-th element is obtained via the following **soft thresholding**:

$$\mathrm{prox}_{\check{\lambda}g^\star}(x_l) = \begin{cases} -\check{\lambda} & \text{if } x_l < -\check{\lambda}\,, \\ x_l & \text{if } |x_l| < \check{\lambda}\,, \\ \check{\lambda} & \text{if } x_l > \check{\lambda}\,. \end{cases} \tag{6.5.17}$$

This is the sufficient condition to employ a solver such as Loris-Verhoeven [148], described in Section 2.3.2. By denoting with $\mathbf{x}^{(q)}$ the estimation of $\mathbf{x}_{:m}$ at the q-th iteration (and an associated dual variable $\mathbf{u}^{(q)}$), the updates are performed with the iteration:

$$\begin{cases} \mathbf{u}^{(q+\frac{1}{2})} & = \text{prox}_{\check{\lambda} g^\star}\left(\mathbf{u}^{(q)} + \check{\sigma}\,\mathbb{L}\left(\left(\mathbf{x}^{(q)}\right) - \check{\tau}\left(\mathbf{A}^\mathsf{T}\left(\mathbf{A}\mathbf{x}^{(q)} - \mathbf{y}\right) + \mathbb{L}^*\left(\mathbf{u}^{(q)}\right)\right)\right)\right), \\ \mathbf{x}^{(q+1)} & = \mathbf{x}^{(q)} - \check{\rho}\check{\tau}\left(\mathbf{A}^\mathsf{T}\left(\mathbf{A}\mathbf{x}^{(q)} - \mathbf{y}\right) + \mathbb{L}^*\left(\mathbf{u}^{(q+\frac{1}{2})}\right)\right), \\ \mathbf{u}^{(q+1)} & = \mathbf{u}^{(q)} + \check{\rho}\left(\mathbf{u}^{(q+\frac{1}{2})} - \mathbf{u}^{(q)}\right), \end{cases} \tag{6.5.18}$$

where $\mathbb{L}^*$ is the adjoint operator of $\mathbb{L}$, $\check{\sigma}$ and $\check{\tau}$ are convergence parameters such that $\check{\sigma}\check{\tau} \leq 1/\|\,\mathbb{L}\,\|_{op}$ and $1 \leq \check{\rho} < 2$ is the over-relaxation parameter. The adjoint operator for the LDCT and LDWT are equivalent with the inverse discrete cosine transform (IDCT) and inverse discrete wavelet transform (IDWT), respectively, and their operator norm is $\|\,\mathbb{L}\,\|_{op} = 1$, as both transformations are orthogonal. For the LTV, the adjoint operator $\mathbb{L}^*(\mathbf{u})$ is $\{u_{l-1} - u_l\}_{l\in[1,\dots,N_b]}$ (assuming $u_0 = 0$) and its operator norm is $\|\,\mathbb{L}\,\|_{op} = \sqrt{8}$. An implementation-friendly summary of the procedure, including a suggested initialization and common choices for the convergence parameters, is proposed in Algorithm 8.

Algorithm 8: Loris-Verhoeven algorithm for the solution of a LASSO problem.

Result:

- Spectrum Estimation: $\widehat{\mathbf{x}}_{:m} \in \mathbb{R}^{N'_b}$

Input:

- Direct model of the device: $\mathbf{A} \in \mathbb{R}^{N_k \times N'_b}$
- Linear operator: $\mathbb{L}$ (with adjoint $\mathbb{L}^*$ and operator norm $\| \mathbb{L} \|_{op}$)
- m-th acquisitions: $\mathbf{y}_{:m} \in \mathbb{R}^{N_k}$
- Maximum number of iterations: N_q

Procedure:

Define the function: $\mathrm{prox}_{\check{\lambda} g^*}(\mathbf{x}) = \min\left(\max\left(\mathbf{x}, -\check{\lambda}\right), \check{\lambda}\right)$
Define the convergence parameters $\check{\tau} = \frac{0.99}{\| \mathbb{L} \|_{op}}$ and $\check{\sigma} = \frac{1}{\| \mathbb{L} \|_{op}}$
Define the over-relaxation parameter: $\check{\rho} = 1.9$
Initialize $\mathbf{x}^{(0)} = \mathbf{A}^\mathsf{T} \mathbf{y}_{:m}$, $\boldsymbol{\ell}^{(0)} = \mathbf{x}^{(0)}$, and $\mathbf{u}^{(0)} = \mathbb{L}\left(\mathbf{x}^{(0)}\right)$
for $q = 0, ..., N_q - 1$ **do**
    $\mathbf{e}^{(q)} = \mathbf{A}^\mathsf{T}\left(\mathbf{A}\mathbf{x}^{(q)} - \mathbf{y}_{:m}\right)$
    $\mathbf{u}^{(i+\frac{1}{2})} = \mathrm{prox}_{\check{\lambda} g^*}\left(\mathbf{u}^{(q)} + \check{\sigma}\, \mathbb{L}\left(\mathbf{x}^{(q)} - \check{\tau}\left(\mathbf{e}^{(q)} + \boldsymbol{\ell}^{(q)}\right)\right)\right)$
    $\boldsymbol{\ell}^{(q+\frac{1}{2})} = \mathbb{L}^*\left(\mathbf{u}^{(q+\frac{1}{2})}\right)$
    $\mathbf{x}^{(q+1)} = \mathbf{x}^{(q)} - \check{\rho}\check{\tau}\left(\mathbf{e}^{(q)} + \boldsymbol{\ell}^{(q+\frac{1}{2})}\right)$
    $\boldsymbol{\ell}^{(q+1)} = \mathbf{u}^{(q)} + \check{\rho}\left(\boldsymbol{\ell}^{(q+\frac{1}{2})} - \boldsymbol{\ell}^{(q)}\right)$
    $\mathbf{u}^{(q+1)} = \mathbf{u}^{(q)} + \check{\rho}\left(\mathbf{u}^{(q+\frac{1}{2})} - \mathbf{u}^{(q)}\right)$
end
$\widehat{\mathbf{x}}_{:m} \leftarrow \mathbf{x}^{(N_q)}$

6.5.3 Experimental results

Experimental setup

The experimental setup for the validation of the inversion protocol is made up of the following phases [5]:

- **Input pre-processing**: The first operations aim at generating an ideally sampled set of N_a high resolution spectra arranged in a matrix $\mathbf{X} \in \mathbb{R}^{N'_b \times N_a}$, with an amount of samples N'_b sufficiently high to simulate a continuous wavenumber range;
- **Bandpass filtering**: the input is filtered to simulate the limited bandwidth of the instrument;
- **Simulated acquisition**: we simulate an ideal transfer function $\mathbf{A}' \in \mathbb{R}^{N_k \times N'_b}$ whose coefficients are obtained as sampling of an ∞-wave model; the simulated acquisition is then obtained as a matrix multiplication $\mathbf{Y} = \mathbf{A}\mathbf{X}$;
- **Noise perturbation**: if indicated, AWGN is added to the acquisition, targeting a certain SNR;
- **Model description**: Our knowledge of the system is supposed to be model with a transfer matrix $\mathbf{A} \in \mathbb{R}^{N_k \times N_b}$, which has in general less coefficients than $\mathbf{A}'$. These coefficients are also obtained by sampling the interferometer optical transfer models of table 6.5, but with generally different parameters with respect to $\mathbf{A}'$;
- **Inversion model testing**: a selection of inversion methods are tested to obtain a matrix of reconstructed spectra $\widehat{\mathbf{X}} \in \mathbb{R}^{N_b \times N_a}$, at a lower spectral resolution, such that $N_b \leq N'_b$;
- **Validation**: a validation is performed comparing $\widehat{\mathbf{X}}$ against a spectrally degraded version $\widetilde{\mathbf{X}}^{\downarrow} \in \mathbb{R}^{N_b \times N_a}$ of the input spectra, which allows for dimensional consistency; this is obtained by appropriately downsampling $\mathbf{X}$ (low pass filtering with a cutoff frequency of $N_b/(2N'_b)$ and decimating with a step of N'_b/N_b). In our experiment, the mean and STD the reconstructed spectra are equalized with respect to the reference before the comparison, as we are just interested in the shape and not the absolute intensity of the reconstructed spectra.

[5] In this experimental section, the operations of filtering are performed through the convolution with a 23-rd order Butterworth digital filter

Their m-th columns of $\mathbf{X}$ and $\widetilde{\mathbf{X}}^{\downarrow}$ are compared by evaluating the RMSE and the spectral angle mapper (SAM), defined as:

$$\mathrm{SAM}_m = \arccos\left(\frac{\left\langle \widehat{\mathbf{x}}_{:m}, \widetilde{\mathbf{x}}^{\downarrow}_{:m} \right\rangle}{\|\widehat{\mathbf{x}}_{:m}\|_2 \left\|\widetilde{\mathbf{x}}^{\downarrow}_{:m}\right\|_2}\right), \tag{6.5.19}$$

which are then averaged across all available spectra to provide a global index.

Baseline test

The first part of the experiment involves describing the system with a given standard set of parameter and we suppose our modeled knowledge of the optical device matches perfectly the one of the system. This test is defined here as **baseline**, to assess the performances of the methods in close to ideal conditions.

- **Input pre-processing**: the input is obtained from a set of $N_a = 22$ solar spectra measured at the ground level at different times of the day. The intensity samples were first transformed from the wavelength to the wavenumber domain and interpolated with a cubic spline kernel over a regularly spaced interval. The obtained input $\mathbf{X}$ has a spatial resolution of 1500 samples/um^{-1} (i.e. 1501 samples in the range $[1, 2]$ um^{-1}, if both ends are included). Some more advanced domain transformations could have been considered, but it was decided for this method as a good compromise between common practice [58] and ease of reproducibility.
- **Bandpass filtering**: the input is filtered to limit the available samples to the range $[1, 2]$ um^{-1};
- **Simulated acquisition**: the transfer matrix is chosen to simulate the parameters of the PROTO-1 prototype. Specifically, we choose a set of $N_k = 215$ OPDs increasing from zero with a step of 200 nm. The reflectivity is set to $\mathcal{R} = 0.13$, the phase shift to zero and the system gain to 1;
- **Noise simulation**: no noise is added to the simulation (this is equivalent to choosing an SNR of ∞);
- **Model description**: We assume here that the parameters of the matrix $\mathbf{A}$ fully match the ones of the simulation description of the optical system. We also assume that the reconstruction samples are equal to the amount of interferometers ($N_b = N_k = 215$);

- **Reconstruction test**: From the SVD class, we test the RR, and TSVD, while from the LASSO framework, we LASSO-DCT (LDCT), LASSO-DWT (LDWT), and LASSO-TV (LTV). We also include the previously unmentioned TSVDL and RRL, which are simple variation of the TSVD and RR, but an automatic determination of their characteristic penalization parameters with the L-curve method [113]. For the non-automatic parameter determination setups, the best regularization term $\check{\lambda}$ for the LASSO methods and for the RR, as well as the amount of singular value (s.v.) for the TSVD have been decided in terms of the best measured RMSE in the validation. For the LASSO based methods, we experienced better performances by normalizing each column of $\mathbf{Y}$ and $\mathbf{A}$ in terms of their mean; this pre-processing phase is assumed in every test where it can apply.

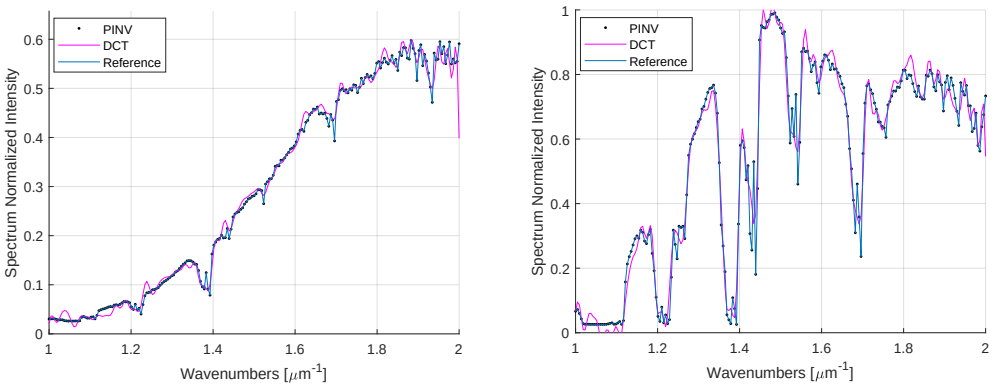
In this experiment, the solar spectra were chosen as ideal inputs as they are a typical product to reconstruct for ImSPOC acquisitions. In the envisioned application setup, the user can analyze the absorption ranges in the reconstructed spectra to detect the concentration of a given gas component.

A first preliminary test, intended as a sanity check, is performed by matching the amount of simulation and reconstruction samples ($N_b' = N_b = 215$). It can be easily verified that in ideal conditions it is possible to obtain a perfect reconstruction with a PINV, as shown in Fig. 6.22. All PMD-based and variational techniques were tested but not visualized as the most appropriate choice of their respective parameters reduces exactly to the PINV technique. The analysis can also provide a quick assessment of the accuracy of the design principles of the FTS through inversion technique described in Section 6.5.2, labeled as DCT. This is unfortunately the only setup that allows this analysis, as the DCT does not allow for particular versatility.

The results of the true baseline test are shown in Table 6.12 for the quantitative analysis and in Fig. 6.23 for the visual comparison.

The analysis of the first group of results of Table 6.12 shows that the performances of PINV method already degrade critically in the baseline simulation. As the reconstructed spectra barely resembles the reference, it was decided not to show the PINV results in any visual comparison, nor in the following section.

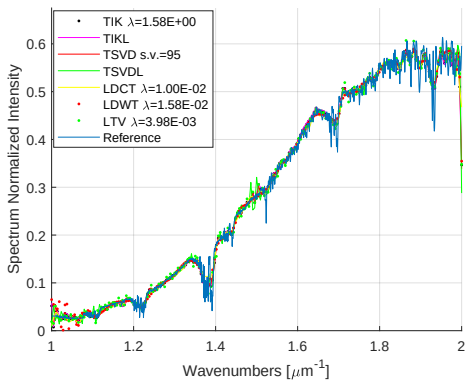
The LDCT outperforms all of its competitors; this is a reasonable result, since solar spectra are likely to show sparsity in the DCT domain, as a cosine oscillation is a reasonable first order approximation of its characteristic curve. Some other methods,



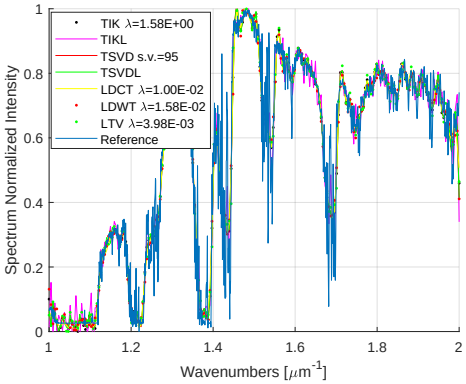
(a) 16-th spectrum **(b)** 22-th spectrum

Fig. 6.22. Reconstruction of 2 sample input spectrum with perfect match between the simulation and the reconstruction model. Black dots mark reconstructed points with PINV, which perfectly follow the blue reference (within machine rounding errors).

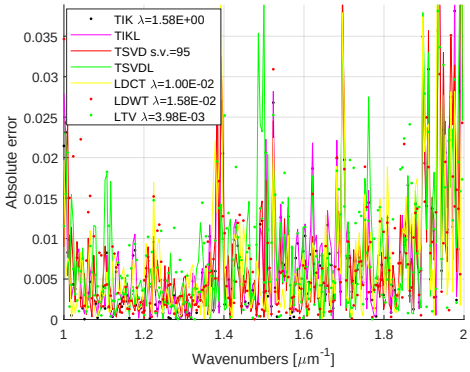
such LDWT and those based on the PMD, present unwanted oscillations, especially for low intensities.



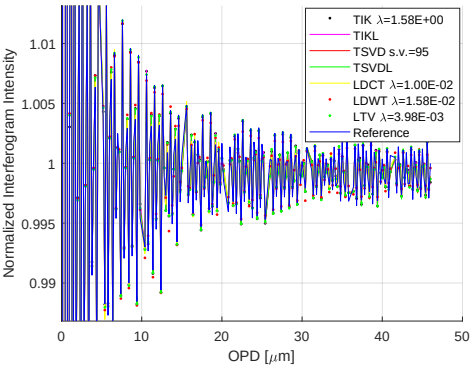
(a) Reconstruction of the 16-th spectrum



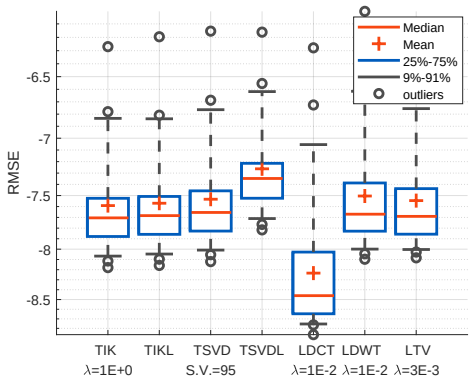
(b) Reconstruction of the 22-nd spectrum



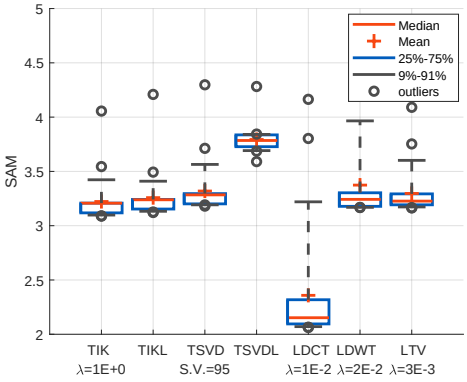
(c) Absolute error on the 16-th spectrum



(d) Acquired and reconstruction model generated 16-th interferogram



(e) RMSE comparison



(f) SAM comparison

Fig. 6.23. Visual comparison of the inversion of simulated interferograms, for the baseline setup described in Section 6.5.3.

Table 6.12. Reconstruction results of the 22 simulated acquisitions with input solar spectra, for the baseline setup (Section 6.4.4) and the model mismatches described in the sideline bars. Best results are in bold, second best ones are underlined.

		Method	Parameter	RMSE $\times 10^{-3}$	SAM
		Ideal		0	0
Baseline		**PINV**		37.6044 ± 18.9494	26.707 ± 0.503
		DCT		0.8817 ± 0.5185	4.022 ± 0.193
		RR	$\check{\lambda} = 1.58E+00$	$\underline{0.5759 \pm 0.3740}$	$\underline{3.223 \pm 0.210}$
		RRL		0.5906 ± 0.3965	3.258 ± 0.228
		TSVD	s.v.$= 95$	0.6155 ± 0.4200	3.319 ± 0.246
		TSVDL		0.7723 ± 0.4066	3.794 ± 0.129
		LDCT	$\check{\lambda} = 1.00E-02$	$\mathbf{0.3615 \pm 0.4123}$	$\mathbf{2.370 \pm 0.493}$
		LDWT	$\check{\lambda} = 1.58E-02$	0.6570 ± 0.5102	3.374 ± 0.341
		LTV	$\check{\lambda} = 3.98E-03$	0.6047 ± 0.3935	3.297 ± 0.219
Transfer model	2-wave	**RR**	$\check{\lambda} = 2.51E+00$	0.5536 ± 0.3786	3.144 ± 0.237
		TSVD	s.v.$= 90$	$\underline{0.5187 \pm 0.4155}$	$\underline{2.998 \pm 0.329}$
		LDCT	$\check{\lambda} = 1.00E-02$	$\mathbf{0.3698 \pm 0.4068}$	$\mathbf{2.409 \pm 0.482}$
		LDWT	$\check{\lambda} = 6.31E-03$	0.5611 ± 0.4706	3.098 ± 0.369
		LTV	$\check{\lambda} = 2.51E-03$	0.5476 ± 0.3596	3.130 ± 0.216
	3-wave	**RR**	$\check{\lambda} = 1.58E+00$	$\underline{0.5749 \pm 0.3757}$	$\underline{3.218 \pm 0.213}$
		TSVD	s.v.$= 95$	0.6172 ± 0.4227	3.323 ± 0.249
		LDCT	$\check{\lambda} = 1.00E-02$	$\mathbf{0.3593 \pm 0.4124}$	$\mathbf{2.360 \pm 0.497}$
		LDWT	$\check{\lambda} = 1.58E-02$	0.6514 ± 0.5034	3.362 ± 0.337
		LTV	$\check{\lambda} = 3.98E-03$	0.6065 ± 0.3959	3.298 ± 0.226
OPD standard deviation	10 nm	**RR**	$\check{\lambda} = 1.00E+00$	$\underline{0.7076 \pm 0.5454}$	$\underline{3.511 \pm 0.361}$
		TSVD	s.v.$= 105$	0.7603 ± 0.5700	3.650 ± 0.352
		LDCT	$\check{\lambda} = 1.00E-02$	$\mathbf{0.6243 \pm 0.6433}$	$\mathbf{3.169 \pm 0.565}$
		LDWT	$\check{\lambda} = 6.31E-03$	0.8299 ± 0.7020	3.759 ± 0.461
		LTV	$\check{\lambda} = 1.00E-03$	0.7128 ± 0.5441	3.525 ± 0.363
	20 nm	**RR**	$\check{\lambda} = 3.98E+00$	$\mathbf{1.2906 \pm 0.7069}$	$\mathbf{4.886 \pm 0.172}$
		TSVD	s.v.$= 95$	1.8649 ± 0.8639	5.945 ± 0.188
		LDCT	$\check{\lambda} = 2.51E-02$	$\underline{1.5917 \pm 1.2386}$	$\underline{5.244 \pm 0.551}$
		LDWT	$\check{\lambda} = 2.51E-03$	1.8516 ± 0.9407	5.888 ± 0.172
		LTV	$\check{\lambda} = 6.31E-04$	1.8112 ± 0.8249	5.862 ± 0.174
	30 nm	**RR**	$\check{\lambda} = 1.00E+01$	$\mathbf{2.2721 \pm 1.6962}$	$\mathbf{6.305 \pm 0.652}$
		TSVD	s.v.$= 95$	4.6225 ± 2.0674	9.367 ± 0.184
		LDCT	$\check{\lambda} = 2.51E-02$	$\underline{4.0214 \pm 2.5657}$	$\underline{8.489 \pm 0.555}$
		LDWT	$\check{\lambda} = 3.98E-03$	4.5066 ± 2.0480	9.238 ± 0.161
		LTV	$\check{\lambda} = 1.00E-03$	4.5466 ± 2.0774	9.274 ± 0.156
Wavelength range	$[400, 1000]$ nm	**RR**	$\check{\lambda} = 6.31E+00$	$\underline{1.2320 \pm 0.6602}$	4.785 ± 0.162
		TSVD	s.v.$= 85$	1.2972 ± 0.5859	4.970 ± 0.216
		LDCT	$\check{\lambda} = 2.51E-02$	$\mathbf{0.5843 \pm 0.5632}$	$\mathbf{3.084 \pm 0.502}$
		LDWT	$\check{\lambda} = 6.31E-02$	1.2622 ± 0.9451	$\underline{4.709 \pm 0.440}$
		LTV	$\check{\lambda} = 1.58E-02$	9.9067 ± 3.4245	14.066 ± 1.742
	$[380, 1050]$ nm	**RR**	$\check{\lambda} = 6.31E+00$	$\underline{1.2653 \pm 0.6630}$	4.859 ± 0.161
		TSVD	s.v.$= 85$	1.3198 ± 0.5936	5.016 ± 0.223
		LDCT	$\check{\lambda} = 2.51E-02$	$\mathbf{0.6262 \pm 0.5432}$	$\mathbf{3.245 \pm 0.431}$
		LDWT	$\check{\lambda} = 6.31E-02$	1.2657 ± 0.9598	$\underline{4.712 \pm 0.453}$
		LTV	$\check{\lambda} = 1.58E-02$	10.1759 ± 3.5561	14.231 ± 1.650

Model mismatches

In this section, we investigate the robustness of the inversion framework when facing specific sources of nonideality, which correspond to mismatches between the real expression of the transfer matrix and our reconstruction model. To this end, the parameters of our possible tests are modified from our baseline test setup; these changes include:

- **Bandwidth of the instrument**: for which the input spectrum is filtered in a larger bandwidth than that of the instrument. Compared to the baseline setup ($[500, 1000]$ nm), we also consider wavelength bandwidth ranges of $[400, 1000]$ and $[380, 1050]$ nm;
- **Reconstruction model**: where we consider different transfer models for the construction of the reconstruction transfer matrix $\mathbf{A}$. Other than the baseline (∞-wave model), we also consider the 2-wave and 3-wave;
- **Uncertain knowledge of the OPD**: where we suppose that the OPD of the reconstruction matrix $\mathbf{A}$ is known with some uncertainty, which is simulated by adding a zero mean Gaussian noise to the nominal value of the OPD. Compared to the baseline setup (perfect knowledge of the OPD), we also consider that the Gaussian noise perturbs these values with a STD of 10, 20 and 30 nm.

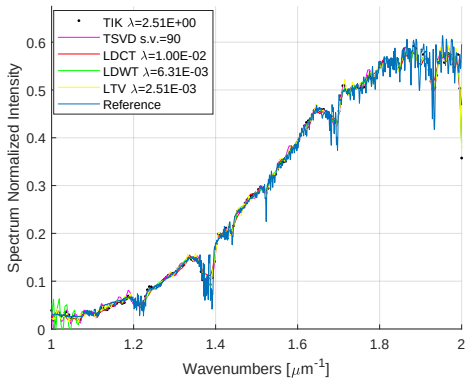
For these tests, the results of the reconstruction are shown in Table 6.12, together with the baseline results. The experiments show a significant degradation in the quality of the reconstruction already with mismatches of the order of 20 nm, with the RR method featuring an outstanding robustness to this class of mismatch (Fig. 6.24c to 6.24d). A very accurate knowledge of the OPD associated with each interferometer is then crucial for a good reconstruction, justifying the efforts on the research devoted to an accurate transfer model characterization in Section 6.4. No particular degradation in quality was shown by employing a 2-wave model (Fig. 6.24a to 6.24b), which is a reasonable consequence of simulating the response of an optical instrument with low finesse/low reflectivity. More severe, but still controllable, distortions are caused by the wavelength bandwidth mismatch (Fig. 6.24e to 6.24f).

We also repeat the baseline experiment adding a **noise contribution** to the acquisitions, such that we target a SNR of 40 and 50 dB (the baseline can be considered with an SNR equal to ∞). Compared to the baseline test, where the number of reconstruction samples was fixed to $N_b = 215$, we also test for 151 and 301 samples, and the quantitative results are show in Table 6.13.

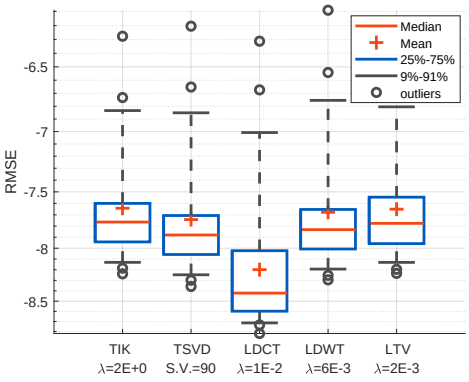
Fig. 6.25a to 6.25b show that, in noiseless environments, a very accurate reconstruction is possible even for 301 samples, which is equivalent to a smaller step size than the limit imposed by the Nyquist-Shannon sampling theorem. The missing information in the data term is in fact compensated by the prior. The sensitivity of the device to the noise is however still a relevant issue; the results show that the algorithms are well performing just for very high SNRs (e.g. equal to 50 dB in Fig. 6.25c to 6.25d). To deal with this scenario, the regularization term is more predominant and the performances degrades almost immediately for lower SNR.

The effect of the noise added to interferogram has to be kept under control, as the relevant information contained within the oscillation around a given mean value must be extremely accurate. For low finesse devices this oscillation is quite limited (e.g., Fig. 6.23d), so the interferogram is easily distorted by the uncertainty introduced by the additive noise.

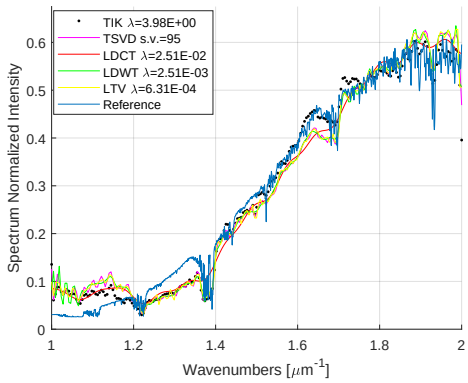
Is it however worth noting that the low finesse has a complementary beneficial effect that has been ignored within this simulation framework. As described in Section 5.4.6, low finesse interferometers allow to collect a larger amount of photons on the detectors due to their higher throughput, increasing the amount of energy of the desired signal and in turn the SNR.



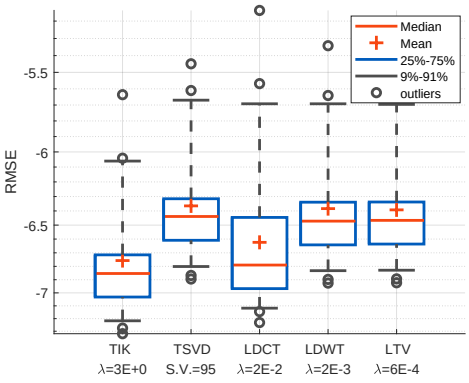
(a) 2-waves model: reconstruction of the 16-th interferogram



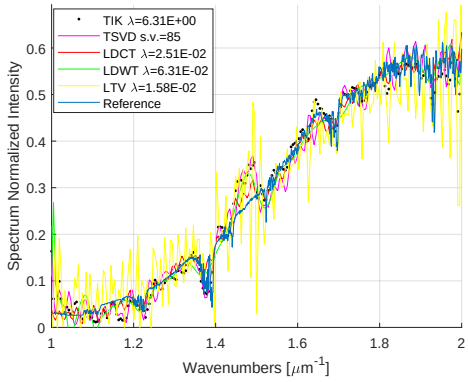
(b) 2-Wave model: RMSE comparison



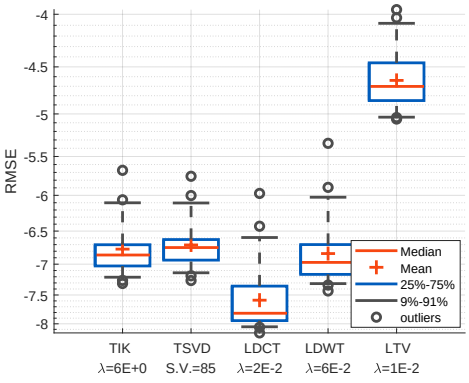
(c) OPD STD 20 nm: reconstruction of the 16-th interferogram



(d) OPD STD 20 nm: RMSE comparison

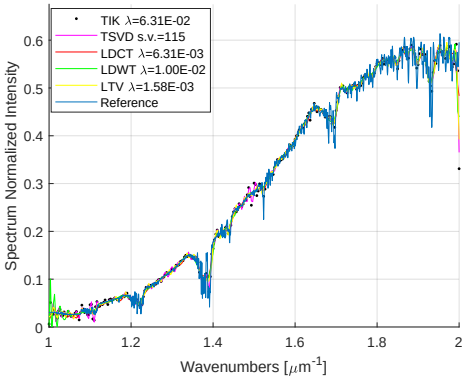


(e) Instrument range $[380, 1050]$ nm: reconstruction of the 16-th Interferogram

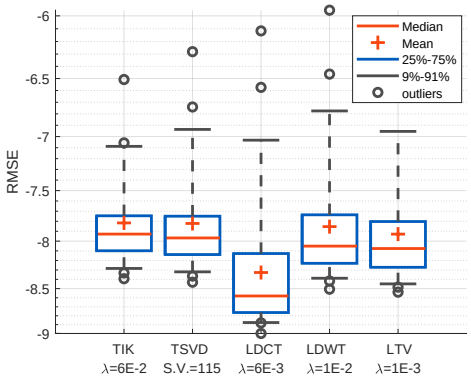


(f) Instrument range $[380, 1050]$ nm: RMSE comparison

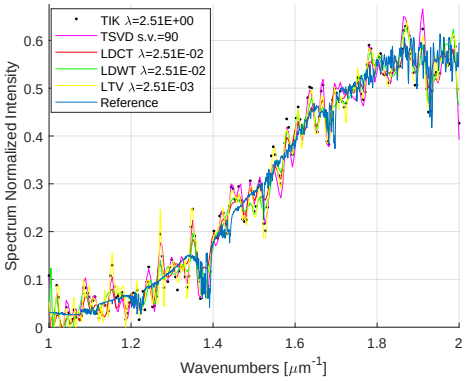
Fig. 6.24. Visual comparison results for the inversion procedures described in Section 6.5.3, in the case of model mismatches.



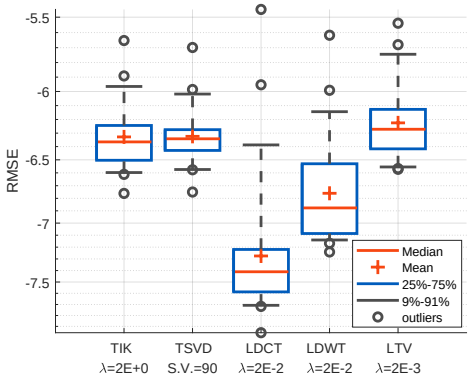
(a) SNR ∞ dB, $N_s = 301$: reconstruction of the 16-th interferogram



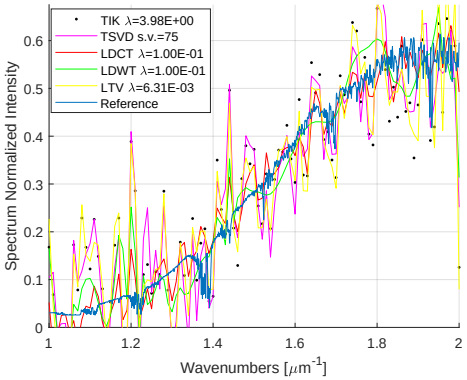
(b) SNR ∞ dB, $N_s = 301$: RMSE comparison



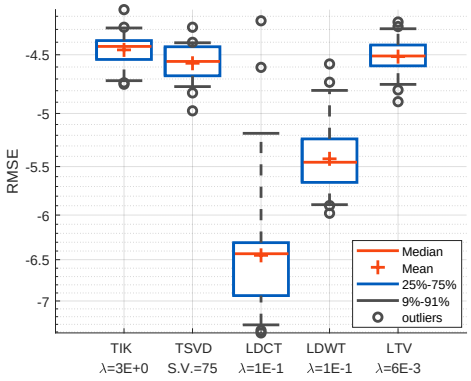
(c) SNR 50 dB, $N_s = 215$: reconstruction of the 16-th interferogram



(d) SNR 50 dB, $N_s = 215$: RMSE comparison



(e) SNR 40 dB, $N_s = 101$: reconstruction of the 16-th interferogram



(f) SNR 40 dB, $N_s = 101$: RMSE comparison

Fig. 6.25. Visual comparison results of the inversion for noisy interferograms and different amount N_s of reconstruction samples, with the setup described in Section 6.5.3

Table 6.13. Validation comparison of SNR for the inversion of 22 simulated interferograms obtained from solar input spectra. When non specified, the variables are chosen according to the baseline setup described in Section 6.4.4. The cross-comparison between setups separated by double horizontal line has to be taken with caution, as they employ different references. Best results for each category are in bold, second best ones are underlined.

		Method	Parameter	RMSE $\times 10^{-3}$	SAM
		Ideal		0	0
Rec. samples: 215	SNR: 50 dB	**RR**	$\check{\lambda} = 2.51E+00$	1.8404 ± 0.5087	6.098 ± 0.751
		TSVD	s.v.= 90	1.8380 ± 0.4522	6.111 ± 0.748
		LDCT	$\check{\lambda} = 2.51E-02$	$\mathbf{0.8680 \pm 0.8877}$	$\mathbf{3.827 \pm 0.742}$
		LDWT	$\check{\lambda} = 2.51E-02$	$\underline{1.2727 \pm 0.6781}$	$\underline{4.898 \pm 0.477}$
		LTV	$\check{\lambda} = 2.51E-03$	2.0571 ± 0.6411	6.422 ± 0.797
	SNR: 40 dB	**RR**	$\check{\lambda} = 6.31E+00$	11.0452 ± 1.2430	15.290 ± 2.672
		TSVD	s.v.= 75	9.8897 ± 1.4565	14.385 ± 2.261
		LDCT	$\check{\lambda} = 1.58E-01$	$\mathbf{2.6077 \pm 4.3156}$	$\mathbf{5.856 \pm 2.318}$
		LDWT	$\check{\lambda} = 1.58E-01$	$\underline{5.1909 \pm 1.8597}$	$\underline{10.131 \pm 1.056}$
		LTV	$\check{\lambda} = 6.31E-03$	14.6268 ± 2.5058	17.532 ± 3.124
Reconstruction samples: 151	SNR: ∞ dB	**RR**	$\check{\lambda} = 1.58E+00$	$\underline{0.7491 \pm 0.3288}$	$\underline{3.783 \pm 0.173}$
		TSVD	s.v.= 90	0.7716 ± 0.3265	3.848 ± 0.194
		LDCT	$\check{\lambda} = 1.58E-02$	$\mathbf{0.3531 \pm 0.3883}$	$\mathbf{2.357 \pm 0.465}$
		LDWT	$\check{\lambda} = 1.00E-02$	0.7719 ± 0.3310	3.839 ± 0.132
		LTV	$\check{\lambda} = 2.51E-03$	0.7662 ± 0.3102	3.844 ± 0.208
	SNR: 50 dB	**RR**	$\check{\lambda} = 2.51E+00$	1.9653 ± 0.5022	6.312 ± 0.782
		TSVD	s.v.= 90	1.9286 ± 0.4070	6.275 ± 0.777
		LDCT	$\check{\lambda} = 2.51E-02$	$\mathbf{0.7823 \pm 0.8800}$	$\mathbf{3.582 \pm 0.758}$
		LDWT	$\check{\lambda} = 2.51E-02$	$\underline{1.4454 \pm 0.9789}$	$\underline{5.122 \pm 0.558}$
		LTV	$\check{\lambda} = 2.51E-03$	2.0046 ± 0.5933	6.355 ± 0.837
	SNR: 40 dB	**RR**	$\check{\lambda} = 6.31E+00$	11.1782 ± 1.3750	15.380 ± 2.732
		TSVD	s.v.= 75	10.0233 ± 1.4539	14.486 ± 2.308
		LDCT	$\check{\lambda} = 1.58E-01$	$\mathbf{2.4822 \pm 4.7844}$	$\mathbf{5.328 \pm 2.539}$
		LDWT	$\check{\lambda} = 1.58E-01$	$\underline{4.6960 \pm 2.1472}$	$\underline{9.549 \pm 1.081}$
		LTV	$\check{\lambda} = 6.31E-03$	13.0440 ± 2.2480	16.581 ± 3.128
Reconstruction samples: 301	SNR: ∞ dB	**RR**	$\check{\lambda} = 6.31E-02$	0.4577 ± 0.2915	2.879 ± 0.175
		TSVD	s.v.=165	0.4727 ± 0.3096	2.919 ± 0.193
		LDCT	$\check{\lambda} = 6.31E-03$	$\mathbf{0.3608 \pm 0.4940}$	$\mathbf{2.282 \pm 0.618}$
		LDWT	$\check{\lambda} = 1.00E-02$	0.5124 ± 0.5530	2.858 ± 0.528
		LTV	$\check{\lambda} = 1.58E-03$	$\underline{0.4465 \pm 0.3986}$	$\underline{2.741 \pm 0.380}$
	SNR: 50 dB	**RR**	$\check{\lambda} = 3.98E+00$	1.7950 ± 0.5694	5.999 ± 0.703
		TSVD	s.v.=90	1.8031 ± 0.5344	6.031 ± 0.740
		LDCT	$\check{\lambda} = 3.98E-02$	$\mathbf{0.9326 \pm 1.2280}$	$\mathbf{3.750 \pm 0.959}$
		LDWT	$\check{\lambda} = 2.51E-02$	$\underline{1.3531 \pm 0.7382}$	$\underline{5.076 \pm 0.608}$
		LTV	$\check{\lambda} = 2.51E-03$	2.3046 ± 0.6463	6.833 ± 0.936
	SNR: 40 dB	**RR**	$\check{\lambda} = 1.00E+01$	10.9892 ± 1.2353	15.244 ± 2.625
		TSVD	s.v.=75	9.8406 ± 1.5099	14.344 ± 2.242
		LDCT	$\check{\lambda} = 1.58E-01$	$\mathbf{2.7090 \pm 3.6938}$	$\mathbf{6.382 \pm 2.086}$
		LDWT	$\check{\lambda} = 1.58E-01$	$\underline{5.9992 \pm 1.7326}$	$\underline{10.985 \pm 1.325}$
		LTV	$\check{\lambda} = 6.31E-03$	16.6753 ± 2.4960	18.774 ± 3.410

6.6 Conclusions and future perspectives

In this chapter, we described the signal processing operations that are required for the treatment of a raw acquisition with an image spectrometer based on simultaneous acquisitions with an array of FP interferometers, aimed at the reconstruction of the input radiance spectra.

The first step includes a co-registration to align each portion of the acquisition assigned to each element, whose position on the focal plane can be estimated with a set of region estimation protocols based on mathematical morphology. The image registration was approached via a point mapping technique assisted by laser spot acquisitions. The spatial coordinates transformation function was modelled as a bidimensional polynomial and validated with the SSIM index over real acquisition. We verified that the quadratic or cubic degrees offer the best accuracy, avoiding overfitting on the available illuminated spots. The algorithm is also not particularly computationally intensive and just requires interpolation over regular grids.

The problem of model characterization, aimed at estimating the expression of the transfer function which characterizes the device in terms of wavelength, was approached by imposing a physical model based on the energy balance of the incident and emerging rays under ideal conditions and by estimating its parameters through Bayesian inference. A variety of different approaches aimed at this estimation were proposed and tested over both simulated and real data, obtained in heterogeneous conditions, showing the advantages of employing the Airy's distribution models for a more accurate description of the behaviour of the device. If computational speed is the priority, solving the ML problem provides reasonable performances, which can be iteratively refined with procedures based on the GNA, if necessity arises.

Finally, some initial consideration aimed at spectrum inversion were provided for acquisition in a fixed incident direction, providing a basic framework to approach it systematically; for the inversion of typical spectra, a LASSO based inversion with DCT regularization seems to provide the most consistent and robust results with whichever source of mismatch between the acquisition and the reconstruction model, except for the case of OPD uncertainty, for which a more straightforward Tikhonov regularization exhibited less overall degradation in the reconstructed spectrum.

This chapter is mostly intended to provide some baseline protocols to provide user-ready products that can be exploited to recover information on the scene. Possible evolutions may either involve keeping each processing layer separate or to merge them to take advantage of the joint optimization. For the first set of techniques,

which are better suited to be implemented in modular fashion, one desirable goal is to make the process independent from any laboratory characterization of the device, e.g. by forcing the registration of the subimages through the detection and matching of specific characteristic features. Additionally, one could setup possible inversions with partial or missing knowledge of the transfer function of the device, which could be characterized dynamically with a black box approach. This procedure may exploit a mixed approach of data-driven and machine/deep learning [179, 15, 159, 150].

The model to characterize can also be refined quite straightforwardly by introducing a wavelength dependency on the reflectivity term, as well as considering additional a priori, e.g. introducing the information that adjacent interferometers in the staircase design of ImSPOC usually feature close thicknesses. It may also be useful not to be subject to the liabilities of a faulty optimization of the non-convex objective function used for the GNA, by relaxing the problem over a convex hull.

The inversion procedure was considered for a fixed direction of incidence, although much better performances are expected to be achieved if the inversion introduces a spatial regularization term, e.g. by imposing any form of TV [48, 65] that was considered in the rest of the thesis.

With regard to merging steps together, a more encompassing framework may employ a joint optimization of both the coefficients of the transfer matrix and of the spectrum to reconstruct, with an appropriate procedure of alternating optimization, such as alternating least squares (ALS) [45] or majorization-minimization (MM) [212]. Similarly, the spatial calibration may be considered jointly to the inversion by introducing a spatial transformation relationship within the functional to minimize.

7 Conclusions

7.1 Summary

This thesis presented a series of approaches for the processing of data acquired with nonconventional optical devices; the investigated prototypes required approaches based on computational imaging as the desired products and the acquisitions are in different domains. The problem was addressed with a physical based approach, where we separated the analysis into a characterization, that defines the model of the optical transformation performed by the instrument, and an inversion, to recover the datacube of hyperspectral (HS) images, which describes the spectral radiance incident to the instrument. The thesis was focused on the data processing of acquisitions captured by two innovative optical device concepts:

- the multiresolution color filter array acquisition (MRCA), a compressed acquisition system based on the color filter array (CFA) technology, whose focal plane array (FPA) is composed by detectors with different characteristic resolution.
- the image spectrometer on chip (ImSPOC), a snapshot spectro-imaging system, made up of an array of Fabry-Pérot (FP) disposed over a staircase pattern, overlaid to an imaging system.

The analysis was laid out as follows:

- In Chapter 1, we presented the context of compressed imaging, of HS imaging, and the envisioned domains of applications;
- In Chapter 2, we provided a review of the inversion techniques based on proximal operators [181], describing the importance of the regularization process for an accurate reconstruction of the desired products;
- In Chapter 3, we introduced the problems related to multimodality of the data, whose main challenges are an accurate registration of data and a fusion of the information with different characteristics;

- In Chapter 4, we proposed a model for the optical transformations performed by the MRCA device, analyzing the performances of a set of inversion algorithms based on a variational approach and proposing some initial designs for the dispositions of the elements on the FPA;
- In Chapter 5, we introduced the optical concepts necessary for the interferometry and proposed a model of the acquisitions taken with the ImSPOC prototype;
- In Chapter 6, we proposed a full pipeline of procedures for the inversion of ImSPOC acquisitions, which include the stacking and co-registration of subimages, the calibration of the direct model and the spectrum inversion.

In order to improve the versatility of the proposed approaches, the inversion problems that were setup in this work were based on the theory of proximal operators [181]. While the analysis of the speed of convergence was not a focus of this thesis, the recent results of the over-relaxation of the algorithms [50], allowed to run the test relatively quickly and the choice of the convergence parameters a relatively painless process. On the other hand, the choice of the regularization parameter is strongly reliant on the intensity of the noise and may require an analysis of the measurement conditions to be set up properly. Some more advanced techniques of inversion, such as elastic nets [41] and total generalized variation (TGV) [31] require to set up more than one parameter, which may further complicate the analysis.

For the MRCA we proposed:

- A mathematical formulation [187] of the optical transformation of the device, to be exploited for a Bayesian inversion. This framework is versatile for any composed design of the CFA pattern that makes up the FPA;
- An inversion scheme whose reconstruction scheme based on a collaborative total variation (CTV) regularization [65, 66]. The scheme is able to deal jointly with the problem of reconstruction and fusion of the virtual images at different resolutions;
- A preliminary assessment of the most effective mosaicing patterns for the distribution of the sensors over the FPA [188], based on non-redundancy principles [46].

The proposed framework has state-of-the-art performances for periodic HS masks and the total variation (TV) reconstruction allows for an impressive suppression

of the artifacts introduced by the mosaicing. As the structure of the data is internally redundant, the take-home message for the regularization is that, if the mean square error (MSE) is used as a metric for the data fidelity term, the reconstruction has to provide some mechanism for noise whitening in the multi-dimensional domain spanned by the final product, as the noise is generally distributed unevenly across different dimensions, while the Bayesian interpretation supposes additive white Gaussian noise (AWGN). The employment of a nuclear norm is a good example of the benefits of this whitening effect both in the spectral and in the spatial domains [65].

While the proposed approach is applicable to every composed design of the CFA pattern that makes up the FPA, the performances are not even across all its designs. We tested that deterministic patterns tend to outperform random distributions, and the sensors have to be spaced evenly to reach the best performances. An appropriate ratio of the high resolution image (HRI) and low resolution image (LRI) sensors is also necessary to reach a good balance between spatial and spectral accuracy of the reconstructed products.

The effectiveness of the proposed approach is limited in comparison to more specialized demosaicing procedures in the case of masks with a dominant band, such as in the case where we are facing a classical problem of demosaicing of Bayer patterns, which are already accompanied by extremely specialized high performances reconstruction methods [166]. For this cases, we experienced a limited effectiveness of the inversion procedure along curve borders, which is a limitation of the demosaicing part of the algorithm; to face this issue we propose instead to use a cascade of classical methods for the inversion, while limiting our framework to the fusion of the obtained products. Additionally, the compression ratio obtained of MRCA products is not comparable to classic software compression approaches, which suggests that proposed physical structure still has margins of improvement to further reduce the redundancy of the acquisitions.

For the ImSPOC concept, we proposed:

- A classical framework of operations for the alignment of subimages, based on a polynomial geometry transformation function;
- Three approaches for the spectral calibration of the interferometers which compose the staircase pattern of the device, based on the estimations of the parameters of an Airy's distribution [59, 190];
- A framework for the inversion of a single interferogram, employing simple regularization schemes, with an analysis of the mismatches between our

knowledge of the parameters of the model and the real ones which characterize the device;

The procedure of co-registration showed that the degree of the polynomial of the analytical function that describes the geometry transformation must be chosen as a good compromise between flexibility in shifting the structures of the subimages, and avoiding overfitting of the training data. While the procedure gave very accurate results, a possible limitation of this approach is that it requires a complicated ad-hoc calibration procedure which allows to map points across different subimages, as we did not provide a procedure to match features in the case of a generic acquisition.

For the spectral calibration, the procedure can be used both to highlight faults in the manufacture of the device, such as interferometer thicknesses that do not match nominal values and to refine the direct model to be used for the inversion. In our analysis, we showed the advantages of employing a more accurate description of the optical transformation model, as the Airy distribution is a more accurate representation of the interferences of the rays within the FP etalons. Additionally, technique based on nonlinear regression allow to explore a continuous space for the parameters to infer, which yields better performances in the reconstruction. However, some nonidealities are not taken into consideration in our considered approach, which may be limiting for the accuracy of the inversion; nonlinear regression is also reliant on a proper initialization, which may cause the solution not to converge.

The proposed inversion methods were chosen as a compromise between accuracy of the reconstructed products and computational speed, and the best results were obtained with sparse-inducing discrete cosine transform (DCT) regularizations. For this method, the memory storage requirement, even for iterative-based approaches, is relatively limited, as the current estimation is overwritten at every iteration. The experiments proved a relative robustness of the proposed algorithms to an "imperfect" knowledge of the parameters of the system, except for the thickness of the interferometers, which has to be known with a relatively high degree of precision. Unfortunately, the analysis also showed that interferograms are relatively sensitive to additive noise, demanding for more sophisticated regularization techniques and technological control of the measurements.

7.2 Perspectives for future works

In this section we list some macro-subjects of interest for the extension of this work that employ different approaches compared to the rationale that was followed in

this work; these considerations are intended as suggestions both to extend the scope of this thesis and to kickstart future projects.

We separate these approaches in device-specific applications and general direction of research. Some possible perspectives to investigate for the MRCA device include:

- **Regularization protocols**: the regularization protocols chosen in our experiments were thought to minimize the amount of parameters to manually set up. However, some advanced approaches include some variations on the TV [31, 48], the Elastica minimization [41, 158], and more [128, 243];
- **Mask design**: The design of the masks employed in this work is just based on general considerations that act as rule of thumb guidelines. One more in depth discussion may involve a deeper understanding of compression sensing, i.e. by employing the framework of compressing statistical learning [131, 100];
- **Sensor area-dependent design**: As the MRCA is composed by sensors with different characteristics, it is expected that the area of the sensors at different resolutions may cover different a different area on the FPA. It would be interesting to extend the current framework to consider this case, for which one can model CFA patterns like the Quad Bayer.

Some possible developments on the data processing of ImSPOC protocols include instead:

- **More sophisticated regularization**: In our inversion analysis, the analysis was limited to a simplified case in which the interferogram relative pixel was inverted separately. This approach has the intrinsic limitation that it ignores the correlation of the information between adjacent pixels; the model can be extended to take into account the spatial information by adding a spatial regularization, even with simple approaches like the TV [48, 65];

- **Acquisition model refinement**: The spectral calibration acquisition showed some optical effects that were not taken in consideration in our model, which mostly involve a dependence of the parameters by the wavelength. Additionally some parameters share some common features between different interferometers, while our characterization considered them as separate;

- **Initialization of nonlinear algorithms**: Some of the inversion techniques that were developed for this thesis involve some form of linear regression; with our proposed formalism, the convergence is reliant on an appropriate initialization. More sophisticated methods allow however to avoid this issue, which may involve some sort of nonlinear programming [22].

Furthermore, some general purpose research directions are described in this list:

- **Multimodal Fusion**: The main target of employing specialized technologies aims at resolving images with high spatial and spectral resolution. The limitation of a given technology can be overcome with the support of an additional acquisition system, which is in charge of acquiring complementary data in the domain where they are lacking. One possible setup may consists in supporting the ImSPOC acquisition with a traditional camera, e.g. which employs a CFA technology [@3];
- **Machine Learning Approaches**: While inverse problems have a solid mathematical interpretation, its practical implementation is heavily reliant on a good characterization of the direct model and the products to reconstructs. While physical models are able to provide an initial interpretation of many phenomena, some more advanced effect may become cumbersome to characterize rigorously. Recent approaches introduce mixed model and data-driven models [179, 15, 221] to train both the models and prior when acquisitions are limited.
- **Joint Optimization**: In this thesis, the inversion problem and the determination of the direct model were approached separately, with the model inferred from controlled acquisitions; this approach is efficient in terms of computational speed, yet there may be situations in which calibration data are either absent or not sufficient to characterize the behaviour of the optical device for in situ captures. This situation may be approached with protocols of alternate minimization [45, 212], which aims to jointly optimize model and reconstructed products.
- **Super-resolution**: The scenarios considered in this thesis suppose the utilization of snapshot acquisitions, and target the highest resolution available in the raw acquisition. Such limitation can be theoretically lifted in the case multiple replicas are available (such as in the case of the multiple sub-images in ImSPOC), generating a high resolution from a set of low resolution images [182].
- **User-ready Applications**: The scope of this thesis is limited to the reconstruction of hyperspectral datacubes; an additional effort should be spent to identify if the reconstructed products are sufficient for the envisioned applications. I.e. for the case of ImSPOC, those applications are defined by the final applications (e.g., Appendix A.2), which typically is a form of gas detection and estimation of its concentration.

- **Choice of the inversion parameters**: the recent results of the over-relaxation of the algorithms [50], allowed to run the test relatively quickly and the choice of the convergence parameters a relatively painless process. On the other hand, the choice of the regularization parameter is strongly reliant on the intensity of the noise and may require an analysis of the measurement conditions to be set up properly, which would benefit from some more sophisticated procedure of automatic procedure.

Appendix A

A.1 Linear operators

In the context of inverse problems, direct models and regularization domains are often described by a given linear operator $\mathbb{A}(\cdot)$. We describe in this section a series of linear operator that are used throughout this thesis.

The algorithms based on proximal operators, such as the ones described in Section 2.3.2, are able to solve a wide variety of problems framed as Bayesian inferences, but require the knowledge of their adjoint $\mathbb{A}^*(\cdot)$ and of their operator norm $\| \mathbb{A} \|_{op}$.

In this section, we describe a set of operators that are commonly employed in this thesis, by showing that they can be interpreted as a matrix multiplication.

This interpretation allows to derive the expression of its adjoint operator, which is equivalent to the Hermitian of the matrix, while the operator norm is its largest singular value (s.v.).

An implementation of the adjoint operator as a matrix multiplication is often too computationally heavy, as the involved matrices often contain too many coefficients, for which we aim to find equivalent expressions that may be easily implemented.

A.1.1 Operator properties

Let $A : \mathtt{E}_x \rightarrow \mathtt{E}_y$ define a generic bounded linear operator operating between two real Hilbert spaces with scalar products $\langle \cdot, \cdot \rangle_x$ and $\langle \cdot, \cdot \rangle_y$, respectively.

In this thesis, we are mostly interested in two properties of linear operators, the adjoint operator and the operator norm, whose definitions are given in the following subsections.

Adjoint operator

The adjoint operator of $\mathbb{A}$, denoted with $\mathbb{A}^*(\cdot) : \mathtt{E}_y \rightarrow \mathtt{E}_x$ is a generally unbounded operator that satisfies the following condition:

$$\langle \mathbb{A}(\mathbf{x}), \mathbf{u} \rangle_y = \langle \mathbf{u}, \mathbb{A}^*(\mathbf{u}) \rangle_x \qquad \forall \mathbf{x} \in \mathtt{E}_x, \mathbf{u} \in \mathtt{E}_y \, . \tag{A.1.1}$$

For a matrix multiplication, in particular, where the original operator is given by $\mathbf{y} = \mathbf{A}\mathbf{x}$, with $\mathbf{x} \in \mathbb{R}^{N_x}$ and $\mathbf{y} \in \mathbb{R}^{N_y}$ and $\mathbf{A} \in \mathbb{R}^{N_y \times N_x}$ it can be easily shown that the adjoint operator is equivalent to the Hermitian conjugate $\mathbf{A}^*$, or its transpose $\mathbf{A}^\mathsf{T}$ if $\mathbf{A}$ is real.

Operator norm

The operator norm of $\mathbb{A}$ is defined as the largest scalar by which $\mathbb{A}$ stretches the elements of $\mathtt{E}_x$:

$$\| \mathbb{A} \|_{op} = \inf \{ \check{\alpha} \geq 0 : \| \mathbb{A}(\mathbf{x}) \|_y \leq \check{\alpha} \|\mathbf{x}\|_x, \forall \mathbf{x} \in \mathtt{E}_x \} \, , \tag{A.1.2}$$

which in the case of a matrix multiplication studied in the previous section, the operator norm $\|\mathbf{A}\|_{op} = \check{\zeta}_{\mathbf{A}}^{[max]}$ is given by the largest singular value of the matrix $\mathbf{A}$, which is equivalent to the square root of the largest eigenvalue associated with the Gram matrix $\mathbf{A}^*\mathbf{A}$.

A.1.2 Convolution product

Matrix multiplication interpretation

Let $\mathbf{U}^{[x]} = \left\{ u^{[x]}_{i_1,i_2} \right\}_{\substack{i_1 \in [1,\ldots,N_{i_1}] \\ i_2 \in [1,\ldots,N_{i_2}]}}$ be a matrix representing a monochromatic image and $\mathbf{U}^{[b]} \in \mathbb{R}^{N_{d_1} \times N_{d_2}}$ be a bidimensional filter, whose index in the coefficients $u^{[b]}_{i_1,i_2}$ span the quasi-symmetrical ranges $i_1 \in \left[-\left\lfloor \frac{N_{d_1}}{2} \right\rfloor, \ldots, \left\lfloor \frac{N_{d_1}-1}{2} \right\rfloor \right]$ and $i_2 \in \left[-\left\lfloor \frac{N_{d_2}}{2} \right\rfloor, \ldots, \left\lfloor \frac{N_{d_2}-1}{2} \right\rfloor \right]$.

We define in this thesis a circulant convolution product $\mathbf{U}^{[y]} = \mathbf{U}^{[x]} \ **\ \mathbf{U}^{[b]}$ a matrix $\mathbf{U}^{[y]} \in \mathbb{R}^{N_{i_1} \times N_{i_2}}$, whose coefficients are:

$$u^{[y]}_{i_1,i_2} = \sum_{m_1=-\left\lfloor \frac{N_{d_1}}{2} \right\rfloor}^{\left\lfloor \frac{N_{d_1}-1}{2} \right\rfloor} \sum_{m_2=-\left\lfloor \frac{N_{d_2}}{2} \right\rfloor}^{\left\lfloor \frac{N_{d_2}-1}{2} \right\rfloor} u^{[b]}_{m_1,m_2} u^{[x]}_{(i_1-m_1)_{N_{i_1}},(i_2-m_2)_{N_{i_2}}}, \tag{A.1.3}$$

where $(i)_{N_i} := ((i-1) \mod N_i) + 1$ and $\lfloor \cdot \rfloor$ stands for integer part.

If the input and output matrices $\mathbf{U}^{[x]}$ and $\mathbf{U}^{[y]}$ are arranged in lexicographic order, it can be shown that this operation is equivalent to a multiplication by a doubly block circulant matrix $\check{\mathbf{B}}$, which is a particular case of a Toeplitz matrix. A Toeplitz matrix co the elements of each descending diagonal from left to right are the same; in our case the matrix $\check{\mathbf{B}}$ is generally sparse, and the non-zero elements are given by the coefficients of $\mathbf{U}^{[b]}$.

A formal proof of this statement, known as **circulant convolution theorem**, is available in the dedicated literature [196], but we will show an example to illustrate the main concept behind it. Let us consider a monochromatic image $\mathbf{U}^{[x]} \in \mathbb{R}^{3\times3}$ and a convolution filter $\mathbf{U}^{[b]} \in \mathbb{R}^{2\times2}$, then the circulant convolution product is given by:

$$\mathbf{U}^{[x]} \ **\ \mathbf{U}^{[b]} = \begin{bmatrix} x_1 & x_4 & x_7 \\ x_2 & x_5 & x_8 \\ x_3 & x_6 & x_9 \end{bmatrix} \ **\ \begin{bmatrix} b_4 & b_2 \\ b_3 & b_1 \end{bmatrix} \tag{A.1.4a}$$

$$= \begin{bmatrix} \begin{matrix} x_1b_1+x_4b_3 \\ +x_2b_2+x_5b_4 \end{matrix} & \begin{matrix} x_4b_1+x_7b_3 \\ +x_5b_2+x_8b_4 \end{matrix} & \begin{matrix} x_7b_1+x_1b_3 \\ +x_8b_2+x_2b_4 \end{matrix} \\ \begin{matrix} x_2b_1+x_5b_3 \\ +x_3b_2+x_6b_4 \end{matrix} & \begin{matrix} x_5b_1+x_8b_3 \\ +x_6b_2+x_9b_4 \end{matrix} & \begin{matrix} x_8b_1+x_2b_3 \\ +x_9b_2+x_3b_4 \end{matrix} \\ \begin{matrix} x_3b_1+x_6b_3 \\ +x_1b_2+x_4b_4 \end{matrix} & \begin{matrix} x_6b_1+x_9b_3 \\ +x_4b_2+x_7b_4 \end{matrix} & \begin{matrix} x_9b_1+x_3b_3 \\ +x_7b_2+x_1b_4 \end{matrix} \end{bmatrix}. \tag{A.1.4b}$$

which can be also interpreted as a correlation product if $\mathbf{U}^{[b]}$ flipped both horizontally and vertically, which is easier to mentally visualize, imagining the filter as a sliding window over the image.

If the resulting coefficients are rewritten in their lexicographic order $\mathbf{y} = \mathrm{matr}\left(\mathbf{U}^{[y]}\right)$, they can be also obtained multiplying $\mathbf{x} = \mathrm{matr}\left(\mathbf{U}^{[x]}\right)$ by a doubly block circulant matrix $\check{\mathbf{B}}$, as shown below:

$$\mathbf{y} = \check{\mathbf{B}}\mathbf{x} = \begin{bmatrix} b_1 & b_2 & 0 & b_3 & b_4 & 0 & 0 & 0 & 0 \\ 0 & b_1 & b_2 & 0 & b_3 & b_4 & 0 & 0 & 0 \\ 0 & 0 & b_1 & b_2 & 0 & b_3 & b_4 & 0 & 0 \\ 0 & 0 & 0 & b_1 & b_2 & 0 & b_3 & b_4 & 0 \\ 0 & 0 & 0 & 0 & b_1 & b_2 & 0 & b_3 & b_4 \\ b_4 & 0 & 0 & 0 & 0 & b_1 & b_2 & 0 & b_3 \\ b_3 & b_4 & 0 & 0 & 0 & 0 & b_1 & b_2 & 0 \\ 0 & b_3 & b_4 & 0 & 0 & 0 & 0 & b_1 & b_2 \\ b_2 & 0 & b_3 & b_4 & 0 & 0 & 0 & 0 & b_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{bmatrix} \tag{A.1.5a}$$

$$= \begin{bmatrix} x_1 b_1 + x_4 b_3 + x_2 b_2 + x_5 b_4 \\ x_4 b_1 + x_7 b_3 + x_5 b_2 + x_8 b_4 \\ x_7 b_1 + x_1 b_3 + x_8 b_2 + x_2 b_4 \\ x_2 b_1 + x_5 b_3 + x_3 b_2 + x_6 b_4 \\ x_5 b_1 + x_8 b_3 + x_6 b_2 + x_9 b_4 \\ x_8 b_1 + x_2 b_3 + x_9 b_2 + x_3 b_4 \\ x_3 b_1 + x_6 b_3 + x_1 b_2 + x_4 b_4 \\ x_6 b_1 + x_9 b_3 + x_4 b_2 + x_7 b_4 \\ x_9 b_1 + x_3 b_3 + x_7 b_2 + x_1 b_4 \end{bmatrix}. \tag{A.1.5b}$$

Properties

The adjoint operator of a convolution product can be obtained by noticing that the effect of transposing the circulant matrix $\check{\mathbf{B}}$ from Appendix A.1.2 returns another circulant matrix but with flipped coefficients. I.e., the adjoint operator of $\check{\mathbf{B}}$ in eq. (A.1.5) is given by:

$$\check{\mathbf{B}}^{\mathsf{T}} = \begin{bmatrix} b_1 & 0 & 0 & b_4 & b_3 & 0 & 0 & b_2 \\ b_2 & b_1 & 0 & 0 & b_4 & b_3 & 0 & 0 \\ 0 & b_2 & b_1 & 0 & 0 & b_4 & b_3 & 0 \\ 0 & 0 & b_2 & b_1 & 0 & 0 & b_4 & b_3 \\ b_3 & 0 & 0 & b_2 & b_1 & 0 & 0 & b_4 \\ b_4 & b_3 & 0 & 0 & b_2 & b_1 & 0 & 0 \\ 0 & b_4 & b_3 & 0 & 0 & b_2 & b_1 & 0 \\ 0 & 0 & b_4 & b_3 & 0 & 0 & b_2 & b_1 \end{bmatrix}, \tag{A.1.6}$$

from which we can notice (e.g. from the sixth row) that the filter coefficients are flipped with respect to the original arrangement. Consequently, the adjoint of a convolution product is equivalent correlation with the same kernel, that is, a convolution with the same filter, whose coefficients are flipped both horizontally and vertically.

Regarding the operator norm, we can make use here of a well known property of circulant matrices, which allows to evaluate the eigenvalues of a circulant matrix [54]. Specifically, given a square circulant matrix $\check{\mathbf{B}} \in \mathbb{R}^{N_i \times N_i}$ its i-th s.v. $\check{\zeta}^{(i)}_{\check{\mathbf{B}}}$ is given by:

$$\check{\zeta}^{(i)}_{\check{\mathbf{B}}} = \sqrt{\sum_{k=1}^{N_i} \check{b}_{k,1} \exp\left(j\frac{2\pi}{N_i}(N_i - k + 1)i\right)}\,, \qquad \forall i \in [1, \dots, N_i]\,, \tag{A.1.7}$$

where $\exp\left(j\frac{2\pi}{N_i}\right)$ is the N_i-th root of unity, j is the imaginary unit, and $\check{b}_{k,1}$ is the k-th element of the first column of $\check{\mathbf{B}}$. The relation (A.1.7) allows to define an upper bound for the operator norm, as any s.v., including the largest one $\check{\zeta}^{[max]}_{\check{\mathbf{B}}}$, has to satisfy the following relationship:

$$\check{\zeta}^{[max]}_{\check{\mathbf{B}}} \leq \sqrt{\sum_{i=1}^{N_i} |\check{b}_{i,1}|} = \sqrt{\sum_{i=1}^{N_{d_1} N_{d_2}} |b_i|}\,, \tag{A.1.8}$$

where $\{b_i\}_{i \in [1,\dots,N_{d_1} N_{d_2}]}$ are the elements of the convolution matrix $\mathbf{b} = \mathrm{matr}\left(\mathbf{U}^{[b]}\right)$.

A.1.3 Masking

Matrix multiplication interpretation

Let $\mathbf{X} \in \mathbb{R}^{N_i \times N_b}$ and $\mathbf{H} \in \mathbb{R}^{N_i \times N_b}$ represent the original image and the mask, respectively, represented in lexicographic order. A masking operation can be seen as:

$$\mathbf{y} = \sum_{k=1}^{N_b} \mathbf{x}_{:k} \odot \mathbf{h}_{:k}\,, \tag{A.1.9}$$

where $\odot$ denotes an element by element multiplication, while $\mathbf{x}_{:k}$ and $\mathbf{h}_{:k}$ are the k-th column of $\mathbf{X}$ and $\mathbf{H}$, respectively. To find the equivalent matrix multiplication

expression, let $\mathbf{v}^{[x]} = \mathrm{vec}(\mathbf{x})$ be the vector representation of $\mathbf{X}$ (obtained concatenating all of its columns), for which we obtain that eq. (A.1.9) is equivalent to:

$$\mathbf{y} = \left[\mathrm{diag}\left(\mathbf{h}_{:1}\right), ..., \mathrm{diag}\left(\mathbf{h}_{:N_b}\right)\right] \mathbf{v}^{[x]}\,, \tag{A.1.10}$$

where $[\cdot,\cdot]$ stands for row concatenation and $\mathrm{diag}(\mathbf{h}_{:k})$ is a diagonal matrix whose elements are on the main diagonal are given by the k-th column of $\mathbf{H}$.

Properties

Let $\mathbf{y} \in \mathbb{R}^{N_i}$, be a monochromatic acquisition, expressed in lexicographic order. The adjoint operator $\mathbf{X} = \mathbb{A}^*(\mathbf{y})$ of a masking operation (A.1.9) can be derived from its expression as matrix multiplication (A.1.10), transposing the transformation matrix. This yields:

$$\mathrm{vec}\left(\mathbb{A}^*(\mathbf{y})\right) = \left[\mathrm{diag}\left(\mathbf{h}_{:1}\right); ...; \mathrm{diag}\left(\mathbf{h}_{:N_b}\right)\right] \mathbf{y}\,, \tag{A.1.11}$$

where the diagonal matrices are column concatenated instead of row concatenated as it was in eq. (A.1.10). By unwrapping the result, this operation is equivalent to a multiband image $\mathbf{X} \in \mathbb{R}^{N_i \times N_b}$, whose k-th column $\mathbf{x}_{:k}$ is given by:

$$\mathbf{x}_{:k} = \mathbf{y} \odot \mathbf{h}_{:k}\,. \tag{A.1.12}$$

With respect to the operator norm, the matrix that describes the masking is obtained as column concatenation of $\mathbf{A}_{[k]} = \mathrm{diag}(\mathbf{h}_{:k})$, hence the associated Gram matrix is given by $\sum_{k=1}^{N_b} \left(\mathbf{A}_{[k]}^* \mathbf{A}_{[k]}\right)$. Given the diagonal nature of its components, its eigenvalues are given by the sum of the eigenvalues of the Gram matrix associated with $\mathbf{A}_{[k]}$, from which it follows that the operator norm of the masking operation is:

$$\| \mathbb{A} \|_{op} = \max_{i \in [1,\dots,N_i]} \sqrt{\sum_{k=1}^{N_b} h_{ik}^2}\,. \tag{A.1.13}$$

A.1.4 Decimation

Matrix multiplication interpretation

In image processing, the operation of decimation by a factor ρ is equivalent to take one every ρ-th sample both in the horizontal and vertical direction. For this operation it is useful to introduce the so-called **selection matrix**, which is in charge

of selecting a set of columns from a matrix. Let $\mathbf{e}_{(i)} \in \mathbb{R}^{N_{i_1}}$ denote a column vector with all components equal to zero, except a 1 in the i-th position. This matrix allows to select specific columns from a given monochromatic image $\mathbf{U}^{[x]} \in \mathbb{R}^{N_{i_1} \times N_{i_2}}$.

I.e., if we want to select the second and fifth column from a $5 \times N_{i_2}$ image, we can perform the following operation:

$$\left[\mathbf{e}_{(3)}^{\mathsf{T}}; \mathbf{e}_{(5)}^{\mathsf{T}}\right] \mathbf{U}^{[x]} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{U}^{[x]}, \tag{A.1.14}$$

where $[\cdot;\cdot]$ stands for column concatenation and $(\cdot)^{\mathsf{T}}$ is the transposition operator.

The full decimation operation can then be seen as applying a column selection operator both in the horizontal and vertical direction:

$$\mathbf{U}^{[y]} = \mathbf{U}^{[v^{\downarrow}]} \left(\mathbf{U}^{[h^{\downarrow}]} \mathbf{U}^{[x]}\right)^{\mathsf{T}}, \tag{A.1.15}$$

where $\mathbf{U}^{[h^{\downarrow}]} \in \mathbb{R}^{\frac{N_{i_1}}{\rho} \times N_{i_1}}$ is a matrix whose i-th row is $\mathbf{u}_{i:}^{[h^{\downarrow}]} = \mathbf{e}_{(\lfloor N_{i_1}/2 \rfloor + 1 + (i-1)\rho)}^{\mathsf{T}}$, with the shift $\lfloor N_{i_1}/2 \rfloor + 1$ being necessary to select the center pixel in the decimation.

Similarly, the i-th row of $\mathbf{U}^{[v^{\downarrow}]} \in \mathbb{R}^{\frac{N_{i_2}}{\rho} \times N_{i_2}}$ is $\mathbf{u}_{i:}^{[v^{\downarrow}]} = \mathbf{e}_{(\lfloor N_{i_2}/2 \rfloor + 1 + (i-1)\rho)}^{\mathsf{T}}$, where $\mathbf{e}_{(i)}$ is a column vector of length N_{i_2}.

Properties

The adjoint of a decimation can be promptly derived by noticing that the effect of transposing a selection matrix is equivalent to placing the columns back in their original positions. I.e., the adjoint of eq. (A.1.14), applied to an image $\mathbf{U}^{[y]}$ of sizes $2 \times N_{i_2}$ is given by:

$$\left[\mathbf{e}_{(3)}, \mathbf{e}_{(5)}\right] \mathbf{U}^{[x]} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{U}^{[y]} = \left[\mathbf{0}_{[5\times1]}, \mathbf{0}_{[5\times1]}, \mathbf{u}_{:1}^{[y]}, \mathbf{0}_{[5\times1]}, \mathbf{u}_{:2}^{[y]}\right], \tag{A.1.16}$$

where $\mathbf{0}_{[N_{i_1}\times1]} \in \mathbb{R}^{N_{i_1}}$ is a column vector of all zeros. Consequently, the adjoint operator of a decimation by a factor ρ is equivalent to an expansion by the same factor. The operator norm of a decimation operator is simply equal to 1, as the

decimation matrix is simply equivalent to an identity matrix with interposed rows made of all zeros.

A.1.5 Total variation

Matrix multiplication interpretation

For a given monochromatic image $\mathbf{U}^{[x]} = \left\{u^{[x]}_{i_1,i_2,k}\right\}_{\substack{i_1\in\left[1,\dots,N_{i_1}\right]\\ i_2\in\left[1,\dots,N_{i_2}\right]}}$, the digital total variation (TV) defines the concatenation of the gradients $\mathbf{U}^{[v]}$ and $\mathbf{U}^{[h]}$ in the vertical and horizontal direction, respectively. Specifically, their elements are given by:

$$u^{[v]}_{i_1,i_2} = u^{[x]}_{\max(i_1+1,N_{i_1}),i_2} - u^{[x]}_{i_1,i_2}\,, \tag{A.1.17a}$$

$$u^{[h]}_{i_1,i_2} = u^{[x]}_{i_1,\max(i_2+1,N_{i_2})} - u^{[x]}_{i_1,i_2}\,, \tag{A.1.17b}$$

for all $i_1 \in [1,\dots,N_{i_1}]$ and $i_2 \in [1,\dots,N_{i_2}]$. Each of these operations can be seen as a special case of a circular convolution, except that the operator does not wrap around the boundaries of the image. Consequently, each of these operations is equivalent to a multiplication by a circulant matrix with non zero elements equal to $[1,-1]$ and with the elements of its last row set equal to zero. I.e., if $\mathbf{U}^{[x]}$ is a 5×4 px image, then the i_2-th column $\mathbf{u}^{[v]}_{:i_2}$ of vertical gradient $\mathbf{U}^{[v]}$ and the i_1-th row $\mathbf{u}^{[h]}_{i_1:}$ of the horizontal gradient $\mathbf{U}^{[h]}$ can be obtained as:

$$\mathbf{u}^{[v]}_{i_1:} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0\\ 0 & 1 & -1 & 0 & 0\\ 0 & 0 & 1 & -1 & 0\\ 0 & 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{u}^{[x]}_{i_1:} \tag{A.1.18a}$$

$$\left(\mathbf{u}^{[h]}_{:i_2}\right)^{\mathsf{T}} = \begin{bmatrix} 1 & -1 & 0 & 0\\ 0 & 1 & -1 & 0\\ 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 0 \end{bmatrix} \left(\mathbf{u}^{[x]}_{:i_2}\right)^{\mathsf{T}} \tag{A.1.18b}$$

The two results are typically concatenated together along a new dimension; in this thesis, this concatenation is performed along the fourth one, to keep the third one for the channels, so that the full TV operator is $\mathbb{A}\left(\mathbf{U}^{[x]}\right) = \left[\mathbf{U}^{[v]}, \mathbf{U}^{[h]}\right]_4$.

Properties

As the TV is given by a concatenation of a horizontal and vertical gradient, its combined adjoint is given by the sum of the adjoint of each of the two components, similarly to how the masking, which is a sum of operations applied to each bands, can be seen as the adjoint of a concatenation.

That is, given the vertical and horizontal gradient components $\mathbf{U}^{[v]} \in \mathbb{R}^{N_{i_1} \times N_{i_2}}$ and $\mathbf{U}^{[h]} \in \mathbb{R}^{N_{i_1} \times N_{i_2}}$, the adjoint operator of the TV can be rewritten as $\mathbb{A}^* \left(\left[\mathbf{U}^{[v]}, \mathbf{U}^{[h]} \right]_4 \right) = \mathbb{A}_v^* \left(\mathbf{U}^{[v]} \right) + \mathbb{A}_h^* \left(\mathbf{U}^{[h]} \right) = \mathbf{V}^{[v]} + \mathbf{V}^{[h]}$, where $\mathbf{V}^{[v]} = \mathbb{A}_v^* \left(\mathbf{U}^{[v]} \right)$ and $\mathbf{V}^{[h]} = \mathbb{A}_v^* \left(\mathbf{U}^{[h]} \right)$ are the adjoint operators of the vertical and horizontal gradients respectively.

By transposing the matrix multiplications in eq. A.1.18, it is easy to show that their coefficients are given by:

$$v_{i_1,i_2}^{[v]} = u_{i_1,i_2}^{[v]} - u_{\min(i_1-1,1),i_2}^{[v]} \tag{A.1.19a}$$

$$v_{i_1,i_2}^{[h]} = u_{i_1,i_2}^{[h]} - u_{i_1,\min(i_2-1,1)}^{[v]} \tag{A.1.19b}$$

for all $i_1 \in [1, \dots, N_{i_1}]$ and $i_2 \in [1, \dots, N_{i_2}]$. The expression assumes that $\mathbf{u}_{N_{i_1}:}^{[v]}$ and $\mathbf{u}_{:N_{i_2}}^{[h]}$ are vectors made of all zeros, which is the case if those gradients are obtained from eq. (A.1.17).

The operator norm $\| \mathbb{A} \|_{op}$ must satisfy the following relationship:

$$\| \mathbb{A} \|_{op}^2 = \| \mathbb{A}_v^* + \mathbb{A}_h^* \|_{op}^2 \leq \| \mathbb{A}_v \|_{op}^2 + \| \mathbb{A}_h \|_{op}^2 =\leq (|1| + |-1|)^2 + (|1| + |-1|)^2 = 8 \tag{A.1.20}$$

where we have used Cauchy's inequality to upper bound the norm to the sum of the operator norm of each of its components and we substituted eq. (A.1.8) with the assumption that each gradient is approximately described by a multiplication by a circulant matrix.

A.2 ImSPOC research projects

Four different projects are currently associated with the ImSPOC concept:

- **Horizon 2020 (H2020) Space CARBon Observatory (SCARBO)** [@8] (2018-2020), whose aim is to measure the effects of the greenhouse effect in the atmosphere. The target of this project is to allow the finest resolution of spectra within the absorption window of greenhouse gases (GHGs), such as carbon dioxide CO_2 or methane CH_2. The device is mostly used in the visible (VIS), infrared (IR) and short wave infrared (SWIR) bandwidths [33, 96, 97, 76, 98].
- **Fonds Unique Interministériel (FUI) ImaGAZ** (2017-2020) is aimed at the monitoring of industrial sites at risk of environment, health and safety (EHS); the main application is for the detection of gas leakages or anomalies in their composition. This detector works in the IR domain of wavelengths and its ideal final product is the concentration of each gas, through spectral unmixing of the recovered spectrum, eventually supervised by the prior information of the spectral signatures of the known gases involved in the acquisition.
- **FUI ImSPOC-ultraviolet (UV)** (2018-2021)'s main application is envisioned to be for climate and air quality monitoring, or to track natural environments such as for the monitoring of the aurora borealis, and providing better joint spectral and temporal resolution than current commercial available alternatives. Its spectral bands include the UV, VIS and near infrared (NIR) bandwidths (250-950 nm);
- **Agence Nationale de Recherche (ANR) Astrid FUsion MULTIspectral-ImSPOC (FuMultiSPOC)** [@3] (2021-2024), which main aim is the development of tools for the data processing of raw acquisition and the fusion of acquisition taken simultaneously by an ImSPOC device and a traditional camera.

While the model described in Section 5.4.2 is applicable to every prototype, the necessity for different applications leads to different designs and priorities into compensation of different undesired effects. The details of the manufacturing of each prototype is still under constant discussion and fine-tuning; two main different design philosophies arose for the manufacturing of the devices:

- The **NanoCARB**, which is the basic technology developed for SCARBO, conceptualizes configurations able to recover spectrum samples within different absorption windows; for each of such windows, the device has an associated

set of interferometers, which is in charge to detect the interferogram samples associated with the window of interest. For this design, the current proposition is to realize he interferometers in silicon: as the refractive index of silicon is pretty high compared to air ($n \approx 3.18$), no additional coating is necessary to the reflectivity of the surface, which is instead generated naturally by the boundary. Another beneficial effect is related to the angle of acceptance (AoA), as the leading optics may allow a relatively large input solid angle, as, according to Snell's law, as they still map to relatively small internal reflection angles. However, the reflective index of silicon shows a strong dependency of the optical path difference (OPD) with the wavenumber.

- For the other designs, intended for projects such as ImaGAZ and ImSPOC-UV working over a wider range of wavelengths, the interferometers are planned to be manufactured in glass. As the latter has a refraction index which is extremely close to that of the air ($n \approx n_0 = 1$), the medium would be transparent to incident rays. Consequently, the surfaces need a reflective coating to allow the interference between replicas of the rays; at the current point they are is planned to be realized in $\mathrm{TiO_2}$. One disadvantage of such configuration is that the internal reflection angle is very close to the incident angle, hence the former's order of magnitude is similar to that of the AoA, and more cross-talking between adjacent interferometers arises. On the other hand, as glass shows almost no dependency with the wavenumber, the OPD will not show any particular variation with σ, except for non uniform effects of the coating.

Glossary

Q^2n $Q2^n$ index

ADMM alternating direction method of multipliers

AFN average Fourier norm

ALS alternating least squares

ANR Agence Nationale de Recherche

AoA angle of acceptance

ARI adaptative residual interpolation

ATV anisotropic total variation

ATWT à trous wavelet transform

AWGN additive white Gaussian noise

BayesNaive Bayesian with naive regularization

BDSD band-dependent spatial detail

BIN radiometric binning

BT binary tree

BTES binary tree-based edge-Sensing

CASSI compressive coded aperture spectral imaging

CBD context based decision

CCE centroid center estimation

CCSDS Consultative Committee for Space Data Systems

CFA color filter array

CNMF coupled nonnegative matrix factorization

CNN convolutional neural network

COVE panchromatic coverage

CS component substitution

CTV collaborative total variation

CYM cyan yellow magenta

DBT dominant binary tree

DCT discrete cosine transform

DFT discrete Fourier transform

DIAG diagonal

DIRI Dirichlet distribution based

DMD digital micromirror device

DOAS differential optical absorption spectroscopy

DWT discrete wavelet transform

EHS environment, health and safety

EM electro-magnetic

ERGAS relative dimensionless global error in synthesis

ES exhaustive search

EXP interpolated image

FFT fast Fourier transform

FIR finite impulse response

FoV field of view

FP Fabry-Pérot

FPA focal plane array

FPGA field programmable gate array

FSR free spectral range

FT Fourier transform

FTIR Fourier transform infrared spectroscopy

FTS Fourier transform spectrometer

FUI Fonds Unique Interministériel

FuMultiSPOC FUsion MULTIspectral-ImSPOC

FWHM full width at half maximum

GCE Gaussian-fit center estimation

GCV generalized cross validation

GD gradient descent

GE-1 GeoEye-1
GHG greenhouse gas
GLP generalized Laplacian pyramid
GNA Gauss-Newton algorithm
GSA Gram-Schmidt adaptive
GSA Gram-Schmidt
GSD ground sample distance
GT ground truth
H2020 Horizon 2020
HPM high pass modulation
HRI high resolution image
HS hyperspectral
HSI hyperspectral imaging
HySure hyperspectral superresolution
i.i.d. independent and identically distributed
ID intensity difference
IDCT inverse discrete cosine transform
IDW inverse distance weighting
IDWT inverse discrete wavelet transform
ImSPOC image spectrometer on chip
IPAG Institut de Planétologie et d'Astrophysique de Grenoble
IR infrared
ItID iterative intensity difference
ItSD iterative spectral difference
ITV isotropic total variation
JPEG Joint Photographic Experts Group
L0 level 0
L1 level 1
L2 level 2
L3 level 3
LASSO least absolute shrinkage and selection operator
LDCT LASSO-DCT
LDWT LASSO-DWT
LPF low pass filter
LRI low resolution image
LTV LASSO-TV
MAE mean absolute error
MAUE mean absolute unbiased error
MAXDIS maximum distance
MED mean Euclidean distance
ML maximum likelihood
MLE maximum likelihood estimation
MLRI minimized-Laplacian residual interpolation
MM majorization-minimization
MRA multiresolution analysis
MRCA multiresolution color filter array acquisition
MS multispectral
MSE mean square error
MSFA multispectral filter array
MSG multiscale gradients
MTF modulation transfer function
MTF-GLP MTF-matched generalized Laplacian pyramid
NASA National Aeronautics and Space Administration
NIR near infrared
NN nearest neighbour
OMP orthogonal matching pursuit
ONERA Office National d'Etudes et de Recherches Aérospatiales
OPD optical path difference
OPL optical path length
PAN panchromatic
PCA principal component analysis
PCHIP piecewise cubic Hermite interpolating polynomial
pdf probability density function
PERI periodic
PINV pseudo-inversion
PMD penalized matrix decomposition
pmf probability mass function

PRESTO Precursory Research for Embryonic Science and Technology
PRISMA Hyperspectral Precursor of the Application Mission
PROTO-1 prototype ImSPOC-UV/VIS
PROTO-2 prototype ImSPOC-UV-drone
PROTO-3 prototype ImaGAZ-1
PROTO-4 prototype NanoCarb-1
PSD power spectral density
PSF point spread function
PSNR peak signal to noise ratio
QB QuickBird
QMF quadrature mirror filter
r.v. random variable
RAND random pick
RBF radial basis function
RGB red green blue
RI residual interpolation
RIP restricted isometry property
RMSE root mean square error
ROF Rudin-Osher-Fatemi
RoI region of interest
RR ridge regression
s.v. singular value
SAM spectral angle mapper
SAR synthetic aperture radar
SCARBO Space CARBon Observatory
sCC spatial cross-covariance coefficient
SCE scanline center estimation
SD spectral difference
SNR signal to noise ratio
SSIM structural similarity
STD standard deviation
STV Shannon total variation
SURE Stein's unbiased risk estimate
SVD singular value decomposition
SWIR short wave infrared
SWT stationary wavelet transform
TE transverse electric
TEM transverse electro-magnetic
TGV total generalized variation
TIN triangulated irregular network
TM transverse magnetic
TPS thin plate spline
TSVD truncated singular value decomposition
TV total variation
UAV unmanned aerial vehicle
UBT uniform binary tree
UIQI universal image quality index
UTV upwind total variation
UV ultraviolet
VERT vertical
VIS visible
VTV vector total variation
WB weighted bilinear
WV2 WorldView-2
WV3 WorldView-3

Symbols

λ	wavelength
σ	wavenumber in the vacuum
$\mathbf{k}$	wavevector
ν	optical frequency
θ	polar angle
ϕ	azimuth angle
$\boldsymbol{\theta}$	spherical coordinate angle
$\mathcal{R}$	reflectivity
$\mathcal{T}$	transmissivity
n	refractive index
n_0	refractive index in the air/-vacuum
ε	absolute permittivity
μ	absolute permeability
c	phase velocity
c_0	speed of light in the vacuum
$\mathcal{L}$	radiance
$\mathcal{L}_\sigma$	spectral radiance
$\mathcal{E}$	irradiance
$\mathcal{E}_\sigma$	spectral irradiance
$\mathbb{A}$	direct model (linear operator)
$\mathbb{L}$	regularizer (linear operator)
$\mathbb{L}_{tv}$	total variation operator
$\mathbb{L}_{utv}$	upwind total variation operator
J	objective function
f	data fidelity metric
g	regularization metric
$\check{\lambda}$	regularization parameter
$\check{\zeta}$	singular value
$\check{\rho}$	over-relaxation parameter
$\check{\tau}$	main convergence parameter
$\check{\sigma}$	secondary convergence parameter
$\mathbf{P}$, $\mathbf{p}$, $\mathbf{U}^{[p]}$	high resolution image
$\mathbf{M}$, $\mathbf{U}^{[m]}$	low resolution image
$\widetilde{\mathbf{M}}^{\uparrow}$	upsampled LRI
$\mathbf{H}$, $\mathbf{U}^{[h]}$	mask
$\mathbf{B}$, $\mathbf{U}^{[b]}$	blurring kernel
ρ	scale ratio

ρ_c	compression ratio
$\mathbb{A}_b$	HRI blur operator
$\mathbb{A}_c$	compression operator
$\mathbb{A}_d$	degradation operator
$\mathbb{A}_m$	spatial degradation operator
$\mathbb{A}_{m^\downarrow}$	downsampling operator
$\mathbb{A}_p$	spectral degradation operator
$\mathbb{E}_x$	space of inputs
$\mathbb{E}_y$	space of acquisitions
$\mathbb{E}_b$	space of parameters
N_a	number of acquisitions
N_b	number of bands
N_{b_p}	number of HRI bands
N_i	number of incidence angle
N_k	number of interferometers/-subimages
N_m	number of modes or order
N_o	number of photodetectors
N_x	number of input samples
N_y	number of output samples
$\boldsymbol{\mathcal{X}}$, $\mathbf{X}$, $\mathbf{U}^{[x]}$	ideal input
$\boldsymbol{\mathcal{Y}}$, $\mathbf{Y}$, $\mathbf{U}^{[y]}$	acquisition
$\boldsymbol{\mathcal{A}}$, $\mathbf{A}$	direct model
$\widehat{\boldsymbol{\mathcal{X}}}$, $\widehat{\mathbf{X}}$	reconstructed input
$\boldsymbol{\mathcal{B}}$, β	parameters
$\mathbf{R}^{[0]}$	subimage centers
$\mathbf{R}^{[t]}$, $\check{\mathbf{R}}^{[t]}$	photodetector centers
$\check{\mathbf{R}}^{[f]}$	feature positions
$\check{\mathbf{R}}^{[d]}$	misalignments
vec	column concatenation
matr	lexicographic order reshaping
$\mathrm{reshape}$	permuted lexicographic order reshaping
stack	datacube stacking
$*$, $**$	convolution product
$\odot$	Hadamard product
$\max$	maximum operator
$\min$	minimum operator
$(\cdot)^*$	adjoint operator
$(\cdot)^\intercal$	transpose operator
$(\cdot)^\uparrow$	extension
$(\cdot)^\downarrow$	decimation
$(\cdot)^\dagger$	Moore-Penrose pseudo-inverse
$(\cdot)^\boxdot$	sparse channel
$(\cdot)^\star$	Fenchel conjugate
$\|\cdot\|_1$	ℓ_1 norm
$\|\cdot\|_2$	ℓ_2 norm
$\|\cdot\|_\infty$	ℓ_∞ norm
$\langle\cdot,\cdot\rangle$	scalar product

Bibliography

[1] Rémy Abergel and Lionel Moisan. "The Shannon total variation". In: *Journal of Mathematical Imaging and Vision* 59.2 (May 2017), pp. 341–370 (cit. on pp. 37, 102, 104, 105, 143).

[2] Elhadi Adam, Onisimo Mutanga, and Denis Rugege. "Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: A review". In: *Wetlands Ecology and Management* 18 (June 2010), pp. 281–296 (cit. on pp. 8, 181).

[3] N. Ahmed, T. Natarajan, and K. R. Rao. "Discrete cosine transform". In: *IEEE Transactions on Computers* C-23.1 (Jan. 1974), pp. 90–93 (cit. on p. 272).

[4] B. Aiazzi, L. Alparone, S. Baronti, R. Carla, A. Garzelli, L. Santurri, and M. Selva. "Effects of multitemporal scene changes on pansharpening fusion". In: *2011 6th International Workshop on the Analysis of Multi-temporal Remote Sensing Images (Multi-Temp)*. IEEE, July 2011, pp. 73–76 (cit. on p. 49).

[5] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva. "MTF-tailored multiscale fusion of high-resolution MS and pan imagery". In: *Photogrammetric Engineering & Remote Sensing* 72.5 (May 2006), pp. 591–596 (cit. on pp. 59, 60).

[6] B. Aiazzi, S. Baronti, and M. Selva. "Improving component substitution pansharpening through multivariate regression of MS + Pan data". In: *IEEE Transactions on Geoscience and Remote Sensing* 45.10 (Oct. 2007), pp. 3230–3239 (cit. on pp. 65, 118).

[7] Bruno Aiazzi, Luciano Alparone, Stefano Baronti, Andrea Garzelli, and Massimo Selva. "Advantages of Laplacian pyramids over "à trous" wavelet transforms for pansharpening of multispectral images". In: *Image and Signal Processing for Remote Sensing XVIII*. Ed. by Lorenzo Bruzzone. SPIE, Nov. 2012 (cit. on pp. 66, 118).

[8] Bruno Aiazzi, Stefano Baronti, Massimo Selva, and Luciano Alparone. "Bi-cubic interpolation for shift-free pan-sharpening". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 86.6 (Dec. 2013), pp. 65–76 (cit. on p. 60).

[9] Ali Naci Akansu, Richard A. Haddad, and Hakan Caglar. "Perfect reconstruction binomial QMF-wavelet transform". In: *Visual Communications and Image Processing'90: Fifth in a Series*. Vol. 1360. International Society for Optics and Photonics. 1990, pp. 609–618 (cit. on p. 33).

[10] Naveed Akhtar, Faisal Shafait, and Ajmal Mian. "Sparse spatio-spectral representation for hyperspectral image super-resolution". In: *Computer Vision – ECCV 2014*. Springer International Publishing, 2014, pp. 63–78 (cit. on p. 68).

[11] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini. “A global quality measurement of pan-sharpened multispectral imagery”. In: *IEEE Geoscience and Remote Sensing Letters* 1.4 (Oct. 2004), pp. 313–317 (cit. on p. 86).

[12] Prakhar Amba, Jérôme Dias, and David Alleysson. “Random color filter arrays are better than regular ones”. In: *Journal of Imaging Science and Technology* 60.5 (Sept. 2016), pp. 504061–504066 (cit. on pp. 74, 96).

[13] Prakhar Amba, Jean Baptiste Thomas, and David Alleysson. “N-LMMSE demosaicing for spectral filter arrays”. In: *Color and Imaging Conference* 2017.25 (Sept. 2017), pp. 130–140 (cit. on p. 74).

[14] Gonzalo R. Arce, David J. Brady, Lawrence Carin, Henry Arguello, and David S. Kittle. “Compressive coded aperture spectral imaging: An introduction”. In: *IEEE Signal Processing Magazine* 31.1 (Jan. 2014), pp. 105–115 (cit. on pp. 10, 76, 82, 103, 111).

[15] Simon Arridge, Peter Maass, Ozan Öktem, and Carola-Bibiane Schönlieb. “Solving inverse problems using data-driven models”. In: *Acta Numerica* 28 (May 2019), pp. 1–174 (cit. on pp. 22, 290, 296).

[16] Josselin Aval, Xavier Briottet, Sophie Fabre, Pierre-Yves Foucher, Véronique Carrère, Rodolphe Marion, Kuniaki Uto, Christiane Weber, and Mauro Dalla Mura. “Applications in remote sensing - anthropogenic activities”. In: *Hyperspectral Imaging*. Vol. 32. Elsevier, 2020, pp. 411–452 (cit. on p. 8).

[17] Touria Bajjouk, Florian de Boissieu, Jocelyn Chanussot, et al. “Applications in remote sensing - natural landscapes”. In: *Hyperspectral Imaging*. Vol. 32. Elsevier, 2020, pp. 371–410 (cit. on p. 8).

[18] Constantine A. Balanis. *Antenna theory: analysis and design*. John wiley & sons, 2015 (cit. on p. 154).

[19] Stefano Baronti, Bruno Aiazzi, Massimo Selva, Andrea Garzelli, and Luciano Alparone. “A theoretical analysis of the effects of aliasing and misregistration on pansharpened imagery”. In: *IEEE Journal of Selected Topics in Signal Processing* 5.3 (June 2011), pp. 446–453 (cit. on p. 66).

[20] Heinz H. Bauschke and Patrick L. Combettes. “Fenchel–Rockafellar duality”. In: *CMS Books in Mathematics*. Springer International Publishing, 2017, pp. 247–262 (cit. on p. 275).

[21] Bryce E. Bayer. *Color imaging array*. US Patent 3,971,065. July 1976 (cit. on pp. 52, 73, 75).

[22] Mokhtar S. Bazaraa, Hanif D. Sherali, and Chitharanjan M. Shetty. *Nonlinear programming: theory and algorithms*. John Wiley & Sons, 2013 (cit. on pp. 47, 295).

[23] E. Ben-Dor, S. Chabrillat, J. A. M. Demattê, G. R. Taylor, J. Hill, M. L. Whiting, and S. Sommer. “Using imaging spectroscopy to study soil properties”. In: *Remote Sensing of Environment* 113 (Sept. 2009), S38–S55 (cit. on pp. 8, 181).

[24] James Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. “Algorithms for hyper-parameter optimization”. In: *Advances in neural information processing systems* 24 (2011) (cit. on p. 251).

[25]José M. Bioucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gader, and Jocelyn Chanussot. “Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5.2 (Apr. 2012), pp. 354–379 (cit. on p. 148).

[26]Patrizia Boccacci and Mario Bertero. *Introduction to inverse problems in imaging*. Taylor & Francis, Jan. 1998 (cit. on p. 21).

[27]Giuseppe Bonifazi, Giuseppe Capobianco, and Silvia Serranti. “Asbestos containing materials detection and classification by the use of hyperspectral imaging”. In: *Journal of hazardous materials* 344 (2018), pp. 981–993 (cit. on p. 8).

[28]Clémentine Bouyé, Thierry Robin, and Benoît d’Humiéres. *Spectral imaging end-users needs, market and trends*. Tech. rep. Tematys, Exploration of photonics markets, Jan. 2018 (cit. on pp. 6–8, 181).

[29]Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, Mar. 8, 2004. 727 pp. (cit. on pp. 21, 39).

[30]Johannes Brauers and Til Aach. “A color filter array based multispectral camera”. In: *12. Workshop Farbbildverarbeitung*. Ilmenau. 2006 (cit. on p. 77).

[31]Kristian Bredies, Karl Kunisch, and Thomas Pock. “Total generalized variation”. In: *SIAM Journal on Imaging Sciences* 3.3 (2010), pp. 492–526 (cit. on pp. 35, 37, 148, 292, 295).

[32]Xavier Bresson and Tony F. Chan. “Fast dual minimization of the vectorial total variation norm and applications to color image processing”. In: *Inverse Problems and Imaging* 2.4 (2008), pp. 455–484 (cit. on p. 38).

[33]Laure Brooker Lizon-Tati, Heinrich Bovensmann, Cyril Crevoisier, Etienne Le Coarer, Andrzei Klonecki, Udrivolf Pica, and Aaldert Van Amerongen. “A constellation of small satellites for the monitoring of greenhouse gases”. In: *International Astronautical Congress (IAC) 2018*. 2018 (cit. on p. 308).

[34]Antoni Buades, Bartomeu Coll, Jean-Michel Morel, and Catalina Sbert. “Self-similarity driven demosaicking”. In: *Image Processing On Line* 1 (June 2011), pp. 51–56 (cit. on p. 82).

[35]Martin D. Buhmann. *Radial basis functions: theory and implementations*. Vol. 12. Cambridge university press, 2003 (cit. on p. 61).

[36]E. J. Candes, J. Romberg, and T. Tao. “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information”. In: *IEEE Transactions on Information Theory* 52.2 (Feb. 2006), pp. 489–509 (cit. on p. 94).

[37]Oscar Carrasco, Richard B. Gomez, Arun Chainani, and William E. Roper. “Hyperspectral imaging applied to medical diagnoses and food safety”. In: *Geo-Spatial and Temporal Image and Data Exploitation III*. Ed. by Nickolas L. Faust and William E. Roper. SPIE, Aug. 2003, 5097:215–222 (cit. on p. 8).

[38]Ayan Chakrabarti and Todd Zickler. “Statistics of real-world hyperspectral images”. In: *CVPR 2011*. IEEE, June 2011 (cit. on p. 54).

[39] Antonin Chambolle, Stacey E. Levine, and Bradley J. Lucier. "An upwind finite-difference method for total variation–based image smoothing". In: *SIAM Journal on Imaging Sciences* 4.1 (Jan. 2011), pp. 277–299 (cit. on pp. 37, 102, 104, 143).

[40] Antonin Chambolle and Thomas Pock. "A first-order primal-dual algorithm for convex problems with applications to imaging". In: *Journal of Mathematical Imaging and Vision* 40.1 (Dec. 2010), pp. 120–145 (cit. on p. 39).

[41] Antonin Chambolle and Thomas Pock. "An introduction to continuous optimization for imaging". In: *Acta Numerica* 25 (May 2016), pp. 161–319 (cit. on pp. 102, 148, 292, 295).

[42] Chen I. Chang. *Hyperspectral data exploitation: Theory and applications*. Ed. by John Wiley & Sons. 2007 (cit. on pp. 1, 7).

[43] Shun Lien Chuang. *Physics of photonic devices*. Vol. 80. John Wiley & Sons, 2012 (cit. on p. 160).

[44] Vittorio Colao and Giuseppe Marino. "Krasnoselskii-Mann method for non-self mappings". In: *Fixed Point Theory and Applications* 2015.1 (Mar. 2015) (cit. on p. 41).

[45] P. Comon, X. Luciani, and A. L. F. de Almeida. "Tensor decompositions, alternating least squares and other tales". In: *Journal of Chemometrics* 23.7-8 (July 2009), pp. 393–405 (cit. on pp. 290, 296).

[46] Laurent Condat. "A new random color filter array with good spectral properties". In: *2009 16th IEEE International Conference on Image Processing (ICIP)*. IEEE, Nov. 2009 (cit. on pp. 71, 292).

[47] Laurent Condat. "Color filter array design using random patterns with blue noise chromatic spectra." In: *Image and Vision Computing* 28.8 (2010), pp. 1196–1202 (cit. on p. 75).

[48] Laurent Condat. "Discrete total variation: New definition and minimization". In: *SIAM Journal on Imaging Sciences* 10.3 (Jan. 2017), pp. 1258–1290 (cit. on pp. 35, 37, 148, 290, 295).

[49] Laurent Condat. "Fast projection onto the simplex and the l1 ball". In: *Mathematical Programming* 158.1-2 (Sept. 2015), pp. 575–585 (cit. on p. 42).

[50] Laurent Condat, Daichi Kitahara, Andrés Contreras, and Akira Hirabayashi. "Proximal splitting algorithms: A tour of recent advances, with new twists". In: (2019). arXiv: `1912.00137 [math.OC]` (cit. on pp. 41, 45, 292, 297).

[51] M. G. Cox. *Reliable numerical computation*. Oxford New York: Clarendon Press Oxford University Press, 1990 (cit. on p. 24).

[52] M. Dalla Mura, S. Prasad, F. Pacifici, P. Gamba, J. Chanussot, and J. A. Benediktsson. "Challenges and opportunities of multimodality and data fusion in remote sensing". In: *Proceedings of the IEEE* 103.9 (Sept. 2015), pp. 1585–1601 (cit. on pp. 1, 49).

[53] Mark A. Davenport. "Random observations on random observations: Sparse signal acquisition and processing". PhD thesis. 2010 (cit. on p. 94).

[54]Philip Davis. *Circulant matrices*. New York: Wiley, 1979 (cit. on p. 303).

[55]Charles-Alban Deledalle, Samuel Vaiter, Gabriel Peyre, Jalal Fadili, and Charles Dossal. "Unbiased risk estimation for sparse analysis regularization". In: *2012 19th IEEE International Conference on Image Processing*. IEEE, Sept. 2012 (cit. on p. 30).

[56]Ronald A. DeVore. "Deterministic constructions of compressed sensing matrices". In: *Journal of complexity* 23.4-6 (2007), pp. 918–925 (cit. on p. 95).

[57]Luc Devroye. *Nonuniform random variate generation*. Ed. by Springer-Vlg. Springer Science+Business Media, 1986 (cit. on p. 96).

[58]Di Di, Min Min, Jun Li, and Mathew M. Gunshor. "The radiance differences between wavelength and wavenumber spaces in convolving hyperspectral infrared sounder spectrum to broadband for intercomparison". In: *Remote Sensing* 11.10 (2019), p. 1177 (cit. on pp. 159, 279).

[59]Aneline Dolet, Daniele Picone, Mauro Dalla Mura, Didier Voisin, Silvère Gousset, Sylvain Douté, and Etienne Le Coarer. "Gas characterization based on a snapshot interferometric imaging spectrometer". In: *Image and Signal Processing for Remote Sensing XXV*. Ed. by Lorenzo Bruzzone, Francesca Bovolo, and Jon Atli Benediktsson. SPIE, Oct. 2019 (cit. on pp. 18, 248, 259, 293).

[60]Aneline Dolet, Daniele Picone, Silvère Gousset, Mauro Dalla Mura, Etienne Le Coarer, and Didier Voisin. "A new snapshot interferometric imaging spectrometer: a first comparison with a classical grating spectrometer." In: *European Geosciences Union (EGU) General Assembly Conference Abstracts*. 2020, p. 8353 (cit. on p. 19).

[61]Aneline Dolet, Daniele Picone, Silvère Gousset, Mauro Dalla Mura, Etienne Le Coarer, and Didier Voisin. "Using zenith observations for evaluation of an improved interferometric imaging spectrometer". In: *European Geosciences Union (EGU) General Assembly Conference Abstracts*. 2021, EGU21–2536 (cit. on p. 19).

[62]D. L. Donoho. "Compressed sensing". In: *IEEE Transactions on Information Theory* 52.4 (Apr. 2006), pp. 1289–1306 (cit. on pp. 32, 93).

[63]David L. Donoho and Michael Elad. "Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization". In: *Proceedings of the National Academy of Sciences* 100.5 (2003), pp. 2197–2202 (cit. on p. 32).

[64]J. Duran, A. Buades, B. Coll, C. Sbert, and G. Blanchet. "A survey of pansharpening methods with a new band-decoupled variational model". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 125 (Mar. 2017), pp. 78–105 (cit. on pp. 124, 148).

[65]Joan Duran, Michael Moeller, Catalina Sbert, and Daniel Cremers. "Collaborative total variation: A general framework for vectorial TV models". In: *SIAM Journal on Imaging Sciences* 9.1 (2016), pp. 116–151 (cit. on pp. 35, 38, 102, 143, 290, 292, 293, 295).

[66] Joan Duran, Michael Moeller, Catalina Sbert, and Daniel Cremers. “On the implementation of collaborative TV regularization: Application to cartoon+ texture decomposition”. In: *Image Processing On Line* 6 (2016), pp. 27–74 (cit. on pp. 35, 38, 102, 104, 105, 143, 292).

[67] Thibaud Ehret and Gabriele Facciolo. “A study of two CNN demosaicking algorithms”. In: *Image Processing On Line* 9 (Sept. 2019), pp. 220–230 (cit. on p. 82).

[68] Jason M. Eichenholz and John Dougherty. “Ultracompact fully integrated megapixel multispectral imager”. In: *Integrated Optics: Devices, Materials, and Technologies XIII*. Ed. by Jean-Emmanuel Broquin and Christoph M. Greiner. SPIE, Feb. 2009 (cit. on p. 92).

[69] Yonina C. Eldar and Gitta Kutyniok. *Compressed sensing: Theory and applications*. Cambridge University Press, Feb. 16, 2016. 558 pp. (cit. on pp. 10, 93, 94).

[70] Óscar Espitia, Sergio Castillo, and Henry Arguello. “Compressive hyperspectral and multispectral imaging fusion”. In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXII*. Ed. by Miguel Velez-Reyes and David W. Messinger. SPIE, May 2016 (cit. on pp. 92, 106, 111).

[71] von F. Zernike. “Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode”. In: *Physica* 1.7-12 (May 1934), pp. 689–704 (cit. on p. 229).

[72] M. D. Farrell and R. M. Mersereau. “On the impact of PCA dimension reduction for hyperspectral detection of difficult targets”. In: *IEEE Geoscience and Remote Sensing Letters* 2.2 (Apr. 2005), pp. 192–195 (cit. on p. 68).

[73] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. “Fast and robust multiframe super resolution”. In: *IEEE Transactions on Image Processing* 13.10 (Oct. 2004), pp. 1327–1344 (cit. on p. 190).

[74] Yao-Ze Feng and Da-Wen Sun. “Application of hyperspectral imaging in food safety inspection and control: A review”. In: *Critical Reviews in Food Science and Nutrition* 52.11 (Nov. 2012), pp. 1039–1058 (cit. on p. 8).

[75] Donald Fenna. *Cartographic science: A compendium of map projections, with derivations*. CRC Press, 2006 (cit. on p. 63).

[76] Yann Ferrec, Guillaume Bonnery, Laure Brooker, Laurence Croizé, Silvère Gousset, and Etienne Le Coarer. “NanoCarb part 1: Compact snapshot imaging interferometer for CO2 monitoring from space”. In: *International Conference on Space Optics — ICSO 2018*. CHANIA, Greece, Oct. 2018 (cit. on pp. 15, 181, 308).

[77] Christodoulos A. Floudas. *Nonlinear and mixed-integer optimization: Fundamentals and applications*. Oxford University Press, 1995 (cit. on p. 47).

[78] Christodoulos A. Floudas and Panos M. Pardalos. *Encyclopedia of optimization*. Springer Science & Business Media, 2008 (cit. on pp. 46, 226).

[79] Marion Foare, Nelly Pustelnik, and Laurent Condat. “Semi-linearized proximal alternating minimization for a discrete Mumford–Shah model”. In: *IEEE Transactions on Image Processing* 29 (2020), pp. 2176–2189 (cit. on p. 148).

[80]Simon Foucart and Holger Rauhut. *A mathematical introduction to compressive sensing (Applied and numerical harmonic analysis)*. Birkhäuser, 2013 (cit. on p. 89).

[81]Giorgio Franceschetti and Riccardo Lanari. *Synthetic aperture radar processing*. CRC press, 2018 (cit. on pp. 9, 186).

[82]Jerome Friedman, Trevor Hastie, and Rob Tibshirani. "Regularization paths for generalized linear models via coordinate descent". In: *Journal of statistical software* 33.1 (2010), p. 1 (cit. on p. 104).

[83]Ying Fu, Yinqiang Zheng, Hua Huang, Imari Sato, and Yoichi Sato. "Hyperspectral image super-resolution with a mosaic RGB image". In: *IEEE Transactions on Image Processing* 27.11 (Nov. 2018), pp. 5539–5552 (cit. on p. 107).

[84]Claudio Galeazzi, Andrea Sacchetti, Andrea Cisbani, and Gianni Babini. "The PRISMA program". In: *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2008 (cit. on p. 51).

[85]A. Garzelli and F. Nencini. "Hypercomplex quality assessment of multi/hyperspectral images". In: *IEEE Geoscience and Remote Sensing Letters* 6.4 (Oct. 2009), pp. 662–665 (cit. on p. 86).

[86]Andrea Garzelli, Filippo Nencini, and Luca Capobianco. "Optimal MMSE pan sharpening of very high resolution multispectral images". In: *IEEE Transactions on Geoscience and Remote Sensing* 46.1 (Jan. 2008), pp. 228–236 (cit. on pp. 66, 118).

[87]Todor G. Georgiev. *Plenoptic camera*. US Patent 7,620,309. Nov. 2009 (cit. on pp. 190, 229).

[88]Anthony Gerrard and James M. Burch. *Introduction to matrix methods in optics*. Courier Corporation, 1994 (cit. on p. 166).

[89]Frédéric Gillard. "Design and realization of a compact infrared spectrometer". Theses. Université Paris Sud - Paris XI, Mar. 2012 (cit. on p. 199).

[90]Didier A. Girard. "A fast 'Monte-Carlo cross-validation' procedure for large least squares problems with noisy data." In: *Numerische Mathematik* 56.1 (1989), pp. 1–24 (cit. on p. 142).

[91]Andrew S. Glassner. *An introduction to ray tracing*. Morgan Kaufmann, 1989 (cit. on p. 183).

[92]Gene H. Golub, Per Christian Hansen, and Dianne P. O'Leary. "Tikhonov regularization and total least squares". In: *SIAM Journal on Matrix Analysis and Applications* 21.1 (Jan. 1999), pp. 185–194 (cit. on p. 274).

[93]Gene H. Golub, Michael Heath, and Grace Wahba. "Generalized cross-validation as a method for choosing a good ridge parameter". In: *Technometrics* 21.2 (1979), pp. 215–223 (cit. on pp. 30, 274).

[94]M. González-Audícana, X. Otazu, O. Fors, and A. Seco. "Comparison between Mallat's and the 'à trous' discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images". In: *International Journal of Remote Sensing* 26.3 (Feb. 2005), pp. 595–614 (cit. on p. 59).

[95] Arthur Ardeshir Goshtasby. *2-D and 3-D image registration: for medical, remote sensing, and industrial applications*. John Wiley & Sons, 2005 (cit. on pp. 62, 229, 230).

[96] S. Gousset, N. Guérineau, L. Croizé, E. Le Coarer, T. Laveille, and Y. Ferrec. "NANOCARB-21: a miniature Fourier-transform spectro-imaging concept for a daily monitoring of greenhouse gas concentration on the Earth surface". In: *International Conference on Space Optics — ICSO 2016*. Ed. by Nikos Karafolas, Bruno Cugny, and Zoran Sodnik. SPIE, Sept. 2017 (cit. on pp. 181, 308).

[97] Silvère Gousset, Laurence Croizé, Etienne Le Coarer, Yann Ferrec, Laure Brooker, and SCARBO consortium. "NanoCarb part 2: Performance assessment for total column CO2 monitoring from a nano-satellite". In: *International Conference on Space Optics — ICSO 2018*. CHANIA, Greece, Oct. 2018 (cit. on pp. 181, 182, 213, 308).

[98] Silvère Gousset, Laurence Croizé, Etienne Le Coarer, Yann Ferrec, Juana Rodrigo-Rodrigo, Laure Brooker, et al. "NanoCarb hyperspectral sensor: On performance optimization and analysis for greenhouse gas monitoring from a constellation of small satellites". In: *CEAS Space Journal* 11.4 (2019), pp. 507–524 (cit. on pp. 181, 308).

[99] Hans Grahn and Paul Geladi. *Techniques and applications of hyperspectral image analysis*. John Wiley & Sons, 2007 (cit. on p. 1).

[100] Rémi Gribonval, Gilles Blanchard, Nicolas Keriven, and Yann Traonmilin. "Compressive statistical learning with random feature moments". In: *arXiv preprint arXiv:1706.07180* (June 22, 2017). arXiv: `1706.07180v2 [stat.ML]` (cit. on pp. 148, 295).

[101] K. Gröchenig and T. Strohmer. "Numerical and theoretical aspects of nonuniform sampling of band-limited images". In: *Nonuniform Sampling*. Springer US, 2001, pp. 283–324 (cit. on p. 62).

[102] C. W. Groetsch. "Comments on Morozov's discrepancy principle". In: *Improperly posed problems and their numerical treatment*. Springer, 1983, pp. 97–104 (cit. on p. 30).

[103] France Grum and Richard J. Becherer. "Optical radiation measurements. Volume 1 - Radiometry". In: *nyap* (1979) (cit. on p. 155).

[104] Nicolas Guerineau, Etienne Le Coarer, Yann Ferrec, and Florence De La Barriere. *Spectro-imageur multivoie à transformée de Fourier*. French, English. FR patent 1,656,162. Jan. 2018 (cit. on pp. 2, 13, 181, 182).

[105] Manuel Guizar-Sicairos, Samuel T. Thurman, and James R. Fienup. "Efficient subpixel image registration algorithms". In: *Optics letters* 33.2 (Jan. 2008), pp. 156–158 (cit. on pp. 109, 110).

[106] Jacques Hadamard. "Sur les problèmes aux dérivées partielles et leur signification physique". In: *Princeton university bulletin* (1902), pp. 49–52 (cit. on pp. 12, 23, 39, 209).

[107] Nathan Hagen. "Snapshot advantage: A review of the light collection improvement for parallel high-dimensional measurement systems". In: *Optical Engineering* 51.11 (June 2012), p. 111702 (cit. on p. 7).

[108]Nathan Hagen and Michael W. Kudenov. “Review of snapshot spectral imaging technologies”. In: *Optical Engineering* 52.9 (Sept. 2013), p. 090901 (cit. on p. 7).

[109]J. F. Jr Hamilton and J. E. Jr Adams. *Adaptive color plan interpolation in single sensor color electronic camera*. US Patent 5629734. 1976 (cit. on p. 78).

[110]Martin Hanke. “Limitations of the L-curve method in ill-posed problems”. In: *BIT Numerical Mathematics* 36.2 (June 1996), pp. 287–301 (cit. on p. 30).

[111]Per Christian Hansen. “Analysis of discrete ill-posed problems by means of the L-curve”. In: *SIAM Review* 34.4 (Dec. 1992), pp. 561–580 (cit. on p. 142).

[112]Per Christian Hansen. *Discrete inverse problems: Insight and algorithms*. Society for Industrial and Applied Mathematics, Jan. 2010 (cit. on p. 21).

[113]Per Christian Hansen. “The L-curve and its use in the numerical treatment of inverse problems”. In: (1999) (cit. on pp. 30, 142, 274, 280).

[114]Per Christian Hansen. “Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank”. In: *SIAM Journal on Scientific and Statistical Computing* 11.3 (May 1990), pp. 503–518 (cit. on p. 274).

[115]Robert M. Haralick, Stanley R. Sternberg, and Xinhua Zhuang. “Image analysis using mathematical morphology”. In: *IEEE transactions on pattern analysis and machine intelligence* 4 (1987), pp. 532–550 (cit. on p. 220).

[116]R. C. Hardie, M. T. Eismann, and G. L. Wilson. “MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor”. In: *IEEE Transactions on Image Processing* 13.9 (Sept. 2004), pp. 1174–1184 (cit. on p. 67).

[117]Parameswaran Hariharan. *Basics of interferometry*. Elsevier, 2010 (cit. on pp. 173, 209).

[118]Xiyan He, Laurent Condat, Jose M. Bioucas-Dias, Jocelyn Chanussot, and Junshi Xia. “A new pansharpening method based on spatial and spectral sparsity priors”. In: *IEEE Transactions on Image Processing* 23.9 (Sept. 2014), pp. 4160–4174 (cit. on p. 80).

[119]M. R. Hestenes and E. Stiefel. “Methods of conjugate gradients for solving linear systems”. In: *Journal of Research of the National Bureau of Standards* 49.6 (Dec. 1952), p. 409 (cit. on p. 40).

[120]Norman Hodgson and Horst Weber. *Optical resonators: Fundamentals, advanced concepts, applications*. Vol. 108. Springer Science & Business Media, 2005 (cit. on p. 166).

[121]Matthias Holschneider, Richard Kronland-Martinet, Jean Morlet, and Ph Tchamitchian. “A real-time algorithm for signal analysis with the help of the wavelet transform”. In: *Wavelets*. Springer, 1990, pp. 286–297 (cit. on p. 33).

[122]Bormin Huang, ed. *Satellite data compression*. Springer New York, 2011 (cit. on p. 87).

[123]J. Idier. *Bayesian approach to inverse problems*. ISTE. Wiley, 2013 (cit. on p. 272).

[124] Nur Ismail, Cristine Calil Kores, Dimitri Geskus, and Markus Pollnau. “Fabry-Pérot resonator: spectral line shapes, generic and related Airy distributions, linewidths, finesses, and performance at low or frequency-dependent reflectivity”. In: *Optics Express* 24.15 (July 2016), p. 16366 (cit. on p. 2).

[125] John David Jackson. *Classical electrodynamics*. John Wiley & Sons, 2007 (cit. on p. 150).

[126] Sunil Prasad Jaiswal, Lu Fang, Vinit Jakhetiya, Jiahao Pang, Klaus Mueller, and Oscar Chi Au. “Adaptive multispectral demosaicking based on frequency-domain analysis of spectral correlation”. In: *IEEE Transactions on Image Processing* 26.2 (Feb. 2017), pp. 953–968 (cit. on p. 81).

[127] Jie Jia, Kenneth J. Barnard, and Keigo Hirakawa. “Fourier spectral filter array for optimal multispectral imaging”. In: *IEEE Transactions on Image Processing* 25.4 (Apr. 2016), pp. 1530–1543 (cit. on p. 209).

[128] Kyong Hwan Jin, Dongwook Lee, and Jong Chul Ye. “A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank Hankel matrix”. In: *IEEE Transactions on Computational Imaging* 2.4 (Dec. 2016), pp. 480–495 (cit. on pp. 148, 295).

[129] R. R. Jones, D. W. Schumacher, T. F. Gallagher, and P. H. Bucksbaum. “Bound-state interferometry using incoherent light”. In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 28.13 (1995), p. L405 (cit. on p. 169).

[130] Steven M. Kay. *Fundamentals of statistical processing, Volume I*. Prentice Hall, Mar. 26, 1993. 608 pp. (cit. on p. 249).

[131] Nicolas Keriven, Anthony Bourrier, Remi Gribonval, and Patrick Perez. “Sketching for large-scale learning of mixture models”. In: *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, Mar. 2016 (cit. on pp. 148, 295).

[132] Muhammad Jaleed Khan, Hamid Saeed Khan, Adeel Yousaf, Khurram Khurshid, and Asad Abbas. “Modern trends in hyperspectral image analysis: A review”. In: *IEEE Access* 6 (2018), pp. 14118–14129 (cit. on pp. 7, 8).

[133] Daisuke Kiku, Yusuke Monno, Masayuki Tanaka, and Masatoshi Okutomi. “Beyond color difference: Residual interpolation for color image demosaicking”. In: *IEEE Transactions on Image Processing* (2016), pp. 1–1 (cit. on pp. 80, 118).

[134] Daisuke Kiku, Yusuke Monno, Masayuki Tanaka, and Masatoshi Okutomi. “Residual interpolation for color image demosaicking”. In: *2013 IEEE International Conference on Image Processing*. IEEE, Sept. 2013 (cit. on p. 13).

[135] Yong Chan Kim, Hyeong-Geun Yu, Jae-Hoon Lee, Dong-Jo Park, and Hyun-Woo Nam. “Hazardous gas detection for FTIR-based hyperspectral imaging system using DNN and CNN”. In: *In Electro-Optical and Infrared Systems: Technology and Applications XIV, International Society for Optics and Photonics* (2017), 10433:1043317 (cit. on pp. 8, 181).

[136]Samuel Kotz, N. Balakrishnan, and Norman L. Johnson. *Continuous multivariate distributions, Volume 1: Models and applications*. Wiley-Interscience, 2000 (cit. on p. 96).

[137]Jonas Kuhn, Ulrich Platt, Nicole Bobrowski, and Thomas Wagner. "Towards imaging of atmospheric trace gases using Fabry–Pérot interferometer correlation spectroscopy in the UV and visible spectral range". In: *Atmospheric Measurement Techniques* 12.1 (2019), pp. 735–747 (cit. on p. 202).

[138]Saroj Kumar, Christine Desmedt, Denis Larsimont, Christos Sotiriou, and E. Goormaghtigh. "Change in the microenvironment of breast cancer studied by FTIR imaging". In: *The Analyst* 138 (May 2013) (cit. on p. 8).

[139]Lev Davidovich Landau, J. S. Bell, M. J. Kearsley, L. P. Pitaevskii, E. M. Lifshitz, and J. B. Sykes. *Electrodynamics of continuous media*. Vol. 8. elsevier, 2013 (cit. on p. 152).

[140]Pierre-Jean Lapray, Xingbo Wang, Jean-Baptiste Thomas, and Pierre Gouton. "Multispectral filter arrays: Recent advances and practical implementation". In: *Sensors* 14.11 (Nov. 2014), pp. 21626–21659 (cit. on p. 96).

[141]Reimar Lenz. *Optoelectronic colored image converter*. US Patent 5,877,807. Sept. 1997 (cit. on p. 54).

[142]Kenneth Levenberg. "A method for the solution of certain non-linear problems in least squares". In: *Quarterly of applied mathematics* 2.2 (1944), pp. 164–168 (cit. on p. 46).

[143]Marc Levoy. "Light fields and computational imaging". In: *Computer* 39.8 (2006), pp. 46–55 (cit. on p. 2).

[144]Muxingzi Li, Peihan Tu, and Wolfgang Heidrich. "Robust joint image reconstruction from color and monochrome cameras". In: BMVC. 2019 (cit. on p. 51).

[145]Xin Li, Bahadir Gunturk, and Lei Zhang. "Image demosaicing: A systematic survey". In: *Visual Communications and Image Processing 2008*. Vol. 6822. International Society for Optics and Photonics. 2008, 68221J (cit. on p. 13).

[146]Daniele Liciotti, Marina Paolanti, Emanuele Frontoni, and Primo Zingaretti. "People detection and tracking from and RGB-D camera in top-view configuration: Review of challenges and applications". In: 2017, pp. 207–18 (cit. on p. 7).

[147]Laetitia Loncan, Luis B. de Almeida, Jose M. Bioucas-Dias, et al. "Hyperspectral pansharpening: A review". In: *IEEE Geoscience and Remote Sensing Magazine* 3.3 (Sept. 2015), pp. 27–46 (cit. on pp. 13, 52).

[148]Ignace Loris and Caroline Verhoeven. "On a generalization of the iterative soft-thresholding algorithm for the case of non-separable penalty". In: *Inverse Problems* 27.12 (Nov. 2011), p. 125007 (cit. on pp. 43, 44, 103, 276).

[149]Yue M. Lu, Mina Karzand, and Martin Vetterli. "Demosaicking by alternating projections: Theory and fast one-step implementation". In: *IEEE Transactions on Image Processing* 19.8 (Aug. 2010), pp. 2085–2098 (cit. on p. 118).

[150] Alice Lucas, Michael Iliadis, Rafael Molina, and Aggelos K. Katsaggelos. “Using deep neural networks for inverse problems in imaging: beyond analytical methods”. In: *IEEE Signal Processing Magazine* 35.1 (2018), pp. 20–36 (cit. on p. 290).

[151] Valentina Luciani, Giuseppe Bonifazi, Peter Rem, and Silvia Serranti. “Upgrading of PVC rich wastes by magnetic density separation and hyperspectral imaging quality control”. In: *Waste Management* 45 (Nov. 2015), pp. 118–125 (cit. on p. 8).

[152] R. Lukac and K. N. Plataniotis. “Color filter arrays: Design and performance analysis”. In: *IEEE Transactions on Consumer Electronics* 51.4 (Nov. 2005), pp. 1260–1267 (cit. on pp. 73, 75, 96).

[153] Andrew Lumsdaine and Todor Georgiev. “The focused plenoptic camera”. In: *2009 IEEE International Conference on Computational Photography (ICCP)*. IEEE, Apr. 2009 (cit. on pp. 190, 191).

[154] Joseph N. Mait, Gary W. Euliss, and Ravindra A. Athale. “Computational imaging”. In: *Advances in Optics and Photonics* 10.2 (May 2018), p. 409 (cit. on p. 9).

[155] Stéphane Mallat. *A wavelet tour of signal processing*. Elsevier, 2009 (cit. on p. 33).

[156] Donald W. Marquardt. “An algorithm for least-squares estimation of nonlinear parameters”. In: *Journal of the Society for Industrial and Applied Mathematics* 11.2 (June 1963), pp. 431–441 (cit. on p. 46).

[157] Giuseppe Masi, Davide Cozzolino, Luisa Verdoliva, and Giuseppe Scarpa. “Pansharpening by convolutional neural networks”. In: *Remote Sensing* 8.7 (July 2016), p. 594 (cit. on p. 69).

[158] Simon Masnou and Jean-Michel Morel. “On a variational theory of image amodal completion”. In: *Rendiconti del Seminario Matematico della Universita di Padova* 116 (2006), pp. 211–252 (cit. on pp. 148, 295).

[159] Michael T. McCann, Kyong Hwan Jin, and Michael Unser. “Convolutional neural networks for inverse problems in imaging: A review”. In: *IEEE Signal Processing Magazine* 34.6 (2017), pp. 85–95 (cit. on p. 290).

[160] Shahar Mendelson, Alain Pajor, and Nicole Tomczak-Jaegermann. “Uniform uncertainty principle for Bernoulli and subgaussian ensembles”. In: *Constructive Approximation* 28.3 (Feb. 2008), pp. 277–289 (cit. on p. 95).

[161] L. Miao, H. Qi, R. Ramanath, and W. E. Snyder. “Binary tree-based generic demosaicking algorithm for multispectral filter arrays”. In: *IEEE Transactions on Image Processing* 15.11 (Nov. 2006), pp. 3550–3558 (cit. on p. 78).

[162] Lidan Miao and Hairong Qi. “The design and evaluation of a generic method for generating mosaicked multispectral filter arrays”. In: *IEEE Transactions on Image Processing* 15.9 (Sept. 2006), pp. 2780–2791 (cit. on pp. 71, 72, 78, 110, 137).

[163] Sofiane Mihoubi, Olivier Losson, Benjamin Mathon, and Ludovic Macaire. “Multispectral demosaicing using intensity-based spectral correlation”. In: *2015 International Conference on Image Processing Theory, Tools and Applications (IPTA)*. IEEE, Nov. 2015 (cit. on pp. 81, 118).

[164]Sofiane Mihoubi, Olivier Losson, Benjamin Mathon, and Ludovic Macaire. “Multispectral demosaicing using pseudo-panchromatic image”. In: *IEEE Transactions on Computational Imaging* 3.4 (Dec. 2017), pp. 982–995 (cit. on pp. 81, 82, 118).

[165]Junya Mizutani, Shu Ogawa, Kazuma Shinoda, Madoka Hasegawa, and Shigeo Kato. “Multispectral demosaicking algorithm based on inter-channel correlation”. In: *2014 IEEE Visual Communications and Image Processing Conference*. IEEE, Dec. 2014 (cit. on p. 79).

[166]Yusuke Monno, Daisuke Kiku, Masayuki Tanaka, and Masatoshi Okutomi. “Adaptive residual interpolation for color and multispectral image demosaicking”. In: *Sensors* 17.12 (Dec. 2017), p. 2787 (cit. on pp. 80, 118, 293).

[167]Yusuke Monno, Sunao Kikuchi, Masayuki Tanaka, and Masatoshi Okutomi. “A practical one-shot multispectral imaging system using a single image sensor”. In: *IEEE Transactions on Image Processing* 24.10 (Oct. 2015), pp. 3048–3059 (cit. on pp. 54, 71, 137).

[168]Jorge J. Moré. “The Levenberg-Marquardt algorithm: implementation and theory”. In: *Numerical analysis*. Springer, 1978, pp. 105–116 (cit. on p. 231).

[169]Jean Jacques Moreau. “Décomposition orthogonale d’un espace hilbertien selon deux cônes mutuellement polaires”. In: *Comptes rendus hebdomadaires des séances de l’Académie des sciences* 255 (1962), pp. 238–240 (cit. on p. 42).

[170]V. A. Morozov. *Methods for solving incorrectly posed problems*. New York, NY: Springer New York, 1984 (cit. on p. 21).

[171]Vladimir Alekseevich Morozov. “On the solution of functional equations by the method of regularization”. In: *Doklady Akademii Nauk*. Vol. 167. 3. Russian Academy of Sciences. 1966, pp. 510–512 (cit. on p. 30).

[172]A. Moshtaghpour, J. M. Bioucas-Dias, and L. Jacques. “Compressive single-pixel Fourier transform imaging using structured illumination”. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 12-17 May, Brighton, United Kingdom*. 2019, pp. 7810–7814 (cit. on p. 209).

[173]Bouthayna Msellmi, Daniele Picone, Mauro Dalla Mura, Zouhaier Ben Rabah, and Imed Riadh Farah. “Isotropic total variation minimization for sub-pixel mapping”. In: *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, July 2019 (cit. on p. 18).

[174]Bouthayna Msellmi, Daniele Picone, Zouhaier Ben Rabah, Mauro Dalla Mura, and Imed Riadh Farah. “Sub-pixel mapping model based on total variation regularization and learned spatial dictionary”. In: *Remote Sensing* 13.2 (Jan. 2021), p. 190 (cit. on p. 18).

[175]Laurent Najman and Hugues Talbot. *Mathematical morphology: From theory to applications*. John Wiley & Sons, 2013 (cit. on p. 220).

[176]Junichi Nakamura. *Image sensors and signal processing for digital still cameras (Optical science and engineering)*. CRC Press, 2005 (cit. on pp. 17, 89).

[177] Mila Nikolova and Pauline Tan. “Alternating structure-adapted proximal gradient descent for nonconvex block-regularised problems”. In: (Jan. 2018) (cit. on p. 190).

[178] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol. “Multiresolution-based image fusion with additive wavelet decomposition”. In: *IEEE Transactions on Geoscience and Remote Sensing* 37.3 (May 1999), pp. 1204–1211 (cit. on pp. 66, 118).

[179] Gregory Ongie, Ajil Jalal, Christopher A. Metzler, Richard G. Baraniuk, Alexandros G. Dimakis, and Rebecca Willett. “Deep learning techniques for inverse problems in imaging”. In: *IEEE Journal on Selected Areas in Information Theory* 1.1 (May 2020), pp. 39–56 (cit. on pp. 290, 296).

[180] Nobuyuki Otsu. “A threshold selection method from gray-level histograms”. In: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (Jan. 1979), pp. 62–66 (cit. on p. 222).

[181] Neal Parikh and Stephen Boyd. “Proximal algorithms”. In: *Foundations and Trends® in Optimization* 1.3 (2014), pp. 127–239 (cit. on pp. 40, 43, 275, 291, 292).

[182] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang. “Super-resolution image reconstruction: A technical overview”. In: *IEEE Signal Processing Magazine* 20.3 (May 2003), pp. 21–36 (cit. on p. 296).

[183] I. Pekkucuksen and Y. Altunbasak. “Multiscale gradients-based color filter array interpolation”. In: *IEEE Transactions on Image Processing* 22.1 (Jan. 2013), pp. 157–165 (cit. on p. 118).

[184] Marcelo Pereyra, Jose M. Bioucas-Dias, and Mario A. T. Figueiredo. “Maximum-a-posteriori estimation with unknown regularisation parameters”. In: *2015 23rd European Signal Processing Conference (EUSIPCO)*. IEEE, Aug. 2015 (cit. on pp. 30, 31).

[185] A. Perot and Charles Fabry. “On the application of interference phenomena to the solution of various problems of spectroscopy and metrology”. In: *The Astrophysical Journal* 9 (1899), p. 87 (cit. on p. 173).

[186] James A. De Haseth Peter Griffiths. *Fourier transform infrared spectrometry*. John Wiley & Sons, Aug. 20, 2007. 704 pp. (cit. on p. 10).

[187] Daniele Picone, Laurent Condat, Florian Cotte, and Mauro Dalla Mura. “Image fusion and reconstruction of compressed data: A joint approach”. In: *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, Oct. 2018 (cit. on pp. 2, 12, 18, 292).

[188] Daniele Picone, Laurent Condat, and Mauro Dalla Mura. “Analysis of masks for compressed acquisitions in variational-based pansharpening”. In: *CoSeRa*. Citeseer, Sept. 2018 (cit. on pp. 18, 96, 292).

[189] Daniele Picone, Mauro Dalla Mura, and Laurent Condat. “Pansharpening of images acquired with color filter arrays”. In: *Unconventional Optical Imaging*. Ed. by Corinne Fournier, Marc P. Georges, and Gabriel Popescu. International Society for Optics and Photonics. SPIE, May 2018 (cit. on p. 18).

[190]Daniele Picone, Aneline Dolet, Silvere Gousset, Didier Voisin, Mauro Dalla Mura, and Etienne Le Coarer. “Characterisation of a snapshot Fourier transform imaging spectrometer based on an array of Fabry-Perot interferometers”. In: *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, May 2020 (cit. on pp. 18, 251, 259, 293).

[191]Daniele Picone, Rocco Restaino, Gemine Vivone, Paolo Addesso, and Jocelyn Chanussot. “Pansharpening of hyperspectral images: Exploiting data acquired by multiple platforms”. In: *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, July 2016 (cit. on p. 49).

[192]Daniele Picone, Rocco Restaino, Gemine Vivone, Paolo Addesso, Mauro Dalla Mura, and Jocelyn Chanussot. “Band assignment approaches for hyperspectral sharpening”. In: *IEEE Geoscience and Remote Sensing Letters* 14.5 (May 2017), pp. 739–743 (cit. on p. 68).

[193]Ulrich Platt and Jochen Stutz. “Differential optical absorption spectroscopy”. In: *Physics of Earth and Space Environments*. Springer Berlin Heidelberg, 2008, pp. 135–174 (cit. on p. 208).

[194]Armande Pola Fossi. “Miniaturization of an infrared hyperspectral camera”. Theses. Université Paris-Saclay, Dec. 2016 (cit. on pp. 7, 181).

[195]John Henry Poynting. “XV. On the transfer of energy in the electromagnetic field”. In: *Philosophical Transactions of the Royal Society of London* 175 (1884), pp. 343–361 (cit. on p. 155).

[196]John G. Proakis and Dimitris G. Manolakis. *Digital signal processing*. Pearson Education, 2013 (cit. on p. 301).

[197]S. Ramani, Zhihao Liu, J. Rosen, J. Nielsen, and J. A. Fessler. “Regularization parameter selection for nonlinear iterative image restoration and MRI reconstruction using GCV and SURE-based methods”. In: *IEEE Transactions on Image Processing* 21.8 (Aug. 2012), pp. 3659–3672 (cit. on p. 30).

[198]K. Ramamohan Rao and Ping Yip. *Discrete cosine transform: Algorithms, advantages, applications*. Academic press, 2014 (cit. on pp. 32, 272).

[199]Gabriele Reich. “Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications”. In: *Advanced Drug Delivery Reviews* 57.8 (June 2005), pp. 1109–1143 (cit. on p. 8).

[200]Leonid I. Rudin, Stanley Osher, and Emad Fatemi. “Nonlinear total variation based noise removal algorithms”. In: *Physica D: Nonlinear Phenomena* 60.1-4 (1992), pp. 259–268 (cit. on pp. 33, 89, 102).

[201]J. Saastamoinen. “Atmospheric correction for the troposphere and stratosphere in radio ranging satellites”. In: *The use of artificial satellites for geodesy* 15 (1972), pp. 247–251 (cit. on p. 208).

[202]Bahaa E. A. Saleh and Malvin Carl Teich. *Fundamentals of photonics*. John Wiley & Sons, Inc., Aug. 1991 (cit. on pp. 11, 151, 160, 163).

[203] G. Sapiro and D. L. Ringach. “Anisotropic diffusion of multivalued images with applications to color filtering”. In: *IEEE Transactions on Image Processing* 5.11 (1996), pp. 1582–1586 (cit. on p. 38).

[204] Robert A. Schowengerdt. *Remote sensing: Models and methods for image processing*. Second. Orlando, FL, USA: Elsevier, 2006 (cit. on p. 208).

[205] Larry Schumaker. *Spline functions: Basic theory*. Cambridge University Press, 2007 (cit. on p. 61).

[206] Jean Serra. “Introduction to mathematical morphology”. In: *Computer vision, graphics, and image processing* 35.3 (1986), pp. 283–305 (cit. on p. 220).

[207] Silvia Serranti, Roberta Palmieri, and Giuseppe Bonifazi. “Hyperspectral imaging applied to demolition waste recycling: innovative approach for product quality control”. In: *Journal of Electronic Imaging* 24.4 (July 2015), p. 043003 (cit. on p. 8).

[208] Adi Shay, Isaac Y. August, and Adrian Stern. “Compressive and classical hyperspectral systems: a fundamental comparison”. In: *Compressive Sensing IV*. Ed. by Fauzia Ahmad. SPIE, May 2015 (cit. on pp. 7, 181).

[209] Miguel Simoes, Jose Bioucas-Dias, Luis B. Almeida, and Jocelyn Chanussot. “A convex formulation for hyperspectral image superresolution via subspace-based regularization”. In: *IEEE Transactions on Geoscience and Remote Sensing* 53.6 (June 2015), pp. 3373–3388 (cit. on pp. 67, 68).

[210] Daniel-Chen Soncco, Clara Barbanson, Mila Nikolova, Andres Almansa, and Yann Ferrec. “Fast and accurate multiplicative decomposition for fringe removal in interferometric images”. In: *IEEE Transactions on Computational Imaging* 3.2 (June 2017), pp. 187–201 (cit. on p. 190).

[211] Charles M. Stein. “Estimation of the mean of a multivariate normal distribution”. In: *The annals of Statistics* (1981), pp. 1135–1151 (cit. on p. 142).

[212] Ying Sun, Prabhu Babu, and Daniel P. Palomar. “Majorization-minimization algorithms in signal processing, communications, and machine learning”. In: *IEEE Transactions on Signal Processing* 65.3 (2016), pp. 794–816 (cit. on pp. 31, 290, 296).

[213] George Susanu, Stefan Petrescu, Florin Nanu, Adrian Capata, and Peter Corcoran. *RGBW sensor array*. US Patent 8,878,967. 2012 (cit. on pp. 75, 93, 97).

[214] Runjie Tan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. “Color image demosaicking via deep residual learning”. In: *IEEE Int. Conf. Multimedia and Expo (ICME)*. Vol. 2. 4. 2017, p. 6 (cit. on p. 82).

[215] Hayato Teranaka, Yusuke Monno, Masayuki Tanaka, and Masatoshi Ok. “Single-sensor RGB and NIR image acquisition: Toward optimal performance by taking account of CFA pattern, demosaicking, and color correction”. In: *Electronic Imaging* 2016.18 (Feb. 2016), pp. 1–6 (cit. on p. 80).

[216] Claire Thomas and Lucien Wald. “Analysis of changes in quality assessment with scale”. In: *2006 9th International Conference on Information Fusion*. IEEE, July 2006, pp. 1–5 (cit. on p. 5).

[217]Robert Tibshirani. “Regression shrinkage and selection via the lasso”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1 (Jan. 1996), pp. 267–288 (cit. on p. 32).

[218]Robert Tibshirani. “Regression shrinkage and selection via the lasso: A retrospective”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73.3 (Apr. 2011), pp. 273–282 (cit. on pp. 103, 274).

[219]A. N. Tikhonov. *Numerical methods for the solution of ill-posed problems*. Dordrecht: Springer Netherlands, 1995 (cit. on pp. 23, 27, 29).

[220]Adam K. Tilling, Garry J. O’Leary, Jelle G. Ferwerda, Simon D. Jones, Glenn J. Fitzgerald, Daniel Rodriguez, and Robert Belford. “Remote sensing of nitrogen and water stress in wheat”. In: *Field Crops Research* 104.1-3 (Oct. 2007), pp. 77–85 (cit. on pp. 8, 181).

[221]Francesco Tonolini, Ashley Lyons, Piergiorgio Caramazza, Daniele Faccio, and Roderick Murray-Smith. “Variational inference for computational imaging inverse problems”. In: *CoRR* abs/1904.06264 (2019). arXiv: 1904.06264 (cit. on p. 296).

[222]Joel A. Tropp and Anna C. Gilbert. “Signal recovery from random measurements via orthogonal matching pursuit”. In: *IEEE Transactions on Information Theory* 53.12 (Dec. 2007), pp. 4655–4666 (cit. on p. 32).

[223]Rodney G. Vaughan, Neil L. Scott, and D. Rod White. “The theory of bandpass sampling”. In: *IEEE Transactions on signal processing* 39.9 (1991), pp. 1973–1984 (cit. on p. 199).

[224]Gemine Vivone, Luciano Alparone, Jocelyn Chanussot, Mauro Dalla Mura, Andrea Garzelli, Giorgio A. Licciardi, Rocco Restaino, and Lucien Wald. “A critical comparison among pansharpening algorithms”. In: *IEEE Transactions on Geoscience and Remote Sensing* 53.5 (May 2015), pp. 2565–2586 (cit. on pp. 13, 52, 64, 65).

[225]Gemine Vivone, Miguel Simoes, Mauro Dalla Mura, Rocco Restaino, Jose M. Bioucas-Dias, Giorgio A. Licciardi, and Jocelyn Chanussot. “Pansharpening based on semiblind deconvolution”. In: *IEEE Transactions on Geoscience and Remote Sensing* 53.4 (Apr. 2015), pp. 1997–2010 (cit. on p. 143).

[226]L. Wald. *Data fusion: Definitions and architectures — Fusion of images of different spatial resolutions*. Paris, France: Les Presses de l’École des Mines, 2002 (cit. on p. 84).

[227]Lucien Wald, Thierry Ranchin, and Marc Mangolini. “Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images”. In: *Photogrammetric engineering and remote sensing* 63.6 (June 1997), pp. 691–699 (cit. on pp. 83, 108).

[228]Xingbo Wang, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Pierre Gouton. “Discrete wavelet transform based multispectral filter array demosaicking”. In: *2013 Colour and Visual Computing Symposium (CVCS)*. IEEE, Sept. 2013 (cit. on p. 82).

[229] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. “Image quality assessment: from error visibility to structural similarity”. In: *IEEE Transactions on Image Processing* 13.4 (Apr. 2004), pp. 600–612 (cit. on pp. 85, 238).

[230] Zhou Wang and A. C. Bovik. “A universal image quality index”. In: *IEEE Signal Processing Letters* 9.3 (Mar. 2002), pp. 81–84 (cit. on p. 85).

[231] Qi Wei, José Bioucas-Dias, Nicolas Dobigeon, and Jean-Yves Tourneret. “Hyperspectral and multispectral image fusion based on a sparse representation”. In: *IEEE Transactions on Geoscience and Remote Sensing* 53.7 (2015), pp. 3658–3668 (cit. on p. 67).

[232] Qi Wei, Nicolas Dobigeon, and Jean-Yves Tourneret. “Bayesian fusion of multi-band images”. In: *IEEE Journal of Selected Topics in Signal Processing* 9.6 (Sept. 2015), pp. 1117–1127 (cit. on pp. 68, 118).

[233] Qi Wei, Nicolas Dobigeon, and Jean-Yves Tourneret. “Fast fusion of multi-band images based on solving a Sylvester equation”. In: *IEEE Transactions on Image Processing* 24.11 (Nov. 2015), pp. 4109–4121 (cit. on p. 68).

[234] Emil Wolf. *Introduction to the theory of coherence and polarization of light*. Cambridge University Press, 2007 (cit. on p. 169).

[235] Clair Wyatt. *Radiometric calibration: theory and methods*. Elsevier, 2012 (cit. on p. 155).

[236] Seisuke Yamanaka. *Solid state color camera*. US Patent 4,054,906. Oct. 1977 (cit. on pp. 73, 75).

[237] Fumihito Yasuma, Tomoo Mitsunaga, Daisuke Iso, and Shree K. Nayar. “Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum”. In: *IEEE Transactions on Image Processing* 19.9 (Sept. 2010), pp. 2241–2253 (cit. on p. 54).

[238] Naoto Yokoya, Jocelyn Chanussot, and Akira Iwasaki. “Hyperspectral and multispectral data fusion based on nonlinear unmixing”. In: *2012 4th Workshop on Hyperspectral Image and Signal Processing (WHISPERS)*. 1-4. IEEE, June 2012, pp. 147–149 (cit. on p. 68).

[239] Naoto Yokoya, Claas Grohnfeldt, and Jocelyn Chanussot. “Hyperspectral and multispectral data fusion: A comparative review of the recent literature”. In: *IEEE Geoscience and Remote Sensing Magazine* 5.2 (June 2017), pp. 29–56 (cit. on p. 52).

[240] Guoxia Yu, Tanya Vladimirova, and Martin N. Sweeting. “Image compression systems on board satellites”. In: *Acta Astronautica* 64.9-10 (May 2009), pp. 988–1005 (cit. on pp. 87, 112).

[241] Qiangqiang Yuan, Yancong Wei, Xiangchao Meng, Huanfeng Shen, and Liangpei Zhang. “A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11.3 (Mar. 2018), pp. 978–989 (cit. on p. 69).

[242]Roberta H. Yuhas, Alexander F. H. Goetz, and Joe W. Boardman. “Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm”. In: *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*. Vol. 1. 1992, pp. 147–149 (cit. on p. 85).

[243]Mingli Zhang and Christian Desrosiers. “High-quality image restoration using low-rank patch regularization and global structure sparsity”. In: *IEEE Transactions on Image Processing* 28.2 (Feb. 2019), pp. 868–879 (cit. on pp. 148, 295).

[244]Dale Zimmerman, Claire Pavlik, Amy Ruggles, and Marc P. Armstrong. “An experimental comparison of ordinary and universal kriging and inverse distance weighting”. In: *Mathematical Geology* 31.4 (1999), pp. 375–390 (cit. on p. 62).

Webpages

[@1]ABB Ltd. *ABB Ltd*. 2010. URL: http://www.abb.com/ (visited on July 15, 2021) (cit. on p. 8).

[@2]Computer Vision Laboratory of Columbia University. *CAVE dataset*. 2010. URL: https://www.cs.columbia.edu/CAVE/databases/multispectral/ (visited on Aug. 12, 2020) (cit. on p. 54).

[@3]Mauro Dalla Mura. *FUsion MULTIspectral-imSPOC (FuMultiSPOC)*. 2021. URL: https://anr.fr/Project-ANR-20-ASTR-0006 (visited on June 28, 2021) (cit. on pp. 296, 308).

[@4]Foveon. *Foveon X3 sensor*. 2020. URL: http://www.foveon.com/article.php?a=247 (visited on Aug. 12, 2020) (cit. on pp. 53, 54).

[@5]Harvard University. *Harvard dataset*. 2020. URL: http://vision.seas.harvard.edu/hyperspec/download.html (visited on Aug. 12, 2020) (cit. on p. 54).

[@6]Interuniversity Microelectronics Centre. *Interuniversity Microelectronics Centre*. 2010. URL: http://www.imec-int.com/ (visited on July 15, 2021) (cit. on p. 8).

[@7]Joint Photographic Experts Group. *JPEG 2000 suite of standards*. 2020. URL: https://jpeg.org/jpeg2000/index.html (visited on Aug. 12, 2020) (cit. on pp. 32, 111).

[@8]Etienne Le Coarer. *Space CARBon Observatory - SCARBO*. 2018. URL: https://scarbo-h2020.eu/project (visited on June 28, 2021) (cit. on p. 308).

[@9]MAXAR Technologies. *MAXAR product samples*. 2020. URL: https://www.maxar.com/product-samples (visited on Aug. 12, 2020) (cit. on pp. 51, 109).

[@10]Silios Technologies. *Color SHADES by Silios*. 2020. URL: https://www.silios.com/multispectral-imaging (visited on Aug. 12, 2020) (cit. on pp. 93, 96).

[@11]Smart Vision Lights. *Smart Vision Lights*. 2021. URL: https://smartvisionlights.com/ (visited on July 18, 2021) (cit. on p. 6).

[@12]Teledyne e2v. *Onyx 1.3m - EV76C664 - CMOS image sensor*. 2020. URL: https://imaging.teledyne-e2v.com/products/standard-image-sensors/cmos-standard-image-sensors/onyx/ (visited on Aug. 12, 2020) (cit. on pp. 75, 93, 97).

[@13]Tokyo Institute of Technology. *TokyoTech dataset*. 2020. URL: http://www.ok.sc.e.titech.ac.jp/res/MSI/MSIdata31.html (visited on Aug. 12, 2020) (cit. on p. 54).

[@14]Kuniaki Uto. *Estimation of crop vitality based on tensor decomposition and data fusion of multimodal, multitemporal leaf-scale aerial images*. 2010. URL: https://www.jst.go.jp/kisoken/presto/en/project/1112075/1112075_14.html (visited on Aug. 12, 2020) (cit. on p. 19).

Declaration

I solemnly declare to have completed this work solely and only with the help of the references I mentioned.

Saint Martin d'Hères, November 25, 2021

Daniele Picone