Dématriçage et démélange conjoints d'images multispectrales

Kinan Abbas

Thèse de Doctorat

Mention : Sciences et technologies de l'information et de la communication
Spécialité : Traitement du Signal et des Images

présentée à *l'École Doctorale en Sciences Technologie et Santé (ED 585)*

de l'Université du Littoral Côte d'Opale

par

Kinan ABBAS

pour obtenir le grade de Docteur de l'Université du Littoral Côte d'Opale

Dématriçage et démélange conjoints d'images multispectrales

Soutenue le 3 septembre 2024, après avis des rapporteurs, devant le jury de thèse :

M. Y. DEVILLE, Professeur, Université Toulouse 3 Paul Sabatier	**Président**
M. P. DESBARATS, Professeur, Université de Bordeaux	**Rapporteur**
M. J. CHANUSSOT, Directeur de Recherche, INRIA	**Rapporteur**
M. R. BOYER, Professeur, Université de Lille	**Examinateur**
Mme H. CHATOUX, Maîtresse de Conférences, Université de Bourgogne	**Examinatrice**
M. M. PUIGT, Professeur, Université du Littoral Côte d'Opale	**Co-Directeur de thèse**
M. G. ROUSSEL, Professeur, Université du Littoral Côte d'Opale	**Co-Directeur de thèse**
M. G. DELMAIRE, Maître de Conférences, Université du Littoral Côte d'Opale	**Membre invité**

Doctoral Thesis

Field: Information and Communication Sciences and Technologies
Specialty: Signal and Image Processing

Presented at ***Doctoral School in Sciences, Technology, and Health (ED 585)***

of the University of Littoral Côte d'Opale

by

Kinan ABBAS

To earn the Doctorate degree from University of Littoral Côte d'Opale

Joint Demosaicing and Unmixing of Multispectral Images

Defended on September 3rd 2024, following the assessment of the reviewers, before the thesis committee:

Mr. Y. DEVILLE, Professor, University Toulouse 3 Paul Sabatier	**President**
Mr. P. DESBARATS, Professor, University of Bordeaux	**Reviewer**
Mr. J. CHANUSSOT, Research Director, INRIA	**Reviewer**
Mr. R. BOYER, Professor, University of Lille	**Examiner**
Ms. H. CHATOUX, Associate Professor, University of Burgundy	**Examiner**
Mr. M. PUIGT, Professor, University of Littoral Côte d'Opale	**Co-Supervisor**
Mr. G. ROUSSEL, Professor, University of Littoral Côte d'Opale	**Co-Supervisor**
Mr. G. DELMAIRE, Associate Professor, University of Littoral Côte d'Opale	**Guest Member**

Dedications

To my father Nabeh and my mother Samera whose endless love and support paved my way. Your sacrifices shaped my character and illuminated my life path.

To my lovely sisters, Roaa, Riham, and Areege, for your pure love, encouragement, and support.

To my dearest people, Lama, Patrick, and Danny. I am forever grateful for your love and unconditional support.

To my country, Syria, and my people, who did not have the chance to follow their dreams. May peace and justice prevail, bringing an end to tyranny and pain.

Acknowledgements

Several people have contributed in one way or another to the realization of this PhD thesis.

I will start with **Pr. Matthieu Puigt**. I am really speechless in describing how grateful I am for your effect on my career. Your intelligence, attention to detail, and fantastic problem-solving skills are impressive. You have consistently raised my challenges and encouraged me to surpass them. I am proud and lucky that I worked under your supervision and am a better researcher because of you. I thank you from the bottom of my heart.

My thanks and gratitude to **Pr. Gilles Roussel**, for your support at all levels. You have been supportive from day one, even before I arrived at the lab. Thank you for all the scientific advice and assistance you provided. I have learned a lot from your guidance, wisdom, calmness, and different ways of seeing and solving problems. Thank you very much.

My sincere thanks to Assoc. Pr. Gilles Delmaire, who directly contributed to the success of my research through his creative ideas and deep insight into problems. I will never forget the kindness of your heart and how much you supported me personally during my first days in France when I had no friends.

I want to thank the CSI committee, namely Pr. Yannick Deville, Pr. Nicolas Vandenbroucke, and Pr. Rémi Cozot, for following up on my research and providing advice and guidance that enriched the research and raised its level.

I would also like to thank the thesis committee members, namely Pr. Jocelyn Chanussot, Pr. Pascal Desbarats, Pr. Yannick Deville, Pr. Rémy Boyer, Assoc. Pr. Hermine Chatoux and Assoc. Pr. Gilles Delmaire for agreeing to review my thesis work. I greatly appreciate your insightful comments and suggestions for improving my thesis.

I will never forget the most essential person in this work, Assoc. Pr. Ahed Alboody, who was the main reason I met this fantastic team. His endless support, along with his wife and children, helped me a lot on all levels. They were and still are my family.

I thank all the PhD students, doctors, and interns who stopped by the laboratory. All of them contributed directly or indirectly to the success of my research. In particular, I would like to mention my friends: Pierre Chatelain, Adam El Bergui, Mohamed Alimoussa, Chedly Ben Azizi, Claire Guilloteau, Erick Adje, Zeinab Jaber, and Cheick Tidiani Cissé.

I want to give special thanks to Gaëlle Compiegne in the LISIC secretary office for her valuable help and incredible commitment to work.

Lastly, I thank the Région Hauts-de-France for partly funding my Ph.D. fellowship. Experiments

presented in this thesis were carried out using the CALCULCO computing platform, supported by DSI/ULCO.

Kinan Abbas

Abstract

In this thesis, we consider images sensed by a miniaturized multispectral (MS) snapshot camera. Contrary to classical RGB cameras, MS imaging allows to observe a scene on tens of different wavelengths, allowing a much more precise analysis of the observed content.

While most MS cameras require a scan to generate an image, snapshot MS cameras can instantaneously provide images, or even videos. When the camera is miniaturized, instead of a 3D data cube, it gets a 2D image, each pixel being associated with a filtered version of the theoretical spectrum it should acquire. Post-processing, called "demosaicing", is then necessary to reconstruct a data cube. Furthermore, in each pixel of the image, the observed spectrum can be considered as a mixture of spectra of pure materials present in the pixel. Estimating these spectra named endmembers as well as their spatial distribution (named abundances) is called "unmixing". While a classical pipeline to process MS snapshot images is to first demosaice and then unmix the data, the work introduced in this thesis explores alternative strategies in which demosaicing and unmixing are jointly performed. Extending classical assumptions met in sparse component analysis and in remote sensing MS unmixing, we propose two different frameworks to restore and unmixing the acquired scene, based on low-rank matrix completion and deconvolution, respectively, the latter being specifically designed for Fabry-Perot filters used in the considered camera. The four proposed methods exhibit a far better unmixing enhancement than the variants they extend when the latter are applied to demosaiced data. Still, they allow a similar demosaicing performance as state-of-the-art methods.

The last part of this thesis introduces a deconvolution approach to restore the spectra of such cameras. Our contribution lies in the weights of the penalization term which are automatically set using the entropy of the Fabry-Perot harmonics. The proposed method exhibits a better spectrum restoration than the strategy proposed by the camera manufacturer and than the classical deconvolution technique it extends.

Keywords: Snapshot Spectral Imaging, Unmixing, Demosaicing, Low-Rank Approximation, Sparsity, Deconvolution.

Résumé

Dans cette thèse, nous considérons des images captées par une caméra multispectrale (MS) miniaturisée « snapshot ». Contrairement aux caméras RVB classiques, l'imagerie MS permet d'observer une scène sur des dizaines de longueurs d'onde différentes, permettant une analyse beaucoup plus précise du contenu observé. Alors que la plupart des caméras MS nécessitent un scan pour générer une image, les caméras MS snapshot peuvent fournir instantanément des images, voire des vidéos. Lorsque la caméra est miniaturisée, au lieu d'un cube de données 3D, elle fournit une image 2D, chaque pixel étant associé à une version filtrée du spectre théorique sensé être acquis. Un post-traitement, appelé « dématriçage », est alors nécessaire pour reconstruire le cube de données. De plus, dans chaque pixel de l'image, le spectre observé peut être considéré comme un mélange de spectres de matériaux purs présents dans le pixel. L'estimation de ces spectres nommés endmembers ainsi que leur distribution spatiale (appelée abondances) est appelée « démélange ».

Alors qu'un pipeline classique pour traiter les images MS snapshot consiste d'abord à dématricer puis à démélanger les données, les travaux présentés dans cette thèse explorent des stratégies alternatives dans lesquelles le dématriçage et le démélange sont effectués conjointement. En étendant les hypothèses classiques rencontrées dans l'analyse des composantes parcimonieuses et dans le démélange MS utilisé en télédétection, nous proposons deux cadres différents pour restaurer et démélanger la scène acquise, basés respectivement sur la complétion de matrice de faible rang et la déconvolution, cette dernière étant spécifiquement conçue pour les filtres Fabry-Pérot utilisés dans la caméra considérée. Les quatre méthodes proposées présentent une bien meilleure qualité de démélange que les variantes qu'elles étendent lorsque ces dernières sont appliquées à des données dématricées. Néanmoins, elles permettent des performances de dématriçage similaires à celles des méthodes de l'état de l'art.

La dernière partie de cette thèse introduit une approche de déconvolution pour restaurer les spectres de telles caméras. Notre contribution réside dans les poids du terme de pénalisation qui sont automatiquement fixés en utilisant l'entropie des harmoniques de Fabry-Pérot. La méthode proposée présente une meilleure restauration spectrale que la stratégie proposée par le fabricant de la caméra et que la technique de déconvolution classique qu'elle étend.

Mots clés : Approximation et Complétion de Matrices de Faible Rang, Imagerie Spectrale Snapshot, Démelange, Dématriçage, Déconvolution.

Contents

Dedications **2**

Acknowledgements **3**

Abstract **5**

Résumé **6**

List of Figures **15**

List of Tables **16**

List of Algorithms **17**

List of Acronyms **18**

List of Symbols **23**

Résumé étendu **25**

I.1 Cadre général 25
I.2 Motivation et Objectifs de la Thèse 26
I.3 Structure de la thèse 27
I.3.1 Chapitre 1 : Imagerie spectrale snapshot 27
I.3.2 Chapitre 2 : Dématriçage des images spectrales snapshot 28
I.3.3 Chapitre 3 : Post-traitement des images dématricées : techniques de démélange 30
I.3.4 Chapitre 4 : Méthodes localement de rang-un de démélange et dématriçage conjoints pour les images spectrales snapshot : un formalisme de complétion de matrices 31
I.3.5 Chapitre 5 : Méthodes localement de rang-un de démélange et dématriçage conjoints pour les images spectrales snapshot : Un formalisme basé sur le filtrage 33
I.3.6 Chapitre 6 : Validation expérimentale 35

I.3.7 Chapitre 7 : déconvolution pénalisée par une pondération liée à l'entropie . . 35
I.3.8 Conclusion générale et perspectives 35
I.3.8.1 Conclusion 35
I.3.8.2 Perspectives 37

List of the Author's Publications and Communications During the Ph.D Thesis **39**

General Introduction **41**
II.1 General Framework 41
II.2 Thesis Motivation and Objectives 42
II.3 Thesis Structure 42

I Literature Reviews 45

1 Snapshot Spectral Imaging **46**
1.1 Introduction 46
1.2 Hyperspectral Imaging 46
1.3 Snapshot Spectral Imaging 48
1.3.1 Integrated Field Spectrometry 49
1.3.2 Multispectral Beamsplitting (MSBS) 50
1.3.3 Multiaperture Filtered Camera (MAFC) 50
1.3.4 Computed Tomography Imaging Spectrometer (CTIS) 50
1.3.5 Coded Aperture Snapshot Spectral Imager (CASSI) 51
1.3.6 Image Mapping Spectrometer (IMS) 51
1.3.7 Image-replicating Imaging Spectrometer (IRIS) 52
1.3.8 Fabry-Perot-Based Sensors 52
1.3.8.1 Snapshot Tiled (SST) 53
1.3.8.2 Snapshot Mosaic (SSI) 55
1.3.9 Concluding Insights 55
1.4 Calibration of Fabry-Perot Based Cameras 58
1.4.1 Tracing the Problem's Roots 58
1.4.2 Band pass filters 59
1.4.3 Dark Level Correction (Bias Correction) 59
1.4.4 White Balancing (Reflectance Calculation) 59
1.4.5 Spectral Correction 60
1.4.6 Angularity Correction 60
1.4.7 Non-uniformity Correction 61
1.4.8 Final Key Insights 62
1.5 Conclusion 62

2 Demosaicing Snapshot Spectral Images **64**
2.1 Introduction . . . 64
2.2 Traditional Demosaicing Methods . . . 66
2.2.1 Demosaicing Based on Interpolation . . . 68
2.2.1.1 Weighted Bilinear Interpolation (WB) . . . 68
2.2.1.2 Spectral Difference (SD) . . . 68
2.2.1.3 Median Filtering in Multispectral Filter Array Demosaicing . . . 69
2.2.1.4 A Linear Interpolation Algorithm for Spectral Filter Array Demosacing (LISFAD) . . . 69
2.2.1.5 Iterative Spectral Difference (ItSD) . . . 69
2.2.1.6 Multispectral Local Directional Interpolation (MLDI) . . . 70
2.2.2 Demosaicing Based on Wavelet Analysis . . . 70
2.2.2.1 Discrete Wavelet Trasform (DWT) MSFA Demosaicing . . . 71
2.2.2.2 A Generic Multispectral Demosaicking Method Based on Interchannel Spectrum Correlation (GMDICC) . . . 71
2.2.3 Demosaicing Based on Binary Trees . . . 72
2.2.3.1 Binary Tree-based Edge-sensing (BTES) . . . 73
2.2.3.2 Probability of Appearance based Convolution Filter based Weighted Bilinear (PCWB) . . . 73
2.2.3.3 Probability of Appearance based Convolution Filter based Bilinear Spectral Difference (PCBSD) . . . 73
2.2.3.4 Generic Multispectral Demosaicking Based on Directional Interpolation (GMDDI) . . . 73
2.2.4 Demosaicing Based on Pseudo-panchromatic Analysis . . . 74
2.2.4.1 Pseudo-panchromatic Image Demosaicing (PPID) . . . 74
2.2.4.2 Generic Multispectral Demosaicking Using Spectral Correlation Between Spectral Bands and Pseudo-panchromatic Image (GMDSCPPI) . . . 74
2.2.5 Demosaicing Based on Low-rank Matrix Approximation . . . 75
2.2.5.1 Graph and Rank Regularized Demosaicing (GRMR) . . . 75
2.2.5.2 Multispectral Snapshot Demosaicing Via Non-convex Matrix Completion (MSDMC) . . . 75
2.2.5.3 Generalized MSFA Engineering With Structural and Adaptive Nonlocal Demosaicing (SAND) . . . 76
2.2.6 Discussion and Insights . . . 77
2.3 Deep Learning-based Methods . . . 79
2.3.1 Deep Demosaicking for Multispectral Filter Arrays (In-Net) . . . 80
2.3.2 Hyperspectral Demosaicking and Crosstalk Correction using Deep Learning (DsNet) . . . 80

2.3.3 Deep Panchromatic Image Guided Residual Interpolation for Multispectral Image Demosaicking (DGRI) . . . 80
2.3.4 Deep Convolutional Networks for Snapshot Hyperspectral Demosaicking (Sp-Net) . . . 80
2.3.5 Deep Learning Approach for Hyperspectral Image Demosaicking, Spectral Correction and High-resolution RGB Reconstruction . . . 80
2.3.6 Mosaic Convolution-Attention Network for Demosaicing Multispectral Filter Array Images . . . 81
2.3.7 Hyperspectral Demosaicing of Snapshot Camera Images Using Deep Learning 81
2.3.8 Joint Spatial-spectral Pattern Optimization and Hyperspectral Image Reconstruction . . . 82
2.3.9 MSFA-Frequency-Aware Transformer for Hyperspectral Images Demosaicing (FDM-Net) . . . 82
2.3.10 A Snapshot Multi-Spectral Demosaicing Method for Multi-Spectral Filter Array Images Based on Channel Attention Network . . . 83
2.3.11 Discussion and Insights . . . 83
2.4 Conclusion . . . 85

3 Post-Processing of Demosaiced Images: Unmixing Techniques 86
3.1 Introduction . . . 86
3.2 Fundamentals of Hyperspectral Unmixing . . . 87
3.3 Hyperspectral Unmixing Methods . . . 90
3.3.1 Geometrical Methods . . . 90
3.3.2 Sparse regression-based Methods . . . 92
3.3.3 Deep Learning-Based Methods . . . 92
3.3.4 Statistical Methods . . . 93
3.3.5 Discussion . . . 93
3.4 Nonnegative Matrix Factorization (NMF) . . . 93
3.4.1 Constrained Nonnegative Matrix Factorization . . . 94
3.4.2 Structured Nonnegative Matrix Factorization . . . 95
3.4.3 Generalized Nonnegative Matrix Factorization . . . 96
3.5 Sparse Component Analysis . . . 96
3.6 Conclusion . . . 98

II Proposed Frameworks 99

4 Locally-Rank-One-Based Joint Unmixing and Demosaicing Methods for Snapshot Spectral Images: a Matrix Completion Framework 100
4.1 Introduction . . . 100
4.2 Problem Statement . . . 102

4.3 Proposed Methods . . . 104
4.3.1 Naive Method . . . 104
4.3.2 Locally Rank-1 and Clustering-based Proposed Technique . . . 105
4.3.3 Method with Relaxed Abundance Sparsity Assumption . . . 107
4.3.4 Algorithms . . . 108
4.4 Conclusion . . . 109

5 Locally-Rank-One-Based Joint Unmixing and Demosaicing Methods for Snapshot Spectral Images: a Filtering-Based Framework 111
5.1 Introduction . . . 111
5.2 Problem Statement . . . 112
5.3 Proposed methods . . . 113
5.3.1 Classical Strategies . . . 114
5.3.2 Clustering and Filtering-based Proposed Technique . . . 114
5.3.3 Method with Relaxed Abundance Sparsity Assumption . . . 115
5.3.4 Algorithms . . . 116
5.4 Conclusion . . . 118

6 Experimental Validation 120
6.1 Introduction . . . 120
6.2 Experimental Validation of Matrix-Completion Framework . . . 120
6.2.1 Experimental Setup . . . 120
6.2.2 Performance Evaluation on Synthetic Images For The Ideal Case . . . 122
6.2.3 Performance Evaluation on CAVE Dataset . . . 126
6.2.4 Performance Evaluation on Hyko 2 Dataset . . . 127
6.3 Experimental Validation of Filtering-based Framework . . . 128
6.3.1 Experiment Setup . . . 128
6.3.2 The Impact of Regularization Parameter . . . 130
6.3.3 Performance Evaluation On Synthetic Images For The Real Case . . . 131
6.3.4 Performance Evaluation On Hyko 2 Dataset . . . 136
6.4 Conclusion . . . 136

7 Improved deconvolution with Entropy-weighted penalization 138
7.1 Introduction . . . 138
7.2 Preface . . . 138
7.3 Understanding the Role of Entropy . . . 141
7.4 Proposed Spectral Correction Method . . . 141
7.5 Experimental Validation . . . 143
7.5.1 Experimental Setup . . . 143
7.5.2 Synthetic Data Experiment . . . 144
7.5.3 Jasper Ridge image Experiment . . . 144

7.5.4 Real Data Experiment 145
7.5.5 Super Spectral Resolution 146
7.6 Conclusion 147

8 General Conclusion and Perspectives **148**
8.1 Conclusion 148
8.2 Perspectives 149
8.2.1 Running Time and Hyperspectral Video Processing 149
8.2.2 Spectral Variability 149
8.2.3 Fabry-Perot Filter Variability 150
8.2.4 Angularity Correction Integration 150
8.2.5 Integrating Deep Learning 150

List of Figures

1 Vue d'ensemble de l'acquisition et du traitement sur une scène contenant trois membres finaux (sable, roches et œufs). L'image SSI peut être traitée soit par une approche en deux étapes, soit en appliquant conjointement le démélange et le dématriçage pour restaurer le cube de données, extraire les membres finaux, et trouver les cartes d'abondance. 27
2 Disposition de l'arrangement des filtres couleur de Bayer 28
3 Une approche typique consiste à regrouper ensemble des groupes de pixels, connus sous le nom de "super-pixels", pour produire des images à résolution spatiale plus basse avec plus (milieu) ou toutes (en haut à droite) les bandes spectrales. Tandis que les approches de dématriçage visent à repeupler le cube complet tout en conservant la pleine résolution spatiale (en bas à droite). Adapté de [143]. 29
4 Exemple de démélange hyperspectral (source : adapté de [99]). 31
5 Etapes de traitement requises par le fabricant de la caméra pour fournir des images SSI de qualité à démélanger. 34
6 Overview of acquisition and processing on a scene contains three endmemebers (sand, rocks, and eggs). The SSI image can be processed either by a two-stage approach or by applying unmixing and demosaicing jointly to restore the data cube, extract the endmembers, and find the abundance maps. 43

1.1 Standard Scanning Methods (Source [154]). 47
1.2 The system layout for a computed tomography imaging spectrometer (Source [69]). . 51
1.3 Fabry-Perot Structure (Source IMEC) 53
1.4 Two packaged mosaic spectral snapshot sensors, alongside two tiny spectral cameras and a Euro coin for scale (Source [61]) 54
1.5 Filter organization in Linescan, Snapshot Tiled, and Snapshot Mosaic layouts, respectively. 54
1.6 Microscope images of snapshot mosaic imagers with different filter configurations. Each pixel (size 5.5×5.5 µm) within the cell is a different spectral filter. As can be seen, although current commercial designs use a 4×4 or 5×5 cell layout, other configurations are possible (Source [61]). 55
1.7 Snapshot mosaic 4×4 filter layout (Source IMEC) 56
1.8 Snapshot mosaic 5×5 filter layout with example image (Source [74]) 56

1.9 Spectral response of Fabry-Perot filter of the 5×5 mosaic snapshot camera (Source IMEC). 59
1.10 The measured spectrum of a sample appears shifted if placed at different distances from the optical axis. Using spectral shift correction, the spectra can be realigned (Source [62]). 61
1.11 Two measurement setups for obtaining the non-uniformity correction cube (Source [148]). 62
1.12 Processing steps as required by the camera manufacturer to ensure high-quality snapshot spectral image restoration and unmixing. 63

2.1 Bayer color filter array arrangement 65
2.2 A typical approach involves binning together groups of pixels, known as "superpixels", to produce lower spatial resolution images with more (middle) or all (top right) spectral bands. While demosaicing approaches aim to repopulate the full cube and keeping the full spatial resolution (bottom right). Adapted from [143]. 66
2.3 Different MSFAs with a spatial resolution 16×16 and spectral resolution 16 (Source [147]) 67
2.4 Main flow of MLDI 70
2.5 Comparative Analysis of Demosaicing Techniques through Venn Diagram 76
2.6 Mosaic convolution module (MCM) for one channel case 82

3.1 Example of hyperspectral unmixing (source: adapted from [99]). 86
3.2 Mixed and pure pixels in hyperspectral image (Source [33]) 87
3.3 Linear and Nonlinear mixing models (Source [134]) 88
3.4 Example of atmospheric interferers, multiple scattering, and shadows and variable Illumination effects (Source [33]) 90
3.5 Classification of Hyperspectral Unmixing Algorithms. Supervised methods are highlighted in orange, and unsupervised methods in blue. 91

4.1 Overview of acquisition and processing on a scene contains three endmemebers (sand, rocks, and eggs). The SSI image can be processed either by a two-stage approach or by applying unmixing and demosaicing jointly to restore the data cube, extract the endmembers, and find the abundance maps. 101
4.2 Principles of the two-step (in green) and joint (in red) strategies. 102

5.1 The spectral response of the 25 spectral filters of the 5×5 mosaic Photon Focus SSI camera covers the wavelength range from 400 to 1000 nm [119]. 113

6.1 Image 1, SSI image where assumptions 1 and assumption 2 are satisfied with abundance maps of the three endmembers. 121
6.2 Image 2, SSI image where assumptions 1 and assumption 3 are satisfied with abundance maps of the three endmembers. 121

6.3 Mean PSNR, SAM, SIR, MER and RMSE—obtained for Image 1 (Assumption 1&2) with 4×4 and 5×5 filters—relative to input SNR. 122
6.4 Mean PSNR, SAM, SIR, MER and RMSE—obtained for Image 2 (Assumption 1&3) with 4×4 and 5×5 filters—relative to input SNR. 123
6.5 Estimated spectra for the Image 1 with mosaic filter of size 5×5. 124
6.6 Estimated spectra for the Image 2 with mosaic filter of size 5×5. 124
6.7 Estimated abundance maps for the Image 1 with 5×5 mosaic filter. 125
6.8 Estimated abundance maps for the Image 2 with 5×5 mosaic filter. 126
6.9 Comparison of the demosaicing performance of all the methods on the Cloths image from the CAVE dataset with 4×4 mosaic filter. 128
6.10 Comparison of the segmentation performance of all the methods on the Hyko 2 Dataset Image. 129
6.11 Influence of the regularization parameter α on the achieved Spectral Angle Mapper (SAM) value for each noise level 130
6.12 From left to right: mean PSNR, SAM, SIR, MER and RMSE—obtained for Image 1 (Assumption 1&2) with 5×5 real filter—relative to input SNR. 131
6.13 From left to right: mean PSNR, SAM, SIR, MER and RMSE—obtained for Image 2 (Assumption 1&3) with 5×5 Real filter—relative to input SNR. 132
6.14 Estimated spectra for the Image 1 with Real filter of size 5×5. 133
6.15 Estimated spectra for the Image 2 with Real filter of size 5×5. 133
6.16 Estimated abundance maps for the Image 1 with real filter of size 5×5 134
6.17 Estimated abundance maps for the Image 2 with real filter of size 5×5 135
6.18 Comparison of the segmentation performance of all the methods on the Hyko 2 Dataset Image 137

7.1 The Hyperspectral Linescan camera from XIMEA using a wedge pattern of 192 VIS-NIR filters (Source [159]). 139
7.2 The essential steps in processing images acquired by snapshot spectral and linescan cameras. (Single Image Demosaicing vs. Multi-Image Stitching) 140
7.3 Fabry-Perot filter responses for different bands of the linescan camera. On the left: 805 nm. On the right: 650 nm. 141
7.4 Fabry-Perot filter responses (in blue) for one band from linescan and SSI cameras with highlighted virtual wavelengths (in red) 142
7.5 Reached SAM values vs the value of λ and the input SNR. 144
7.6 Results on synthetic simulation used USGS spectral data 145
7.7 Corrected spectral comparison for the linescan (LS) camera on the real data 146
7.8 Super Resolution Spectral Analysis 147

List of Tables

1.1 Comparison of Snapshot Spectral Imaging Technologies 57

2.1 Traditional demosaicing methods comparison. 77
2.2 Deep Learning Methods for MSFA Demosaicing. 83

6.1 PSNR, SAM, SIR, MER, RMSE, and Time in seconds obtained for the synthetic images with 5x5 and 4x4 (into brackets) ideal filters. In bold, the highest performance value and for RMSE and Time, the lowest value. 122
6.2 Demosaicing performance averaged over all the images in the Cave dataset. Perf. criterion: PSNR (in dB). 127
6.3 PSNR, SAM, SIR, MER, RMSE, and Time in seconds obtained for the synthetic images with 5x5 real filters. In bold, the highest performance value. 131

7.1 Spectral correction using the proposed SDS-Cor and IMEC methods for Jasper Ridge image. Performance criteria: PSNR, SAM, and SIR. 145

List of Algorithms

1 Naive Method 108
2 Rank-one patch detection and spectra estimation method used in both proposed KP-WNMF and VPWNMF 109
3 K-means (resp. K-medians) Patch-based Weighted Nonnegative Matrix Factorization (KPWNMF) 110
4 VCA Patch-based Weighted Nonnegative Matrix Factorization (VPWNMF) 110
5 Rank-one patch detection and spectra estimation method used in both proposed FP-Kmeans and FPVCA 117
6 Filter Patch-based Kmeans (resp. K-medians), (FPKmeans) 118
7 Filter Patch-based Vertex Component Analysis (FPVCA) 118

List of Acronyms

1. ALS - Alternating Least Squares
2. ASD - Alternating Steepest Descent
3. AWCA - Adaptive Weighted Channel Attention Module
4. BN - Blue Noise MSFA Design
5. BRA - Brauers MSFA Design
6. BTES - Binary Tree-based Edge-sensing
7. BSS - Blind Source Separation
8. CASSI - Coded Aperture Snapshot Spectral Imagers
9. CFA - Color Filter Array
10. CGIHT - Conjugate Gradient Iterative Hard Thresholding
11. CNN - Convolutional Neural Network
12. CT - Computed Tomography
13. CTIS - Computed Tomography Imaging Spectrometers
14. DAMCNet - Domain Adapted Multi-scale Channel-attention Network
15. DGRI - Deep Panchromatic Image Guided Residual Interpolation for Multispectral Image Demosaicking
16. DJMDA - Deep Joint Multispectral Demosaicing and Anti-clipping
17. DL - Deep Learning
18. DPI-Net - ResNet-based Deep Learning

19. DS - Down Sampled Image
20. DsNet - Hyperspectral Demosaicking and Crosstalk Correction using Deep Learning
21. DVMF - Direct Vector Median Filtering
22. DW - Discrete Wavelet
23. DWT - Discrete Wavelet Transform
24. FE - Feature Extraction
25. EM - Expectation-Maximization
26. FDM-Net - MSFA-Frequency-Aware Transformer for Hyperspectral Images Demosaicing
27. FPf - Fabry-Perot Filter
28. FPI - Fabry-Perot Interferometer
29. FOV - Field of view
30. FWHM - Full Width at Half Maximum
31. GAP-TV - Generalized Alternating Projection-based Total Variation
32. GMDDI - Generic Multispectral Demosaicking Based on Directional Interpolation
33. GMDICC - Generic Multispectral Demosaicking Method Based on Inter-Channel Spectrum Correlation
34. GMDSCPPI - Generic Multispectral Demosaicking Using Spectral Correlation between Spectral Bands and Pseudo-panchromatic Image
35. GRMR - Graph and Rank Regularized Demosaicing
36. GSAND - Generalized MSFA Engineering With Structural and Adaptive Nonlocal Demosaicing
37. HALS - Hierarchical Alternating Least Squares
38. HSI - Hyperspectral Imaging
39. HSR - High Spectral Reconstruction
40. HU - Hyperspectral Unmixing
41. ICA - Independent Component Analysis

42. IDM - Initial Demosaicing Module
43. IDWT - Inverse Discrete Wavelet Transform
44. IMS - Image Mapping Spectrometer
45. In-Net - Deep Demosaicking for Multispectral Filter Arrays
46. IRIS - Image-replicating Imaging Spectrometers
47. IFS - Integrated Field Spectrometry
48. IFS-F - Integrated Field Spectrometry with Fiber Arrays
49. IFS-L - Integrated Field Spectrometry with Lenslet Arrays
50. IFS-M - Integrated Field Spectrometry with Faceted Mirrors
51. IS - Itakura–Saito
52. ItSD - Iterative Spectral Difference
53. LISFAD - Linear Interpolation Algorithm for Spectral Filter Array Demosaicing
54. LRMA - Low-rank Matrix Approximation
55. LRMC - Low-Rank Matrix Completion
56. LS - Linescan Wedge
57. MAFC - Multiaperture Filtered Camera
58. MAM - Mosaic Attention Module
59. MCM - Mosaic Convolution Module
60. MLDI - Multispectral Local Directional Interpolation
61. MS-MCAN - Multi-Scale Mosaic Channel Attention Network
62. MSDMC - Multispectral Snapshot Demosaicing via Non-convex Matrix Completion
63. MSBS - Multispectral Beamsplitting
64. MSFA - Multispectral Filter Array
65. MU - Multiplicative Updates

66. NA - Numerical Aperture
67. ne - Northeast
68. NLRAN - Non-Local Residual Attention Network
69. NMF - Nonnegative Matrix Factorization
70. NTF - Nonnegative Tensor Factorization
71. nw - Northwest
72. OSP - Optimal Sphere Packing MSFA Design
73. PCBSD - Probability of Appearance based Convolution Filter based Bilinear Spectral Difference
74. PCWB - Probability of Appearance based Convolution Filter based Weighted Bilinear
75. PNMF - Project-based Nonnegative Matrix Factorization
76. PoA - Probability of Appearance
77. PPAN - Pseudo Panchromatic
78. PPI - Pseudo-Panchromatic Image
79. PPID - Pseudo-panchromatic Image Demosaicing
80. RCAN - Residual Channel Attention Network
81. RMVES - Robust Minimum Volume Enclosing Simplex
82. RND - Random MSFA Design
83. SCA - Sparse Component Analysis
84. SD - Spectral Difference
85. SD - Spectral Difference
86. se - Southeast
87. SEQ - Sequential MSFA Design
88. SMACC - Sequential Maximum Angle Convex Cone
89. SP - Sphere Packing

90. SpaDM - Spatial Demosaicing
91. SpeSR - Spectral Super-resolution
92. SpNet - Deep Convolutional Networks for Snapshot Hypercpectral Demosaicking
93. SSI - Snapshot Mosaic
94. SSFs - Spectral Sensitivity Functions
95. SSMF - Spherical Space Median Filtering
96. ST - Snapshot Tiled
97. sw - Southwest
98. UAV - Unmanned Aerial Vehicle
99. UNIF - Uniform MSFA Design
100. VCA - Vertex Component Analysis
101. WB - Weighted Bilinear Interpolation
102. WNMF - Weighted Nonnegative Matrix Factorization
103. WT - Wavelet Transform

List of Symbols

α, β - Parameters controlling the influence of regularization.

δ - A parameter used to adjust the influence of the sum-to-one constraint in the Naive method.

ω_i - Additive noise associated with the i-th pixel's measurement.

$\mathbf{C}$ - The correction matrix.

D - Matrix representing the discrete derivative operator.

F - The endmember matrix.

$\mathbb{F}$ - Matrix holding the collected spectra from patches.

G - The abundance matrix.

$\mathbf{H}$ - The response matrix of the camera.

$\mathbf{H}^{\text{ideal}}$ - The ideal response matrix.

$h_i(\lambda)$ - Fabry-Perot filter response associated with pixel i, showing the camera's response at wavelength λ.

k - Number of bands in an image. Total number of spectral bands captured by the camera. Number of pixels/bands in each patch of snapshot image.

$m \times n$ - Represents the dimensions of the SSI image in terms of pixel counts in the horizontal and vertical dimensions, respectively.

$\text{Max}_{\text{outIter}}$ - The maximum number of iterations allowed for the optimization process.

p - Number of distinct endmembers present in the scene.

$\mathbb{1}$ - Matrix of ones.

$\mathbb{1}_{m \times n}$ - Matrix of ones with size $m \times n$.

U - A predefined spectral library used in Project-based Nonnegative Matrix Factorization (PNMF).

V - Transformation matrix used in Project-based Nonnegative Matrix Factorization (PNMF).

W - Weight matrix used in Weighted Nonnegative Matrix Factorization (WNMF). Weight matrix indicating the wavelengths observed by the snapshot camera.

X - The data matrix in Nonnegative Matrix Factorization (NMF). Unfolded 2D representation of the theoretical 3D data cube, where each row corresponds to a spatial position and each column to a wavelength.

$\mathbb{X}$ - Matrix holding the collected spectra from patches.

$x_i(\lambda_i)$ - The i-th SSI pixel in a patch capturing spectral information at wavelength λ_i.

$\underline{x}$ - A k-dimensional vector representing spectral values of one SSI patch across k wavelengths.

$\hat{Y}$ - Estimated complete data matrix in Weighted Nonnegative Matrix Factorization (WNMF). Unfolded and expanded SSI image, representing the 3D data cube after demosaicing.

$\bar{Y}$, $\bar{F}$ - Augmented matrices of Y and F, respectively, incorporating an additional column of ones to facilitate the application of the abundance sum-to-one constraint.

$\hat{Y}$, $\hat{G}$, $\hat{F}$ - Estimated matrices of Y, G, and F.

$y_i(\lambda)$ - Actual spectrum intended to be observed by pixel i at wavelength λ.

$\|\cdot\|_F^2$ - The Frobenius norm

$\approx$ - Approximately equal

$\|\cdot\|_2^2$ - Squared Euclidean Norm

$\gg$ - Much greater than

$\triangleq$ - Is defined as or is equivalent to

$\circ$ - Hadamard product

Résumé étendu

I.1 Cadre général

L'imagerie hyperspectrale (HSI) est une technique analytique basée sur la spectroscopie, qui implique la capture de centaines d'images à diverses longueurs d'onde pour la même région spatiale. La véritable puissance de l'HSI réside dans sa capacité à détecter, identifier et quantifier des matériaux et des phénomènes dans diverses applications. De la surveillance précise de la santé agricole et de la surveillance environnementale [87], aux applications révolutionnaires dans les diagnostics médicaux et la sécurité alimentaire [48], l'HSI fournit un outil crucial pour la recherche et l'industrie.

Un défi fondamental auquel les capteurs HSI doivent faire face est l'acquisition de données HSI tridimensionnelles, englobant deux dimensions spatiales et une dimension spectrale, à l'aide d'un détecteur unique, tel qu'un détecteur en réseau 1D ou en plan 2D. Ainsi, différentes stratégies dans les conceptions d'acquisition HSI ont émergé de la disparité entre les demandes des détecteurs et la dimensionnalité disponible, menant à des méthodes d'acquisition par balayage de longueur d'onde, balayage ponctuel, balayage de ligne, et des méthodes d'acquisition instantanée [101, 151]. Néanmoins, une caractéristique commune à tous les scénarios est la nécessité de balayer à plusieurs reprises la scène et d'acquérir de nombreux clichés (cadres) pour compiler le cube de données à résolution spatio-spectrale complète, tout en tenant compte des considérations relatives au coût et à la taille de la caméra.

Une nouvelle génération d'architectures d'HSI[1], connue sous le nom d'imagerie spectrale "snapshot" (SSI), a été introduite pour répondre aux défis mentionnés précédemment [54]. La SSI permet l'acquisition efficace du contenu spatio-spectral de scènes dynamiques à l'aide de plateformes miniaturisées et peut acquérir le cube complet à partir d'une seule ou de quelques expositions. Pour atteindre cet objectif, les architectures SSI associent chaque pixel spatial à une bande spectrale spécifique, introduisant ainsi un compromis crucial entre la résolution spatiale et spectrale [143].

Ces caméras peuvent capturer des vidéos à un taux de trame élevé et sont utiles dans les cas où le mouvement est imprévisible, et la caméra ou l'objet se déplace en 2D ou 3D, comme dans les applications robotiques ou le tri des voies [79]. Parmi les stratégies récentes qui ont émergé, les caméras SSI utilisant des filtres Fabry-Perot (FPf) [53] produisent uniquement une image 2D dérivée des données

[1]En réalité, certaines de ces technologies ne permettent que l'acquisition d'images multispectrales, fournissant un nombre de bandes spectrales plus réduit que l'HSI. Néanmoins, dans la suite de ce résumé étendu, nous parlerons d'HSI pour indiquer à la fois des images multi- ou hyperspectrales.

hyperspectrales 3D. Ensuite, une technique de post-traitement appelée "dématriçage" est nécessaire pour estimer le cube de données hyperspectrales complet. Diverses approches ont été proposées pour effectuer le dématriçage. Ces approches peuvent être classées en deux catégories, incluant les méthodes "traditionnelles" [27, 105, 108, 107, 22, 143] et les stratégies basées sur l'apprentissage profond [39, 66, 116, 46, 178, 94].

Après que l'image hyperspectrale 3D ait été reconstruite à partir des données SSI 2D, nous pouvons employer toute technique de post-traitement de notre choix. En particulier, le processus de démélange implique l'extraction des signatures spectrales de tous les membres finaux présents dans une scène observée. Plus spécifiquement, les méthodes de démélange populaires fonctionnent sous l'hypothèse que pour chaque membre final, il existe au moins un pixel spatial où le matériau correspondant existe exclusivement. Cela implique que le spectre observé dans un tel pixel correspond au spectre du membre final. Les algorithmes les plus utilisés sont l'analyse des composants verticaux (VCA) [110], N-FINDR [156], l'indice de pureté des pixels (PPI) et le cône convexe d'angle maximal séquentiel (SMACC) [23]. L'analyse des composants parcimonie (SCA) est l'une des principales approches de la séparation aveugle de sources (BSS). Elle détecte des zones à source unique en exploitant les propriétés de parcimonie des sources dans différents domaines de représentation [35].

I.2 Motivation et Objectifs de la Thèse

Une fois que le cube de données est construit en utilisant le dématriçage, toutes les applications de post-traitement, par exemple, le démélange ou la classification, peuvent être appliquées pour traduire les données spectrales brutes en informations interprétables et exploitables comme le montre la Fig. 1.

Toutefois, comme l'ont indiqué Tsagkatakis *et al.* [143], appliquer directement la classification sur les images SSI après le dématriçage conduit souvent à des performances insatisfaisantes. Par ailleurs, la plupart des méthodes de démélange existantes sont conçues pour fonctionner sur des cubes de données pleinement reconstruits et ne prennent pas en compte les entrées manquantes inhérentes aux scénarios d'imagerie "snapshot". De plus, dans le contexte d'un problème d'étalonnage *in situ* de capteurs mobiles, il a été démontré qu'une approche combinant complétion et factorisation matricielle de faible rang est nettement plus performante qu'un processus séquentiel en deux étapes de complétion matricielle suivie de factorisation [42]. Ces résultats soulignent les limites du pipeline séquentiel proposé par les fabricants de caméra snapshot.

Reconnaissant cette problématique, notre approche proposée vise à effectuer le dématriçage et le démélange de manière conjointe, en se basant sur la factorisation de matrices de faible rang et la complétion. Cette méthode intégrée a pour but non seulement d'améliorer les résultats du démélange mais également de maintenir une performance optimale du dématriçage.

En nous appuyant sur ce cadre, nous analysons en outre la chaîne de traitement pour restaurer le cube de données HSI introduit par le fabricant de la caméra et explorons l'utilisation de la matrice de réponse harmonique des filtres Fabry-Perot (FPf) pour un dématriçage une déconvolution et un démélange conjoints. Notre objectif est de démontrer que l'intégration de ces techniques conduit non seulement à des résultats améliorés par rapport aux étapes séparées de dématriçage, démélange, et

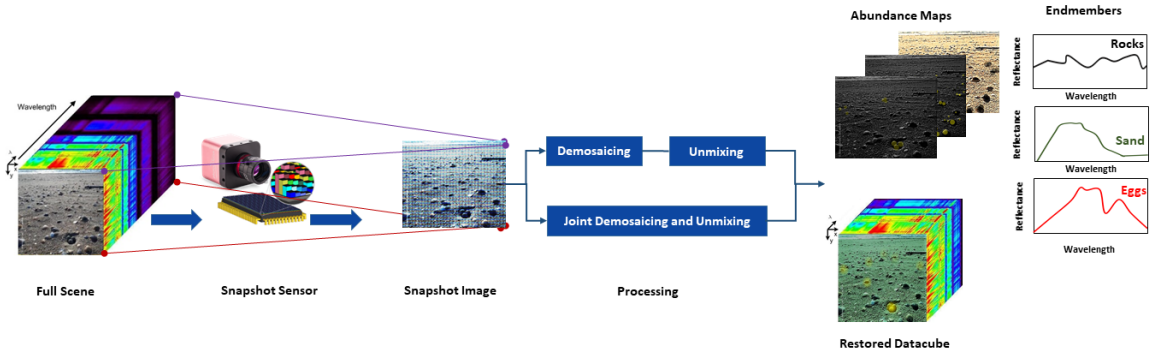


Figure 1 – Vue d'ensemble de l'acquisition et du traitement sur une scène contenant trois membres finaux (sable, roches et œufs). L'image SSI peut être traitée soit par une approche en deux étapes, soit en appliquant conjointement le démélange et le dématriçage pour restaurer le cube de données, extraire les membres finaux, et trouver les cartes d'abondance.

correction spectrale – communément appelée déconvolution – mais simplifie également l'ensemble du processus de traitement en intégrant la correction spectrale au sein du cadre conjoint.

I.3 Structure de la thèse

I.3.1 Chapitre 1 : Imagerie spectrale snapshot

L'imagerie spectrale snapshot (SSI) représente une avancée innovante dans le domaine de l'imagerie hyperspectrale (HSI), permettant la capture de l'information spectrale complète en une seule exposition. Cette technologie se distingue des méthodes traditionnelles qui collectent les données de manière séquentielle. L'introduction des technologies SSI, telles que la spectrométrie de champ intégré (IFS), le fractionnement de faisceau multispectral (MSBS) et l'imagerie utilisant des filtres de Fabry-Perot (FPfs), parmi d'autres, a ouvert de nouvelles perspectives dans divers domaines d'application allant de la surveillance environnementale et l'agriculture de précision jusqu'au diagnostic médical et la sécurité alimentaire.

Dans ce chapitre, nous présentons une étude complète de l'imagerie spectrale snapshot. Nous commençons par un aperçu de l'HSI et ses vastes applications à travers divers domaines. Nous explorons les complexités et les avancées des méthodes d'acquisition de données HSI, en mettant en lumière la transition des techniques de balayage traditionnelles vers les technologies d'imagerie snapshot. Ensuite, nous examinons en profondeur diverses méthodologies d'imagerie snapshot.

Enfin, puisque cette thèse se concentrera sur les caméras snapshot utilisant des filtres de Fabry-Perot, nous approfondissons les étapes cruciales de la correction spectrale, reconnaissant que chaque caméra nécessite des étapes d'étalonnage préalables et postérieures à l'acquisition de données, afin de garantir une performance optimale.

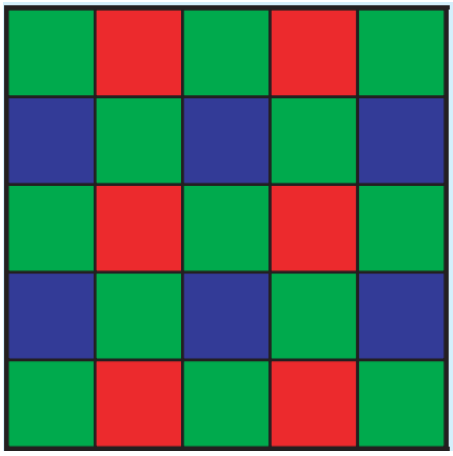


Figure 2 – Disposition de l'arrangement des filtres couleur de Bayer

I.3.2 Chapitre 2 : Dématriçage des images spectrales snapshot

Le dématriçage est un problème bien établi dans la communauté de l'imagerie. Il est crucial pour les images couleur RGB, compte tenu de la présence de filtres structurés de Bayer. Dans une matrice typique de filtres de Bayer, la moitié des pixels sont verts, tandis que le rouge et le bleu représentent un quart des pixels totaux, comme le montre la Fig 2. Cette disposition reflète l'accent mis par le motif de Bayer sur le vert, reconnaissant son importance dans la perception des couleurs par l'humain. Chaque pixel de cette mosaïque capture uniquement l'une des trois couleurs, ce qui signifie que le capteur manque des deux tiers des informations de couleur pour cet emplacement particulier à tout moment donné.

Les algorithmes de dématriçage visent à reconstruire des images en couleur complète à partir de ces données incomplètes. Ces algorithmes fonctionnent en interpolant les informations de couleur manquantes pour chaque pixel. Par exemple, un pixel qui capte la lumière verte aura ses valeurs rouge et bleue estimées en fonction des pixels adjacents qui capturent ces couleurs. Le défi réside dans la prédiction précise de ces valeurs manquantes pour créer une image homogène et réaliste.

La complexité du processus de dématriçage varie selon l'algorithme utilisé. Les méthodes simples pourraient employer une interpolation bilinéaire ou bicubique [65], qui calcule les valeurs de couleur manquantes sur la base d'une moyenne pondérée des pixels voisins. Des techniques plus avancées, comme l'interpolation basée sur le gradient ou dirigée par les bords [65], sont conçues pour mieux préserver les bords nets et les détails fins, réduisant ainsi les artefacts. Ces algorithmes analysent les motifs de pixels environnants pour faire des estimations plus éclairées, en particulier autour des bords ou des zones de fort contraste, conduisant à une reconstruction plus précise et visuellement satisfaisante.

Dans le domaine de l'imagerie spectrale snapshot (SSI), également connue sous le nom de matrice de filtres multispectraux (MSFA), le dématriçage adopte un rôle unique et plus complexe. Contrairement à l'imagerie RGB conventionnelle, SSI/MSFA associe chaque pixel à une bande spectrale

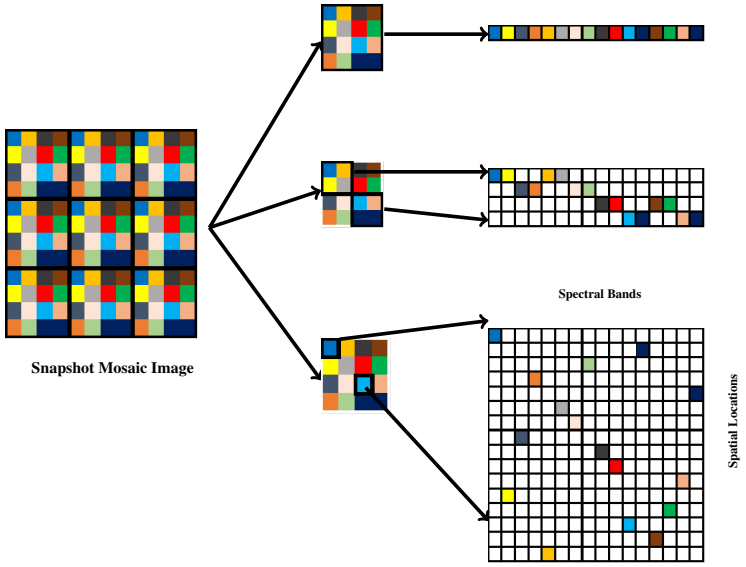


Figure 3 – Une approche typique consiste à regrouper ensemble des groupes de pixels, connus sous le nom de “super-pixels”, pour produire des images à résolution spatiale plus basse avec plus (milieu) ou toutes (en haut à droite) les bandes spectrales. Tandis que les approches de dématriçage visent à repeupler le cube complet tout en conservant la pleine résolution spatiale (en bas à droite). Adapté de [143].

spécifique, nécessitant une méthode différente pour la reconstruction d’image. Dans des scénarios comme le cas à 16 bandes, une technique courante implique la formation de “super-pixels”. Cette méthode, essentiellement une forme de sous-échantillonnage spatial, rassemble des groupes de pixels 4×4 en un seul super-pixel. Bien que cette approche simplifie le processus d’échantillonnage, elle entraîne une réduction notable de la résolution spatiale de jusqu’à 93,75%, comme illustré dans la Fig. 3 [143]. Ces super-pixels représentent une version condensée du tableau de pixels original, capturant une gamme spectrale plus complète au détriment des informations spatiales détaillées. Cependant, le dématriçage vise à générer l’ensemble du cube de données HS et à conserver la résolution spatiale. Par conséquent, diverses approches pour estimer les informations manquantes ont été proposées. Ces approches peuvent être classifiées en deux catégories, c’est-à-dire les méthodes “traditionnelles” et les stratégies basées sur l’apprentissage profond.

Les méthodes traditionnelles représentent un éventail de stratégies algorithmiques développées au fil des années pour relever le défi de la reconstruction de cubes de données dans le MSFA sans recourir aux architectures complexes de réseaux neuronaux qui caractérisent les techniques modernes d’apprentissage profond. Les méthodes traditionnelles de dématriçage sont ancrées dans diverses approches algorithmiques et heuristiques, chacune conçue pour reconstruire un cube de données HSI 3D à partir des informations limitées capturées par les pixels individuels dans un MSFA.

Les méthodes traditionnelles de dématriçage peuvent être globalement divisées en cinq catégories distinctes :

- Dématriçage basé sur l'interpolation.
- Dématriçage basé sur l'analyse par ondelettes.
- Dématriçage basé sur les arbres binaires.
- Dématriçage basé sur l'analyse pseudo-panchromatique.
- Dématriçage basé sur l'approximation de matrices de faible rang.

D'autre part, ces dernières années, l'apprentissage profond (DL) a influencé de nombreux domaines de la science et de la technologie, offrant des améliorations substantielles par rapport aux méthodes traditionnelles en termes de précision et d'efficacité. Dans le traitement d'images, les techniques d'apprentissage profond ont particulièrement transformé l'approche du dématriçage des images MSFA. Cette section explore diverses méthodes basées sur l'apprentissage profond développées pour relever les défis complexes associés au dématriçage des images MSFA.

En explorant ces méthodes avancées, nous pouvons les catégoriser en fonction de leurs architectures sous-jacentes – telles que les réseaux neuronaux convolutionnels (CNN) [13], les réseaux résiduels (ResNets) [72] ou ceux intégrant des mécanismes d'attention [146] – ou par leurs fonctionnalités spécifiques, comme le dématriçage et le débruitage conjoints ou la correction spectrale. Cependant, il est essentiel de noter que malgré ces distinctions, il existe un chevauchement significatif dans les fonctionnalités offertes par les différentes architectures. Par exemple, les CNN et les ResNets sont souvent utilisés avec des mécanismes d'attention pour améliorer les performances. Ce chevauchement rend difficile la séparation nette de ces méthodes en groupes distincts.

Étant donné ces complexités et la nature intégrée de nombreuses approches d'apprentissage profond, nous n'avons pas encore regroupé les méthodes strictement par architecture ou fonctionnalité dans cette discussion. Au lieu de cela, nous présentons un aperçu de chaque méthode, permettant une comparaison fluide et une meilleure appréciation de la manière dont différentes techniques peuvent être combinées ou adaptées pour répondre à des défis spécifiques de dématriçage.

I.3.3 Chapitre 3 : Post-traitement des images dématricées : techniques de démélange

Dans le domaine de l'imagerie spectrale snapshot, la construction d'un cube de données 3D à partir des images capturées n'est que le début de l'extraction d'informations significatives. Une fois qu'un cube de données complet a été dématriqué, une gamme de tâches de post-traitement peut être appliquée, telles que le démélange, la classification, la détection d'anomalies ou la détection de changements. Ces processus sont fondamentaux pour traduire les données spectrales brutes en informations interprétables.

Le démélange hyperspectral (HU) est le processus qui sépare les spectres de pixels d'une image hyperspectrale en une collection de signatures spectrales de matériaux observés (endmembers) et un ensemble d'abondances fractionnelles, comme illustré dans la Fig. 4. Les endmembers représentent

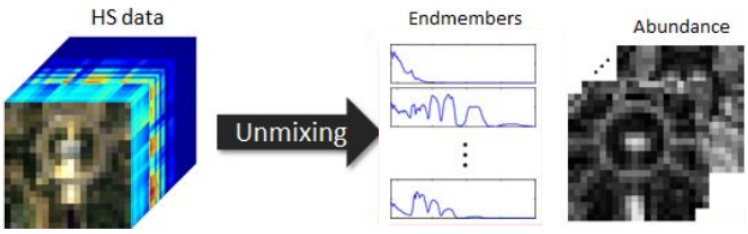


Figure 4 – Exemple de démélange hyperspectral (source : adapté de [99]).

les matériaux purs dans l'image, et les abondances à chaque pixel représentent la proportion de la présence de chaque endmember dans le pixel [23].

Ce chapitre explore les diverses techniques et les défis associés au démélange hyperspectral, en fournissant un bref aperçu des méthodes utilisées pour démêler les données hyperspectrales. Bien que des études et des classifications exhaustives des méthodes de démélange existent dans la littérature [23, 21, 47, 73, 26, 58], ce chapitre n'a pas pour but de revisiter de manière exhaustive ces méthodologies. Au lieu de cela, nous esquisserons brièvement une classification générale des techniques de démélange pour préparer le terrain à un examen ciblé de la factorisation en matrices non-négatives (NMF) et de l'analyse en composantes parcimonieuses (SCA), car ces deux approches sont particulièrement pertinentes pour nos cadres proposés.

I.3.4 Chapitre 4 : Méthodes localement de rang-un de démélange et dématriçage conjoints pour les images spectrales snapshot : un formalisme de complétion de matrices

Dans les chapitres précédents, nous avons introduit les bases de l'imagerie spectrale snapshot et le processus de dématriçage. Le dématriçage est une étape préliminaire cruciale qui reconstruit un cube de données 3D complet à partir des données brutes acquises par les caméras spectrales snapshot. Une fois ce cube de données construit, diverses applications de post-traitement, telles que le démélange ou la classification, peuvent être appliquées pour traduire les données spectrales brutes en informations interprétables et exploitables. Cependant, comme l'ont indiqué Tsagkatakis *et al.* [143], l'application directe de la classification sur les images SSI après le dématriçage conduit souvent à des performances insatisfaisantes. En revanche, la plupart des méthodes de démélange existantes sont conçues pour fonctionner sur des cubes de données entièrement reconstruits et ne tiennent pas compte des entrées manquantes inhérentes aux scénarios d'imagerie snapshot. De plus, dans le contexte d'étalonnage *in situ* de capteurs mobiles, il a été démontré qu'une approche combinée de complétion de matrices de faible rang et de factorisation est nettement plus efficace qu'un processus séquentiel en deux étapes de complétion de matrices suivi de factorisation [42]. Cette lacune méthodologique souligne une limitation cruciale dans les pratiques actuelles, où l'équilibre délicat entre le dématriçage et le démélange n'est pas correctement abordé.

Reconnaissant ce problème, notre approche proposée cherche à effectuer le dématriçage et le

démélange conjointement. Cette méthodologie intégrée vise non seulement à améliorer les résultats du démélange mais aussi à maintenir une performance optimale de dématriçage.

Dans ce chapitre[2], nos contributions au domaine de l'imagerie spectrale snapshot (SSI) sont détaillées comme suit :

- Nous nous concentrons sur l'établissement d'un cadre général pour le dématriçage des images, applicable à diverses matrices de filtres multispectraux (MSFA) [90]. Ce cadre à large portée pose les bases de nos méthodes complètes de « dématriçage » et de « démélange » des images hyperspectrales capturées par les caméras SSI.

- Notre investigation est motivée par l'hypothèse que la combinaison de la complétion de matrices de faible rang et de la factorisation est plus efficace qu'un processus en deux étapes impliquant ces composants séparément.

- Trois méthodes innovantes sont proposées pour le dématriçage et le démélange des images brutes SSI :

  – La première est une approche dite "naïve", directement dérivée de la factorisation de matrices non-négatives pondérée (WNMF).

  – Tout en reprenant le formalisme de l'approche naïve, les deuxième et troisième méthodes opèrent sous l'hypothèse de parcimonie des abondances au niveau des "patchs" du capteur. Un patch correspond à un motif de la MFSA appliquée au capteur. Plus particulièrement, ces approches supposent qu'il existe pour chaque endmember au moins un patch dominé par cet unique endmember. Ces méthodes suivent la même structure, à savoir (i) l'approximation des données de chaque patch par WNMF de rang-1, (ii) une métrique de confiance d'activité mono-source inspirée par la SCA, (iii) une phase d'extraction des endmembers, et (iv) une étape finale d'estimation des abondances.

    Ces méthodes diffèrent par leurs hypothèses et leur traitement des cas impliquant plusieurs endmembers.

    1. Une méthode suppose que tous les patchs de rang-1 sont purs, c'est à dire qu'ils sont dominés par un endmember. Cette hypothèse correspond au formalisme de la SCA et l'étape (iii) ci-dessus est réalisée par clustering des spectres estimés. L'approche proposée se nomme alors KPWNMF pour *K-means Patch-based Weighted Nonnegative Matrix Factorization*.
    2. Au contraire, l'autre méthode suppose qu'il peut aussi exister des patchs dans lesquels les abondances sont relativement constantes. Dans ce cas, l'approximation de rang-1 permet d'estimer des spectres purs ou des mélanges de ces spectres. L'étape (iii) est alors réalisée en appliquant une méthode de démélange fondée sur l'hypothèse de pixel pur (ou NMF séparable / convexe). La méthode proposée se nomme alors VPWNMF pour *VCA Patch-based Weighted Nonnegative Matrix Factorization*.

[2]Les méthodes présentées dans ce chapitre ont été publiées dans [6, 4, 3].

I.3.5 Chapitre 5 : Méthodes localement de rang-un de démélange et dématriçage conjoints pour les images spectrales snapshot : Un formalisme basé sur le filtrage

Dans le Chapitre 4, nous avons proposé des méthodes de dématriçage et de démélange conjoints pour les systèmes d'imagerie spectrale snapshot (SSI). En plus d'une approche naïve, dérivée directement de la factorisation de matrices non négatives pondérées (WNMF), nous avons proposé deux nouvelles méthodes de dématriçage : KPWNMF et VPWNMF. Ces méthodes incorporaient l'hypothèse de parcimonie des abondances au sein des patchs du capteur, où un seul endmember domine principalement chaque patch. Les méthodes employaient des calculs de WNMF de rang-1, une mesure de confiance d'activité mono-source spécifiquement proposée pour les données SSI, l'extraction des endmembers et l'estimation des abondances. Cependant, elles divergeaient dans le traitement des scénarios impliquant plusieurs endmembers, offrant des solutions distinctes pour aborder ces cas.

Dans ce chapitre[3], nous analysons davantage la chaîne de traitement pour restaurer le cube de données HSI – introduite précédemment en section 1.4 et illustrée dans la Fig. 5 – et explorons l'utilisation de la matrice de réponse harmonique des filtres Fabry-Perot (FPf) pour la déconvolution, le démélange et le dématriçage conjoints. Notre objectif est de démontrer que l'intégration de ces techniques conduit non seulement à des résultats améliorés par rapport aux étapes séparées de dématriçage, démélange et correction spectrale – communément appelée déconvolution – mais simplifie également la chaîne de traitement globale en intégrant la correction spectrale dans le cadre conjoint. De plus, le pipeline de traitement d'image actuel des images snapshot implique l'application d'une matrice de correction après le dématriçage afin de supprimer certaines harmoniques indésirables. En revanche, nos approches proposées éliminent le besoin d'appliquer la matrice de correction et initient le processus de déconvolution directement à partir de l'image brute SSI. Grâce à cette nouvelle méthodologie, nous visons à démontrer l'efficacité de nos approches pour obtenir des résultats de démélange et de dématriçage améliorés tout en simplifiant la chaîne de traitement globale.

Il est important de noter les différences entre les deux cadres. Premièrement, dans le Chapitre 4, nous établissons un cadre général pour le dématriçage des images. Cette approche est large en portée et potentiellement applicable à diverses matrices de filtres multispectraux (MSFA) [90]. Ce cadre pose les bases de notre discussion sur les méthodes conjointes pour effectuer le dématriçage et le démélange des images hyperspectrales capturées par des caméras SSI. Basée sur des constatations antérieures, notre investigation ici est motivée par l'hypothèse que l'incorporation conjointe de la complétion de matrices de faible rang et de la factorisation est plus efficace qu'une approche en deux étapes impliquant une complétion de matrices de faible rang suivie d'une factorisation de matrices.

En nous appuyant sur le cadre général établi dans le Chapitre 4, ce chapitre explore la chaîne complète fournie par le fabricant de la caméra, en se concentrant sur de nouvelles approches de démélange pour les images brutes SSI qui prennent en compte cette chaîne. Ce cadre est conçu spécifiquement pour les applications impliquant des filtres Fabry-Perot ou tout capteur MSFA où les filtres ne sont pas suffisamment sélectifs. La différenciation entre les deux cadres réside dans leur

[3]Les méthodes présentées dans ce chapitre ont été publiées dans [7, 2, 5].

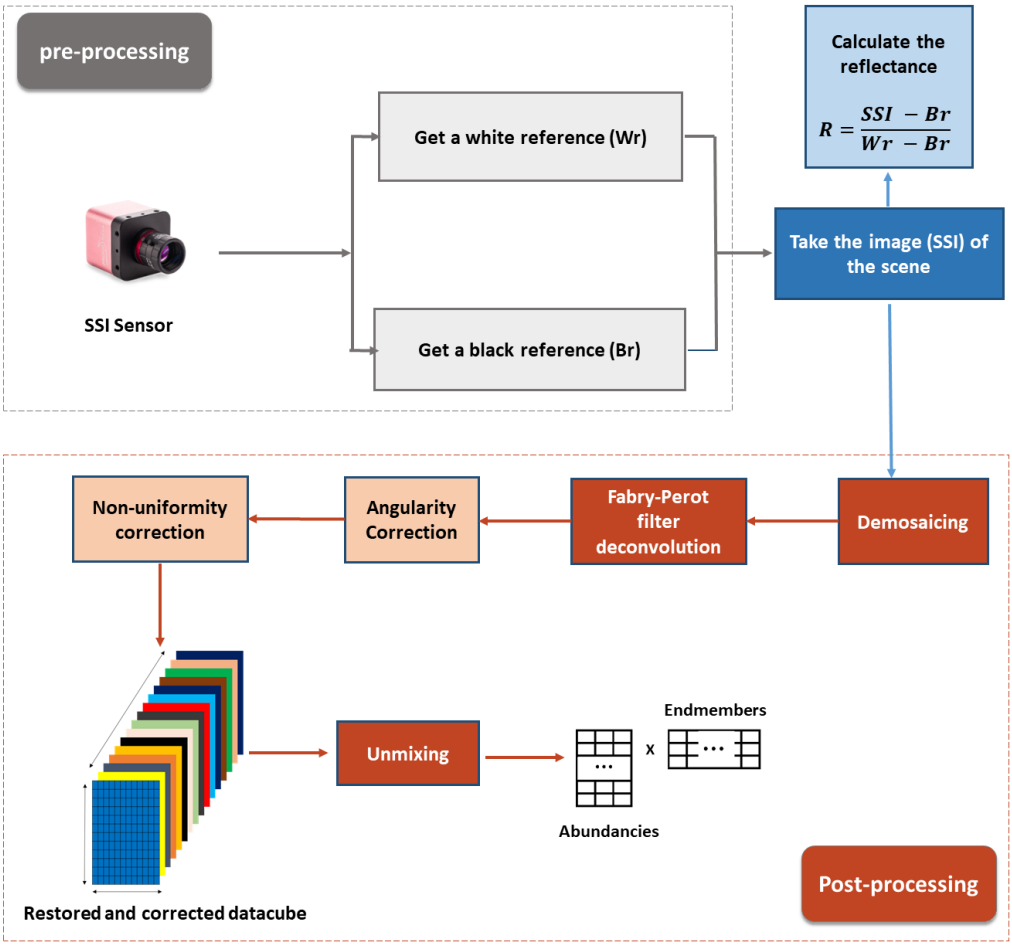


Figure 5 – Etapes de traitement requises par le fabricant de la caméra pour fournir des images SSI de qualité à démélanger.

focus et leur applicabilité : tandis que le premier jette les bases pour une approche générale, le second s'attaque aux scénarios complexes spécifiques souvent rencontrés dans la pratique.

I.3.6 Chapitre 6 : Validation expérimentale

Dans ce chapitre, nous évaluons les performances des cadres proposés pour le dématriçage et le démélange conjoints des images spectrales snapshot : les cadres basés sur la complétion de matrices et sur le filtrage. L'évaluation est réalisée à l'aide d'images synthétiques et réelles capturées par des caméras snapshot réelles. Nous examinons également les performances dans diverses conditions de bruit pour garantir la robustesse de nos méthodes. Nos résultats expérimentaux montrent que les méthodes proposées surpassent les approches en deux étapes, consistant d'abord à appliquer le dématriçage, puis le démélange.

I.3.7 Chapitre 7 : déconvolution pénalisée par une pondération liée à l'entropie

Dans ce chapitre, nous introduisons une nouvelle méthode de déconvolution spectrale pour les caméras hyperspectrales basées sur Fabry-Perot. Le fabricant de la caméra a proposé une matrice de correction indépendante de la scène pour éliminer les harmoniques présentes dans l'image SSI. Par conséquent, nous exploitons la réponse du filtre Fabry-Perot – comme nous l'avons fait dans le Chapitre. 5 – et proposons une nouvelle méthode de correction et de calibration spectrale dépendante de la scène[4]. Notre approche utilise la déconvolution avec régularisation de Tikhonov pondérée par l'entropie des harmoniques de Fabry-Perot pour éliminer les artefacts générés et restaurer les spectres originaux. Elle s'adapte aux caractéristiques uniques de la scène, réduisant les harmoniques et améliorant la qualité des données hyperspectrales. Les expériences avec des données synthétiques et des images réelles provenant de deux capteurs IMEC – i.e., *linescan wedge* et *snapshot mosaique* – que notre méthode surpasse la solution du fabricant en éliminant les harmoniques et en améliorant les informations spectrales.

I.3.8 Conclusion générale et perspectives

I.3.8.1 Conclusion

Cette thèse a exploré de manière exhaustive les technologies d'imagerie spectrale snapshot (SSI), en se concentrant sur l'intégration conjointe du dématriçage et du démélange directement à partir des données brutes plutôt que de les traiter comme des étapes séparées. Nous avons commencé au Chapitre 1 par une compréhension fondamentale de la SSI en tant qu'approche avancée dans l'imagerie hyperspectrale (HSI), en soulignant sa capacité à capturer l'information spectrale complète en une seule exposition. Le chapitre a également introduit les processus de calibration essentiels pour optimiser la performance de la technologie SSI basée sur Fabry-Perot, en soulignant leur importance pour produire des données spectrales fiables.

[4]Cette méthode est soumise à IEEE Sensors Letters [1].

Le Chapitre 2 a exploré de manière approfondie les diverses méthodologies appliquées au dématriçage des images spectrales snapshot, en passant en revue les méthodes traditionnelles et celles basées sur l'apprentissage profond. Nous avons analysé de manière critique leurs capacités et leurs limitations, en notant particulièrement les défis posés par les méthodes d'apprentissage profond dans des scénarios avec des données d'entraînement limitées et très variables.

Le Chapitre 3 a exploré diverses méthodologies de démélange, en les catégorisant en méthodes *géométriques, statistiques, basées sur la régression parcimonieuse et basées sur l'apprentissage profond*. Nous nous sommes concentrés sur la factorisation de matrices non négatives (NMF) et l'analyse de composants parcimonieux (SCA), car ces approches sont particulièrement pertinentes pour nos cadres proposés. Nous avons également identifié trois principales catégories de NMF : *NMF contrainte, structurée et généralisée*, chacune dédiée à relever des défis spécifiques de l'imagerie hyperspectrale.

Le Chapitre 4 a introduit les méthodes que nous avons proposé pour effectuer conjointement le dématriçage et le démélange. Nous avons proposé deux nouvelles approches en plus de la méthode naïve dérivée de la WNMF (Weighted Nonnegative Matrix Factorization). La première, KPWNMF (K-means Patch-based Weighted Nonnegative Matrix Factorization), suppose que les abondances sont parcimonieuses dans quelques patchs à trouver, de sorte qu'un endmember domine chacun de ces patchs. Une telle hypothèse est similaire à la SCA, sauf que nous considérons des données partiellement observées. En revanche, la seconde, VPWNMF (VCA Patch-based Weighted Nonnegative Matrix Factorization), assouplit l'hypothèse de parcimonie nécessaire dans la première. Ces méthodes ont été conçues pour améliorer le traitement des images SSI et réduire les erreurs associées aux processus traditionnels en deux étapes.

Le Chapitre 5 s'est concentré sur la réponse du filtre Fabry-Perot pour effectuer le dématriçage et le démélange conjoints. Plus précisément, nous avons réalisé la déconvolution, le dématriçage et le démélange conjoints, où nous avons intégré la déconvolution directement dans le processus de démélange. Cette intégration utilise efficacement les réponses harmoniques des filtres Fabry-Perot, améliorant la précision et l'efficacité du traitement des données spectrales. Nous avons proposé deux nouvelles méthodes pour cet objectif : FPVCA (Filter Patch-based Vertex Component Analysis) et FPKmeans (Filter Patch-based Kmeans).

Nos évaluations expérimentales approfondies, discutées dans le Chapitre 6, ont démontré la performance supérieure de nos méthodes proposées par rapport aux techniques traditionnelles, notamment dans leur capacité à traiter efficacement les données brutes SSI.

Enfin, le Chapitre 7 a introduit SDS-Cor, une méthode de correction spectrale dépendante de la scène qui représente une avancée significative par rapport aux matrices de correction traditionnelles indépendantes de la scène. SDS-Cor s'adapte aux caractéristiques spécifiques de la scène en utilisant une régularisation de Tikhonov pondérée par l'entropie. Nous avons évalué SDS-Cor et démontré qu'il réduisait significativement les distorsions harmoniques et améliorait la précision spectrale, surpassant la méthode de correction du fabricant, en particulier dans des environnements bruyants.

I.3.8.2 Perspectives

Temps de Calcul et Traitement Vidéo Hyperspectral

Les caméras SSI peuvent capturer des vidéos avec un taux de trame élevé et sont utiles dans des cas où le mouvement est imprévisible et où la caméra/l'objet se déplace en 2D ou en 3D, comme dans les applications robotiques ou le tri des voies [79]. Cependant, une limitation notable de nos cadres actuels – les cadres basés sur la complétion et sur le filtrage – est leur demande computationnelle. Le temps nécessaire pour traiter chaque trame est trop élevé pour les applications d'analyse vidéo en temps réel. Pour relever ce défi, les travaux futurs se concentreront sur l'amélioration de l'efficacité computationnelle de ces cadres. Une approche prometteuse est l'adoption de techniques d'apprentissage compressé, qui peuvent réduire significativement le volume de données à traiter en extrayant et utilisant des représentations compressées des données.

Variabilité Spectrale

L'un des principaux défis identifiés lors de notre exploration des technologies d'imagerie spectrale snapshot (SSI) est la variabilité spectrale, qui impacte significativement les performances de nos cadres proposés. La variabilité spectrale fait référence aux différences dans les signatures spectrales des mêmes matériaux sous des conditions d'éclairage, des angles de vue et des facteurs environnementaux variés. Ces variations conduisent à des inexactitudes dans les tâches de post-traitement. Les deux cadres peuvent être améliorés pour gérer la variabilité spectrale. Cela implique de développer l'étape d'estimation des endmembers en incorporant des modèles, tels que les extensions multicouches/profondes de la NMF [130, 45], qui détectent dynamiquement la variabilité des endmembers.

Variabilité des Filtres Fabry-Perot

La matrice de réponse qui enregistre la réponse réelle de la bande pour chaque FPf peut être sujette au bruit en raison des erreurs potentielles dans son estimation. En tant que composant crucial des méthodes proposées FPVCA, FPKmeans et SDS-Cor, elle peut affecter leurs performances. Par conséquent, l'effet de la matrice de réponse bruitée doit être étudié en profondeur et le cadre proposé doit être mis à jour pour être robuste à ce bruit.

Intégration de la Correction d'Angularity

La correction d'angularité, introduite par Goossens *et al.* [62], ajuste la variabilité de l'angle de la lumière. Notre méthode proposée, basée sur le filtrage et la méthode SDS-Cor, peut être étendue pour traiter cette question. Cela améliorera la généralisabilité de nos méthodes dans divers scénarios pratiques.

Intégration de l'Apprentissage Profond

Les approches d'apprentissage profond démontrent généralement des performances supérieures par rapport aux méthodes traditionnelles de dématriçage, notamment dans le traitement des données de haute dimension et la préservation des détails et de la précision des couleurs. Malgré ces avancées, les méthodes d'apprentissage profond rencontrent des défis significatifs, notamment en termes de généralisation et de besoins en données. Elles nécessitent souvent de grandes quantités de données d'entraînement annotées pour être performantes. De plus, ces modèles peuvent avoir du mal à généraliser à de nouvelles données ou à des configurations MSFA différentes de celles incluses dans l'ensemble d'entraînement. Pour relever ces défis, les recherches futures sur l'apprentissage profond pour le dématriçage MSFA pourraient se concentrer sur le développement de modèles plus robustes nécessitant moins de données d'entraînement et capables de mieux généraliser à travers différents motifs MSFA. Des techniques telles que l'apprentissage par transfert [77], l'apprentissage par few-shot [139], et la génération de données synthétiques [50] pourraient jouer des rôles significatifs dans la réalisation de ces objectifs. De plus, l'intégration de paradigmes d'apprentissage non supervisé ou semi-supervisé pourrait réduire la dépendance aux grands ensembles de données annotées et améliorer la capacité du modèle à généraliser à de nouvelles conditions.

List of the Author's Publications and Communications During the Ph.D Thesis

Submitted Papers

1. **K. Abbas**, P. Chatelain, M. Puigt, G. Delmaire, G. Roussel, *Fabry-Perot Spectral Deconvolution with Entropy-weighted Penalization*, IEEE Sensors Letters, vol. 8, no. 9, pp. 1-4, Sept. 2024.

In Peer-reviewed Journals

1. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Locally-Rank-One-Based Joint Unmixing and Demosaicing Methods for Snapshot Spectral Images. Part II: a Filtering-Based Framework*, IEEE Trans. Computational Imaging 10 (2024), pp. 806-817.

2. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Locally-Rank-One-Based Joint Unmixing and Demosaicing Methods for Snapshot Spectral Images. Part I: a Matrix-Completion Framework*, IEEE Trans. Computational Imaging 10 (2024), pp. 848-862.

In Proceedings of Peer-reviewed International Conferences

1. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Joint unmixing and demosaicing methods for snapshot spectral images*, in Proc. IEEE ICASSP'23, Rhodes, Greece, June 2023.

2. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Filtering-based endmember identification method for snapshot spectral images*, in Proc. of IEEE WHISPERS'22, Roma, Italy, September 2022. Received the José Bioucas Dias WHISPERS paper award

In Proceedings of Peer-reviewed National Conferences

1. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Méthodes parcimonieuses de déconvolution et démélange pour les images multispectrales «snapshot»*, in Actes du XXIXème colloque GRETSI, Grenoble, France, Août-Septembre 2023

2. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Méthode de démélange et dématriçage conjoints fondée sur la complétion de rang un pour les images multispectrales «snapshot»*, in Actes du XXVIIIème colloque GRETSI, Nancy, France, Septembre 2022

In International Conferences without Proceedings

1. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Filtering-based endmember estimation from snapshot spectral images*, in 2nd Workshop on Low-Rank Models and Applications (LRMA'22), Mons, Belgium, September 2022.

In National Conferences without Proceedings

1. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Méthodes parcimonieuses de déconvolution et démélange conjoints pour les images multispectrales "snapshot"*, Journée du LISIC, 12 Juillet 2023.

2. **K. Abbas**, Ma thèse en 180 s, Journée du pôle MTE de l'ULCO, Dunkerque, 12 Juillet 2022.

3. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Rank-One Clustering-Based Joint Unmixing and Demosaicing Method for Snapshot Spectral Images*, Journée du pôle MTE de l'ULCO, Dunkerque, 12 Juillet 2022.

4. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Rank-One Clustering-Based Joint Unmixing and Demosaicing Method for Snapshot Spectral Images*, Journée du LISIC, 11 Juillet 2022.

5. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *A dedicated joint unmixing and demosaicing method for snapshot spectral images*, Journée Régionale des Doctorants en Automatique, Douai, France, 14 juin 2022.

6. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Joint Demosaicing and Unmixing in Snapshot Spectral Images*, GdR ISIS et GdR ONDES, journées imagerie non-conventionnelle, Paris, 17 mars 2022.

7. **K. Abbas**, M. Puigt, G. Delmaire, G. Roussel, *Joint Demosaicing and Unmixing in Hyperspectral Videos*, Journée Régionale des Doctorants en Automatique, Compiègne, France, 19 novembre 2021.

General Introduction

II.1 General Framework

Hyperspectral Imaging (HSI) is an analytical technique based on spectroscopy, which involves capturing hundreds of images at various wavelengths for the same spatial region. The real power of HSI lies in its ability to detect, identify, and quantify materials and phenomena across diverse applications. From precise monitoring of agricultural health and environmental surveillance [87], to revolutionary applications in medical diagnostics and food safety [48], HSI provides a critical tool for research and industry.

A fundamental challenge HSI sensors must address is acquiring three-dimensional HSI data, encompassing two spatial and one spectral dimension, using a single detector, such as a 1D-array or 2D-plane detector. Thus, different strategies in HSI acquisition designs have emerged from the disparity between detector demands and available dimensionality, leading to wavelength-scanning, point-scanning, line-scanning, and snapshot acquisition methods [101, 151]. Nonetheless, a common characteristic across all scenarios is the need for repeated scanning of the scene and the acquisition of numerous exposures (frames) to compile the full spatio-spectral resolution data cube, coupled with considerations regarding the cost and size of the camera.

A new generation of HSI imaging architectures, known as Snapshot Spectral Imaging (SSI), has been introduced to address the challenges mentioned earlier [54]. SSI enables the efficient acquisition of the spatio-spectral content of dynamic scenes using miniaturized platforms and can acquire the complete cube from a single or a few exposures. To accomplish this objective, SSI architectures associate each spatial pixel with a specific spectral band, hence introducing a crucial trade-off between spatial and spectral resolution [143]. These cameras can capture videos with a high frame rate and are helpful in cases where motion is unpredictable, and the camera or object is moving in 2D or 3D, such as in robotic applications or lane sorting [79]. Among the recent strategies that have emerged, SSI cameras utilizing Fabry-Perot filters (FPf) [53] systems only produce a 2-D image derived from the 3-D hyperspectral data. Then, a post-processing technique called "demosaicing" is required to estimate the complete hyperspectral data cube. Various approaches have been proposed to perform demosaicing. These approaches can be classified into two categories, including "traditional" methods [27, 105, 108, 107, 22, 143] and deep-learning-based strategies [39, 66, 116, 46, 178, 94].

After the 3-D hyperspectral image has been reconstructed from the 2-D SSI data, we can employ any post-processing technique of choice. In particular, the unmixing process involves extracting the

spectral signatures of all endmembers present in an observed scene. Specifically, popular unmixing methods operate under the assumption that for each endmember, there is at least one spatial pixel where the corresponding material exclusively exists. This implies that the observed spectrum in such a pixel matches the endmember spectrum. The most widely used algorithms are the vertex component analysis (VCA) [110], N-FINDR [156], the pixel purity index (PPI), and the sequential maximum angle convex cone (SMACC) [23]. Sparse component analysis (SCA) is one of the main approaches to Blind Source Separation (BSS). It detects single-source zones by exploiting the source sparsity properties in different representation domains [35].

II.2 Thesis Motivation and Objectives

Once the datacube is constructed using demosaicing, any post-processing applications, e.g., unmixing or classification, can be applied to translate the raw spectral data into interpretable and actionable insights as shown in Fig. 6. However, as Tsagkatakis *et al.* [143] have stated that applying classification directly on SSI images post-demosaicing often leads to unsatisfactory performance. Contrastingly, most existing unmixing methods are designed to work on fully reconstructed data cubes and do not account for the inherent missing entries typical in snapshot imaging scenarios. Furthermore, in the context of *in-situ* mobile sensor calibration, it has been demonstrated that a combined approach of low-rank matrix completion and factorization is significantly more efficient than a sequential two-stage process of matrix completion followed by factorization [42]. This gap in methodology underscores a crucial limitation in current practices, where the intricate balance between demosaicing and unmixing is not adequately addressed.

Recognizing this issue, our proposed approach seeks to perform demosaicing and unmixing jointly based on low-rank matrix factorization and completion framework. This integrated methodology aims not only to enhance unmixing results but also to maintain optimal demosaicing performance.

Building upon this framework, we further analyze the pipeline to restore the HSI data cube—introduced by the camera manufacturer—and explore the utilization of the harmonic response matrix of Fabry-Perot filters (FPf) for joint demosaicing, deconvolution and unmixing. Our objective is to demonstrate that integrating these techniques not only leads to enhanced outcomes compared to the separate stages of demosaicing, unmixing, and spectral correction—commonly referred to as deconvolution—but also simplifies the overall processing pipeline by embedding the spectral correction within the joint framework.

II.3 Thesis Structure

This Ph.D thesis is structured as follows:

Chapter 1 lays the foundation for the thesis by providing an introduction to Snapshot Spectral Imaging (SSI) as an advanced hyperspectral imaging technique. It details the ability of SSI to capture complete spectral information in a single exposure, which is very important for dynamic and real-time

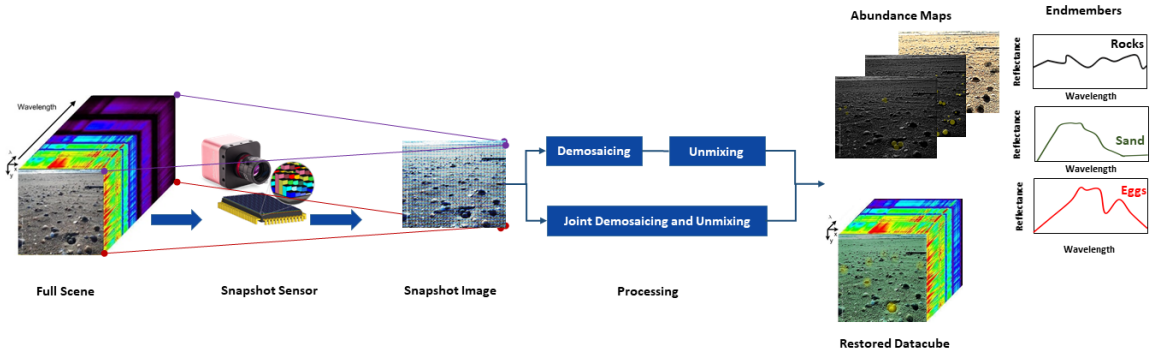


Figure 6 – Overview of acquisition and processing on a scene contains three endmemebers (sand, rocks, and eggs). The SSI image can be processed either by a two-stage approach or by applying unmixing and demosaicing jointly to restore the data cube, extract the endmembers, and find the abundance maps.

applications. This chapter also discusses the essential calibration processes that optimize the performance of Fabry-Perot-based SSI technologies, emphasizing the critical role of accurate calibration in ensuring the reliability of spectral data.

Chapter 2 delves into the diverse methodologies utilized in demosaicing snapshot spectral images. The chapter thoroughly reviews traditional and deep learning-based methods, critically analyzing their capabilities and limitations.

Chapter 3 focuses on the unmixing process, which is crucial for interpreting hyperspectral images. The chapter categorizes unmixing methodologies into *Geometrical, Statistical, Sparse Regression-based, and Deep Learning-based* methods. It highlights nonnegative matrix factorization (NMF) and sparse component analysis (SCA) as key techniques due to their relevance to the thesis proposed frameworks.

Chapter 4 introduces our novel methods for jointly performing demosaicing and unmixing, moving beyond traditional two-step processes. We propose two new approaches alongside a naive method derived from Weighted NMF (WNMF), namely KPWNMF and VPWNMF.

Chapter 5 explores using Fabry-Perot filter responses for joint demosaicing, deconvolution and unmixing. We introduce two innovative methods, FPVCA and FPKmeans, that integrate deconvolution into the unmixing process. This integration significantly enhances spectral data processing by utilizing the harmonic responses of Fabry-Perot filters.

The experimental evaluations are discussed in Chapter 6, where we demonstrate the superior performance of our proposed methods compared to traditional techniques. The experiments particularly focus on the methods ability to handle raw SSI data effectively.

Finally in Chapter 7 we present SDS-Cor, a novel spectral correction method that adapts to scene-specific characteristics through entropy-weighted Tikhonov regularization. We also provide a detailed assessment of SDS-Cor, showing its efficacy improving spectral accuracy, especially in challenging noisy environments.

Chapter 8 is the last chapter of the thesis. Here, we make a general conclusion, recapping methods, experiments, results, contributions, and perspectives for future work.

Part I

Literature Reviews

Chapter 1

Snapshot Spectral Imaging

"To photograph truthfully and effectively is to see beneath the surfaces and record the qualities of nature and humanity which live or are latent in all things."

— Ansel Adams

1.1 Introduction

Snapshot Spectral Imaging (SSI) represents an innovative advancement in the field of Hyperspectral Imaging (HSI), enabling the capture of complete spectral information in a single exposure. This technology differs from traditional methods that collect data in a sequence. The introduction of SSI technologies, such as Integrated Field Spectrometry (IFS), Multispectral Beamsplitting (MSBS), and Fabry-Perot-based imaging (FPfs), among others, has opened new views in various application domains ranging from environmental monitoring and precision agriculture to medical diagnostics and food safety. In this chapter, we present a comprehensive study of snapshot spectral imaging. We begin with an overview of hyperspectral imaging and its vast applications across various fields. We delve into the complexities and advancements in HSI data acquisition methods, highlighting the transition from traditional scanning techniques to snapshot imaging technologies. Then, we do an in-depth examination of various snapshot imaging methodologies. Finally, since this thesis will focus on Fabry-Perot-based snapshot cameras, we delve into the critical steps of spectral correction, acknowledging that every camera requires comprehensive preprocessing and postprocessing calibration steps to ensure optimal performance.

1.2 Hyperspectral Imaging

Hyperspectral Imaging is an analytical technique based on spectroscopy, which involves capturing hundreds of images at various wavelengths for the same spatial region. Hyperspectral imaging records the complete light spectrum for every pixel within a scene, providing high-quality wavelength resolu-

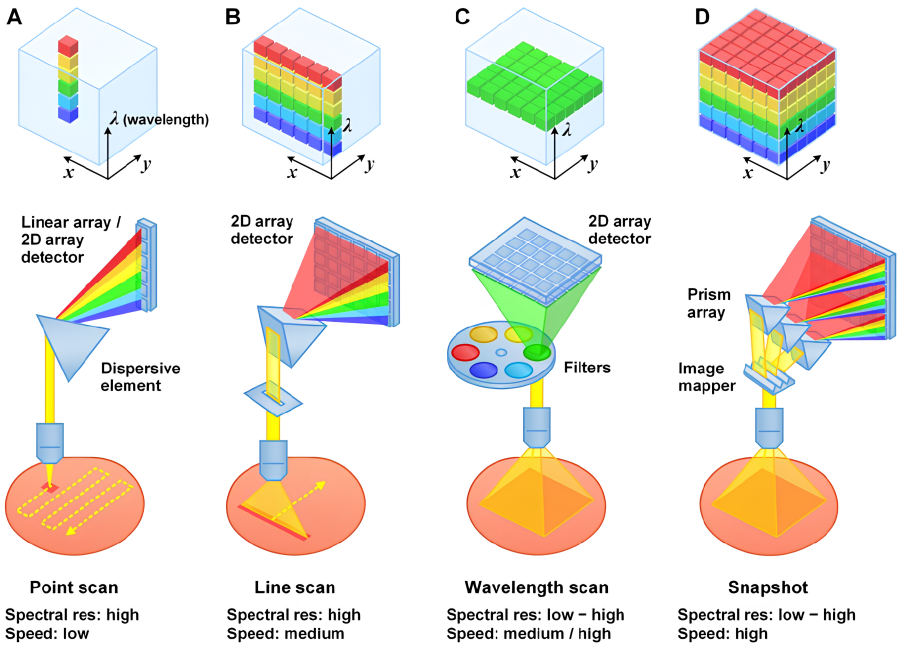


Figure 1.1 – Standard Scanning Methods (Source [154]).

tion, encompassing both the visible and near-infrared spectra. The gathered data forms a hyperspectral cube, where two dimensions define the spatial dimensions of the scene, and the third dimension represents its spectral content [23, 101, 87].

The real power of HSI lies in its ability to detect, identify, and quantify materials and phenomena across diverse applications. From the precise monitoring of agricultural health, environmental surveillance [87], and mineral exploration to its revolutionary applications in medical diagnostics and food safety [48], HSI provides a critical tool for research and industry. This technology enables the detection of subtle differences in material composition, changes in vegetation health, and even the early detection of diseases through tissue analysis, demonstrating versatility across the scientific spectrum [87, 82, 48].

A fundamental challenge that HSI sensors must address is how to acquire three-dimensional HSI data, encompassing two spatial and one spectral dimensions, using a single detector, such as a 1D-array or 2D-plane detector. Thus, different strategies in HSI acquisition designs have emerged from the disparity between detector demands and available dimensionality, leading to wavelength-scanning, point-scanning, line-scanning and snapshot acquisition methods [101, 151] as shown in Fig. 1.1.

The wavelength-scanning design—commonly referred to as area-scanning, focal plane scanning, or tunable filter—involves acquiring hyperspectral images by keeping the image field of view (FoV) fixed and capturing images at different wavelengths, one after the other. This approach provides an entire image at each wavelength and allows for efficient data collection and analysis. This method benefits applications requiring high spectral resolution over a large area, such as environmental monitoring and material characterization. However, its sequential nature can be a drawback for capturing

dynamic scenes or moving objects, as changes during the scanning process can introduce artifacts in the data [82, 121, 101].

On the other hand, the point-scanning methods—sometimes referred to as whisk-broom—involves acquiring hyperspectral data by measuring the spectrum of a single point and then moving the sample to measure another spectrum. This method is typically used in laboratory settings or for applications that require high spectral resolution and low spatial resolution. The primary limitation is its slow data acquisition rate, making it less suitable for large-area or fast-moving subjects [82, 121].

The third method, known as line-scanning or push-broom, involves acquiring spectral measurements from a line of samples simultaneously recorded by an array detector. This method is commonly used for remote sensing applications, where the platform is moving and linearly collecting data, such as from an airborne or satellite platform. The line-scanning approach provides high spatial resolution data and is particularly well-suited for mapping large areas of the earth surface. Although it can cover large areas efficiently, it requires the sensor or the target to be in constant, linear motion, which can be a limitation in stationary settings [82, 121, 101].

Finally, snapshot acquisition refers to the ability to acquire a complete hyperspectral image in a single instant, without the need for scanning. If traditional scanning techniques are like painting a picture stroke by stroke, snapshot imaging is like taking a photograph. The former builds the image slowly over time, while the latter captures it all at once, offering a more immediate and complete view. Snapshot systems offer significant advantages in scenarios where movement or variation in the scene must be captured with minimal distortion. Examples include aerial surveillance, medical imaging during surgeries, and dynamic scene analysis [106]. The challenge lies in the complexity of the optical and computational methods needed to reconstruct the hyperspectral data from the snapshot image, often requiring advanced algorithms and processing techniques [68]. In the following discussion, we will focus on the snapshot imaging technologies, delving deeper into their technical aspects and innovative solutions.

1.3 Snapshot Spectral Imaging

Snapshot spectral imaging aims to capture the complete spectral information of a scene in a single snapshot, without the need for mechanical scanning or sequential data acquisition. This concept finds its roots in the plenoptic function, a term first introduced to describe the vast array of light information accessible to an observe [10]. The plenoptic function is fully characterized by simultaneously accounting for seven key dimensions:

- The spatial dimensions are represented by two Cartesian coordinates (2-D).
- The spectral dimension denoted by the wavelength of light (1-D).
- Temporal changes over time (1-D).

- The multi-angular perspectives from which light is captured, including horizontal, vertical, and depth variations (3-D).

Expanding upon this foundational definition of the plenoptic function and applying it to the realm of snapshot hyperspectral imaging, a "snapshot" imaging spectrometer efficiently captures the spatial (2-D) and spectral (1-D) characteristics of the light within a scene in a single exposure. This results in a 3-D datacube, a fundamental component of hyperspectral data analysis [106, 68].

Several innovative technologies have made the move from theoretical concepts to practical applications in snapshot spectral imaging possible. Each method offers a unique approach to resolving the challenge of capturing an entire hyperspectral cube instantaneously without the need for traditional scanning mechanisms. The following discussion will explore these technologies.

1.3.1 Integrated Field Spectrometry

Integrated Field Spectrometry (IFS) represents a class of snapshot imaging technologies designed to capture spatial and spectral information without scanning simultaneously. IFS achieves this through innovative optical designs that effectively dissect the incoming light into multiple spectra, each corresponding to a distinct spatial element of the scene. This section explores three primary configurations of IFS, highlighting their unique approaches and applications [68].

- **Integrated Field Spectrometry with Faceted Mirrors (IFS-M)**

  IFS-M utilizes an array of small, precisely angled mirrors to direct different portions of the incoming light to separate areas of a detector. Each mirror is aligned to reflect light from a specific part of the scene, ensuring that the spectral information from each spatial region is captured simultaneously. This method is particularly effective for capturing high-resolution spectral data across a wide field of view, making it suitable for applications in astronomy and remote sensing where detailed spectral mapping of large areas is required [14, 68].

- **Integrated Field Spectrometry with Fiber arrays (IFS-F)**

  In the IFS-F configuration, a fiber optic array transports light from different spatial locations to a shared detector. Each fiber acts as a light guide, preserving spectral information from discrete points across the scene during transmission. This setup allows for a flexible arrangement of the input aperture, accommodating various shapes and sizes of scenes. IFS-F is highly adaptable and can be used in laboratory settings and field applications, including environmental monitoring and precision agriculture, where detailed analysis of spatially heterogeneous samples is necessary [117, 68].

- **Integrated Field Spectrometry with Lenslet arrays (IFS-L)**

  IFS-L employs a lenslet array, where each lenslet focuses light from a specific portion of the scene onto the detector, effectively creating multiple small images side by side. The lenslet array efficiently captures spectral data across diverse samples, enabling detailed analysis of biological tissues and engineered materials [68].

1.3.2 Multispectral Beamsplitting (MSBS)

MSBS technology uses specialized optical elements, such as dichroic mirrors or prism arrays, designed to reflect or transmit light based on its wavelength. These elements are arranged to partition the incoming light into predetermined spectral bands, each directed toward a specific sensor or detector segment. This method ensures that multiple spectral images are captured in parallel, reducing the time needed for data acquisition. MSBS technology finds applications in fields where speed and spectral selectivity are crucial. Despite its advantages, MSBS faces challenges in terms of optical complexity and the need for precise calibration. The design and alignment of the beam-splitting components are critical for ensuring accurate spectral separation and imaging performance. Furthermore, requiring multiple detector arrays can increase the system cost and complexity [68].

1.3.3 Multiaperture Filtered Camera (MAFC)

MAFC technology employs an array of microfilters, each designed to transmit a specific spectral band. These filters are placed directly over the sensor array, with each filter corresponding to a unique set of sensor pixels. This configuration allows the camera to capture different spectral bands in parallel, with each aperture collecting light from the scene filtered at a distinct wavelength. The result is a composite image comprising multiple spectral channels, each representing a narrow spectrum segment. Due to its capacity for rapid, simultaneous acquisition of multispectral data, the MAFC finds applications in fields as diverse as precision agriculture, environmental monitoring, and medical imaging. Despite its benefits, the MAFC faces filter design and image processing challenges. Creating microfilters that accurately transmit specific spectral bands while maintaining high optical throughput is complex [102, 68, 138].

1.3.4 Computed Tomography Imaging Spectrometer (CTIS)

Computed Tomography Imaging Spectrometer (CTIS) employs advanced gratings to create overlapping spectral cube projections on a 2-D sensor [51]. CTIS is inspired by X-ray computed tomography (CT), a commercial technique with a rich history spanning over half a century. CTIS opts for spatial division, splitting the imaging sensor into an array of regions, each capturing a distinct projection of the target data cube from a particular angle [34]. Subsequently, complex algorithms, akin to those used in CT are applied to extract the spectral cube from these overlapping projections[67]. However, this approach comes with certain limitations:

1. **Complex Reconstruction Algorithms**: the reconstruction algorithms are complex, making real-time visualization impractical. These complexities in data processing can hinder the efficiency of CTIS.

2. **Limited Spectral Cube Resolution**: the spectral cubes generated using CTIS have restricted resolution. Depending on the specific system, the number of pixels in the unrolled cubes can

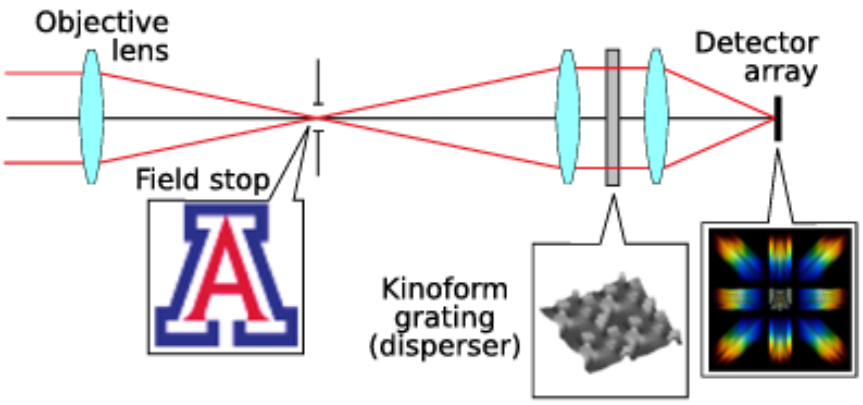


Figure 1.2 – The system layout for a computed tomography imaging spectrometer (Source [69]).

range from approximately 4.2% to $22 - 64\%$ of the sensor size. This limited resolution may affect the precision of spectral data [67, 84].

3. **Snapshot Limitations**: CTIS faces challenges as a snapshot instrument due to the inherently ill-posed nature of its acquisition process. This limitation is associated with limited-angle projection tomography, which affects the completeness of the spectral information captured in a single snapshot [67, 84].

1.3.5 Coded Aperture Snapshot Spectral Imager (CASSI)

CASSI is another snapshot imaging technology that depends on compressive sensing to capture the full datacube in a few measurements and in some cases, a single shot [17]. It tries to overcome the spatial versus spectral resolution multiplexing trade-off limitations by undersampling the scene spatially in each band and using compressive sensing to rebuild the full spatial resolution. It enhances the randomness in measurement to preserve more encoded information, facilitating a more efficient and faster acquisition compared to traditional Nyquist sampling methods [168]. The compressive sensing concept replaces the spatial versus spectral resolution trade-off with a signal-dependent spatial resolution versus image quality trade-off, which makes the resulting quality unpredictable and typically introduces spatial and spectral reconstruction artifacts [55].

1.3.6 Image Mapping Spectrometer (IMS)

IMS employs a micromachined remapping mirror positioned at the image plane of an objective lens. This mirror, akin to image slicers used in astronomy, consists of numerous small striped facets oriented in various directions (k different directions). Each facet redirects a column of pixels to one of the k sub-images, leaving the remaining $k - 1$ columns empty in each sub-image. Subsequently, a

prism is utilized to spectrally decompose each pixel within each column, filling the other $k-1$ empty columns in each subview with spectral data. Finally, collecting optics are used to refocus the light onto the sensor [52]. However, this approach has some limitations:

1. **Micromachined Mirror Precision**: Achieving high accuracy in micromachined mirrors is essential, and manufacturing defects can lead to issues such as crosstalk and reduced light throughput, affecting the quality of spectral data.

2. **Spectral Resolution Limitation**: The spectral resolution is limited by the number of tilt angles, which is constrained by the mirror accuracy and the numerical aperture (NA) of the collecting lens.

3. **Optical Complexity**: The complete optical architecture of IMS can be quite complex, potentially increasing the challenges associated with system design and maintenance [86].

1.3.7 Image-replicating Imaging Spectrometer (IRIS)

IRIS utilizes "Lyot filters" which are constructed using a combination of polarizers and waveplates [71]. These successive polarizers are responsible for implementing spectral filtering. In the case of IRIS, the polarizers take the form of "Wollaston prism polarizing beam splitters," serving a dual purpose of spectral filtration and image duplication. However, this system has its limitations, including reduced light throughput due to the polarizing filters and a restricted number of spectral bands. The use of Wollaston prisms to redirect the filtered subimages contributes to this limitation, impacting the system's spectral versatility.

1.3.8 Fabry-Perot-Based Sensors

The devices mentioned above share common limitations due to the complexity of their optical systems, involving custom and non-standard components. These complexities contribute to higher costs, reduced optical throughput, alignment difficulties, manufacturing challenges, reconstruction artifacts, and portability. These issues collectively impact the device efficiency, accuracy, and application versatility [57, 53].

Therefore, technological advances have recently enabled fast, miniaturized, and low-cost spectral cameras by monolithically integrating optical interference filters on top of standard CMOS image sensors. It operates by integrating a set of Fabry-Perot filters (FPfs) organized as tiled filters, mosaic, or any other custom design on top of a standard, off-the-shelf CMOS sensor [57].

The FPf is a well-developed and broadly utilized tool to direct and estimate wavelengths of light in systems with compact, integrated, and widely tunable spectral filtering capabilities. Theoretically, the FPf filter acts like sunglasses that only let through certain colors. These sensors selectively "see" the world through spectral "lenses," filtering out all but the desired wavelengths. Technically, the FPf is made of a transparent layer (or cavity) with two mirrors on each side of that layer, as shown in Fig. 1.3. The cavity length and the mirror reflectivity determine the filter's selected wavelength

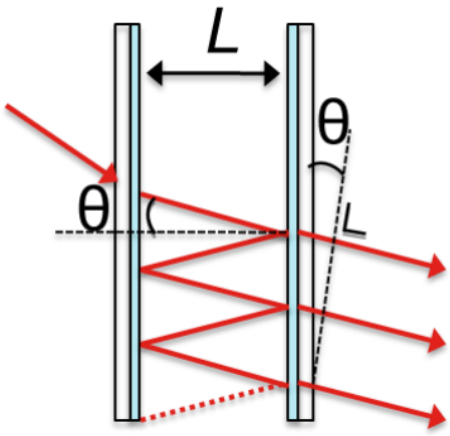


Figure 1.3 – Fabry-Perot Structure (Source IMEC)

and the spectral bandwidth (or full width at half maximum), respectively [57, 53, 78]. By using the monolithic integration of the filter on the sensor, the stray light in the system is heavily decreased, and the sensor sensitivity and speed are increased. In addition, using CMOS process technology heavily reduces the cost and improves the compactness of the hyperspectral camera as shown in Fig. 1.4.

However, FPfs are limited when applied in snapshot spectral imaging despite their precision and versatility. For instance, they can exhibit additional harmonic responses around each desired wavelength, leading to challenges such as spectral leakage and shifting. These imperfections require the comprehensive calibration steps outlined in the Section 1.4. The filter layout describes the pattern in which the filters are deposited on the sensor, and currently, there are two snapshot designs: **Snapshot Tiled, and Snapshot Mosaic** as shown in Fig. 1.5. There are variations for each layout, depending on the type, the active range of the sensor and the number of different filters in the layout [78].

1.3.8.1 Snapshot Tiled (SST)

The snapshot tiled filter layout employs an area design, arranging the k filters into a matrix of tiles with uniform width and height. Each filter corresponds to a band image, represented by one tile on the sensor active area. A comprehensive data cube requires the optical system to replicate the image across all k tiles. This approach mitigates the limitations of scanning-based systems by facilitating snapshot acquisition, capturing the entire multispectral data cube in a single shot. However, the snapshot-tiled filter layout necessitates optics with a more complex design compared to other Fabry-Perot-based sensors, incorporating advanced engineering to achieve its functionality and performance [56].

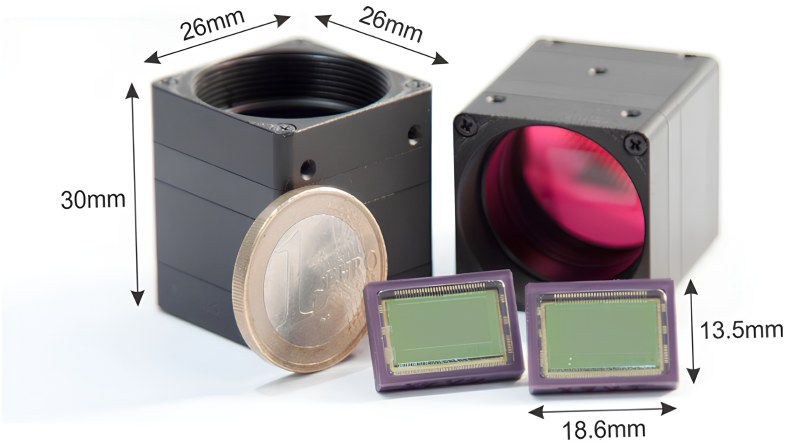


Figure 1.4 – Two packaged mosaic spectral snapshot sensors, alongside two tiny spectral cameras and a Euro coin for scale (Source [61])

(a) Snapshot tiled layout

(b) Snapshot mosaic layout

Figure 1.5 – Filter organization in Linescan, Snapshot Tiled, and Snapshot Mosaic layouts, respectively.

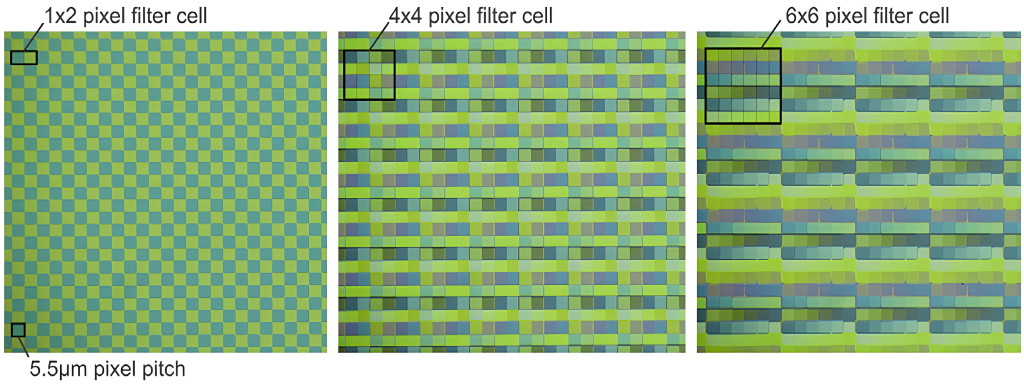


Figure 1.6 – Microscope images of snapshot mosaic imagers with different filter configurations. Each pixel (size 5.5×5.5 µm) within the cell is a different spectral filter. As can be seen, although current commercial designs use a 4×4 or 5×5 cell layout, other configurations are possible (Source [61]).

1.3.8.2 Snapshot Mosaic (SSI)

The snapshot mosaic filter layout has a per-pixel design. The basic FPf structure is extended into a set of filters by varying the cavity length for each pixel-level within a $\sqrt{k} \times \sqrt{k}$ filter cell ($\sqrt{k}$ could be 3, 4, or 5) [57, 53]. It is worth noting that the most commonly employed filter patterns are 5×5 and 4×4, but various other configurations are also feasible as shown in Fig. 1.6 [61]. This mosaic pattern, also known as a sensor patch, is replicated across the entire active area of the sensor surface. The organization of the filters in the mosaic pattern is shown in Fig. 1.7 and Fig. 1.8. Due to its design and contrary to the tiled filter, it can be used with a standard optical system.

These cameras can also capture videos with a high frame rate, and they are helpful in the case where the motion is unpredictable, and the camera/object is moving in 2-D or 3-D, such as in robotic applications or lane sorting [79].

Since SSI cameras associate each spatial pixel with a specific spectral band, thus introduce a critical trade-off between spatial and spectral resolution, so a post-processing technique is known as "demosaicing" must be applied to estimate the entire HS data cube and conserve the spatial resolution. Demosaicing will be discussed later in the next chapter.

In this work, we will specifically focus on snapshot mosaic cameras that are based on Fabry-Perot interferometers, delving into their unique capabilities and applications.

1.3.9 Concluding Insights

In this section, we discussed snapshot spectral imaging technologies in detail, from integrated field spectrometry imaging techniques to the precision and efficiency of Fabry-Perot-Based sensors. Each technology presents unique advantages and challenges. A synthesized comparison of these technologies, detailed in a Table 1.1, highlights their benefits and limitations, providing a clear overview for selecting the appropriate method based on specific imaging requirements and constraints.

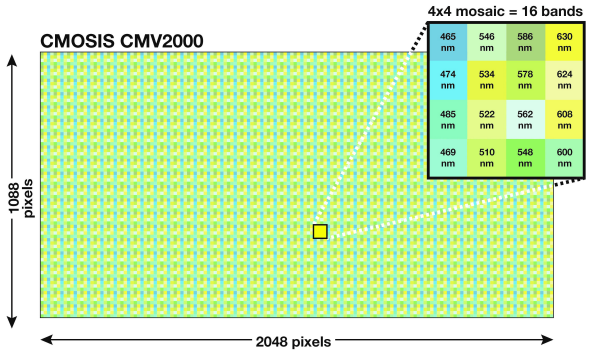


Figure 1.7 – Snapshot mosaic 4×4 filter layout (Source IMEC)

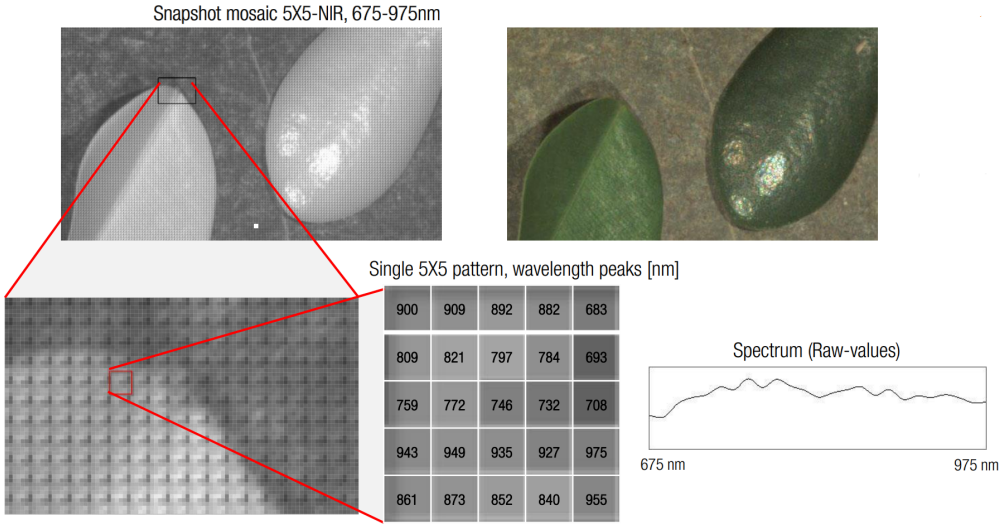


Figure 1.8 – Snapshot mosaic 5×5 filter layout with example image (Source [74])

Table 1.1 – Comparison of Snapshot Spectral Imaging Technologies

Technology	**Technique**	**Pros**	**Cons**
Integrated Field Spectrometry (IFS)	Utilizes arrays (mirrors, fibers, lenslets) to split incoming light into spectra	Efficient for capturing high-resolution data and adaptable for various scenes	Complexity in design and manufacturing challenges
Multispectral Beam-splitting (MSBS)	Employs dichroic mirrors or prism arrays for spectral partitioning	Rapid, simultaneous acquisition and high spectral selectivity	Optical complexity and precise calibration needed
Multiaperture Filtered Camera (MAFC)	Uses microfilters over sensor arrays for parallel spectral band capture	Real-time data acquisition and enhances image quality	Challenges in filter design and image processing
Computed Tomography Imaging Spectrometers (CTIS)	Advanced gratings create spectral cube projections on a 2D sensor	Does not require mechanical scanning	Complex reconstruction algorithms and limited spectral cube resolution
Coded Aperture Snapshot Spectral Imagers (CASSI)	Compressive sensing with coded apertures for efficient data capture	Fast acquisition	Unpredictable image quality and some reconstruction artifacts
Image Mapping Spectrometer (IMS)	Micromachined mirror for remapping and spectral decomposition	High spatial resolution	Precision in micromachined mirrors and optical complexity
Image-replicating Imaging Spectrometers (IRIS)	Utilizes 'Lyot filters' for spectral filtering and image duplication	Effective spectral filtering	Limited spectral bands and reduced light throughput due to filters
Fabry-Perot-Based Sensors (FPf)	Integrates optical interference filters on CMOS sensors for spectral imaging	Fast, miniaturized, cost-effective and high spatial resolution	Challenges with filter design, spectral leaking, shifting and harmonics

1.4 Calibration of Fabry-Perot Based Cameras

As we delve into the steps of calibrating Fabry-Perot-based cameras, it is nice to remember that, much like in music, "**Fine-tuning the instrument improves the melody.**" This proverb holds particularly true in spectral imaging, where the precision of the "instrument"—the spectral camera—directly influences the clarity and accuracy of the "melody" it produces, namely the spectral images. The following subsections will explore the essential calibration steps, each serving as a fine-tuning process to ensure the camera performs at its best.

1.4.1 Tracing the Problem's Roots

Miniaturized spectral cameras utilizing FPf technology exhibit certain imperfections due to physical constraints and the type of used lenses, necessitating calibration steps. In an ideal scenario, an FPf would efficiently transmit light within a narrow spectral range to the sensor while blocking light outside this range. However, in practical applications, FPfs tend to exhibit additional harmonic responses around each desired wavelength.

Figure 1.9 presents the spectral response of a Fabry-Perot filter employed in a 5×5 mosaic snapshot camera. The primary peak, labeled as the "central wavelength," represents the desired first-order response of the filter which corresponds to the specific wavelength the FPf is designed to transmit. We expect most of the light energy to be focused on this peak, indicating the filter optimal operational wavelength. However, the figure also shows additional unintended peaks. The one marked as a "second-order response" is an example of a harmonic response. These harmonics arise due to the constructive interference of light at these higher-order wavelengths within the FPf cavity. While ideally, the FPf should only transmit the central wavelength, physical imperfections, and design limitations can result in these additional responses. Another notable feature in the graph is the presence of "spectral leaking" and "crosstalk." Spectral leaking refers to the transmission of light outside the specified spectral range, which can degrade the quality of the image data. Crosstalk, indicated by smaller peaks within the active range of the sensor, refers to the mixing of signals from different wavelengths, which can lead to inaccurate color representation in the spectral image. Finally, "FWHM" stands for Full Width at Half Maximum, which measures the bandpass filter's selectivity. It indicates the width of the peak at half its maximum height and is a critical factor in determining the filter's resolving power. Furthermore, the cameras often encounter various challenges such as spectral shifting, and variations in CMOS imager performance. As a result, they require both pre-processing and post-processing steps to ensure the production of high-quality images [161, 155, 78, 148]. In the following subsections, we will delve into these calibration steps, highlighting their significance and their substantial impact on the fidelity of spectral imaging.

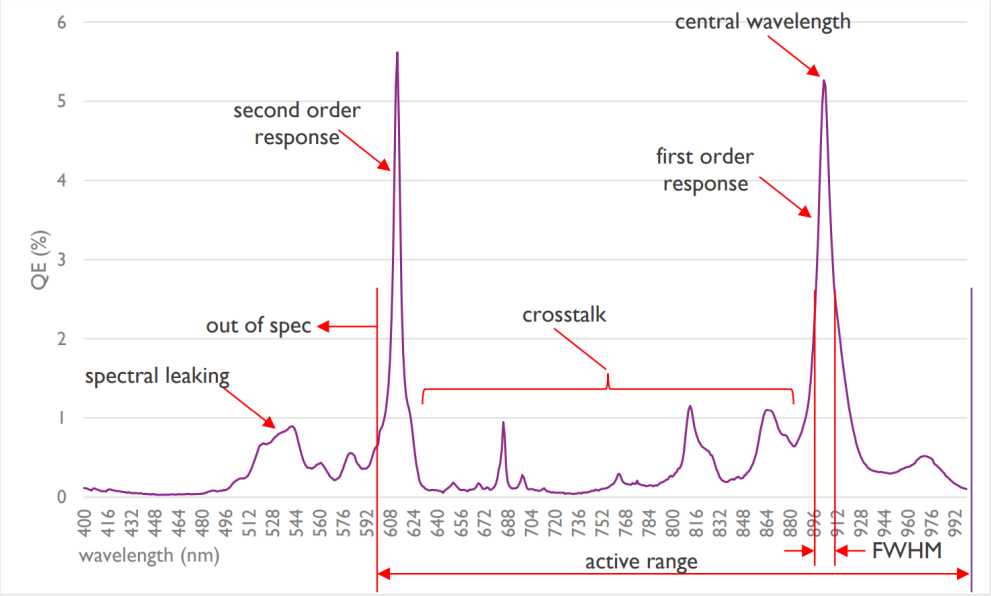


Figure 1.9 – Spectral response of Fabry-Perot filter of the 5×5 mosaic snapshot camera (Source IMEC).

1.4.2 Band pass filters

First, incident light experiences particular conditioning through the application of a specialized *rejection filter*, strategically placed before the sensor, blocking all wavelengths outside the sensor's active range. This step acts as a robust defense against spectral leaking and undesirable second-order responses [143, 78].

1.4.3 Dark Level Correction (Bias Correction)

The Dark Level Correction, also known as Bias Correction, is fundamental in pre-processing spectral images. The correction is crucial because it accounts for the camera sensor's noise floor—the base level of electrical disturbance or "noise" without illumination. We can effectively isolate and remove this noise by subtracting a dark reference image, which is captured with the lens cap on and no light entering the sensor. This step is essential for enhancing the signal-to-noise ratio of the final image, thereby improving the accuracy of subsequent spectral analysis.

1.4.4 White Balancing (Reflectance Calculation)

White Balancing, or Reflectance Calculation, is necessary to normalize the variations introduced by the sensor sensitivity and the illumination conditions. This process uses a white reference to standardize the image data, enabling the camera to represent colors and intensities accurately.

1.4.5 Spectral Correction

While the rejection filters remove most undesired wavelengths, some persist within the sensor active range, thus requiring *spectral correction*. The latter is performed by applying a fixed—i.e., scene-independent—correction matrix denoted $\mathbf{C}$ which is determined through a process of minimizing the difference between the actual band response, denoted as the response matrix $\mathbf{H}$, and the desired or ideal band response, represented as $\mathbf{H}^{\text{ideal}}$ [120, 161]. This minimization process is expressed mathematically as:

$$\min_{\mathbf{C}} \|\mathbf{H}^{\text{ideal}} - \mathbf{C} \cdot \mathbf{H}\|_{\mathbb{F}}^2, \tag{1.1}$$

where $\|\cdot\|_{\mathbb{F}}^2$ denotes the Frobenius norm. To apply this correction matrix effectively, it typically requires the calculation of reflectance using a white reference [78]. However, it is noteworthy that Vunckx and Charle [148] have introduced an approach that enables spectral correction directly on raw data that has already undergone dark-level correction.

The correction matrix $\mathbf{C}$ serves to transform physical wavelengths—as captured by the camera sensor—into virtual wavelengths, representing the spectrally-corrected data. The number of virtual bands is often less than that of the physical bands. Such discrepancies may arise from strong correlations among the responses of specific physical bands or from a lack of signal captured by one or more physical bands [78].

Furthermore, we note that the response matrix has crucial information about filter characteristics—including first and second-order responses—and the full width of response peaks at half the maximum of the peak (FWHM). The number of rows of the response matrix equals the number of filters on the sensor and the number of columns equals the number of measurement points used during calibration (i.e., 601 samples from 400-1000 nm) [78]. By multiplying this matrix with an irradiance spectrum, we simulate the sensor response, thereby facilitating evaluations under realistic conditions.

1.4.6 Angularity Correction

Angularity correction proposed by Goossens *et al.* [62] addresses the challenge posed by the angular sensitivity of thin-film interference filters. When integrated into camera sensors, these filters are designed to be sensitive to light at specific wavelengths. However, their spectral response can shift depending on the angle at which light strikes the filter. This phenomenon appeared from using non-telecentric lenses often found in compact, lightweight remote sensing applications.

As the light enters the FPf at angles other than the normal, the effective optical path length within the filter changes, leading to a shift in the spectral response toward shorter wavelengths, as shown in Fig. 1.10. In practical terms, this means that the same target, when imaged at different positions across the camera's field of view, will produce spectra that appear to be from different wavelengths due to these angular variations.

The calibration process to correct this spectral shift requires knowledge of the lens f-number and the distance between the lens exit pupil and the image sensor. With these parameters, the shift in the central wavelength can be calculated for each pixel position. This is achieved by applying a model

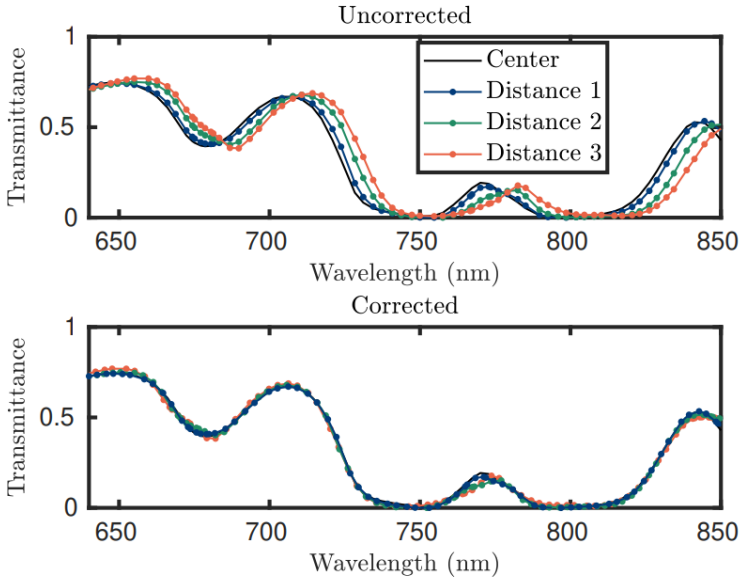


Figure 1.10 – The measured spectrum of a sample appears shifted if placed at different distances from the optical axis. Using spectral shift correction, the spectra can be realigned (Source [62]).

that adjusts the central wavelengths of each filter based on the chief ray angle, which is influenced by the off-axis distance of the pixel and the lens f-number.

1.4.7 Non-uniformity Correction

Non-uniformity correction addresses the variations in the pixel response of a CMOS imager across the sensor [148]. These variations can arise due to several factors, including inherent sensor non-uniformities and optical effects such as vignetting caused by the lens. Vignetting leads to a decrease in image brightness or saturation at the periphery compared to the image center.

The procedure for non-uniformity correction involves the creation of a non-uniformity correction cube, which characterizes the spatial variation in pixel response throughout the image sensor. This cube is obtained through calibration measurements using a uniform light source, such as an integrating sphere or a uniformly illuminated white reflectance tile, as shown in Fig. 1.11.

Once the non-uniformity correction cube has been created, it can be used to correct the images captured by the camera. This is done by dividing the reconstructed spectral cube—point-by-point—by the non-uniformity correction cube. The result is a set of images where the non-uniformities have been removed, leading to a more accurate and consistent representation of the scene spectral characteristics.

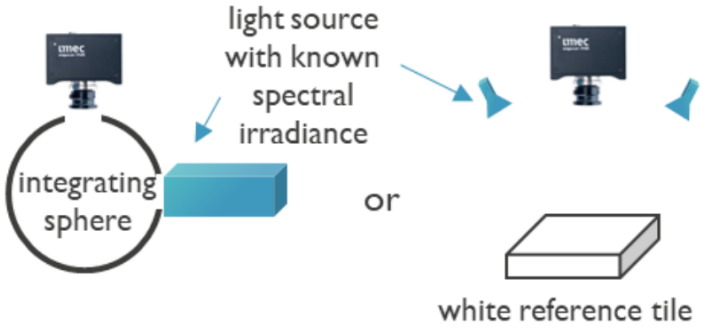


Figure 1.11 – Two measurement setups for obtaining the non-uniformity correction cube (Source [148]).

1.4.8 Final Key Insights

Camera calibration is crucial in achieving the highest fidelity in snapshot spectral imaging (SSI). As outlined throughout this section, the complex process encompasses a series of systematic steps designed to refine the raw spectral data captured by the camera's sensor. These steps range from bias correction to sophisticated angularity and non-uniformity adjustments, each serving a unique purpose in increasing the spectral accuracy.

Fig. 1.12 shows the overall processing pipeline of the SSI images with the required steps. It is important to note that some of these steps are optional and can be adapted based on specific requirements. For instance, using white and black references is essential to obtain reflectance values. However, these steps can be deleted if radiance values suffice for the intended analysis. Similarly, angularity correction has its applicability tied to the availability of the f-number of the optics. This correction is most relevant for fixed lenses, excluding those with autofocus mechanisms.

1.5 Conclusion

In this chapter, we explored SSI as an advanced approach within HSI, emphasizing its ability to capture complete spectral information in a single exposure. We began by introducing typical scanning technologies used in hyperspectral imaging, detailing their operational mechanisms and limitations. Then, we moved forward by exploring snapshot imaging technologies with a focus on Fabry-Perot-based ones highlighting their innovative contributions and the critical need for precise calibration to ensure data accuracy. The chapter also introduced the essential calibration processes for optimizing SSI technology performance, underlining their importance in producing reliable spectral data. Moving forward, Chapter 2 will delve into the complexities of image processing, specifically focusing on "Demosaicing," an essential step in building the 3-D datacube from the raw data collected by snapshot spectral cameras.

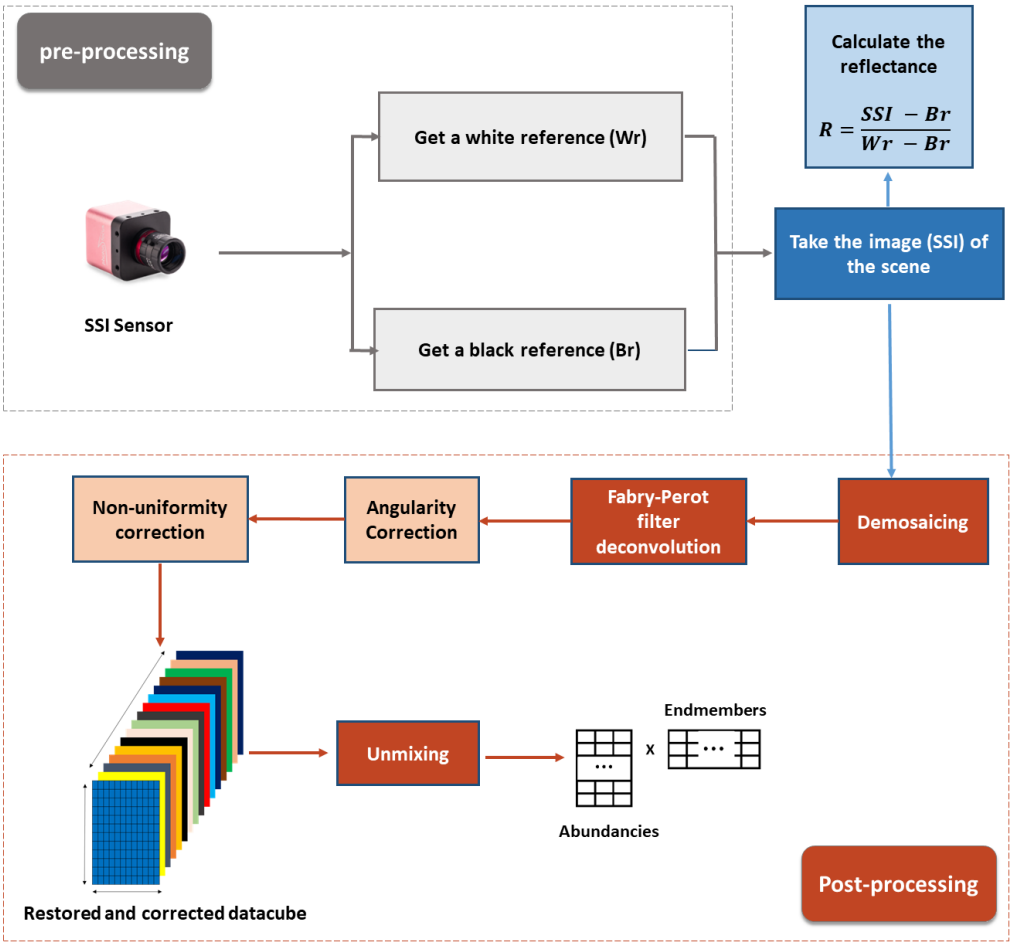


Figure 1.12 – Processing steps as required by the camera manufacturer to ensure high-quality snapshot spectral image restoration and unmixing.

Chapter 2

Demosaicing Snapshot Spectral Images

"Only with the heart can one see rightly; what is essential is invisible to the eye."

— Antoine de Saint-Exupéry, *The Little Prince*

2.1 Introduction

Demosaicing[1] is a well-established challenge in the imaging community. It is pivotal for RGB color images, considering the presence of Bayer-structured filters that sample the red, green and blue. In a typical Bayer filter array, half of the pixels are green, while red and blue represent a quarter of the total pixels, as shown in Fig 2.1. This arrangement reflects the Bayer pattern emphasis on green, acknowledging its significance in human color perception. Each pixel in this mosaic captures only one of the three colors, meaning the sensor lacks two-thirds of the color information for that particular location at any given point.

Demosaicing algorithms aim at reconstructing full-color images from this incomplete data. These algorithms work by interpolating the missing color information for each pixel. For instance, a pixel that captures green light will have its red and blue values estimated based on the adjacent pixels that capture these colors. The challenge lies in accurately predicting these missing values to create a seamless and realistic image.

The complexity of the demosaicing process varies depending on the algorithm used. Simple methods might employ bilinear or bicubic interpolation [65], which calculates the missing color values based on a weighted average of neighboring pixels. More advanced techniques, like gradient-based or edge-directed interpolation [65], are designed to preserve sharp edges and fine details better, thereby reducing artifacts. These sophisticated algorithms analyze the surrounding pixel patterns to make more informed estimations, particularly around edges or areas of high contrast, leading to a more accurate and visually satisfactory reconstruction.

[1] This chapter uses the terms "demosaicing" and "demosaicking" interchangeably. Both terms refer to the same process of reconstructing a 3-D datacube from an SSI image.

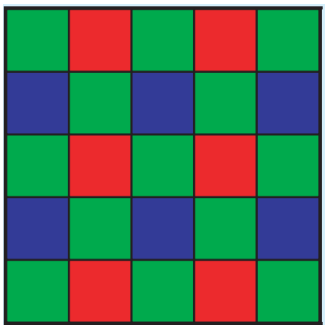


Figure 2.1 – Bayer color filter array arrangement

In the domain of snapshot spectral imaging (SSI), also known as Multispectral Filter Array (MSFA), demosaicing adopts a unique and more complex role. Unlike conventional RGB imaging, SSI/MSFA links each pixel with a specific spectral band, necessitating a different method for image reconstruction. In scenarios like the 16-band case, a common technique involves the formation of "super-pixels". This method—essentially a form of spatial down-sampling—gathers groups of $\sqrt{k} \times \sqrt{k}$—with common values for k being 9, 16 or 25—pixels into a single super-pixel. Here, k represents the number of wavelengths supported by the cameras. Although this approach simplifies the sampling process, it results in a notable reduction in spatial resolution by as much as 93.75% in the case when k=16, as illustrated in Fig. 2.2 [143]. These super-pixels represent a condensed version of the original pixel array, capturing a more comprehensive spectral range at the expense of detailed spatial information. However, demosaicing aims to generate the entire HS data cube and to conserve the spatial resolution. Therefore, various approaches for estimating the missing information have been proposed. These approaches can be classified into two categories, i.e., "Traditional" methods and "Deep Learning-based" strategies.

It is noteworthy that demosaicing is inherently influenced by the arrangement of wavelengths on the sensor surface, which varies according to the design of the MSFA pattern employed [38]. Several MSFA patterns are designed with specific goals, such as minimizing reconstruction artifacts, improving spatial and spectral resolution, and ensuring efficient data capture within hardware constraints. The most notable designs are illustrated in Fig. 2.3.

In this chapter, we focus on demosaicing methods for multispectral filter arrays that utilize Fabry-Perot interferometer technology (IMEC), Brauers (BRA) and Binary Tree-Based Edge-Sensing (BTES). We acknowledge the presence of alternative techniques suited to different snapshot imaging systems like those used in CASSI, as mentioned in Chapter 1. However, these methods fall outside our scope due to their specific use cases and the lack of commercially available CASSI cameras. Therefore, we limit our investigation to demosaicing techniques specific to the Fabry-Perot (IMEC) pattern due to its commercial presence and the practical applicability of our research. The study will also cover BTES and BRA methods since they are also applied in the literature to Fabry-Perot design.

Finally, it is essential to note that some of the approaches are adapted from techniques initially developed for the super-resolution problem. Super-resolution is designed to enhance the resolution of an

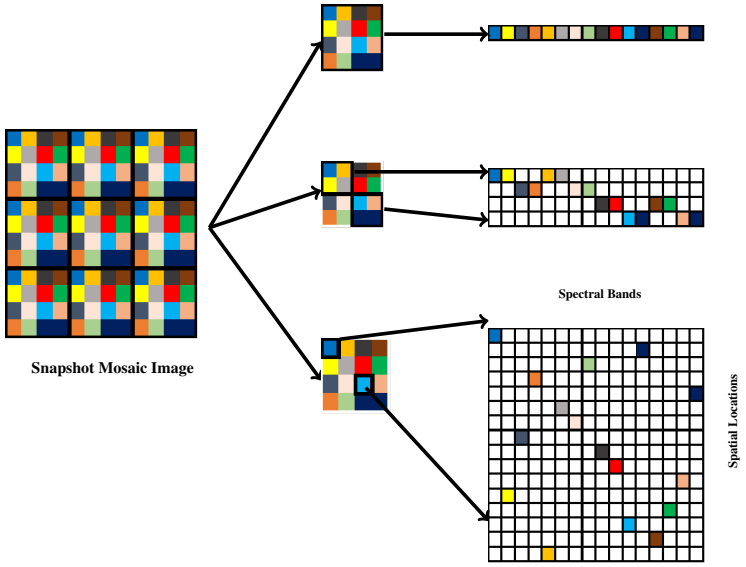


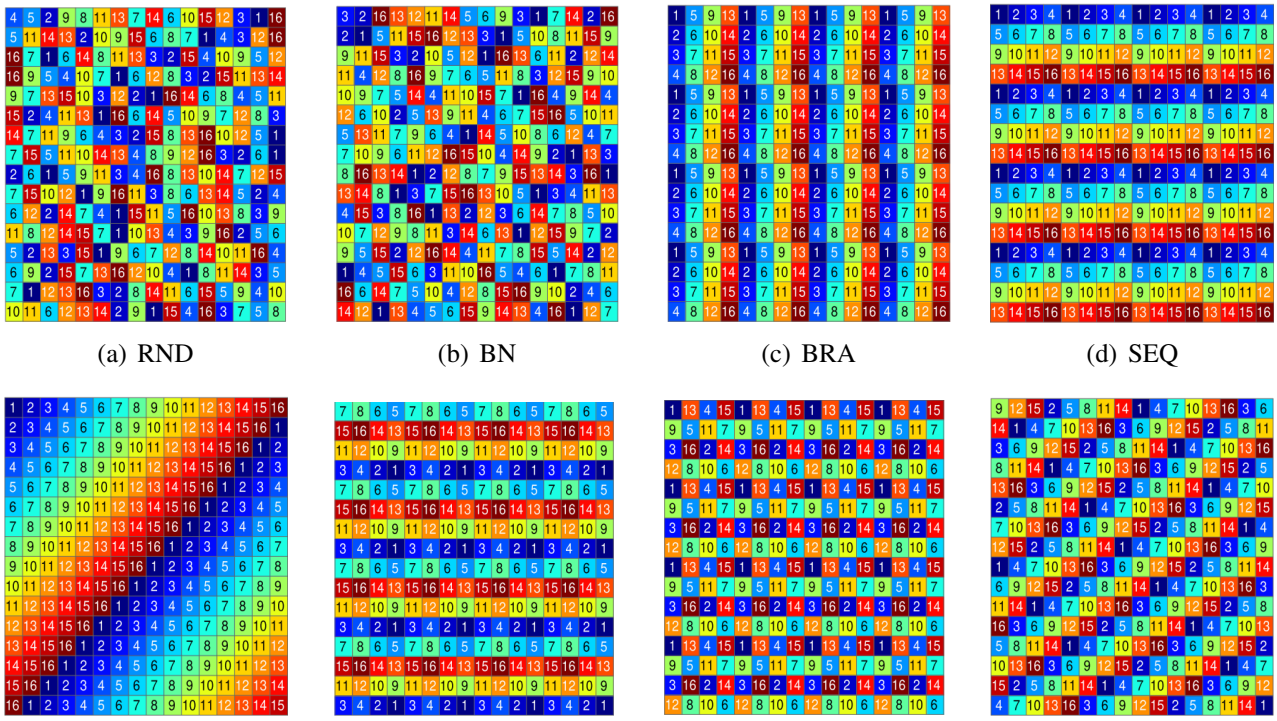
Figure 2.2 – A typical approach involves binning together groups of pixels, known as “super-pixels”, to produce lower spatial resolution images with more (middle) or all (top right) spectral bands. While demosaicing approaches aim to repopulate the full cube and keeping the full spatial resolution (bottom right). Adapted from [143].

image, either spatially or spectrally, by interpolating additional pixels or spectral bands not originally captured by the sensor [163]. However, it is crucial to distinguish between spectral spectral/spatial super-resolution and the demosaicing tasks addressed in this section. Demosaicing focuses on reconstructing missed spectral information in each spatial pixel without altering the image resolution. In contrast, spectral super-resolution aims to increase the spectral resolution by generating new bands not captured by the imaging sensor. Similarly, spatial super-resolution enhances the spatial details by increasing the image resolution.

2.2 Traditional Demosaicing Methods

This section focuses on a range of established methodologies distinct from contemporary deep-learning approaches. These traditional methods represent a range of algorithmic strategies developed over the years to address the challenge of datacube reconstruction in MSFA without relying on the complex neural network architectures that characterize modern deep learning techniques. Traditional demosaicing methods are rooted in various algorithmic and heuristic approaches, each designed to reconstruct a 3-D HSI datacube from the limited information captured by individual pixels in an MSFA. Traditional demosaicing methods can be broadly divided into five distinct categories:

- Demosaicing Based on Interpolation.



(a) RND (b) BN (c) BRA (d) SEQ

(e) UNIF (f) IMEC (g) BTES (h) OSP

Figure 2.3 – Different MSFAs with a spatial resolution 16×16 and spectral resolution 16 (Source [147])

- Demosaicing Based on Wavelet Analysis.
- Demosaicing Based on Binary Trees.
- Demosaicing Based on Pseudo-panchromatic Analysis.
- Demosaicing Based on Low-rank Matrix Approximation.

In this section, we will delve into each category, exploring the unique approaches and techniques they encompass.

2.2.1 Demosaicing Based on Interpolation

Interpolation-based demosaicing represents a foundational approach to image reconstruction from MSFA images. It estimates the missing spectral information at each pixel by leveraging the known color values of neighboring pixels, assuming that adjacent pixels share similar spectral characteristics.

These methods are based on various interpolation strategies, ranging from simple bilinear interpolation to more sophisticated adaptive and weighted schemes.

2.2.1.1 Weighted Bilinear Interpolation (WB)

This demosaicing method employs a two-step process for reconstructing each spectral band in a multispectral image. Initially, it generates a sparse, raw image for each spectral band, utilizing only the observed measurements. Demosaicing itself is achieved through a convolution process with a low-pass filter. This filter is strategically normalized to accommodate the missing measurements, ensuring that the interpolation is accurate and maintains the integrity of the original spectral information. The importance of the WB method lies in its role as a fundamental base initialization step, often serving as the starting point for most advanced demosaicing methods [27].

2.2.1.2 Spectral Difference (SD)

As outlined in [27], the Spectral Difference method leverages the disparities between spectral channels to enhance the sharpness in interpolated images. This approach is characterized by its utilization of differences in spectral channels and a series of operations that refine the image quality. The key steps of this method are:

- The process begins by subtracting one raw channel from another bilinear interpolated channel, with the raw channel containing non-zero entries exclusively where actual measurement data is present. This process is repeated over all the channels.
- To refine the chromatic color difference layer, a bilinear interpolation is applied. This step can be interpreted as a smoothing operation, contributing to the overall enhancement of the image.

- The final step involves an inverse transformation to generate the full interpolated image.

Although this method appears computationally complex, it is strong in maintaining image sharpness while managing spectral disparities.

2.2.1.3 Median Filtering in Multispectral Filter Array Demosaicing

In [153], Wang *et al.* delve into the adaptation of color filter array (CFA) demosaicing techniques to multispectral filter arrays, focusing on vector-based median filtering methods. Their study evaluates two approaches: direct vector median filtering (DVMF) and spherical space median filtering (SSMF). The former solves demosaicing problems by means of vector median filters, and the latter applies median filtering to the demosaiced image in spherical space as a subsequent refinement process to reduce artifacts introduced by demosaicing. The experiments show that while DVMF shows limited success, especially with natural images, the refinement process using SSMF consistently enhances image quality by reducing estimation errors.

2.2.1.4 A Linear Interpolation Algorithm for Spectral Filter Array Demosacing (LISFAD)

In [149] Wang *et al.* present a novel approach to demosaice MSFA images. The proposed method is based on a linear interpolation strategy incorporating residual interpolation, where the difference between the original measurements and the tentative estimation pixel values is interpolated to capture the finer details and improve the overall estimation. The proposed method effectiveness relies significantly on spectral correlation for optimal performance. It was also initially designed and tested on 4 and 5-band MSFAs, but it can be adapted to the IMEC-MSFA design.

2.2.1.5 Iterative Spectral Difference (ItSD)

Mizutani *et al.* [108] introduced an advanced extension to the SD method, known as Iterative Spectral Difference (ItSD). This method is particularly effective in scenarios with significant cross-correlation among images captured at different wavelengths. The ItSD method enhances the cross-correlation of demosaiced channels through an iterative process involving repeated interpolations considering inter-channel correlations, using previously interpolated pixel values as tentative data for subsequent interpolations.

A key aspect of ItSD is its dynamic regulation of the number of repetitions based on the spectral sensitivity of the color filters. If the optical wavelengths of two bands are closely aligned, implying a higher cross-correlation, the number of repetitions increases. This adaptive control mechanism is crucial for preserving and enhancing inter-band relationships during the demosaicing process, improving image quality. The method efficiency lies in its ability to iteratively interpolate bands by considering the spectral proximity and cross-correlation, thereby significantly improving the reconstructed image quality.

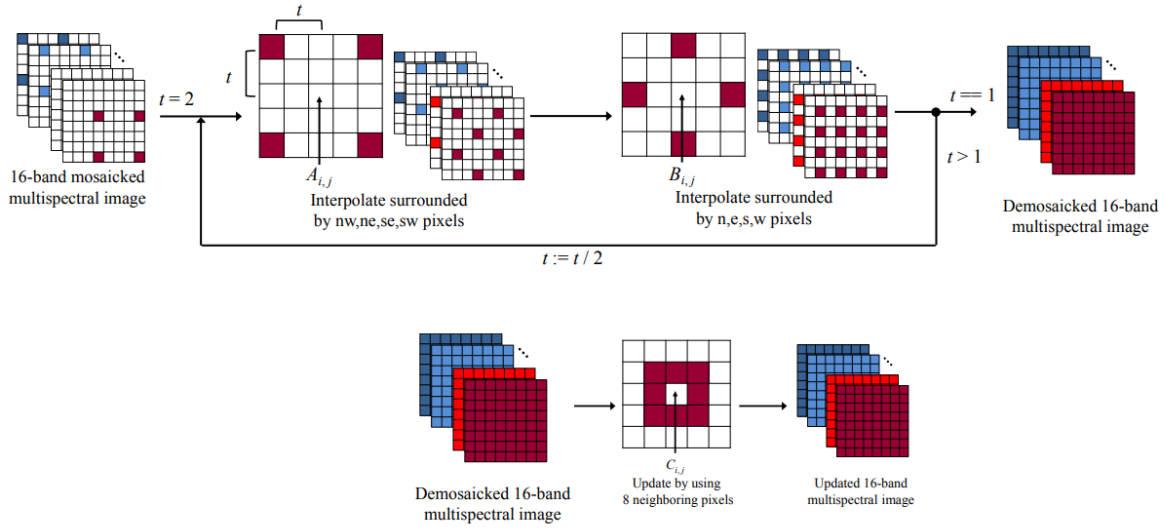


Figure 2.4 – Main flow of MLDI

2.2.1.6 Multispectral Local Directional Interpolation (MLDI)

The Multispectral Local Directional Interpolation (MLDI) [136] is a demosaicing method that emphasizes using local directional information in the interpolation process. Initially, the center pixel of an image is interpolated by considering the gradients among its four neighboring pixels: northwest (nw), northeast (ne), southeast (se), and southwest (sw). This first step focuses on capturing the local directional variations around the central pixel, thereby helping in a more accurate estimation of its value.

After this initial interpolation, the method progresses to interpolate the center pixel surrounded by the north, east, south, and west pixels. This further interpolation step refines the image by incorporating additional surrounding pixel information, thereby enhancing the overall accuracy of the interpolation.

Once all pixels have been interpolated through these steps, the MLDI method enters its second phase, where the interpolated pixels are modified using information from eight neighboring pixels, expanding upon the initial four-pixel gradient calculation. This step is crucial as it adjusts the initially interpolated pixels by incorporating a broader range of local information, leading to a more accurate image reconstruction, as shown in Fig 2.4.

Through this two-step process, MLDI effectively utilizes local directional gradients and expanded neighborhood information to achieve a good demosaicing of multispectral images.

2.2.2 Demosaicing Based on Wavelet Analysis

Wavelet Transform (WT) is a mathematical technique that decomposes a signal into components at various scales or resolutions. It uses wavelet functions localized in both time and frequency [11].

Meanwhile, discrete wavelet transform (DWT) is a specific implementation of the WT where both the scale and translation parameters are discretized. This means that the signal is analyzed at discrete intervals for both scale (frequency or resolution) and position (time) [40]. DWT is applied successfully to color filter array (CFA) demosaicing [43, 32]. In this section, we will explore its application in MSFA demosaicing,

2.2.2.1 Discrete Wavelet Trasform (DWT) MSFA Demosaicing

The DWT process involves decomposing the image into various subbands, including an approximation subband and detail subbands representing high-frequency information at different scales [152]. This subband coding forms the basis for DWT-based demosaicing, exploiting the inter-channel correlation inherent in most images. The DWT decomposition is achieved by applying high-pass and low-pass filters to the image through the following steps:

- Initiate the process by performing DWT decomposition on the MSFA image. This step includes splitting the image into different subbands, each characterized by varying frequency components.
- Employ linear estimation to determine the high-frequency coefficients of the decomposed image. This step involves estimating the details present in the high-frequency components.
- Simultaneously, employ WB to estimate the low-frequency coefficients. This step focuses on approximating the broader, low-frequency features within the image.
- Bring together the estimated low-frequency and high-frequency components, creating a comprehensive representation of the multispectral image.
- Apply the inverse discrete wavelet transform (IDWT) to the recombined components. This final step reconstructs the final datacube, leveraging the inverse transformation to restore the image original form.

The selected wavelet and the desired level of decomposition control the number of decomposition levels in DWT. Stopping the decomposition process at a specific level involves a trade-off between capturing sufficient frequency information and avoiding excessive computational complexity.

2.2.2.2 A Generic Multispectral Demosaicking Method Based on Inter-channel Spectrum Correlation (GMDICC)

The authors in [181] propose a novel demosaicing method for MSFA based on DWT and inter-channel spectrum correlation. The method is structured in three main parts:

1. **Image Pre-processing:**

   This step deals with the grayscale differences between adjacent pixels in the raw image. Significant grayscale differences can lead to noticeable horizontal and vertical bright bars, blurring

the structure and edges of objects in the image. To overcome this issue, the method employs a gray correction technique to normalize the gray level across all pixels, thus eliminating sharp changes in the raw image. By using wavelet transform for image fusion enhancement processing, the original spatial structure information of the image can be recovered, and the texture, contour, and structure information in the image can be enhanced.

2. **Spatial Reconstruction:**

   In this stage, the process distinguishes between low and high-frequency information in the raw image. For the low-frequency components, which generally contain broader image features and smoother gradients, the method employs weighted bilinear (WB) interpolation. This technique effectively reconstructs these broader image aspects by averaging adjacent pixel values, thus ensuring that the general structure and form of the image are preserved.

   On the other hand, high-frequency details, which include finer textures and edges, require a more accurate reconstruction approach. The process begins by decomposing the pre-processed image into N down-sampled (DS) images to isolate and focus on particular pixel groups for more detailed processing. After the down-sampling, each DS image is further decomposed using the DWT. This transform splits the images into various spatial frequency sub-bands, effectively isolating different frequency components of the image. In the final step, the corresponding DS images from the pre-processed image are used to replace the spatial high-frequency sub-bands in the down-sampled images. This means that the high-frequency details in the down-sampled images are substituted with those from the pre-processed images, ensuring that these finer details are accurately represented in the final reconstructed image.

3. **Demosaicing Process:**

   The demosaicing process involves estimating a fully defined channel difference and reconstructing each channel. This is done by summing the adjusted channel differences with the initial image estimations across all channels. The method employs a combination of initial estimations, channel difference calculations, and iterative refinement to achieve the final demosaicied image.

In summary, the method addresses grayscale inconsistencies and enhances low and high-frequency details. It combines gray correction, wavelet transform, weighted bilinear interpolation, and iterative channel reconstruction to improve the quality of the spatial and spectral reconstruction of multispectral images.

2.2.3 Demosaicing Based on Binary Trees

As previously discussed, the Binary Tree-based Edge-Sensing (BTES) method represents a strategic approach to organizing the spectral bands on the sensor surface [105]. Several demosaicing methods have been proposed to leverage the inherent advantages of the BTES structure, aiming to enhance the quality of image reconstruction of MSFA images. In this section, we will explore these methods.

2.2.3.1 Binary Tree-based Edge-sensing (BTES)

Miao *et al.* [105] proposed the BTES (Binary Tree-based Edge-Sensing) method exploiting the binary tree structure of the MSFA. The method starts with the band selection step, which determines the interpolation order for different spectral bands during the demosaicing process. The selection order adheres to the probability of appearance (PoA), where bands with higher PoAs have more detailed information. Initiating interpolation with bands possessing higher PoAs ensures the preservation and utilization of detailed information in the reconstruction process. The reconstructed image of a selected band helps interpolate other spectral bands with lower PoAs, enhancing the overall quality of the reconstructed image. Band selection happens by choosing leaf nodes at various levels of the binary tree, with deeper levels corresponding to bands with smaller PoAs. Then, the interpolation algorithm utilizes edge correlation information from the binary tree structure to perform the interpolation process.

2.2.3.2 Probability of Appearance based Convolution Filter based Weighted Bilinear (PCWB)

The PCWB method [131] exploits the probability of appearance (PoA) of spectral bands in MSFA. It employs a weighted bilinear interpolation algorithm that uses convolution filters specifically designed based on the PoA of each band within the MSFA pattern. The design of these filters is guided by several principles, including filter symmetry, accommodation of the minimum required known pixel values for interpolation, and normalization of weights for interpolation. The PCWB algorithm selects the appropriate filter for a band considering its PoA, facilitating the effective interpolation of missing pixel values. This method is beneficial for handling images from binary-tree-based MSFA.

2.2.3.3 Probability of Appearance based Convolution Filter based Bilinear Spectral Difference (PCBSD)

Building upon the initial estimates provided by the PCWB method, the PCBSD technique [131] further refines the demosaicing process by incorporating spectral correlation through the SD approach. This method calculates sparse band differences based on the initial multispectral image estimates and then fully defines these band differences using the PCWB interpolation method. Subsequently, it estimates the complete multispectral image by adjusting the interpolated values based on the calculated spectral differences. The PCBSD method enhances the quality of the generated datacube by leveraging the spectral correlation, especially for BTES MSFAs.

2.2.3.4 Generic Multispectral Demosaicking Based on Directional Interpolation (GMDDI)

The MSFA demosaicing method outlined in [132] depends on the spectral and spatial correlations in MSFA images. This method uses binary tree-based MSFA patterns, which are essential in determining the pixel arrangement for all bands, regardless of the number of bands in the raw image. Each band probability of appearance (PoA) within these binary tree patterns guides the selection of the appropriate interpolation scheme.

The method employs a progressive approach, initially estimating parts of missing pixel values before using these estimated values along with initially known pixel values to interpolate other unknown pixel values. It calculates four directional estimates of the band around an unknown pixel location and combines these estimates based on weights calculated for each direction as MLDI method. These weights are derived from the edge magnitude information of the band chosen for interpolation, differing from other methods that might use both the band chosen for interpolation and the band at the target interpolation location.

It is worth noting that MLDI and GMDDI methods utilize directional information for interpolation. However, while MLDI focuses on local directional gradients using neighboring pixels, GMDDI takes a more holistic approach by considering the binary tree-based MSFA patterns and utilizing progressive interpolation that integrates spectral and spatial correlations.

2.2.4 Demosaicing Based on Pseudo-panchromatic Analysis

Pseudo-Panchromatic Image (PPI) refers to an image generated from a multispectral or hyperspectral image, where the pixel values are derived by averaging or summing across several or all spectral bands. This process creates a grayscale image that closely represents the overall luminance or brightness of the scene. In MSFA demosaicing, PPIs serve as a reference to help reconstruct full-resolution multispectral images from the sparsely sampled data captured by MSFA sensors. This section explores the application of PPI in demosaicing MSFA images.

2.2.4.1 Pseudo-panchromatic Image Demosaicing (PPID)

The demosaicing method proposed in [107] employs a Pseudo-Panchromatic Image (PPI) to construct the 3-D datacube. This method assumes that the PPI is strongly correlated with all channels. Then, the average correlation coefficient between each channel and the PPI is computed to validate this. This correlation is a key factor allowing PPI-based demosaicing methods to enhance fidelity by exploiting inter-channel differences. The initial estimation of the PPI can be obtained through an averaging filter. However, local directional information is considered for a sharper PPI estimation. This approach involves computing weights based on the scale-adjusted raw image and integrating the estimated PPI into the demosaicing procedure to enhance the overall quality of the reconstructed datacube.

2.2.4.2 Generic Multispectral Demosaicking Using Spectral Correlation Between Spectral Bands and Pseudo-panchromatic Image (GMDSCPPI)

The demosaicing method presented in [133] employs a novel approach where a Pseudo-Panchromatic (PPAN) image is first estimated from the raw image by calculating a weighted average of pixel values from neighboring bands at each pixel location. Unique spatial filters designed for each band in the raw image are then applied to ensure the inclusion of at least one-pixel value from each spectral band. The reconstructed PPAN image is subsequently used to interpolate each spectral band in the raw image. This interpolation is achieved by computing differences between the raw and PPAN

images and applying convolution filters based on the probability of the appearance of each band. The process involves iterative refinement, where the PPAN image and the spectral bands are progressively enhanced to improve the overall quality of the reconstructed multispectral image.

2.2.5 Demosaicing Based on Low-rank Matrix Approximation

Low-rank matrix approximation (LRMA) is an essential technique to refine complex matrices, effectively preserving their important attributes while minimizing data complexity. LRMA mainly aims to reduce a matrix to a lower-rank version that closely resembles the original, based on specific norms such as the Frobenius or spectral norm. LRMA is helpful in various domains, including image processing, data compression, dimensionality reduction, and noise suppression [169, 88]. Additionally, Low-Rank Matrix Completion (LRMC) extends LRMA capabilities by accurately estimating missing or incomplete data [112, 81]. This section explores the application of LRMA and LRMC in demosaicing MSFA images.

2.2.5.1 Graph and Rank Regularized Demosaicing (GRMR)

Tsagkatakis *et al.* [143] present Graph and Rank Regularized Matrix Recovery (GRMR), a novel approach to recovering high spatial and spectral resolution content from SSI images within the low matrix completion framework. This approach advances beyond traditional sampling paradigms by incorporating realistic sampling strategies considering spectral profiles associated with each pixel. The method can be divided into two steps:

1. **Spectral Measurements Matrix Sampling**

   The SSI camera captures a 2-D image with each pixel representing a different spectral band, forming a subsampled matrix of the full-resolution hyperspectral tensor. An unfolding operator converts this 3-D tensor into a 2-D spectral measurements matrix by collapsing the spatial dimensions into a single pixel index. Sampling matrices/masks are created using binary, random, or spectral filter profile sampling methods.

2. **Spectral Measurements Matrix Recovery**

   The demosaicing problem is formulated as a regularized recovery of the spectral measurements matrix from limited measurements. Then, the nuclear norm with graph regularization is used to estimate the missing information. After that, the Proximal Gradient method uses the nuclear norm proximal operator to threshold the matrix singular values. Finally, the recovered matrix is folded again into a 3-D restored datacube

2.2.5.2 Multispectral Snapshot Demosaicing Via Non-convex Matrix Completion (MSDMC)

The authors in [16] proposed a method based on the matrix completion framework. The process involves solving an optimization problem to minimize the discrepancy between the observed data and

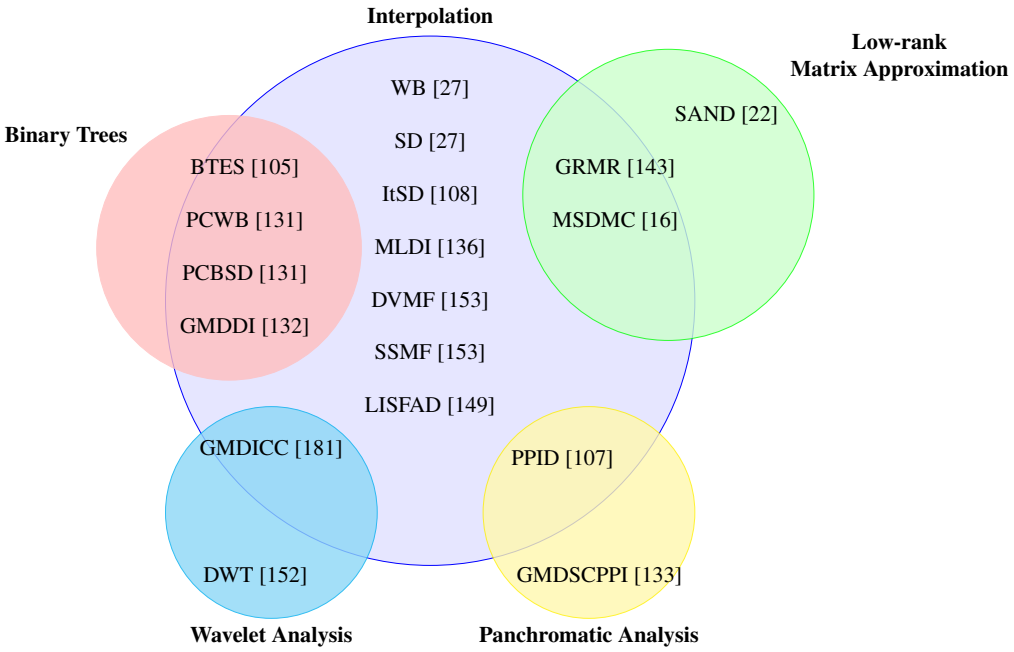


Figure 2.5 – Comparative Analysis of Demosaicing Techniques through Venn Diagram

a matrix representing an unfolding of the three-dimensional data cube, subject to a rank constraint. Then, algorithms like conjugate gradient iterative hard thresholding (CGIHT) [24] and alternating steepest descent (ASD) [140] are used to solve the optimization problem initialized with estimates from the spectral difference (SD) method.

2.2.5.3 Generalized MSFA Engineering With Structural and Adaptive Nonlocal Demosaicing (SAND)

The demosaicing method described in the article [22] combines structural and adaptive nonlocal optimization. Firstly, Generalized Alternating Projection-based Total Variation (GAP-TV) [172] is used to reconstruct multispectral images from the captured mosaic images. GAP-TV works by iteratively refining the image reconstruction. Then, the structural block matching step groups similar image blocks based on structural similarity across different spectral bands. This step allows more coherent and effective optimization in later steps, as operations can be customized to the specific characteristics of each group. After that, the image details are reconstructed by minimizing the rank of each block information and using the latent information extracted from similar blocks. Finally, the recovery steps are iterated adaptively and repeated several times with adjustments based on previous results until the best image quality is reached.

2.2.6 Discussion and Insights

In the realm of demosaicing for snapshot spectral imaging, we discovered various traditional methodologies, each with unique advantages and challenges. We classified them into five categories: Interpolation, Wavelet Analysis, Binary Trees, Pseudo-panchromatic Analysis, and Low-rank Matrix Approximation.

Interpolation methods, such as Weighted Bilinear Interpolation and Spectral Difference, stand out for their simplicity and computational efficiency. However, they can struggle with artifacts in areas of high contrast or fine detail. Wavelet Analysis, on the other hand, offers a more sophisticated approach to handling spatial and spectral resolution through its multi-scale decomposition. However, its performance can heavily depend on the choice of wavelets and decomposition levels. Binary Tree methods leverage hierarchical structures to optimize spectral band sampling, enhancing image reconstruction quality. However, these methods can become complex, requiring careful design and optimization of the tree structure. Pseudo-panchromatic techniques, which generate a grayscale image reflecting the overall scene luminance, can effectively preserve spatial details but may oversimplify the spectral information. Lastly, Low-rank Matrix Approximation and its extensions stand out for their ability to recover high-quality images from undersampled data, offering a robust solution to noise and missing data. However, these methods often involve intensive computations and complex optimization procedures.

The Venn diagram in Fig. 2.5 provides a comprehensive overview of how these demosaicing families intersect. Additionally, Table 2.1 summarizes all methods ordered by their publication date.

Table 2.1 – Traditional demosaicing methods comparison.

Method Name	**Approach Type**	**Key Features**	**Publication Date**
Weighted Bilinear Interpolation [27]	Interpolation	Fundamental base for other methods, uses convolution with a low-pass filter	2006
Spectral Difference (SD) [27]	Interpolation	Utilizes disparities between channels, refines image quality	2006
Binary Tree-based Edge-sensing (BTES) [105]	Binary Trees	Employs binary tree structure, focuses on edge correlation	2006
Discrete Wavelet Transform (DWT) [152]	Wavelet Analysis	Decomposes image into subbands, captures high-frequency details	2013
Direct Vector Median Filtering (DVMF) [153]	Interpolation	Direct vector-based median filtering for demosaicing	2013
Continued on next page			

Table 2.1 – continued from previous page

Method Name	Approach Type	Key Features	Publication Date
Spherical Space Median Filtering (SSMF) [153]	Interpolation	Refines vector median filtering to reduce errors	2013
Iterative Spectral Difference (ItSD) [108]	Interpolation	Enhances cross-correlation, adaptive iterations based on spectral sensitivity	2014
A Linear Interpolation Algorithm for Spectral Filter Array Demosaicing (LISFAD) [149]	Interpolation	Linear interpolation strategy incorporating residual interpolation	2014
Multispectral Local Directional Interpolation (MLDI) [136]	Interpolation	Utilizes local directional gradients and expanded neighborhood information for improved accuracy	2015
Pseudo-panchromatic Image Demosaicing (PPID) [107]	Pseudo-panchromatic Analysis	Utilizes mean value over all channels, emphasizes inter-channel differences	2017
Graph and Rank Regularized Demosaicing (GRMR) [143]	Low-rank Matrix Approximation	Combines rank and graph regularization, uses proximal gradient method	2019
Multispectral Snapshot Demosaicing via Non-convex Matrix Completion (MSDMC) [16]	Low-rank Matrix Approximation	Utilizes optimization and algorithms like CGIHT and ASD	2019
Generalized MSFA Engineering With Structural and Adaptive Nonlocal Demosaicing (SAND) [22]	Structural and Adaptive Optimization	Combines GAP-TV, structural block matching, and adaptive iteration	2021
Probability of Appearance based Convolution Filter based Weighted Bilinear (PCWB) [131]	Binary Trees	Uses convolution filters designed based on the probability of appearance	2021
Probability of Appearance based Convolution Filter based Bilinear Spectral Difference (PCBSD) [131]	Binary Trees	Refines PCWB by incorporating spectral differences	2021

Continued on next page

Table 2.1 – continued from previous page

Method Name	Approach Type	Key Features	Publication Date
Generic Multispectral Demosaicking Based on Directional Interpolation (GMDDI) [132]	Interpolation	Progressive approach, combines spectral and spatial correlations	2022
Generic Multispectral Demosaicking Using Spectral Correlation between spectral bands and pseudo-panchromatic image [133]	Pseudo-panchromatic Analysis	Uses Pseudo-Panchromatic image, iterative refinement	2023
A Generic Multispectral Demosaicking Method Based on Inter-Channel Spectrum Correlation (GMDICC) [181]	Wavelet Analysis	Focuses on inter-channel spectrum correlation, pre-processing, and spatial reconstruction	2023

2.3 Deep Learning-based Methods

In recent years, deep learning (DL) has influenced numerous fields of science and technology, offering substantial improvements over traditional methods in accuracy and efficiency. In image processing, deep learning techniques have particularly transformed the approach of demosaicing MSFA images. This section delves into various deep learning-based methods developed to address the complex challenges associated with demosaicing MSFA images.

As we explore these advanced methods, we can categorize them based on their underlying architectures—such as convolutional neural networks (CNNs) [13], residual networks (ResNets) [72], or those incorporating attention mechanisms [146]—or by their specific functionalities, like joint demosaicing and denoising or spectral correction. However, it is essential to note that despite these distinctions, there is a significant overlap in the functionalities offered by different architectures. For instance, CNNs and ResNets are often used with attention mechanisms to boost performance. This overlap makes it challenging to separate these methods into discrete groups cleanly.

Given these complexities and the integrated nature of many deep learning approaches, we have yet to group the methods strictly by architecture or functionality in this discussion. Instead, we present an overview of each method, allowing for a fluid comparison and a better appreciation of how different techniques may be combined or adapted to meet specific demosaicing challenges.

2.3.1 Deep Demosaicking for Multispectral Filter Arrays (In-Net)

The authors in [137] proposed a deep learning method using residual network (ResNet) [72]. It begins with preliminary interpolating the SSI image using a bilinear interpolation technique. Subsequently, the method employs a ResNet that integrates multiple three-dimensional convolutional layers and fixed linear units (ReLUs). This network structure refines the initially interpolated images by reducing artifacts in the final 3-D datacube.

2.3.2 Hyperspectral Demosaicking and Crosstalk Correction using Deep Learning (DsNet)

The method presented in the article [39] leverages an end-to-end deep learning framework to correct crosstalk and perform demosaicing in hyperspectral images. This network is based on CNN and is trained to identify and correct the interference between different spectral bands (crosstalk) and reconstruct full-resolution images from the SSI images. The network is trained with datasets containing known crosstalk and resolution characteristics, enabling it to apply learned patterns to new images.

2.3.3 Deep Panchromatic Image Guided Residual Interpolation for Multispectral Image Demosaicking (DGRI)

Pan *et al.* [115] proposed DGRI, a two-stage process utilizing deep learning for MSFA demosaicing. The initial stage involves using a ResNet-based deep learning model, DPI-Net, to create a high-quality panchromatic image from the raw MSFA data. This panchromatic image serves as a guide in the subsequent stage, where a two-pass guided residual interpolation method is applied to estimate and refine the missing spectral information for each pixel.

2.3.4 Deep Convolutional Networks for Snapshot Hyperspectral Demosaicking (SpNet)

The paper [66] employs deep convolutional neural networks (CNNs) based on the HSCNN+ architecture [162]. It includes 16 residual blocks and convolutional layers at both the input and output stages, each accompanied by ReLU activation layers. The entire system is trained on simulated datasets of hyperspectral images.

2.3.5 Deep Learning Approach for Hyperspectral Image Demosaicking, Spectral Correction and High-resolution RGB Reconstruction

The method in [95] primarily focuses on enhancing medical imaging applications by generating high-resolution RGB images from restored hyperspectral data cubes. This process is important for surgical imaging, where precise visual clarity and color accuracy are essential for effective decision-making

during surgeries. The method employs a CNN that incorporates the response of the Fabry-Perot filter (FPf) filter into the simulation of data. After restoring the image through demosaicing, spectral correction is applied, as Chapter 1 of the thesis outlines. Notably, applying the spectral correction matrix is integrated into the neural network training stage. This integration ensures that the network learns to reconstruct the hyperspectral image and simultaneously corrects spectral inaccuracies. The final output is an RGB image derived from the corrected and enhanced hyperspectral data optimized for medical applications.

2.3.6 Mosaic Convolution-Attention Network for Demosaicing Multispectral Filter Array Images

Feng *et al.* [46] developed Mosaic Convolution-Attention Network (MCAN), a novel approach for demosaicing SSI images. The network employs the Mosaic Convolution Module (MCM) and the Mosaic Attention Module (MAM).

1. Mosaic Convolution Module (MCM): This module handles the splitting of the periodic spectral mosaic in raw images. It does so while preserving the complete spatial information. The MCM unique feature is its position-sensitive weight-sharing strategy, which assigns weights based on the pixel relative position in the spectral mosaic pattern, as shown in Fig. 2.6.

2. Mosaic Attention Module (MAM): Complementing the MCM, the MAM targets the reduction of mosaic distortion within the spectral feature maps. It employs a novel position-sensitive feature aggregation strategy, termed "mosaic pooling", to aggregate features in the same relative position of the MSFA pattern. This module refines the spectral feature maps by focusing on the mosaic patterns, reducing distortion, and enhancing the overall image quality.

2.3.7 Hyperspectral Demosaicing of Snapshot Camera Images Using Deep Learning

The article [157] introduces a DL method particularly useful in fields requiring rapid image capture such as agricultural inspection. This technique uses a parallel neural network architecture trained on a newly developed ground truth dataset captured in a controlled environment with a hyperspectral snapshot camera. The dataset combines real-world captured scenes with simulated images adapted to the camera 4×4 mosaic pattern. The proposed demosaicing network employs feature extraction layers and two deconvolution layers. It operates by first converting the mosaic to a sparse cube and then enhancing features through residual blocks to efficiently handle different spectral and spatial information.

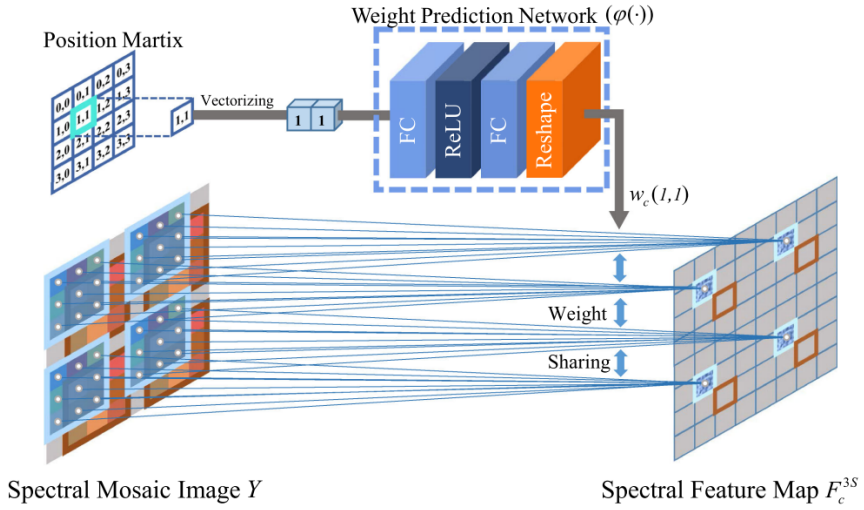


Figure 2.6 – Mosaic convolution module (MCM) for one channel case

2.3.8 Joint Spatial-spectral Pattern Optimization and Hyperspectral Image Reconstruction

Zhang *et al.* [179] introduce a deep-learning-based approach adopted from a super-resolution problem to optimize spatial and spectral patterns jointly. This method employs a CNN that involves joint optimization of the MSFA and spectral sensitivity functions (SSFs) patterns, which are critical for capturing high-quality spatial and spectral details. The network learns to optimize these patterns directly during training, allowing for a more integrated and efficient hyperspectral image reconstruction process. It is worth noting that the article introduces the concepts of spatial demosaicing (SpaDM) and spectral super-resolution (SpeSR), which are traditionally performed sequentially. The method employs them simultaneously to reduce the errors accumulated when these processes are separated.

2.3.9 MSFA-Frequency-Aware Transformer for Hyperspectral Images Demosaicing (FDM-Net)

Zeng *et al.* [174] introduce FDM-Net. A novel demosaicing framework incorporating a unique MSFA-frequency-aware multi-head self-attention mechanism called MaFormer and a Fourier zero-padding method [44] to effectively separate and reconstruct high-frequency and low-frequency components of the HSI. The model leverages the MSFA pattern information and non-local dependencies to enhance the demosaicing process, particularly improving the recovery of high-frequency details, which are often challenging to restore using conventional methods.

2.3.10 A Snapshot Multi-Spectral Demosaicing Method for Multi-Spectral Filter Array Images Based on Channel Attention Network

The article [180] introduces a novel approach that leverages a channel attention network integrated with a CNN to improve the demosaicing of multi-spectral images. The network utilizes a CNN as the fundamental building block. CNNs are particularly effective in extracting hierarchical features. Then, the channel attention mechanism focuses on enhancing specific features within the data channel dimension. The network prioritizes which features are most relevant for reconstructing the missing values by weighting the import of each channel.

2.3.11 Discussion and Insights

In this section, we explored deep learning methods for demosaicing multispectral filter array (MSFA) images. These methods leverage complex neural network architectures, including Convolutional Neural Networks (CNNs), Residual Networks (ResNets), Attention Networks, and Transformers, to address the challenges associated with reconstructing high-quality images from MSFA images. Table 2.2 summarizes all methods ordered by their publication date.

Deep learning approaches generally demonstrate superior performance over traditional methods, especially in handling high-dimensional data and preserving details and color accuracy. Despite these advancements, deep learning methods face significant challenges, notably in generalization and data requirements. They often require large amounts of labeled training data to perform effectively. Furthermore, these models may struggle to generalize to new, unseen data or different MSFA configurations not included in the training set. To address these challenges, future research in deep learning for MSFA demosaicing could focus on developing more robust models that require less training data and can generalize better across different MSFA patterns. Techniques such as transfer learning [77], few-shot learning [139], and synthetic data generation [50] could play significant roles in achieving these goals. Additionally, integrating unsupervised or semi-supervised learning paradigms could reduce the dependency on large labeled datasets and help improve the model ability to generalize to new conditions.

Table 2.2 – Deep Learning Methods for MSFA Demosaicing.

Method Name	**Key Characteristics**	**Architecture**	**Publication Date**
Deep demosaicking for multispectral filter arrays (In-Net) [137]	Uses ResNet for refining preliminary interpolations	ResNet, CNN	2018
Hyperspectral demosaicking and crosstalk correction using deep learning (DsNet) [39]	Corrects crosstalk and performs demosaicing	CNN	2018
Continued on next page			

Table 2.2 – continued from previous page

Method Name	Key Characteristics	Architecture	Publication Date
Deep Panchromatic Image Guided Residual Interpolation for Multispectral Image Demosaicking (DGRI) [115]	Guides demosaicing with a high-quality panchromatic image	ResNet, CNN	2019
Deep convolutional networks for snapshot hyperspectral demosaicking (SpNet) [66]	Utilizes HSCNN+ architecture for demosaicing	CNN	2019
Deep learning approach for hyperspectral image demosaicing, spectral correction, and high-resolution RGB reconstruction [95]	Optimizes medical imaging applications	CNN	2021
Mosaic Convolution-Attention Network for Demosaicing Multispectral Filter Array Images (MCAN) [46]	Incorporates attention mechanisms to focus on relevant features	Attention Network, CNN	2021
Hyperspectral demosaicing of snapshot camera images using deep learning [157]	Uses a parallel neural network trained on a new ground truth dataset	ResNet, CNN	2022
Joint spatial-spectral pattern optimization and hyperspectral image reconstruction [179]	Employs CNN for optimizing MSFA and SSFs patterns	CNN	2022
MSFA-Frequency-Aware Transformer for Hyperspectral Images Demosaicing (FDM-Net) [174]	Incorporates MaFormer and Fourier zero-padding to enhance high and low-frequency details	Transformer, CNN	2023
A Snapshot Multi-Spectral Demosaicing Method for Multi-Spectral Filter Array Images Based on Channel Attention Network [180]	Uses a channel attention network integrated with a CNN for enhanced feature extraction	CNN, Channel Attention Network	2024

2.4 Conclusion

This chapter extensively explored the diverse methodologies applied in demosaicing snapshot spectral images. Traditional methods remain relevant, showing persistent development with recent studies extending into 2023. Deep learning-based methods have also made significant strides, providing advanced solutions that often surpass traditional methods of detail preservation and computational efficiency. However, despite their advanced capabilities, these methods face challenges such as high data requirements and limited generalization. This can hinder their applicability when training data is limited and highly variable.

Finally, it is worth noting that Tsagkatakis *et al.* [143] stated that applying classification directly on SSI images post-demosaicing often leads to unsatisfactory performance. This limitation is primarily attributed to the estimation errors introduced during the demosaicing process, which can accumulate and propagate into subsequent image-processing tasks. This insight is essential for directing future research efforts toward developing new approaches that bridge the gap between demosaicing and subsequent image processing tasks like classification or unmixing.

Chapter 3

Post-Processing of Demosaiced Images: Unmixing Techniques

"We can only see a short distance ahead, but we can see plenty there that needs to be done."

— Alan Turing

3.1 Introduction

In the field of snapshot spectral imaging, constructing a 3-D datacube from captured images is only the beginning of extracting meaningful information. Once a complete datacube has been demosaiced, a range of post-processing tasks can be applied, e.g., unmixing, classification, anomaly detection, or change detection. These processes are fundamental in translating the raw spectral data into interpretable insights.

Hyperspectral unmixing (HU) is the process that separates the pixel spectra from a hyperspectral image into a collection of endmembers and a set of fractional abundances. The endmembers represent the pure materials in the image, and the abundances at each pixel represent the percentage proportion of the presence of each endmember in the pixel [23].

This chapter delves into the diverse techniques and challenges associated with hyperspectral unmixing, providing a brief overview of the methods utilized to unmix hyperspectral data. While ex-

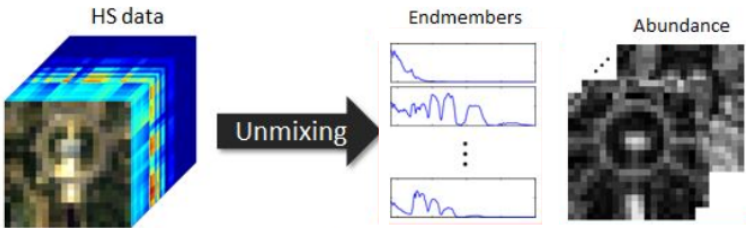


Figure 3.1 – Example of hyperspectral unmixing (source: adapted from [99]).

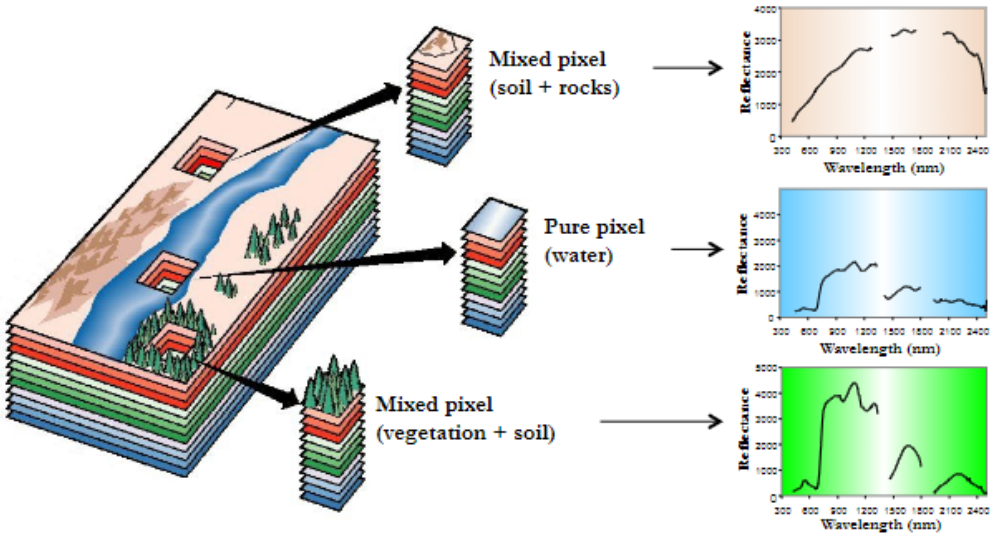


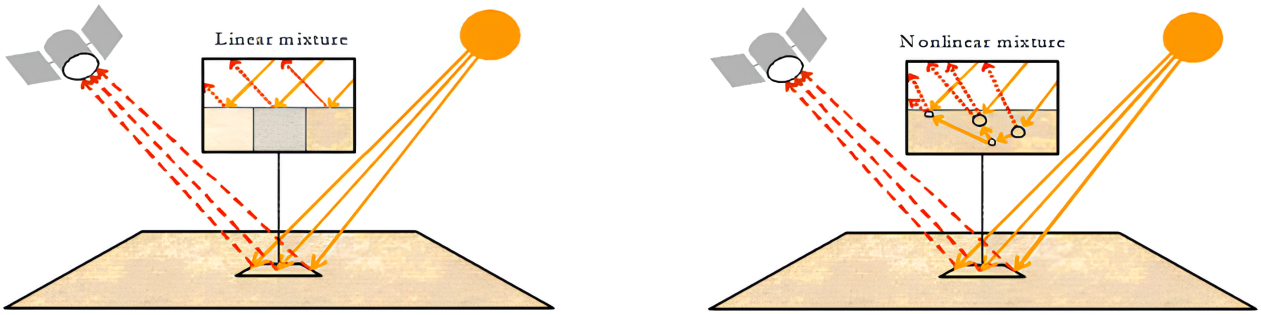
Figure 3.2 – Mixed and pure pixels in hyperspectral image (Source [33])

tensive studies and classifications of unmixing methods exist in the literature [23, 21, 47, 73, 26, 58], this chapter will not aim to exhaustively revisit these methodologies. Instead, we will briefly outline a general classification of unmixing techniques to set the stage for a focused examination of Nonnegative Matrix Factorization (NMF) and Sparse Component Analysis (SCA) because these two approaches are particularly relevant to our proposed frameworks.

3.2 Fundamentals of Hyperspectral Unmixing

Hyperspectral unmixing (HU) refers to any process that separates the pixel spectra from a hyperspectral image into a collection of constituent spectra, or spectral signatures, called endmembers, and a set of fractional abundances, one set per pixel. The endmembers are generally assumed to represent the pure materials present in the image and the set of abundances, or simply abundances, at each pixel to represent the percentage of each endmember that is present in the pixel [70, 23].

In hyperspectral imagery, pixels typically fall into two categories: mixed and pure. Pure pixels are those where the observed spectrum corresponds exclusively to a single material without interference from other materials within the sensor field of view. These are ideal for the straightforward identification of endmembers because they represent the spectral signatures of the materials. Conversely, mixed pixels contain spectral signatures from multiple materials. These occur due to the spatial resolution limits of the hyperspectral sensors, where the area represented by a single pixel may include several different materials, as illustrated in Fig. 3.2. The challenge in hyperspectral unmixing lies in accurately decomposing these mixed pixels into their constituent endmembers and determining the proportion of each endmember, respectively, the abundances.



(a) Linear Mixing (b) Nonlinear Mixing

Figure 3.3 – Linear and Nonlinear mixing models (Source [134])

The mixing model in mixed pixels can be characterized as either linear or nonlinear [58, 23, 70]. Linear mixing holds when the mixing scale is macroscopic, and the incident light interacts with just one material, as shown in Fig. 3.3(a). Conversely, nonlinear mixing is usually due to physical interactions between the light scattered by multiple materials in the scene as shown in Fig. 3.3(b). These interactions can be at a classical, multilayered, a microscopic or intimate level [25, 23].

In addition to spectral mixing effects, many other interferences can significantly affect the process of analyzing hyperspectral data as shown in Fig. 3.4. These external factors can introduce errors and spectral variability [26] that complicate the unmixing process:

- Atmospheric Interferences: Atmospheric conditions such as the presence of water vapor and various gases can act as potential sources of errors in spectral unmixing. These atmospheric constituents can absorb and scatter electromagnetic radiation as it travels from the earth surface to the sensor, altering the observed spectral signatures [101].

- Multiple Scattering Effects: The phenomenon of multiple scattering—where light rays bounce off multiple surfaces before reaching the sensor—can also lead to model inaccuracies. This effect mainly exists in complex environments like urban areas or dense forests, where light interacts with multiple surfaces at different angles. Multiple scattering can result in nonlinear mixture which complicate the unmixing process [12].

- Shadows and Variable Illumination Conditions: Shadows can cause underestimation of reflectance values, which may lead to incorrect interpretations of the material properties. Variable illumination—due to changes in sunlight angle and cloud cover—can similarly alter the appearance of the landscape in hyperspectral images taken at different times or dates [26].

Finally, the hyperspectral unmixing processing chain typically involves atmospheric correction, dimensionality reduction, and the unmixing process. Dimensionality reduction is essential since the data dimensionality is usually much lower than the number of available bands. Identifying the appropriate subspaces is essential for reducing dimensionality, which can enhance algorithms performance and reduce both computational complexity and data storage needs [23, 58].

In this chapter, we focus exclusively on methods related to the Linear Mixing Model (LMM), which can be mathematically expressed as follows:

$$x_i = \sum_{j=1}^{p} g_{ij} \cdot f_j + \omega_i \tag{3.1}$$

where:

- $\cdot$ denotes the dot product.
- x_i is the observed spectral vector for the i-th pixel.
- g_{ij} is the abundance of the j-th endmember in the i-th pixel.
- f_j is the spectral signature of the j-th endmember.
- ω_i is an additive noise vector for the i-th pixel.
- p is the number of endmembers.

To ensure the physical meaning of the abundances, two main constraints are imposed:

1. **Abundance Nonnegativity Constraint (ANC):**

$$g_{ij} \geq 0, \quad \forall i, j \tag{3.2}$$

   This constraint ensures that the abundances are non-negative, meaning that the proportion of each endmember cannot be less than zero.

2. **Abundance Sum-to-One Constraint (ASC):**

$$\sum_{j=1}^{p} g_{ij} = 1, \quad \forall i \tag{3.3}$$

   This constraint ensures that the abundances sum to one, meaning the total proportion of all endmembers in a pixel is exactly one.

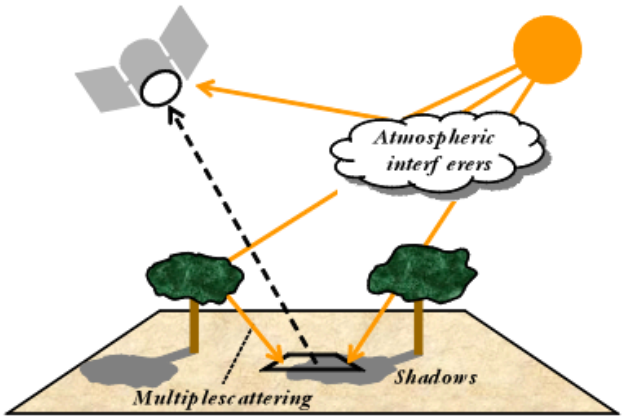


Figure 3.4 – Example of atmospheric interferers, multiple scattering, and shadows and variable Illumination effects (Source [33])

3.3 Hyperspectral Unmixing Methods

Hyperspectral unmixing aims at decomposing mixed pixel spectra into a collection of spectral signatures or endmembers and their corresponding abundance fractions. This decomposition is critical for identifying and quantifying the materials present within each pixel of a hyperspectral image. Various unmixing methods have been developed, each designed to leverage different aspects of the data and assumptions about the scene. These methods fall into four categories: *geometrical*[1]*, sparse regression-based, deep learning-based, and statistical* [47]. Each category encompasses a range of algorithms designed to address specific challenges posed by hyperspectral data, such as nonlinearity, noise, and high dimensionality. This section will explain each category briefly and the main methods that it encompasses. Fig. 3.5 presents a visual classification of hyperspectral unmixing algorithms.

3.3.1 Geometrical Methods

Geometrical approaches to hyperspectral unmixing rely on the fact that linearly mixed vectors form a simplex set or a positive cone. These approaches are typically executed in two steps. First, they aim to estimate the endmember matrix. Then, the abundance matrix can be derived as a nonnegative regression problem. These methods are categorized into two main subcategories: Pure Pixel (PP) based and Minimum Volume (MV) based. The pure-pixel-based[2] algorithms operate under the "pure-pixel assumption," which states that for each endmember present in the image, at least one pixel consists entirely of that material without mixing with other materials. When this assumption is satisfied, the endmembers can be found in the data matrix. The challenge of these methods lies in how they iden-

[1]Most methods in this category are statistical methods with geometrical interpretations.

[2]This family of methods was discovered in the mathematics, computer science, and signal processing communities under the names of "separable" or "convex" Nonnegative Matrix Factorization (NMF).

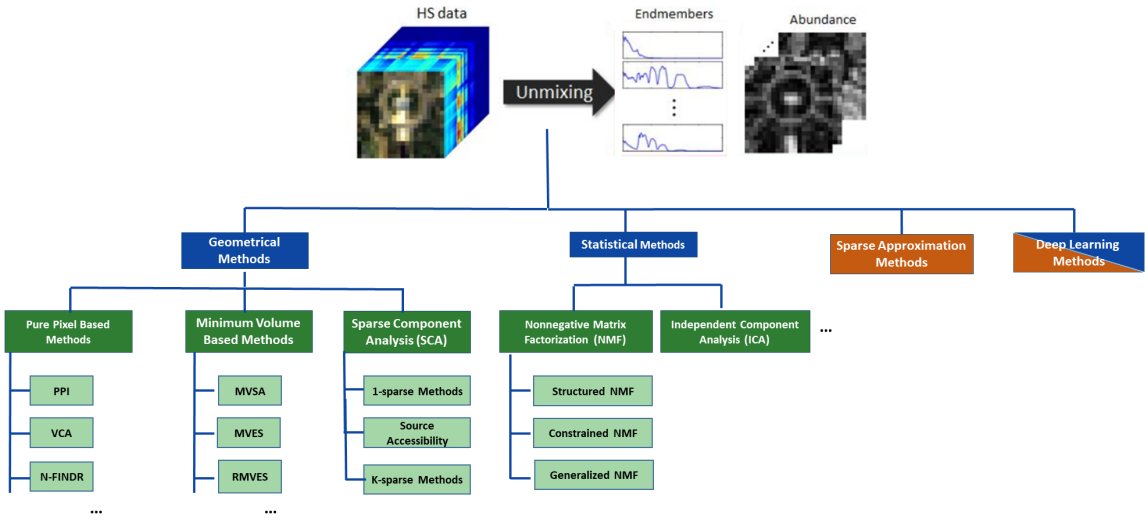


Figure 3.5 – Classification of Hyperspectral Unmixing Algorithms. Supervised methods are highlighted in orange, and unsupervised methods in blue.

tify these pure pixels. The classical methods include the pixel purity index, N-FINDR, iterative error analysis, vertex component analysis (VCA), simplex growing algorithm, and the sequential maximum angle convex cone [23].

PPI projects spectral vectors onto numerous random vectors to identify extreme points, assuming the existence of pure pixels. N-FINDR inflates a simplex within the data to find the set of pixels that define the maximum volume. IEA implements a series of linear constrained unmixing, each time choosing those pixels as endmembers to minimize the remaining error in the unmixed image. VCA projects data in directions orthogonal to the subspace spanned by already identified endmembers. The new endmember signature corresponds to the extreme of the projection. The algorithm iterates until all endmembers are exhausted. SGA iteratively grows a simplex by adding vertices that maximize the simplex volume, which can handle moderate mixing but struggles without pure pixels. SMACC constructs a convex cone and selects new endmembers based on the maximum angle formed with the existing cone.

Minimum Volume-Based algorithms focus on identifying the smallest simplex or cone that can contain all the data points. The motivation behind these methods is that the true endmembers should define a volume as small as possible, assuming that smaller volumes reduce the chances of including mixed or outlier pixels within the endmember set. Minimum Volume Simplex Analysis (MVSA) [93] is a notable MV method, which aims to minimize the volume of the simplex while ensuring that all data points are contained within or are very close to the simplex boundaries. Minimum Volume Enclosing Simplex (MVES) [28] and Robust Minimum Volume Enclosing Simplex (RMVES) [15] further refine the approach by adding robustness to noise and outliers, ensuring that the simplex accurately represents the pure materials.

Sparse Component Analysis (SCA) is one of the main approaches to Blind Source Separation

(BSS). When applied to hyperspectral data, while a few SCA methods assume the endmembers to be locally sparse [123], most works focused on the sparsity of the abundances [35] . Finally, while SCA is not truly a geometrical approach, it incorporates some geometrical constraints and utilizes similar assumptions to those found in geometrical methods. We will discuss SCA further in section 3.5.

3.3.2 Sparse regression-based Methods

Sparse regression-based methods depend on a spectral signature library and suppose each pixel in the image as is a linear combination of a set of pure spectral signatures in this library, i.e., contrary to the above methods, they are supervised. Multiple methods incorporating weights have been proposed to enhance the sparsity of endmembers in the spectral library and boost the sparsity of fractional abundances [176, 92, 128, 80]. Sparse regression-based optimization problem can be expressed as:

$$\min_{G} \|X - GF\|_2^2 + \lambda \|G\|_1 \tag{3.4}$$

where:

- X is the observed spectral data matrix.
- F is the dictionary matrix containing spectral signatures (endmembers).
- G is the abundance matrix to be estimated.
- $\| \cdot \|_2^2$ denotes the Frobenius norm, which measures the reconstruction error.
- $\| \cdot \|_1$ denotes the ℓ_1-norm, which promotes sparsity in the abundance matrix G.
- λ is a regularization parameter that controls the trade-off between the data fitting term and the sparsity term.

The constraints for the optimization problem are:

$$G_{ij} \geq 0, \quad \forall i, j \tag{3.5}$$

$$\sum_{j=1}^{p} G_{ij} = 1, \quad \forall i \tag{3.6}$$

3.3.3 Deep Learning-Based Methods

Recently, deep learning has shown remarkable power and potential in pattern recognition. Therefore, multiple methods have been proposed using autoencoder and its variants for hyperspectral unmixing, achieving more competitive unmixing performance [114, 118, 113, 21]. However, these approaches still suffer from drawbacks because they require a lot of training and network parameters to achieve acceptable unmixing performance [47].

3.3.4 Statistical Methods

When the mixtures are highly mixed, the geometrical-based methods generate unsatisfactory results because there are insufficient spectral vectors in the simplex facets. Instead, the statistical methods can be used to simultaneously identify the endmembers and their corresponding abundances by utilizing the HSI statistical properties. Independent component analysis (ICA) is a popular statistical method, but its use is not applicable in its standard way because the independence assumption is not valid due to sum-to-one constraints [111]. Still, modified ICA can be used for initializing other unmixing methods [20]. Nonnegative matrix factorization (NMF) is one of the most used statistical methods for unmixing. NMF provides a good fit for hyperspectral unmixing due to its nonnegativity and interpretability. We will discuss NMF further in section 3.4.

3.3.5 Discussion

In addition to the aforementioned categories, unmixing methods can also be differentiated based on their assumed mixing models: linear and nonlinear. In the linear mixing model (LMM), the observed spectrum at each pixel is considered as a linear combination of a set of endmemebers weighted by their abundance fractions. Contrarily, nonlinear mixing is due to the physical interactions between the light scattered by multiple materials in the scene. Most of the methods mentioned above have a low generalization ability in the nonlinear case. While bilinear and linear-quadratic models [103] can tackle the nonlinearity in the mixtures and have a satisfactory unmixing performance simultaneously. Many methods have been proposed to that end, e.g., [171, 104, 167]. Further, unmixing techniques can be classified based on their approach to the availability of prior knowledge: unsupervised methods, which do not rely on known endmembers and typically include geometrical and statistical approaches, and supervised methods, which utilize known spectral libraries as in sparse regression-based methods.

Finally, the endmembers can be significantly affected by atmospheric, illumination, or environmental variations within an image. Unfortunately, traditional spectral unmixing algorithms disregard the spectral variability of the endmembers, which reproduces significant errors throughout the whole unmixing process and compromises the quality of its results. Therefore, multiple methods have been proposed to take spectral variability into account [26]. Zhang *et al.* [175] proposed a convex optimization-based method for spectral variability-augmented reconstruction. Liu *et al.* [97] introduced a wavelet domain approach with Bayesian methods for more accurate spectral variability handling.

3.4 Nonnegative Matrix Factorization (NMF)

NMF is a powerful analytical tool with extensive applications in image processing, data mining, and signal processing, particularly within hyperspectral unmixing [47]. NMF aims to decompose a

nonnegative data matrix X into two lower-rank nonnegative matrices G and F, as follows

$$X \approx G \cdot F \tag{3.7}$$

where $\cdot$ denotes the dot product. Since an exact decomposition is generally not achievable, the approximation is typically assessed using several cost functions. The classical NMF optimization problem using the Frobenius norm is defined as:

$$\min_{G,F \geq 0} \|X - G \cdot F\|_F^2, \tag{3.8}$$

where $\| \cdot \|_F$ denotes the Frobenius norm. To solve the NMF problem, multiple standard methods can be employed, such as Multiplicative Updates (MU), Hierarchical Alternating Least Squares (HALS), and Alternating Least Squares (ALS) [59].

In hyperspectral imaging, X represents the observed data matrix, consisting of the spectral signatures of all pixels in the image, G is the abundance matrix, containing the proportions of each material present in every pixel, and F is the endmember matrix, representing the pure spectral signatures of the materials found within the scene. The fundamental assumption in NMF is that the endmembers and their corresponding abundances in the data are nonnegative. Several NMF-based HU methods have been proposed to seek better unmixing performance, which can be classified into three groups [47]:

- Constrained NMF
- Structured NMF
- Generalized NMF

For a comprehensive overview of NMF and its extensive applications beyond hyperspectral unmixing, interested readers are encouraged to refer to [164]. In this section, we focus specifically on applying NMF in hyperspectral unmixing.

3.4.1 Constrained Nonnegative Matrix Factorization

Constrained Nonnegative Matrix Factorization is a variant of the basic NMF that incorporates additional constraints into the optimization process to improve the unmixing performance. These constraints can be related to the properties of the endmembers and their abundances, helping to guide the factorization towards more physically meaningful solutions [47].

The general form of the Constrained NMF optimization problem can be expressed as:

$$\min_{G,F \geq 0} D(X, G \cdot F) + \alpha J_1(G) + \beta J_2(F), \tag{3.9}$$

subject to $G \geq 0$ and $F \geq 0$. Here, $J_1(G)$ and $J_2(F)$ are regularization terms for the endmembers and abundances, respectively, while α and β are nonnegative parameters that control the influence of these regularization terms.

The constraints for endmembers are integrated into NMF to account for endmember variability [135], smoothness [83], manifold constraints [100] and introducing prior spectral information [142].

Abundance constraints ensure that the abundance coefficients meet certain criteria, sum-to-one [47], sparsity [129, 183, 83], and manifold constraints [100].

3.4.2 Structured Nonnegative Matrix Factorization

Structured Nonnegative Matrix Factorization is a class of methods that incorporate additional structures into the factorization to handle specific problems in hyperspectral unmixing, such as accounting noise or accommodating missing data.

- **Weighted Nonnegative Matrix Factorization (WNMF)** is an extension of NMF that is particularly useful for dealing with data matrices X where some entries are missing or unreliable. WNMF aims to reconstruct the missing parts of X by introducing a weight matrix W where each entry w_{ij} indicates the reliability or availability of the corresponding entry in X. This approach modifies the NMF objective as follows:

$$\min_{G,F\geq 0} \|W \circ (X - G \cdot F)\|_F^2, \tag{3.10}$$

  where $\circ$ denotes the Hadamard product, and $\|\cdot\|_F$ denotes the Frobenius norm.

  WNMF can be sovled using different strategies:

  - Directly incorporating the weights into the update rules [75],
  - Using an Expectation-Maximization (EM) framework [41],
  - Employing stochastic gradient descent focusing on the available data points [89].

  Among these strategies, the EM approach is reported to be particularly effective, offering faster convergence and greater accuracy [41]. The EM strategy for WNMF includes two iterative steps: the Expectation (E-step) and the Maximization (M-step).

  In the E-step, the expected complete data matrix $\hat{Y}$ is estimated considering both the known data and the best estimates for the missing data, given by:

$$\hat{Y} = W \circ X + (\mathbb{1} - W) \circ (\hat{G}^{(t-1)} \cdot \hat{F}^{(t-1)}), \tag{3.11}$$

  where $\mathbb{1}$ is an matrix of ones with the same size as W, and $\hat{G}^{(t-1)}$ and $\hat{F}^{(t-1)}$ are the estimates of G and F from the previous iteration $(t-1)$.

  The M-step then consists of applying standard NMF update rules to $\hat{Y}$ to derive $\hat{G}^t$ and $\hat{F}^t$. Once NMF converged to a given solution [177] or after a given number $\text{Max}_{\text{outIter}}$ of iterations [41], $\hat{Y}$ is updated in another E-step using the last estimates of G^t and F^t [177, 41].

- **Project-based Nonnegative Matrix Factorization (PNMF)** incorporates a predefined spectral library within the NMF framework using the same principles of sparse regression-based methods [173]. This method introduces a transformation matrix $V$to project relevant spectra from the spectral library U onto a subspace, thereby obtaining the projected endmembers expressed as $F = UV$. The optimization problem can then be formulated as follows:

$$\min_{V \geq 0, G \geq 0} \|X - UVG\|_F^2 \tag{3.12}$$

  This approach generates endmembers from the spectral library and directly ties them to the hyperspectral images.

- **Robust Nonnegative Matrix Factorization (Robust NMF)** adapts the classic NMF framework to better handle the diverse noise types that can affect hyperspectral images e.g., Gaussian noise. The standard NMF is particularly sensitive to such noise, which can significantly degrade the unmixing performance. To enhance the robustness of NMF against these noises, various models have been developed that employ different metrics and regularizers designed to be less sensitive to outliers, Gaussian, and non-Gaussian noise distributions. These include Bounded Itakura–Saito (IS) Divergence [91], $L_{2,1}$-Norm Regularizer [49, 76] and Cauchy Function [150].

3.4.3 Generalized Nonnegative Matrix Factorization

Generalized Nonnegative Matrix Factorization extends the traditional NMF framework to adapt to various specific challenges by modifying the factorization approach. This generalization encompasses several subcategories:

- **Multilayer/Deep Extensions of NMF** incorporate multiple layers or levels of factorization, aiming to capture more complex structures and relationships within the data [130, 45].

- **Nonnegative Tensor Factorization (NTF)** deals with third-order tensors, which are the high-dimensional extensions of matrices [160]. This is particularly useful in hyperspectral imaging, where data naturally forms a three-dimensional tensor.

- **Linear-quadratic NMF** considers nonlinear interactions among the endmemebers by incorporating quadratic terms in the factorization model, allowing for the modeling of more complex spectral signatures [171].

- **Kernelized NMF** is introduced for nonlinear hyperspectral unmixing without the need of estimating the nonlinear mixture model [96, 182].

3.5 Sparse Component Analysis

Sparse Component Analysis (SCA) is a mathematical framework initially developed to address the challenges of audio source separation and has since been expanded to encompass a broader range

of applications, including hyperspectral unmixing and (post-)nonlinear mixtures [141, 145, 36, 126, 127]. SCA works primarily on signal sparsity, which means that signals can often be represented as sparse in some domain, whether temporal, frequency, or time-frequency. In HSI, "signal sparsity" can be either endmember sparsity [127] or abundance sparsity [85]. The advantage of SCA lies in its ability to decompose complex mixtures into their sources without requiring prior information about the mixing process. This characteristic is useful in environments where traditional methods might fail due to the complexity of source overlaps or when the mixing elements are unknown or complex. The methodology leverages mathematical techniques that transform signals into domains where their sparse characteristics are most evident, such as through the Fourier transform. SCA effectively exploits the sparse nature of data in this domains and isolates the individual components [8, 37, 124, 63, 31]. SCA leverages several sparsity assumptions to isolate individual sources:

- **1-sparse sources:** a source is considered 1-sparse if, within a given observation vector at a specific time or transformed domain—such as time-frequency—only one source is active while all others are not active. This high sparsity level is beneficial in environments where sources do not overlap significantly in the time-frequency domain, such as in certain types of audio processing where different sounds appear distinctly over time. This assumption was key in early applications, such as the audio separation method described in [170].

- **Sources are accessible:** A source is considered accessible when there are zone within the observation space—referred to as "single-source zones"—where the source is active while all other sources remain inactive [35]. This assumption simplifies isolating and identifying each audio source from a complex mixture. Methods like TiFROM [9]—later extended in [8, 124]—utilize this concept to enhance the effectiveness of SCA. The notion of accessibility of sources is similar to the "pure-pixel" assumption in hyperspectral unmixing. This assumption states that for each endmember present in the image, at least one pixel consists entirely of that material without mixing from other materials. Karoui *et al.* [85] extend the current SCA for applying hyperspectral unmixing to hyperspectral images by assuming abundance sparsity. The process involves identifying zones within the data where only one endmember is active. These zones are essential for estimating tentative endmembers based on observed data. Then, a clustering stage is employed to refine these estimates into a definitive set of unique endmembers. Finally, the abundances of each endmember in the hyperspectral images are estimated by solving an inverse problem.

- **q-sparse sources:** assume that up to q sources may be active simultaneously at any given observation point. This scenario often arises in complex signal environments where multiple sources contribute to the observed data simultaneously [109]. A similar concept is applied through methods that involve constraints like the Minimum Volume Constraint in hyperspectral unmixing, which focus on minimizing the volume of the convex hull formed by the endmembers in feature space [122]. Benachir *et al.* [19] proposed a geometrical method for blind unmixing of hyperspectral data using a two-source sparsity constraint, assuming at most two endmembers per local region. The methodology involves selecting local regions, extracting pairs of

endmembers based on their geometrical structure, and estimating their abundances with least squares.

3.6 Conclusion

In this chapter, we explored the fundamentals of the unmixing process, which attempts to decompose hyperspectral images into endmembers and abundances. We discovered a set of hyperspectral unmixing methods categorized into four groups: *Geometrical, Statistical, Sparse Regression-Based, and Deep Learning-Based methods*.

Geometrical methods are particularly useful in scenarios where pure pixels exist within the image, allowing for a straightforward identification and extraction of endmembers. While *Statistical methods* are useful when geometrical approaches fail, especially in highly mixed scenarios where the separation of spectral signatures becomes more complex. *Sparse regression methods* are supervised approaches that utilize a predefined spectra library to facilitate the unmixing process. In contrast, *Deep Learning methods* require extensive training data to achieve good performance but are increasingly favored for their ability to handle complex scenarios where training data are available.

As our focus is on Nonnegative Matrix Factorization (NMF) for our frameworks, we discussed its usage in unmixing. We identified three main categories of NMF: *Constrained, Structured, and Generalized NMF*, each dedicated to addressing specific challenges in hyperspectral imaging.

Finally, we explored Sparse Component Analysis (SCA), a framework initially introduced for audio source separation and later adapted to hyperspectral unmixing.

By the end of this chapter, we conclude the first part of the thesis, in which we conducted an extensive study of snapshot spectral imaging and post-processing technologies, including spectral correction, demosaicing, and unmixing. In the next part, we will discuss our frameworks for performing the post-processing techniques for snapshot spectral images.

Part II

Proposed Frameworks

Chapter 4

Locally-Rank-One-Based Joint Unmixing and Demosaicing Methods for Snapshot Spectral Images: a Matrix Completion Framework

"Innovation is seeing what everybody has seen and thinking what nobody has thought."

— Albert Szent-Györgyi

4.1 Introduction

In the preceding chapters, we introduced the fundamentals of snapshot spectral imaging and the demosaicing process. Demosaicing is a crucial preliminary step that reconstructs a full 3-D datacube from the raw data acquired by snapshot spectral cameras. Once this datacube is constructed, any post-processing applications, e.g., unmixing or classification, can be applied to translate the raw spectral data into interpretable and actionable insights as shown in Fig. 4.1. However, as Tsagkatakis *et al.* [143] have stated that applying classification on demosaiced SSI images often leads to unsatisfactory performance. Contrastingly, most existing unmixing methods are designed to work on the full 3-D datacube and do not account for the inherent missing entries typical in snapshot imaging scenarios. Furthermore, in the context of *in-situ* mobile sensor calibration, it has been demonstrated that a combined approach of low-rank matrix completion and factorization is significantly more efficient than a sequential two-stage process of matrix completion followed by factorization [42]. This gap in methodology underscores a crucial limitation in current practices, where the intricate balance between demosaicing and unmixing is not adequately addressed.

Recognizing this issue, our proposed approach seeks to perform demosaicing and unmixing jointly. This integrated methodology aims not only to enhance unmixing results but also to maintain optimal demosaicing performance.

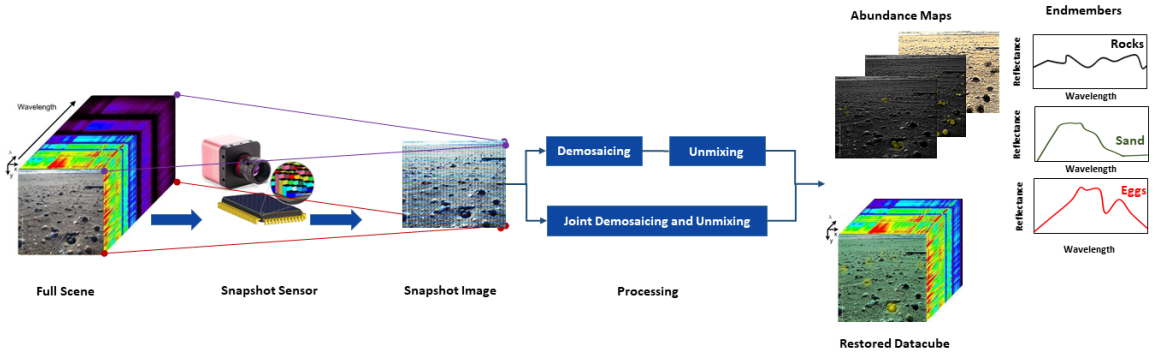


Figure 4.1 – Overview of acquisition and processing on a scene contains three endmemebers (sand, rocks, and eggs). The SSI image can be processed either by a two-stage approach or by applying unmixing and demosaicing jointly to restore the data cube, extract the endmembers, and find the abundance maps.

In this chapter[1], our contributions to the field of Snapshot Spectral Imaging (SSI) are outlined as follows:

- We concentrate on establishing a general framework for demosaicing images, applicable to various Multispectral Filter Arrays (MSFAs) [90]. This broad-scope framework sets the foundation for our comprehensive methods in "demosaicing" and "unmixing" hyperspectral images captured by SSI cameras.
- Our investigation is driven by the hypothesis that combining low-rank matrix completion and factorization is more efficient than a two-stage process involving these components separately.
- Three innovative methods are proposed for RAW SSI image demosaicing and unmixing:
  - The first is a Naive approach directly derived from Weighted NMF (WNMF).
  - While building on the Naive approach, the second and third methods operate under the assumption of sparsity within sensor "patches" mainly dominated by a single unique endmember. These methods differ in their assumptions and treatment in cases involving multiple endmembers.
  - The second and the third methods incorporate (i) rank-1 WNMF within the "patches", (ii) a single-source confidence metric, (iii) an endmember extraction stage, and (iv) a final abundance estimation step.

Building on the general framework established in this chapter, Chapter 5 will explore the complete pipeline provided by the camera manufacturer, focusing on new unmixing approaches for RAW SSI images. This chapter is designed specifically for applications involving Fabry-Perot filters or MSFA sensors with less selective filters, addressing complex scenarios often encountered in practice.

[1] Methods introduced in this chapter were published in [6, 4, 3].

4.2 Problem Statement

As previously mentioned, the considered SSI technology operates on a mosaic of Fabry-Perot filters, allowing each camera pixel only to capture one unique narrow band of wavelengths in the ideal case. Unmixing the SSI images requires processing the images to build the 3-D data cube. We aim to apply the unmixing and demosaicing jointly, starting from the raw SSI data (see Fig. 4.2). As a consequence, in this section we define the SSI system and the joint demosaicing and unmixing problem we aim to solve.

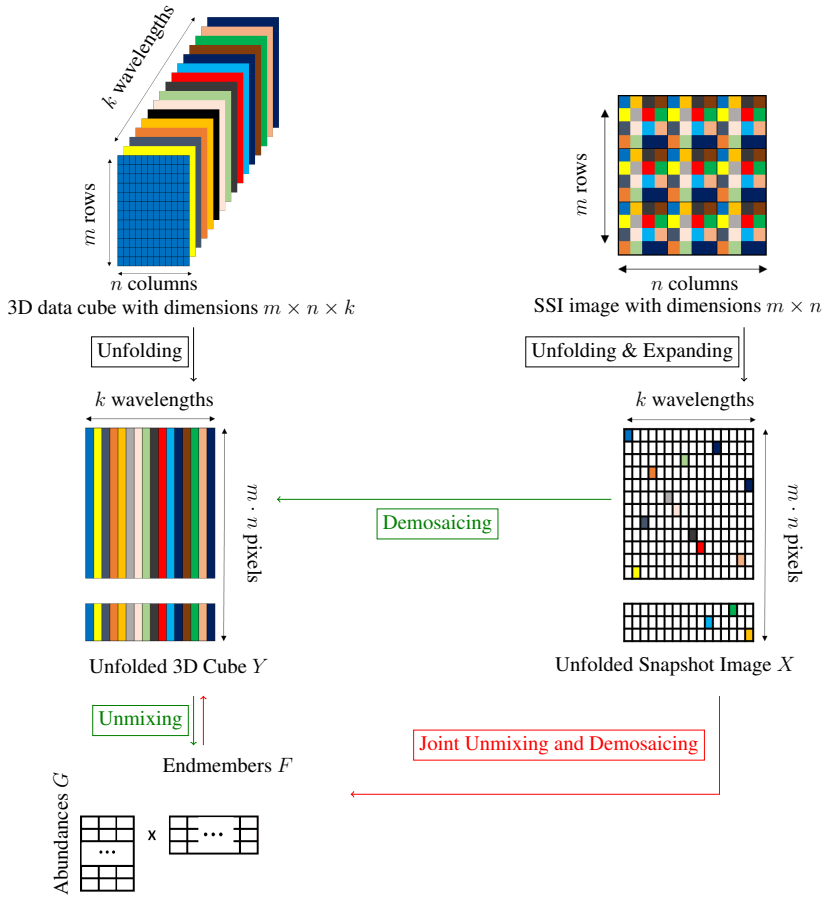


Figure 4.2 – Principles of the two-step (in green) and joint (in red) strategies.

Formally, an SSI camera captures a two-dimensional image consisting of $m \times n$ pixels for each exposure, where m and n represent the pixel counts in the horizontal and vertical dimensions, respectively. Each of the spatial pixels among the $m \cdot n$ available ones corresponds to a distinct spectral band from a total of k bands[2] expected to be acquired by the camera. More specifically, this signifies

[2]Typical values of k are 16 or 25, such that the patch is of size 4×4 or 5×5, respectively.

that an SSI image can be seen as a two-dimensional projection of a theoretical 3-D data cube with dimensions $m \times n \times k$, which we will now formalize and model.

Following an unfolding strategy commonly utilized in unmixing, we express the theoretical 3-D data cube as an $(m \cdot n) \times k$ matrix whose rows correspond to spatial positions while column indices are linked to wavelengths. As suggested in [143], it becomes feasible to create a data matrix of dimensions $(m \cdot n) \times k$ denoted as X, which contains missing entries. Both matrices X and Y are linked through

$$W \circ X = W \circ Y, \tag{4.1}$$

where W represents a binary weight matrix, with its non-zero entries indicating the wavelengths observed by the camera, Y is the unfolded 3-D data cube, X is the unfolded and expanded SSI image—so that X and Y share the same dimensions, see Fig. 4.2—and the symbol $\circ$ signifies the Hadamard product between matrices. Similarly, applying the unfolding process to the $m \times n$ SSI image, as we did for the data cube Y, results in an $(m \cdot n) \times 1$ vector denoted as $\underline{z}$. In this vector, the i-th entry corresponds to the non-null value found in the i-th row of X. Since the binary weight matrix W is known, obtaining X from $\underline{z}$ is a straightforward process and vice-versa. Hence, we can assume that X represents the original data matrix derived from the SSI acquisition process. The process of retrieving Y from X corresponds to "demosaicing" the SSI image. In practice, this can be addressed by, e.g., incorporating the assumption that Y is low-rank [143].

Additionally, we make the assumption that every row of Y can be represented as a linear mixture of the spectra corresponding to the materials observed by the camera—a.k.a. endmembers—i.e.,

$$Y \approx G \cdot F, \tag{4.2}$$

where F represents the $p \times k$ matrix of endmembers, G represents the $(m \cdot n) \times p$ abundance matrix, and p represents the number of endmembers present in the scene. Equation (4.2) is not only a very classical model met in hyperspectral unmixing [23] but also a low-rank approximation model, provided $p < \min\{(m \cdot n), k\}$. Combining Eqs. (4.1) and (4.2) provides the considered joint "demosaicing" and "unmixing" model, i.e.,

$$W \circ X \approx W \circ (G \cdot F). \tag{4.3}$$

Indeed, if it was possible to fully estimate both G and F from the partially observed matrix X, then their product $G \cdot F$ is an estimation of Y and one may derive a more accurate one as

$$\hat{Y} = W \circ X + (\mathbb{1}_{(m \cdot n) \times k} - W) \circ (G \cdot F), \tag{4.4}$$

where $\mathbb{1}_{(m \cdot n) \times k}$ represents the $(m \cdot n) \times k$ matrix of ones. Moreover, the information contained within G and F can be valuable for various applications, such as spectral library learning through the use of F, or for land use/cover analysis derived from G. Once more, it is important to emphasize our objective of comparing the performance between a two-stage strategy—including a demosaicing step where Y is estimated, followed by an unmixing step where G and F are derived (referred to as the green framework in Fig.4.2)—and a joint demosaicing and unmixing strategy (depicted in red in Fig.4.2). In Section 4.3, we introduce three different approaches for this joint strategy.

4.3 Proposed Methods

We now present our proposed approaches. In fact, we are proposing three joint demosaicing and unmixing methods. To start with, we introduce a "naïve" approach, derived from weighted NMF, following the strategy presented in [42] for a different application. Subsequently, we propose two novel methods, expressly developed for SSI data.

4.3.1 Naive Method

Initially, we introduce a naive method aimed at solving Eq. (4.3), which is a specific case of Weighted NMF (WNMF). Due to its similarity with low-rank matrix completion [41], the naive approach below can be seen as an alternative to [143] in which the structure of the low-rank matrix to recover is more interpretable. As introduced in Subsect. 3.4.2, WNMF can be addressed using various strategies, including (i) incorporating the weights into the update rules as proposed in [75], (ii) applying an Expectation-Maximization (EM) Framework as discussed in [41], or (iii) employing stochastic gradient descent while focusing on the available data points [89]. The second strategy, as identified in [41], was reported to be significantly faster and more accurate than the first one. Therefore, we choose this approach.

In the EM strategy, there are two steps—i.e., the E-step and the M-step—which are alternatingly and iteratively run. The E-step aims to estimate the expected complete matrix $\hat{Y}$ with respect to the known data $W \circ X$, the uncertain or unknown matrix data $(\mathbb{1}_{(m\cdot n)\times k} - W) \circ Y$, and the estimates of G and F at the $(t-1)$ iteration, denoted $\hat{G}^{(t-1)}$ and $\hat{F}^{(t-1)}$, respectively. Its solution reads [177]

$$\hat{Y} = W \circ X + (\mathbb{1}_{(m\cdot n)\times k} - W) \circ (\hat{G}^{t-1} \cdot \hat{F}^{t-1}). \tag{4.5}$$

The M-step then consists of applying standard NMF update rules to $\hat{Y}$ to derive $\hat{G}^t$ and $\hat{F}^t$. Once NMF converged to a given solution [177] or after a given number $\text{Max}_{\text{outIter}}$ of iterations [41], $\hat{Y}$ is updated in another E-step using the last estimates of G^t and F^t [177, 41]. In this work, we choose the Nesterov NMF (NeNMF) method [64] which is run until one of the following stopping criteria is reached: a maximum number of NMF iterations has been run—i.e., 1000 NMF iterations in our experiments—or the approximation error between the complete matrix derived in the E-step and its NMF approximation is below a threshold (i.e., 10^{-5} in our experiments). Meanwhile, when the abundance matrix $\hat{G}^t$ is updated, the sum to one constraint is satisfied by updating $\hat{Y}$ and $\hat{F}^t$ using the same strategy as in [47]. To apply this constraint, we modify the matrices $\hat{Y}$ and $\hat{F}$ by adding an additional column of ones. This modification allows us to control the Abundance Sum-to-One Constraint (ASC) using the parameter δ, which adjusts the influence of the sum-to-one constraint in the optimization process[3]. The augmented matrices $\bar{Y}$ and $\bar{F}$ are represented as follows:

$$\bar{Y} \triangleq \begin{bmatrix} \hat{Y} & \delta \mathbb{1}^T_{(m\cdot n)\times 1} \end{bmatrix}, \quad \bar{F} \triangleq \begin{bmatrix} \hat{F} & \delta \mathbb{1}^T_{p\times 1} \end{bmatrix} \tag{4.6}$$

[3] In our implementation, the value of δ is set to be 15, as this value has been found to provide an effective balance in the optimization process.

The whole method is provided in Subsect. 4.3.4.

4.3.2 Locally Rank-1 and Clustering-based Proposed Technique

We now introduce our first proposed method, which originates from the fundamental approach to restoring the data cube. Specifically, the image sensor is divided into patches that are replicated across the sensor surface. Each patch has dimensions $\sqrt{k} \times \sqrt{k}$, with common values for k being 16 or 25 [4]. Therefore, a patch typically measures either 4×4 or 5×5. In this context, it is reasonable to assume that each patch corresponds to a "super-pixel," meaning that each patch is associated with a unique endmember. In practice, such an assumption is not correct, which is why various demosaicing methods have been developed. Nevertheless, this assumption could still be valid for certain patches to find where one endmember significantly dominates over the others. This concept aligns with the pure-pixel assumption [23], or the concept of abundance sparsity in SCA, as described in [35]. However, the primary difference between our problem and the classical unmixing problem lies in the fact that we only have partial observations of the data cube within a patch, and our goal is to estimate it based on a limited number of available samples. This permits us to declare our first assumption.

Assumption 1 (Pure patch assumption)

For each endmember, there exists at least one sensor patch where only this endmember is present.

Our proposed method is outlined as follows. We denote by X_i, $\hat{Y}_i$, and W_i the $k \times k$ sub-matrices of X, $\hat{Y}$, and W, respectively, corresponding to Patch i. We obtain a rank-1 approximation of $\hat{Y}_i$ from X_i using the aforementioned WNMF strategy[5], which can be expressed as follows:

$$W_i \circ X_i \approx W_i \circ (\underline{g}_i \cdot f_i), \tag{4.7}$$

where $\underline{g}_i$ [6] represents a $k \times 1$ column vector and f_i represents a $1 \times k$ row vector.

When a patch contains only one dominant endmember, thereby satisfying Assumption 1, the patch itself becomes rank-1. Consequently, the rank-1 approximation derived from partial data in X_i enables us to estimate the endmember f_i. However, when a patch contains multiple endmembers, we must not detect the patch as pure. This leads us to our second assumption.

[4] As the patch size directly equals the number of wavelength bands, k is consistently used to denote both elements.

[5] We employ WNMF with a Nesterov solver, using Expectation Maximization on each patch, to calculate the rank-1 approximation.

[6] In pure patches, $\underline{g}_i$ is a vector of ones. While we could fix it and only estimate f_i, our proposed methods performed better when we also allowed $\underline{g}_i$ to be estimated.

Assumption 2

In the patches where several endmembers are present, their abundances should significantly vary over each patch.

Assumption 2 is classically stated in SCA [35]. In practical situations, it is expected that the SSI camera is positioned close enough to the observed scene, so that one may not expect multiple abundances to remain in constant proportions over a patch. Consequently, if the i-th patch under consideration is approximately pure, then

$$\|W_i \circ X_i - W_i \circ (\underline{g}_i \cdot f_i)\|_F^2 \approx 0. \tag{4.8}$$

Conversely, if this patch is not pure, then

$$\|W_i \circ X_i - W_i \circ (\underline{g}_i \cdot f_i)\|_F^2 \gg 0. \tag{4.9}$$

In this context, this type of error can be regarded as a "single-source confidence measure," corresponding to those commonly employed in SCA. Hence, from each patch, we obtain a tentative estimate of a "true" endmember. The retrieved spectra that satisfy Eq. (4.8) are gathered and organized into a matrix denoted as $\mathbb{F}$. These estimates are further assumed to be organized as clusters of spectra distributed around the "true" source spectra. Therefore, we can employ any clustering technique, such as K-means or K-medians, initialized with K-means++ [18], to extract the actual endmembers. A refined approach involves identifying the patches where the squared Frobenius norm mentioned above is small-enough[7]. This results in a smaller set of spectra, each of which is closer to the "true" ones. This approach is similar to the *Selective* K-means or K-medians methods proposed in [125] and was found to notably enhance the performance of our proposed method preliminary test[8].

Once the true endmembers are extracted and stored in the matrix F, the abundance matrix G is re-estimated using the Naive method described in Subsect. 4.3.1, with fixed F and the sum-to-one constraint on the abundance matrix G in the M-step. The whole strategy is provided in in Subsect. 4.3.4, and we name it K-means (resp. K-medians) Patch-based Weighted Nonnegative Matrix Factorization (KPWNMF). It is worth noting that the abundance estimation can also be accomplished using the least-squares method, similar to what we applied in [4]. However, our experiments discovered that running the Naive WNMF yields more accurate results.

[7] In practice, we keep all the patches where the rank-one approximation error is below a user-defined threshold. Estimating the optimal threshold value is out of the scope of this work. In the experiments provided in Chapter 6, it is set as the median of all the patch norm errors.

[8] It is noteworthy that K-medians with ℓ_1 norm is used for the clustering stage because it performs slightly better than K-means. It tries to find p median points in the data such that the ℓ_1 distances from each data point to its closest median are minimized. Moreover, the clustering process is repeated several times—e.g., 10 times—and the solution with the lowest within-cluster sums of point-to-centroid distances is selected as the optimal solution.

4.3.3 Method with Relaxed Abundance Sparsity Assumption

We now introduce our third approach. It may be seen as an extension of the previous one as it also assumes Assumption 1 to be valid. However, it significantly relaxes Assumption 2 which is replaced by the following one.

Assumption 3

In the patches where several endmembers are present, their abundances may or may not vary over each patch.

This assumption[9] states that in the patches where multiple endmembers are present, we do not require any constraint on their abundances. Consequently, Eq. (4.8) is satisfied if the patch is pure (Assumption 1) or, when several endmembers are present in a patch, their abundances do not vary within the considered patch. Still, Eq. (4.9) occurs in patches where several endmembers are present and have their abundance proportions to vary over the considered patches. As a consequence, the KPWNMF method cannot be applied with such assumptions and a refined strategy must be proposed.

As for KPWNMF, we collect all the spectra f_i which are estimated in patches where Eq. (4.8) holds[10] and we arrange them in a matrix denoted $\mathbb{X}$. As explained above, each row vector of this matrix is either an estimate of an endmember of F or a mixture of them. As we assumed an LMM, this matrix can be written as

$$\mathbb{X} = \mathbb{G} \cdot F. \tag{4.10}$$

Due to Assumption 1, each row of F exists at least once in $\mathbb{X}$ and the estimation of F can be done by applying a pure-pixel-based method[11], e.g., VCA [110] or SPA [60]. Moreover, $\mathbb{G}$ represents a specific abundance matrix associated with the spectral data in $\mathbb{X}$. It differs from the final abundance matrix G but is derived by selecting rows corresponding to the chosen patch indices in $\mathbb{X}$.

Once the true endmembers are extracted and stored in the matrix F, the abundance matrix G is re-estimated using the Naive method described in Subsect. 4.3.1, with fixed F and the sum-to-one constraint on the abundance matrix G in the M-step. The whole strategy is provided in in Subsect. 4.3.4, and we name it VCA Patch-based Weighted Nonnegative Matrix Factorization (VPWNMF).

In conclusion, we proposed two frameworks, KPWNMF and VPWNMF. They both use spectra estimated in patches where the rank-1 approximation error is low, making their first stages similar. However, they differ in the final stage:

[9] Stating Assumption 3 is not strictly necessary, as we permit the selected rank-1 patches to be either pure or mixed under certain constraints. This approach aligns the processing of the extracted endmembers with pure-pixel-based unmixing methods. Additionally, Assumption 3 can be interpreted as the negation of Assumption 2. However, including Assumption 3 clarifies that we no longer utilize a sparse component analysis framework.

[10] As for KPWNMF, we only keep spectra estimated in patches where the approximation error is below the median of all the patch norm errors.

[11] We selected VCA because it performed slightly better than the other methods in some preliminary tests.

- Thanks to Assumption 2, we know that all the spectra kept in $\mathbb{F}$ correspond to tentative estimates of the actual endmembers in F. Consequently, we can apply any clustering technique to derive F from $\mathbb{F}$, as it is classically done in SCA.
- When Assumption 2 is not met—i.e., when Assumption 3 is stated—some kept spectra might be linear mixtures of the actual endmembers. This is why we here named the pool of kept spectral $\mathbb{X}$. We could propose an outlier robust clustering method to process these data to derive F, provided the number of mixed spectra in $\mathbb{X}$ is low. Instead, as its assumptions are satisfied in $\mathbb{X}$ and as it is not sensitive to the number of mixed spectra, we use VCA instead.

4.3.4 Algorithms

This section introduces the pseudo code for our three proposed methods. While these algorithms share several procedural steps, they fundamentally differ in their methodologies, as previously discussed. Algorithm 1 details our Naive WNMF approach, serving as the fundamental framework for the following methodologies. Algorithm 2 introduces the shared patch processing steps essential to both the K-means Patch-based WNMF (KPWNMF) and the VCA Patch-based WNMF (VPWNMF), focusing on the extraction of tentative endmembers. Building upon this, Algorithm 3 and Algorithm 4 diverge to apply specific endmember extraction techniques needing Assumption 2 or Assumption 3, respectively.

Algorithm 1 Naive Method

Input:
X represents the unfolded SSI image with a rank of p and is weighted by matrix W.
Output:
Y signifies the reconstructed unfolded data cube.
$[G, F]$ represents the final abundances and endmembers.
Processing:

1: **Initialize** G and F
2: **for** $\text{Counter}_1 = 1$ **to** $\text{Max}_{\text{outerIter}}$ **do**
3: $\quad Y = W \circ X + (\mathbb{1} - W) \circ (G \cdot F)^{\text{Counter}_1 - 1}$
4: $\quad$ **for** $\text{Counter}_2 = 1$ **to** $\text{Max}_{\text{innerIter}}$ **do**
5: $\quad\quad$ **Update** G from Y and F using NeNMF
6: $\quad\quad$ **Update** F from Y and G using NeNMF

Algorithm 2 Rank-one patch detection and spectra estimation method used in both proposed KP-WNMF and VPWNMF

Inputs:

X represents the unfolded SSI image with a rank of p and is weighted by matrix W.
nb_patches denotes the number of patches to be processed.

Output:

Matrix $\mathbb{M}$ containing the "best" vectors—according to Eqs. (4.8) and (4.9)—
for further processing.

Processing:

for $i = 1$ **to** nb_patches **do**
  Let the submatrices X_i and W_i linked to Patch i
  Initialize $\underline{g}_i$ and f_i
  for $t = 1$ **to** $\text{Max}_{\text{outerIter}}$ **do**
    $X_i^{\text{Comp}} = W_i \circ X_i + (\mathbb{1}_{k\times k} - W_i) \circ (\underline{g}_i \cdot f_i)$
    for Counter $= 1$ **to** $\text{Max}_{\text{innerIter}}$ **do**
      Update $\underline{g}_i$ from X_i^{Comp} and f_i using NeNMF
      Update f_i from X_i^{Comp} and $\underline{g}_i$ using NeNMF
Keep the 50% best vectors f_i—according to Eqs. (4.8) and (4.9)—and organize them into matrix $\mathbb{M}$
return $\mathbb{M}$

4.4 Conclusion

In this chapter, we aimed to perform "demosaicing" and "unmixing" jointly for the hyperspectral images acquired by the SSI camera. Therefore, we proposed two novel approaches in addition to the naive method derived from Weighted NMF (WNMF). The first one KPWNMF (K-means Patch-based Weighted Nonnegative Matrix Factorization) assumes that the abundances are sparse in a few patches to find, so that each of these patches is dominated by one endmember. Such an assumption is similarly met with SCA, except that we consider partially observed data. In contrast, the second one VPWNMF (VCA Patch-based Weighted Nonnegative Matrix Factorization) relaxes the sparsity assumption needed in the latter. The performance of these proposed methods is assessed in Chapter 6, where we comprehensively evaluate their effectiveness through various experiments. In the next chapter, we will introduce our filtering-based framework to solve joint unmixing and demosaicing problems. This framework is designed specifically for applications involving Fabry-Perot filters or MSFA sensors with less selective filters, addressing complex scenarios often encountered in practice.

Algorithm 3 K-means (resp. K-medians) Patch-based Weighted Nonnegative Matrix Factorization (KPWNMF)

Inputs:
Matrix $\mathbb{F}$ the output from Algorithm 2.
Rank p.
Outputs:
$\hat{Y}^{\text{final}}$ signifies the reconstructed unfolded data cube.
$[G, F]$ represents the final abundances and endmembers.
Processing:

1: Call Algorithm 2 to get $\mathbb{F}$
2: $F = \text{K-means}(\mathbb{F}, p)$ (resp. $F = \text{K-medians}(\mathbb{F}, p)$)
3: **Initialize** G of size $(m \cdot n) \times p$
4: **Compute** G and $\hat{Y}^{\text{final}}$ using Algorithm. 1 with fixed F

Algorithm 4 VCA Patch-based Weighted Nonnegative Matrix Factorization (VPWNMF)

Inputs:
Matrix $\mathbb{X}$ the output from Algorithm 2.
Rank p.
Outputs: $\hat{Y}^{\text{final}}$ signifies the reconstructed unfolded data cube.
$[G, F]$ represents the final abundances and endmembers.
Processing:

1: Call Algorithm 2 to get $\mathbb{X}$
2: $F = \text{VCA}(\mathbb{X}, p)$
3: **Initialize** G of size $(m \cdot n) \times p$
4: **Compute** G and $\hat{Y}^{\text{final}}$ using Algorithm. 1 with fixed F

Chapter 5

Locally-Rank-One-Based Joint Unmixing and Demosaicing Methods for Snapshot Spectral Images: a Filtering-Based Framework

"You never change things by fighting the existing reality. To change something, build a new model that makes the existing model obsolete."

— Buckminster Fuller

5.1 Introduction

In Chapter 4, we proposed joint demosaicing and unmixing methods for snapshot spectral imaging (SSI) systems. In addition to a naive approach—straightforwardly derived from Weighted Nonnegative Matrix Factorization (WNMF)—we proposed two novel demosaicing methods, i.e., KPWNMF (K-means Patch-based Weighted Nonnegative Matrix Factorization) and VPWNMF (VCA Patch-based Weighted Nonnegative Matrix Factorization). These methods incorporated the assumption of sparse abundances within sensor "patches," where a single endmember primarily dominates each patch. The methods employed rank-1 WNMF computations, a specialized single-source confidence measure, endmember extraction, and abundance estimation. However, they diverged in treating scenarios involving multiple endmembers, providing distinct solutions to address such cases.

Building upon these methods, in this chapter[1], we further analyze the pipeline to restore the HSI data cube—introduced previously in Sect. 1.4 and illustrated in Fig. 1.12—and explore the utilization of the harmonic response matrix of Fabry-Perot filters (FPf) for joint demosaicing, deconvolution and unmixing. Our objective is to demonstrate that integrating these techniques not only

[1] Methods introduced in this chapter were published in [7, 2, 5]

leads to enhanced outcomes compared to the separate stages of demosaicing, unmixing, and spectral correction—commonly referred to as deconvolution—but also simplifies the overall processing pipeline by embedding the spectral correction within the joint framework. Furthermore, the current image processing pipeline of snapshot images involves applying a correction matrix after demosaicing in order to remove some unwanted harmonics. In contrast, our proposed approaches obviate the need for applying the correction matrix and initiates the deconvolution process directly from the raw SSI image. Through this novel methodology, we aim to demonstrate the efficacy of our approaches in achieving improved unmixing and demosaicing results while simplifying the overall processing pipeline.

It is worth noting the differences between the two frameworks. Firstly, in Chapter 4, we established a general framework for demosaicing images. It is broad in scope and potentially applicable to various Multispectral Filter Arrays (MSFAs) [90]. This framework sets the foundation for our discussion on the joint methods for conducting "demosaicing" and "unmixing" for hyperspectral images captured by SSI cameras. Based on prior findings, our investigation here is driven by the hypothesis that incorporating low-rank matrix completion and factorization together is more efficient than a two-stage approach involving low-rank matrix completion followed by matrix factorization.

Building on the general framework established in Chapter 4, this chapter explores the pipeline provided by the camera manufacturer, focusing on new unmixing approaches for RAW SSI images that consider this pipeline. This framework is designed specifically for applications involving Fabry-Perot filters or any MSFA sensors where the filters are not selective enough. The differentiation between the two frameworks lies in their focus and applicability: while the first one lays the groundwork for a general approach, the second one works on specific complex scenarios often encountered in practice.

5.2 Problem Statement

In this section, we provide a definition of the SSI acquisition system and outline the problem that we aim to address. Let us first recall that the SSI camera acquires a two-dimensional image consisting of $m \times n$ pixels for each exposure, where m and n represent the pixel counts in the horizontal and vertical dimensions, respectively. Additionally, it is assumed that the camera observes k spectral bands. In practice, the SSI technology relies on a mosaic of Fabry-Perot filters comprised of $\sqrt{k} \times \sqrt{k}$ patches that are replicated across the sensor surface. In an ideal scenario, an FPf permits only light within a minimal spectral range to reach the sensor while blocking light outside this range. However, in real implementations, these filters exhibit additional harmonics around each wavelength of interest, as illustrated in Fig. 5.1. Fortunately, these filters are known and provided by the camera manufacturer [53].

For the rest of this section, our focus will be on a single patch of Fabry-Perot filters. Denoting $x_i(\lambda_i)$ as the i-th SSI pixel in the patch, which theoretically captures spectral information at λ_i nm, we have:

$$x_i(\lambda_i) = \sum_{j=1}^{k} h_i(\lambda_j) \cdot y_i(\lambda_j) + \omega_i, \tag{5.1}$$

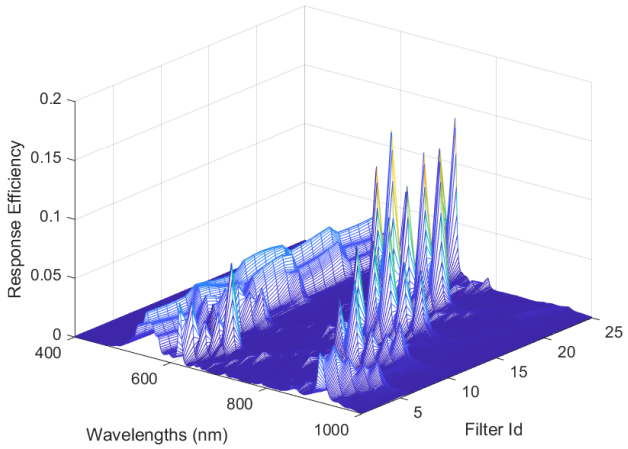


Figure 5.1 – The spectral response of the 25 spectral filters of the 5×5 mosaic Photon Focus SSI camera covers the wavelength range from 400 to 1000 nm [119].

where $h_i(\lambda)$ represents the Fabry-Perot filter associated with Pixel i, $y_i(\lambda)$ is the actual spectrum intended to be observed by Pixel i, and ω_i denotes some additive noise. Furthermore, assuming a linear mixture model, the observed spectrum can be expressed as a combination of endmembers, i.e.,

$$y_i(\lambda) = \sum_{l=1}^{p} g_{il} f_l(\lambda), \tag{5.2}$$

where p represents the number of endmembers in the observed scene, $f_l(\lambda)$ signifies the spectrum of the l-th endmember, and g_{il} is the corresponding abundance proportion in Pixel i, i.e.,

$$\forall l = 1, \ldots, p,\ 0 \leq g_{il} \leq 1 \quad \text{and} \quad \sum_{l=1}^{p} g_{il} = 1. \tag{5.3}$$

By combining Eqs. (5.1) and (5.2), we obtain

$$x_i(\lambda_i) = \sum_{l=1}^{p} g_{il} \left(\sum_{j=1}^{k} h_i(\lambda_j) \cdot f_l(\lambda_j) \right) + \omega_i. \tag{5.4}$$

In this chapter, we aim to estimate the p endmembers $f_l(\lambda)$ with their associated abundance coefficients g_{il} using Eqs. (5.4).

5.3 Proposed methods

We now introduce our proposed methods. We actually propose two novel methods for SSI data exploiting the filter characteristics and using the same sparsity assumptions introduced in Chapter 4. The latter are briefly recalled in the subsections below. Before introducing these methods, we re-emphasize the classical strategy provided by the camera manufacturer.

5.3.1 Classical Strategies

The traditional approach recommended by the camera manufacturer involves applying demosaicing, deconvolution, and then unmixing. The deconvolution step corrects the spectral distortions caused by the sensor and filters, typically using a correction matrix provided by the manufacturer. To estimate the endmembers, one can apply a naive approach starting from Eqs. (5.4). We can define:

$$\tilde{f}_l(\lambda) = \sum_{j=1}^{k} h_i(\lambda_j) \cdot f_l(\lambda_j) \tag{5.5}$$

where $\tilde{f}_l(\lambda)$ represents the convolution of the filter response with the endmember spectrum.

We can now write Eqs. (5.4) as follows:

$$x_i(\lambda_i) = \sum_{l=1}^{p} g_{il} \tilde{f}_l(\lambda_i) + \omega_i. \tag{5.6}$$

$\tilde{f}_l(\lambda)$ can be estimated using the approaches from Chapter 4. Then, deconvolve it to retrieve the original endmember spectra $f_l(\lambda)$ using the inversion matrix proposed by IMEC.

5.3.2 Clustering and Filtering-based Proposed Technique

We now introduce our approach and assumptions to solve the problem described in the previous section. Based on Eq. (5.1), the set of k observed SSI values and the k complete spectra of the patch can be expressed:

$$\underline{x} \triangleq \begin{bmatrix} x_1(\lambda_1), \\ \vdots \\ x_k(\lambda_k) \end{bmatrix}, \quad Y \triangleq \begin{bmatrix} y_1(\lambda_1) & \dots & y_1(\lambda_k) \\ \vdots & & \vdots \\ y_k(\lambda_1) & \dots & y_k(\lambda_k) \end{bmatrix}. \tag{5.7}$$

where $\underline{x}$ is defined as a k-dimensional vector representing spectral values of one SSI patch across k wavelengths, while Y is a matrix representing the full spectral values for this patch. Each row in Y corresponds to a different pixel, and each column represents a specific wavelength. The dimensions of Y are $k \times k$, where the first dimension represents the number of pixels and the second represents the number of spectral bands.

Estimating Y from $\underline{x}$ and Eq. (5.1) is an ill-posed problem. Therefore, we consider the same set of assumptions that we presented in Chapter 4.

In a patch where Assumption 1 is verified, the matrix Y is *approximately* rank-1 and can be replaced by a vector $\underline{y} = [y(\lambda_1), \dots, y(\lambda_k)]$ corresponding to an endmember. By noting $\underline{\omega} \triangleq [\omega_1, \dots, \omega_k]^T$ the vector associated with the noise in the patch and

$$\mathbb{H} \triangleq \begin{bmatrix} h_1(\lambda_1) & \dots & h_1(\lambda_k) \\ \vdots & & \vdots \\ h_k(\lambda_1) & \dots & h_k(\lambda_k) \end{bmatrix} \tag{5.8}$$

is the matrix of Fabry-Perot filters (response matrix), Eq. (5.1) can be written as

$$\underline{x} = \mathbb{H} \cdot \underline{y}^T + \underline{\omega}. \tag{5.9}$$

For each patch, we first aim to recover a tentative spectrum $\underline{y}$ from Eq. (5.9). In practice, as the matrix $\mathbb{H}$ can be ill-conditionned[2], we aim to solve a penalized optimization problem, i.e.,

$$\min_{\underline{y} \geq 0} \frac{1}{2} \|\underline{x} - \mathbb{H} \cdot \underline{y}^T\|_2^2 + \frac{\alpha}{2} \|D \cdot \underline{y}^T\|_2^2, \tag{5.10}$$

where D is a square matrix that accounts for the discrete derivative of the spectrum $\underline{y}$, and α represents the penalization term. Eq (5.10) is a quadratic problem which can be rewritten as [30]

$$\min_{\underline{y} \geq 0} \frac{1}{2} \left\| \begin{pmatrix} \underline{x} \\ 0 \end{pmatrix} - \begin{pmatrix} \mathbb{H} \\ \sqrt{\alpha}\, D \end{pmatrix} \cdot \underline{y}^T \right\|_2^2. \tag{5.11}$$

Using the same assumptions as in the previous chapter, we look for patches with low-error rank-one approximations$\|\underline{x} - \mathbb{H} \cdot \underline{\hat{y}}\|_2$, where $\underline{\hat{y}}$ is the estimated spectrum obtained from Eqs. (5.10) or (5.11).

If Assumption 2 is satisfied, then all the selected vectors are the tentative rows of F, and we can apply any clustering technique. As for KPWNMF, we here choose *Selective* K-means or K-medians proposed in [125], initialized with K-means++ [18][3].

Once the actual endmembers are derived and stored in the matrix F, a final step involves estimating the abundance in each pixel of the SSI image. We adopt the low-rank matrix completion framework introduced in Chapter 4 for this purpose. At the patch level, we consider the observed values $y_i(\lambda_i)$ to be part of a partially observed $k \times k$ matrix Y. Denoting S as the matrix of endmembers convolved by $\mathbb{H}$, we obtain the matrix form of Eq. (5.4), i.e.,

$$\mathbb{H} \circ Y \approx \mathbb{H} \circ (G \cdot S), \tag{5.12}$$

where G is the matrix of abundances in the considered patch, and $\mathbb{H}$ is a scaled version of H. In practice, we initialize G through least-squares regression using F and the patch demosaiced by Weighted Bilinear interpolation (WB) approach [27]. Subsequently, we aggregate all these matrices Y to update the abundance matrix using Naive WNMF in Chapter 4 globally. The whole strategy is provided in Subsect. 5.3.4, and we name it Filter Patch-based Kmeans (resp. K-medians), (FPKmeans).

5.3.3 Method with Relaxed Abundance Sparsity Assumption

We now introduce our second approach. It may be seen as an extension of the previous one as it also assumes Assumption 1 to be valid. However, it significantly relax Assumption 2 which is replaced by the following Assumption 3.

[2]For example, the 5×5 filter matrix $\mathbb{H}$ using real FPf in [143] has two rows which are almost null.

[3]Similarly to KPWNMF, this method utilizes K-medians clustering with the ℓ_1 norm. The clustering process is repeated 10 times, and the solution with the lowest within-cluster sums of point-to-centroid distances is selected as the optimal solution.

This assumption states that in the patches where multiple endmembers are present, we do not require any constraint on their abundances. Consequently, the approximation error $\|\underline{x} - \mathbb{H} \cdot \underline{\hat{y}}\|_2$ tends to be low when a patch is pure (in accordance with Assumption 1) or when the abundances of multiple endmembers remain consistent within the patch. However, the approximation error increases significantly in patches with multiple endmembers and their abundance proportions vary across the patch. As for FPKmeans, we collect the proportion[4] ρ of tentative endmembers in the "best" patches—i.e., those which provide the smallest approximation errors—and we arrange them in a matrix denoted $\mathbb{Y}$. As explained above, each row vector of this matrix is either an estimate of an endmember of F or a mixture of them. As we assumed a linear mixture model (LMM), this matrix can be written as

$$\mathbb{Y} = \mathbb{G} \cdot F. \tag{5.13}$$

Then, as we did in VPWNMF in Chapter 4, VCA is used to extract the final set of endmembers. Finally, the abundance matrix G is estimated using the same approach as in FPKmeans. The whole strategy is provided in Subsect. 5.3.4, and we name it Filter Patch-based Vertex Component Analysis (FPVCA)).

5.3.4 Algorithms

This section introduces the pseudo code for our proposed methods. While these algorithms share several procedural steps, they fundamentally differ in their methodologies as previously discussed. Algorithm 5 introduces the shared patch processing steps essential to both the Filter Patch-based Kmeans (FPKmeans) and the Filter Patch-based Vertex Component Analysis (FPVCA), focusing on the extraction of tentative endmembers. Building upon this, Algorithm 6 and Algorithm 7 diverge to apply specific endmember extraction techniques needing Assumption 2 or Assumption 3, respectively.

[4] As for KPWNMF, we only keep spectra estimated in patches where the approximation error is below the median of all the patch norm errors, i.e., $\rho = 0.5$.

Algorithm 5 Rank-one patch detection and spectra estimation method used in both proposed FP-Kmeans and FPVCA

Inputs:

The SSI matrix

p: the number of endmembers

$\mathbb{H}$: the response matrix

α: the regularization parameter

nb_patches: the number of patches to be processed

Output:

- Matrix $\mathbb{M}$ containing the "best" vectors for further processing.

Processing:

1: **for** $r = 1$ **to** nb_patches **do**
2: Let $\underline{x}_r$ be the SSI vector linked to Patch r
3: Estimate $\hat{\underline{y}}$ using Eq (5.11)
4: Keep the 50% best vectors $\hat{\underline{y}}$—according to $\|\underline{x}_r - \mathbb{H}\hat{\underline{y}}\|_2$—and organize them as a matrix $\mathbb{M}$
5: **return** $\mathbb{M}$

Algorithm 6 Filter Patch-based Kmeans (resp. K-medians), (FPKmeans)

Inputs:
Matrix $\mathbb{F}$ from Algorithm 5
p: the number of endmembers

Outputs:
$[G, F]$ represents the final abundances and endmembers.

Processing:
1: Call Algorithm 5 to obtain $\mathbb{F}$
2: $F = \text{K-means}(\mathbb{F}, p)$ (resp. $\text{K-medians}(\mathbb{F}, p)$)
3: **Compute** G using Eq. (5.12) and Naive WNMF

Algorithm 7 Filter Patch-based Vertex Component Analysis (FPVCA)

Inputs:
Matrix $\mathbb{Y}$ from Algorithm 5
p: the number of endmembers

Outputs:
$[G, F]$ represents the final abundances and endmembers.

Processing:
1: Call Algorithm 5 to obtain $\mathbb{Y}$
2: $F = \text{VCA}(\mathbb{Y}, p)$
3: **Compute** G using equation (5.12) and Naive WNMF

5.4 Conclusion

In this chapter we focused on the Fabry-Perot filter response to perform the joint demosaicing and unmixing. Specifically, we performed joint demosaicing, deconvolution and unmixing, where we integrated deconvolution directly into the unmixing process. This integration effectively utilizes the harmonic responses of Fabry-Perot filters, enhancing the accuracy and efficiency of the spectral data processing. We proposed two novel methods for this goal: FPVCA (Filter Patch-based Vertex Component Analysis) and FPKmeans (Filter Patch-based Kmeans). These methods leverage Fabry-Perot filter deconvolution and extend the "pure pixel" framework to the SSI sensor patch level, enabling improved unmixing accuracy and introducing the concept of localized spectral purity. It is worth noting that our proposed methods eliminate the need for the manufacturer correction matrix, as the

deconvolution is integrated directly into the framework. This integration enhances the overall spectral accuracy and efficiency. The performance of these proposed methods is assessed in Chapter 6, where we comprehensively evaluate their effectiveness through various experiments.

Chapter 6

Experimental Validation

"It does not matter how beautiful your theory is, it does not matter how smart you are. If it does not agree with experiment, it is wrong."

— Richard P. Feynman

6.1 Introduction

In this chapter, we assess the performance of the proposed frameworks for joint demosaicing and unmixing of snapshot spectral images: Matrix-Completion and Filtering-based frameworks. The evaluation is conducted using synthetic and real images captured by real snapshot cameras. We also examine the performance under various noisy conditions to ensure the robustness of our methods. Our experimental results show that the proposed methods outperform the two-stage approaches, which consist of applying demosaicing and, after that, unmixing.

6.2 Experimental Validation of Matrix-Completion Framework

6.2.1 Experimental Setup

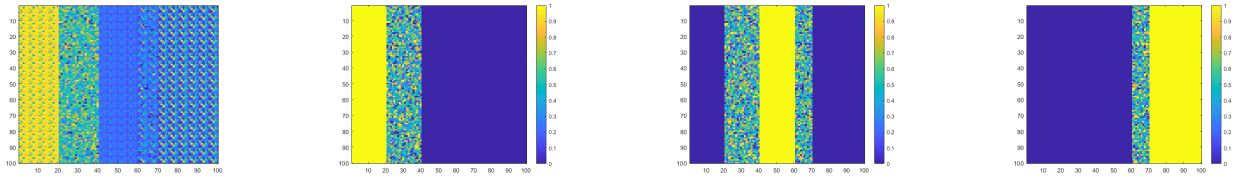
To evaluate the effectiveness of the proposed methods, we carry out experiments using SSI simulations generated from synthetic images as well as real SSI images captured by SSI cameras. For the synthetic images, we did two experiments. In the first one, we create one image where assumptions one and two are satisfied, and another where assumptions 1 and 3 are satisfied. Each one has 100×100 pixels with three endmembers, i.e., water, metal, and concrete, whose signatures are taken from [144].

We compare the performance reached with the naive method with our proposed approaches KP-WNMF and VPWNMF and seven 2-step demosaicing-then-unmixing methods. For the latter, we consider seven SotA demosaicing methods—i.e., GRMR [143], BTES [105], WB [27], PPID [107], ItSD [108], SAND [22], and PCWB [131]—while in the second step, we unmix the restored data

cube Y using VCA for estimating the endmembers and Fully Constrained Least Squares (FCLS) for abundance estimation. To measure the effectiveness of the tested methods, we assess their demosaicing improvements by comparing the estimated Y matrices to the ground truth, utilizing the Peak Signal-to-Noise Ratio (PSNR). For unmixing enhancement, we employ the Signal-to-Interference Ratio (SIR), Spectral Angle Mapper (SAM) for endmember estimation, and Mixing Error Ratio (MER) to measure the quality of abundance maps. We additionally incorporated RMSE for abundance estimation accuracy and running time in seconds to assess computational efficiency

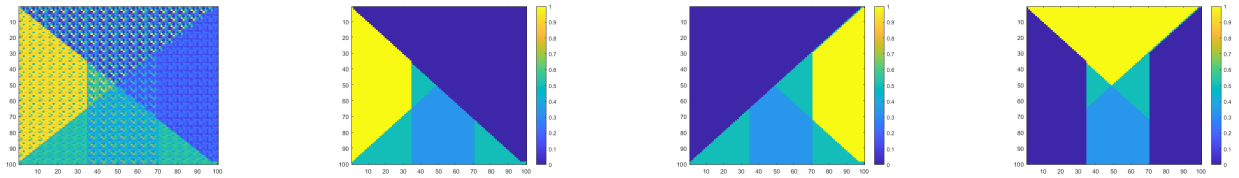
For the second experiment, we aim to validate the demosaicing enhancement provided by the proposed approaches on more challenging data. We provide comparative results with the SoTA methods on the CAVE dataset after simulating the 4×4 and 5×5 SSI images using the same strategy as in [143] with all the images in the dataset. Moreover, using real multispectral images allow us to experimentally investigate the validity of the stated assumptions.

Finally, we used real SSIs taken from the Hyko 2 dataset [155]. The images in the dataset are captured using two snapshot mosaic cameras. It is the first dataset to capture hyperspectral data from a moving vehicle, enabling hyperspectral scene analysis for road scene understanding. The data span the visible and near-infrared spectral ranges, from 400 to 1000 nm.



(a) SSI Image (b) Water (c) Metal (d) Concrete

Figure 6.1 – Image 1, SSI image where assumptions 1 and assumption 2 are satisfied with abundance maps of the three endmembers.



(a) SSI Image (b) Water (c) Metal (d) Concrete

Figure 6.2 – Image 2, SSI image where assumptions 1 and assumption 3 are satisfied with abundance maps of the three endmembers.

Table 6.1 – PSNR, SAM, SIR, MER, RMSE, and Time in seconds obtained for the synthetic images with 5x5 and 4x4 (into brackets) ideal filters. In bold, the highest performance value and for RMSE and Time, the lowest value.

Method	Image 1 (Assumptions 1 & 2)						Image 2 (Assumptions 1 & 3)					
	PSNR	SAM	SIR	MER	RMSE	Time	PSNR	SAM	SIR	MER	RMSE	Time
GRMR	27.2 (27.5)	0.3 (0.3)	14.3 (65)	2.1 (1.1)	0.4 (0.4)	1.6 (1.57)	29.7 (33.2)	0.4 (0.2)	12.1 (67)	4.9 (4)	0.3 (0.3)	1.5 (0.3)
BTES	24.1 (24.2)	0.4 (0.3)	13.41 (65.8)	-1.2 (-5.4)	0.5 (0.5)	**0.2 (0.3)**	30 (29.1)	0.4 (0.3)	12.3 (65.5)	-4.5 (-2)	0.4 (0.5)	0.4 (0.3)
WB	26.3 (28.9)	0.3 (0.2)	15.6 (70.2)	-0.45 (1.8)	0.4 (0.4)	**0.2 (0.3)**	30.2 (31.2)	0.4 (0.3)	18.2 (63.2)	4.8 (1.5)	0.3 (0.2)	0.3 (0.3)
PPID	30.3 (30.2)	0.06 (0.06)	34.7 (85.9)	9.9 (9.8)	0.3 (0.3)	0.3 (0.3)	34.8 (38)	0.08 (0.05)	31.7 (86)	9.2 (14.9)	0.3 (0.2)	0.3 (0.3)
ItSD	24.5 (23.9)	0.7 (0.6)	8.2 (55.7)	0.5 (-7.4)	0.4 (0.4)	**0.2 (0.3)**	26.4 (26.1)	0.4 (0.5)	8 (56)	-2.4 (-2.08)	0.4 (0.2)	0.3 (0.3)
SAND	30 (30.1)	0.1 (0.1)	31.2 (80.1)	11.3 (9.1)	0.2 (0.2)	3720 (2800)	35.4 (36.9)	0.1 (0.1)	32 (82)	10 (13.2)	0.2 (0.1)	3660 (2500)
PCWB	28.6 (20.3)	0.1 (0.2)	31.7 (70.3)	9.7(9.4)	0.2 (0.3)	0.2 (0.2)	31 (20.2)	0.4 (0.3)	15.5 (71.8)	1.2 (4)	0.4 (0.2)	0.2 (0.2)
Naive	32.5 (32.7)	0.2 (0.2)	19.8 (68.7)	9.2 (8.5)	0.3 (0.3)	15 (12)	34.5 (34.4)	0.2 (0.3)	21.5 (67.2)	10.7 (5.4)	0.3 (0.2)	15 (12)
KPWNMF	**36.5 (35.5)**	0.01 (0.007)	50.2 (103.5)	**16.0 (14.8)**	**0.2(0.2)**	6 (6)	40.3 (40.4)	0.003 (0.006)	50.3 (102.5)	20.9 (19.7)	**0.1 (0.1)**	6 (6)
VPWNMF	35.9 (33.7)	**0.002 (0.003)**	**59.9 (111.0)**	14.7 (12.8)	**0.2(0.2)**	5 (5)	**41.4 (40.8)**	**0.0006 (0.001)**	**75.8 (117.0)**	**21.0 (19.7)**	**0.1 (0.1)**	5 (5)

6.2.2 Performance Evaluation on Synthetic Images For The Ideal Case

In the first set of experiments, we generated an image (available on Fig. 6.1) that fulfills Assumptions 1 and 2, and another image (see Fig. 6.2) that satisfies Assumptions 1 and 3. Both images have a resolution of 100×100 pixels and are composed of three endmembers—i.e., water, metal, and concrete—with signatures sourced from [144]. In this simulation, we consider the ideal case where each filter associated with each detector element corresponds to an ideal filter with perfect cut-off characteristics, which allow light from a single wavelength to be captured by the detector element. Both images simulate a scene observed from a short distance, implying that many patches are pure. We consider 4×4 and 5×5 spectral filter patterns and investigate the performance of the tested methods under different noise levels.

The PSNR, SAM, SIR, and SAM achieved by all the methods for 4×4 and 5×5 spectral filter patterns are reported in Table 6.1. While the average performance with different noise levels for both images with 4×4 and 5×5 filters is presented in Fig. 6.3 and Fig. 6.4, respectively.

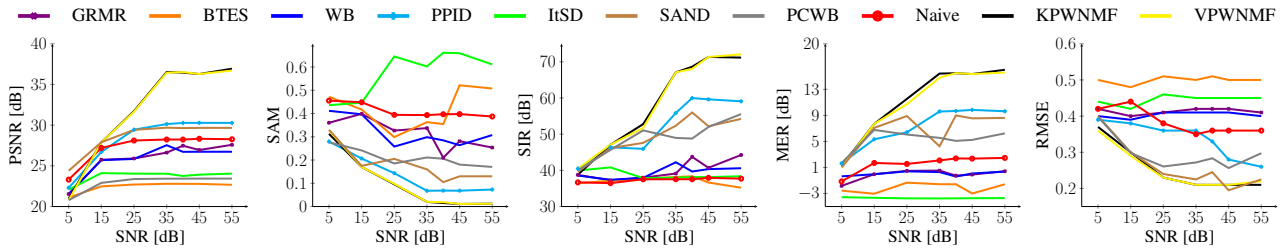


Figure 6.3 – Mean PSNR, SAM, SIR, MER and RMSE—obtained for Image 1 (Assumption 1&2) with 4×4 and 5×5 filters—relative to input SNR.

A number of key observations can be made from the results presented in the table and the figures:

- The proposed KPWNMF and VPWNMF methods exhibit the highest PSNRs for both images, i.e., the best demosaicing quality. While the demosaicing performance reached with PPID and

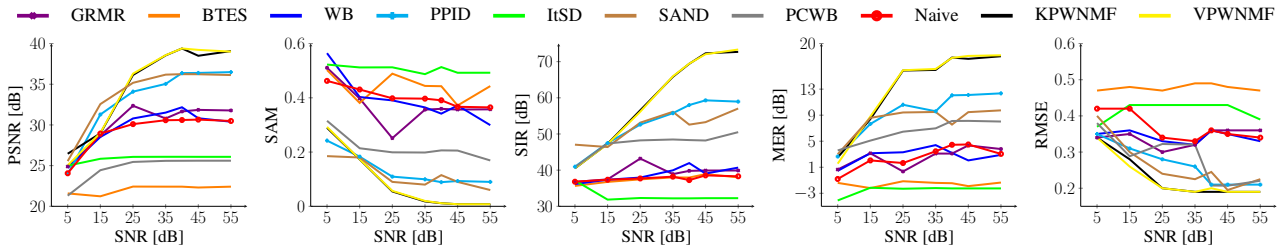


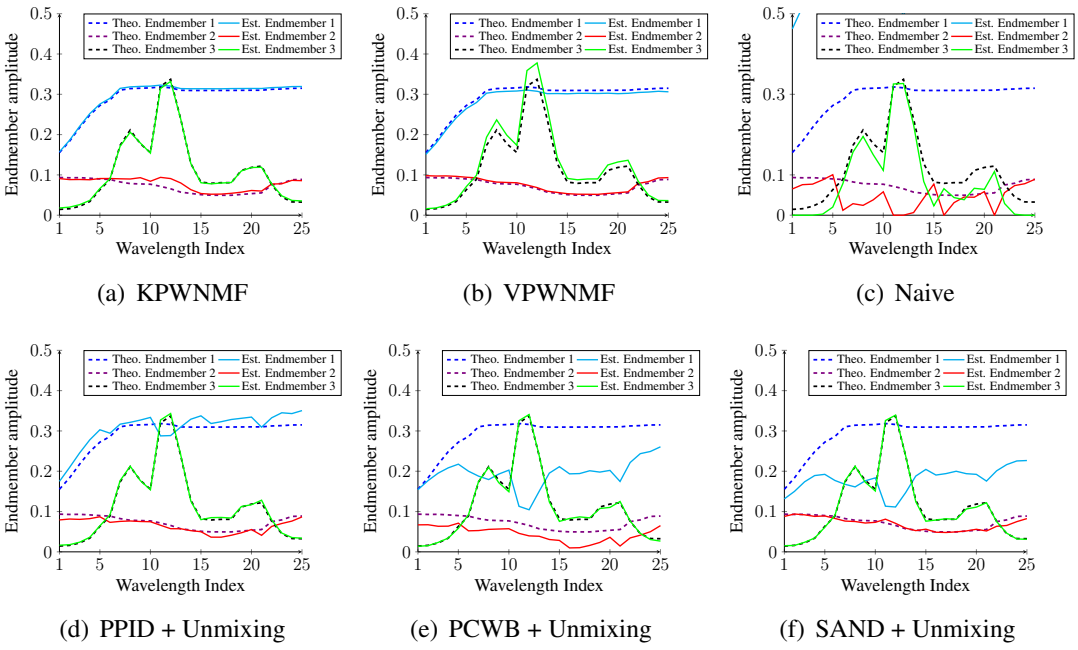
Figure 6.4 – Mean PSNR, SAM, SIR, MER and RMSE—obtained for Image 2 (Assumption 1&3) with 4×4 and 5×5 filters—relative to input SNR.

the Naive method is almost similar, the remaining methods achieve lower reconstruction quality for both images, with WB—i.e., a baseline method—performing similarly to GRMR and ItSD, and with BTES showing the poorest performance. It is important to note that the performance of these methods may vary under different noise levels. Our results show that KPWNMF and VPWNMF continue to significantly outperform the other tested methods until the input SNR is around 25 dB. With lower input SNRS, their performance degrades and both methods provide a similar demosaicing performance as SotA methods.

- In addition to their superior reconstruction quality, the proposed KPWNMF and VPWNMF methods demonstrate the highest performance in endmember identification as measured by the spectral angle mapper (SAM) and signal interference ratio (SIR). Our experimental results show that KPWNMF and VPWNMF achieved the lowest SAM and the highest SIR values among all other methods, indicating their ability to accurately identify endmembers even in the presence of noise. Moreover, our experiments demonstrate that KPWNMF and VPWNMF are robust to noise, maintaining their superior performance even at higher noise levels. Fig. 6.5 and Fig. 6.6 show the true and estimated endmembers for both images with the 5×5 filter in the noiseless case. The quality of the restored spectra with the proposed methods is much higher than the 2-stage approach using PPID for demosaicing, the latter being the SotA method to provide the highest unmixing performance.

- The proposed methods KPWNMF and VPWNMF provide the highest MER values among all other methods, indicating their superior abundance estimation performance. Moreover, our experiments demonstrate that KPWNMF and VPWNMF are robust to noise, maintaining their superior performance even at higher noise levels. The comparison between the restored abundance maps and the true abundance maps for both images using the 5×5 filter is shown in Fig.6.7 and Fig.6.8 in the noiseless case. As can be seen from the figures, the abundance maps estimated by KPWNMF and VPWNMF are visually closer to the true abundance maps compared to the other methods.

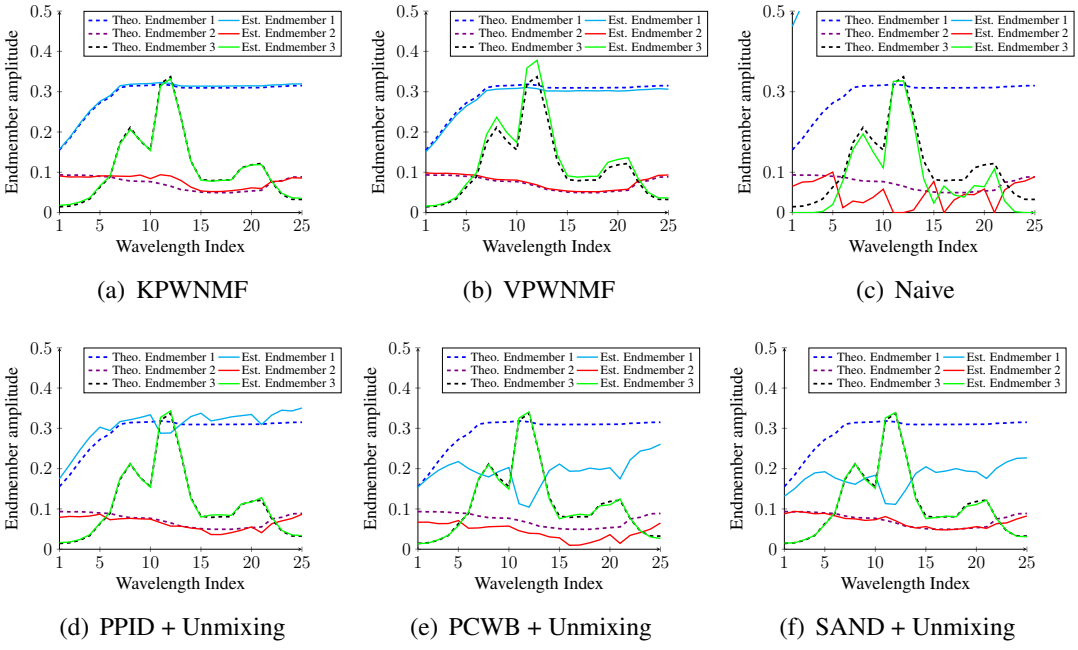
- In addition to MER, our analysis extends to the accuracy of abundance estimation, as measured by RMSE. The RMSE results for the KPWNMF and VPWNMF methods underscore their



(a) KPWNMF (b) VPWNMF (c) Naive

(d) PPID + Unmixing (e) PCWB + Unmixing (f) SAND + Unmixing

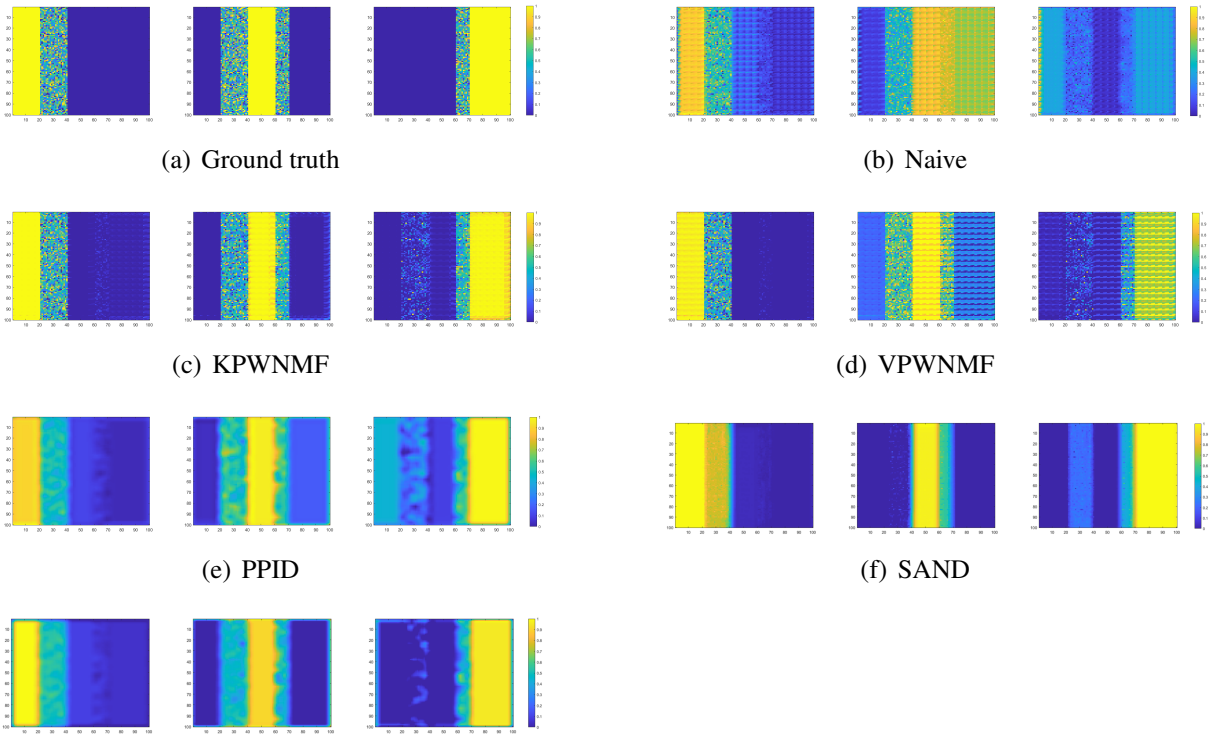
Figure 6.5 – Estimated spectra for the Image 1 with mosaic filter of size 5×5.



(a) KPWNMF (b) VPWNMF (c) Naive

(d) PPID + Unmixing (e) PCWB + Unmixing (f) SAND + Unmixing

Figure 6.6 – Estimated spectra for the Image 2 with mosaic filter of size 5×5.



(a) Ground truth (b) Naive

(c) KPWNMF (d) VPWNMF

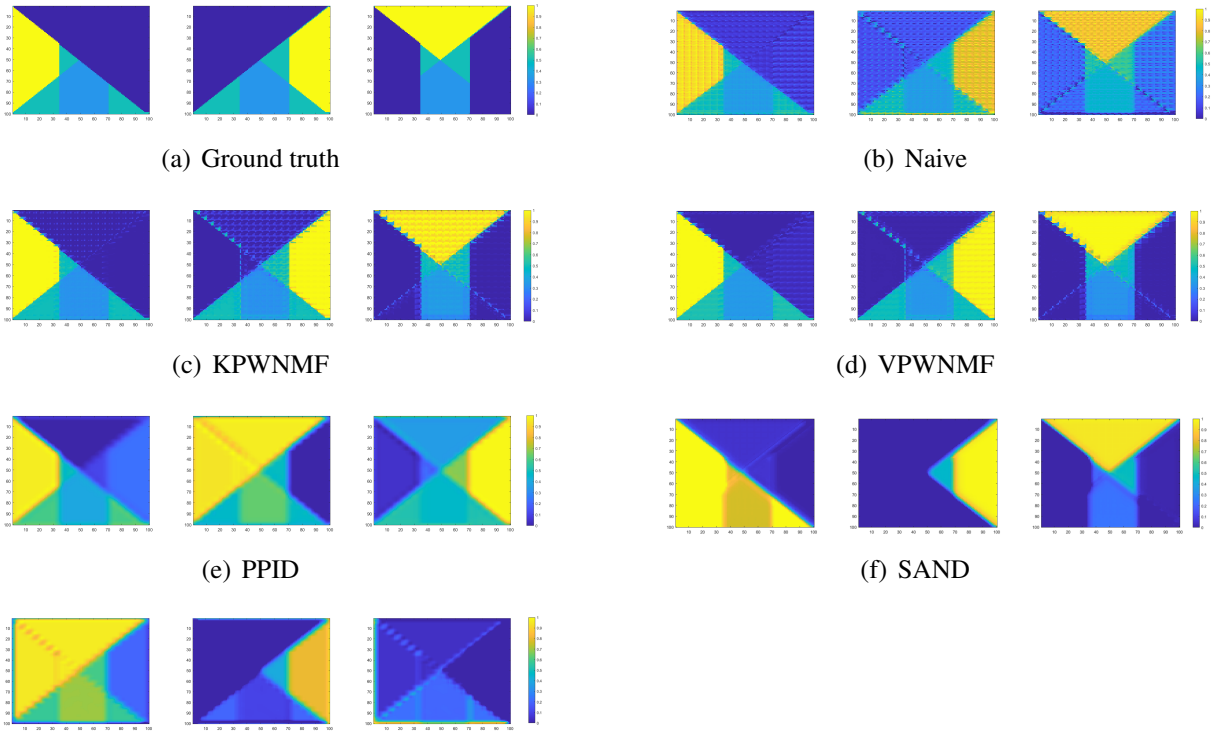
(e) PPID (f) SAND

(g) PCWB

Figure 6.7 – Estimated abundance maps for the Image 1 with 5×5 mosaic filter.

remarkable precision in estimating abundances, shown by the lowest RMSE values among all methods evaluated. While the SAND method exhibits almost similar RMSE performance, a closer examination of the restored abundance maps, as illustrated in Fig.6.7 and Fig.6.8, reveals that our methods achieve superior details and quality.

- While evaluating the effectiveness of our proposed KPWNMF and VPWNMF methods, it is essential to consider the computational time alongside other performance metrics. The experimental results indicate that the running time for both methods is notably higher than other evaluated methods. However, it is important to note that the primary focus of this study is to demonstrate the accuracy and effectiveness of these methods in enhancing the quality of demosaicing and unmixing in snapshot spectral imaging. Developing and optimizing the computational efficiency of these algorithms falls outside the immediate scope of this work. Future work will explore strategies to reduce computational time, making these methods more possible for time-sensitive applications.



(a) Ground truth (b) Naive

(c) KPWNMF (d) VPWNMF

(e) PPID (f) SAND

(g) PCWB

Figure 6.8 – Estimated abundance maps for the Image 2 with 5×5 mosaic filter.

6.2.3 Performance Evaluation on CAVE Dataset

In addition to the experiments conducted on synthetic images, we evaluated the performance of the proposed methods on the CAVE dataset, which consists of 32 spectral images captured by a multispectral camera in the 400-700 nm range with 10 nm steps. The images have a resolution of 512×512 pixels and contain a variety of scenes and objects, making them a suitable dataset for evaluating the performance of multispectral image reconstruction methods. We performed experiments on the CAVE dataset using 4×4 and 5×5 spectral filter patterns using the same strategy as in [143] with all the images in the dataset and evaluated the PSNR achieved by all the methods for both filter patterns. The results are reported in Table 6.2, where our proposed methods KPWNMF and VPWNMF achieved the highest PSNR values, indicating their superior performance in reconstructing the images. To further evaluate the performance of the proposed methods, we also generated a visual comparison of the restored images using the 4×4 spectral filter pattern. The restored images using KPWNMF, VPWNMF, and other SotA methods are shown in Fig. 6.9. Our proposed methods demonstrated remarkable results in restoring the details of the images demonstrating their ability to preserve crucial spatial information in the images, which was not effectively captured by the other methods. These results demonstrate the proposed methods effectiveness in restoring hyperspectral images with complex

Table 6.2 – Demosaicing performance averaged over all the images in the Cave dataset. Perf. criterion: PSNR (in dB).

Method	16 bands	25 bands
GRMR	35.3	33.3
BTES	34.7	33.3
WB	35.0	33.4
PPID	37.1	35.4
ItSD	30.2	27.2
SAND	34.2	33.3
PCWB	33.7	33.7
Naive	35.1	34.1
KPWNMF	**37.7**	**35.5**
VPWNMF	**37.7**	**35.5**

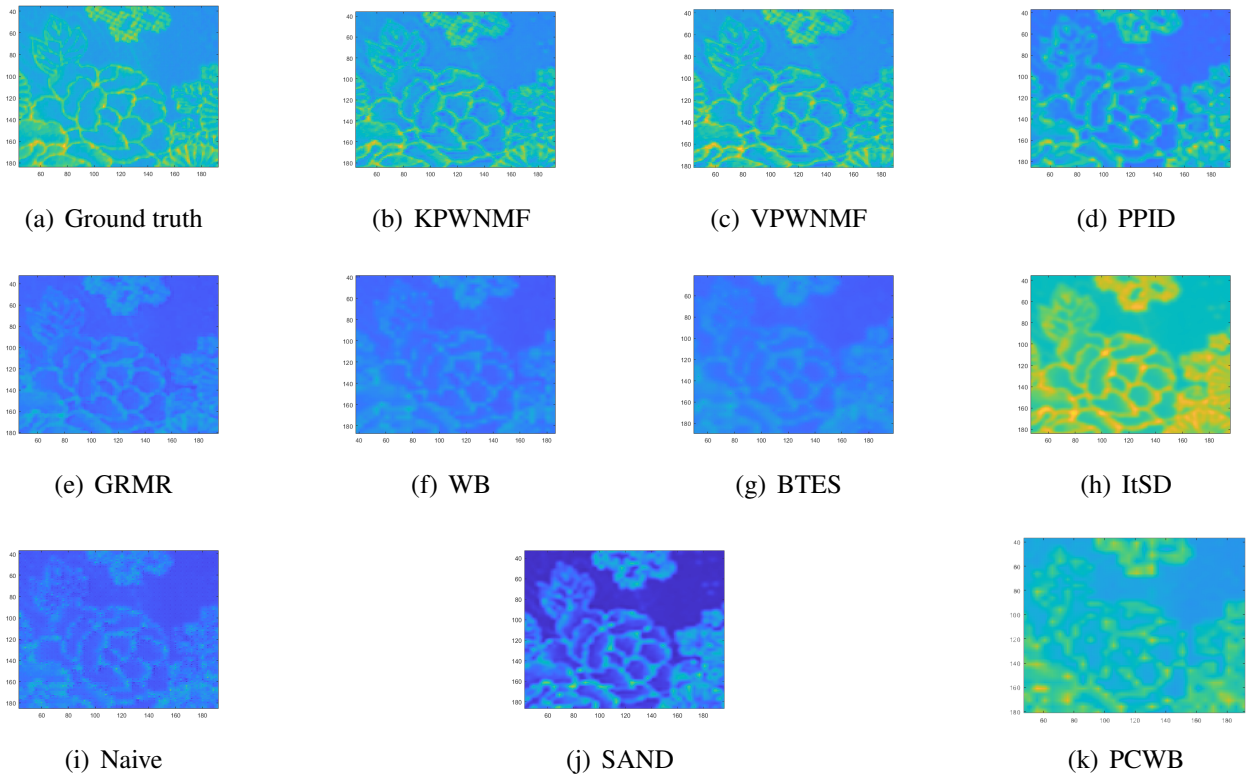
spectral and spatial information and their superiority over other SotA methods.

6.2.4 Performance Evaluation on Hyko 2 Dataset

We conducted a real data experiment using one image from the Hyko 2 dataset [155], which captures hyperspectral data from a moving vehicle. The dataset[1] contains images captured using two snapshot mosaic cameras and spans the visible and near-infrared spectral ranges from 400 to 1000 nm. We selected one image from the dataset that contained a road, trees, sky, and white signs on the road. As we do not know the real endmembers, we cannot estimate the unmixing accuracy. Instead, we use unmixing to perform image segmentation, as shown in Fig 6.10. The segmentation is done using the abundance maps generated by each method, where we select the dominant element in each pixel and draw an image showing the segmentation.

Our proposed VPWNMF method—which is based on Assumption 1 and Assumption 3—performed the best among all the methods. It can detect endmembers better than the other methods by using VCA because of the flexibility of Assumption 3. On the other hand, KPWNMF—which is based on Assumption 1 and Assumption 2—only considers the case where the abundance coefficients are significantly changing when there are several endmembers. Therefore, it did not perform as well on this image since the abundance maps were not significantly changing in certain patches. The other SotA methods, including the Naive method, their performance varied and generally fell short in capturing fine details compared to our proposed methods. While VPWNMF was the best-performing method, its performance could have been better, as in some places, the road was classified as the sky. This

[1]The database is accessible at `https://wp.uni-koblenz.de/hyko/`.



(a) Ground truth (b) KPWNMF (c) VPWNMF (d) PPID

(e) GRMR (f) WB (g) BTES (h) ItSD

(i) Naive (j) SAND (k) PCWB

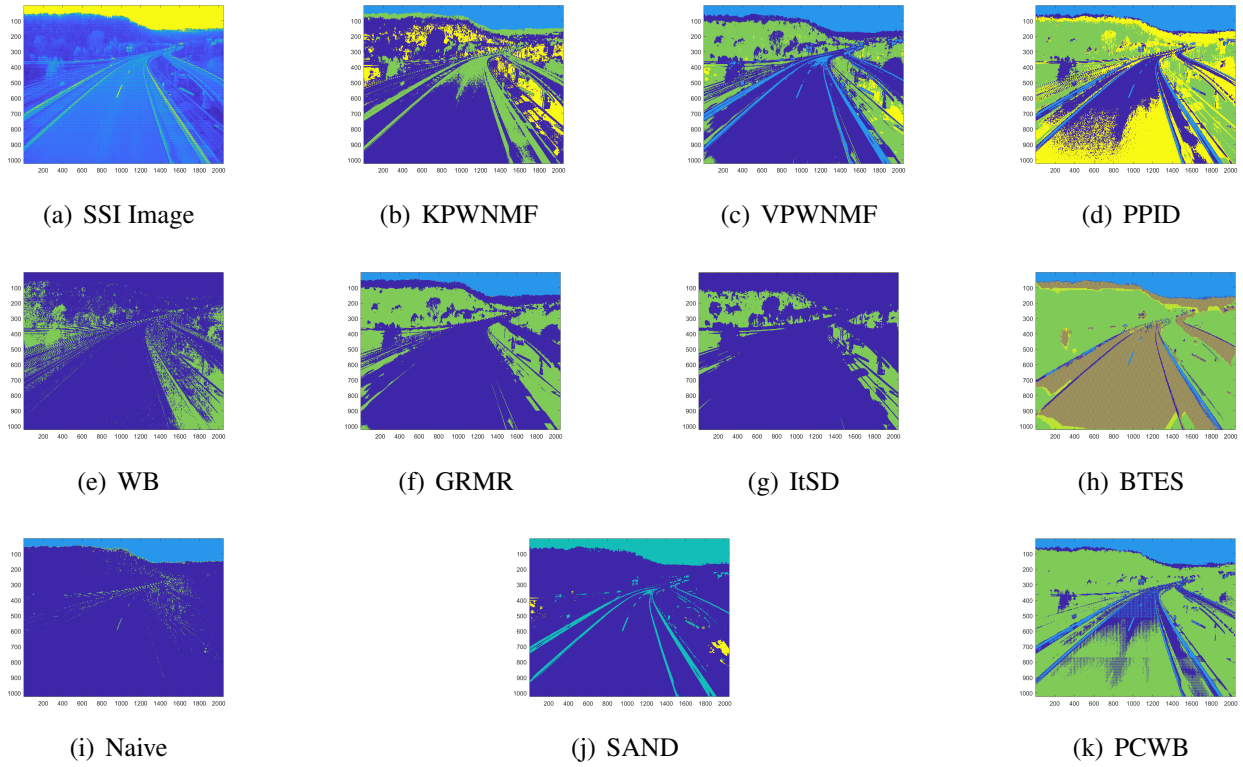
Figure 6.9 – Comparison of the demosaicing performance of all the methods on the Cloths image from the CAVE dataset with 4×4 mosaic filter.

might be due to multiple issues, e.g., spectral variability, sensor impurities, or nonlinear mixtures in some pixels.

6.3 Experimental Validation of Filtering-based Framework

6.3.1 Experiment Setup

To evaluate the effectiveness of our proposed methods FPKmeans and FPVCA, we conducted experiments on synthetic images and real snapshot spectral imaging (SSI) images captured by SSI cameras. We utilized the same set of images introduced in Sect. 6.2 for the synthetic images. Specifically, we considered two scenarios: Assumptions one and two are satisfied Fig. 6.1, and Assumptions one and three are satisfied Fig. 6.2. Each synthetic image consisted of a 100×100 pixel grid with three endmembers, namely water, metal, and concrete, whose spectral signatures were obtained from [144]. Furthermore, we consider the real spectral profiles derived from the calibration files of the photonfocus Snapshot 5x5 spectral camera [119]. Spectral correction is applied after processing the images



(a) SSI Image (b) KPWNMF (c) VPWNMF (d) PPID

(e) WB (f) GRMR (g) ItSD (h) BTES

(i) Naive (j) SAND (k) PCWB

Figure 6.10 – Comparison of the segmentation performance of all the methods on the Hyko 2 Dataset Image.

using the correction matrix provided by the camera manufacturer, as explained in Sect. 1.4.

In our evaluation, we compared the performance of our approaches against several baseline methods, including the Naive method, KPWNMF, VPWNMF, and seven two-step demosaicing-then-unmixing methods. For the latter category, we employed seven state-of-the-art demosaicing methods, namely GRMR [143], BTES [105], WB [27], PPID [107], ItSD [108], SAND[22], and PCWB [131]. In the second step, we unmix the restored data cube Y using VCA for estimating the endmembers and Fully Constrained Least Squares (FCLS) for abundance estimation.

It is important to note that in these experiments, we consider a more complex scenario by simulating realistic filter responses. As a result, unlike the experiments in Sect. 6.2, an extra stage is introduced for all the state-of-the-art (SotA) methods. This additional step involves applying the correction matrix. We have two options for its placement: either after the demosaicing step or equivalently after the unmixing step. However, for consistency and to maximize performance, we apply the correction matrix after demosaicing for all methods because they perform better except PPID, KPWNMF, and VPWNMF, which show similar performance trends in both placements.

We evaluated the tested methods by assessing demosaicing and unmixing enhancements. For demosaicing, we measured the performance using Peak Signal-to-Noise Ratio (PSNR) by comparing

estimated Y matrices with ground truth. While for unmixing, we utilized Signal-to-Interference Ratio (SIR) and Spectral Angle Mapper (SAM) for endmember estimation, along with Mixing Error Ratio (MER) for quantifying abundance map quality. We additionally incorporated RMSE for abundance estimation accuracy and running time in seconds to assess computational efficiency.

We also conducted experiments to study the impact of regularization parameter in the deconvolution step on the effectiveness of the proposed methods.

For our real data experiment, we use images from the Hyko 2 dataset [155]. This dataset consists of images captured by two snapshot mosaic cameras and covers the spectral range from 400 to 1000 nm, encompassing the visible and near-infrared regions.

6.3.2 The Impact of Regularization Parameter

In the first set of experiments we study the impact of the regularization on the deconvolution performance. The regularization parameter α controls the Tikhonov regularization, which helps balance the trade-off between data fidelity and regularization in the deconvolution process. Fig.6.11 shows the influence of the regularization parameter α on the achieved Spectral Angle Mapper (SAM) value for each noise level in the context of the inverse problem when applied to noiseless and noisy scenarios. We can see from the figure that a small value of $\alpha = 0.0005$ results in the best SAM value. However, as noise is introduced, the regularization parameter needs to be increased to improve the robustness of the deconvolution process. This trend can be followed in the plot, where higher noise levels correspond to larger values of α for achieving better SAM. It is important to note that in all the following experiments, 0.0005 was chosen as the value for α for the noiseless case. For the noisy case, we selected the value from the figure where SAM was the lowest, ensuring optimal performance at each noise level.

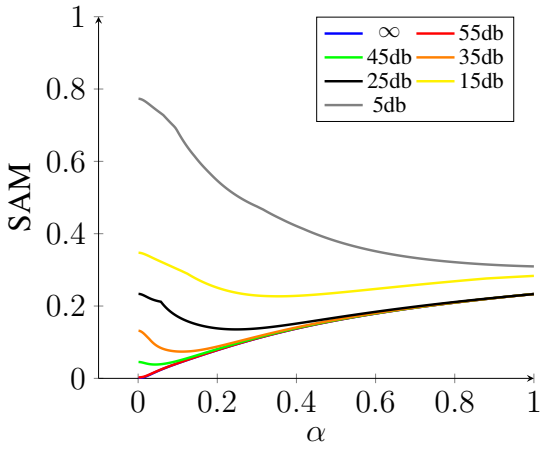


Figure 6.11 – Influence of the regularization parameter α on the achieved Spectral Angle Mapper (SAM) value for each noise level

Table 6.3 – PSNR, SAM, SIR, MER, RMSE, and Time in seconds obtained for the synthetic images with 5x5 real filters. In bold, the highest performance value.

Method	Image 1 (Assumptions 1 & 2)						Image 2 (Assumptions 1 & 3)					
	PSNR	SAM	SIR	MER	RMSE	Time	PSNR	SAM	SIR	MER	RMSE	Time
GMRM	23.0	0.52	8.3	0.7	0.4	1.7	24.3	0.36	10.8	0.2	0.4	1.6
BTES	18.6	0.52	9.4	-6.1	0.5	**0.3**	18.3	0.3	8.4	-3.3	0.5	0.5
WB	23.8	0.38	13.5	-4.2	0.4	**0.3**	26.4	0.3	16.4	-0.3	0.3	**0.4**
PPID	28.6	0.12	23.2	7.1	0.3	0.4	33.9	0.1	30.2	13.0	0.2	**0.4**
ItSD	24.4	0.32	11.8	0.2	0.5	**0.3**	26.3	0.32	14.3	0.6	0.4	**0.4**
SAND	28.9	0.13	24.3	7.9	0.3	3728	33.9	0.1	17.3	9.1	0.2	3700
PCWB	27.3	0.15	19	5.3	0.3	0.3	29.3	0.4	15	2.3	0.2	0.3
Naive	23.9	0.4	11.2	2.5	0.3	10	25.6	0.4	12.0	3.0	0.3	10
KPWNMF	**30.1**	0.06	32.7	**12.5**	0.2	6.1	34.9	0.08	32.2	16.4	0.2	6.1
VPWNMF	30.0	0.06	33.6	12.4	0.2	5.1	**35.6**	0.06	33.1	16.9	0.2	5.1
FPVCA	**30.1**	$9 \cdot 10^{-8}$	**149.0**	12.2	**0.1**	4	**36.5**	$8 \cdot 10^{-8}$	**149.0**	**17.6**	**0.07**	4
FPKmeans	**30.1**	$9 \cdot 10^{-8}$	**149.0**	12.2	**0.1**	4.5	35.8	0.05	104.5	17.1	**0.07**	4.5

6.3.3 Performance Evaluation On Synthetic Images For The Real Case

In the second set of experiments, we utilize images introduced in Sect. 6.2. We consider a real filter with a 5×5 filter pattern and investigate the performance of the tested methods under different noise levels. The PSNR, SAM, SIR, and SAM achieved by all the methods are reported in Table 6.3. While the average performance with different noise levels for both images with is presented in Fig. 6.12 and Fig. 6.13 respectively. The results presented in the table and figures reveal several significant observations:

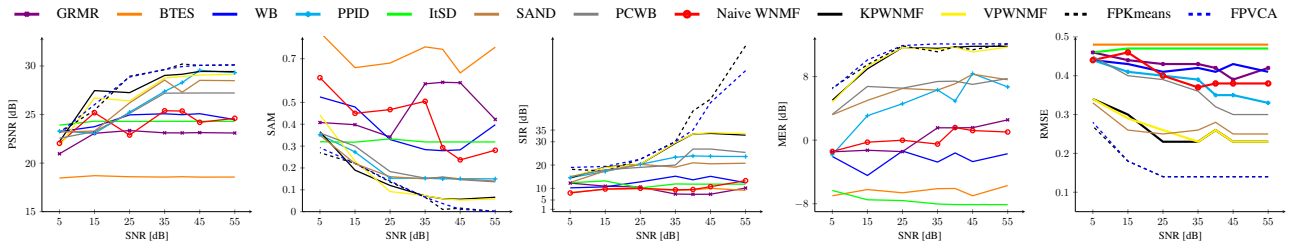


Figure 6.12 – From left to right: mean PSNR, SAM, SIR, MER and RMSE—obtained for Image 1 (Assumption 1&2) with 5×5 real filter—relative to input SNR.

- The performance of both KPWNMF and VPWNMF methods drops when compared to the ideal situation. This decline is mainly due to two reasons: first, the spectral correction step, which causes estimation mistakes, especially when there is a lot of background noise. Second, the weight matrix uses its values from the response matrix during the factorization step for each rank-1 patch.

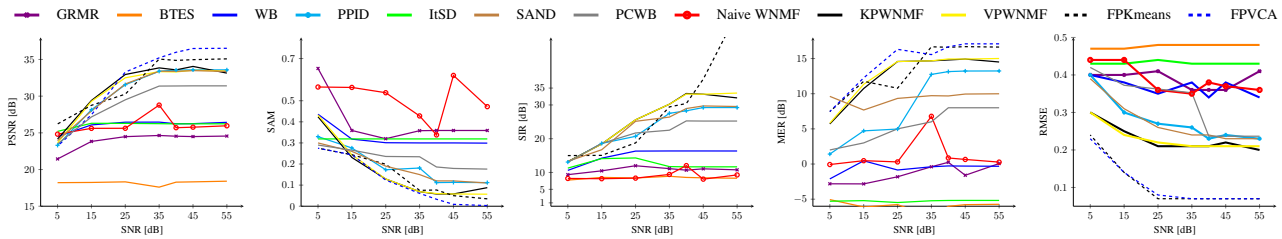


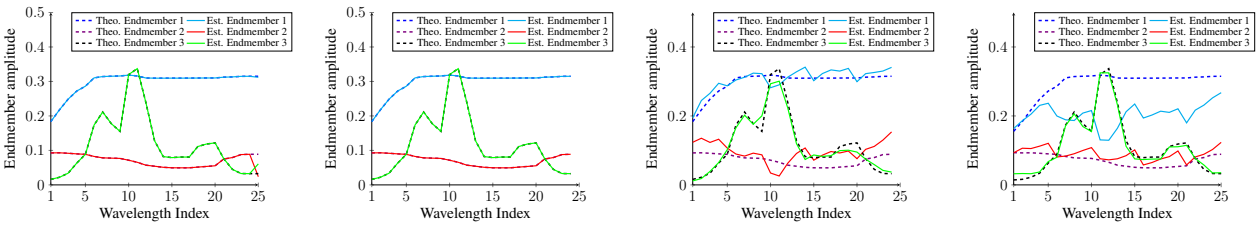
Figure 6.13 – From left to right: mean PSNR, SAM, SIR, MER and RMSE—obtained for Image 2 (Assumption 1&3) with 5×5 Real filter—relative to input SNR.

- The proposed FPKmeans and FPVCA methods exhibit superior performance in PSNR compared to our previous methods and the two-step approaches. The performance of these methods may vary under different noise levels. FPKmeans and FPVCA consistently outperform the other methods. However, their performance declines at lower input SNRs, and they exhibit similar demosaicing performance to the state-of-the-art methods. It is worth mentioning that the advantage of FPKmeans and FPVCA is that they eliminate the need for spectral correction steps required in the other methods.

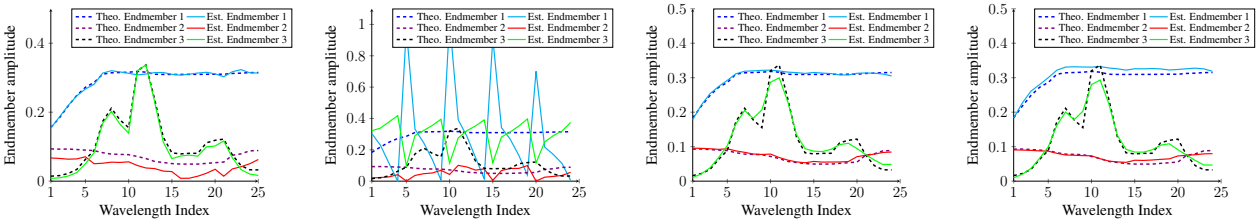
- In addition to their superior reconstruction quality, the proposed FPKmeans and FPVCA methods outperform all other methods regarding endmember identification as shown in Figures 6.14 and 6.15. They achieve the lowest spectral angle mapper (SAM) values and the highest signal interference ratio (SIR) values, indicating their accurate identification of endmembers even in the presence of noise. These methods exhibit robustness to noise, maintaining their superior performance even at higher noise levels. However, it is worth noting an exception observed in Image 2 with SNRs below 25 dB, where even PPID outperforms FPKmeans in terms of SIR. In the case of Image 1, where the assumptions of FPKmeans are satisfied, the latter exhibits slightly better performance than FPVCA.

- The proposed methods FPKmeans and FPVCA demonstrate comparable performance in terms of the Mixing Error Ratio (MER), to the joint unmixing and demosaicing approaches KPWNMF and VPWNMF. These methods consistently achieve the highest MER values among all other methods, indicating their superior accuracy in estimating the abundance maps. Furthermore, the experiments reveal the robustness of FPKmeans and FPVCA to noise, as they maintain their exceptional performance even at higher noise levels. Fig. 6.16 and Fig. 6.17 illustrate the comparison between the restored abundance maps and the true abundance maps for both images, using the 5×5 filter in the noiseless case.

- The RMSE results, implying precision in abundance estimation, show FPKmeans and FPVCA as superior, even under various noise levels. Their lowest RMSE values, complemented by the visual quality of abundance maps in figures Fig. 6.16 and Fig. 6.17 , demonstrate their robustness and accuracy.

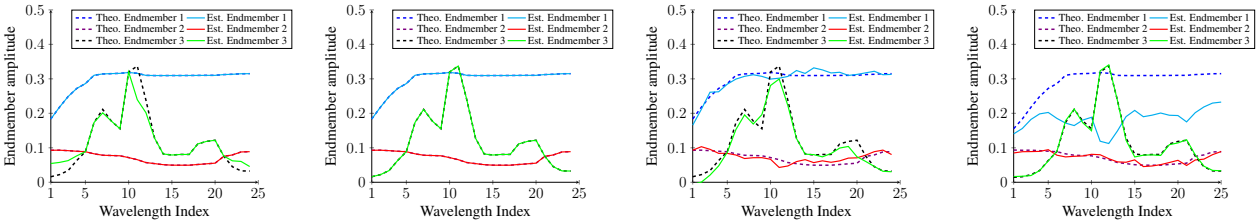


(a) FPKmeans (b) FPVCA (c) PPID (d) SAND

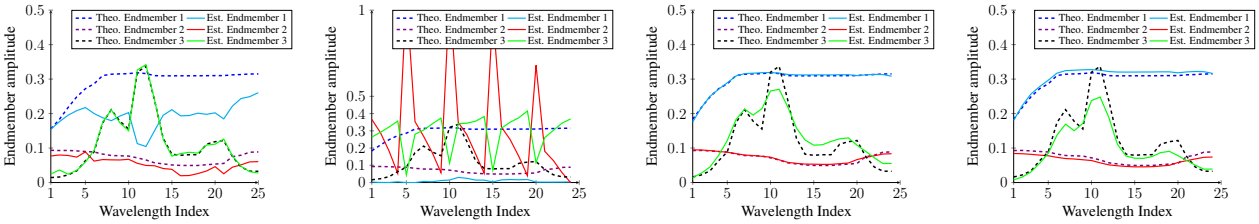


(e) PCWB (f) Naive (g) KPWNMF (h) VPWNMF

Figure 6.14 – Estimated spectra for the Image 1 with Real filter of size 5×5.

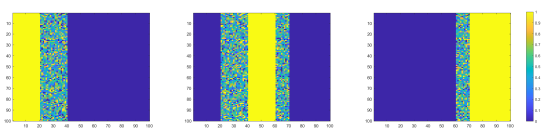


(a) FPKmeans (b) FPVCA (c) PPID (d) SAND

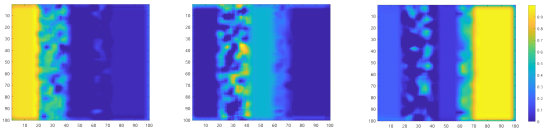


(e) PCWB (f) Naive (g) KPWNMF (h) VPWNMF

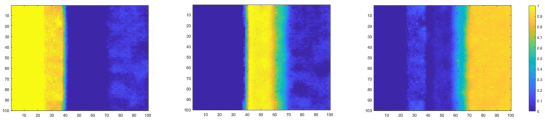
Figure 6.15 – Estimated spectra for the Image 2 with Real filter of size 5×5.



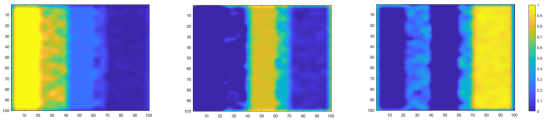
(a) Ground truth



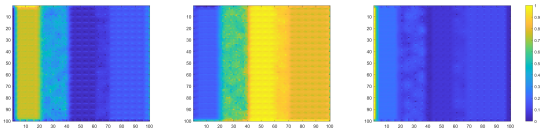
(b) PPID



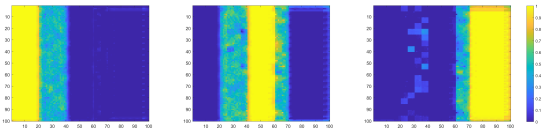
(c) SAND



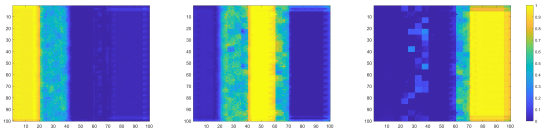
(d) PCWB



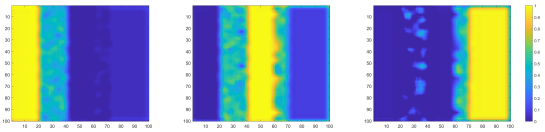
(e) Naive



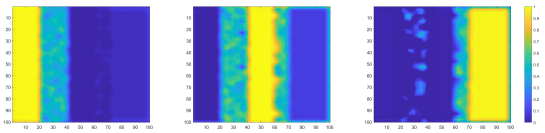
(f) KPWNMF



(g) VPWNMF

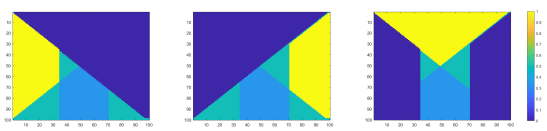


(h) FPVCA

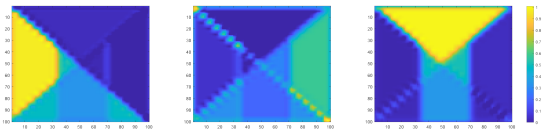


(i) FPKmeans

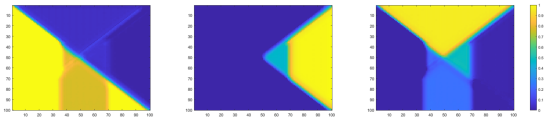
Figure 6.16 – Estimated abundance maps for the Image 1 with real filter of size 5×5



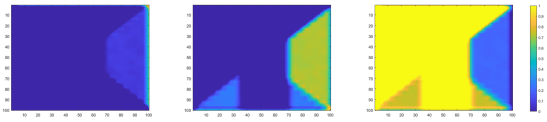
(a) Ground truth



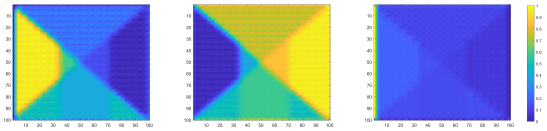
(b) PPID



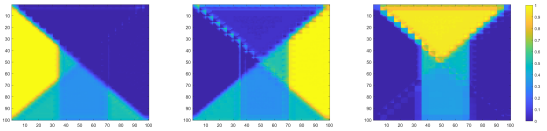
(c) SAND



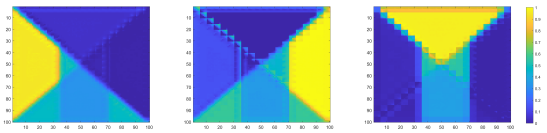
(d) PCWB



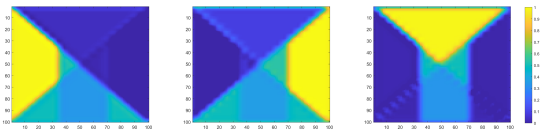
(e) Naive



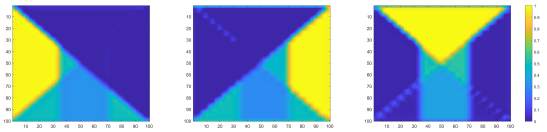
(f) KPWNMF



(g) VPWNMF



(h) FPVCA



(i) FPKmeans

Figure 6.17 – Estimated abundance maps for the Image 2 with real filter of size 5×5

- In assessing the new methods FPKmeans and FPVCA, we note their computational time is improved compared to KPWNMF and VPWNMF. However, it remains higher than some SotA methods. Despite this, our study focuses on demonstrating the accuracy and effectiveness of these methods in snapshot spectral imaging rather than on computational efficiency. Optimizing running time is considered future work to enhance these methods for more time-sensitive applications

6.3.4 Performance Evaluation On Hyko 2 Dataset

In the first part of the experiments, the Hyko dataset was utilized to evaluate various methods for image segmentation, including VPWNMF and KPWNMF. The results demonstrated that VPWNMF achieved the best segmentation performance, followed by KPWNMF. Spectral correction was applied to the output of these methods after the demosaicing step to improve the accuracy[2]. In the second part of the journal, a similar experiment was conducted using the same dataset. However, in this case, the proposed methods, FPVCA, and FPKmeans, exhibited comparable performance to VPWNMF and KPWNMF, respectively, without spectral correction. The results in Fig. 6.18 show that FPVCA exhibited comparable performance to VPWNMF, while FPKmeans outperformed all other methods, including VPWNMF. It is worth noting that the spectral correction step did not affect the abundance estimation for VPWNMF, and KPWNMF as their abudndances remained consistent even without correction.

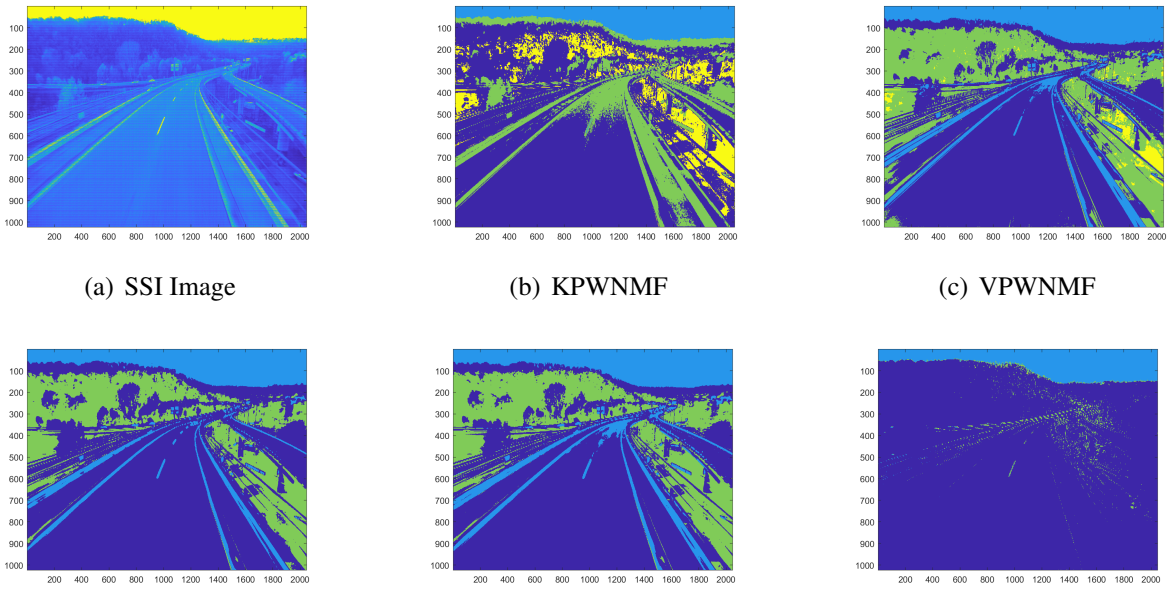
6.4 Conclusion

In this chapter, we evaluated our two novel frameworks for performing joint demosaicing and unmixing of hyperspectral images acquired by snapshot spectral imaging (SSI) cameras. In the matrix completion framework, we aimed to integrate the "demosaicing" and "unmixing" processes within a single step. Our comprehensive experiments with both synthetic and real SSI datasets revealed that these proposed methods surpass traditional two-stage approaches, which separate demosaicing and unmixing processes.

Furthermore, we explored the integration of deconvolution with unmixing in our second framework, which proved highly effective. This approach utilizes the harmonic responses of the Fabry-Perot filters from the spectral response matrix. It improves the accuracy of segmentation, demosaicing, and unmixing and simplifies the processing pipeline by eliminating the need for post-demosaicing spectral correction. Our experimental results indicated that incorporating these advanced sparsity and deconvolution techniques could significantly enhance the performance of SSI systems, particularly when compared to traditional methods and existing joint demosaicing-unmixing approaches.

Despite these advancements, our methods face ongoing challenges, such as errors introduced by spectral variability and sensor impurities, which we plan to address in future research.

[2]The correction matrix was kindly provided to us by the Hyko 2 dataset authors.



(a) SSI Image (b) KPWNMF (c) VPWNMF

(d) FPKmeans (e) FPVCA (f) Naive

Figure 6.18 – Comparison of the segmentation performance of all the methods on the Hyko 2 Dataset Image

Chapter 7

Improved deconvolution with Entropy-weighted penalization

"Measurement is the first step that leads to control and eventually to improvement. If you can't measure something, you can't understand it. If you can't understand it, you can't control it. If you can't control it, you can't improve it." — H. James Harrington

7.1 Introduction

In this chapter, we introduce a novel method for spectral deconvolution in Fabry-Perot-based hyperspectral cameras. The camera manufacturer proposed a scene-independent correction matrix to remove the harmonics presented in the SSI image. Therefore, we exploit the response of the Fabry-Perot filter—as we did in Chapter 5—and propose a novel scene-dependent spectral correction and calibration method[1]. Our approach utilizes deconvolution with Tikhonov regularization weighted by the entropy of the Fabry-Perot harmonics to remove the generated artifacts and restore the original spectra. It adapts to the scene unique characteristics, reducing harmonics and improving hyperspectral data quality. Experiments with synthetic data and real images from two IMEC sensors—i.e., linescan wedge and snapshot mosaic FPI designs——show our method outperforms the manufacturer solution in eliminating harmonics and enhancing spectral information.

7.2 Preface

As we previously explained in Sect. 1.4, the camera manufacturer proposed a correction matrix to remove the harmonics presented in the SSI image. The latter is determined through a process of minimizing the difference between the actual band response, denoted as the response matrix $\mathbf{H}$, and

[1]This method is submitted to IEEE Sensors Letters [1].

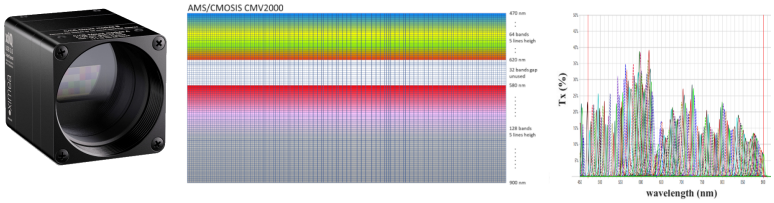


Figure 7.1 – The Hyperspectral Linescan camera from XIMEA using a wedge pattern of 192 VISNIR filters (Source [159]).

the desired or ideal band response, represented as $\mathbf{H}^{\text{ideal}}$ [120, 161]. This minimization process is expressed mathematically as:

$$\min_{\mathbf{C}} \|\mathbf{H}^{\text{ideal}} - \mathbf{C} \cdot \mathbf{H}\|_{\mathbb{F}}^2, \tag{7.1}$$

where $\|\cdot\|_{\mathbb{F}}^2$ denotes the Frobenius norm.

This matrix is scene-independent and does not adapt to the scene characteristics. Therefore we exploit the response of the Fabry-Perot filter—using the same concept in our filtering-based framework—and propose a novel scene-dependent spectral correction and calibration method. Our approach utilizes deconvolution with Tikhonov regularization weighted by the entropy of the Fabry-Perot harmonics to remove the generated artifacts and restore the original spectra. It adapts to the scene unique characteristics, reducing harmonics and improving hyperspectral data quality. This approach is designed and tested on two Fabry-Perot-based cameras: linescan wedge and snapshot mosaic.

The linescan wedge filter configuration employs a distinctive wedge-shaped design. The sensor is segmented into k distinct bandlets in this layout, each corresponding to a unique Fabry-Perot filter characterized by a specific cavity height. This differentiation enables the capture of varied spectral information from the scene, as depicted in Fig. 7.1. These bandlets are typically arranged in an ascending order based on their cavity heights, forming a pattern akin to a staircase, also known as a discrete wedge.

This setup requires relative movement—placing the object on a conveyor belt or mobilizing the camera, such as mounting it on an Unmanned Aerial Vehicle (UAV). It is also crucial to fine-tune the camera frame rate to ensure synchronization and address the potential discrepancies arising from capturing images at varying viewpoints, which may require correction [29, 106]. Consequently, the raw images collected need the process of stitching and alignment to construct an initial, spectrally uncorrected data cube. Then, spectral correction is required to build the final data cube.

While the linescan wedge filter configuration requires specific handling, such as image-stitching and alignment, the process for the snapshot mosaic camera revolves primarily around demosaicing. However, both the linescan wedge and snapshot mosaic cameras follow a similar image processing pipeline, with the main difference being replacing the stitching step with demosaicing in the snapshot mosaic setup. Fig. 7.2 illustrates the image processing pipeline—previously presented in Fig. 1.12—adapted to integrate both camera configurations seamlessly.

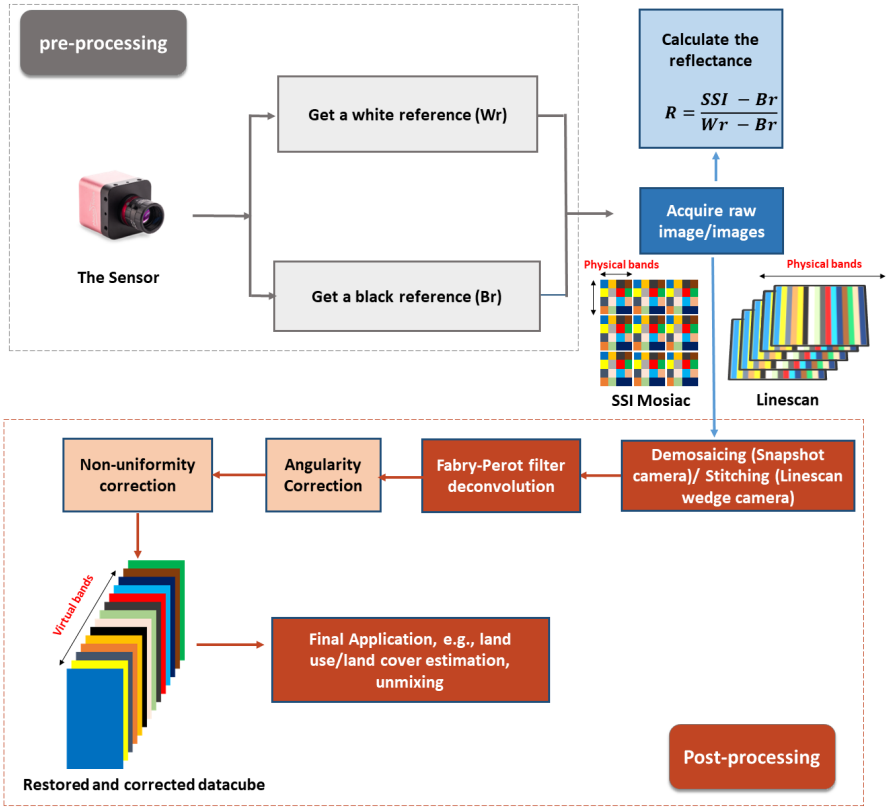


Figure 7.2 – The essential steps in processing images acquired by snapshot spectral and linescan cameras. (Single Image Demosaicing vs. Multi-Image Stitching)

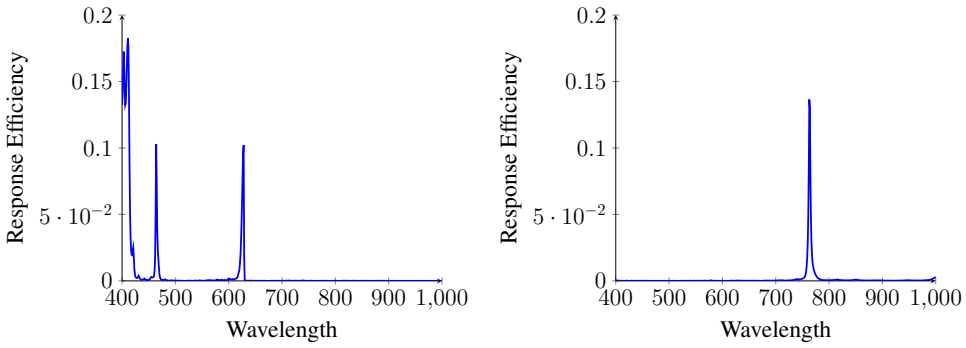
7.3 Understanding the Role of Entropy

Entropy allows to quantify the level of information in a given system. Consider a d-dimensional probability vector, denoted as $\mathbf{u} = (u_1, u_2, \ldots, u_d)^T$ where each u_i represents the probability of the i^{th} event occurring in a d-dimensional probability space. The entropy of the vector $\mathbf{u}$ is defined as:

$$E(\mathbf{u}) = -\sum_{i}^{d} u_i \cdot \log(u_i). \tag{7.2}$$

Here, the logarithmic function can be either the natural logarithm or the binary logarithm, depending on the context. The entropy here serves to quantify the extent of uncertainty or disorder inherent in the values of $\mathbf{u}$. One important characteristic of entropy is that it reaches its maximum value when the probability distribution of $\mathbf{u}$ is equally distributed among all events, and conversely, it reaches its minimum value when a single event dominates with near certainty [98].

Given the harmonics of the FPf shown in Fig. 7.3, when the band response exhibits multiple peaks, the resulting entropy will be very high. Conversely, when there is only a single response peak, the entropy is near zero. Incorporating this information into the deconvolution problem allows us to effectively control the level of smoothness required for the corrected spectrum.



(a) Entropy = 0.54 (b) Entropy = 0.04

Figure 7.3 – Fabry-Perot filter responses for different bands of the linescan camera. On the left: 805 nm. On the right: 650 nm.

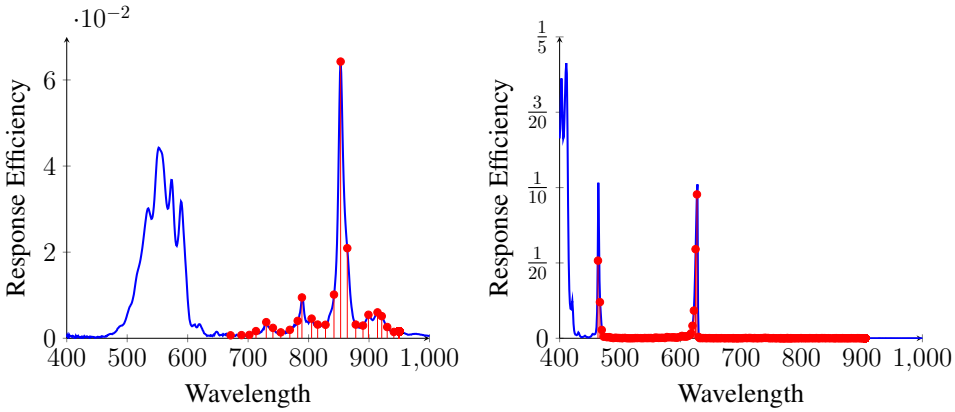
7.4 Proposed Spectral Correction Method

In our spectral imaging context, after the process of demosaicing or stitching, we encounter a datacube where each spatial pixel is a $k \times 1$ vector denoted as $\mathbf{z}$—where k is the number of physical

wavelengths—which reads

$$\mathbf{z} = \mathbf{H} \cdot \mathbf{s}. \tag{7.3}$$

Here, $\mathbf{H}$ is the response matrix with dimensions $k \times q$ where q is the number of measurement points used during the calibration of the sensor and $\mathbf{s}$ is the original spectrum that we intend to see in each pixel. In practice, $q = 601$ as the manufacturer samples responses at every 1 nm interval within the range of 400 nm – 1000 nm. The primary challenge we encounter is the estimation of the vector $\mathbf{s}$ based on the observed spectral pixel vector $\mathbf{z}$ and the response matrix $\mathbf{H}$. However, due to the available dimensions, estimating $\mathbf{s}$ is an under-determined problem. Additionally, the use of the cut-off filter results in the blocking of harmonics outside the sensor active range, making the matrix $\mathbf{H}$ not suitable for the retrieval process. To address this, we propose to reduce the size of the matrix $\mathbf{H}$ to retain only the applicable information. This reduction is achieved through a strategic sampling of the matrix $\mathbf{H}$, keeping values corresponding to virtual wavelengths that carry the most relevant information of the sensor, while accounting for contributions from harmonics and cross-talk, as depicted in Fig. 7.4.



(a) Photonfocus SSI 5×5 [1]

(b) Ximea linescan 192 bands [2]

Figure 7.4 – Fabry-Perot filter responses (in blue) for one band from linescan and SSI cameras with highlighted virtual wavelengths (in red)

Subsequently, we perform a scaling operation on the rows of matrix $\mathbf{H}$ to ensure that they sum to one. The sampled response matrix is denoted as $\mathbb{H}^{\mathbb{S}}$ and has dimensions $k \times m$, where m represents the number of virtual wavelengths. Thereafter, the size of the vector $\mathbf{s}$ is adjusted to $m \times 1$ and is referred to as $\mathbf{v}$ for clarity. In order to obtain $\mathbf{v}$, we aim to solve a nonnegative least-squares problem, which is formulated as

$$\min_{\mathbf{v} \geq 0} \frac{1}{2} \|\mathbf{z} - \mathbb{H}^{\mathbb{S}} \cdot \mathbf{v}\|_2^2 + \frac{\lambda}{2} \cdot \|D \cdot \mathbf{v}\|_2^2 + \frac{\lambda}{2} \cdot \mathcal{R}(\mathbb{H}^{\mathbb{S}}, \mathbf{v}). \tag{7.4}$$

The deconvolution incorporates regularization with a penalty term $\mathcal{R}(\mathbb{H}^{\mathbb{S}}, \mathbf{v})$, controlled by the regularization parameter λ. It reads

$$\mathcal{R}(\mathbb{H}^{\mathbb{S}}, \mathbf{v}) = \|E(\mathbb{H}^{\mathbb{S}}) \cdot D \cdot \mathbf{v}\|_2^2, \tag{7.5}$$

where $E(\mathbb{H}^{\mathbb{S}})$ is diagonal matrix accounts for the entropy of each column of the matrix $\mathbb{H}^{\mathbb{S}}$ as follows:

$$E(\mathbb{H}^{\mathbb{S}}) = \begin{cases} -\sum_{l=1}^{k} \mathbb{H}^{\mathbb{S}}_{l,j} \cdot \log(\mathbb{H}^{\mathbb{S}}_{l,j}) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{7.6}$$

Here, i and j are the row and column indices. The entropy is applied to each column of the matrix $\mathbb{H}^{\mathbb{S}}$ to automatically determine the degree of smoothness required in the estimation. In practice, it can be interpreted as follows: when the band response exhibits multiple peaks, the resulting matrix $\mathbb{H}^{\mathbb{S}}$ can be ill-conditioned, as most of its values will tend to be near zero[2]. This situation results in a high entropy value indicating a need for increased smoothness regularization. Conversely, when there is only a single response peak, the entropy is zero, signifying that no additional penalization for smoothness is necessary. In summary, this method combines entropy regularization on the matrix $\mathbb{H}^{\mathbb{S}}$ and Tikhonov regularization on the vector $\mathbf{v}$, effectively balancing data fidelity and the desired level of smoothness in the estimated spectrum. This approach ensures an adaptive optimization process for recovering $\mathbf{v}$ from the observed data, and we refer to it as "Scene-Dependent Spectral Correction (SDS-Cor)"

Finally, in the context of our proposed method, the possibility of doing super-resolution appears when considering scenarios without cut-off filters. In such cases, the full informational content of the harmonics in the response matrix $\mathbf{H}$ can be exploited for "super-spectral" resolution enhancement. Our inverse problem framework is well-suited to exploit this opportunity by changing the dimensions of $\mathbf{H}$ and adjusting the spectral estimation process accordingly.

7.5 Experimental Validation

7.5.1 Experimental Setup

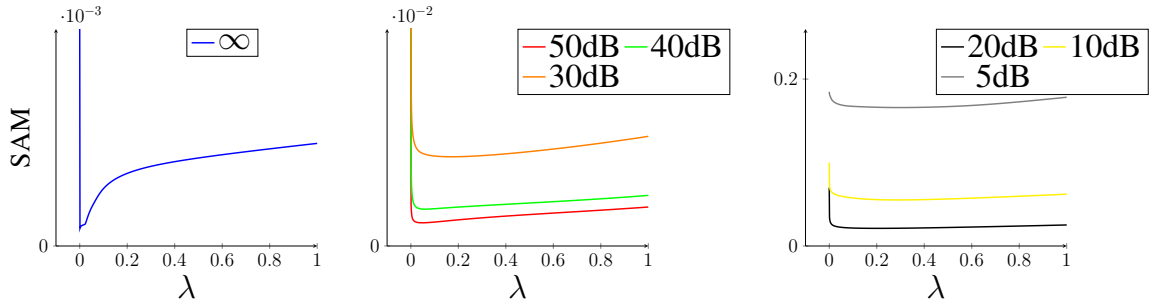
We conducted experiments to precisely evaluate the effectiveness of our proposed spectral correction method SDS-Cor. Our assessment contained both synthetic simulations, generated using response matrices, and real-world image acquisitions obtained from two distinct physical cameras. To further illustrate the versatility of our method, we included a specialized experiment involving a satellite image. This experiment aimed to demonstrate how our spectral correction method performs in remote sensing and spaceborne imaging. Finally, we conducted an experiment specifically designed to showcase the super-resolution capabilities of our method.

[2]For example, the 5×5 filter matrix $\mathbf{H}$ using real FPfs in [143] has two rows which are almost null.

7.5.2 Synthetic Data Experiment

The synthetic experiment evaluated the spectral responses of water, metal, and concrete using spectral signatures from [144] and the response matrices of the Snapshot Mosaic [119] and Linescan Wedge cameras [158]. We simulated sensor responses to these materials, applied both our proposed and IMEC's correction methods [120], and assessed corrections by comparing the Spectral Angle Mapper (SAM) scores between corrected and original spectra under varying noise levels.

However, we first study the impact of the regularization parameter λ on the NNLS problem. Fig. 7.5 illustrates the influence of the regularization parameter λ on the SAM value for each noise level in the context of the inverse problem for the linescan camera. Optimal SAM values are obtained with $\lambda = 0.0001$ in noiseless scenarios, while higher noise requires larger λ values to maintain deconvolution robustness. Throughout this section, $\lambda = 0.0001$ was used for noiseless synthetic experiments, with noise-specific optimal values determined from the plot. For the real data experiment, we chose the value $\lambda = 0.0014$, assuming the cameras to generate images with an input SNR equal to 40 dB.



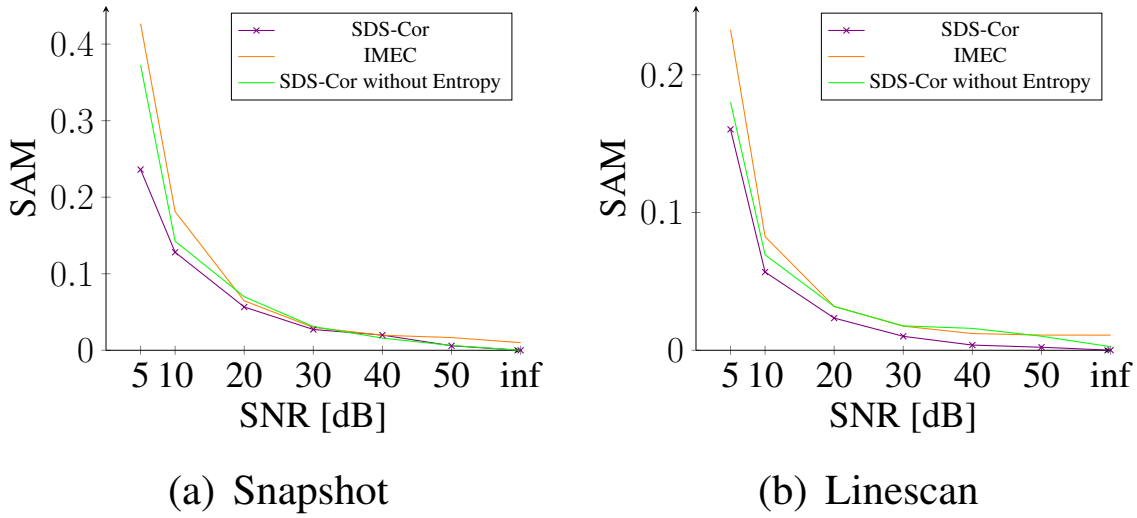
(a) Noiseless (b) 50, 40, and 30 dB (c) 20, 10, and 5 dB

Figure 7.5 – Reached SAM values vs the value of λ and the input SNR.

Our comparative analysis shown in Fig. 7.6, evaluates the performance across various noise levels and highlights the advantage of incorporating entropy weighting into SDS-Cor method. The comparison highlights SDS-Cor superiority to the IMEC method in noisy and noise-free settings, showing greater fidelity to the original spectra even in high-noise conditions.

7.5.3 Jasper Ridge image Experiment

The second experiment assessed SDS-Cor on the Jasper Ridge image, simulating a real-world satellite image acquired with an SSI camera featuring a 5×5 mosaic pattern. Post-simulation, we enhanced spatial resolution through demosaicing and endmembers were estimated using the Vertex Component Analysis (VCA) method [110]. Then, spectral correction of endmembers was performed using either



(a) Snapshot (b) Linescan

Figure 7.6 – Results on synthetic simulation used USGS spectral data

Table 7.1 – Spectral correction using the proposed SDS-Cor and IMEC methods for Jasper Ridge image. Performance criteria: PSNR, SAM, and SIR.

	SDS-Cor			**IMEC**		
Method	**PSNR**	**SAM**	**SIR**	**PSNR**	**SAM**	**SIR**
GRMR	20.9	0.573	54.9	16.0	1.001	17.2
WB	21.7	0.455	55.8	16.3	1.023	18.6
KPWNMF	24.1	0.113	60.7	22.9	0.159	29.2

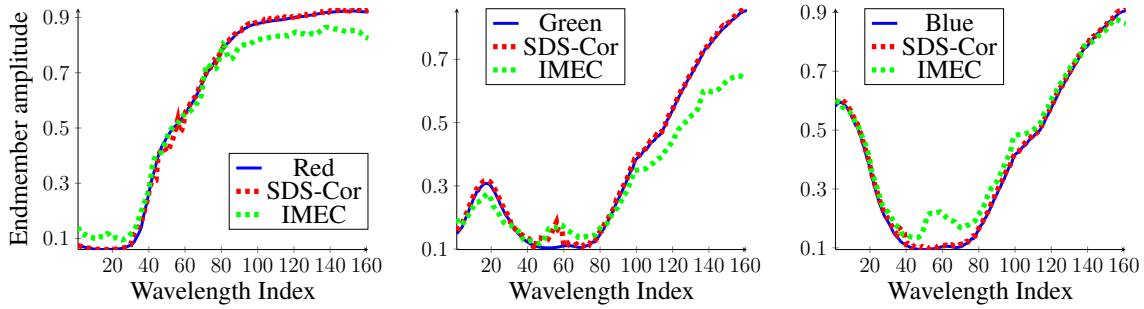
the IMEC correction or our SDS-Cor method. We compared against the KPWNMF, and analyzed performance alongside two demosaicing methods, GRMR [143] and WB [27]. The performance of each method was assessed using Peak Signal-to-Noise Ratio (PSNR), Spectral Information Ratio (SIR), and SAM. The results of these evaluations are presented in Table 7.1. For all the tested methods, the results showcase SDS-Cor superior performance over the IMEC correction approach.

7.5.4 Real Data Experiment

To further evaluate the effectiveness of our proposed method, we conducted experiments using linescan camera [158]. The camera captured images of colored papers in red, green, and blue under natural sunlight. Subsequently, we measure the spectral information of the colored papers using an ASD FieldSpec 4 spectrometer[3], serving as the reference dataset for our assessment. We then applied

[3]See, e.g., https://www.malvernpanalytical.com/en/products/product-range/asd-range/fieldspec-range/fieldspec4-hi-res-high-resolution-spectroradiometer.

our method and IMEC correction matrix to the images and compared the corrected spectra with those obtained from the spectrometer measurements. Fig. 7.7 display the restored spectra for the linescan camera. Our method outperforms the manufacturer solution despite facing challenges in the spectral range between wavelength numbers 30 to 60, attributed to harmonics from the Fabry-Perot filters.

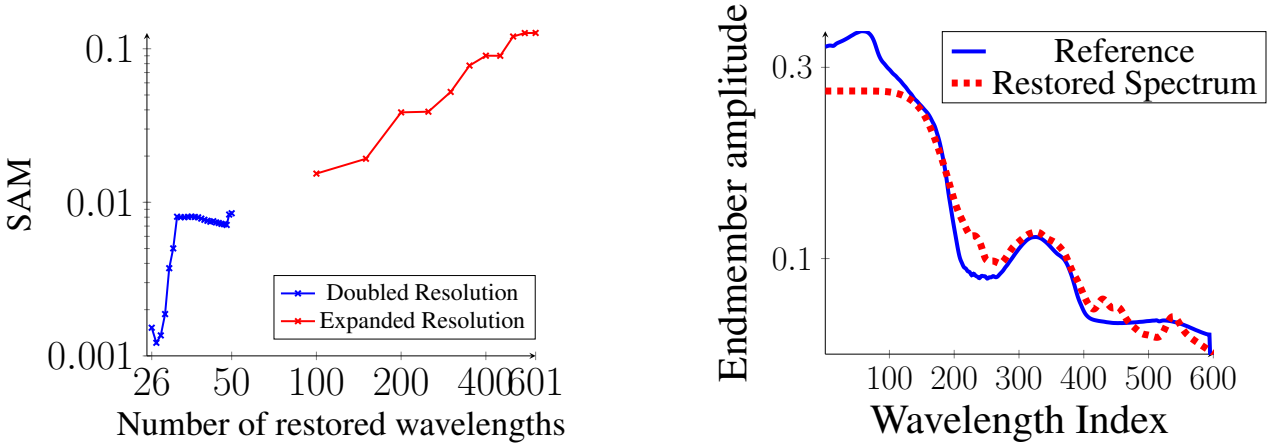


(a) Red Paper (LS) (b) Green Paper (LS) (c) Blue Paper (LS)

Figure 7.7 – Corrected spectral comparison for the linescan (LS) camera on the real data

7.5.5 Super Spectral Resolution

Finally, our proposed method can be used to enhance the resolution of hyperspectral imaging by performing super-resolution, starting directly from the response matrix of the sensor. The snapshot sensor initially supported 25 distinct wavelengths within its active range. Our primary objective was to interpolate and restore additional spectral bands, mainly focusing on the midpoints between each pair of successive bands. For instance, given that the sensor supports bands at 650 nm and 680 nm, our first step was to accurately restore the band at 665 nm, which lies precisely in the middle of these two bands. This process was iteratively conducted for all available bands, gradually increasing the number of restored bands. As we progressed, the number of bands incrementally increased, reaching 49 bands. This enhancement effectively doubled the original spectral resolution of the sensor. To evaluate the effectiveness of this super-resolution process, we utilized the SAM for each stage of the reconstruction as the blue curve shows in Fig. 7.8(a). As the number of restored wavelengths increases, we observe a corresponding increase in the SAM angle, implying a reduction in the spectral reconstruction accuracy as shown by the red curve in Fig. 7.8(a). Moreover, Fig. 7.8(b) shows an example of the super-resolution restored spectrum of the water, expanding from 25 to 601 bands with a SAM value of 0.1. The results show the potential of our method in significantly enhancing spectral resolution.



(a) SAM Angle Variations Across Expanded Band Ranges

(b) Water spectrum

Figure 7.8 – Super Resolution Spectral Analysis

7.6 Conclusion

In this chapter we introduced SDS-Cor, a new spectral correction method that use the same concept in our Filtering-based framework. We exploited the response of the Fabry-Perot filter and propose a novel scene-dependent spectral correction and calibration method. Our approach utilizes deconvolution with Tikhonov regularization weighted by the entropy of the Fabry-Perot harmonics to remove the generated artifacts and restore the original spectra. The experimental results on synthetic and real data demonstrated that the proposed method outperformed the manufacturer solution even in the presence of noise. Lastly, it should be noticed that the proposed SDS-Cor method could also be incorporated in the filtering-based joint unmixing, deconvolution, and demosaicing methods proposed in Chapter 5.

Chapter 8

General Conclusion and Perspectives

"The important thing is not to stop questioning. Curiosity has its own reason for existing."

— Albert Einstein

8.1 Conclusion

This thesis comprehensively explored snapshot spectral imaging (SSI) technologies, focusing on integrating demosaicing and unmixing jointly directly from raw data rather than treating them as separate stages. We started in Chapter 1 with a foundational understanding of SSI as an advanced approach within hyperspectral imaging (HSI), emphasizing its ability to capture complete spectral information in a single exposure. The chapter also introduced the essential calibration processes for optimizing Fabry-Perot-based SSI technology performance, underlining their importance in producing reliable spectral data.

Chapter 2 extensively explored the diverse methodologies applied in demosaicing snapshot spectral images, reviewing traditional and deep learning-based methods. We critically analyzed their capabilities and limitations, particularly noting the challenges posed by deep learning methods in scenarios with limited and highly variable training data.

Chapter 3 explored various unmixing methodologies, categorizing them into *Geometrical, Statistical, Sparse Regression-based, and Deep Learning-based* methods. We focused on nonnegative matrix factorization (NMF) and sparse component analysis (SCA) because these approaches are particularly relevant to our proposed frameworks. We also identified three main categories of NMF: *Constrained, Structured, and Generalized NMF*, each dedicated to addressing specific challenges in hyperspectral imaging.

Chapter 4 introduced our novel methods for jointly performing "demosaicing" and "unmixing". We proposed two novel approaches in addition to the naive method derived from Weighted NMF (WNMF). The first one, KPWNMF (K-means Patch-based Weighted Nonnegative Matrix Factorization), assumes that the abundances are sparse in a few patches to find, so that one endmember dominates each of these patches. Such an assumption is similar to SCA, except that we consider

partially observed data. In contrast, the second one, VPWNMF (VCA Patch-based Weighted Non-negative Matrix Factorization), relaxes the sparsity assumption needed in the latter. These methods were designed to enhance the processing of SSI images and reduce errors associated with traditional two-step processes.

Chapter 5 focused on the Fabry-Perot filter response to perform the joint demosaicing and unmixing. Specifically, we performed joint demosaicing, deconvolution and unmixing, where we integrated deconvolution directly into the unmixing process. This integration effectively utilizes the harmonic responses of Fabry-Perot filters, enhancing the accuracy and efficiency of the spectral data processing. We proposed two novel methods for this goal: FPVCA (Filter Patch-based Vertex Component Analysis) and FPKmeans (Filter Patch-based Kmeans).

Our extensive experimental evaluations, discussed in Chapter 6, demonstrated the superior performance of our proposed methods over traditional techniques, particularly in their ability to handle raw SSI data effectively.

Finally, Chapter 7 introduced SDS-Cor, a scene-dependent spectral correction method that significantly advances traditional scene-independent correction matrices. SDS-Cor adapts to specific scene characteristics by employing entropy-weighted Tikhonov regularization. We assessed SDS-cor and showed that it significantly reduced harmonic distortions and improved the spectral accuracy, outperforming the manufacturer correction method, particularly in noisy environments.

8.2 Perspectives

8.2.1 Running Time and Hyperspectral Video Processing

SSI cameras can capture videos with a high frame rate, and they are helpful in cases where the motion is unpredictable and the camera/object is moving in 2-D or 3-D, such as in robotic applications or lane sorting [79]. However, a notable limitation of our current frameworks—the completion-based and filtering-based frameworks—is their computational demand. The time required to process each frame is too high for real-time video analysis applications. To address this challenge, future work will focus on improving the computational efficiency of these frameworks. One promising approach is the adoption of compressed learning techniques [166, 165], which can significantly reduce the data volume that needs to be processed by extracting and using compressed representations of the data.

8.2.2 Spectral Variability

One of the key challenges identified during our exploration of snapshot spectral imaging (SSI) technologies is spectral variability, which significantly impacts the performance of our proposed frameworks. Spectral variability refers to the differences in spectral signatures of the same materials under varying illumination conditions, viewing angles, and environmental factors. These variations lead to inaccuracies in the post-processing tasks. Both frameworks can be enhanced to handle spectral variability. This involves developing the endmemeber estimation step by incorporating models, e.g.,

Multilayer/Deep Extensions of NMF [130, 45], that dynamically detect the variability in the end-memebers.

8.2.3 Fabry-Perot Filter Variability

The response matrix that registers the actual band response for each FPf can be subject to noise due to potential errors in its estimation. As a crucial component of the proposed methods FPVCA, FPKmeans, and SDS-Cor, it can affect their performance. Therefore, the effect of the noisy response matrix should be studied deeply and update the proposed framework to be robust to this noise.

8.2.4 Angularity Correction Integration

Angularity correction, as introduced by Goossens *et al.* [62], adjusts for light angle variability. Our proposed method filtering-based framework and SDS-Cor method can be extended to tackle this issue. This will improve the generalizability of our methods in diverse practical scenarios.

8.2.5 Integrating Deep Learning

Deep learning approaches generally demonstrate superior performance over traditional methods for demosaicing, especially in handling high-dimensional data and preserving details and color accuracy. Despite these advancements, deep learning methods face significant challenges, notably in generalization and data requirements. They often require large amounts of labeled training data to perform effectively. Furthermore, these models may struggle to generalize to new, unseen data or different MSFA configurations not included in the training set. To address these challenges, future research in deep learning for MSFA demosaicing could focus on developing more robust models that require less training data and can generalize better across different MSFA patterns. Techniques such as transfer learning [77], few-shot learning [139], and synthetic data generation [50] could play significant roles in achieving these goals. Additionally, integrating unsupervised or semi-supervised learning paradigms could reduce the dependency on large labeled datasets and help improve the model ability to generalize to new conditions.

Bibliography

[1] K. Abbas, P. Chatelain, M. Puigt, G. Delmaire, and G. Roussel. “Fabry–Perot Spectral Deconvolution With Entropy-Weighted Penalization”. In: *IEEE Sensors Letters* 8.9 (2024), pp. 1–4. DOI: 10.1109/LSENS.2024.3439209.

[2] K. Abbas, M. Puigt, G. Delmaire, and G. Roussel. “Filtering-based endmember identification method for snapshot spectral images”. In: *Proc. IEEE WHISPERS’22*. Roma, Italy, Sept. 2022.

[3] K. Abbas, M. Puigt, G. Delmaire, and G. Roussel. “Méthode de démélange et dématriçage conjoints fondée sur la complétion de rang un pour les images multispectrales "snapshot"”. In: *Actes du GRETSI*. Nancy, France, Sept. 2022.

[4] K. Abbas, M. Puigt, G. Delmaire, and G. Roussel. “Joint unmixing and demosaicing methods for snapshot spectral images”. In: *Proc. IEEE ICASSP’23*. Rhodes, Greece, 2023.

[5] K. Abbas, M. Puigt, G. Delmaire, and G. Roussel. “Méthodes parcimonieuses de déconvolution et démélange pour les images multispectrales ”snapshot””. In: *XXIXème Colloque Francophone de Traitement du Signal et des Images (GRETSI 2023)*. Grenoble, France, Aug. 2023, pp. 749–752. URL: https://ulco.hal.science/hal-04129211.

[6] K. Abbas, M. Puigt, G. Delmaire, and G. Roussel. “Locally-Rank-One-Based Joint Unmixing and Demosaicing Methods for Snapshot Spectral Images. Part I: A Matrix-Completion Framework”. In: *IEEE Trans. Comput. Imaging* 10 (2024), pp. 848–862. DOI: 10.1109/TCI.2024.3402322.

[7] K. Abbas, M. Puigt, G. Delmaire, and G. Roussel. “Locally-Rank-One-Based Joint Unmixing and Demosaicing Methods for Snapshot Spectral Images. Part II: A Filtering-Based Framework”. In: *IEEE Trans. Comput. Imaging* 10 (2024), pp. 806–817. DOI: 10.1109/TCI.2024.3402441.

[8] F. Abrard and Y. Deville. “A time–frequency blind signal separation method applicable to underdetermined mixtures of dependent sources”. In: *Signal processing* 85.7 (2005), pp. 1389–1403.

[9] F. Abrard, Y. Deville, and P. White. “A new source separation approach based on time-frequency analysis for instantaneous mixtures”. In: *Proc. ECM2S* (2001), pp. 259–267.

[10] E. H. Adelson, J. R. Bergen, et al. “The plenoptic function and the elements of early vision”. In: *Computational models of visual processing* 1.2 (1991), pp. 3–20.

[11] A. N. Akansu and R. A. Haddad. “Chapter 6 - Wavelet Transform”. In: *Multiresolution Signal Decomposition (Second Edition)*. Ed. by A. N. Akansu and R. A. Haddad. Second Edition. San Diego: Academic Press, 2001, pp. 391–442. ISBN: 978-0-12-047141-6.

[12] N. Al Makdessi, M. Ecarnot, P. Roumet, and G. Rabatel. “A spectral correction method for multi-scattering effects in close range hyperspectral imagery of vegetation scenes: application to nitrogen content assessment in wheat”. In: *Precision agriculture* 20 (2019), pp. 237–259.

[13] S. Albawi, T. A. Mohammed, and S. Al-Zawi. “Understanding of a convolutional neural network”. In: *2017 international conference on engineering and technology (ICET)*. Ieee. 2017, pp. 1–6.

[14] J. Allington-Smith. “Basic principles of integral field spectroscopy”. In: *New Astronomy Reviews* 50.4 (2006). Integral Field Spectroscopy: Techniques and Data Production, pp. 244–251. ISSN: 1387-6473.

[15] A. Ambikapathi, T.-H. Chan, W.-K. Ma, and C.-Y. Chi. “Chance-Constrained Robust Minimum-Volume Enclosing Simplex Algorithm for Hyperspectral Unmixing”. In: *IEEE Trans. Geosci. Remote Sens.* 49.11 (2011), pp. 4194–4209. DOI: 10.1109/TGRS.2011.2151197.

[16] G. A. Antonucci, S. Vary, D. Humphreys, R. A. Lamb, J. Piper, and J. Tanner. “Multispectral Snapshot Demosaicing Via Non-Convex Matrix Completion”. In: *2019 IEEE Data Science Workshop (DSW)*. 2019, pp. 227–231. DOI: 10.1109/DSW.2019.8755561.

[17] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle. “Compressive coded aperture spectral imaging: An introduction”. In: *IEEE Signal Process. Mag.* 31.1 (2013), pp. 105–115.

[18] D. Arthur and S. Vassilvitskii. “K-Means++: The Advantages of Careful Seeding”. In: *Proc. SODA’07*. SIAM, 2007, pp. 1027–1035.

[19] D. Benachir, Y. Deville, S. Hosseini, and M. S. Karoui. “Blind unmixing of hyperspectral remote sensing data: A new geometrical method based on a two-source sparsity constraint”. In: *Remote Sensing* 12.19 (2020), p. 3198.

[20] D. Benachir, S. Hosseini, Y. Deville, M. S. Karoui, and A. Hameurlain. “Modified independent component analysis for initializing non-negative matrix factorization: An approach to hyperspectral image unmixing”. In: *Proc. IEEE ECMSM’13*. IEEE. 2013, pp. 1–6.

[21] J. S. Bhatt and M. V. Joshi. “Deep Learning in Hyperspectral Unmixing: A Review”. In: *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*. 2020, pp. 2189–2192. DOI: 10.1109/IGARSS39084.2020.9324546.

[22] L. Bian, Y. Wang, and J. Zhang. “Generalized MSFA Engineering With Structural and Adaptive Nonlocal Demosaicing”. In: *IEEE Trans. Image Process.* 30 (2021), pp. 7867–7877. DOI: 10.1109/tip.2021.3108913. URL: https://doi.org/10.1109/tip.2021.3108913.

[23] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot. "Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches". In: *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 5.2 (Apr. 2012), pp. 354–379. DOI: 10.1109/jstars.2012.2194696. URL: https://doi.org/10.1109/jstars.2012.2194696.

[24] J. D. Blanchard, J. Tanner, and K. Wei. "CGIHT: conjugate gradient iterative hard thresholding for compressed sensing and matrix completion". In: *Information and Inference: A Journal of the IMA* 4.4 (2015), pp. 289–327.

[25] C. C. Borel and S. A. Gerstl. "Nonlinear spectral mixing models for vegetative and soil surfaces". In: *Remote sensing of environment* 47.3 (1994), pp. 403–416.

[26] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, C. Richard, J. Chanussot, L. Drumetz, J.-Y. Tourneret, A. Zare, and C. Jutten. "Spectral Variability in Hyperspectral Data Unmixing: A comprehensive review". In: *IEEE Trans. Geosci. Remote Sens.* 9.4 (2021), pp. 223–270.

[27] J. Brauers and T. Aach. "A Color Filter Array Based Multispectral Camera". In: *12. Workshop Farbbildverarbeitung*. Ed. by G. C. Group. Ilmenau, Oct. 2006.

[28] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma. "A Convex Analysis-Based Minimum-Volume Enclosing Simplex Algorithm for Hyperspectral Unmixing". In: *IEEE Trans. Signal Process.* 57.11 (2009), pp. 4418–4432. DOI: 10.1109/TSP.2009.2025802.

[29] P. Chatelain, G. Delmaire, A. Alboody, M. Puigt, and G. Roussel. "Semi-Automatic Spectral Image Stitching for a Compact Hybrid Linescan Hyperspectral Camera towards Near Field Remote Monitoring of Potato Crop Leaves". In: *Sensors* 21.22 (2021), p. 7616.

[30] P. Chatelain, G. Delmaire, M. Puigt, and G. Roussel. "Inversion de réseaux de filtres de Fabry-Perot pour la restauration de cubes hyperspectraux". In: *Actes du GRETSI*. Nancy, France, Sept. 2022.

[31] P. Comon and C. Jutten. *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

[32] S. Courroux, S. Guyetant, S. Chevobbe, and M. Paindavoine. "A wavelet-based demosaicking algorithm for embedded applications". In: *2010 Conference on Design and Architectures for Signal and Image Processing (DASIP)*. 2010, pp. 90–96.

[33] M. Dalla Mura, J. Chanussot, and A. Plaza. *An Overview on Hyperspectral Unmixing*. PowerPoint Presentation. Presented at GIPSA-lab, Grenoble-INP on November 12, 2014. 2014.

[34] M. Descour and E. Dereniak. "Computed-tomography imaging spectrometer: experimental calibration and reconstruction results". In: *Appl. Opt.* 34.22 (Aug. 1995), pp. 4817–4826. DOI: 10.1364/AO.34.004817. URL: https://opg.optica.org/ao/abstract.cfm?URI=ao-34-22-4817.

[35] Y. Deville. “Sparse component analysis: A general framework for linear and nonlinear blind source separation and mixture identification”. In: *Blind Source Separation*. Springer, 2014, pp. 151–196.

[36] Y. Deville and S. Hosseini. “Blind identification and separation methods for linear-quadratic mixtures and/or linearly independent non-stationary signals”. In: *2007 9th International Symposium on Signal Processing and Its Applications*. IEEE. 2007, pp. 1–4.

[37] Y. Deville and M. Puigt. “Temporal and time-frequency correlation-based blind source separation methods. Part I: Determined and underdetermined linear instantaneous mixtures”. In: *Signal Processing* 87.3 (2007), pp. 374–407.

[38] N. Diaz, A. Alvarado, P. Meza, F. Guzmán, and E. Vera. “Multispectral Filter Array Design by Optimal Sphere Packing”. In: *IEEE Trans. Image Process.* 32 (2023), pp. 3634–3649. DOI: 10.1109/TIP.2023.3288414.

[39] K. Dijkstra, J. van de Loosdrecht, L. R. B. Schomaker, and M. A. Wiering. “Hyperspectral demosaicking and crosstalk correction using deep learning”. In: *Machine Vision and Applications* 30.1 (July 2018), pp. 1–21. DOI: 10.1007/s00138-018-0965-4. URL: https://doi.org/10.1007/s00138-018-0965-4.

[40] D. Donald, Y. Everingham, L. McKinna, and D. Coomans. “3.23 - Feature Selection in the Wavelet Domain: Adaptive Wavelets”. In: *Comprehensive Chemometrics*. Ed. by S. D. Brown, R. Tauler, and B. Walczak. Oxford: Elsevier, 2009, pp. 647–679. ISBN: 978-0-444-52701-1.

[41] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel. “Fast Nonnegative Matrix Factorization and Completion Using Nesterov Iterations”. In: *Proc. LVA/ICA’17*. Vol. 10169. LNCS. Feb. 2017, pp. 26–35. URL: https://hal.archives-ouvertes.fr/hal-01469366.

[42] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel. “Informed nonnegative matrix factorization methods for mobile sensor network calibration”. In: *IEEE Trans. Signal Inf. Process. Netw.* 4.4 (2018), pp. 667–682.

[43] J. Driesen and P. Scheunders. “Wavelet-based color filter array demosaicking”. In: *2004 International Conference on Image Processing, 2004. ICIP ’04.* Vol. 5. 2004, 3311–3314 Vol. 5.

[44] C. E. Duchon. “Lanczos filtering in one and two dimensions”. In: *Journal of Applied Meteorology and Climatology* 18.8 (1979), pp. 1016–1022.

[45] H. Fang, A. Li, H. Xu, and T. Wang. “Sparsity-Constrained Deep Nonnegative Matrix Factorization for Hyperspectral Unmixing”. In: *IEEE Geosci. Remote Sens. Lett.* 15.7 (2018), pp. 1105–1109. DOI: 10.1109/LGRS.2018.2823425.

[46] K. Feng, Y. Zhao, J. C.-W. Chan, S. Kong, X. Zhang, and B. Wang. “Mosaic Convolution-Attention Network for Demosaicing Multispectral Filter Array Images”. In: *IEEE Trans. Comput. Imaging* 7 (2021), pp. 864–878. DOI: 10.1109/tci.2021.3102052. URL: https://doi.org/10.1109/tci.2021.3102052.

[47] X.-R. Feng, H.-C. Li, R. Wang, Q. Du, X. Jia, and A. J. Plaza. *Hyperspectral unmixing based on nonnegative matrix factorization: A comprehensive review*. 2022.

[48] Y.-Z. Feng and D.-W. Sun. "Application of Hyperspectral Imaging in Food Safety Inspection and Control: A Review". In: *Critical Reviews in Food Science and Nutrition* 52.11 (2012), pp. 1039–1058. DOI: 10.1080/10408398.2011.651542.

[49] C. Févotte and N. Dobigeon. "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization". In: *IEEE Trans. Image Process.* 24.12 (2015), pp. 4810–4819.

[50] A. Figueira and B. Vaz. "Survey on synthetic data generation, evaluation methods and GANs". In: *Mathematics* 10.15 (2022), p. 2733.

[51] B. K. Ford, C. E. Volin, S. M. Murphy, R. M. Lynch, and M. R. Descour. "Computed tomography-based spectral imaging for fluorescence microscopy". In: *Biophysical Journal* 80.2 (2001), pp. 986–993.

[52] L. Gao, R. T. Kester, N. Hagen, and T. S. Tkaczyk. "Snapshot Image Mapping Spectrometer (IMS) with high sampling density for hyperspectral microscopy". In: *Opt. Express* 18.14 (July 2010), pp. 14330–14344. DOI: 10.1364/OE.18.014330. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-18-14-14330.

[53] B. Geelen, C. Blanch, P. Gonzalez, N. Tack, and A. Lambrechts. "A tiny VIS-NIR snapshot multispectral camera". In: *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics VIII*. Ed. by G. von Freymann, W. V. Schoenfeld, R. C. Rumpf, and H. Helvajian. SPIE, Mar. 2015. DOI: 10.1117/12.2077583. URL: https://doi.org/10.1117/12.2077583.

[54] B. Geelen, C. Blanch, P. Gonzalez, N. Tack, and A. Lambrechts. "A tiny VIS-NIR snapshot multispectral camera". In: *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics VIII*. Vol. 9374. International Society for Optics and Photonics. 2015, p. 937414.

[55] B. Geelen, N. Tack, and A. Lambrechts. "A snapshot multispectral imager with integrated tiled filters and optical duplication". In: *SPIE Proceedings*. Ed. by G. von Freymann, W. V. Schoenfeld, and R. C. Rumpf. SPIE, Mar. 2013. DOI: 10.1117/12.2004072. URL: https://doi.org/10.1117/12.2004072.

[56] B. Geelen, N. Tack, and A. Lambrechts. "A snapshot multispectral imager with integrated tiled filters and optical duplication". In: *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics VI*. Ed. by G. von Freymann, W. V. Schoenfeld, and R. C. Rumpf. Vol. 8613. International Society for Optics and Photonics. SPIE, 2013, p. 861314. DOI: 10.1117/12.2004072. URL: https://doi.org/10.1117/12.2004072.

[57] B. Geelen, N. Tack, and A. Lambrechts. "A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic". In: *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics VII*. Ed. by G. von Freymann, W. V. Schoenfeld, and R. C. Rumpf. SPIE, Mar. 2014. DOI: 10.1117/12.2037607. URL: https://doi.org/10.1117/12.2037607.

[58] P. Ghamisi, N. Yokoya, J. Li, W. Liao, S. Liu, J. Plaza, B. Rasti, and A. Plaza. “Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art”. In: *IEEE Geoscience and Remote Sensing Magazine* 5.4 (2017), pp. 37–78.

[59] N. Gillis. “The why and how of nonnegative matrix factorization”. In: *Regularization, optimization, kernels, and support vector machines* 12.257 (2014), pp. 257–291.

[60] N. Gillis and S. A. Vavasis. “Fast and Robust Recursive Algorithmsfor Separable Nonnegative Matrix Factorization”. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 36.4 (2014), pp. 698–714. DOI: 10.1109/TPAMI.2013.226.

[61] P. Gonzalez, B. Geelen, C. Blanch, K. Tack, and A. Lambrechts. “A CMOS-compatible, monolithically integrated snapshot-mosaic multispectral imager”. In: *NIR news* 26.4 (2015), pp. 6–11.

[62] T. Goossens, K. Vunckx, A. Lambrechts, and C. V. Hoof. “Spectral Shift Correction for Fabry-Perot Based Spectral Cameras”. In: *Proc. IEEE WHISPERS’19*. IEEE, Sept. 2019. DOI: 10.1109/whispers.2019.8920890. URL: https://doi.org/10.1109/whispers.2019.8920890.

[63] R. Gribonval and S. Lesage. “A survey of sparse component analysis for blind source separation: principles, perspectives, and new challenges”. In: *ESANN’06 proceedings-14th European Symposium on Artificial Neural Networks*. d-side publi. 2006, pp. 323–330.

[64] N. Guan, D. Tao, Z. Luo, and B. Yuan. “NeNMF: An optimal gradient method for nonnegative matrix factorization”. In: *IEEE Trans. Signal Process.* 60.6 (June 2012), pp. 2882–2898.

[65] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau. “Demosaicking: color filter array interpolation”. In: *IEEE Signal Process. Mag.* 22.1 (2005), pp. 44–54.

[66] T. A. Habtegebrial, G. Reis, and D. Stricker. “Deep Convolutional Networks For Snapshot Hypercpectral Demosaicking”. In: *Proc. IEEE WHISPERS’19*. Sept. 2019. DOI: 10.1109/whispers.2019.8921273. URL: https://doi.org/10.1109/whispers.2019.8921273.

[67] N. Hagen and E. Dereniak. “New grating designs for a CTIS imaging spectrometer - art. no. 65650N”. In: *Proceedings of SPIE - The International Society for Optical Engineering* 6565 (Apr. 2007). DOI: 10.1117/12.719533.

[68] N. Hagen, R. T. Kester, L. Gao, and T. S. Tkaczyk. “Snapshot advantage: a review of the light collection improvement for parallel high-dimensional measurement systems”. In: *Optical Engineering* 51.11 (2012), pp. 111702–111702.

[69] N. A. Hagen and M. W. Kudenov. “Review of snapshot spectral imaging technologies”. In: *Optical Engineering* 52.9 (2013), p. 090901.

[70] B. Hapke. *Theory of reflectance and emittance spectroscopy*. Cambridge university press, 2012.

[71] A. Harvey, D. Fletcher-Holmes, S. Kudesia, and C. Beggan. "Imaging spectrometry at visible and infrared wavelengths using image replication. Electro-optical and infrared systems: Technology and applications". In: *Proceedings of SPIE - The International Society for Optical Engineering* 5612 (Oct. 2004). DOI: 10.1117/12.580059.

[72] K. He, X. Zhang, S. Ren, and J. Sun. "Deep residual learning for image recognition". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

[73] R. Heylen, M. Parente, and P. Gader. "A review of nonlinear hyperspectral unmixing methods". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7.6 (2014), pp. 1844–1868.

[74] J. Hillmann. *Hyperspectral Imaging: Data Correction and Standardization, Mobile Applications*. SpectroNet Collaboration Forum. Presented on 23rd March 2016. Dresden, Germany: XIMEA GmbH, Mar. 2016. URL: https://www.ximea.com/support/attachments/5981/SpectroNet-2016-03-Ximea-V02.pdf (visited on 05/03/2024).

[75] N.-D. Ho. "Non negative matrix factorization algorithms and applications". Phd Thesis. Université Catholique de Louvain, 2008.

[76] R. Huang, X. Li, and L. Zhao. "Spectral–spatial robust nonnegative matrix factorization for hyperspectral unmixing". In: *IEEE Trans. Geosci. Remote Sens.* 57.10 (2019), pp. 8235–8254.

[77] M. Iman, H. R. Arabnia, and K. Rasheed. "A review of deep transfer learning and recent advancements". In: *Technologies* 11.2 (2023), p. 40.

[78] IMEC. *Hyperspectral Sensors User Manual*. Available online at Support Portal of imec https://www.imec-int.com/. 2019.

[79] Imec. *Real-time hyperspectral snapshot cameras*. https://www.imechyperspectral.com/en/real-time-hyperspectral-snapshot-cameras, Last accessed: 2022-06-24.

[80] T. Ince. "Superpixel-Based Graph Laplacian Regularization for Sparse Hyperspectral Unmixing". In: *IEEE Geosci. Remote Sens. Lett.* 19 (2022), pp. 1–5. DOI: 10.1109/LGRS.2020.3027055.

[81] P. Jain, P. Netrapalli, and S. Sanghavi. "Low-rank matrix completion using alternating minimization". In: *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*. 2013, pp. 665–674.

[82] G. Jaiswal, A. Sharma, and S. K. Yadav. "Critical insights into modern hyperspectral image applications through deep learning". In: *WIREs Data Mining and Knowledge Discovery* 11.6 (2021), e1426.

[83] S. Jia and Y. Qian. "Constrained nonnegative matrix factorization for hyperspectral unmixing". In: *IEEE Trans. Geosci. Remote Sens.* 47.1 (2008), pp. 161–173.

[84] W. Johnson, D. Wilson, and G. Bearman. “Spatial-spectral modulating snapshot hyperspectral imager”. In: *Applied optics* 45 (Apr. 2006), pp. 1898–908. DOI: 10.1364/AO.45.001898.

[85] M. S. Karoui, Y. Deville, S. Hosseini, and A. Ouamri. “Blind spatial unmixing of multispectral images: New methods combining sparse component analysis, clustering and non-negativity constraints”. In: *Pattern Recognition* 45.12 (2012), pp. 4263–4278.

[86] R. T. Kester, L. Gao, and T. S. Tkaczyk. “Development of image mappers for hyperspectral biomedical imaging applications”. In: *Appl. Opt.* 49.10 (Apr. 2010), pp. 1886–1899. DOI: 10.1364/AO.49.001886. URL: https://opg.optica.org/ao/abstract.cfm?URI=ao-49-10-1886.

[87] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas. “Modern Trends in Hyperspectral Image Analysis: A Review”. In: *IEEE Access* 6 (2018), pp. 14118–14129. DOI: 10.1109/ACCESS.2018.2812999.

[88] N. Kishore Kumar and J. Schneider. “Literature survey on low rank approximation of matrices”. In: *Linear and Multilinear Algebra* 65.11 (2017), pp. 2212–2244.

[89] Y. Koren, R. Bell, and C. Volinsky. “Matrix factorization techniques for recommender systems”. In: *Computer* 42.8 (2009), pp. 30–37.

[90] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton. “Multispectral filter arrays: Recent advances and practical implementation”. In: *Sensors* 14.11 (2014), pp. 21626–21659.

[91] C. Li, X. Chen, and Y. Jiang. “On diverse noises in hyperspectral unmixing”. In: *IEEE Trans. Geosci. Remote Sens.* 53.10 (2015), pp. 5388–5402.

[92] H. Li, R. Feng, L. Wang, Y. Zhong, and L. Zhang. “Superpixel-Based Reweighted Low-Rank and Total Variation Sparse Unmixing for Hyperspectral Remote Sensing Imagery”. In: *IEEE Trans. Geosci. Remote Sens.* 59.1 (2021), pp. 629–647. DOI: 10.1109/TGRS.2020.2994260.

[93] J. Li and J. M. Bioucas-Dias. “Minimum Volume Simplex Analysis: A Fast Algorithm to Unmix Hyperspectral Data”. In: *Proc. IEEE IGARSS 2008*. Vol. 3. 2008, pp. III - 250-III – 253. DOI: 10.1109/IGARSS.2008.4779330.

[94] P. Li, M. Ebner, P. Noonan, C. Horgan, A. Bahl, S. Ourselin, J. Shapey, and T. Vercauteren. “Deep learning approach for hyperspectral image demosaicking, spectral correction and high-resolution RGB reconstruction”. In: *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 10.4 (2022), pp. 409–417. DOI: 10.1080/21681163.2021.1997646.

[95] P. Li, M. Ebner, P. Noonan, C. Horgan, A. Bahl, S. Ourselin, J. Shapey, and T. Vercauteren. “Deep learning approach for hyperspectral image demosaicking, spectral correction and high-resolution RGB reconstruction”. In: *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 10.4 (2022), pp. 409–417.

[96] X. Li, J. Cui, and L. Zhao. “Blind nonlinear hyperspectral unmixing based on constrained kernel nonnegative matrix factorization”. In: *Signal, Image and Video Processing* 8.8 (Oct. 2012), pp. 1555–1567. DOI: 10.1007/s11760-012-0392-3. URL: https://doi.org/10.1007/s11760-012-0392-3.

[97] H. Liu, Y. Lu, Z. Wu, Q. Du, J. Chanussot, and Z. Wei. “Spectral variability Bayesian unmixing for hyperspectral sequence in wavelet domain”. In: *IEEE Trans. Geosci. Remote Sens.* (2023).

[98] J. Liu, S. Yuan, X. Zhu, Y. Huang, and Q. Zhao. “Nonnegative matrix factorization with entropy regularization for hyperspectral unmixing”. In: *International Journal of Remote Sensing* 42.16 (2021), pp. 6359–6390.

[99] L. Loncan. “Fusion hyperspectrale et Panchromatiques avec des Hautes résolutions spatiales”. PhD thesis. University of Grenoble, 2016.

[100] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li. “Manifold Regularized Sparse NMF for Hyperspectral Unmixing”. In: *IEEE Trans. Geosci. Remote Sens.* 51.5 (2013), pp. 2815–2826. DOI: 10.1109/TGRS.2012.2213825.

[101] D. Manolakis, R. Lockwood, and T. Cooley. *Hyperspectral Imaging Remote Sensing: Physics, Sensors, and Algorithms*. Cambridge University Press, 2016. ISBN: 9781316017876. URL: https://books.google.fr/books?id=1dLynQAACAAJ.

[102] A. McClung, S. Samudrala, M. Torfeh, M. Mansouree, and A. Arbabi. “Snapshot spectral imaging with parallel metasystems”. In: *Science Advances* 6.38 (2020), eabc7646. DOI: 10.1126/sciadv.abc7646.

[103] I. Meganem, P. Déliot, X. Briottet, Y. Deville, and S. Hosseini. “Linear–quadratic mixing model for reflectances in urban environments”. In: *IEEE Trans. Geosci. Remote Sens.* 52.1 (2013), pp. 544–558.

[104] I. Meganem, Y. Deville, S. Hosseini, P. Deliot, and X. Briottet. “Linear-quadratic blind source separation using NMF to unmix urban hyperspectral images”. In: *IEEE Trans. Signal Process.* 62.7 (2014), pp. 1822–1833.

[105] L. Miao, H. Qi, R. Ramanath, and W. Snyder. “Binary Tree-based Generic Demosaicking Algorithm for Multispectral Filter Arrays”. In: *IEEE Trans. Image Process.* 15.11 (Nov. 2006), pp. 3550–3558. DOI: 10.1109/tip.2006.877476. URL: https://doi.org/10.1109/tip.2006.877476.

[106] J. G. Michael West and C. Galvan. *Commercial Snapshot Spectral Imaging: The Art of the Possible*. Available online at https://www.mitre.org/sites/default/files/publications/pr-18-3832-commercial-snapshot-spectral-imaging-art-of-possible.pdf. Last access: 08/26/2022. 2018.

[107] S. Mihoubi, O. Losson, B. Mathon, and L. Macaire. "Multispectral Demosaicing Using Pseudo-Panchromatic Image". In: *IEEE Trans. Comput. Imaging* 3.4 (Dec. 2017), pp. 982–995. DOI: 10.1109/tci.2017.2691553. URL: https://doi.org/10.1109/tci.2017.2691553.

[108] J. Mizutani, S. S. Ogawa, K. Shinoda, M. Hasegawa, and S. Kato. "Multispectral demosaicking algorithm based on inter-channel correlation". In: *Proc. IEEE VCIP'14*. 2014, pp. 474–477.

[109] F. Movahedi Naini, G. Hosein Mohimani, M. Babaie-Zadeh, and C. Jutten. "Estimating the mixing matrix in Sparse Component Analysis (SCA) based on partial k-dimensional subspace clustering". In: *Neurocomputing* 71.10 (2008), pp. 2330–2343.

[110] J. M. P. Nascimento and J. M. Bioucas-Dias. "Vertex component analysis: a fast algorithm to unmix hyperspectral data". In: *IEEE Trans. Geosci. Remote Sens.* 43.4 (2005), pp. 898–910. DOI: 10.1109/TGRS.2005.844293.

[111] J. M. Nascimento and J. M. Dias. "Does independent component analysis play a role in unmixing hyperspectral data?" In: *IEEE Trans. Geosci. Remote Sens.* 43.1 (2005), pp. 175–187.

[112] L. T. Nguyen, J. Kim, and B. Shim. "Low-Rank Matrix Completion: A Contemporary Survey". In: *IEEE Access* 7 (2019), pp. 94215–94237. DOI: 10.1109/ACCESS.2019.2928130.

[113] B. Palsson, J. Sigurdsson, J. R. Sveinsson, and M. O. Ulfarsson. "Hyperspectral Unmixing Using a Neural Network Autoencoder". In: *IEEE Access* 6 (2018), pp. 25646–25656. DOI: 10.1109/ACCESS.2018.2818280.

[114] B. Palsson, J. R. Sveinsson, and M. O. Ulfarsson. "Blind Hyperspectral Unmixing Using Autoencoders: A Critical Comparison". In: *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 15 (2022), pp. 1340–1372. DOI: 10.1109/JSTARS.2021.3140154.

[115] Z. Pan, B. Li, Y. Bao, and H. Cheng. "Deep panchromatic image guided residual interpolation for multispectral image demosaicking". In: *2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. IEEE. 2019, pp. 1–5.

[116] Z. Pan, B. Li, H. Cheng, and Y. Bao. "Joint Demosaicking and Denoising for CFA and MSFA Images Using a Mosaic-Adaptive Dense Residual Network". In: *Computer Vision – ECCV 2020 Workshops*. Springer International Publishing, 2020, pp. 647–664. DOI: 10.1007/978-3-030-67070-2_39. URL: https://doi.org/10.1007/978-3-030-67070-2_39.

[117] I. R. Parry. "Optical fibres for integral field spectroscopy". In: *New Astronomy Reviews* 50.4 (2006). Integral Field Spectroscopy: Techniques and Data Production, pp. 301–304. ISSN: 1387-6473.

[118] J. R. Patel, M. V. Joshi, and J. S. Bhatt. "Spectral Unmixing Using Autoencoder with Spatial and Spectral Regularizations". In: *Proc. IEEE IGARSS'21*. 2021, pp. 3321–3324. DOI: 10.1109/IGARSS47720.2021.9553799.

[119] Photonfocus. *Photonfocus snapshot mosaic camera CMV2K-SM5x5*. https://www.photonfocus.com/products/camerafinder/camera/mv0-d2048x1088-c01-hs02-160-g2/, Last accessed: 2023-08-07. Switzerland.

[120] J. Pichette, T. Goossens, K. Vunckx, and A. Lambrechts. "Hyperspectral calibration method For CMOS-based hyperspectral sensors". In: *Photonic Instrumentation Engineering IV*. Ed. by Y. G. Soskind and C. Olson. Vol. 10110. International Society for Optics and Photonics. SPIE, 2017, 101100H. DOI: 10.1117/12.2253617. URL: https://doi.org/10.1117/12.2253617.

[121] H. Pu, L. Lin, and D.-W. Sun. "Principles of Hyperspectral Microscope Imaging Techniques and Their Applications in Food Quality and Safety Detection: A Review." In: *Comprehensive reviews in food science and food safety* 18 4 (2019), pp. 853–866. DOI: 10.1111/1541-4337.12432.

[122] M. Puigt. *Méthodes d'analyse en variables latentes fondées sur la parcimonie et/ou la faiblesse de rang. Application aux problèmes inverses multivariés*. HDR Thesis, Université du Littoral Côte d'Opale. Mar. 2023. URL: https://ulco.hal.science/tel-04119788.

[123] M. Puigt, O. Berné, R. Guidara, Y. Deville, S. Hosseini, and C. Joblin. "Cross-validation of blindly separated interstellar dust spectra". In: *9th International Workshop on Electronics, Control, Modelling, Measurement and Signals (ECMS 2009)*. 2009, pp–41.

[124] M. Puigt and Y. Deville. "Time–frequency ratio-based blind separation methods for attenuated and time-delayed sources". In: *Mechanical Systems and Signal Processing* 19.6 (2005), pp. 1348–1379.

[125] M. Puigt and Y. Deville. "Iterative-shift cluster-based time-frequency BSS for fractional-time-delay mixtures". In: *Proc. ICA'09*. 2009, pp. 306–313.

[126] M. Puigt, A. Griffin, and A. Mouchtaris. "Post-nonlinear speech mixture identification using single-source temporal zones & curve clustering". In: *2011 19th European Signal Processing Conference*. IEEE. 2011, pp. 1844–1848.

[127] M. Puigt, A. Griffin, and A. Mouchtaris. "Nonlinear blind mixture identification using local source sparsity and functional data clustering". In: *2012 IEEE 7th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. IEEE. 2012, pp. 481–484.

[128] L. Qi, J. Li, Y. Wang, Y. Huang, and X. Gao. "Spectral–Spatial-Weighted Multiview Collaborative Sparse Unmixing for Hyperspectral Images". In: *IEEE Trans. Geosci. Remote Sens.* 58.12 (2020), pp. 8766–8779. DOI: 10.1109/TGRS.2020.2990476.

[129] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly. "Hyperspectral Unmixing via $L_{1/2}$ Sparsity-Constrained Nonnegative Matrix Factorization". In: *IEEE Trans. Geosci. Remote Sens.* 49.11 (2011), pp. 4282–4297. DOI: 10.1109/TGRS.2011.2144605.

[130] R. Rajabi and H. Ghassemian. "Spectral Unmixing of Hyperspectral Imagery Using Multilayer NMF". In: *IEEE Geosci. Remote Sens. Lett.* 12.1 (Jan. 2015), pp. 38–42. DOI: 10.1109/lgrs.2014.2325874.

[131] V. Rathi and P. Goyal. "Convolution Filter based Efficient Multispectral Image Demosaicking for Compact MSFAs." In: *VISIGRAPP (4: VISAPP)*. 2021, pp. 112–121.

[132] V. Rathi and P. Goyal. "Generic multispectral demosaicking based on directional interpolation". In: *IEEE Access* 10 (2022), pp. 64715–64728.

[133] V. Rathi and P. Goyal. "Generic multispectral demosaicking using spectral correlation between spectral bands and pseudo-panchromatic image". In: *Signal Processing: Image Communication* 110 (2023), p. 116893.

[134] A. Remón, S. Sánchez, S. Bernabé, E. S. Quintana-Ortí, and A. Plaza. "Performance versus energy consumption of hyperspectral unmixing algorithms on multi-core platforms". In: *EURASIP Journal on Advances in Signal Processing* 2013 (2013), pp. 1–15.

[135] C. Revel, Y. Deville, V. Achard, X. Briottet, and C. Weber. "Inertia-constrained pixel-by-pixel nonnegative matrix factorisation: A hyperspectral unmixing method dealing with intra-class variability". In: *Remote Sensing* 10.11 (2018), p. 1706.

[136] K. Shinoda, S. Ogawa, Y. Yanagi, M. Hasegawa, S. Kato, M. Ishikawa, H. Komagata, and N. Kobayashi. "Multispectral filter array and demosaicking for pathological images". In: *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*. IEEE. 2015, pp. 697–703.

[137] K. Shinoda, S. Yoshiba, and M. Hasegawa. "Deep demosaicking for multispectral filter arrays". In: *arXiv preprint arXiv:1808.08021* (2018).

[138] R. Shogenji, Y. Kitamura, K. Yamada, S. Miyatake, and J. Tanida. "Multispectral imaging using compact compound optics". In: *Opt. Express* 12.8 (Apr. 2004), pp. 1643–1655.

[139] Y. Song, T. Wang, P. Cai, S. K. Mondal, and J. P. Sahoo. "A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities". In: *ACM Computing Surveys* 55.13s (2023), pp. 1–40.

[140] J. Tanner and K. Wei. "Low rank matrix completion by alternating steepest descent methods". In: *Applied and Computational Harmonic Analysis* 40.2 (2016), pp. 417–429.

[141] F. J. Theis and S.-i. Amari. "Postnonlinear overcomplete blind source separation using sparse sources". In: *Independent Component Analysis and Blind Signal Separation: Fifth International Conference, ICA 2004, Granada, Spain, September 22-24, 2004. Proceedings 5*. Springer. 2004, pp. 718–725.

[142] L. Tong, J. Zhou, Y. Qian, X. Bai, and Y. Gao. "Nonnegative-matrix-factorization-based hyperspectral unmixing with partially known endmembers". In: *IEEE Trans. Geosci. Remote Sens.* 54.11 (2016), pp. 6531–6544.

[143] G. Tsagkatakis, M. Bloemen, B. Geelen, M. Jayapala, and P. Tsakalides. "Graph and Rank Regularized Matrix Recovery for Snapshot Spectral Image Demosaicing". In: *IEEE Trans. Comput. Imaging* 5.2 (June 2019), pp. 301–316.

[144] *U.S. Geological Survey (USGS) spectral library*. https://www.usgs.gov/labs/spectroscopy-lab/science/spectral-library, Last accessed: 2022-10-24. USA.

[145] S. Van Vaerenbergh and I. Santamaría. "A spectral clustering approach to underdetermined postnonlinear blind source separation of sparse sources". In: *IEEE Trans. Neural Netw.* 17.3 (2006), pp. 811–814.

[146] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. "Attention is all you need". In: *Advances in neural information processing systems* 30 (2017).

[147] E. Vera, F. Guzmán, and N. Díaz. "Shuffled rolling shutter for snapshot temporal imaging". In: *Optics Express* 30.2 (2022), pp. 887–901.

[148] K. Vunckx and W. Charle. "Accurate Video-Rate Multi-Spectral Imaging Using IMEC Snapshot Sensors". In: *Proc. IEEE WHISPERS'21*. 2021, pp. 1–7. DOI: 10.1109/WHISPERS52202.2021.9483975.

[149] C. Wang, X. Wang, and J. Y. Hardeberg. "A linear interpolation algorithm for spectral filter array demosaicking". In: *Image and Signal Processing: 6th International Conference, ICISP 2014, Cherbourg, France, June 30–July 2, 2014. Proceedings 6*. Springer. 2014, pp. 151–160.

[150] H. Wang, W. Yang, and N. Guan. "Cauchy sparse NMF with manifold regularization: A robust method for hyperspectral unmixing". In: *Knowledge-Based Systems* 184 (2019), p. 104898.

[151] L. Wang and C. Zhao. *Hyperspectral Image Processing*. Springer Berlin Heidelberg, 2015. ISBN: 9783662474563. URL: https://books.google.fr/books?id=6kYwCgAAQBAJ.

[152] X. Wang, J.-B. Thomas, J. Y. Hardeberg, and P. Gouton. "Discrete wavelet transform based multispectral filter array demosaicking". In: *2013 Colour and Visual Computing Symposium (CVCS)*. IEEE. 2013, pp. 1–6.

[153] X. Wang, J.-B. Thomas, J. Y. Hardeberg, and P. Gouton. "Median filtering in multispectral filter array demosaicking". In: *Digital Photography IX*. Vol. 8660. SPIE. 2013, pp. 103–112.

[154] Y. W. Wang, N. P. Reder, S. Kang, A. K. Glaser, and J. T. Liu. "Multiplexed Optical Imaging of Tumor-Directed Nanoparticles: A Review of Imaging Systems and Approaches". In: *Nanotheranostics* 1.4 (2017), pp. 369–388. DOI: 10.7150/ntno.21136. URL: https://doi.org/10.7150/ntno.21136.

[155] C. Winkens, F. Sattler, V. Adams, and D. Paulus. “HyKo: A Spectral Dataset for Scene Understanding”. In: *Proc. IEEE ICCVW'17*. IEEE, Oct. 2017. DOI: 10.1109/iccvw.2017.39. URL: https://doi.org/10.1109/iccvw.2017.39.

[156] M. E. Winter. “N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data”. In: *Imaging Spectrometry V*. Vol. 3753. International Society for Optics and Photonics. 1999, pp. 266–275.

[157] E. L. Wisotzky, C. Daudkane, A. Hilsmann, and P. Eisert. “Hyperspectral demosaicing of snapshot camera images using deep learning”. In: *DAGM German Conference on Pattern Recognition*. Springer. 2022, pp. 198–212.

[158] XIMEA. *XIMEA - Hyperspectral Linescan USB3 camera 150 bands 470-900nm*. https://www.ximea.com/en/products/hyperspectral-cameras-based-on-usb3-xispec/mq022hg-im-ls150-visnir. 2023.

[159] XIMEA. *XIMEA Hyperspectral Cameras Based on USB3*. Accessed 2024. URL: https://www.ximea.com/en/products/hyperspectral-cameras-based-on-usb3-xispec/mq022hg-im-ls150-visnir.

[160] F. Xiong, Y. Qian, J. Zhou, and Y. Y. Tang. “Hyperspectral Unmixing via Total Variation Regularized Nonnegative Tensor Factorization”. In: *IEEE Trans. Geosci. Remote Sens.* 57.4 (2019), pp. 2341–2357. DOI: 10.1109/TGRS.2018.2872888.

[161] F. Xiong, J. Zhou, and Y. Qian. “Material Based Object Tracking in Hyperspectral Videos”. In: *IEEE Trans. Image Process.* 29 (2020), pp. 3719–3733. DOI: 10.1109/tip.2020.2965302. URL: https://doi.org/10.1109/tip.2020.2965302.

[162] Z. Xiong, Z. Shi, H. Li, L. Wang, D. Liu, and F. Wu. “Hscnn: Cnn-based hyperspectral image recovery from spectrally undersampled projections”. In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2017, pp. 518–525.

[163] J. Xue, Y.-Q. Zhao, Y. Bu, W. Liao, J. C.-W. Chan, and W. Philips. “Spatial-spectral structured sparse low-rank representation for hyperspectral image super-resolution”. In: *IEEE Trans. Image Process.* 30 (2021), pp. 3084–3097.

[164] F. Yahaya. “Compressive informed (semi-)non-negative matrix factorization methods for incomplete and large-scale data : with application to mobile crowd-sensing data”. PhD thesis. Université du Littoral Côte d'Opale, Nov. 2021. URL: https://theses.hal.science/tel-03616665.

[165] F. Yahaya, M. Puigt, G. Delmaire, and G. Roussel. “How to Apply Random Projections to Nonnegative Matrix Factorization with Missing Entries?” In: *Proc. EUSIPCO'19*. IEEE, Sept. 2019. DOI: 10.23919/eusipco.2019.8903036. URL: https://doi.org/10.23919/eusipco.2019.8903036.

[166] F. Yahaya, M. Puigt, G. Delmaire, and G. Roussel. "Random Projection Streams for (Weighted) Nonnegative Matrix Factorization". In: *Proc. IEEE ICASSP'21*. IEEE, June 2021. DOI: 10.1109/icassp39728.2021.9413496. URL: https://doi.org/10.1109/icassp39728.2021.9413496.

[167] B. Yang, B. Wang, and Z. Wu. "Unsupervised nonlinear hyperspectral unmixing based on bilinear mixture models via geometric projection and constrained nonnegative matrix factorization". In: *Remote Sensing* 10.5 (2018), p. 801.

[168] S. Yang, H. Qin, X. Yan, S. Yuan, and Q. Zeng. "Mid-Wave Infrared Snapshot Compressive Spectral Imager with Deep Infrared Denoising Prior". In: *Remote Sensing* 15.1 (2023). ISSN: 2072-4292.

[169] J. Ye. "Generalized low rank approximations of matrices". In: *Proceedings of the twenty-first international conference on Machine learning*. 2004, p. 112.

[170] O. Yilmaz and S. Rickard. "Blind separation of speech mixtures via time-frequency masking". In: *IEEE Trans. Signal Process.* 52.7 (2004), pp. 1830–1847.

[171] N. Yokoya, J. Chanussot, and A. Iwasaki. "Nonlinear unmixing of hyperspectral data using semi-nonnegative matrix factorization". In: *IEEE Trans. Geosci. Remote Sens.* 52.2 (2013), pp. 1430–1437.

[172] X. Yuan. "Generalized alternating projection based total variation minimization for compressive sensing". In: *2016 IEEE International Conference on Image Processing (ICIP)*. 2016, pp. 2539–2543. DOI: 10.1109/ICIP.2016.7532817.

[173] Y. Yuan, Y. Feng, and X. Lu. "Projection-based NMF for hyperspectral unmixing". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8.6 (2015), pp. 2632–2643.

[174] H. Zeng, K. Feng, S. Huang, J. Cao, Y. Chen, H. Zhang, H. Luong, and W. Philips. "MSFA-Frequency-Aware Transformer for Hyperspectral Images Demosaicing". In: *arXiv preprint arXiv:2303.13404* (2023).

[175] G. Zhang, S. Mei, B. Xie, M. Ma, Y. Zhang, Y. Feng, and Q. Du. "Spectral variability augmented sparse unmixing of hyperspectral images". In: *IEEE Trans. Geosci. Remote Sens.* 60 (2022), pp. 1–13.

[176] S. Zhang, J. Li, H.-C. Li, C. Deng, and A. Plaza. "Spectral–Spatial Weighted Sparse Regression for Hyperspectral Image Unmixing". In: *IEEE Trans. Geosci. Remote Sens.* 56.6 (2018), pp. 3265–3276. DOI: 10.1109/TGRS.2018.2797200.

[177] S. Zhang, W. Wang, J. Ford, and F. Makedon. "Learning from incomplete ratings using non-negative matrix factorization". In: *Proc. SIAM ICDM'06*. SIAM. 2006, pp. 549–553.

[178] T. Zhang, Z. Liang, and Y. Fu. "Joint Spatial-Spectral Pattern Optimization and Hyperspectral Image Reconstruction". In: *IEEE J. Sel. Topics Signal Process.* 16.4 (2022), pp. 636–648. DOI: 10.1109/JSTSP.2022.3179806.

[179] T. Zhang, Z. Liang, and Y. Fu. “Joint spatial-spectral pattern optimization and hyperspectral image reconstruction”. In: *IEEE Journal of Selected Topics in Signal Processing* 16.4 (2022), pp. 636–648.

[180] X. Zhang, Y. Dai, G. Zhang, X. Zhang, and B. Hu. “A Snapshot Multi-Spectral Demosaicing Method for Multi-Spectral Filter Array Images Based on Channel Attention Network”. In: *Sensors* 24.3 (2024), p. 943.

[181] X. Zhang, G. Zhang, and B. Hu. “A generic multispectral demosaicking method based on inter-channel spectrum correlation”. In: *Third International Computing Imaging Conference (CITA 2023)*. Vol. 12921. SPIE. 2023, pp. 293–299.

[182] F. Zhu and P. Honeine. “Biobjective Nonnegative Matrix Factorization: Linear Versus Kernel-Based Models”. In: *IEEE Trans. Geosci. Remote Sens.* 54.7 (July 2016), pp. 4012–4022. DOI: 10.1109/tgrs.2016.2535298. URL: https://doi.org/10.1109/tgrs.2016.2535298.

[183] F. Zhu, Y. Wang, B. Fan, S. Xiang, G. Meng, and C. Pan. “Spectral unmixing via data-guided sparsity”. In: *IEEE Trans. Image Process.* 23.12 (2014), pp. 5412–5427.