

Image Recovery Algorithm from Random Spectral Projections via Low-Rank Regularizations

Tatiana Carolina Gélvez Barrera

Doctoral thesis to qualify for the Doctor of Engineering degree, electronic area

Advisor

PhD. Henry Arguello Fuentes

Department of Systems Engineering, and Informatics

Universidad Industrial de Santander

School of Physical-Mechanical Engineering

Department of Electrical, Electronics, and Telecommunications Engineering

Bucaramanga

2022

### **Dedication**

I lovingly dedicate my dissertation to,

- My almighty Lord God, for the guidance, strength, protection and received gifts and skills.
- My beloved parents, Carmenza and José, who have been my source of inspiration and provided the moral, spiritual, emotional, and financial support unconditionally with every decision I have made.
- My brother and my sister, Harley and Mayra, who have shared their voice of advice and encouragement to follow my dreams.
- My treasured aunt Rosa Delia and my family, for her love and words of encouragement.
- My dear friend Ana Remolina and her family, all my gratitude for your love and for always being present.
- My darling Fabián and his family, for his love, patience, and being a source of strength.
- Colombians who are pursuing an academic goal.

### **Acknowledgements**

I express my gratitude to my Alma Mater, the Universidad Industrial de Santander. Thanks to my country Colombia and compatriots, this could not be possible without the public university and the contributions of the Colombian people. Thanks to COLCIENCIAS and COLFUTURO for the financial and administrative support.

I gratefully acknowledge the scientific and administrative staff from the High Dimensional Signal Processing (HDSP) research group for their support during my doctoral training. Thanks to my advisor, professor Henry Arguello for his supervision and advice.

I would like to offer my special thanks to Professor Alessandro Foi for his technical support in my research, and financial support during my stay in Tampere, Finland.

My gratitude extends to my friends, lab mates, colleagues, and research team for the hard work and a cherished time spent together in the lab, trips, and social meetings.

Thanks to the examination committee for the dedicated time, effort, and feedback.

**Table of Contents**

<b>Abbreviations</b>	<b>12</b>
<b>Symbols</b>	<b>15</b>
<b>Introduction</b>	<b>21</b>
<b>1. Spectral imaging theoretical background</b>	<b>29</b>
1.1. Hyperspectral and multispectral imaging . . . . .	30
1.2. Compressive spectral imaging . . . . .	31
1.3. Spectral imaging recovery problems . . . . .	32
1.3.1. Spectral image denoising . . . . .	33
1.3.2. Single hyperspectral image super-resolution . . . . .	33
1.3.3. Multispectral-hyperspectral image fusion . . . . .	33
1.3.4. Compressive spectral imaging reconstruction . . . . .	34
1.3.5. Compressive spectral imaging fusion . . . . .	34
1.4. Low-rank regularization . . . . .	35
1.4.1. Global spectral low-rank prior . . . . .	35
1.4.2. Local and non-local spatial low-rank prior . . . . .	37
<b>2. Global low-rank regularization for compressive spectral imaging reconstruction</b>	<b>38</b>

2.1.	Compressive spectral imaging reconstruction related works . . . . .	39
2.2.	Global spectral low-rank prior analysis . . . . .	40
2.3.	Optimization problem formulation for compressive spectral imaging reconstruction	43
2.4.	Joint sparse and low-rank compressive spectral imaging algorithm . . . . .	44
2.4.1.	Alternating minimization . . . . .	45
2.4.2.	Analysis of the computational complexity . . . . .	49
2.4.3.	Convergence analysis . . . . .	50
2.5.	Simulations and results . . . . .	52
2.5.1.	Regularization parameters analysis . . . . .	53
2.5.2.	Performance evaluation . . . . .	54
2.6.	Conclusions . . . . .	61
<b>3.</b>	<b>Non-local low-rank abundance prior for compressive spectral imaging fusion</b>	<b>63</b>
3.1.	Compressive spectral imaging fusion related works . . . . .	64
3.2.	Global and non-local low-rank regularization analysis . . . . .	65
3.2.1.	Global low-rank theoretical premise . . . . .	65
3.2.2.	Non-local low-rank empirical study . . . . .	66
3.3.	Optimization problem formulation for compressive spectral imaging fusion . . . . .	69
3.4.	Numerical algorithm . . . . .	71
3.4.1.	Abundance estimation algorithm . . . . .	72
3.4.2.	Endmember estimation algorithm . . . . .	78

3.4.3.	General scheme . . . . .	79
3.5.	Simulations and results . . . . .	81
3.5.1.	Implementation details . . . . .	81
3.5.2.	Performance evaluation . . . . .	85
3.6.	Conclusions . . . . .	89
<b>4.</b>	<b>Joint non-local, spectral, and similarity low-rank priors for spectral imaging fusion</b>	<b>90</b>
4.1.	Hyperspectral-multispectral image fusion related works . . . . .	91
4.2.	Rank-one similarity prior . . . . .	93
4.2.1.	Concept . . . . .	93
4.2.2.	Concept test . . . . .	94
4.3.	Proposed hyperspectral-multispectral image fusion . . . . .	98
4.4.	PnP-ADMM algorithm . . . . .	98
4.4.1.	Initialization . . . . .	99
4.4.2.	Precomputation . . . . .	100
4.4.3.	Plug-and-Play of the Multichannel BM3D . . . . .	100
4.5.	Simulations and results . . . . .	104
4.5.1.	Hyperspectral dataset acquisition . . . . .	104
4.5.2.	Comparative benchmarking . . . . .	106
4.5.3.	Parameters selection . . . . .	109
4.5.4.	Rank analysis . . . . .	111

4.6. Discussion . . . . .	112
4.7. Conclusions . . . . .	113
<b>5. Mixture-Net: Low-Rank Deep Image Prior Inspired by Mixture Models for Spectral Image Recovery</b>	<b>114</b>
5.1. Spectral image recovery related works . . . . .	115
5.2. Proposed spectral image recovery method . . . . .	116
5.2.1. Input structure component . . . . .	117
5.2.2. Mixture-Net: Network architecture component . . . . .	118
5.2.3. Customized loss function component . . . . .	121
5.3. Simulations and Results . . . . .	121
5.3.1. Characterization of Mixture-Net . . . . .	123
5.3.2. Mixture-Net Performance for Spectral Image Recovery . . . . .	128
5.3.3. High-level Tasks Experiments . . . . .	134
5.4. Conclusions . . . . .	136
<b>Conclusions</b>	<b>139</b>
<b>References</b>	<b>142</b>
<b>References</b>	<b>142</b>
<b>Appendices</b>	<b>165</b>

**List of figures**

Figure 1. Spectral imaging acquisition techniques. . . . . 29

Figure 2. Compressive optical systems. . . . . 33

Figure 3. Example of the Pearson correlation matrix analysis. . . . . 41

Figure 4. Pearson correlation matrix analysis for nine datasets. . . . . 42

Figure 5. Parameter analysis of the proposed CSI reconstruction algorithm. . . . . 54

Figure 6. Recovered spectral distributions visual comparison . . . . . 60

Figure 7. Reconstructed spatial distribution visual comparison. . . . . 61

Figure 8. HRI visual representations. . . . . 65

Figure 9. Singular value decay for non-local patch-groups. . . . . 68

Figure 10. Scheme of the CSI fusion approach with non-local low-rank prior. . . . . 76

Figure 11. Selection of patch-size experiment. . . . . 82

Figure 12. Algorithm parameters experiment. . . . . 84

Figure 13. Visual comparison of the spectral reconstructions. . . . . 87

Figure 14. RGB representation visual comparison. . . . . 88

Figure 15. Illustration of the rank-one similarity prior in the Colombia dataset. . . . . 95

Figure 16. Empirical validation of the rank-one similarity prior. . . . . 97

Figure 17. Propagation of the block-matching over the PCs. . . . . 102

Figure 18. Hyperspectral dataset acquisition sytem . . . . . 104

Figure 19. RGB mapping of the acquired datasets. . . . . 105

Figure 20. Spectral signature comparison at a random spatial location P1. . . . . 107

Figure 21. RGB mapping of the recovered datasets. . . . . 108

Figure 22. Analysis of fusion quality vs block contrast. . . . . 109

Figure 23. Parameters Analysis for each dataset. . . . . 110

Figure 24. Analysis of the number of preserved PCs during the iterations of R1BM3D. . . 111

Figure 25. Proposed SI recovery method scheme. . . . . 117

Figure 26. Input structure quantitative performance box-plot. . . . . 125

Figure 27. Architectures quantitative performance. . . . . 126

Figure 28. Non-linearity block-layer performance. . . . . 127

Figure 29. Multiple deep-blocks quantitative performance. . . . . 129

Figure 30. Visual RGB mapping of the denoised Pavia University. . . . . 131

Figure 31. RGB representation of the reconstructed composite Pavia Center dataset . . . . 132

Figure 32. Recovery quality comparison. . . . . 135

Figure 33. Unmixing application. . . . . 136

Figure 34. Material identification application . . . . . 137

Figure 35. RGB mapping of the ARAD database. . . . . 166

Figure 36. RGB mapping of the CAVE database. . . . . 167

Figure 37. RGB mapping of the KAIST database. . . . . 167

Figure 38. RGB mapping of the remote sensing datasets. . . . . 170

**List of tables**

Table 1.	Reconstruction quality - CASSI architecture . . . . .	55
Table 2.	Reconstruction quality - SSCSI architecture . . . . .	56
Table 3.	Reconstruction quality vs captured data - CASSI architecture . . . . .	57
Table 4.	Reconstruction quality vs captured data - SSCSI architecture . . . . .	57
Table 5.	Reconstruction quality vs spatial-spectral resolution ratio - CASSI architecture .	58
Table 6.	Reconstruction quality vs spatial-spectral resolution ratio - SSCSI architecture .	59
Table 7.	Summary variable computation in the abundance estimation . . . . .	73
Table 8.	Algorithm Parameters . . . . .	83
Table 9.	Fusion metrics for Jasper database . . . . .	85
Table 10.	Fusion metrics for Urban database . . . . .	85
Table 11.	Fusion metrics for Pavia database . . . . .	86
Table 12.	Summary of variables computation . . . . .	100
Table 13.	HSI-MSI Fusion Metrics: RMSE, ERGAS, SAM (degrees), UIQI, PSNR (dB). .	106
Table 14.	Denoising Quantitative Results for Pavia University . . . . .	130
Table 15.	Single Hyperspectral Super-Resolution Quantitative Results for Pavia Center . .	133
Table 16.	Single Hyperspectral Super-Resolution Quantitative Results for Stuff-Toys . . .	134

### Abbreviations

2D	bidimensional 29, 35, 39, 65
3D	three-dimensional 21, 29, 118, 129
ADMM	alternating direction method of multipliers 26, 71, 72, 77– 79, 90
AVIRIS	airborne visible/infrared imaging spectrometer 30, 168
BCDM	block-coordinate descent method 26, 38, 63, 71, 80, 89
C-CASSI	colored-CASSI system 27, 31–33, 81–83, 85, 86
CASSI	coded aperture snapshot spectral imaging system 27, 31– 33, 40, 53, 55, 56, 58–63, 89, 140
CNN	convolutional neural network 36
CSI	compressive spectral imaging 21, 24, 26, 27, 31, 38–40, 42, 43, 48, 49, 51–55, 59, 61–64, 69, 75, 79, 81, 85, 89, 114, 116, 122, 128, 132, 137
DIP	deep image prior 22, 24, 26, 115, 116, 124, 130, 133, 134, 136
DL	deep-learning 22, 24, 114, 115, 130, 136, 137, 140

ERGAS	Relative dimensionless global error 81, 85, 86
FPA	focal plane array 31, 32
GPSR	gradient projection for sparse reconstruction 39, 50, 52, 53
HRI	high spatial-spectral resolution image 21, 30, 33, 34, 37, 64–67, 69, 71, 75, 79, 81–83, 85, 86, 90–93, 98–100, 111– 113
HSI	hyperspectral image 21, 30, 33, 34, 64, 81, 83, 91, 97, 98, 105, 172
HSI-MSI	hyperspectral-multispectral image 21, 22, 24, 26, 33, 34, 90–92, 97, 98, 104, 105, 113, 141
HSI-SR	hyperspectral super-resolution 21, 24, 27, 33, 122, 123, 130, 131, 137, 141
LMM	linear mixture model 35, 36, 63–66, 69, 70, 75, 89, 116, 125, 140
MSE	mean square error 170
MSI	multispectral image 21, 30, 31, 34, 64, 69, 75, 81–83, 85, 91, 97, 98, 110
NLMM	non-linear mixture model 36, 117, 120

PnP	plug-and-play 26, 45, 90
PSNR	peak signal to noise ratio 53–59, 61, 62, 81–83, 85, 86, 170, 171
ROSI	reflective optics system imaging spectrometer 30, 168, 169
SAM	Spectral angle mapper 81, 85, 86
SI	spectral image 20–27, 29, 31–36, 38–40, 43, 45–48, 50, 57, 58, 61, 89, 91, 114–116, 118–122, 129, 130, 136–141, 166
SSCSI	spatial-spectral encoded compressive hyperspectral imager 27, 32, 33, 40, 53, 55, 56, 58–63, 81–83, 85, 86, 89, 140
SSIM	structural similarity 53–59, 171
SVD	singular value decomposition 46, 50, 62, 74, 81, 89
TV	total-variation 22, 24
UIQI	universal image quality index 81, 85, 86, 171

### Symbols

$\mathbf{a}_i$	Fractional proportions at the $i^{th}$ spatial location 35–37
$\bar{\mathbf{a}}$	Abundances 36, 37
$\mathbf{B}$	$\in \mathbb{R}^{N^2 \times N^2}$ Spatial blurring convolution matrix 15, 30, 33, 34
$\bar{\mathbf{B}}$	$\in \mathbb{R}^{N^2 L \times N^2 L}$ Spatial blurring convolution matrix 30, 33, 34
$d_h$	Spatial decimation factor 30, 33, 172
$d_m$	Spectral degradation factor 31, 33
$\mathbf{E}$	Endmembers 35–37, 65, 66, 69–73, 75, 78–80
$\mathbf{F}$	$\in \mathbb{R}^{N^2 \times L}$ Matrix notation of a spectral image 35, 43–47, 49, 65, 66
$\mathbf{f}$	$\in \mathbb{R}^n, n = N^2 L$ Vector notation of a spectral image 29, 30, 32, 34–36, 39, 40, 43–49, 170–172
$\mathbf{f}_i$	$\in \mathbb{R}^L$ Spectral distribution of the $i^{th}$ spatial pixel 29
$\mathbf{f}^\ell$	$\in \mathbb{R}^{N^2}$ Spatial distribution of the $\ell^{th}$ spectral band 29, 35, 170
$\mathcal{F}$	$\in \mathbb{R}^{N \times N \times L}$ Tensor notation of a spectral image 29, 65
$\mathbf{f}_h$	$\in \mathbb{R}^{N_h^2 L}$ Vector notation of an HSI 30, 33

- $\hat{\mathbf{f}}$   $\in \mathbb{R}^n, n = N^2L$  Vector notation of the estimated spectral image 34, 43–45, 47, 49, 170–172
- $\mathbf{f}_m$   $\in \mathbb{R}^{N^2L_m}$  Vector notation of a MSI 30, 33
- $\gamma$  Compression ratio 34
- $\mathbf{H}$   $\in \mathbb{R}^{m \times n}$  Compressive sensing matrix 34, 43–45, 47–51
- $\mathbf{I}_L$   $L \times L$  identity matrix 30
- $\mathbf{I}_{N^2}$   $N^2 \times N^2$  identity matrix 30
- $\otimes$  Kronecker product 30, 31, 36
- $L$  Spectral resolution indicating the number of spectral bands. 15, 29–31, 35, 36, 43–47, 50, 57, 170
- $\lambda$  Spectral dimension 29, 32
- $L_m$  MSI spectral resolution indicating the number of spectral bands. 16, 17, 30, 31
- $\mathbf{M}$   $\in \mathbb{R}^{N_h^2 \times N^2}$  Uniform spatial sub-sampling matrix. 30
- $m$  Number of observed measurements 17, 31, 33, 34, 50
- $\bar{\mathbf{M}}$   $\in \mathbb{R}^{N_h^2 L \times N^2 L}$  Uniform spatial sub-sampling matrix. 30, 33, 34
- $N$  Number of pixels at one spatial dimension 29, 31, 50

- $n$  spatial-spectral resolution of a spectral image 29, 33, 34, 39, 43, 50, 51
- $N^2$  HRI spatial resolution indicating the number of spatial pixels. 15, 16, 29–31, 35, 36, 41, 43–47, 49–51, 57
- $N_h^2$  HSI spatial resolution indicating the number of spatial pixels. 15, 30
- $\omega$   $\in \mathbb{R}^m$  Acquisition process noise 32, 33
- $\omega_h$  HSI additive Gaussian noise 30
- $\omega_m$  MSI additive Gaussian noise 30
- $\Phi$   $\in \mathbb{R}^{m \times n}$  Sensing matrix 32–34
- $\Psi$   $\in \mathbb{R}^{n \times n}$  Orthonormal representation basis 39, 43–45, 47, 49
- $Q$  Number of acquired snapshots 31, 32
- $\mathbf{R}$   $\in \mathbb{R}^{L_m \times L}$  Multispectral sensor response 30
- $r$  Rank of the spectral image 35, 36, 53, 65, 70–74
- $\bar{\mathbf{R}}$   $\in \mathbb{R}^{N_m^2 L_m \times N^2 L}$  Multispectral sensor response 30, 31, 33, 34
- $s$  Sparsity level of a spectral image 39
- $\theta$   $\in \mathbb{R}^n$  Spectral image sparse representation 39, 44, 45, 47, 49

$(x, y)$  Spatial dimension 29, 32

$\mathbf{y} \in \mathbb{R}^m$  Observed measurements 32–34, 43–45, 47, 49

## Resumen

**Título:** Algoritmo de reconstrucción de imágenes a partir de proyecciones espectrales aleatorias usando regularizadores de bajo rango<sup>1</sup>

**Autor:** TATIANA CAROLINA GÉLVEZ BARRERA \*\*

**Palabras Clave:** Adquisición de imágenes espectrales, problemas de recuperación de imágenes espectrales, información previa de bajo rango, términos regularizadores.

**Descripción:** El uso de información previa es crucial en la solución de problemas inversos mal condicionados en el procesamiento de imágenes. Esta tesis estudia el denominado bajo rango como información previa de una imagen espectral (IE) abordando la pregunta de investigación ¿cómo aprovechar la propiedad de bajo rango para resolver problemas de recuperación de imágenes espectrales? Una IE puede ser modelada como un arreglo tridimensional con dos dimensiones espaciales y una dimensión espectral. Típicamente, las escenas naturales contienen redundancia en las respuestas espectrales y similitudes en las estructuras espaciales, tal que la propiedad de bajo rango indica que una IE se encuentra en un subespacio de baja dimensión. Sin embargo, la literatura omite comúnmente un aspecto crítico relacionado con la distinta connotación que tiene la propiedad de bajo rango en las dimensiones espaciales y espectral. Por lo tanto, esta tesis analiza la propiedad de bajo rango por medio de premisas teóricas y estudios empíricos para determinar cómo emplear el bajo rango, obteniendo las siguientes contribuciones principalmente. Un enfoque que incorpora la propiedad de bajo rango *implícitamente* en la arquitectura de una red neuronal, proponiendo una alternativa a la formulación típica de una función de regularización. Además, se introduce el concepto de similitud de rango-uno asumiendo que las correlaciones estructurales en una IE son preservadas a lo largo de las bandas espectrales. Este concepto determina una dimensión adicional a las dimensiones espaciales y espectral para emplear la propiedad de bajo rango. Finalmente, se presenta la implementación y desarrollo de cuatro algoritmos de recuperación de imágenes espectrales, cubriendo formulaciones de optimización convexa y aprendizaje profundo que consideran la propiedad de bajo rango en la dimensión espacial, espectral, de similitud, y desde el punto de vista global y no-local. En particular, se abordan los problemas de eliminación de ruido, superresolución, fusión, fusión compresiva, y reconstrucción compresiva, superando los métodos de referencia comparativa en métricas de calidad objetivas y visuales.

---

<sup>1</sup> Tesis doctoral

\*\* Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingenierías Eléctrica, Electrónica y telecomunicaciones.  
Director: Henry Arguello Fuentes.

## Abstract

**Title:** Image Recovery Algorithm from Random Spectral Projections via Low-Rank Regularizations \*

**Author:** TATIANA CAROLINA GELVEZ BARRERA \*\*

**Keywords:** Spectral imaging, spectral image recovery problems, low-rank prior, regularization terms.

**Description:** Using prior information is crucial for solving ill-posed inverse problems in image processing. This thesis aims to study the so-called low-rank prior exhibited by a spectral image (SI), addressing the research question of how to take advantage of the low-rank prior for solving spectral imaging recovery problems. A SI can be modeled as a three-dimensional array with two spatial and one spectral dimension. Natural scenes typically contain redundant spectral responses and self-similar spatial structures so that the low-rankness indicates that a SI lies in a low-dimensional subspace. However, the literature commonly neglects a critical issue related to the different low-rank prior connotations across the spatial and spectral dimensions. Therefore, this thesis studies the low-rank prior through theoretical premises and extensive empirical studies to determine how to apply the low-rank prior, obtaining the following main contributions. First, I present a strategy that introduces the low-rank regularization *implicitly* in the architecture of a neural network, providing an alternative to the traditional formulation of a regularization function. Furthermore, I introduce the concept of rank-one similarity prior under the idea that the structural correlations in a SI are preserved across the spectral bands. This concept determines an additional dimension apart from the traditional spatial and spectral dimensions to consider the low-rank prior. Finally, I present the deployment and development of four practical spectral imaging recovery algorithms, covering convex optimization and deep learning formulations that consider the low-rank prior in the spatial, spectral, and similarity dimensions and from the global and non-local points of view. Notably, I address SI denoising, single hyperspectral image super-resolution, spectral imaging fusion, compressive spectral imaging fusion, and compressive spectral imaging reconstruction, overcoming the comparative benchmark methods in quality objective metrics and visual quality.

---

\* Doctoral Thesis

\*\* Department of Electrical, Electronics, and Telecommunication Engineering. Department of Electrical, Electronics, and Telecommunication Engineering. Advisor: Henry Arguello Fuentes.

## Introduction

**Spectral imaging** sensors collect spatial-spectral information of a scene through specialized optical systems. The information is commonly organized in a three-dimensional (3D) datacube, referred to as a *spectral image* (SI), with two spatial and one spectral dimension (Khan et al., 2018). A SI can be called a multispectral image (MSI) or a hyperspectral image (HSI) associating dozens or hundreds of spectral bands, respectively (Lu et al., 2019). Each SI spatial location contains a spectral distribution, dubbed *spectral signature*, that uniquely characterizes the captured material. Hence, spectral imaging supports remote sensing applications such as precision agriculture, classification, and material identification (Vargas et al., 2019).

SI acquisition is commonly affected by systematic artifacts linked with external conditions and optical aberrations that corrupt, degrade, or diminish the amount of captured **random spectral projections**, referred to as *measurements* (Amigo and Santos, 2020). For instance, the presence of such artifacts generates many challenges in the SI analysis, so that correct post-processing to recover the underlying scene is a critical step (Chang et al., 2020a). The photon efficiency, light source nature, and atmospheric conditions can add Gaussian noise, outlines, and striping, leading to a denoising problem (Xue et al., 2019). Moreover, SI sensing technologies face a trade-off between the spatial and spectral resolution, limiting the acquisition of a simultaneously high spatial-spectral resolution image (HRI), which has led to a single HSI super-resolution (HSI-SR) or HSI-MSI fusion problem (Guilloteau et al., 2020). The compressive spectral imaging (CSI) paradigm provides few compressed measurements, leading to a reconstruction problem (Barajas et al., 2020). Going

further, some methods aim to perform high-level imaging tasks from the observations, yielding to problems such as clustering (Hinojosa et al., 2018), classification (Sanchez et al., 2019) and compressive HSI-MSI fusion (Vargas et al., 2018a).

**SI recovery** is an active research field addressing the aforementioned ill-posed inverse problems, where there exist several feasible scenes that match the known measurements. Therefore, using **prior information** related to the SI natural characteristics is crucial to counter the ill-posing and effectively find a solution. The mainstream of the employed SI prior information includes (i) *Wavelet-based representation* assuming particular SI properties in the Wavelet transformation (Vaish et al., 2019); (ii) *joint-sparsity* assuming that the SI can be represented as the product of two matrices where only a few rows contain nonzero entries (Gao et al., 2020); (iii) *sparsity* assuming a SI sparse representation with few nonzero coefficients in a given orthonormal basis (Correa et al., 2016); (iv) *total-variation* (TV) assuming smooth transitions within the SI spatial locations (Zeng et al., 2020); and (v) *low-rankness* assuming a SI low-dimensional subspace representation by cause of high spatial-spectral structural correlations. (Fu et al., 2016).

The SI prior information can be included in greedy algorithms, model-based optimization, deep-learning (DL), and deep image prior (DIP) approach. The Greedy strategy considers the prior information through analytic, heuristic, and systematic procedures to recover the SI (Mäkinen et al., 2020). The model-based optimization includes the prior information through hand-crafted regularization terms in an optimization problem (Chen et al., 2018). The data-driven DL strategy learns the prior information or the mapping operator by processing a vast amount of available datasets (Ongie et al., 2020). The non-data-driven DIP approach captures the low-level statistics of

the SI via a single generator network structure without data-training (Ulyanov et al., 2018).

SI recovery strategies taking advantage of priors based on a representation basis have obtained satisfactory performance. Nonetheless, the choice of the representation basis is a critical issue to obtain such good results (León-López and Fuentes, 2020). Using sparsity implies the prior knowledge of an orthonormal basis or a dictionary learning with a large amount of data to be processed (Tu et al., 2020). On the other hand, despite the availability of traditional basis such as the Cosine, Wavelet or Karhunen-Loeve transforms, the optimal representation basis varies according to the scene. For instance, the representation basis may vary if the scene contains landscape or city objects, or if the scene is static or dynamic (Correa-Pugliese et al., 2016).

Incorporating the low-rank prior through efficient strategies has attracted widespread interest in the SI analysis as an alternative to the traditional approaches based on a representation basis. Specifically, a natural and noiseless SI commonly contains a few different materials represented with unique spectral signatures repeated at various spatial locations, generating high redundancy across the spectral domain (Ongie and Jacob, 2017). Besides, the same structure is commonly repeated at various sub-regions, generating high local and nonlocal spatial self-similarities. These spatial-spectral correlations enable the use of a low-rank prior where the SI can be low-rank approximated (Chang et al., 2020a). In general, the inclusion of low-rank regularizations for SI recovery problems from random spectral projections have faced the following **challenges**:

1. **Analysis of the SI low-rank prior:** The low-rank prior has been applied in different modalities to promote the low-rankness across the spatial or spectral dimension. For instance, (Liu et al., 2018) adopt the low-rank recovery with a local neighborhood weighted spatial-spectral

TV model. (Zeng et al., 2020) adopt a low-rank tensor modeling with a global sparse and TV regularizer. (Xue et al., 2017b) combine the sparsity and non-local low-rank property to eliminate the noise. Still, these works do not consider the analysis and validation of the low-rank property along the spatial and spectral dimensions and from the global and non-local point of view over multiple spectral datasets.

2. **Incorporation of the low-rank prior in SI recovery:** There exist various strategies in which the low-rank prior can be incorporated in the solution of SI recovery problems. Model-based optimization employs hand-crafted regularization terms to introduce the low-rank prior in an optimization problem (Chen et al., 2019; Chang et al., 2020b). However, the regularizations usually do not cover the wide variety of SIs. On the other hand, DL-based approaches combine the learning of a black-box operator from available datasets with a low-rank property (Song et al., 2019; Cao et al., 2019b). However, DL approaches require a vast amount of training datasets, which results challenging in the spectral imaging field. Still, developing alternative model-based, DL-based or DIP-based strategies to incorporate the low-rank prior across the spatial and spectral dimension is an active research field.
3. **Development of SI recovery methods using the low-rank prior:** SI recovery such as SI denoising, single HSI-SR, CSI reconstruction, HSI-MSI fusion, and compressive HSI-MSI fusion are commonly highly ill-posed problems. Therefore, improving the recovery quality of these problems by developing methodologies that incorporate the studied low-rank property with the proposed regularizations is a challenge.

This thesis addressed the aforementioned challenges by achieving the following **objectives**.

**General objective.** To develop low-rank regularizations for recovering SIs from encoded random spectral projections in order to improve the reconstructed image quality.

**Specific objectives.**

1. To analyze the low-rank properties of SIs to validate the assumption that a low-rank regularization could improve the performance of the spectral inverse problem from encoded random projections.
2. To design two regularization terms to measure the low-rank characteristic of spectral images.
3. To formulate an optimization problem which includes the low-rank regularization terms to solve the spectral inverse problem from encoded random projections.
4. To design a numerical algorithm to solve the proposed spectral inverse optimization problem.
5. To verify the developed algorithm by using two state-of-the-art random encoder projectors.

**Contributions.** The thesis' contribution consists of an extensive research for incorporating the low-rank prior to solve ill-posed SI recovery problems, as detailed below.

1. The research discusses theoretical premises and builds empirical studies to analyze the SI low-rank property. Section 2.2 analyzes the low-rank property across the spectral domain, under which a few materials uniquely represented by a spectral signature are repeated at various SI spatial locations, leading to a global spectral low-rank prior. Section 3.2 analyzes the low-rank property across the spatial domain, under which similar structures are repeated

- at different locations, leading to a non-local low-rank prior. Section 4.2 analyzes the low-rank property across a proposed similarity domain, under which the sorting of similar blocks can be shared among the spectral bands, leading to a novel *rank-one similarity* prior.
2. The research formulates various regularizations to take advantage of the intrinsic SI spatial-spectral correlations. Section 2.3 incorporates the global spectral low-rank prior through a hand-crafted regularization known as the nuclear norm. Section 3.3 incorporates the global spectral low-rank prior through the modeling of a SI as a linear combination of two matrices, and incorporates the non-local low-rank prior through an explicit hand-crafted regularization term that promotes self-similarities between small extracted patches. Differently, Section 4.3 combines the global spectral low-rank, non-local self-similarities and the rank-one similarity prior through an implicit regularization term. Section 5.2 proposes an alternative strategy to incorporate the global spectral low-rank prior intrinsically in a neural network architecture, without requiring an explicit function.
  3. The research develops practical algorithms to solve various ill-posed SI recovery problems following convex optimization formulations to guarantee finding a global optimal. Specifically, Section 2.4 presents an alternating minimization with the block-coordinate descent method (BCDM) for CSI reconstruction. Section 3.4 presents a compressive HSI-MSI fusion algorithm iteratively solved with the alternating direction method of multipliers (ADMM). Section 4.4 presents a HSI-MSI fusion algorithm following the plug-and-play with ADMM (PnP-ADMM) approach. Section 5.3 presents a non-data-driven DIP-based approach for sol-

ving single HSI-SR, SI denoising, and CSI reconstruction.

4. The developed algorithms are evaluated over several publicly available daily objects, multispectral, and hyperspectral datasets. Furthermore, two hyperspectral datasets were acquired in the Optics Laboratory from Universidad Industrial de Santander as described in Section 4.5.
5. The developed methods for CSI reconstruction were evaluated by using three implementable state-of-the-art random encoder projectors: CASSI, SSCSI, and C-CASSI.

## **Publications**

The contributions of this dissertation have been published in following journals and conferences.

### **Journal papers.**

1. Gelvez, T., Rueda, H., Arguello, H. (2017). Joint sparse and low rank recovery algorithm for compressive hyperspectral imaging. *Applied optics*, 56(24), 6785-6795. (Published).
2. Gelvez-Barrera, T., Arguello, H. (2020). Nonlocal low-rank abundance prior for compressive spectral image fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 59(1), 415-425. (Published).
3. Gelvez-Barrera, T., Arguello, H., Foi A. Joint Nonlocal, Spectral, and Similarity Low-Rank Priors for Hyperspectral-Multispectral Image Fusion. *IEEE Transactions on Geoscience and Remote Sensing*. (Submitted).
4. Gelvez-Barrera, T., Bacca, J., Arguello, H. Mixture-Net: Low-Rank Deep Image Prior Inspired by Mixture Models for Spectral Image Recovery. *IEEE Transactions on Image Proces-*

sing. (Submitted).

**Conference papers.**

1. Gelvez, T., Bacca, J., Arguello, H. (2021, September). Interpretable Deep Image Prior Method Inspired In Linear Mixture Model For Compressed Spectral Image Recovery. In 2021 IEEE International Conference on Image Processing (ICIP) (pp. 1934-1938). (Published).
2. Ortiz, I., Rivera, S., Gelvez, T., Rojas, F., Arguello, H. (2021, September). (Published). Hyperspectral-Multispectral Image Fusion with Rank Estimation by using a Joint-sparse Regularizer. In 2021 XXIII IEEE Symposium on Image, Signal Processing and Artificial Vision (STSIVA) (pp. 1-6). (Published).
3. Gelvez, T., Bacca, J., Arguello, H. (2021, August). Interpretable Multiple Loss Functions in A Low-Rank Deep Image Prior Based Method For Single Hyperspectral Image Super-Resolution. In 2021 29th European Signal Processing Conference (EUSIPCO) (pp. 696-700). (Published).
4. Gelvez, T., and Arguello, H. (2018, September). Spectral Image Fusion from Compressive Projections Using Total-Variation and Low-Rank Regularizations. In 2018 26th European Signal Processing Conference (EUSIPCO) (pp. 1985-1989). (Published).

### 1. Spectral imaging theoretical background

Spectral imaging sensors capture spectral information ( $\lambda$ ) at various bidimensional (2D)  $(x, y)$  spatial locations from a scene, forming a discrete 3D SI representing a SI (Rueda et al., 2016). Capturing rich spectral information supports diverse remote sensing applications such as disaster management, material identification, and precision agriculture (Vargas et al., 2019). The SI can be denoted as a third-order tensor  $\mathcal{F} \in \mathbb{R}^{N \times N \times L}$  with  $N \times N$  spatial pixels and  $L$  spectral bands, whose vector notation is given by  $\mathbf{f} \in \mathbb{R}^n, n = N^2L$ .

Figure 1 (a) depicts a 3D SI, and four common acquisition techniques: (b) The whisker-broom consists of a point-scanning capturing the spectral distribution of a spatial pixel  $\mathbf{f}_i \in \mathbb{R}^L$ , for  $i = 1, \dots, N^2$  at a time; (c) The push-broom consists of a line-scanning capturing 2D slices with spatial-spectral data at a time; (d) The tunable spectral filter consists of an area-scanning capturing the spatial distribution of a spectral band  $\mathbf{f}^\ell \in \mathbb{R}^{N^2}$ , for  $\ell = 1, \dots, L$  at a time; and (e) The snapshot technique captures the whole spatial-spectral data at one single snapshot.

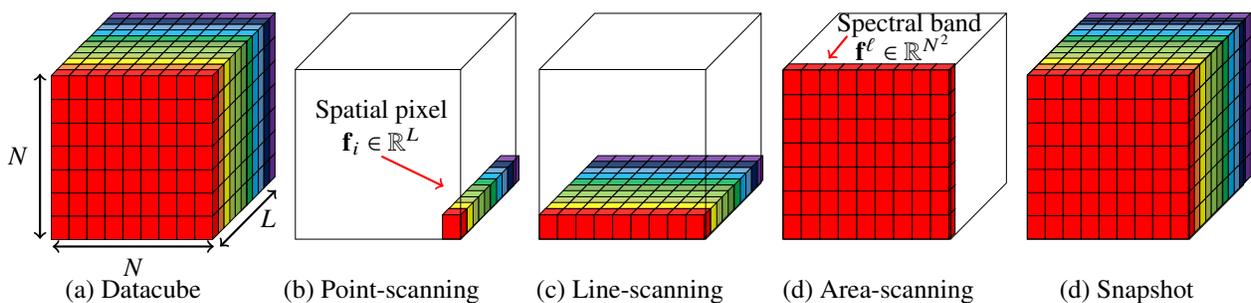


Figure 1. Spectral imaging acquisition techniques.

### 1.1. Hyperspectral and multispectral imaging

Current imaging sensors face a trade-off between the spatial/spectral resolution and the signal-to-noise ratio by cause of limited amount of incident energy (Guilloteau et al., 2020). Hyperspectral sensors such as Hyperion, reflective optics system imaging spectrometer (ROSIS), and airborne visible/infrared imaging spectrometer (AVIRIS) can record an HSI with hundreds of low-resolution narrow spectral bands (Lu et al., 2019). Meanwhile, multispectral sensors such as Landsat, Sentinel-2, and RedEdge, can integrate tens to hundreds of nanometers into one high-resolution spectral band to obtain an MSI.

Let  $\mathbf{M} \in \mathbb{R}^{N_h^2 \times N^2}$  denote a uniform spatial sub-sampling,  $\mathbf{B} \in \mathbb{R}^{N^2 \times N^2}$  denote a spatial blurring convolution, and  $\boldsymbol{\omega}_h \in \mathbb{R}^{N_h^2 L}$  denote additive Gaussian noise. The HSI, denoted by  $\mathbf{f}_h \in \mathbb{R}^{N_h^2 L}$ , can be modeled as a spatial blurred and sub-sampled version of an underlying HRI  $\mathbf{f} \in \mathbb{R}^{N^2 L}$  as

$$\mathbf{f}_h = (\mathbf{I}_L \otimes \mathbf{M})(\mathbf{I}_L \otimes \mathbf{B})\mathbf{f} + \boldsymbol{\omega}_h = \bar{\mathbf{M}}\bar{\mathbf{B}}\mathbf{f} + \boldsymbol{\omega}_h, \quad (1)$$

where the Kronecker product  $\otimes$  applies the sub-sampling and blurring across the spectral bands maintaining a vector notation, so that  $\bar{\mathbf{M}} \in \mathbb{R}^{N_h^2 L \times N^2 L}$  and  $\bar{\mathbf{B}} \in \mathbb{R}^{N^2 L \times N^2 L}$ . The HSI contains  $N_h^2 = N^2/d_h^2$  spatial pixels, with  $d_h > 1$  being a spatial sub-sampling factor.

Similarly, let  $\mathbf{R} \in \mathbb{R}^{L_m \times L}$  denote a multispectral sensor response, and  $\boldsymbol{\omega}_m \in \mathbb{R}^{N^2 L_m}$  denote additive Gaussian noise. The MSI, denoted by  $\mathbf{f}_m \in \mathbb{R}^{N^2 L_m}$ , can be modeled as a spectrally degraded version of the HRI  $\mathbf{f}$  as

$$\mathbf{f}_m = (\mathbf{R} \otimes \mathbf{I}_{N^2})\mathbf{f} + \boldsymbol{\omega}_m = \bar{\mathbf{R}}\mathbf{f} + \boldsymbol{\omega}_m, \quad (2)$$

where  $\otimes$  applies the spectral degradation along all spatial pixels, so that  $\bar{\mathbf{R}} \in \mathbb{R}^{N^2 L_m \times N^2 L}$ . The MSI contains  $L_m = L/d_m$  spectral bands, where  $d_m > 1$  is a spectral degradation factor.

## 1.2. Compressive spectral imaging

The CSI paradigm reduces the time-consuming effort of conventional scanning techniques by incorporating the compressive sensing theory to spectral imaging, where a high-dimensional SI can be acquired and compressed simultaneously into few random projections if satisfying sparsity and incoherence principles (Candès and Wakin, 2008; Brady, 2009). Some remarkable compressive optical systems are described below.

***Coded aperture snapshot spectral imaging.*** The coded aperture snapshot spectral imaging system (CASSI) is composed of three main optical elements, a binary coded aperture that spatially modulates the incoming light, a dispersive element that smears the encoded light, and a detector that multiplexes the encoded smeared light in a focal plane array (FPA) (Wagadarikar et al., 2008). The number of measurements acquired with one CASSI snapshot is equal to  $m = N(N+L-1)$ . For rich spatial-spectral acquisitions, additional snapshots can be acquired by varying the coded aperture spatial distribution, so that  $m = QN(N+L-1)$  for  $Q$  snapshots (Rueda et al., 2016).

***Colored coded aperture snapshot spectral imaging.*** The colored-CASSI system (C-CASSI) replaces the binary coded apertures in CASSI by an array of optical filters known as colored coded apertures. This replacement provides a higher degree of randomness in the spectral domain, which in turn, reduces the amount of required compressed measurements (Arguello and Arce, 2014). The number of measurements acquired with C-CASSI is given by  $QN(N+L-1)$

for  $Q$  snapshots (Rueda et al., 2016).

***Spatial-spectral encoded compressive hyperspectral imager.*** The spatial-spectral encoded compressive hyperspectral imager (SSCSI) uses a single static coded aperture in front of the sensor. The spatial-spectral modulation varies according to the distance between the coded aperture and the sensor. If the coded aperture is mounted directly on the sensor, the shear vanishes and it only modulates the spatial dimensions, and if the coded aperture is placed in the spectral plane, the coding pattern for the spatial dimension will be the same (Lin et al., 2014).

Figure 2 shows an schematic representation of CASSI, C-CASSI, and SSCSI. ***Left:*** CASSI: the scene  $f_o(x, y, \lambda)$  is encoded by a coded aperture  $T(x, y)$  which block/unblock the energy in the  $(x, y)$  spatial location across  $\lambda$  wavelengths. The spectrum components of the coded field are separated by a dispersive element  $V(\lambda)$ . The dispersed field is then integrated over the FPA. ***Middle:*** C-CASSI: the binary coded aperture is replaced by a colored coded aperture  $T_\lambda(x, y)$  which filters the incoming light at the spatial and spectral domains. ***Right:*** SSCSI: the scene  $f_o(x, y, \lambda)$  passes through a diffraction grating to disperse its light into the spectral plane. The resulting dispersed field is encoded by an attenuation coded aperture mounted at a slight offset in front of the sensor.

### 1.3. Spectral imaging recovery problems

The general discrete SI forward model can be mathematically described as

$$\mathbf{y} = \Phi \mathbf{f} + \boldsymbol{\omega}, \quad (3)$$

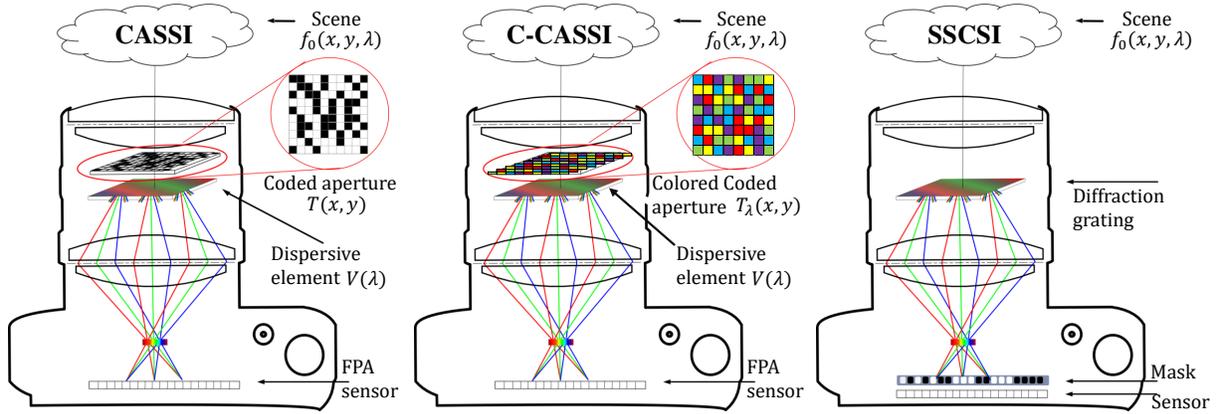


Figure 2. Compressive optical systems.

where  $\mathbf{y} \in \mathbb{R}^m$  stands for observed measurements,  $\Phi \in \mathbb{R}^{m \times n}$  stands for the sensing matrix, and  $\boldsymbol{\omega} \in \mathbb{R}^m$  stands for the acquisition process noise. According to the nature of the sensing matrix, the SI recovery from the observed measurements is referred to as a different recovery problem.

**1.3.1. Spectral image denoising.** Inverse problem when  $m = n$ ,  $\Phi \in \mathbb{R}^{n \times n}$  denotes an identity matrix, and  $\mathbf{y}$  models noisy measurements. This problem appears by cause of the optical system external factors and environment conditions during the acquisition process, such as the illumination, or atmospheric conditions (Chen et al., 2017).

**1.3.2. Single hyperspectral image super-resolution.** Inverse problem when  $m = n/d_h^2$ ,  $\Phi = \bar{\mathbf{M}}\bar{\mathbf{B}}$ , and  $\mathbf{y} = \mathbf{f}_h$  models the spatially degraded HSI in (1). The single HSI-SR problem appears by cause of the technology limitations, so that the spatial resolution is affected when acquiring a SI with several spectral bands (Guilloteau et al., 2020).

**1.3.3. Multispectral-hyperspectral image fusion.** Inverse problem when  $m = n/d_h^2 + n/d_m$ ,  $\Phi = [(\bar{\mathbf{M}}\bar{\mathbf{B}})^T \bar{\mathbf{R}}^T]^T$ , and  $\mathbf{y} = [\mathbf{f}_h^T \mathbf{f}_m^T]^T$ . The HSI-MSI fusion problem aims to synthesize an HRI by combining the useful information of the HSI in (1) with low-spatial-and-high-spectral

resolution; and the MSI in (2) with high-spatial-and-low-spectral resolution (Dian et al., 2021).

**1.3.4. Compressive spectral imaging reconstruction.** Inverse problem when the number of measurements is less than the SI dimension, i.e.,  $m \ll n$ ,  $\Phi = \mathbf{H} \in \mathbb{R}^{m \times n}$  denotes a sensing matrix modeling a compressive optical system, and  $\mathbf{y} \in \mathbb{R}^m$  models compressed measurements, where the ratio  $\gamma = m/n$  is known as the compression ratio.

**1.3.5. Compressive spectral imaging fusion.** Inverse problem when  $m = m_h + m_m$ ,  $\Phi = [(\mathbf{H}_h \bar{\mathbf{M}} \bar{\mathbf{B}})^T \quad \mathbf{H}_m \bar{\mathbf{R}}^T]^T$ , and  $\mathbf{y} = [\mathbf{y}_h^T \quad \mathbf{y}_m^T]^T$ . The compressive HSI-MSI fusion problem aims to synthesize an HRI by combining the useful information of an HSI and an MSI from the corresponding compressed observations.

The use of prior information such as sparsity, smoothness or low-rankness is a crucial idea to effectively recover the SI from the noisy, MSI, HSI or compressed measurements. The prior information is typically introduced through hand-crafted regularization functions in an inverse problem formulation given by

$$\hat{\mathbf{f}} \in \underset{\mathbf{f} \in \mathbb{R}^n}{\operatorname{argmin}} F(\mathbf{f}|\mathbf{y}) + \iota R(\mathbf{f}), \quad (4)$$

where  $\hat{\mathbf{f}} \in \mathbb{R}^n$  denotes the estimated SI,  $F(\cdot) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  denotes the fidelity term to the observations  $\mathbf{y}$ ,  $R(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$  denotes an explicit regularization function that promotes different SI prior information, with  $\iota > 0$  being a regularization parameter to balance the terms.

## 1.4. Low-rank regularization

The low-rank prior assumes that a high-dimensional SI lies in a low-dimensional subspace because of the presence of high spatial-spectral correlations (Mei et al., 2018; Cao et al., 2019a). Low-rank modeling aims to recover the underlying SI from degraded observations by finding a low-rank approximation. Notably, the low-rank prior can be differently regularized as if considering the spatial or spectral dimensions (He et al., 2020; Yokoya et al., 2017).

**1.4.1. Global spectral low-rank prior.** The global spectral low-rank is based on the physical observation that natural scenes commonly contain a few number ( $r \ll L$ ) of different materials uniquely represented by its spectral signature, i.e., each spatial pixel can be modeled as a linear combination of few spectral distributions repeated at several spatial locations (Cohen and Gillis, 2017; Mei et al., 2018).

The global spectral low-rank prior can be regularized unfolding the SI as a 2D matrix  $\mathbf{F} \in \mathbb{R}^{N^2 \times L} = [\mathbf{f}^1 \ \dots \ \mathbf{f}^\ell \ \dots \ \mathbf{f}^L]$ , where each column contain a spectral band. In this formulation, the SI can be well approximated with a low-rank matrix where only  $r \ll L$  different spectral distributions span the scene i.e.,  $\text{rank}(\mathbf{F}) = r$ , where  $\text{rank}(\cdot)$  is referred to as the rank of a matrix determining the number of non-zero singular values.

In addition, the global spectral low-rank can be implicitly regularized in the observation model through a SI mixture model. The linear mixture model (LMM) assumes that the  $i^{\text{th}}$  spatial pixel  $\mathbf{f}_i$  can be modeled as a linear combination of the form  $\mathbf{f}_i = \mathbf{E}\mathbf{a}_i$ , where  $\mathbf{E} \in \mathbb{R}^{L \times r} = [\mathbf{e}_1 \ \dots \ \mathbf{e}_\zeta \ \dots \ \mathbf{e}_r]$  denotes an endmember matrix, whose columns  $\mathbf{e}_\zeta \in \mathbb{R}^L$  contain a unique spectral distribution,

for  $\zeta = 1, \dots, r$ , and  $\mathbf{a}_i \in \mathbb{R}^r$  denotes an abundance vector, whose elements contain the fractional proportions of each endmember at the  $i^{\text{th}}$  spatial pixel, for  $i = 1, \dots, N^2$ . Consequently,  $\mathbf{f}$  can be represented with the following decomposition

$$\mathbf{f} = (\mathbf{E} \otimes \mathbf{I}_{N^2}) \bar{\mathbf{a}} = \bar{\mathbf{E}} \bar{\mathbf{a}}, \quad (5)$$

where  $\bar{\mathbf{a}} \in \mathbb{R}_+^{N^2 r} = [\mathbf{a}_1^T \dots \mathbf{a}_i^T \dots \mathbf{a}_{N^2}^T]^T$  stacks the abundances;  $\bar{\mathbf{E}} \in \mathbb{R}_+^{N^2 L \times N^2 r}$  encompasses the endmembers spanning the SI;  $\mathbf{I}_{N^2}$  refers to an identity matrix of size  $N^2$ ; and  $\otimes$  denotes the Kronecker product, introduced to apply the endmembers along the spatial locations.

The LMM usually accurately characterizes a SI where the endmembers do not interact with each other. Nonetheless, the LMM does not reliably describe a SI where the endmembers interact in complex spatial-spectral scenarios (Bioucas-Dias et al., 2012). Therefore, the non-linear mixture model (NLMM) aims to take into account the non-linear interactions and scattering factors through the use of physical models based on the nature of the environment, or flexible models that approximate the non-linearity with a convolutional neural network (CNN), Kernels, or post-non-linear transformations (Altmann et al., 2014; Heylen et al., 2014). The NLMM is given by

$$\mathbf{f} = \mathcal{N}(\mathbf{E}, \bar{\mathbf{a}}), \quad (6)$$

where  $\mathcal{N}(\cdot)$  stands for an implicit function that defines non-linear interactions between the endmembers and the abundances (Wang et al., 2019b). The aforementioned mixture models impose

additional physical constraints over components  $\mathbf{E}$  and  $\bar{\mathbf{a}}$ . Precisely, the non-negative constraint that considers the nature of reflected light in spectral signatures, such that the entries of the endmember matrix  $\mathbf{E}$  have to be non-negative, and the sum-to-one constraint that considers the entire composition of the pixels, such that the sum of non-negative fractional proportions of each endmember at each pixel has to be one, i.e.,  $\mathbf{a}_i[a] \geq 0$ , and  $\sum_{a=1}^r \mathbf{a}_i[a] = 1, \forall i$ . (Jia and Qian, 2009).

**1.4.2. Local and non-local spatial low-rank prior.** Typically, the spatial distribution of an HRI is highly structured, such that, there exist high self-similarities between small collections of spatial pixels, dubbed patches. Thus, an exemplar patch, or proportional versions of an exemplar patch, may appear several times at different spatial locations of the HRI. This feature generates high redundancy, where each matrix grouping similar versions of an exemplar patch can be assumed to be well approximated with a low-rank matrix. This prior can be referred to as local or non-local low-rank approximation, according to the grouping of patches located at local or non-local positions (Zhang et al., 2018b; Fu et al., 2016).

The non-local low-rank prior has been applied over spectral imaging inverse problems (Xu et al., 2019; Chang et al., 2020a) following four main steps: (i) dis-aggregation of the whole HRI into small spatial overlapped patches of fixed size; (ii) grouping of the patches according to their similarity degree (Kervrann and Boulanger, 2006); (iii) low-rank approximation of each patch-group stacked in a matrix whose rows and columns are assumed to be all proportional to each other (Zhang et al., 2014); and (iv) aggregation of the patches extracted from the low-rank approximation of the patch-groups to form the HRI.

## 2. Global low-rank regularization for compressive spectral imaging reconstruction

*Part of this chapter has been adapted from the journal paper (Gélvez et al., 2017) (published) and conference (Gélvez et al., 2016) (published).*

This chapter presents the first approach developed in this thesis for the inclusion of low-rank regularizations in spectral imaging recovery, where I propose a CSI reconstruction method to estimate a SI from compressed measurements, that combines the sparsity and the global spectral low-rank prior, thus outperforming previous literature approaches. Section 2.1 presents the most related works for CSI reconstruction employing sparsity and low-rank priors, and the identified gaps that motivated the proposed approach, with the following main contributions.

1. Validation of the global spectral low-rank prior through an extensive empirical study of the Pearson correlation matrix along several datasets detailed in Section 2.2.
2. Formulation of a convex optimization inverse problem that unlike previous CSI reconstruction methods employing sparse representations, seeks for a solution that simultaneously satisfies the sparse and global spectral low-rank priors, as shown in Section 2.3.
3. Development of a practical algorithm that addresses the solution of the formulated convex optimization inverse problem following a BCDM, described in Section 2.4.
4. A quantitative gain in the quality of the CSI reconstruction along various datasets and compressive optical systems. The simulations and experiments evaluating the performance of the proposed method including the global spectral low-rank prior are reported in Section 2.5.

### 2.1. Compressive spectral imaging reconstruction related works

Compressive optical systems encode and disperse a SI to sense its spatial-spectral data with few 2D compressed measurements. Traditional CSI techniques reconstruct the full-dimension SI taking advantage of the sparsity prior as a SI is highly compressible when represented in a proper orthonormal basis. The sparsity prior assumes that an  $n$ -dimensional SI has a  $s$ -sparse representation,  $s \ll n$ , in some proper orthonormal basis  $\Psi \in \mathbb{R}^{n \times n}$ , such that,  $\mathbf{f} = \Psi\boldsymbol{\theta}$ , where  $\boldsymbol{\theta} \in \mathbb{R}^n$  is a vector that contains  $s \ll n$  non-zero coefficients, i.e.  $\|\boldsymbol{\theta}\|_0 = s \ll n$ , where  $\|\cdot\|_0$  is referred to as the  $\ell_0$ -pseudo-norm determining the number of non-zero values in an array.

Most of CSI reconstruction algorithms are based on an  $\ell_2 - \ell_1$  unconstrained convex optimization problem, whose cost function includes the  $\ell_2$ -norm as the data-fidelity term and the  $\ell_1$ -norm as the regularization to estimate the sparse approximation (Bioucas-Dias and Figueiredo, 2007; Afonso et al., 2010). Unlike, the remarkable gradient projection for sparse reconstruction (GPSR) algorithm reformulates the  $\ell_2 - \ell_1$  inverse problem as a bound-constrained quadratic programming optimization (Figueiredo et al., 2007b). On the other hand, (Oymak et al., 2015) showed that the performance of multi-objective optimization, when a SI satisfies simultaneously low-rank and sparse structures, is not better than the performance of individually using a single structure. Nonetheless, the main assumption in such study is that the SI is given by the sum of a low-rank matrix with a sparse matrix i.e., the sparsity and low-rankness conditions are promoted in same domain. In contrast, the simultaneous sparsity and low-rank priors can be effectively used by representing the SI in different domains. For instance, (Golbabaee and Vandergheynst, 2012) com-

bined the joint-sparse and low-rank priors to recover the Wavelet coefficient matrix of a SI with a proximal gradient algorithm; (Gogna et al., 2014) presented an efficient Split Bregman algorithm, combining the joint-sparse/sparse and low-rank priors; and (Jia et al., 2015) took further advantage incorporating the smoothness prior, low-rank prior, and a structural self-similarity prior, improving the CSI reconstruction quality through the joint of various priors.

The development of CSI recovery methods combining various priors is then an active research field. Previous mentioned approaches assume the compressive sensing matrix as a convolutional operator representing the theoretical single-pixel camera, exposing a gap in the CSI reconstruction when using implementable and practical optical systems such as the CASSI and SSCSI remarkable systems described in Section 1.2. Therefore, this chapter studies the incorporation of the global spectral low-rank prior for CSI reconstruction, including an empirical analysis, a convex problem formulation, and a numerical solution for the studied implementable architectures.

## 2.2. Global spectral low-rank prior analysis

Recalling the global spectral low-rank prior in Section 1.4.1 based on the high spectral redundancies, this subsection presents the analysis of the global spectral low-rank prior, supported by an extensive empirical study of the Pearson correlation matrix for various publicly available datasets.

The global spectral low-rank prior can be analyzed by calculating the internal correlations within the spatial pixels of a SI. Mathematically, the Pearson correlation coefficient between the spectral distributions  $\mathbf{f}_u$  and  $\mathbf{f}_v$  corresponding to the  $u^{th}$  and  $v^{th}$  spatial pixels can be calculated as

$$\mathbf{C}(u, v) = \frac{\text{cov}(\mathbf{f}_u, \mathbf{f}_v)}{\varrho_{\mathbf{f}_u} \varrho_{\mathbf{f}_v}}, \quad (7)$$

where  $\mathbf{C} \in \mathbb{R}^{N^2 \times N^2}$  denotes the Pearson correlation matrix,  $\text{cov}(\cdot)$  denotes the covariance, and  $\sigma$  denotes the standard deviation of the spectral distributions. The Pearson correlation coefficients take values in the range  $[-1, 1]$ , where positive values indicate that the spectral distributions are directly correlated, negative values indicate that the spectral distributions are counter correlated, and zero indicates that there is no correlation. The higher the absolute value of the coefficient, the stronger the direct or counter correlation. As an example, Fig. 3 left illustrates the Pearson correlation matrix for the Samson dataset (see Appendix 1), measuring the linear dependence between each pair of the  $N^2$  spatial pixels. The white square shows that the 56.91% of pairs present a correlation coefficient higher than 0.9. The color histogram shows the distribution of all entries of the Pearson correlation matrix. Furthermore, Fig. 3 right shows the zoomed entries of two pairs of spatial pixels. The red one presents a direct correlation (0.9816) indicating similar spectral distribution of the material at that location, and the blue one presents a counter correlation ( $-0.6485$ ) indicating different spectral distribution of the materials as shown in the plots.

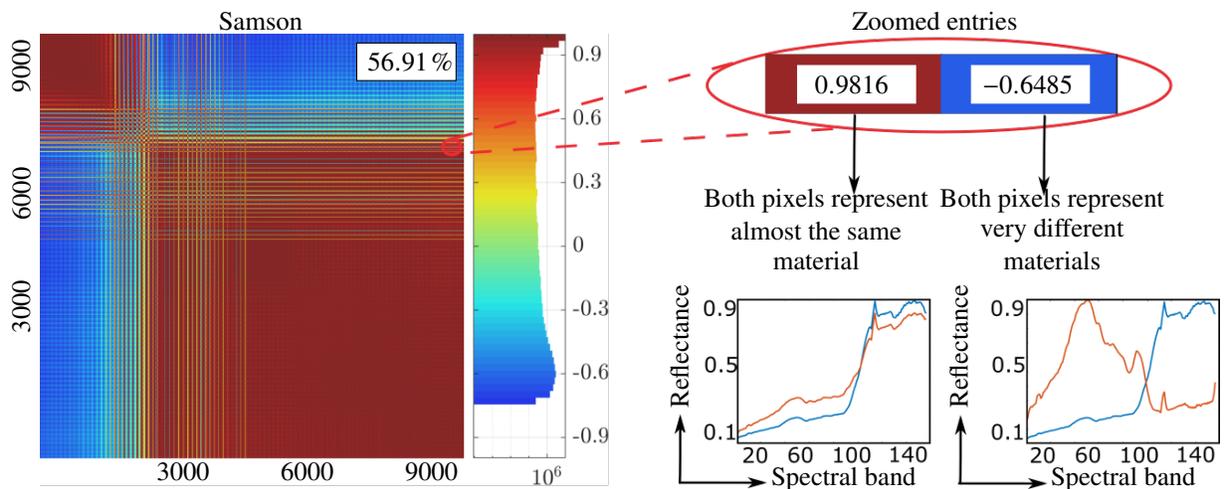


Figure 3. Example of the Pearson correlation matrix analysis.

The global spectral low-rank prior is then validated through the calculation of the Pearson correlation matrix for nine datasets (see Appendix 1) shown in Fig. 4. The percentage of spatial pixel pairs is shown for each correlation matrix as well as the color histogram. Notice that the spectral correlations above 0.9 are higher for hyperspectral than for multispectral scenes, supporting the idea of taking advantage of the global spectral low-rank prior for solving hyperspectral imaging recovery problems, such as the CSI reconstruction problem addressed in this chapter.

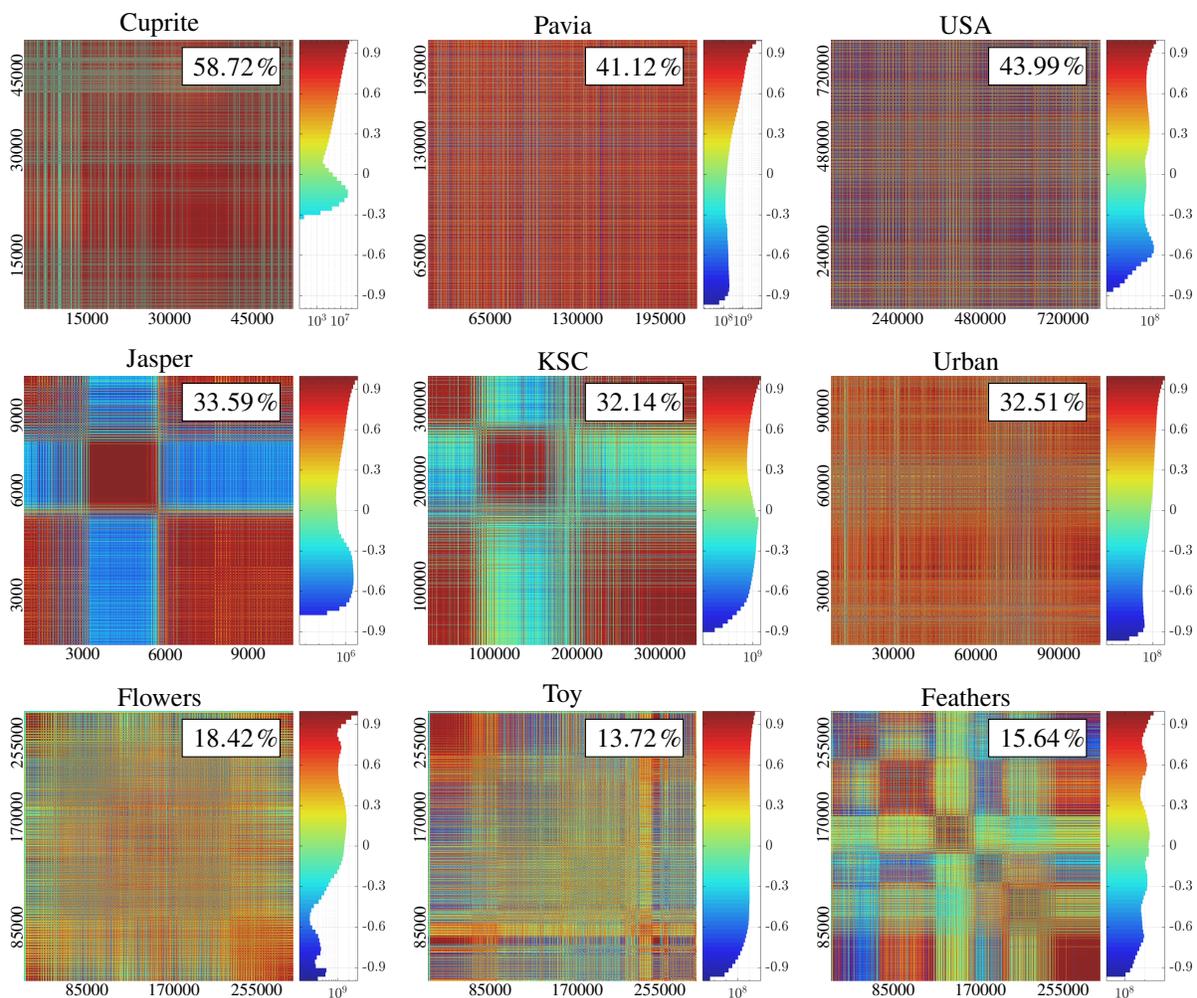


Figure 4. Pearson correlation matrix analysis for nine datasets.

### 2.3. Optimization problem formulation for compressive spectral imaging reconstruction

Based on the observation model in (3), I formulate a CSI reconstruction optimization problem that combines the sparsity and the global spectral low-rank prior. The estimated SI  $\hat{\mathbf{f}}$  from the compressed measurements is enforced to satisfy the sparsity prior in the orthonormal representation basis  $\Psi$ , and the global spectral low-rank property when represented as a matrix  $\mathbf{F} = \text{vec}^{-1}(\mathbf{f})$ , where  $\text{vec}(\cdot)^{-1} : \mathbb{R}^{N^2L} \mapsto \mathbb{R}^{N^2 \times L}$  unfolds a vector as a matrix, by the following minimization

$$\begin{aligned} & \underset{\mathbf{f} \in \mathbb{R}^{N^2L}}{\text{minimize}} && \|\mathbf{y} - \mathbf{H}\mathbf{f}\|_2^2, \\ & \text{subject to} && \|\Psi^{-1}\mathbf{f}\|_0 \leq \epsilon_s; \quad \text{rank}(\text{vec}^{-1}(\mathbf{f})) \leq \epsilon_r, \end{aligned} \tag{8}$$

where  $\epsilon_s$  represents the maximum number of non-zero coefficients of the sparse representation, and  $\epsilon_r$  represents the maximum rank of the matrix representation. The  $\ell_0$ -pseudo-norm and the  $\text{rank}(\cdot)$  have been shown to be non-convex constraints, so that (8) is a non-deterministic polynomial-time hard (NP-hard) problem and its solution is difficult to approximate (Candes and Recht, 2008). To overcome such limitation, the problem is reformulated by employing the convex relaxation of the sparsity and low-rank constraints through the  $\ell_1$ -norm and nuclear-norm  $\ell_*$ -norm, obtaining equivalent solutions with high probability at a polynomial time (Wright et al., 2008), i.e., the minimization of  $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$  promotes an array with few non-zero coefficients, and the minimization of  $\|\mathbf{X}\|_* = \sum_i \sigma_i$ , where  $\sigma_i$  is the  $i^{\text{th}}$ -singular value of  $\mathbf{X}$ , promotes a low-rank approximation of

**X** (Candes and Recht, 2008). Then, the convex relaxation of (8) is given by

$$\begin{aligned} & \underset{\mathbf{f} \in \mathbb{R}^{N^2L}}{\text{minimize}} \quad \|\mathbf{y} - \mathbf{H}\mathbf{f}\|_2^2, \\ & \text{subject to} \quad \|\Psi^{-1}\mathbf{f}\|_1 \leq \epsilon_{z'}; \quad \|\text{vec}^{-1}(\mathbf{f})\|_* \leq \epsilon_{r'}, \end{aligned} \quad (9)$$

where  $\epsilon_{z'}$  represents the maximum value for the sum of the coefficients in the sparse representation, and  $\epsilon_{r'}$  represents the maximum value for the sum of the eigenvalues of the matrix representation.

The sparsity and low-rank constraints in (9) reduce the feasible set of the problem, increasing the complexity for finding the optimal solution. Therefore, the explicit constraints are incorporated into the objective function penalizing big values for the  $\ell_1$ -norm and  $\ell_*$ -norm but, without restricting the search domain as follows

$$\hat{\mathbf{f}} \in \underset{\mathbf{f} \in \mathbb{R}^{N^2L}}{\text{argmin}} \quad \|\mathbf{y} - \mathbf{H}\mathbf{f}\|_2^2 + \tau \|\Psi^{-1}\mathbf{f}\|_1 + \beta \|\text{vec}^{-1}(\mathbf{f})\|_*, \quad (10)$$

where  $\tau > 0$ , and  $\beta > 0$  denote the regularization parameters to balance the terms.

#### 2.4. Joint sparse and low-rank compressive spectral imaging algorithm

To effectively solve problem in (10), the proposed strategy begins with the inclusion of two auxiliary variables  $\boldsymbol{\theta} \in \mathbb{R}^{N^2L}$  and  $\mathbf{F} \in \mathbb{R}^{N^2 \times L}$  into the optimization problem, modeling the sparse and the low-rank representations, respectively, as follows

$$\begin{aligned} & \underset{\mathbf{f}, \boldsymbol{\theta}, \mathbf{F}}{\text{minimize}} \quad \|\mathbf{y} - \mathbf{H}\mathbf{f}\|_2^2 + \tau \|\boldsymbol{\theta}\|_1 + \beta \|\mathbf{F}\|_* \\ & \text{subject to} \quad \boldsymbol{\theta} = \Psi^{-1}\mathbf{f}; \quad \mathbf{F} = \text{vec}^{-1}(\mathbf{f}). \end{aligned} \quad (11)$$

**2.4.1. Alternating minimization.** The inclusion of the auxiliary variables allows to decouple the problem, so that, the sparse and low-rank regularizations and the data fidelity term can be addressed independently by alternating the minimization of the Lagrangian function given by

$$\{\hat{\mathbf{f}}, \hat{\mathbf{F}}, \hat{\boldsymbol{\theta}}\} \in \underset{\mathbf{f}, \mathbf{F}, \boldsymbol{\theta}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{f}\|_2^2 + \tau\|\boldsymbol{\theta}\|_1 + \beta\|\mathbf{F}\|_* + \eta\|\boldsymbol{\theta} - \Psi^{-1}\mathbf{f}\|_2^2 + \mu\|\mathbf{F} - \operatorname{vec}^{-1}(\mathbf{f})\|_2^2, \quad (12)$$

where  $\eta > 0$  and  $\mu > 0$  are adjustable Lagrangian parameters.

*Solving for the sparse representation.* Given a previous estimation of the SI  $\mathbf{f}^k$ , the sparse representation is updated by solving

$$\hat{\boldsymbol{\theta}}^{k+1} \in \underset{\boldsymbol{\theta} \in \mathbb{R}^{N^2L}}{\operatorname{argmin}} \eta\|\boldsymbol{\theta} - \Psi^{-1}\mathbf{f}^k\|_2^2 + \tau\|\boldsymbol{\theta}\|_1. \quad (13)$$

Problem in (13) results in the well-known  $\ell_2 - \ell_1$  inverse problem. Therefore, the sparse representation can be estimated by the PnP of any conventional  $\ell_2 - \ell_1$  solver such as those described in (Figueiredo et al., 2007a; Wright et al., 2009; Afonso et al., 2011).

*Solving for the low-rank representation.* Given an updated version of the SI defined as  $\mathbf{f}^k = \Psi\hat{\boldsymbol{\theta}}^{k+1}$ , the low-rank matrix representation is updated by solving

$$\hat{\mathbf{F}}^{k+1} \in \underset{\mathbf{F} \in \mathbb{R}^{N^2 \times L}}{\operatorname{argmin}} \mu\|\mathbf{F} - \operatorname{vec}^{-1}(\mathbf{f}^k)\|_2^2 + \beta\|\mathbf{F}\|_*. \quad (14)$$

Problem in (14) results in the well-known  $\ell_2 - \ell_*$  inverse problem that minimizes the  $\ell_2$ -

norm to maintain data-fidelity, and the  $\ell_*$ -norm regularization to enforce a low-rank approximation. Therefore, the low-rank representation can be estimated by implementing the element-wise soft-shrinkage to the singular value decomposition (SVD) of the matrix  $\mathbf{F}$ , described in Algorithm 1 (Zhao and Yang, 2015).

Algorithm 1 line 5 minimizes the squared  $\ell_2$ -norm term in (14). There, the matrix  $\mathbf{L}^{k_1+1}$  is a temporal matrix approximation which subtracts  $\mu$  times the residual between the matrix representation obtained from the initial estimation  $\mathbf{f}^k$ , and the low-rank approximation at the current iteration  $\mathbf{F}^{k_1}$ . In Algorithm 1, line 6 matrices  $\mathbf{U}^{k_1+1} \in \mathbb{R}^{N^2 \times N^2}$ ,  $\mathbf{S}^{k_1+1} \in \mathbb{R}^{N^2 \times L}$  and  $\mathbf{V}^{k_1+1} \in \mathbb{R}^{L \times L}$  represent the left-singular vectors, the singular values and the right-singular vectors of the matrix  $\mathbf{L}^{k_1+1}$ , respectively. Then, the singular values in  $\mathbf{S}^{k_1+1}$  are used to reduce the rank of the matrix which turns in minimizing the  $\ell_*$ -norm term in (14). Remark that this method does not require a

---

#### Algorithm 1. Low-rank approximation

---

```

1: procedure LRA( $\mathbf{f}^k, \mu, \beta, K_1$ )  $\triangleright \mathbf{f}^k$  : previous estimation of the SI,  $\mu > 0, \beta > 0$ : regularization
   parameters, and  $K_1$  : maximum number of iterations.


---


2:    $\mathbf{F}^0 \leftarrow \mathbf{0}$ 
3:    $k_1 \leftarrow 0$ 
4:   while  $k_1 < K_1$  do  $\triangleright$  while stopping criteria is not achieved
5:      $\mathbf{L}^{k_1+1} \leftarrow \mathbf{F}^{k_1} - \mu(\text{vec}^{-1}(\mathbf{f}^k) - \mathbf{F}^{k_1})$ 
6:      $[\mathbf{U}^{k_1+1} \mathbf{S}^{k_1+1} \mathbf{V}^{k_1+1}] \leftarrow \text{svd}(\mathbf{L}^{k_1+1})$   $\triangleright$  Singular value decomposition
7:      $\alpha^{k_1+1} \leftarrow \frac{\sum_{i=2}^{\Sigma} \sigma_i}{\Sigma}$   $\triangleright \sigma_i$  is the  $i^{\text{th}}$  singular value of  $\mathbf{S}^{k_1+1}$ , for  $i = 0, \dots, \Sigma - 1$ .
8:      $\mathbf{s}^{k_1+1} \leftarrow \max(0, \text{diag}(\mathbf{S}^{k_1+1}) - \beta \alpha^{k_1+1})$ 
9:      $\mathbf{F}^{k_1+1} \leftarrow \mathbf{U}^{k_1+1} \text{diag}(\mathbf{s}^{k_1+1}) \mathbf{V}^{k_1+1T}$   $\triangleright$  elementwise soft shrinkage
10:     $k_1 \leftarrow k_1 + 1$ 
11:  return  $\mathbf{F}^{K_1}$   $\triangleright$  low-rank SI matrix representation


---



```

prior knowledge about the rank of the matrix, and I set the parameter  $\alpha^{k_1+1}$  in Algorithm 1 line 7 for the element-wise soft-shrinkage as the average of the singular values of the previous calculated matrix  $\mathbf{L}^{k_1+1}$ . Any other strategy could be used, producing a different rank condition for the updated matrix  $\mathbf{F}^{k_1+1}$ . For instance, if there exists a prior estimation for the rank of  $\mathbf{F}$  given by  $\hat{w}$ , the operations in Algorithm 1 lines 7 and 8 could be replaced by a hard-thresholding as,  $\mathbf{s}^{k_1+1} \leftarrow \max(0, \text{diag}(\mathbf{S}^{k_1+1}) - \mathbf{S}^{k_1+1}(\hat{w} + 1, \hat{w} + 1))$ .

*Solving for the spectral image.* The SI  $\hat{\mathbf{f}}$  is then updated by considering the estimated sparse representation and low-rank approximation, such that it also fits the observed compressed measurements, by solving the  $\ell_2$  problem,

$$\hat{\mathbf{f}}^{k+1} \in \underset{\mathbf{f} \in \mathbb{R}^{N^2L}}{\text{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{f}\|_2^2 + \eta \|\boldsymbol{\theta}^{k+1} - \Psi^{-1}\mathbf{f}\|_2^2 + \mu \|\mathbf{F}^{k+1} - \text{vec}^{-1}(\mathbf{f})\|_2^2, \quad (15)$$

which can be parsed as a quadratic optimization problem that minimizes the function  $\xi(\mathbf{f})$ ,

$$\begin{aligned} \hat{\mathbf{f}}^{k+1} &\in \underset{\mathbf{f} \in \mathbb{R}^{N^2L}}{\text{argmin}} \xi(\mathbf{f}) = \frac{1}{2} \mathbf{f}^T \mathbf{Q} \mathbf{f} + (\mathbf{c}^{k+1})^T \mathbf{f} + b^{k+1}, \\ b^{k+1} &= \mathbf{y}^T \mathbf{y} + \eta (\hat{\boldsymbol{\theta}}^{k+1})^T \hat{\boldsymbol{\theta}}^{k+1} + \mu (\text{vec}(\hat{\mathbf{F}}^{k+1}))^T \text{vec}(\hat{\mathbf{F}}^{k+1}), \\ \mathbf{c}^{k+1} &= \mathbf{H}^T \mathbf{y} + \eta \Psi \hat{\boldsymbol{\theta}}^{k+1} + \mu \text{vec}(\hat{\mathbf{F}}^{k+1}) \\ \mathbf{Q} &= \mathbf{H}^T \mathbf{H} + \eta \Psi^T \Psi + \mu \mathbf{I} = \mathbf{H}^T \mathbf{H} + (\eta + \mu) \mathbf{I}, \end{aligned} \quad (16)$$

Problem in (16) is then solved by using the steepest descent approach described in Algorithm 2. The array  $\nabla^{k_2+1}$  in Algorithm 2 line 5 represents the search direction calculated as the

gradient of the objective function  $\xi(\mathbf{f})$ . The backtracking in Algorithm 2 line 6 is a procedure which chooses the optimum candidate for the step  $\beta^{k_2+1}$  in the gradient descent algorithm. The selection of the proper step leads the algorithm to make reasonable progress along the given search direction  $\nabla^{k_2+1}$  per iteration (Nocedal and Wright, 2006).

Algorithm 2. Steepest descend

---

```

1: procedure STEDES( $\mathbf{f}^k, \mathbf{Q}, \mathbf{c}^{k+1}, \xi, K_2$ )   $\triangleright \mathbf{f}^k$  : previous estimation,  $\mathbf{Q}, \mathbf{c}^{k+1}$ : parameters of the
   quadratic minimization problem,  $\xi$  the objective function in (16), and  $K_2$  : maximum number
   of iterations.
2:    $\mathbf{f}^0 \leftarrow \mathbf{f}^k$ 
3:    $k_2 \leftarrow 0$ 
4:   while  $k_2 < K_2$  do                                 $\triangleright$  while stopping criteria is not achieved
5:      $\nabla^{k_2+1} \leftarrow 2\mathbf{Q}\mathbf{f}^{k_2} + \mathbf{c}^{k+1}$            $\triangleright$  optimum search direction
6:      $\beta^{k_2+1} \leftarrow \text{backtracking}(\mathbf{f}^{k_2}, \nabla^{k_2+1}, \xi)$    $\triangleright$  optimum step
7:      $\mathbf{f}^{k_2+1} \leftarrow \mathbf{f}^{k_2} - \beta^{k_2+1}\nabla^{k_2+1}$ 
8:      $k_2 \leftarrow k_2 + 1$ 
9:   return  $\mathbf{f}^{K_2}$                                         $\triangleright$  joint low-rank and sparse estimation

```

---

The backtracking procedure can be easily implemented following the procedure shown in Algorithm 3, where notation  $\xi(\cdot)$  implies to evaluate the objective function in (16). By last, the estimated SI joining the sparse representation, low-rank approximation and compressed measurements is updated in Algorithm 2 line 7 by subtracting the gradient  $\nabla^{k_2+1}$  from the previous estimation considering the optimum step  $\beta^{k_2+1}$ .

The three previous sparse, low-rank and fidelity-data steps are executed iteratively until a stopping criteria is achieved. The proposed joint sparse and low-rank CSI reconstruction method is summarized in Algorithm 4. There, the initial estimation of the SI in line 2 was chosen to be a low-quality version obtained by using the transpose,  $\mathbf{H}^T$ , as an approximation of the inverse  $\mathbf{H}^{-1}$ .

## Algorithm 3. Backtracking

---

1: **procedure** BACKTRACKING( $\mathbf{f}^{k_2}, \nabla^{k_2+1}, \xi$ )  $\triangleright \mathbf{f}^{k_2}$  : previous estimation of the SI,  $\nabla^{k_2+1}$  : search direction, and  $\xi$  : the objective function in (16).

---

2:  $\beta \leftarrow 1$

3:  $\rho \leftarrow \text{rand}(0, 1)$

4:  $c \leftarrow \text{rand}(0, 1)$

5: **while**  $\xi(\mathbf{f}^{k_2} + \beta \nabla^{k_2+1}) \geq \xi(\mathbf{f}^{k_2}) + c\beta \nabla^{k_2+1 T} \nabla^{k_2+1}$  **do**

6:      $\beta \leftarrow \rho\beta$

7: **return**  $\beta$   $\triangleright$  optimum step

---

## Algorithm 4. Joint sparse and low-rank CSI reconstruction algorithm

---

1: **procedure** JOINTSPARLOW( $\mathbf{y}, \tau, \mu, \beta, \eta, \mathbf{H}, \Psi, K$ )  $\triangleright \mathbf{y}$  :  
compressed measurements,  $\tau > 0$ ,  $\mu > 0$ ,  $\beta > 0$ ,  $\eta > 0$  : regularization parameters,  $\mathbf{H}$  : sensing matrix,  $\Psi$  : representation basis, and  $K$  : maximum number of iterations.

---

2:  $\mathbf{f}^0 \leftarrow \mathbf{H}^T \mathbf{y}$   $\triangleright$  starting low quality estimation of  $\mathbf{f}$

3:  $\mathbf{Q} \leftarrow \mathbf{H}^T \mathbf{H} + (\eta + \mu) \mathbf{I}$

4:  $k \leftarrow 0$

5: **while**  $k < K$  **do**  $\triangleright$  while stopping criteria is not achieved

6:      $\hat{\boldsymbol{\theta}}^{k+1} \leftarrow \underset{\boldsymbol{\theta} \in \mathbb{R}^{N^2 L}}{\text{argmin}} \eta \|\boldsymbol{\theta} - \Psi^{-1} \mathbf{f}^k\|_2^2 + \tau \|\boldsymbol{\theta}\|_1$

7:      $\mathbf{f}^k \leftarrow \Psi \hat{\boldsymbol{\theta}}^{k+1}$

8:      $\hat{\mathbf{F}}^{k+1} \leftarrow \underset{\mathbf{F} \in \mathbb{R}^{N^2 \times L}}{\text{argmin}} \mu \|\mathbf{F} - \text{vec}^{-1}(\mathbf{f}^k)\|_2^2 + \beta \|\mathbf{F}\|_*$

9:      $\mathbf{c}^{k+1} \leftarrow \mathbf{H}^T \mathbf{y} + \eta \Psi \hat{\boldsymbol{\theta}}^{k+1} + \mu \text{vec}(\hat{\mathbf{F}}^{k+1})$

10:      $b^{k+1} \leftarrow \mathbf{y}^T \mathbf{y} + \eta (\hat{\boldsymbol{\theta}}^{k+1})^T \hat{\boldsymbol{\theta}}^{k+1} + \mu (\text{vec}(\hat{\mathbf{F}}^{k+1}))^T \text{vec}(\hat{\mathbf{F}}^{k+1})$

11:      $\hat{\mathbf{f}}^{k+1} \leftarrow \underset{\mathbf{f} \in \mathbb{R}^{N^2 L}}{\text{argmin}} \frac{1}{2} \mathbf{f}^T \mathbf{Q} \mathbf{f} + (\mathbf{c}^{k+1})^T \mathbf{f} + b^{k+1}$

12:      $k \leftarrow k + 1$

13: **return**  $\mathbf{f}^K$   $\triangleright$  estimated datacube

---

**2.4.2. Analysis of the computational complexity.** The complexity of the proposed

iterative joint CSI reconstruction algorithm is as follows: Initialization of  $\mathbf{f}^0$  in Algorithm 4 line 2

is a matrix-vector product of order  $O(2N^2Lm)$ , and calculation of  $\mathbf{Q}$  in Algorithm 4 line 3 is of order  $O(2(N^2L)^2m)$ . Besides, the complexity per iteration in the loop corresponds to the sum of, the sparse estimation of the SI with resolution  $N \times N \times L$  in Algorithm 4 line 6 which has been proved to be of complexity  $O(N^2L)$  using the GPSR solver (Figueiredo et al., 2007a), the update of the SI estimation in Algorithm 4 line 7 which is a matrix-vector product of order  $O(N^4L^2)$ , the low-rank approximation of the matrix with resolution  $N^2 \times L$  in Algorithm 4 line 8 depends on the computation of the SVD which has been proved to be of order  $O(N^2L^2 + L)$  (Zhao and Yang, 2015). Calculation in Algorithm 4 line 9 has complexity  $O(N^4L^2)$ , calculation in Algorithm 4 line 10 has complexity  $O(2N^2L)$ , and the steepest descent in Algorithm 4 line 11 which has been proved to be of complexity  $O(N^2L)$  (Zhao and Yang, 2015). Hence, the overall computational complexity of Algorithm 4 is  $O(2N^4L^2m + 2N^4L^2 + 2N^2Lm + 4N^2L + N^2L^2 + L)$ .

**2.4.3. Convergence analysis.** The convergence of the proposed algorithm is shown by analyzing the convergence of each step per iteration of Algorithm 4. First, step in Algorithm 4 line 6 is solved using the conventional solver GPSR whose convergence has been proved in (Figueiredo et al., 2007a). Also, step in Algorithm 4 line 8 is solved using the element-wise soft shrinkage whose convergence has been proved in (Cai et al., 2010). The steepest descent method to solve the problem in Algorithm 4 line 16 converges if the matrix  $\mathbf{Q}$  is symmetric and positive definite (Nocedal and Wright, 2006). Since  $\mathbf{Q} = \mathbf{H}^T \mathbf{H} + (\eta + \mu) \mathbf{I}$ , is clearly symmetric, to guarantee matrix  $\mathbf{Q}$  being positive definite I rely on the statement that a matrix is positive definite if and only if all of its eigenvalues are positive (Nocedal and Wright, 2006).

**Proposition 1:** Let a matrix  $\mathbf{B} \in \mathbb{R}^{n \times n}$  be the sum of any matrix  $\mathbf{C} \in \mathbb{R}^{n \times n}$  plus a scaled

identity  $\nu \mathbf{I} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{B} = \mathbf{C} + \nu \mathbf{I}$ . The eigenvalues of  $\mathbf{B}$ ,  $\varphi_i, i = 0, \dots, n-1$ , are given by the sum of the eigenvalues of  $\mathbf{C}$ ,  $\hat{\varphi}_i, i = 0, \dots, n-1$ , plus  $\nu$ ,  $\varphi_i = \hat{\varphi}_i + \nu$ .

**Proof:** The natural procedure to find the eigenvalues of  $\mathbf{B}$  is given by,  $\det(\mathbf{B} - \delta \mathbf{I}) = 0$ ,  $\delta \in \mathbb{R}$ .

Replacing  $\mathbf{B}$  by the sum  $\mathbf{C} + \nu \mathbf{I}$ , I obtain,

$$\det(\mathbf{C} + \nu \mathbf{I} - \delta \mathbf{I}) = \det(\mathbf{C} - (\delta - \nu) \mathbf{I}) = \det(\mathbf{C} - \hat{\delta} \mathbf{I}) = 0, \quad (17)$$

where,  $\hat{\delta} = \delta - \nu$ . By solving (17), I obtain the characteristic polynomial of  $\mathbf{C}$  whose roots are its eigenvalues,  $(\hat{\delta} - \hat{\varphi}_0)(\hat{\delta} - \hat{\varphi}_1) \dots (\hat{\delta} - \hat{\varphi}_{n-1})$ . Replacing  $\hat{\delta} = \delta - \nu$  I obtain,

$$(\delta - (\nu + \hat{\varphi}_0))(\delta - (\nu + \hat{\varphi}_1)) \dots (\delta - (\nu + \hat{\varphi}_{n-1})). \quad (18)$$

Observe that (18) corresponds to the characteristic polynomial of the matrix  $\mathbf{B}$ , whose roots are its eigenvalues. Thus,  $\varphi_i = \hat{\varphi}_i + \nu$  and proposition 1 is proved.

A square symmetric matrix  $\mathbf{E} \in \mathbb{R}^{n \times n}$  is positive semi-definite if  $\mathbf{x}^T \mathbf{E} \mathbf{x} \geq 0$  for all  $\mathbf{x} \in \mathbb{R}^n$ . Note that, for all  $\mathbf{z} \in \mathbb{R}^{N^2 L}$ ,  $\mathbf{z}^T (\mathbf{H}^T \mathbf{H}) \mathbf{z} = (\mathbf{H} \mathbf{z})^T (\mathbf{H} \mathbf{z}) = \|\mathbf{H} \mathbf{z}\|_2^2 \geq 0$ , then the matrix  $\mathbf{H}^T \mathbf{H}$  is positive semi-definite and all of its eigenvalues are non-negative. Regarding proposition 1, the eigenvalues of  $\mathbf{Q}$  are given by the sum of the eigenvalues of  $\mathbf{H}^T \mathbf{H}$  plus  $(\eta + \mu)$ . As  $\eta > 0$ , and  $\mu > 0$ , the eigenvalues of  $\mathbf{Q}$  are guaranteed to be always positive. Hence,  $\mathbf{Q}$  is positive definite, and the steepest descend algorithm will converge. As each step of the joint CSI reconstruction algorithm converges and the stopping criteria is a finite number of iterations, the algorithm globally converges.

Observe that the proposed solution follows a convex formulation, such that a local minima is guaranteed to be a global minima (Boyd et al., 2011).

## 2.5. Simulations and results

Simulated experiments were carried out to measure the gain, if any, given by the proposed joint sparse and low-rank CSI reconstruction under diverse experimental setups. The results were compared against the reconstructions obtained by solving the conventional  $\ell_2 - \ell_1$  problem using the GPSR solver. **Remark:** In Algorithm 4 line 6, I solve the  $\ell_2 - \ell_1$ -norm problem at each iteration of the general joint algorithm using an iterative solver, i.e. an iterative algorithm inside another iterative algorithm. In this work, I set 10 iterations to solve the  $\ell_2 - \ell_1$ -norm problem at each iteration of the joint algorithm. In this regard, I make the comparison between the traditional and the proposed algorithm taking into account the total amount of iterations done by the  $\ell_2 - \ell_1$  solver. For instance, if I use 10 iterations in the joint algorithm, I compare the results using 100 iterations with the single traditional solver. Also, a sensitivity analysis of the involved parameters is presented.

In the experiments, the quality was analyzed over the Samson, Jasper, Urban, and Cuprite datasets in Appendix 1 satisfying a low-rank structure, following the setup in (Zhu et al., 2014a). Seven experimental configurations were set by varying the spatial-spectral resolution as follows

**C1**  $\in \mathbb{R}^{16 \times 16 \times 128}$ : From Samson.

**C5**  $\in \mathbb{R}^{16 \times 16 \times 16}$ : From Samson.

**C2**  $\in \mathbb{R}^{32 \times 32 \times 128}$ : From Jasper.

**C6**  $\in \mathbb{R}^{16 \times 16 \times 32}$ : From Samson.

**C3**  $\in \mathbb{R}^{64 \times 64 \times 128}$ : From Urban.

**C7**  $\in \mathbb{R}^{16 \times 16 \times 64}$ : From Samson.

**C4**  $\in \mathbb{R}^{64 \times 64 \times 128}$ : From Cuprite.

The rank for each dataset was assumed to be the known amount of different materials present in the scene, i.e.  $r = 3$ ,  $r = 4$ ,  $r = 6$ ,  $r = 12$ , for Samson, Jasper, Urban, and Cuprite datasets. Furthermore, the acquisition of the compressed measurements was carried out by simulating the CASSI and SSCSI compressive architectures described in Section 1.2.

The  $\ell_2 - \ell_1$  problem in (13) was solved by using the GPSR solver, nonetheless, the choice of any other solver will equally affect both comparison methods, the traditional and the proposed CSI reconstruction algorithm.

Finally, the quality of the reconstructions obtained with the proposed and conventional  $\ell_2 - \ell_1$ -norm approach was measured through the spatial and spectral peak signal to noise ratio (PSNR) and the structural similarity (SSIM) metrics defined in Appendix 2.

**2.5.1. Regularization parameters analysis.** This experiment analyzes the behaviour of parameters  $\tau$ ,  $\mu$ , and  $\beta$  involved in the proposed CSI reconstruction algorithm since they highly affect the reconstruction quality, and the search of the proper parameters can be a tedious task. Parameter  $\tau$  accounts for the performance of the conventional  $\ell_2 - \ell_1$  minimization problem by controlling the speed towards the sparse representation. Parameter  $\mu$  states how far in terms of the  $\ell_2$ -norm the estimated low-rank matrix is allowed to be from the original given matrix, and parameter  $\beta$  controls the speed towards the low-rank approximation.

Figure 5 shows the parameter analysis for **C1** when using CASSI (*top*) and SSCSI (*bottom*). Figure 5 (a) and Fig. 5 (b) show the spatial PSNR varying the number of iterations  $K$ , and  $\tau$  in an equally spaced range  $[1e^{-6} \ 11e^{-4}]$ . Once  $\tau$  is fixed, Fig. 5 (c) and Fig. 5 (d) show the spatial PSNR varying  $\mu$  and  $\beta$  in an equally spaced range of  $[1e^{-2} \ 1]$  and  $[1e^{-5} \ 1e^{-1}]$ . These parameters always

fall within the range  $(0, 1]$ , however, an extended search is required to find the best choice per each configuration. When there is a previous knowledge of the rank, the element-wise soft-shrinkage can be replaced with a hard-thresholding, eliminating  $\beta$ , and only the seek for  $\mu$  is required, varied in an equally spaced range  $[1e^{-2}, 1]$ , as shown in Fig. 5 (e) and Fig. 5(f). For all parameters a smooth approximation towards the best choice can be observed. This, given that the proposed CSI reconstruction algorithm follows a convex formulation. The spatial PSNR was shown as illustration of the smooth behaviour, that holds for the spectral PSNR and the SSIM metrics.

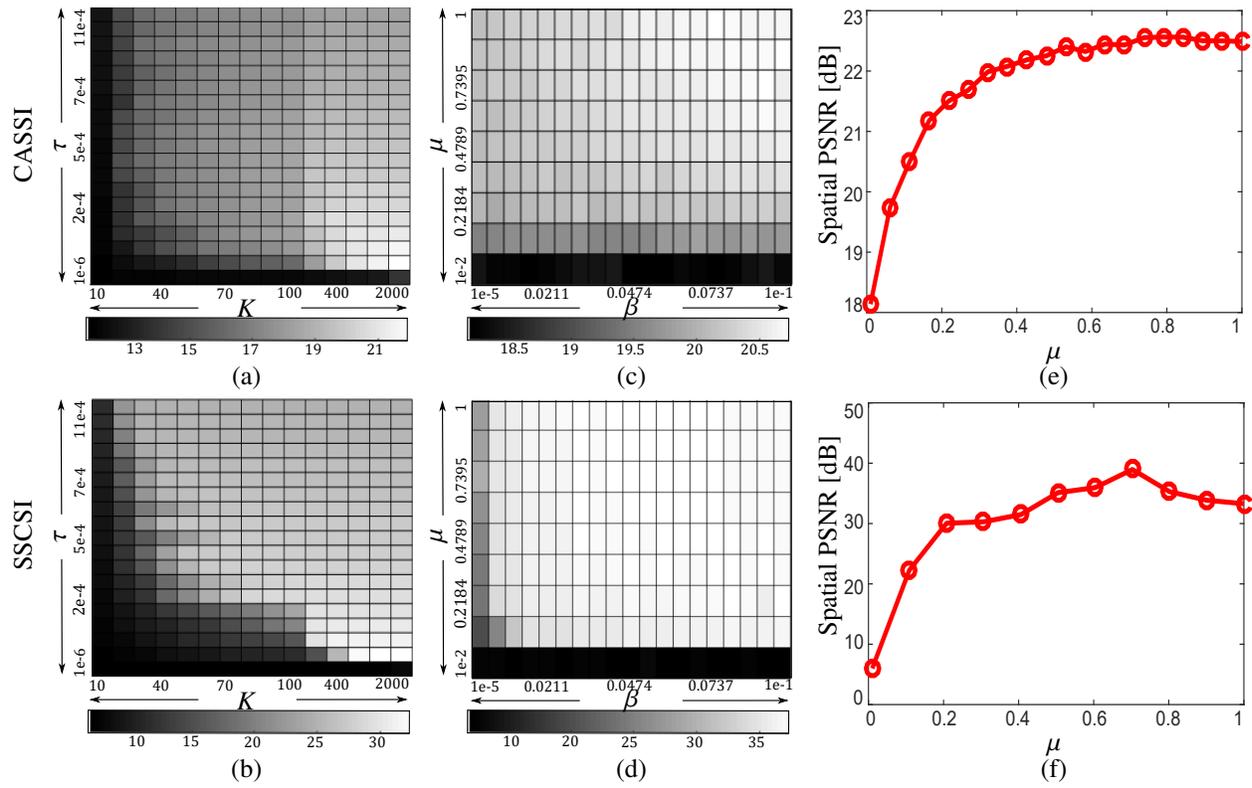


Figure 5. Parameter analysis of the proposed CSI reconstruction algorithm.

**2.5.2. Performance evaluation.** This subsection evaluates the performance of the proposed algorithm through several experiments.

**Quality test along different datasets.** Tables 1 and 2 show the quality comparison between the traditional algorithm labeled as “sparse”, the proposed joint sparse and low-rank CSI reconstruction algorithm without prior information of the rank labeled as “joint”, and the proposed joint sparse and low-rank CSI reconstruction algorithm with prior information of the rank labeled as “joint-r” for the CASSI and the SSCSI architectures. It can be observed that the proposed algorithm improved the reconstruction quality for all tested configurations along the three quantitative metrics. The gain using the CASSI architecture was up to 3[dB], and using the SSCSI architectures was up to 4[dB] which is the expected behaviour due to the SSCSI architecture does not multiplex the spectral information, so that the spectral correlations can be better considered. Notice that a higher gain is obtained when having prior knowledge about the rank of the underlying structure. Nonetheless, even without such prior knowledge about the rank, an improvement is also attained with the proposed CSI reconstruction algorithm.

Table 1  
*Reconstruction quality - CASSI architecture*

Data cube		C1	C2	C3	C4
% of captured data		<b>41.89 %</b>	<b>58.22 %</b>	<b>58.29 %</b>	<b>69.95 %</b>
<b>Spatial PSNR</b> [dB]	sparse	20.9784	24.5606	21.2609	25.7060
	joint	<b>21.0042</b>	<b>26.1277</b>	<b>22.7974</b>	<b>28.5328</b>
	joint-r	<b>22.6785</b>	<b>26.1848</b>	<b>22.7977</b>	<b>28.5339</b>
<b>Spectral PSNR</b> [dB]	sparse	29.1606	21.7445	20.9852	24.5313
	joint	<b>30.0688</b>	<b>22.8637</b>	<b>22.7087</b>	<b>27.5078</b>
	joint-r	<b>31.5125</b>	<b>23.0971</b>	<b>22.7088</b>	<b>27.8055</b>
<b>SSIM</b>	sparse	0.7398	0.8313	0.7511	0.6759
	joint	<b>0.7501</b>	<b>0.8545</b>	<b>0.8249</b>	<b>0.8144</b>
	joint-r	<b>0.8113</b>	<b>0.8580</b>	<b>0.8249</b>	<b>0.8144</b>

Table 2  
*Reconstruction quality - SSCSI architecture*

Data cube		C1	C2	C3	C4
% of captured data		25 %	35.15 %	50 %	39.06 %
Spatial PSNR [dB]	sparse	31.3648	33.9345	35.0364	37.4355
	joint	<b>35.3002</b>	<b>34.6322</b>	<b>36.2072</b>	<b>38.8397</b>
	joint-r	<b>38.3572</b>	<b>36.6847</b>	<b>37.3229</b>	<b>38.8592</b>
Spectral PSNR [dB]	sparse	32.1510	31.9740	34.0564	35.6405
	joint	<b>36.3804</b>	<b>33.1888</b>	<b>36.9949</b>	<b>37.3599</b>
	joint-r	<b>43.2932</b>	<b>34.1520</b>	<b>36.7396</b>	<b>37.6759</b>
SSIM	sparse	0.9523	0.9768	0.9843	0.9699
	joint	<b>0.9828</b>	<b>0.9808</b>	<b>0.9886</b>	<b>0.9767</b>
	joint-r	<b>0.9873</b>	<b>0.9878</b>	<b>0.9907</b>	<b>0.9777</b>

*Quality test along different amount of captured data.* This experiment compares the reconstruction quality varying the amount of captured data. For this, recall that in the CASSI and SSCSI architectures the percentage of captured data depends on the number of acquired snapshots. In particular, they were acquired 2, 6, and 10 snapshots corresponding to 13.96%, 41.89%, and 69.82% of captured information with the CASSI architecture, and 32, 45, and 64 snapshots corresponding to 25%, 35.15%, and 50% of captured information with the SSCSI architecture over configuration C1. Tables 3 and 4 summarize the results for the CASSI and SSCSI architectures, respectively. It can be observed that the larger the amount of information, the higher the obtained reconstruction quality for both, the CASSI and SSCSI architectures. Also, observe that for the CASSI architecture a higher gain is obtained when using more compressed measurements. In contrast, for SSCSI the gain is more significant when capturing less compressed measurements given that the optical setup allows to better consider the spectral correlations.

Table 3

*Reconstruction quality vs captured data - CASSI architecture*

% of captured data		13.96%	41.89%	69,82%
<b>Spatial PSNR</b> [dB]	sparse	13.1288	20.9784	25.7297
	joint	<b>15.1649</b>	<b>21.0042</b>	<b>26.4576</b>
	joint-r	<b>15.4591</b>	<b>22.6785</b>	<b>28.7845</b>
<b>Spectral PSNR</b> [dB]	sparse	20.3882	29.1606	33.4722
	joint	<b>23.1994</b>	<b>30.0688</b>	<b>35.0245</b>
	joint-r	<b>23.2654</b>	<b>31.5125</b>	<b>37.0460</b>
<b>SSIM</b>	sparse	0.3638	0.7398	0.8726
	joint	<b>0.4662</b>	<b>0.7501</b>	<b>0.8893</b>
	joint-r	<b>0.4805</b>	<b>0.8113</b>	<b>0.9172</b>

Table 4

*Reconstruction quality vs captured data - SSCSI architecture*

% of captured data		25%	35.15%	50%
<b>Spatial PSNR</b> [dB]	sparse	31.3648	37.1098	39.7283
	joint	<b>35.3002</b>	<b>37.9113</b>	<b>40.9896</b>
	joint-r	<b>38.3572</b>	<b>41.7885</b>	<b>44.1918</b>
<b>Spectral PSNR</b> [dB]	sparse	32.1510	42.9314	45.5245
	joint	<b>36.3804</b>	<b>43.0472</b>	<b>46.6181</b>
	joint-r	<b>43.2932</b>	<b>46.6506</b>	<b>49.5578</b>
<b>SSIM</b>	sparse	0.9523	0.9872	0.9924
	joint	<b>0.9828</b>	<b>0.9895</b>	<b>0.9951</b>
	joint-r	<b>0.9873</b>	<b>0.9953</b>	<b>0.9971</b>

*Quality test varying the spatial-spectral resolution ratio.* This experiment compares reconstruction quality varying the SI spatial-spectral resolution ratio calculated as  $SSR = L/N^2$ . For this, **C5**, **C6**, **C7**, and **C1** that account for the same dataset with the same spatial resolution but different spectral resolution were employed. Tables 5 and 6 present quality results along the different spatial-spectral ratios. The intuition aimed that the bigger the value for  $SSR$ , the highest

the reconstruction quality gain, since it means more redundancy in the low-rank matrix approximation and hence, more gain when taking advantage of this property. Nevertheless, this is not the case for the CASSI architecture, since the higher gain appears to be when a smaller spatial-spectral ratio is used, and variables as the rank of the low-rank matrix or the ratio between the number of endmembers and the resolutions of the SI must be considered. On the other hand, for the SSCSI architecture, a direct effect of the *SSR* over the reconstruction quality is evident since in general, a higher ratio produces a higher reconstruction quality under a comparative amount of captured data.

Table 5

*Reconstruction quality vs spatial-spectral resolution ratio - CASSI architecture*

Data cube		C5	C6	C7	C1
% of captured data		72.66 %	73.43 %	69.43 %	69.82 %
<i>SSR</i>		6.25 %	12.5 %	25 %	50 %
Spatial PSNR [dB]	sparse	25.3656	24.3692	24.3111	25.7297
	joint	25.0790	25.6642	25.8098	26.4576
	joint-r	30.0120	28.5988	28.1130	28.7845
Spectral PSNR [dB]	sparse	32.2292	31.4208	32.1225	33.4722
	joint	32.4963	32.8737	33.5321	35.0245
	joint-r	37.4498	35.7267	35.5775	37.0460
SSIM	sparse	0.9325	0.8798	0.8530	0.8726
	joint	0.9224	0.8957	0.8787	0.8893
	joint-r	0.9747	0.9404	0.9117	0.9172

*Recovered spectral distributions comparison.* Figure 6 (a) and Fig. 6 (b) show a comparison of the reconstructed spectral distribution at two spatial locations of configuration **C1** by using the “sparse”, “joint”, and “joint-r” algorithms using acquired compressed measurements

Table 6

*Reconstruction quality vs spatial-spectral resolution ratio - SSCSI architecture*

Data cube		C5	C6	C7	C1
% of captured data		25 %	25 %	25 %	25 %
SSR		6.25 %	12.5 %	25 %	50 %
Spatial PSNR [dB]	sparse	19.4493	22.4640	27.9729	31.3648
	joint	<b>23.9089</b>	<b>24.2759</b>	<b>30.8132</b>	<b>35.3002</b>
	joint-r	<b>25.8907</b>	<b>33.7786</b>	<b>37.0001</b>	<b>38.3572</b>
Spectral PSNR [dB]	sparse	25.6091	30.4848	34.2171	32.1510
	joint	<b>28.7140</b>	<b>32.2180</b>	<b>35.4849</b>	<b>36.3804</b>
	joint-r	<b>32.0526</b>	<b>40.3016</b>	<b>41.9582</b>	<b>43.2932</b>
SSIM	sparse	0.8349	0.8505	0.9160	0.9523
	joint	<b>0.9427</b>	<b>0.8946</b>	<b>0.9595</b>	<b>0.9828</b>
	joint-r	<b>0.9529</b>	<b>0.9847</b>	<b>0.9874</b>	<b>0.9873</b>

from the CASSI architecture. It can be observed the comparison of the recovered spectral distribution of two different spatial pixels when using 41.89% of captured data with the proposed CSI reconstruction algorithm against the traditional method and the corresponding ground-truth. The spectral distributions evidence the gain of taking advantage of the global spectral low-rank prior into the proposed CSI reconstruction algorithm inasmuch the recovered spectral distributions are closer to the reference one. Figure 6 (c) and Fig. 6 (d) show the same comparison over configuration **C5** using compressed measurements from the SSCSI architecture. It can be observed the spectral distribution at two spatial locations with 25% of captured data, also evidencing the effectiveness of the proposal in the reconstruction of the spectral distributions by estimating the spectral correlations through the global spectral low-rank regularization.

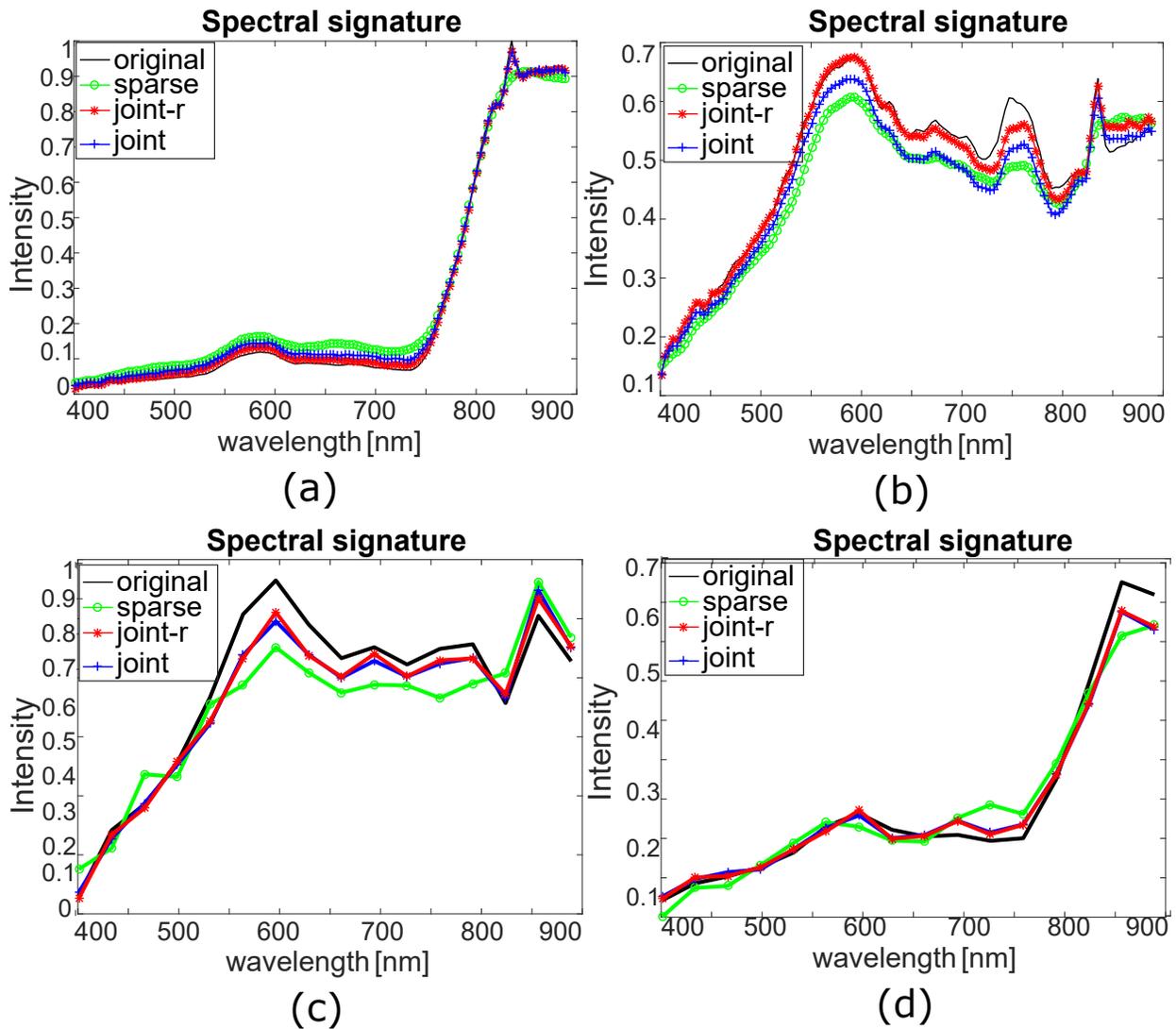


Figure 6. Recovered spectral distributions visual comparison

**Recovered spatial distribution comparison.** Figure 7 shows a visual comparison of the reconstructed spatial distributions over **C3** with the CASSI and SSCSI architectures. In detail, Fig. 7 *left* shows the reconstructions of the spectral bands 37, 70, 74, and 95 from CASSI compressed measurements with 41.89% of captured data, and Fig. 7 *right* shows the reconstructions of the spectral bands 37, 70, 74, and 95 from SSCSI compressed measurements with 50.00% of

captured data. It can be visually evidenced the improvements of the proposal, particularly, for the CASSI reconstructions where unlike the soft texture scenes obtained with the “sparse” algorithm, the “joint” and “joint-r” reconstructions exhibit a sharper appearance which results in more defined details, and in higher reconstruction quality. In the SSCSI reconstructions the gain is not visually clear along the spatial dimension since all PSNR are up to 30[dB], however, the gain was clearly observed along the spectral bands in Fig. 6.

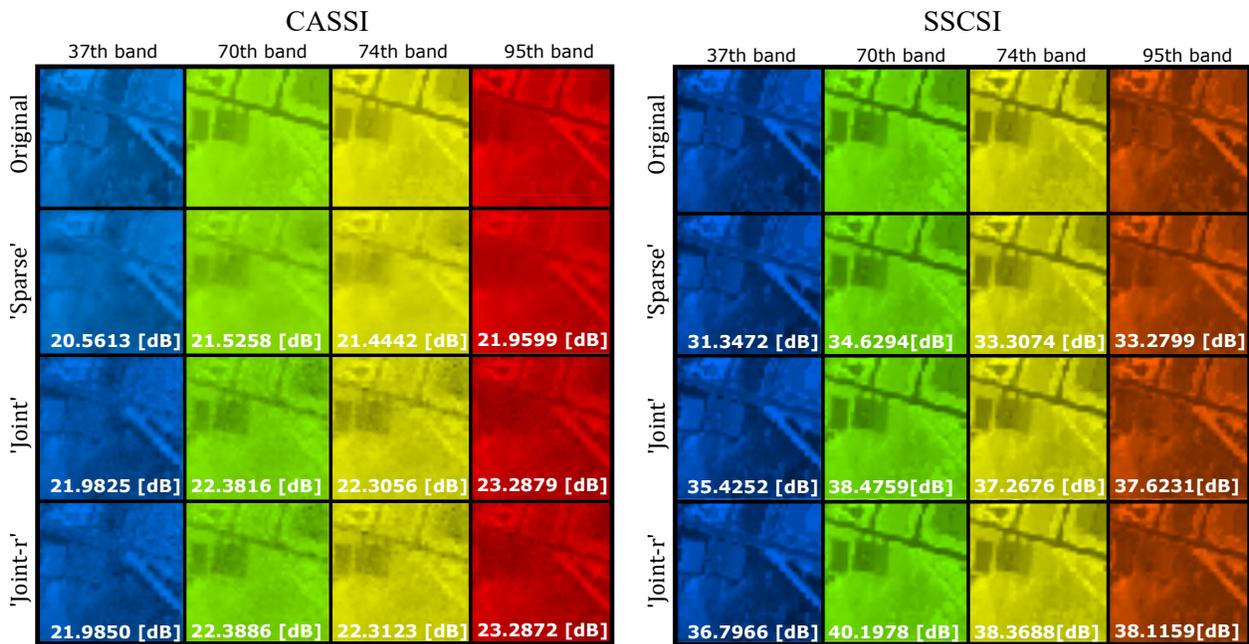


Figure 7. Reconstructed spatial distribution visual comparison.

## 2.6. Conclusions

This chapter presented the first approximation for including the low-rank prior into the SI recovery field. Specifically, this chapter studied the SI low-rankness from a global point of view, considering the spectral redundancy and the CSI reconstruction problem. The global spectral low-rank prior was validated through an extensive empirical study of the Pearson correlation matrix, showing

that it is especially strong for hyperspectral remote sensing datasets cause of its high spectral resolution. Furthermore, the nuclear norm was chosen as the regularization function to measure and promote the global spectral low-rank prior, incorporated into an unconstrained convex optimization problem. An alternating iterative algorithm was developed to solve the formulated inverse problem, presenting its computational complexity and proving its convergence. Notably, the complexity is bounded by the complexity of the SVD computation required by the nuclear norm minimization, so looking for a faster computation or strategy to obtain a low-rank approximation is suggested for future works. Finally, the developed method was tested over several datasets satisfying the low-rank prior when represented as a matrix to perform CSI reconstruction from simulated CASSI, and SSCSI compressed measurements. The proposed algorithm improved the reconstruction quality by up to 4dB in the PSNR and showed robustness to the variation of the spatial-spectral resolution ratio and the compressive acquisition protocol.

### 3. Non-local low-rank abundance prior for compressive spectral imaging fusion

*Part of this chapter has been adapted from the journal paper (Gelvez and Arguello, 2020) [© [2020] IEEE. Reprinted] (published) and conference (Gelvez and Arguello, 2018) (published).*

This chapter presents the second developed approach for including low-rank regularizations in spectral imaging recovery, where I propose a CSI fusion method to combine the spatial-spectral data from CASSI and SSCSI compressed measurements. The method employs the low-rank prior in two ways: (i) the global low-rank through the LMM, and the non-local low-rank prior through an explicit regularization term. Section 3.1 presents the most related works for CSI fusion and the found gaps motivating the proposed approach, with the following main contributions.

1. The analysis of the global low-rank prior through a theoretical premise using the LMM, and the validation of the non-local low-rank prior through an extensive empirical study of the non-local self-similarities, are described in Section 3.2.
2. The formulation of a convex optimization problem that combines the global low-rank with the non-local low-rank, is presented in Section 3.3.
3. The development of a numerical algorithm that solves the formulated convex optimization problem following a BCDM, is detailed in Section 3.4.
4. A quantitative gain in the quality of the CSI fusion along various datasets, validating the effectiveness of the non-local low-rank prior through several simulations and experiments as reported in Section 3.5.

### 3.1. Compressive spectral imaging fusion related works

The CSI fusion problem aims to obtain an HRI fusing the information from two sets of compressed measurements. The first set contains compressed data from an HSI with a fine spectral sampling but low-spatial resolution. The second set contains compressed data from an MSI with a fine spatial sampling but low-spectral resolution (Teodoro et al., 2017). Thus, CSI fusion avoids reconstructing the entire HRI and MSI independently (Chen et al., 2014).

Literature reports few related works to the CSI fusion problem, which mainly take advantage of the HRI smoothness property assuming that natural scenes exhibit smooth transitions within the spatial pixels. The non-smooth pixels commonly correspond to sharp edges or object's boundaries. To name, authors in (Vargas et al., 2017) employed the LMM and promoted a smooth HRI through the total-variation regularization (TV). Then, authors in (Vargas et al., 2019) employed the  $\ell_1$  regularization to consider the sparsity more than smoothness prior benefiting scenarios where the HRI results highly compressible. Finally, (Vargas et al., 2018a) focused the formulation on the design and optimization of the MSI and HRI sensing matrices more than in the HRI regularization, so that the acquisition process of the compressed measurements benefits the CSI fusion. These methods have a gap regarding the low-rank property, particularly, when combining the rich spectral information preserved in the HRI, motivating the use of a spectral low-rank prior, with the rich spatial information preserved in the MSI, motivating the use of a spatial low-rank prior, so that, the HRI can recover its refined spatial-spectral information. Considering that to my knowledge, no study to date has examined the implementation of the low-rank priors into the CSI fusion inverse

problem, this chapter presents an approach to fuse the information from two sets of compressed measurements including the global low-rank and the non-local low-rank priors.

### 3.2. Global and non-local low-rank regularization analysis

The spatial-spectral HRI redundancy is analyzed from two perspectives: the global low-rank through a theoretical premise via the LMM, and the non-local low-rank through an extensive empirical study. To ease the physical understanding and the mathematical formulations, Figure 8 illustrates an HRI as a third-order tensor  $\mathcal{F} \in \mathbb{R}^{N \times N \times L}$ , the unfolding as a 2D matrix  $\mathbf{F} \in \mathbb{R}^{N^2 \times L}$ , where each column contains the vectorization of one spectral band, and the LMM of  $\mathbf{F}$  decomposing the matrix  $\mathbf{F}$  into the product of an abundance matrix  $\mathbf{A}$ , and an endmember matrix  $\mathbf{E}$ .

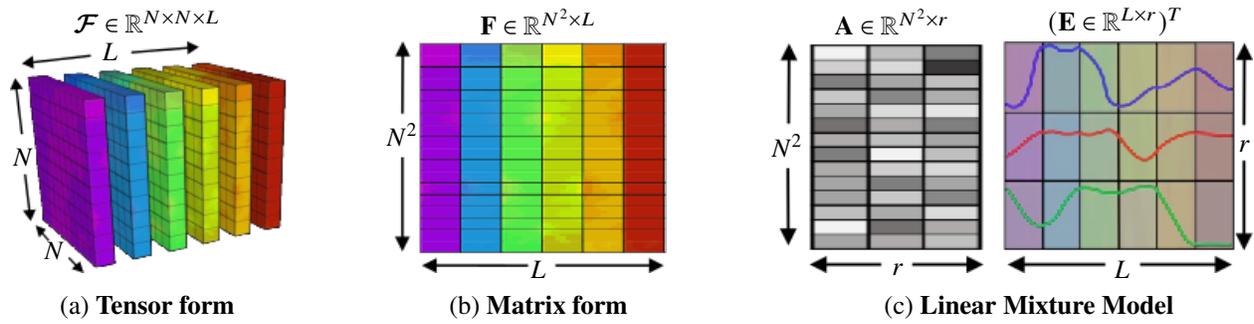


Figure 8. HRI visual representations.

**3.2.1. Global low-rank theoretical premise.** Assuming that an HRI only contains  $r \ll \min(N^2, L)$  different materials, each one characterized by a unique spectral distribution, the LMM states that the matrix representing the HRI can be decomposed as the product  $\mathbf{F} = \mathbf{A}\mathbf{E}^T$ , where  $\mathbf{A} \in \mathbb{R}^{N^2 \times r} = [\mathbf{a}_1 \dots \mathbf{a}_i \dots \mathbf{a}_{N^2}]^T$  represents an abundance matrix containing the spatial fractional proportions  $\mathbf{a}_i \in \mathbb{R}^r$  of each spatial pixel, and  $\mathbf{E} \in \mathbb{R}^{L \times r} = [\mathbf{e}_1 \dots \mathbf{e}_\zeta \dots \mathbf{e}_r]$  represents an endmembers matrix containing  $r$  different spectral distributions  $\mathbf{e}_\zeta \in \mathbb{R}^L$  that span the image.

The rank of the matrix  $\mathbf{F}$  is then mathematically determined by the minimum between the rank of the matrices in the decomposition, i.e between the rank of  $\mathbf{A}$  and the rank of  $\mathbf{E}$ . Given that the rank of a matrix is less or equal to its fewer dimension, the rank of  $\mathbf{A}$  and  $\mathbf{E}$  typically corresponds to the number of spectral distributions  $r$ . Therefore, the  $\text{rank}(\mathbf{F}) = \min(\text{rank}(\mathbf{A}), \text{rank}(\mathbf{E})) \leq r \ll \min(N^2, L)$ . Thus,  $\mathbf{F}$  is shown to be low-rank since  $\text{rank}(\mathbf{F}) \ll \min(N^2, L)$ , supporting the inclusion of the global low-rank prior through the LMM.

**3.2.2. Non-local low-rank empirical study.** Then, this subsection presents an extensive empirical study of the non-local low-rank prior, where the the singular value decay of the matrices grouping similar patches for various spectral datasets is expected to be fast, especially from the first to the second singular value. The empirical study follows the procedure below, where  $\mathbf{p}_i \in \mathbb{R}^{p^2 L}$ ,  $p \in \mathbb{Z}_+ \ll N$  denotes a reference patch aligned at top-left corner at the  $i^{\text{th}}$  location of the HRI containing  $p^2$  spatial pixels.

1. Extraction of  $I = N^2 - p^2 + 1$  overlapped patches modeled as  $\mathbf{p}_i = \mathbf{T}_i \mathbf{f}$ , where  $\mathbf{T}_i \in \{0, 1\}^{p^2 L \times N^2 L}$  is a selection matrix to extract the spatial pixels of the  $i^{\text{th}}$  reference patch. Furthermore, let  $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_i, \dots, \mathbf{p}_I\}$  be the set containing the  $I$  extracted patches, where the same spatial pixel can appear several times in different extracted patches.
2. Patch grouping to build  $G$  patch-groups, through the partition of the set  $\mathcal{P}$  with  $G$  disjoint subsets, such that, each patch can belong only to one patch-group. Thus, let  $\text{Part}(\mathcal{P})$  be the set of partitions of  $\mathcal{P}$ , so that, the set of  $G$  patch-groups  $\bar{\mathcal{P}} \subset \text{Part}(\mathcal{P})$  is given by  $\bar{\mathcal{P}} = \{\mathcal{P}_1, \dots, \mathcal{P}_g, \dots, \mathcal{P}_G\}$ , where  $|\mathcal{P}_g| = I_g$ , satisfying  $\sum_{g=1}^G I_g = I$ .

In particular,  $\bar{\mathcal{P}}$  is the partition in which each subset  $\mathcal{P}_g$  contains similar patches chosen according to a similarity degree metric. This step can be performed by the  $k$ -means method, which aims at partitioning a dataset into distinct non-overlapping clusters where each data point belongs to only one group (Lloyd, 1982).

To ease notation, let  $\mathbf{P}_g \in \mathbb{R}^{p^2 L \times I_g} = \mathcal{C}_g(\mathbf{f})$  be the matrix representation of the  $g^{th}$  patch-group ( $\mathcal{P}_g$ ), where  $\mathcal{C}_g : \mathbb{R}^{N^2 L} \rightarrow \mathbb{R}^{p^2 L \times I_g}$  is the operator to extract the patches belonging to  $\mathbf{P}_g$  from the HRI ( $\mathbf{f}$ ).

3. Analysis of each matrix representing a patch-group ( $\mathbf{P}_g$ ), whose singular value decay is expected to be fast, so that, the patch-group can be well approximated with a low-rank matrix.

Figure 9 illustrates the decay of the singular values of all patch-groups extracted for every analyzed dataset. The color histogram at the right of the singular value decays indicates the amount of patches contained at each build patch-group, given by the  $k$ -means method. Furthermore, the amount of patch-groups varies from one dataset to another in concordance to the resolution of the dataset. It can be observed the expected singular value fast decay from the first to the second singular value. Therefore, the empirical analysis supports the use of the non-local low-rank prior into the solution of spectral imaging inverse problems, where the patch-groups can be approximated with a low-rank matrix. Besides, notice that this behaviour is stronger over the hyperspectral remote sensing datasets than over the multispectral datasets, because of the combination of high spatial-spectral redundancies.

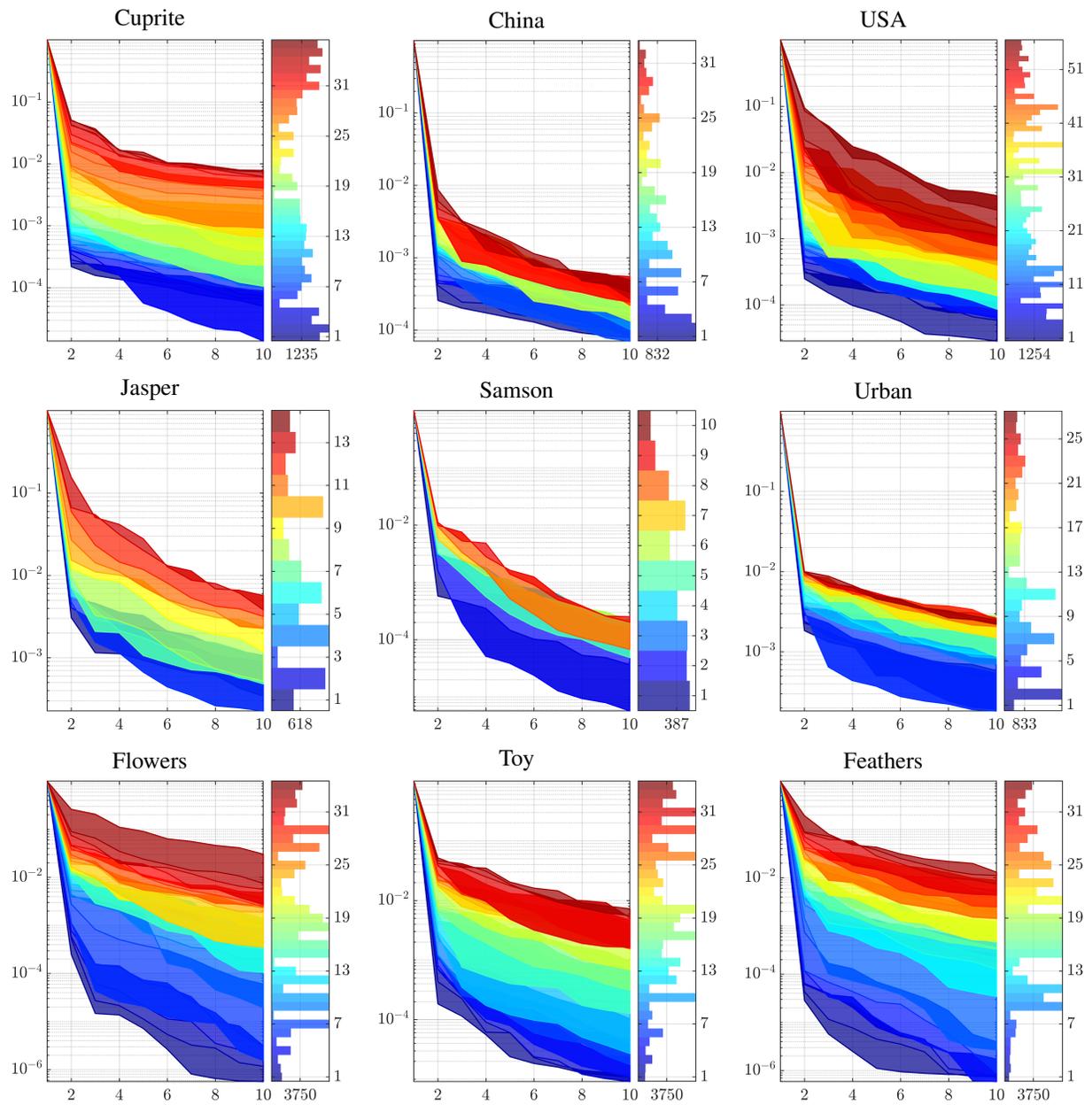


Figure 9. Singular value decay for non-local patch-groups.

### 3.3. Optimization problem formulation for compressive spectral imaging fusion

Given the CSI fusion formulation in 1.3.5, and the LMM formulated in 1.4.1, the CSI fusion observation model is given by

$$\begin{aligned}\mathbf{y}_h &= \mathbf{H}_h \mathbf{f}_h + \mathbf{n}_h = \mathbf{H}_h \bar{\mathbf{M}} \bar{\mathbf{B}} \bar{\mathbf{E}} \mathbf{a} + \mathbf{n}_h = \mathbf{H}_h \mathbf{S}_h \bar{\mathbf{E}} \mathbf{a} + \mathbf{n}_h, \\ \mathbf{y}_m &= \mathbf{H}_m \mathbf{f}_m + \mathbf{n}_m = \mathbf{H}_m \bar{\mathbf{R}} \bar{\mathbf{E}} \mathbf{a} + \mathbf{n}_m = \mathbf{H}_m \mathbf{S}_m \bar{\mathbf{E}} \mathbf{a} + \mathbf{n}_m,\end{aligned}\tag{19}$$

where  $\mathbf{y}_h \in \mathbb{R}^{m_h}$  denotes the HRI compressed measurements acquired with an optical system modeled by the sensing matrix  $\mathbf{H}_h \in \mathbb{R}^{m_h \times N_h^2 L}$ , with  $m_h \ll N_h^2 L$ ;  $\mathbf{y}_m \in \mathbb{R}^{m_m}$  denotes the MSI compressed measurements acquired with an optical system modeled by the sensing matrix  $\mathbf{H}_m \in \mathbb{R}^{m_m \times N^2 L_m}$ , with  $m_m \ll N^2 L_m$ ; and  $\mathbf{n}_h \in \mathbb{R}^{m_h}$ , and  $\mathbf{n}_m \in \mathbb{R}^{m_m}$  denote additive Gaussian noise.

I formulate a CSI fusion that joints the smoothness, global low-rank, and non-local low-rank priors, where the endmembers are directly estimated from the compressed measurements, or initialized from a fast reconstruction of the HRI, and the abundances are iteratively updated by promoting the proposed priors. The global low-rank prior is implicitly included with the estimation of the HRI as the product of variables  $\bar{\mathbf{E}}$  and  $\mathbf{a}$ , and the smoothness and non-local low-rank priors are explicitly included through two regularization terms over the abundances instead of the entire HRI, reducing the computational complexity.

The data fidelity term w.r.t the compressed measurements  $\mathbf{y}_h$  and  $\mathbf{y}_m$  is given by

$$f(\mathbf{a}, \bar{\mathbf{E}}) = \frac{1}{2} \|\mathbf{H}_h \mathbf{S}_h \bar{\mathbf{E}} \mathbf{a} - \mathbf{y}_h\|_2^2 + \frac{1}{2} \|\mathbf{H}_m \mathbf{S}_m \bar{\mathbf{E}} \mathbf{a} - \mathbf{y}_m\|_2^2.\tag{20}$$

The smoothness prior that promotes smooth transitions within the spatial locations in the abundances can be included with the TV regularization, calculated as  $g(\mathbf{a}) = \|\bar{\mathbf{D}}_{\mathbf{v}}\mathbf{a}\|_1 + \|\bar{\mathbf{D}}_{\mathbf{h}}\mathbf{a}\|_1$ , where  $\bar{\mathbf{D}}_{\mathbf{v}} \in \mathbb{R}^{N^2 r \times N^2 r} = \mathbf{D}_{\mathbf{v}} \otimes \mathbf{I}_r$ , and  $\bar{\mathbf{D}}_{\mathbf{h}} \in \mathbb{R}^{N^2 r \times N^2 r} = \mathbf{D}_{\mathbf{h}} \otimes \mathbf{I}_r$  encompass the vertical and horizontal discrete difference operators  $\mathbf{D}_{\mathbf{v}} \in \mathbb{R}^{N^2 \times N^2}$ , and  $\mathbf{D}_{\mathbf{h}} \in \mathbb{R}^{N^2 \times N^2}$ , applied over the  $r$  abundance maps. These operators can be stacked in just one operator as  $g(\mathbf{a}) = \|\mathbf{D}\mathbf{a}\|_1$ .

The non-local low-rank prior that takes advantage of the self-similarities between small patches in the abundances can be included by finding the lower rank approximation of each patch-group built from the abundances. This prior can be formulated as  $h(\mathbf{a}) = \sum_{g=1}^G \|\mathcal{C}_g(\mathbf{a})\|_*$ , where  $\mathcal{C}_g : \mathbb{R}^{N^2 r} \rightarrow \mathbb{R}^{p^2 L \times I_g}$  denotes the operator to extract the patches belonging to  $\mathbf{P}_g$  from  $\mathbf{a}$ .

The proposed problem is then compactly formulated as

$$\begin{aligned} & \underset{\mathbf{a}, \bar{\mathbf{E}}}{\text{minimize}} && s(\mathbf{a}, \bar{\mathbf{E}}) = f(\mathbf{a}, \bar{\mathbf{E}}) + \tau g(\mathbf{a}) + \mu h(\mathbf{a}) \\ & \text{subject to} && \mathbf{a}_i[a] \geq 0, \quad \sum_{a=1}^r \mathbf{a}_i[a] = 1, \forall i, \end{aligned} \tag{21}$$

where,  $\tau > 0$  and  $\mu > 0$  are two parameters to control the relative weights of the regularization terms. Further, the non-negative and sum-to-one LMM constraints in (21) can be restated as

$$\underset{\mathbf{a}, \bar{\mathbf{E}}}{\text{minimize}} \quad f(\mathbf{a}, \bar{\mathbf{E}}) + \tau g(\mathbf{a}) + \mu h(\mathbf{a}) + \phi(\mathbf{a}), \tag{22}$$

where  $\phi(\cdot) : \mathbb{R}^{N^2r} \rightarrow \tilde{\mathbb{R}}$  is the following indicator function

$$\phi(\mathbf{a}) = \begin{cases} 0, & \mathbf{a} \in \mathcal{D} \\ \infty, & \mathbf{a} \notin \mathcal{D} \end{cases}, \quad (23)$$

with  $\mathcal{D} = \{\mathbf{d} \in \mathbb{R}^{N^2r} \mid \mathbf{d}_i[d] \geq 0, \sum_{d=1}^r \mathbf{d}_i[d] = 1, \forall i\}$ . This implies that the cost function can take values in the extended real number line  $\tilde{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ .

On the other hand, the HRI defined as  $\mathbf{f}_h = \mathbf{S}_h \bar{\mathbf{E}} \mathbf{a} = (\mathbf{I}_L \otimes \mathbf{M})(\mathbf{I}_L \otimes \mathbf{B}) \bar{\mathbf{E}} \mathbf{a}$  can be rewritten as

$$\mathbf{f}_h = (\mathbf{I}_L \otimes \mathbf{M}) \bar{\mathbf{E}} (\mathbf{I}_r \otimes \mathbf{B}) \mathbf{a} = \bar{\mathbf{M}} \bar{\mathbf{E}} \bar{\mathbf{B}} \mathbf{a}, \quad (24)$$

where  $\bar{\mathbf{M}} = \mathbf{I}_L \otimes \mathbf{M} \in \mathbb{R}^{N^2L \times N^2L}$  accounts for the downsampling, and  $\bar{\mathbf{B}} = \mathbf{I}_r \otimes \mathbf{B} \in \mathbb{R}^{N^2r \times N^2r}$  accounts for the blurring. This notation will be useful in the solution of the proposed optimization problem, which is lastly formulated by replacing each term of the cost function in (22) as follows

$$\underset{\mathbf{a}, \bar{\mathbf{E}}}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{H}_h \bar{\mathbf{M}} \bar{\mathbf{E}} \bar{\mathbf{B}} \mathbf{a} - \mathbf{y}_h\|_2^2 + \frac{1}{2} \|\mathbf{H}_m \mathbf{S}_m \bar{\mathbf{E}} \mathbf{a} - \mathbf{y}_m\|_2^2 + \tau \|\mathbf{D} \mathbf{a}\|_1 + \mu \sum_{g=1}^G \|\mathcal{C}_g(\mathbf{a})\|_* + \phi(\mathbf{a}). \quad (25)$$

### 3.4. Numerical algorithm

To solve (25), I formulate an algorithm following a BCDM that alternates the minimization with respect to the abundances  $\mathbf{a}$ , and the endmembers  $\bar{\mathbf{E}}$ . Each subproblem is solved using the ADMM as detailed below (Boyd et al., 2011).

**3.4.1. Abundance estimation algorithm.** The abundances update introduces the auxiliary variable  $\mathbf{v} = [\mathbf{v}_1^T \dots \mathbf{v}_7^T]^T$ , obtaining the minimization problem

$$\begin{aligned} \underset{\mathbf{v}, \mathbf{a}}{\text{minimize}} \quad \psi(\mathbf{v}, \mathbf{a}) &= \frac{1}{2} \|\mathbf{H}_h \bar{\mathbf{M}} \mathbf{v}_1 - \mathbf{y}_h\|_2^2 + \frac{1}{2} \|\mathbf{H}_m \mathbf{v}_3 - \mathbf{y}_m\|_2^2 + \\ &\quad \tau \|\mathbf{v}_5\|_1 + \mu \sum_{g=1}^G \|\mathcal{C}_g(\mathbf{v}_6)\|_* + \phi(\mathbf{v}_7), \\ \text{subject to} \quad \mathbf{v}_1 &= \bar{\mathbf{E}} \mathbf{v}_2; \quad \mathbf{v}_2 = \bar{\mathbf{B}} \mathbf{a}; \quad \mathbf{v}_3 = \mathbf{S}_m \bar{\mathbf{E}} \mathbf{v}_4; \quad \mathbf{v}_4 = \mathbf{a}; \\ &\quad \mathbf{v}_5 = \mathbf{D} \mathbf{a}; \quad \mathbf{v}_6 = \mathbf{a}; \quad \mathbf{v}_7 = \mathbf{a}. \end{aligned} \quad (26)$$

Introducing the variable  $\mathbf{v}$  yields to the equality constraints in (26). The ADMM then looks for minimizing the augmented Lagrangian function expressed as

$$\begin{aligned} \hat{\mathbf{a}}, \hat{\mathbf{v}}, \hat{\mathbf{g}} \in \underset{\mathbf{a}, \mathbf{v}, \mathbf{g}}{\text{argmin}} \quad \mathcal{L}(\mathbf{a}, \mathbf{v}, \mathbf{g}) &= \frac{1}{2} \|\mathbf{H}_h \bar{\mathbf{M}} \mathbf{v}_1 - \mathbf{y}_h\|_2^2 + \frac{1}{2} \|\mathbf{H}_m \mathbf{v}_3 - \mathbf{y}_m\|_2^2 + \\ &\quad \tau \|\mathbf{v}_5\|_1 + \mu \sum_{g=1}^G \|\mathcal{C}_g(\mathbf{v}_6)\|_* + \phi(\mathbf{v}_7) + \frac{\rho}{2} \|\mathbf{v}_1 - \bar{\mathbf{E}} \mathbf{v}_2 - \mathbf{g}_1\|_2^2 + \\ &\quad \frac{\rho}{2} \|\mathbf{v}_2 - \bar{\mathbf{B}} \mathbf{a} - \mathbf{g}_2\|_2^2 + \frac{\rho}{2} \|\mathbf{v}_3 - \mathbf{S}_m \bar{\mathbf{E}} \mathbf{v}_4 - \mathbf{g}_3\|_2^2 + \frac{\rho}{2} \|\mathbf{v}_4 - \mathbf{a} - \mathbf{g}_4\|_2^2 + \\ &\quad \frac{\rho}{2} \|\mathbf{v}_5 - \mathbf{D} \mathbf{a} - \mathbf{g}_5\|_2^2 + \frac{\rho}{2} \|\mathbf{v}_6 - \mathbf{a} - \mathbf{g}_6\|_2^2 + \frac{\rho}{2} \|\mathbf{v}_7 - \mathbf{a} - \mathbf{g}_7\|_2^2, \end{aligned} \quad (27)$$

with  $\rho > 0$  being the dual regularization parameter.

Minimization over variables  $\mathbf{a}$ ,  $\mathbf{v}$ , and  $\mathbf{g}$  in (27) leads to solutions summarized in Algorithm 5 with variables computed as detailed in Table 7. In Algorithm 5 line 5 the abundances are efficiently computed with the fast Fourier transform, in which,  $\mathcal{F}(\cdot)$ , and  $\mathcal{F}^{-1}(\cdot)$  denote the Fourier transform, and inverse Fourier transform operators, respectively, and  $\mathbf{w}_0 \in \mathbb{R}^{N^2 r}$  is the first column of  $\mathbf{W}_0$ .

Algorithm 5. Abundance estimation with the ADMM approach

---

1:	<b>procedure</b> ABUE( $\bar{\mathbf{E}}, \mathbf{H}_h, \mathbf{H}_m, \mathbf{y}_h, \mathbf{y}_m, \mathbf{M}, \mathbf{B}, \mathbf{D}, \mathbf{S}_m, \mu, \tau, \rho, \text{IterA}$ )	
2:	$w \leftarrow 0; \mathbf{v}^w \leftarrow \mathbf{0}; \mathbf{g}^w \leftarrow \mathbf{0}$	▷ Initialization
3:	$\mathbf{W}_\alpha \leftarrow \text{Table 7 } \alpha = 1, \dots, 4$	▷ Precomputation
4:	<b>while</b> $w < \text{IterA}$ <b>do</b>	▷ Update
5:	$\mathbf{a}^{w+1} \leftarrow \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{\mathbf{b}_0^w\}}{\mathcal{F}\{\mathbf{w}_0\}} \right\}$	
6:	$\mathbf{v}_\alpha^{w+1} \leftarrow \mathbf{W}_\alpha^{-1} \mathbf{b}_\alpha^w, \alpha = 1, \dots, 4.$	
7:	$\mathbf{v}_5^{w+1} \leftarrow \text{soft}_{\tau/\rho}(\mathbf{D}\mathbf{a}^{w+1} + \mathbf{g}_5^w)$	
8:	$\mathbf{v}_6^{w+1} \leftarrow \text{NonLocal}_{\mu/\rho}(\mathbf{a}^{w+1} + \mathbf{g}_6^w)$	
9:	$\mathbf{v}_7^{w+1} \leftarrow \text{SimplexProjection}(\mathbf{a}^{w+1} + \mathbf{g}_7^w)$	
10:	$\mathbf{g}^{w+1} \leftarrow \mathbf{g}^w + \mathbf{v}^{w+1} - \mathbf{c}^{w+1}$	
11:	$w \leftarrow w + 1$	
12:	<b>return</b> $\mathbf{a}^{\text{IterA}}$	

---

Table 7

Summary variable computation in the abundance estimation

$\mathbf{W}_0$	$\in \mathbb{R}^{N^2 r \times N^2 r} = \bar{\mathbf{B}}^T \bar{\mathbf{B}} + 3\mathbf{I}_{N^2 r} + \mathbf{D}^T \mathbf{D}$
$\mathbf{W}_1$	$\in \mathbb{R}^{N^2 L \times N^2 L} = (\mathbf{H}_h \bar{\mathbf{M}})^T (\mathbf{H}_h \bar{\mathbf{M}}) + \rho \mathbf{I}_{N^2 L}$
$\mathbf{W}_2$	$\in \mathbb{R}^{N^2 r \times N^2 r} = \bar{\mathbf{E}}^T \bar{\mathbf{E}} + \mathbf{I}_{N^2 r}$
$\mathbf{W}_3$	$\in \mathbb{R}^{N^2 L_m \times N^2 L_m} = \mathbf{H}_m^T \mathbf{H}_m + \rho \mathbf{I}_{N^2 L_m}$
$\mathbf{W}_4$	$\in \mathbb{R}^{N^2 r \times N^2 r} = (\mathbf{S}_m \bar{\mathbf{E}})^T (\mathbf{S}_m \bar{\mathbf{E}}) + \mathbf{I}_{N^2 r}$
$\mathbf{b}_0^w$	$\in \mathbb{R}^{N^2 r} = \bar{\mathbf{B}}^T (\mathbf{v}_2^w - \mathbf{g}_2^w) + (\mathbf{v}_4^w - \mathbf{g}_4^w) + \mathbf{D}^T (\mathbf{v}_5^w - \mathbf{g}_5^w) + (\mathbf{v}_6^w - \mathbf{g}_6^w) + (\mathbf{v}_7^w - \mathbf{g}_7^w)$
$\mathbf{b}_1^w$	$\in \mathbb{R}^{N^2 L} = (\mathbf{H}_h \bar{\mathbf{M}})^T \mathbf{y}_h + \rho (\bar{\mathbf{E}} \mathbf{v}_2^w + \mathbf{g}_1^w)$
$\mathbf{b}_2^w$	$\in \mathbb{R}^{N^2 r} = \bar{\mathbf{E}}^T (\mathbf{v}_1^w - \mathbf{g}_1^w) + (\bar{\mathbf{B}} \mathbf{a}^{w+1} + \mathbf{g}_2^w)$
$\mathbf{b}_3^w$	$\in \mathbb{R}^{N^2 L_m} = \mathbf{H}_m^T \mathbf{y}_m + \rho (\mathbf{S}_m \bar{\mathbf{E}} \mathbf{v}_4^w + \mathbf{g}_3^w)$
$\mathbf{b}_4^w$	$\in \mathbb{R}^{N^2 r} = (\mathbf{S}_m \bar{\mathbf{E}})^T (\mathbf{v}_3^{w+1} - \mathbf{g}_3^w) + (\mathbf{a}^{w+1} + \mathbf{g}_4^w)$
$\mathbf{c}^{w+1}$	$= [(\bar{\mathbf{E}} \mathbf{v}_2^{w+1})^T \quad (\bar{\mathbf{B}} \mathbf{a}^{w+1})^T \quad (\mathbf{S}_m \bar{\mathbf{E}} \mathbf{v}_4^{w+1})^T \quad (\mathbf{a}^{w+1})^T \quad (\mathbf{D} \mathbf{a}^{w+1})^T \quad (\mathbf{a}^{w+1})^T \quad (\mathbf{a}^{w+1})^T]^T$

In Algorithm 5 line 8,  $\text{NonLocal}_{\tau/\rho}(\cdot)$  denotes the procedure to promote the non-local low-rank prior with parameter  $\tau/\rho$  over the abundances. For this, let  $\mathbf{z}_6^w = \mathbf{a}^{w+1} + \mathbf{g}_6^w$  be a non-local high-rank version of the abundances. The followed procedure is given below.

1. Extract  $I$  patches as  $\mathbf{p}_i = \mathbf{T}_i \mathbf{z}_6^w$ , where  $\mathbf{T}_i \in \{0, 1\}^{p^2 r \times N^2 r}$ .
2. Build  $G$  patch-groups as  $\mathbf{P}_{6,g} \in \mathbb{R}^{p^2 r \times I_g} = \mathcal{C}_g(\mathbf{z}_6^w)$ .
3. Find the low-rank approximation for each patch-group, assuming a sufficient amount of elements by solving the optimization problems

$$\underset{\mathbf{Q}_{6,g}}{\text{minimize}} \quad \mu \|\mathbf{Q}_{6,g}\|_* + \frac{\rho}{2} \|\mathbf{Q}_{6,g} - \mathbf{P}_{6,g}\|_2^2, \quad \text{for } g = 1, \dots, G, \quad (28)$$

where  $\mathbf{Q}_{6,g} := \mathcal{C}_g(\mathbf{v}_6)$  is estimated with the soft-thresholding operator over the SVD as

$$\hat{\mathbf{Q}}_{6,g} \leftarrow \mathbf{U}_{6,g} \text{soft}_{\mu/\rho}(\sigma(\mathbf{P}_{6,g})) \mathbf{V}_{6,g}^T. \quad (29)$$

$\mathbf{U}_{6,g}$  refers to the left singular vectors,  $\sigma(\cdot)$  indicates the singular values, and  $\mathbf{V}_{6,g}$  refers to the right singular vectors of the matrix  $\mathbf{P}_{6,g}$ .  $\text{soft}_{\xi}(\cdot)$  indicates the soft thresholding operator that for a structure  $\mathbf{o} \in \mathbb{R}^{\Gamma}$  is given by,

$$\text{soft}_{\xi}(\mathbf{o}) = \begin{cases} 0, & o_{\chi} \leq \xi \\ o_{\chi}, & o_{\chi} > \xi \end{cases} \quad \text{for } \chi = 1 \dots \Gamma. \quad (30)$$

4. Estimate the value of each entry of the variable  $\mathbf{v}_6$  through an aggregation step that averages

the individual estimations along the  $G$  low-rank approximations of the patch-groups as

$$\hat{\mathbf{v}}_6 = \frac{1}{\mathbf{u}} \circ \sum_{g=1}^G \mathcal{C}_g^{-1}(\mathbf{Q}_{6,g}), \quad (31)$$

where  $\mathbf{u} \in \mathbb{R}^{N^2r}$  is a weighted array related to the number of times that each entry is repeated along the patch-groups.

Finally, in Algorithm 5 line 9 SimplexProjection( $\cdot$ ) indicates the projection of the argument to the set  $\mathcal{D}$  defined in (23).

Figure 10 sums up the entire scheme, where the HRI ( $\mathbf{f}_h$ ) and the MSI ( $\mathbf{f}_m$ ) are modeled as spatial/spectral degraded versions of an HRI ( $\mathbf{f}$ ). The degradation phenomena is implicit in the sensing process and it cannot be controlled. The measurements  $\mathbf{y}_h$  and  $\mathbf{y}_m$  are acquired by using compressive optical systems denoted by matrices  $\mathbf{H}_h$ , and  $\mathbf{H}_m$ . The information from  $\mathbf{y}_h$  and  $\mathbf{y}_m$  is fused to obtain the HRI with a proposed CSI fusion based on the LMM, such that, the HRI is decomposed as a linear combination of a set of few endmembers ( $\mathbf{E}$ ) to specific abundances ( $\mathbf{a}$ ). Besides, it considers two scenarios: (i) the endmembers can be estimated from the HRI compressed measurements, or (ii) initialized from a fast reconstruction of the HRI. The inverse problem includes two data fidelity terms, a smoothness, and a non-local low-rank prior. The non-local low-rank prior illustrated in Fig. 10 (right) takes advantage of the self-similarities by following four steps, (i) extraction of  $I$  patches of fixed size from a previous estimation of the abundances  $\hat{\mathbf{a}}^w$ ; (ii) grouping of the patches into  $G$  patch-groups by using a similarity degree metric; (iii) finding of the lower-rank approximation for each patch-group; and (iv) aggregation of the  $\mathbf{u}_i$  estimations of each

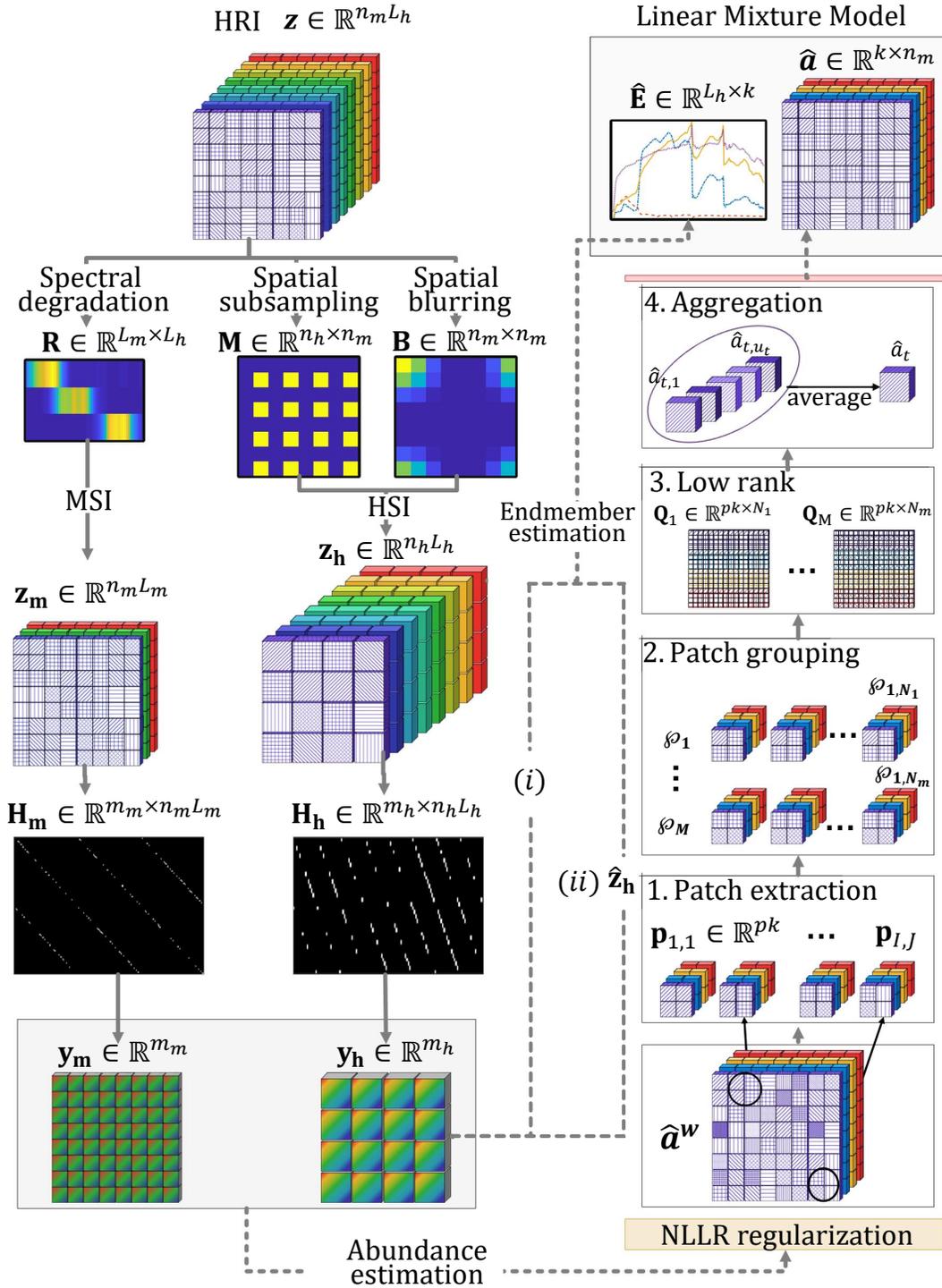


Figure 10. Scheme of the CSI fusion approach with non-local low-rank prior.

entry along the  $G$  low-rank patch-groups. These steps are repeated iteratively in the algorithm until the stopping criteria is achieved.

**Abundance estimation algorithm convergence.** The convergence of Algorithm 5 is guaranteed under the global feasibility, coercivity, Lipschitz, and regularity conditions for the convergence of ADMM optimization described in Theorem 1 of (Wang et al., 2019c). For this purpose, let us define the feasible set of problem (26) as

$$\mathcal{E} := \{(\mathbf{v}, \mathbf{a}) \in \mathbb{R}^{(L+L_m+6)N^2} : \mathbf{J}\mathbf{v} + \mathbf{K}\mathbf{a} = \mathbf{0}\},$$

$$\mathbf{J} = \begin{bmatrix} \mathbf{I} & -\bar{\mathbf{E}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & -\mathbf{S}_m\bar{\mathbf{E}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}; \quad \mathbf{K} = \begin{bmatrix} \mathbf{0} \\ -\bar{\mathbf{B}} \\ \mathbf{0} \\ -\mathbf{I} \\ -\mathbf{D} \\ -\mathbf{I} \\ -\mathbf{I} \end{bmatrix}. \quad (32)$$

Assume that  $\text{Im}(\mathbf{J}) \subseteq \text{Im}(\mathbf{K})$ , such that, the problem is feasible. Subsequently, observe that:

1. All terms of the objective function in (26) are non-negative, such that,  $\psi(\mathbf{v}, \mathbf{a}) \rightarrow \infty$  if  $(\mathbf{v}, \mathbf{a}) \in \mathcal{E}$  and  $\|(\mathbf{v}, \mathbf{a})\| \rightarrow \infty$ . Thus, the coercivity condition is satisfied.
2. For any fixed  $\mathbf{v}$ ,  $\text{argmin}_{\mathbf{a}}\{\psi(\mathbf{v}, \mathbf{a}) : \mathbf{K}\mathbf{a} = r\}$  results in a strongly convex function on closed, convex set, such that, it has a unique minimizer. Likewise, when minimizing w.r.t. each independent variable  $\mathbf{v}_1, \dots, \mathbf{v}_7$ , the resulting subproblems are convex on closed, convex sets,

such that, they have unique minimizers. Thus, the Lipschitz sub-minimization paths condition is satisfied.

3. The objective function can be decoupled as  $\psi(\mathbf{v}, \mathbf{a}) = \sum_f \psi_f(\mathbf{v}_f)$ , where  $\psi_1(\mathbf{v}_1)$  is an  $\ell_2$ -norm, so it is clearly continuous, and  $\psi_f(\mathbf{v}_f)$ , for  $f \neq 1$ , consists on four norms and one indicator function, so that, they are all restricted prox-regularity functions (Wang et al., 2019c).

As a consequence, Algorithm 5 to estimate the abundances converges.

**3.4.2. Endmember estimation algorithm.** The endmembers update introduces the auxiliary variable  $\mathbf{w} = [\mathbf{w}_1^T, \dots, \mathbf{w}_4^T]^T$ , obtaining the minimization problem

$$\begin{aligned} & \underset{\bar{\mathbf{E}}, \mathbf{w}}{\text{minimize}} && \frac{1}{2} \|\mathbf{H}_h \bar{\mathbf{M}} \mathbf{w}_1 - \mathbf{y}_h\|_2^2 + \frac{1}{2} \|\mathbf{H}_m \mathbf{w}_3 - \mathbf{y}_m\|_2^2 \\ & \text{subject to} && \mathbf{w}_1 = \mathbf{w}_2 \mathbf{B} \mathbf{a}; \quad \mathbf{w}_2 = \bar{\mathbf{E}}; \quad \mathbf{w}_3 = \mathbf{w}_4 \mathbf{a}; \quad \mathbf{w}_4 = \bar{\mathbf{E}}. \end{aligned} \quad (33)$$

Notice that variable  $\mathbf{w}$  introduces the equality constraints in (33). The ADMM then looks for minimizing the augmented Lagrangian function of problem in (33) expressed as

$$\begin{aligned} \hat{\bar{\mathbf{E}}}, \hat{\mathbf{w}}, \hat{\mathbf{h}} \in \underset{\bar{\mathbf{E}}, \mathbf{w}, \mathbf{h}}{\text{argmin}} \quad \mathcal{L}(\bar{\mathbf{E}}, \mathbf{w}, \mathbf{h}) = & \frac{1}{2} \|\mathbf{H}_h \bar{\mathbf{M}} \mathbf{w}_1 - \mathbf{y}_h\|_2^2 + \frac{1}{2} \|\mathbf{H}_m \mathbf{w}_3 - \mathbf{y}_m\|_2^2 + \\ & \frac{\gamma}{2} \|\mathbf{w}_1 - \mathbf{w}_2 \mathbf{B} \mathbf{a} - \mathbf{h}_1\|_2^2 + \frac{\gamma}{2} \|\mathbf{w}_2 - \bar{\mathbf{E}} - \mathbf{h}_2\|_2^2 + \\ & \frac{\gamma}{2} \|\mathbf{w}_3 - \mathbf{w}_4 \mathbf{a} - \mathbf{h}_3\|_2^2 + \frac{\gamma}{2} \|\mathbf{w}_4 - \bar{\mathbf{E}} - \mathbf{h}_4\|_2^2, \end{aligned} \quad (34)$$

with  $\gamma > 0$  being the dual regularization parameter.

Minimization over each variable  $\bar{\mathbf{E}}, \mathbf{w}$ , and  $\mathbf{h}$  in (34) leads to Algorithm 6. There, in lines 4

and 5, the subproblems only involve  $\ell_2$ -norms with closed-form solutions, and in line 6 variable

$$\mathbf{e} = [(\mathbf{w}_2^{x+1} \mathbf{B} \mathbf{a})^T (\bar{\mathbf{E}}^{x+1})^T (\mathbf{w}_4^{x+1} \mathbf{a})^T (\bar{\mathbf{E}}^{x+1})]^T.$$

Algorithm 6. Endmembers Estimation with ADMM approach

---

```

1: procedure ENDE( $\mathbf{a}, \mathbf{H}_h, \mathbf{H}_m, \mathbf{y}_h, \mathbf{y}_m, \mathbf{M}, \mathbf{B}, \gamma, \text{IterE}$ )
2:    $x \leftarrow 0; \mathbf{w}^x \leftarrow \mathbf{0}; \mathbf{h}^x \leftarrow \mathbf{0}$  ▷ Initialization
3:   while  $x < \text{IterE}$  do ▷ Update
4:      $\bar{\mathbf{E}}^{x+1} := \underset{\bar{\mathbf{E}}}{\text{argmin}} \mathcal{L}(\bar{\mathbf{E}}, \mathbf{w}^x, \mathbf{h}^x)$ 
5:      $\mathbf{w}^{x+1} := \underset{\mathbf{w}}{\text{argmin}} \mathcal{L}(\bar{\mathbf{E}}^{x+1}, \mathbf{w}, \mathbf{h}^x)$ 
6:      $\mathbf{h}^{x+1} := \mathbf{h}^x + \mathbf{w}^{x+1} - \mathbf{e}^{x+1}$ 
7:      $x \leftarrow x + 1$ 
8:   return  $\bar{\mathbf{E}}^{\text{IterE}}$ 

```

---

*Endmember estimation algorithm convergence.* The objective function in (33) is the sum of two  $\ell_2$ -norms, thus, minimizing a closed and proper convex function with linear constraints. Therefore, the Algorithm 2 is guaranteed to converge (Boyd et al., 2011).

**3.4.3. General scheme.** The general scheme of the proposed CSI fusion numerical algorithm is depicted in Algorithm 7. In line 3, the initialization considers two scenarios. Scenario 1 is used when the compressive optical system does not disperse the spectral information, so that, the endmembers can be rapidly, and accurately estimated from the HRI compressed measurements  $\mathbf{y}_h$ , using strategy in (Vargas et al., 2018b). In this scenario, the variable Iter can be set to 1, and the variable IterE to 0. Scenario 2 is used otherwise. There, the HRI is rapidly recovered ( $\hat{\mathbf{z}}_h$ ) using any literature algorithm such as the gradient projection for sparse reconstruction (GPSR) (Figueiredo et al., 2007b), employed in this chapter. Then,  $\hat{\mathbf{f}}_h$  is used to initialize the endmembers ( $\bar{\mathbf{E}}^0$ ) using algorithm in (Nascimento and Dias, 2005), based on the component vertex analysis.

Algorithm 7. Compressive Spectral Imaging Fusion Approach

---

1:	<b>procedure</b> CSIF( $\mathbf{H}_h, \mathbf{H}_m, \mathbf{y}_h, \mathbf{y}_m, \mathbf{M}, \mathbf{B}, \mathbf{D}, \mathbf{S}_m, \mu, \tau, \rho, \gamma, \text{Iter}, \text{IterA}, \text{IterE}$ )	
2:	$z \leftarrow 0$	
3:	$\bar{\mathbf{E}}^0 \leftarrow \text{Initialization}(\mathbf{H}_h, \mathbf{y}_h)$	▷ See 3.4.3
4:	<b>while</b> $z < \text{Iter}$ <b>do</b>	▷ Update
5:	$\mathbf{a}^{z+1} \leftarrow \text{AbuE}(\bar{\mathbf{E}}^z, \mathbf{H}_h, \mathbf{H}_m, \mathbf{y}_h, \mathbf{y}_m, \mathbf{M}, \mathbf{B}, \mathbf{D}, \mathbf{S}_m, \mu, \tau, \rho, \text{IterA})$	▷ Algorithm 5
6:	$\bar{\mathbf{E}}^{z+1} \leftarrow \text{EndE}(\mathbf{a}^{z+1}, \mathbf{H}_h, \mathbf{H}_m, \mathbf{y}_h, \mathbf{y}_m, \mathbf{M}, \mathbf{B}, \gamma, \text{IterE})$	▷ Algorithm 6
7:	$z \leftarrow z + 1.$	

---

**High-resolution image estimation algorithm convergence.** Algorithm 7 follows a BCDM alternating over variables  $(\mathbf{a}, \bar{\mathbf{E}})$  to minimize the continuous objective function in (21). Note that, the level set  $\mathbf{X}^0 = \{(\mathbf{a}, \bar{\mathbf{E}}) : s(\mathbf{a}, \bar{\mathbf{E}}) \leq s(\mathbf{a}^0, \bar{\mathbf{E}}^0)\}$ , with  $\bar{\mathbf{E}}^0$  as indicated in 3.4.3, and  $\mathbf{a}^0 = \mathbf{0}$ , is bounded. Further, given the convergence guarantees for each subproblem that optimizes  $\mathbf{a}$ , and  $\bar{\mathbf{E}}$ , the sequence  $\{(\mathbf{a}^z, \bar{\mathbf{E}}^z)\}_{z=1, \dots, \text{Iter}}$  generated by the BCDM using the cyclic rule, asymptotically converges to other element in the set. Thus, it is defined and bounded, and the algorithm converges according to Theorem 4.1. in (Tseng, 2001).

**High-resolution image estimation algorithm computational complexity.** The proposed algorithm computational complexity relies upon the strategy to promote the non-local low-rank prior in the abundances update. Specifically, it depends on the complexity of the third and fourth steps consisting on the grouping of the  $I$  patches into  $G$  patch-groups, and the lower-rank approximation for each patch-group.

For the patch grouping, I employ the k-means method, whose complexity is known to be  $O(I^2)$ . Further, the number of overlapped extracted patches is set to  $I = (N^2 - p + 1)^2$ . Thus, the complexity of grouping similar patches per each iteration is  $O((N^2 - p + 1)^4)$  which in turn, is bounded by  $O(N^4)$ . Notice that, this complexity can be reduced by fixing the indices of the patches that belong

to the patch-groups during some few general iterations.

For calculating the lower-rank approximations in (29), the SVD is employed, whose computational complexity is known to be  $O(u^2v + uv + v^3)$  for a  $u \times v$  matrix. Assuming that all patch-groups contain the same amount of elements, the SVD must be calculated  $G$  times per iteration over a matrix of size  $p^2r \times I_G$ ,  $I_G = I/G$ . Thus, the complexity of this step per iteration is given by

$$O(G((p^2r)^2(I_G) + p^2rI_G + (I_G)^3)), \quad (35)$$

which considering that  $p^2 \ll N^2$ , and  $r$  is a very small number compared to the dimension of the HRI,  $r \ll \min(N^2, L)$ , and  $G$  is a fixed few amount of patch-groups, the complexity is bounded by  $I^3$ , and in consequence by  $N^6$ . Then, the complexity of the proposal is of order  $O(N^6)$ .

### 3.5. Simulations and results

Simulated experiments were carried out to evaluate the proposed CSI fusion approach under diverse experimental setups which could affect the performance of the algorithm.

The compressive optical systems correspond to the C-CASSI and SSCSI described in Section 1.2.

The employed databases consist in the Jasper-ridge, Urban and Pavia University described in Appendix 1, with the modified versions as presented in (Zhu et al., 2014b).

The performance was evaluated with the image fusion error metrics described in Appendix 2: the PSNR, the UIQI, the SAM and the ERGAS.

**3.5.1. Implementation details.** In all cases, the 90% and 50% of the data of the HSI/MSI were acquired using the SSCSI and C-CASSI architectures, where the entries of the sen-

sing matrices were generated following a Bernoulli distribution. The blurring matrix has a standard variation of 1.5, and the added Gaussian noises have 30 and 40 [dB] of SNR level for the HRI and MSI acquisition. The selection for the number of patch-groups  $G$ , size of patches ( $p \times p$ ), and parameters  $\mu$ , and  $\tau$ , playing a significant role was addressed as below.

The number of patch-groups ( $G$ ) was set to correspond to the amount of endmembers from each database. Thus, if  $r = 4$ , then  $G = 4$ . This, based on the intuition that if an HRI only contains  $r$  independent spectral signatures, then grouping the extracted patches according to its spectral response will lead to high self-similar patch-groups.

The patch size ( $p \times p$ ) was selected via cross-validation, in which  $p$  was varied in the range  $p = [3 \ 5 \ 7 \ 9 \ 11]$  for each database and optical system. Figure 11 shows the fusion quality in terms of the PSNR as a function of the patch size. It can be seen that this parameter significantly affects the quality when using SSCSI, where the difference in the quality achieves up to 5 [db]. Unlike, it does not significantly affect the quality when using C-CASSI, where the maximum quality difference is 1 [db]. This, considering that SSCSI preserves more spatial details of the image, such that, variations in the patch size can generate very accurate similar patch-groups.

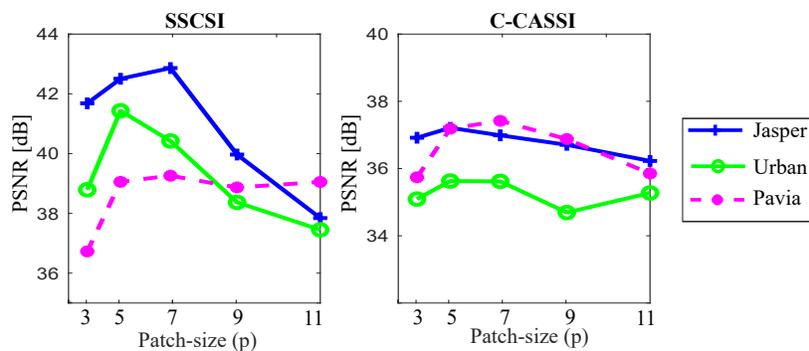


Figure 11. Selection of patch-size experiment.

From this experiment, I selected a patch size of  $p = 5$  which provides good quality results in most of the cases.

The parameter  $\mu$  related to the non-local low-rank regularization term, determining the low-rank approximation of each patch-group, and  $\tau$  related to the total-variation regularization term, determining the smoothness level, were selected via cross-validation in which,  $\mu$  was varied in the interval  $[10^{-6}; 10^{-3}]$  and  $\tau$  was varied in the interval  $[10^{-8}; 10^{-5}]$  for each database and optical system. **Remark:** All experiments set the patch size to  $p = 5$ , the dual parameters to  $\rho = 2.75e^{-6}$  and  $\gamma = 1$ , which were empirically chosen. Figure 12 shows the quality in terms of PSNR when varying the algorithm parameters along the three data bases and the two optical systems. There, it can be observed that the algorithm presents smooth transitions and achieves a peak value (marked with a green asterisk) inside the evaluated intervals.

Table 8 summarizes the experimental setup and chosen algorithm parameters based on the results of this experiment. The Column labeled as **Data** indicates the percentage of compressed data acquired w.r.t. the HRI dimension.

Table 8  
*Algorithm Parameters*

Database	HSI Source			MSI Source			Data [ % ]		Parameters			
	$d_h$	Shots		$d_m$	Shots		SSCSI	C-CASSI	$G$	$p \times p$	$\mu$	$\tau$
		SSCSI	C-CASSI		SSCSI	C-CASSI						
Jasper	4	59	20	11	3	3	10.13	10.46	4	$5 \times 5$	$5e^{-5}$	$1e^{-6}$
Urban	4	73	21	9	5	5	11.81	10.92	6	$5 \times 5$	$5e^{-5}$	$5e^{-7}$
Pavia	4	90	22	10	5	5	10.62	10.98	6	$5 \times 5$	$5e^{-5}$	$5e^{-7}$

Finally, all simulations were carried out using Matlab 2018a in a computer with Intel(R) Core(TM) i7-6700 CPU @ 3.40GHz.

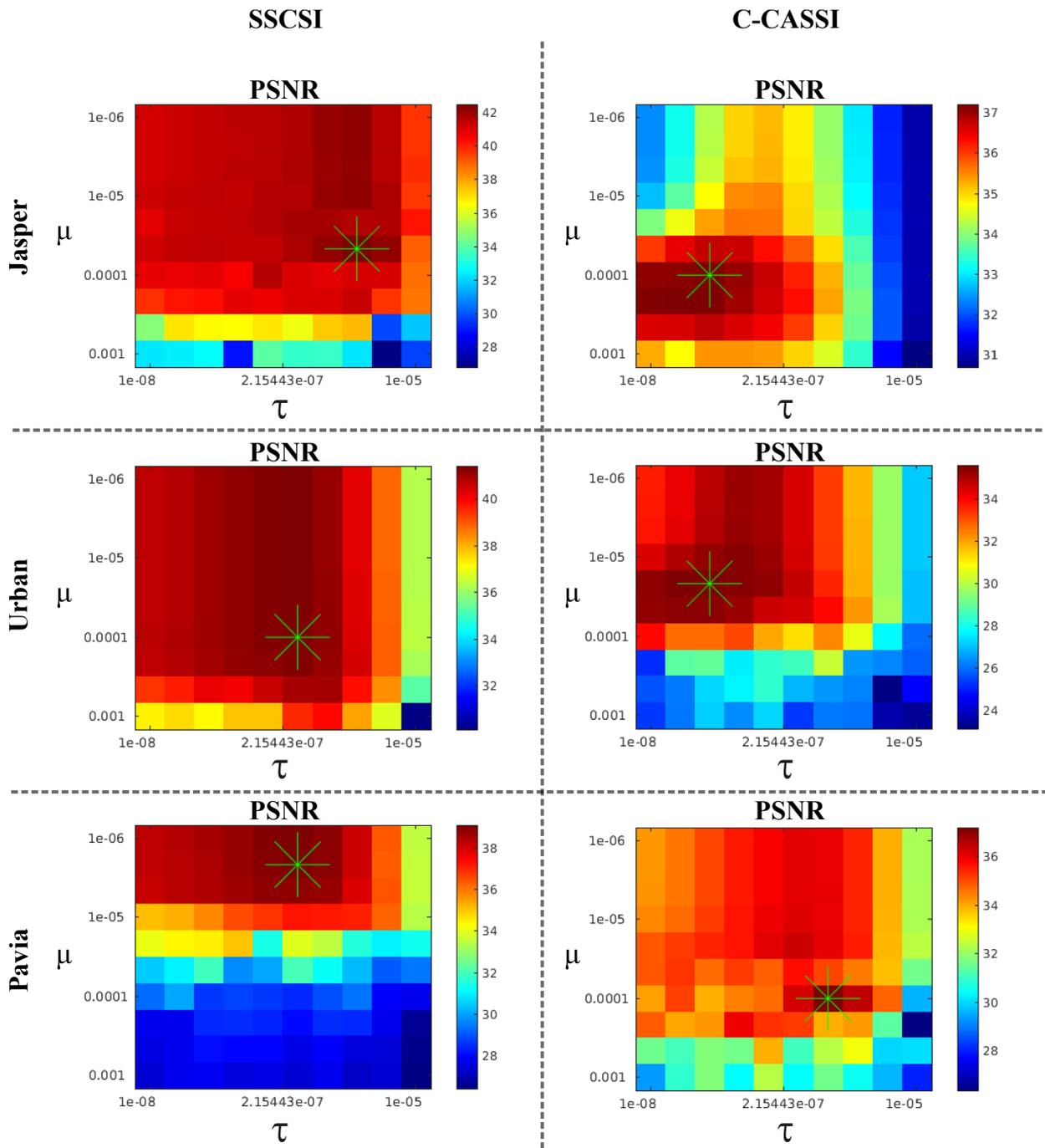


Figure 12. Algorithm parameters experiment.

**3.5.2. Performance evaluation.** The performance of the proposal is compared against the CSI fusion in (Vargas et al., 2019), employing a sparsity prior over the abundances to fuse the compressed measurements. To the best of the author knowledge, this is the most effective method that fuses two sets of compressed measurements. Further, results are compared against the multi-band image fusion approach in (Wei et al., 2016a) using the full HRI and MSI data.

Tables 9, 10, and 11 show the quantitative results in terms of the fusion metrics. Abbreviations **Ours**, **Classic**, and **FUMI** refer to the proposed fusion using the non-local low-rank prior, the literature method using the sparsity prior, and the literature method using full data. The pro-

Table 9

*Fusion metrics for Jasper database*

System	SSCSI		C-CASSI		FUMI
Method	Ours	Classic	Ours	Classic	
PSNR	<b>42.5049</b>	36.3763	<b>37.2109</b>	34.2388	39.9734
UIQUI	<b>0.9898</b>	0.9616	<b>0.9777</b>	0.9607	0.9974
SAM	<b>1.3079</b>	2.4008	<b>2.2177</b>	2.4305	1.7898
ERGAS	<b>2.3070</b>	5.2691	<b>2.6150</b>	2.7907	1.5333
Time[s]	161	<b>16.25</b>	162	<b>23.56</b>	11.9720
Data	10.13 %		10.46 %		100 %

Table 10

*Fusion metrics for Urban database*

System	SSCSI		C-CASSI		FUMI
Method	Ours	Classic	Ours	Classic	
PSNR	<b>41.4325</b>	31.8746	<b>35.6286</b>	35.3664	40.4498
UIQUI	<b>0.9974</b>	0.9748	<b>0.9909</b>	0.9887	0.9972
SAM	<b>1.0554</b>	2.9543	<b>1.8699</b>	2.0026	1.3107
ERGAS	<b>0.6584</b>	2.0968	<b>1.2426</b>	1.3751	0.7668
Time[s]	346	<b>18.45</b>	300	<b>19.65</b>	44.6458
Data	11.81 %		10.92 %		100 %

posal outperforms the **Classic** method, with an average improvement of 4.8 dB in PSNR, and a gain of 0.019 in the UIQUI, while reducing the angle between the reference and estimated spectra

Table 11  
*Fusion metrics for Pavia database*

System	SSCSI		C-CASSI		FUMI
Method	Ours	Classic	Ours	Classic	
PSNR	<b>39.0931</b>	31.2186	<b>37.1879</b>	35.2967	42.3394
UIQI	<b>0.9894</b>	0.9442	<b>0.9827</b>	0.9812	0.9961
SAM	<b>1.1623</b>	8.6189	<b>2.7313</b>	5.6159	2.0523
ERGAS	<b>0.8394</b>	4.6607	<b>1.9365</b>	2.7821	1.0403
Time[s]	407	<b>25.88</b>	417	<b>27.05</b>	31.2606
Data	10.62 %		10.98 %		100 %

in 2.28 of the SAM metric, and a reduction of 1.6 in the ERGAS. In relation to the FUMI method, the proposal obtained comparable performance in terms of the UIQI, SAM and ERGAS metrics, while improving the quality of the fused image in terms of PSNR for the Jasper and Urban datasets. This result demonstrates that using the non-local low-rank effectively restore the structural comprehensive information from the compressed measurements.

Figure 13 shows the spectral response at two spatial positions of each database compared to the proposed and **Classic** estimations along each one of the experimental configurations. There, **F1**, **F2**, and **F3** stands for Jasper, Urban, and Pavia datasets. The zoom versions evidence the visual similarity improvement in the estimated spectrum of the fused image. Figure 14 shows a qualitative RGB representations comparison of the fused HRI along each approach, where the fused HRI obtained with the non-local low-rank prior is visually closer to the original one.

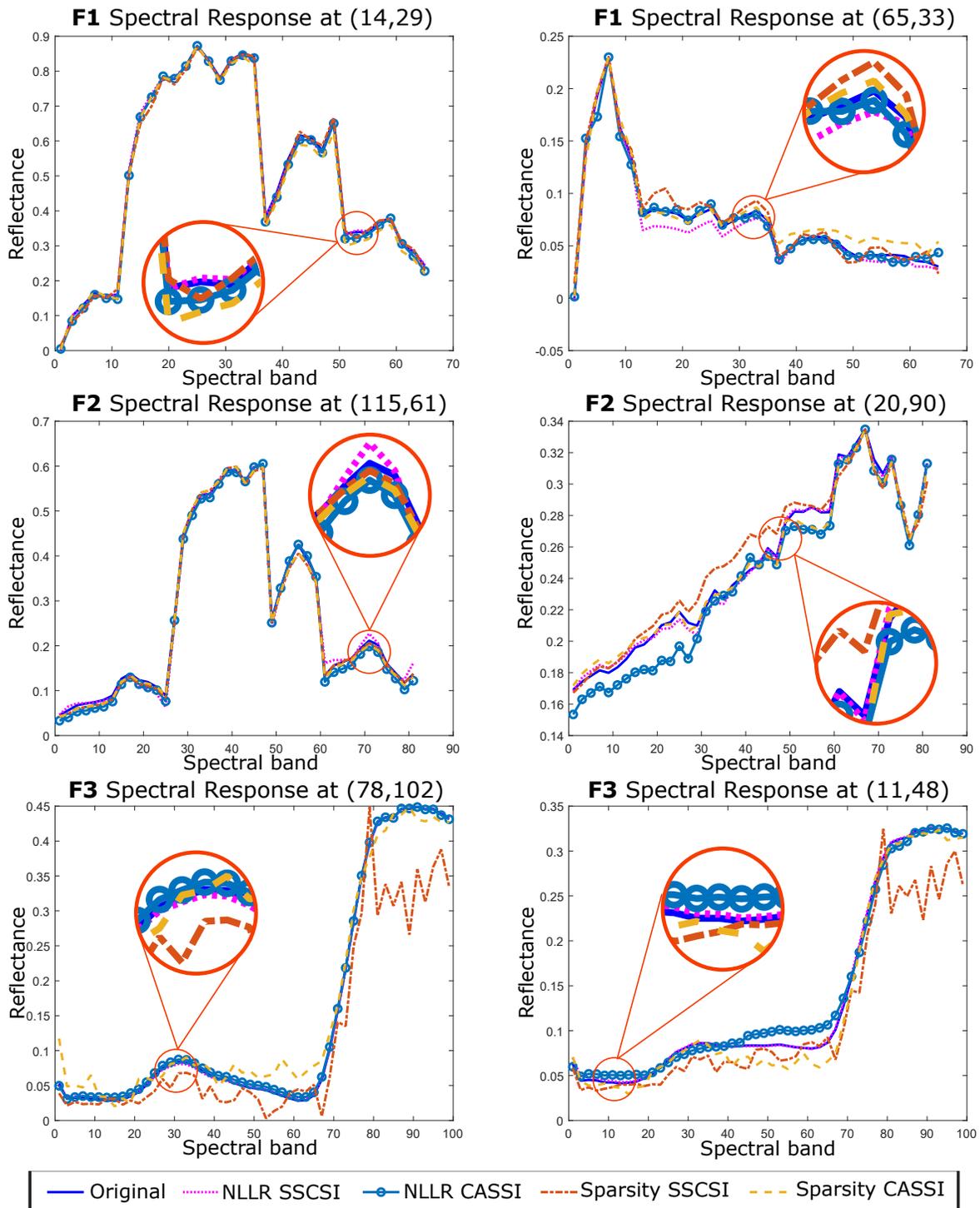


Figure 13. Visual comparison of the spectral reconstructions.

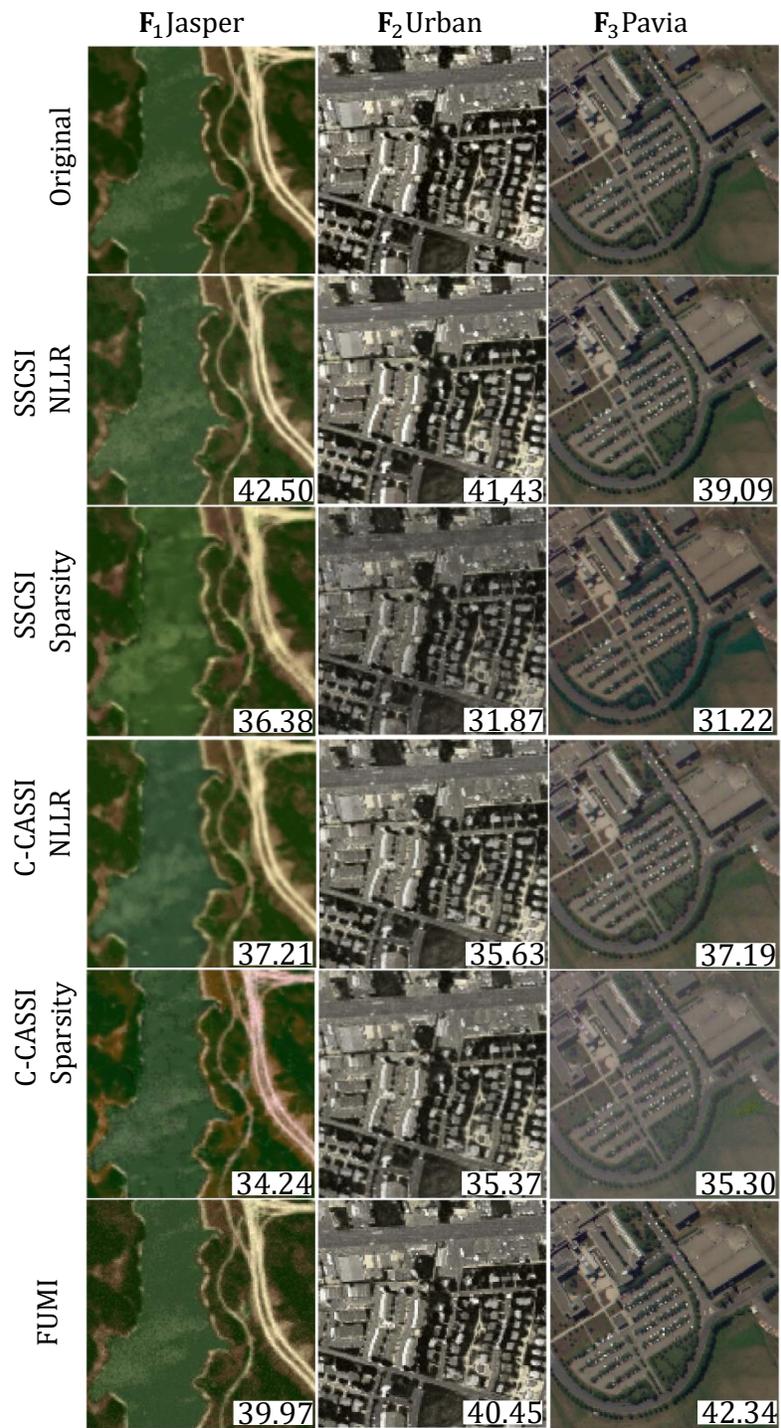


Figure 14. RGB representation visual comparison.

### 3.6. Conclusions

This chapter studied the SI low-rankness from the global and non-local points of view, considering the spectral and spatial correlations for solving the CSI fusion problem. Specifically, the global spectral low-rank was analyzed by a theoretical premise using the LMM, where each pixel was modeled as a linear combination of a few pure spectral responses. The non-local low-rank prior was validated through an extensive empirical study of the SVD decay obtained from matrices stacking similar patches located at non-local positions. The exhibited fast SVD decay supported incorporating the non-local low-rank assumption by a hand-crafted explicit regularization function. The formulated CSI fusion was solved following a BCDM, iteratively updating the abundances and the endmembers. In particular, the abundances were assumed to preserve the non-local low-rank and smoothness priors, which reduced the computational complexity against assuming the low-rankness over the complete image. Simulations over three datasets and the CASSI and SSCSI compressive optical systems showed that the proposal outperformed the CSI fusion state-of-the-art method in up to 2dBs and 8dBs for the CASSI and SSCSI systems. The proposal obtained a comparable spectral quality against methods using complete data, with a significant improvement in the spatial quality.

#### **4. Joint non-local, spectral, and similarity low-rank priors for spectral imaging fusion**

*Part of this chapter has been adapted from the journal paper entitled Joint Nonlocal, Spectral, and Similarity Low-Rank Priors for Hyperspectral-Multispectral Image Fusion Submitted to IEEE Transactions on Geoscience and Remote Sensing.*

The previous chapter adopted the low-rank prior for fusing two sets of compressed measurements. Unlike, this chapter addresses the fusion of two full-dimensional SIs, where I propose an HSI-MSI fusion that jointly promotes various low-rank regularizations. The method synthesizes the HRI through the PnP of a non-local patch-based denoiser in the ADMM. The HRI correlations are considered via the non-local self-similarity, the spectral low-rank, and an introduced rank-one similarity prior. Section 4.1 presents the most related works, where the separate use of low-rank priors in traditional methods motivates the proposal that leads to the following main contributions.

1. The introduction of the concept of rank-one similarity prior, validated through an extensive empirical study can be found in Section 4.2.
2. The joint of the non-local self-similarity, spectral low-rank, and rank-one similarity priors with an implicit regularization for HSI-MSI fusion are described in Section 4.3.
3. The development of a practical algorithm based on PnP-ADMM to solve the fusion problem via the multichannel block-matching and 3D filtering (BM3D) denoising in Section 4.4.
4. A significant improvement in the visual and objective metrics, particularly in the spectral PSNR, with excellent recovery of spectral signatures at low-contrast regions in Section 4.5.

5. The acquisition of two hyperspectral datasets at the Optics Laboratory from the Universidad Industrial de Santander as described in 4.5.1, which will be publicly available to the academic community to be used in the SI recovery research field.

#### 4.1. Hyperspectral-multispectral image fusion related works

Hyperspectral imaging sensors face with a trade-off between the spatial resolution and the signal-to-noise ratio due to limited amount of incident energy (Guilloteau et al., 2020). Hence, HSI-MSI fusion is a computational technique that aims to combine the valuable information of two images obtained from different sensors. Precisely, an HRI is synthesized from a low-spatial-and-high-spectral resolution HSI, and a high-spatial-and-low-spectral resolution MSI (Piella, 2003). Fused information adds robustness and eases subsequent processing in various remote sensing applications (Bioucas-Dias et al., 2013; Li et al., 2019a; Fan et al., 2018).

Particularly, in variational non-blind methods, the point spread function of the spectral imaging cameras is assumed to be known, and the HSI and MSI are modeled as blur and downsampling linear operations of the HRI. Existing variational HSI-MSI fusion methods take advantage of the redundant information of the HRI with prior information as sparsity, smoothness and low-rankness to regularize the resulting ill-posed inverse problem. Some remarkable frameworks taking advantage of such priors include **sparse representation**, modeling the HRI as a linear combination of few atoms in a redundant dictionary (Akhtar et al., 2014; Wei et al., 2015; Dong et al., 2016; Fang et al., 2018; Han et al., 2018; Jian et al., 2020); **Bayesian formulations**, using the observation model to build appropriate posterior distributions (Simoes et al., 2014; Akhtar et al., 2015, 2016; Wei et al., 2016b; Xue et al., 2017a; Lin et al., 2017; Sui et al., 2019); **matrix factorization**, de-

composing the HRI as the linear combination of few endmembers with specific abundances (Wei et al., 2016a; Zhou et al., 2017; Wu et al., 2020b; Liu et al., 2020; Wu et al., 2020a; Yi et al., 2020; Zhou et al., 2020); and **tensor decomposition**, extending the matrix factorization methods to a 3D modeling (Li et al., 2018a, 2019b; Xu et al., 2019; Ma et al., 2020; Dian and Li, 2019; Li et al., 2020b; Wang et al., 2020a).

Among above mentioned priors, the low-rankness taking advantage of the structural local and non-local self-similarities present in the HRI has been demonstrated to significantly improve the recovery quality in general spectral imaging reconstruction (Wang et al., 2020a). Therefore, this chapter focus on the employment of low-rank priors considering the spectral redundancies, and the structural non-local self-similarities. In fact, recent work in (Chang et al., 2020a) established that the low-rank correlations along spatial, spectral, and non-local self-similarity dimensions contribute different in the recovery quality of an HRI, where the joint of non-local self-similarity and spectral low-rank priors is the most beneficial to consider the low-rankness of the HRI. However, with exception of (Wang et al., 2020a; Li et al., 2020b) that combine non-local tensor decomposition with spectral unmixing and sparse representation, respectively. Previous HSI-MSI fusion methods considering low-rankness use the low-rank priors separately. In addition, there is still a gap in (Wang et al., 2020a; Li et al., 2020b) where they neglect the existence and the contribution of a low-rank prior in the similarity dimension.

Motivated by such gaps, I present an HSI-MSI fusion that unlike previous methods synergistically combines the nonlocal self-similarity, the spectral low-rank, and an introduced low-rank similarity prior, dubbed *rank-one similarity*, to recover the HRI. This concept is introduced upon

the idea that intrinsic nonlocal spatial structures occur across the entire spectral dimension.

## 4.2. Rank-one similarity prior

The geometry of the objects in an HRI is typically irrespective of the object reflectance and responsivity at different spectral bands. Therefore, I introduce the following rank-one similarity prior to formalize an intrinsic HRI characteristic under which all spectral bands share common structural similarities of small spatial blocks.

**4.2.1. Concept.** Let  $\mathbf{f} \in \mathbb{R}^{N^2L}$  represent the vector form of an HRI with  $N^2$  spatial pixels and  $L$  spectral bands, and let  $\mathbf{P}_{\lambda,i} \in \mathbb{R}^{p \times p}$  denote a  $p \times p$  HRI block whose top-left corner is at the  $i^{\text{th}}$  spatial position,  $i = 1, \dots, N^2$ , within the  $\lambda^{\text{th}}$  spectral band,  $\lambda = 1, \dots, L$ .

A dissimilarity map is a 2D array containing the dissimilarity measure of a *reference* block to its neighboring blocks. Let  $\mathbf{D}_{\lambda,i} \in \mathbb{R}^{w \times w}$  denote the dissimilarity map for the  $\mathbf{P}_{\lambda,i}$  reference block, where  $w \in \mathbb{Z}_+$  is the neighborhood size for searching similar blocks and whose entries are calculated in terms of the  $\ell_2$ -norm as

$$\mathbf{D}_{\lambda,i}(j) = \|\mathbf{P}_{\lambda,i} - \mathbf{P}_{\lambda,j}\|_2^2, \quad (36)$$

for  $j \subseteq \Omega_i$ , where  $\Omega_i$  is a set containing the indexes of the blocks within the  $w \times w$  neighbourhood of the reference block.

I further denote  $\tilde{\mathbf{D}}_i \in \mathbb{R}^{w^2 \times L}$  the matrix whose columns contain the normalized and vectorized dissimilarity maps for reference blocks at the same position index  $i$  over all spectral bands

$$\tilde{\mathbf{D}}_i = [\text{vec}(\mathbf{D}_{1,i})/s(\mathbf{D}_{1,i}), \dots, \text{vec}(\mathbf{D}_{L,i})/s(\mathbf{D}_{L,i})], \quad (37)$$

where  $s(\mathbf{D}_{\lambda,i})$  denotes the sample standard deviation of  $\mathbf{D}_{\lambda,i}$ .

The *rank-one similarity prior* assumes that, for a high-contrast reference block at position index  $i$ , the dissimilarity maps differ across the spectral bands only by a proportionality factor, i.e.  $\mathbf{D}_{\lambda,i} = \alpha_{\lambda,i} \mathbf{D}_{1,i} \forall \lambda$ . Adopting this prior in practice means approximating  $\tilde{\mathbf{D}}_i$  (37) by a rank-one matrix. While alternative formulations are possible (e.g., adopting different norms for (36), or an histogram equalization of each  $\mathbf{D}_{\lambda,i}(j)$  prior to forming (37)), for the sake of simplicity I adopt the above basic definitions throughout this chapter.

**4.2.2. Concept test.** The rank-one similarity prior is validated through an extensive empirical study that analyzes the singular values of the matrix (37), whose decay is expected to be fast, especially from the first to the second singular value on high-contrast blocks (37). Figure 15 shows a high-contrast (top) and a low-contrast (bottom) reference block of size  $p \times p = 8 \times 8$  from the Colombia dataset acquired in the optics laboratory of Universidad Industrial de Santander (Gelvez-Barrera, 2021). A red square bounds each reference block into the RGB version of the search neighborhood employed to generate a dissimilarity map of size  $39 \times 39$  across the first principal component (PC) and five spectral bands. The dissimilarity maps  $\mathbf{D}_{\lambda,i}, \forall \lambda$ , (36) appear to be all proportional to each other so that  $\tilde{\mathbf{D}}_i$  (37) exhibits a fast singular value decay. The plots at right show the singular-value decays when varying the dissimilarity map size from  $19 \times 19$  to  $239 \times 239$ , while maintaining the reference block size fixed to  $8 \times 8$ . Observe that the decays for the high-contrast reference block are faster than for the low-contrast reference block and that the decays are faster for smaller dissimilarity map size. The fast decay from the first to the second singular value indicates that  $\tilde{\mathbf{D}}_i$  can be approximated as a rank-one matrix.

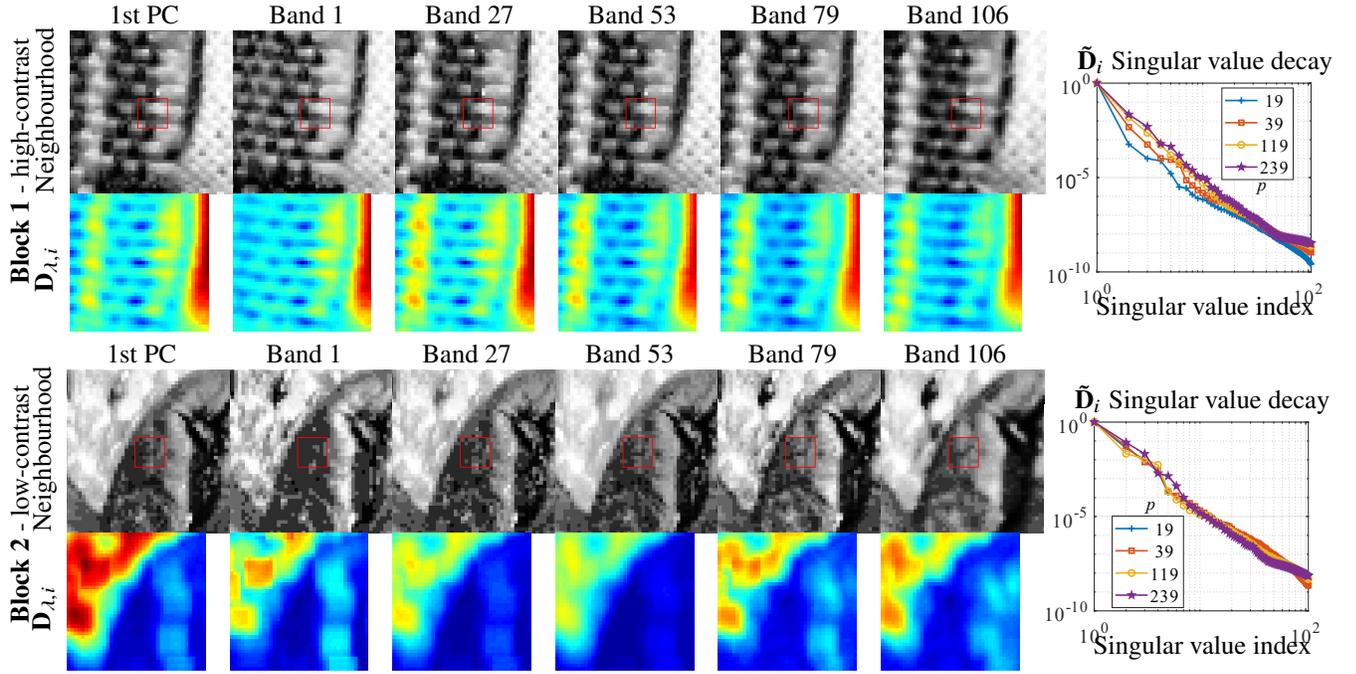


Figure 15. Illustration of the rank-one similarity prior in the Colombia dataset.

The complete empirical study follows the procedure below over ten datasets covering real-world objects, remote sensing, and satellite sources (Yasuma et al., 2008; M Graña, 2008).

1. Calculate the reference block contrast  $\mathcal{C}(\mathbf{P}_{\lambda,i})$  as

$$\mathcal{C}(\mathbf{P}_{\lambda,i}) = \max(\mathbf{P}_{\lambda,i}) - \min(\mathbf{P}_{\lambda,i}). \quad (38)$$

2. Calculate the singular values of  $\tilde{\mathbf{D}}_i$ ,  $\forall i$ .
3. Analyze the singular-value decay, where high-contrast reference blocks are expected to yield a fast decay, i.e. large  $\mathcal{C}(\mathbf{P}_{\lambda,i})$  leads to fast decay of the singular values of  $\tilde{\mathbf{D}}_i$ .

I primarily focus on the case where the block size is  $p \times p = 8 \times 8$  and the neighborhood size

is  $w \times w = 39 \times 39$ , as in (Dabov et al., 2007).

Figure 16 illustrates the singular value decay for all possible reference blocks  $\tilde{\mathbf{D}}_i, \forall i$  (37) from twelve publicly available datasets. I define a high-contrast reference block as a reference block whose contrast is higher than the half of the contrast's range for every analyzed dataset; notice that other definitions are also possible.

The decays are grouped in quantiles ([0%-1%), [1%-5%), [5%-20%), [20%-50%), [50%-80%), [80%-95%), [95%-99%), [99%-100%]) and each singular value is colored according to the maximum contrast of the values that belong to the corresponding quantile; red color indicates high-contrast blocks, and blue color indicates low-contrast blocks. The color histogram on the right-hand side of each plot shows the distribution of the reference blocks according to their contrast; one can observe that in a typical natural image, the most frequent blocks have a relatively low-contrast (Cupitre being an exception, with a nearly uniform histogram). The plots show that only a very slim band of upper quantiles of the singular values (slowest decay) might feature only low-contrast blocks: higher-contrast blocks always have fast decay of the associated singular values and must appear at bottom.

The expected behavior can be observed in the figure, where high-contrast blocks (on average 16.47% of all blocks) lead to a fast decay, i.e. red color must appear at the bottom and must be absent at the top of plots. In particular, the second singular value is on average  $3.2 \times 10^{-3}$  times the first one, i.e., a decay from the first one of 99.7% within the first quartile,  $5.9 \times 10^{-3}$  times the first one, i.e., a 99.4% drop for the first lower half of its distribution. The rate of decay continues through the further singular values, as can be seen in the figure, supporting the incorporation of the

rank-one similarity prior to the solution of spectral imaging problems. For instance, for the HSI-MSI fusion, the rank-one similarity prior is helpful to effectively propagate the detailed spatial information presented in the MSI across the narrow spectral bands presented in the HSI.

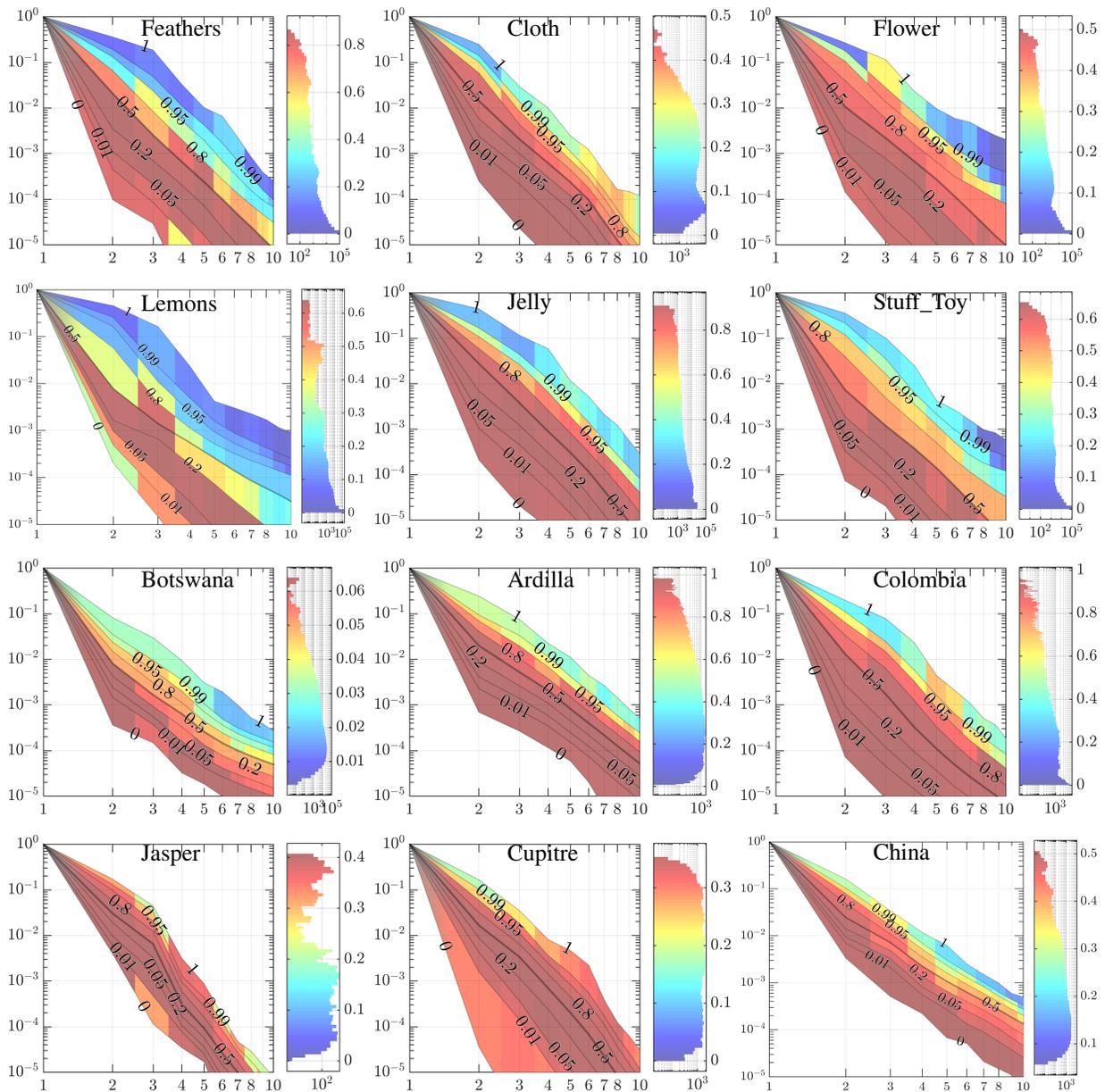


Figure 16. Empirical validation of the rank-one similarity prior.

### 4.3. Proposed hyperspectral-multispectral image fusion

The proposed HSI-MSI fusion follows a subspace-based formulation and a PnP-ADMM algorithm to estimate the HRI from the HSI and MSI degraded observations according to the following linear forward model introduced in 1.1

$$\begin{aligned}\mathbf{f}_h &= \bar{\mathbf{M}}\bar{\mathbf{B}}\mathbf{f} + \boldsymbol{\omega}_h \\ \mathbf{f}_m &= \bar{\mathbf{R}}\mathbf{f} + \boldsymbol{\omega}_m.\end{aligned}\tag{39}$$

The proposed cost function comprises two  $\ell_2$ -norms maintaining the data fidelity to the observations and one implicit regularization function  $\phi : \mathbb{R}^{N^2L} \rightarrow \mathbb{R}$  that aims to promote *jointly* the spectral low-rank, the nonlocal self-similarities, and the rank-one similarity prior, taking full advantage of the intrinsic spatial-spectral structural correlations. The HSI-MSI fusion is thus an inverse problem formulated as

$$\underset{\mathbf{f} \in \mathbb{R}^{N^2L}}{\text{minimize}} \quad f(\mathbf{f}) = \frac{1}{2} \|\mathbf{f}_h - \bar{\mathbf{M}}\bar{\mathbf{B}}\mathbf{f}\|_2^2 + \frac{\mu}{2} \|\mathbf{f}_m - \bar{\mathbf{R}}\mathbf{f}\|_2^2 + \lambda\phi(\mathbf{f}),\tag{40}$$

where  $\mu > 0$  and  $\lambda > 0$  correspond to the regularization parameters that balance the three terms.

### 4.4. PnP-ADMM algorithm

The proposed algorithm to solve (40) follows the established PnP-ADMM framework (Chan et al., 2016), promoting the low-rank priors through a denoiser based on nonlocal regularization (Dabov et al., 2007).

The PnP-ADMM strategy begins with the introduction of three auxiliary variables  $\mathbf{v}_i$ , for

$l = 1, 2, 3$ , upon which (40) becomes

$$\begin{aligned} & \underset{\mathbf{f}, \mathbf{v}_l}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{f}_h - \bar{\mathbf{M}}\mathbf{v}_1\|_2^2 + \frac{\mu}{2} \|\mathbf{f}_m - \bar{\mathbf{R}}\mathbf{v}_2\|_2^2 + \lambda\phi(\mathbf{v}_3), \\ & \text{subject to} \quad \mathbf{v}_1 = \bar{\mathbf{B}}\mathbf{f}; \quad \mathbf{v}_2 = \mathbf{f}; \quad \mathbf{v}_3 = \mathbf{f}, \end{aligned} \quad (41)$$

whose augmented Lagrangian optimization problem with dual variables  $\mathbf{g}_l$ , for  $l = 1, 2, 3$  is given by

$$\begin{aligned} \underset{\mathbf{f}, \mathbf{v}_l, \mathbf{g}_l}{\text{minimize}} \quad \mathcal{L}(\mathbf{f}, \mathbf{v}_l, \mathbf{g}_l) = & \frac{1}{2} \|\mathbf{f}_h - \bar{\mathbf{M}}\mathbf{v}_1\|_2^2 + \frac{\mu}{2} \|\mathbf{f}_m - \bar{\mathbf{R}}\mathbf{v}_2\|_2^2 + \lambda\phi(\mathbf{v}_3) + \\ & \frac{\rho}{2} \|\mathbf{v}_1 - \bar{\mathbf{B}}\mathbf{f} + \mathbf{g}_1\|_2^2 + \frac{\rho}{2} \|\mathbf{v}_2 - \mathbf{f} + \mathbf{g}_2\|_2^2 + \frac{\rho}{2} \|\mathbf{v}_3 - \mathbf{f} + \mathbf{g}_3\|_2^2, \end{aligned} \quad (42)$$

where  $\rho > 0$  is the dual regularization parameter.

Each primal variable is optimized by solving the following iterative alternating process summarized in Algorithm 8 with variables calculated as detailed in Table 12.

$$\begin{aligned} \mathbf{f}^{k+1} & \in \underset{\mathbf{f}}{\text{argmin}} \quad \mathcal{L}(\mathbf{f}, \mathbf{v}_l^k, \mathbf{g}_l^k). \\ \mathbf{v}_l^{k+1} & \in \underset{\mathbf{v}_l}{\text{argmin}} \quad \mathcal{L}(\mathbf{f}^{k+1}, \mathbf{v}_l, \mathbf{g}_l^k). \end{aligned} \quad (43)$$

**4.4.1. Initialization.** In Algorithm 8 line 3, the estimated HRI and the auxiliary and dual variables are initialized as

$$\mathbf{f}^0 \leftarrow \frac{1}{2} (\bar{\mathbf{B}}^T \bar{\mathbf{M}}^T \mathbf{f} + \bar{\mathbf{R}}^T \mathbf{f}); \quad \mathbf{v}_1^0 \leftarrow \bar{\mathbf{B}}\mathbf{f}^0; \quad \mathbf{v}_2^0 \leftarrow \mathbf{f}^0; \quad \mathbf{v}_3^0 \leftarrow \mathbf{f}^0. \quad (44)$$

**4.4.2. Precomputation.** In Algorithm 8 line 4 the matrices that will be used to invert the optimization sub-problems are precomputed as detailed in Table 12.

Algorithm 8. PnP-ADMM HSI-MSI fusion

---

<b>1:</b>	<b>procedure</b> R1BM3D( $\mathbf{f}_h, \mathbf{f}_m, \bar{\mathbf{M}}, \bar{\mathbf{B}}, \mu, \tau, \rho, iters$ )	
<b>2:</b>	$k \leftarrow 0$	
<b>3:</b>	$(\mathbf{f}^k, \mathbf{v}_l^k, \mathbf{g}_l^k) \leftarrow (\text{As in 44}).$	▷ Initialization
<b>4:</b>	$\mathbf{W}_k \leftarrow \text{Table 12}$	▷ Precomputation
<b>5:</b>	<b>while</b> $k < iters$ <b>do</b>	▷ Update
<b>6:</b>	$\mathbf{f}^{k+1} \leftarrow \mathbf{W}_0^{-1} \mathbf{d}_0^k$	
<b>7:</b>	$\mathbf{v}_1^{k+1} \leftarrow \mathbf{W}_1^{-1} \mathbf{d}_1^k$	
<b>8:</b>	$\mathbf{v}_2^{k+1} \leftarrow \mathbf{W}_2^{-1} \mathbf{d}_2^k$	
<b>9:</b>	$\mathbf{v}_3^{k+1} \leftarrow \text{BM3D}_{\tau/\rho}(\mathbf{f}^{k+1} - \mathbf{g}_3^k)$	▷ PnP BM3D
<b>10:</b>	$\mathbf{g}_1^{k+1} \leftarrow \mathbf{v}_1^{k+1} - \bar{\mathbf{B}}\mathbf{f}^{k+1} + \mathbf{g}_1^k.$	
<b>11:</b>	$\mathbf{g}_2^{k+1} \leftarrow \mathbf{v}_2^{k+1} - \mathbf{f}^{k+1} + \mathbf{g}_2^k.$	
<b>12:</b>	$\mathbf{g}_3^{k+1} \leftarrow \mathbf{v}_3^{k+1} - \mathbf{f}^{k+1} + \mathbf{g}_3^k.$	
<b>13:</b>	$k \leftarrow k + 1$	
<b>14:</b>	<b>return</b> $\mathbf{f}^{iters}$	

---

Table 12

Summary of variables computation

$\mathbf{W}_0$	$\in \mathbb{R}^{N^2L \times N^2L}$	$= \bar{\mathbf{B}}^T \bar{\mathbf{B}} + 2\mathbf{I}_{N^2L}.$
$\mathbf{W}_1$	$\in \mathbb{R}^{N^2L \times N^2L}$	$= \bar{\mathbf{M}}\bar{\mathbf{M}}^T + \rho\mathbf{I}_{N^2L}.$
$\mathbf{W}_2$	$\in \mathbb{R}^{N^2L \times N^2L}$	$= \mu\bar{\mathbf{R}}^T\bar{\mathbf{R}} + \rho\mathbf{I}_{N^2L}.$
$\mathbf{d}_0^k$	$\in \mathbb{R}^{N^2L}$	$= \bar{\mathbf{B}}^T(\mathbf{v}_1^k + \mathbf{g}_1^k) + (\mathbf{v}_2^k + \mathbf{g}_2^k) + (\mathbf{v}_3^k + \mathbf{g}_3^k).$
$\mathbf{d}_1^k$	$\in \mathbb{R}^{N^2L}$	$= \bar{\mathbf{M}}^T\mathbf{f}_h + \rho(\bar{\mathbf{B}}\mathbf{f}^{k+1} - \mathbf{g}_1^k).$
$\mathbf{d}_2^k$	$\in \mathbb{R}^{N^2L}$	$= \mu\bar{\mathbf{R}}^T\mathbf{f}_m + \rho(\mathbf{f}^{k+1} - \mathbf{g}_2^k).$

**4.4.3. Plug-and-Play of the Multichannel BM3D.** In Algorithm 8 line 9, the multichannel BM3D is included to take advantage of the redundant information and intrinsic HRI

structure entailing important considerations detailed below.

In practice, the rank-one similarity prior is verified because the structural features that determine the nonlocal block similarity are *repeated* across the different spectral bands. However, over real-world measurements, the structural features may be buried under noise and be distorted by the coarse sampling, depending on the particular band. Therefore, it is challenging to obtain a reliable dissimilarity map to build the block-matching by working over individual bands.

When applying the principal component analysis (PCA) to the wavelengths, the structural features typically end up being represented with a substantially higher signal-to-noise ratio (SNR) in the first principal component (PC). Therefore, it is convenient to operate the block-matching over the first PC; otherwise, noise can disrupt the block-matching as illustrated in Fig. 17. There, the block-matching is executed in the first PC for a low-contrast and a high-contrast reference block that are bounded and zoomed in dashed magenta and cyan windows, respectively. The reference blocks are bounded with a red square and its corresponding matched similar blocks are bounded with a green square. The matched similar blocks found in the first PC that are superimposed over the 3<sup>th</sup> and 7<sup>th</sup> PCs result well suited for other few PCs, except for the noisy case where the last components are mainly driven by noise; therefore, the block-matching found in the first PC can be propagated to filter only other PCs that are considered significant. As the structural features drive the block-matching on the first PC, it can be used by the rank-one prior for all spectral bands.

In principle, the BM3D can be applied separately on each band using the common block-matching inherited from the first PC. However, the spectral decorrelation provided by the PCA is also beneficial to further sparsify the 3D spectrum of a group of blocks within each spectral band

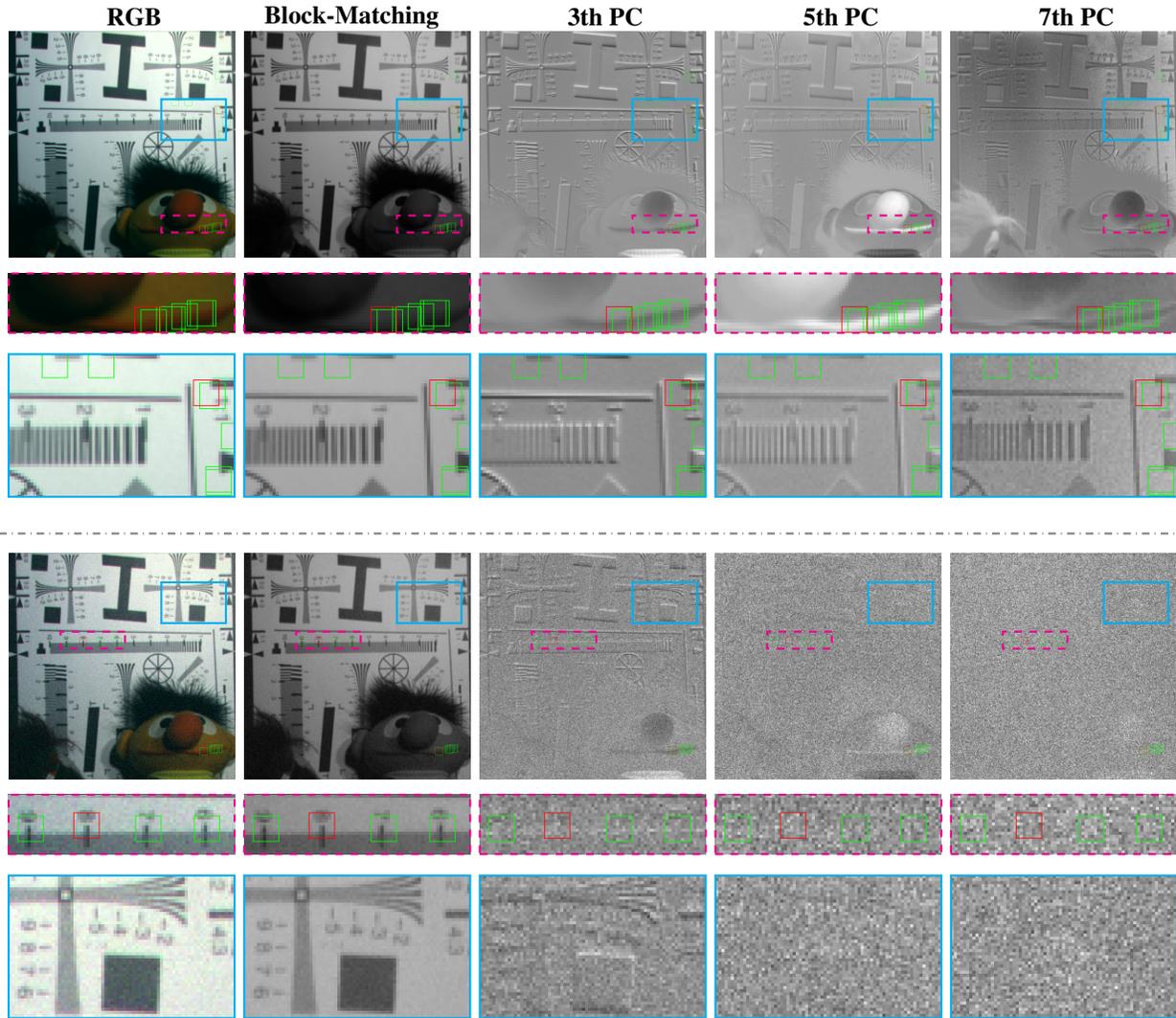


Figure 17. Propagation of the block-matching over the PCs.

in the multichannel BM3D, which is therefore applied upon PCA.

In Algorithm 8 line 9, the update of  $\mathbf{v}_3$  assumes that  $\mathbf{q}_3^k = \mathbf{f}^{k+1} - \mathbf{g}_3^k$  is a noisy version of  $\mathbf{v}_3$  whose effective noise is considered to be white Gaussian with standard deviation given by  $\sigma = \sqrt{\lambda/\rho}$ . Therefore, a denoised version of  $\mathbf{v}_3$  is obtained by the PnP of the multichannel BM3D, where the block-matching step is driven over the first PC as follows.

1. Calculate the PCs of  $\mathbf{q}_3^k$  organized as the matrix  $\mathbf{Q}_3^k \in \mathbb{R}^{L \times N^2}$ .
2. Find the mutual similar blocks by applying the block-matching in the first PC.
3. Propagate the positions of the found similar blocks to filter and preserve a small number  $r \ll L$  of PCs, promoting the rank-one similarity, the spectral low-rank, and the nonlocal self-similarities *jointly*. The amount  $r$  of preserved PCs is updated at each iteration by using the hyperspectral signal subspace identification by minimum error (HySime) method (Bioucas-Dias and Nascimento, 2008).

Notice that the procedure above remarkably reduces the algorithm's computational complexity, avoiding the calculation of multiple block-matching at each PC or at each band. Furthermore, the first PC has the highest inter-pixel variance with well-defined blocks so that the filtering is conservative, and it is expected not to match together blocks that are dissimilar in other PCs, which are typically smoother. Nonetheless, in the rare but possible event where the block-matching on the first PC creates a group of dissimilar blocks in another PC, multichannel BM3D does not break down because the noise attenuation is done after shrinkage of the 3D transform of the group. Finally, matching significantly different blocks (e.g., above the noise level that exists in that PC) will elicit 3D-transform spectrum coefficients larger than the shrinkage threshold. Hence, these coefficients will be preserved, and the blocks' differences will be preserved after the inversion of the 3D transform.

#### 4.5. Simulations and results

To evaluate the proposed R1BM3D HSI-MSI fusion, I conducted extensive experiments across various publicly available multispectral and hyperspectral datasets described in Appendix 1. Furthermore, I acquired two hyperspectral datasets (Ardilla and Colombia) at the Optics Laboratory from the Universidad Industrial de Santander as described in the following section.

**4.5.1. Hyperspectral dataset acquisition.** The testbed prototype shown in Fig. 18 uses an objective lens and the Allied Vision Technologies (AVT) Marlin F-145B2 IRF 15fps IEEE1394 Digital Camera. The light Source consists on the TLS-300XR illuminator (Tunable 300 W Ozone Free Xe, RS232/GPIB) with  $0.3 - 1.8\mu\text{m}$  of spectral range and two fixed slits of  $600\mu\text{m}$  widths located at the input and output of the monochromator.<sup>1</sup>

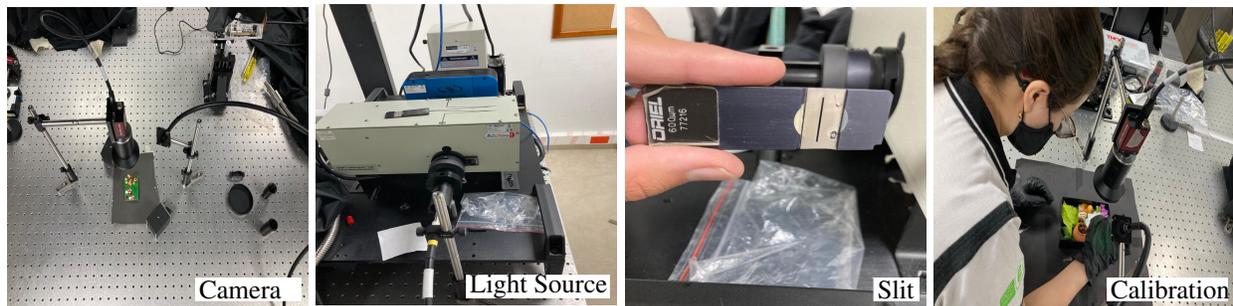


Figure 18. Hyperspectral dataset acquisition system

The Ardilla and Colombia datasets of  $776 \times 1032 \times 121$  spatial-spectral resolution were acquired in the  $0.37 - 0.77\mu\text{m}$  spectral range in intervals of  $0.01\mu\text{m}$  in the ( $0.37\mu\text{m}$  to  $0.45\mu\text{m}$ )

---

<sup>1</sup> I gratefully thanks to the laboratory workers Ph.D.(c) Hans García, Ph.D.(c) Andrés Jeréz, and Ph.D.(c) Miguel Marquez for their contribution in the acquisition setup implementation and description.

and  $(0.65\mu\text{m}$  to  $0.77\mu\text{m})$  spectral ranges, and in intervals of  $0.002\mu\text{m}$  in the  $(0.45\mu\text{m}$  to  $0.65\mu\text{m})$  spectral range. A down-sampled version of size  $388 \times 516 \times 121$  of each dataset is employed to reduce the computational complexity. An RGB image of the scene acquired with a smartphone and the RGB mapping representation from the acquired HSIs are depicted in Fig. 19



Figure 19. RGB mapping of the acquired datasets.

I compare R1BM3D against various state-of-the-art approaches that employ different strategies to obtain the fused image. Specifically, against the subspace based hyperspectral super-resolution (HySure) (Simoes et al., 2014); the HSI-MSI fusion based on Bayesian sparse representation (BSR) (Wei et al., 2015); the non-negative structured sparse representation (NSSR) (Dong et al., 2016); the couple sparse tensor factorization (CSTF) (Li et al., 2018a); and the clustering manifold structure (CMS) (Zhang et al., 2018a). All of comparison methods were implemented by their published code found in (Junjun-Jiang, 2020).

The main parameters of R1BM3D and each comparison method have been tuned separately by trial and error for each dataset to optimize the PSNR. The details are given in Section 4.5.3.

The fusion performance is evaluated in terms of the global, spatial and spectral peak signal-to-noise ratio (PSNR) measured in decibels (dB), the universal image quality index (UIQI), the

spectral angle mapper (SAM) measured in degrees, and the dimensionless global relative error of synthesis (ERGAS) in 2 for 8-bit representations of original and estimated images, i.e., I scale the images to the range [0, 255].

**4.5.2. Comparative benchmarking.** Table 13 summarizes the objective results averaged over five replications changing the random noise for each case. It can be observed that R1BM3D outperforms previous methods by up to 3 dBs in the global PSNR, where the improvement mainly comes from the gain in the spectral domain as shown with the spectral PSNR.

Table 13

*HSI-MSI Fusion Metrics: RMSE, ERGAS, SAM (degrees), UIQI, PSNR (dB).*

Pavia University							
Methods	RMSE ↓	ERGAS ↓	SAM ↓	UIQI ↑	Global PSNR	Spatial PSNR	Spectral PSNR
HySure	2.330 ± 0.016	0.849 ± 0.006	1.506 ± 0.017	$0.992 \pm 8.532 \times 10^{-5}$	39.524 ± 0.058	39.754 ± 0.060	32.706 ± 0.079
BSR	2.363 ± 0.004	0.846 ± 0.001	1.493 ± 0.003	$0.992 \pm 2.995 \times 10^{-5}$	39.399 ± 0.014	39.802 ± 0.012	32.706 ± 0.016
NSSR	2.341 ± 0.010	0.832 ± 0.003	1.495 ± 0.006	$0.992 \pm 8.519 \times 10^{-5}$	41.835 ± 0.482	40.285 ± 0.039	32.753 ± 0.034
CSTF	2.432 ± 0.035	0.870 ± 0.010	1.576 ± 0.039	$0.991 \pm 2.793 \times 10^{-4}$	41.706 ± 0.329	39.915 ± 0.104	32.389 ± 0.183
CMS	3.171 ± 0.010	1.090 ± 0.002	2.109 ± 0.009	$0.984 \pm 8.442 \times 10^{-5}$	39.440 ± 0.502	38.327 ± 0.015	30.387 ± 0.026
NLTD	2.114	0.756	1.298	0.985	41.627	–	–
R1BM3D	<b>1.960</b> ± 0.014	<b>0.697</b> ± 0.007	<b>1.188</b> ± 0.005	<b>0.994</b> ± $1.042 \times 10^{-4}$	<b>43.673</b> ± 0.214	<b>41.858</b> ± 0.099	<b>35.198</b> ± 0.042
Stuff_Toy Dataset							
Methods	RMSE ↓	ERGAS ↓	SAM ↓	UIQI ↑	Global PSNR	Spatial PSNR	Spectral PSNR
HySure	3.858 ± 0.102	0.437 ± 0.013	3.775 ± 0.202	$0.994 \pm 2.547 \times 10^{-4}$	35.294 ± 0.228	37.309 ± 0.149	24.899 ± 0.360
BSR	3.080 ± 0.007	0.318 ± 0.001	2.643 ± 0.008	$0.997 \pm 1.223 \times 10^{-5}$	37.248 ± 0.019	37.209 ± 0.020	27.581 ± 0.016
NSSR	3.108 ± 0.049	0.328 ± 0.005	2.433 ± 0.035	$0.997 \pm 9.114 \times 10^{-5}$	37.170 ± 0.137	37.570 ± 0.168	28.108 ± 0.115
CSTF	2.941 ± 0.070	0.309 ± 0.008	2.094 ± 0.066	$0.997 \pm 2.760 \times 10^{-4}$	37.649 ± 0.205	38.080 ± 0.161	29.846 ± 0.206
CMS	3.197 ± 0.018	0.328 ± 0.002	3.007 ± 0.023	$0.997 \pm 2.520 \times 10^{-5}$	36.922 ± 0.050	36.681 ± 0.053	27.045 ± 0.054
R1BM3D	<b>2.183</b> ± 0.005	<b>0.226</b> ± 0.001	<b>1.252</b> ± 0.006	<b>0.999</b> ± $6.748 \times 10^{-6}$	<b>40.237</b> ± 0.018	<b>40.638</b> ± 0.021	<b>35.401</b> ± 0.037
Urban Dataset							
Methods	RMSE ↓	ERGAS ↓	SAM ↓	UIQI ↑	Global PSNR	Spatial PSNR	Spectral PSNR
HySure	3.119 ± 0.086	0.497 ± 0.014	1.373 ± 0.046	$0.989 \pm 2.257 \times 10^{-4}$	36.173 ± 0.235	35.029 ± 0.185	34.144 ± 0.208
BSR	2.662 ± 0.005	0.422 ± 0.001	1.145 ± 0.003	$0.992 \pm 2.784 \times 10^{-5}$	37.546 ± 0.015	36.354 ± 0.014	35.679 ± 0.019
NSSR	3.088 ± 0.009	0.492 ± 0.002	1.423 ± 0.004	$0.989 \pm 6.836 \times 10^{-5}$	36.255 ± 0.025	35.010 ± 0.028	34.034 ± 0.025
CSTF	3.192 ± 0.096	0.510 ± 0.001	1.412 ± 0.003	$0.988 \pm 3.539 \times 10^{-4}$	35.987 ± 0.016	34.687 ± 0.012	33.919 ± 0.017
CMS	3.195 ± 0.005	0.519 ± 0.001	1.434 ± 0.003	$0.988 \pm 4.027 \times 10^{-5}$	35.960 ± 0.014	34.532 ± 0.013	33.936 ± 0.015
R1BM3D	<b>2.544</b> ± 0.0030	<b>0.402</b> ± 0.001	<b>1.097</b> ± 0.002	<b>0.992</b> ± $1.813 \times 10^{-5}$	<b>37.944</b> ± 0.010	<b>36.738</b> ± 0.012	<b>36.741</b> ± 0.014

To visualize the spectral reconstruction quality improvement reported in Table 13, Fig. 20 shows the spectral signatures and difference in absolute value with respect to the ground-truth for a random spatial location P1 in the reconstructions for each fused image.

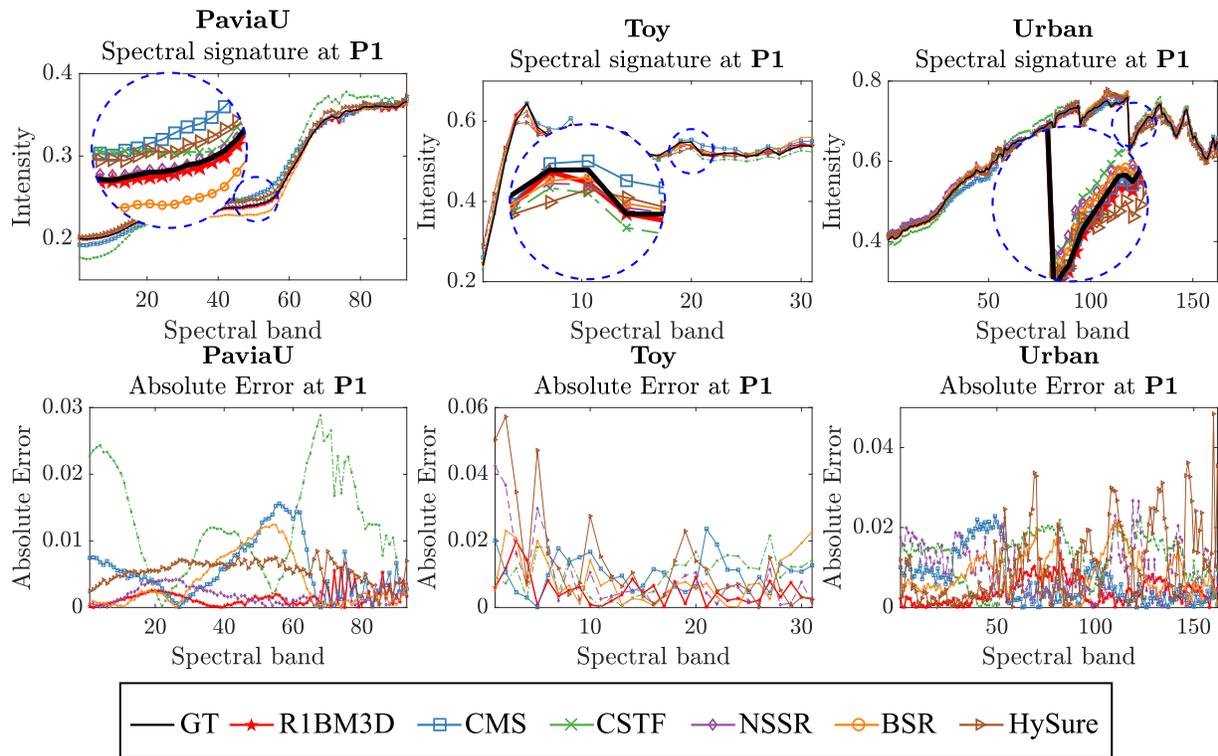


Figure 20. Spectral signature comparison at a random spatial location P1.

Comparison of the recovered spectral signatures at a random spatial location P1 for each dataset. The absolute error plots confirm that the spectral signatures obtained by the R1BM3D method are more accurate than those obtained by the comparison methods.

These results show that the R1BM3D spectral reconstructions are more accurate than those of the comparison methods. Figure 21 shows an RGB mapping of the obtained reconstructions for Pavia University, Stuff\_Toy, and Urban datasets by using the R1BM3D and comparative benchmarking. The spatial improvement can be visualized in the zoomed sub-region, especially in the reconstruction of smooth regions. Each recovered image shows the obtained global PSNR in dBs. Notice in the zoomed versions that the fused image obtained with R1BM3D presents a better re-

construction than the other methods.

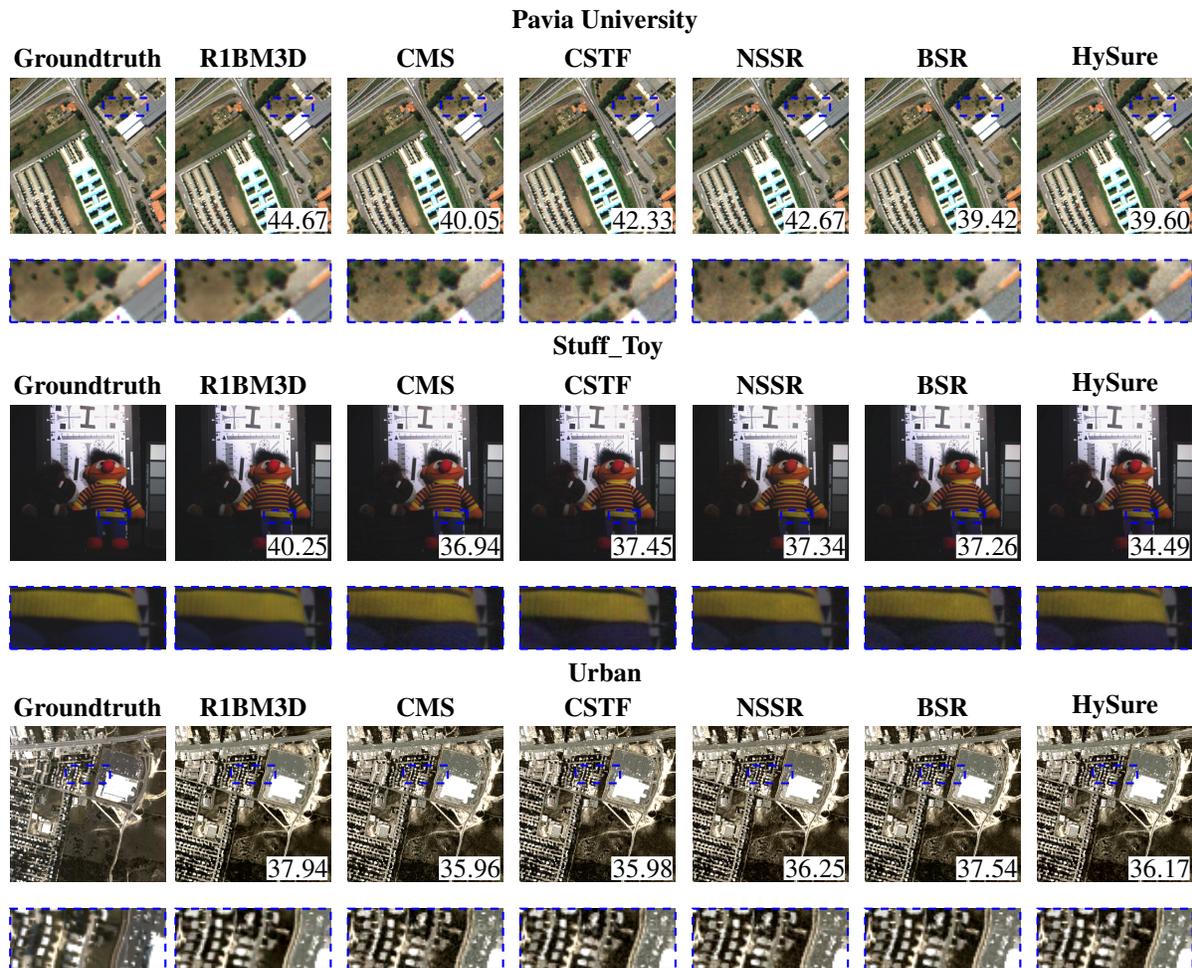


Figure 21. RGB mapping of the recovered datasets.

Figure 22 illustrates a quality comparison taking into account the reference block contrast. It can be observed for the three datasets that the quality improvement obtained at low-contrast reference blocks is more significant than the obtained for high-contrast reference blocks in comparison to the benchmarking methods, indicating that R1BM3D is especially good at recovering the spectral response of low-contrast reference blocks. Analysis of fusion quality vs block contrast.

The PSNR is computed for each block of size  $p \times p = 8 \times 8$  within a fused image and collected into nine bins according to the contrast (38) of the block at the corresponding position in the ground-truth image. The histogram shows the block bin count, confirming that high-contrast blocks are not frequent. The plots show the average PSNR over each contrast bin, comparing the proposed R1BM3D with the best comparison method for each dataset (NSSR for Pavia University, CSTF for Stuff\_Toy, and BSR for Urban). The R1BM3D is superior especially for bins corresponding to low-contrast blocks.

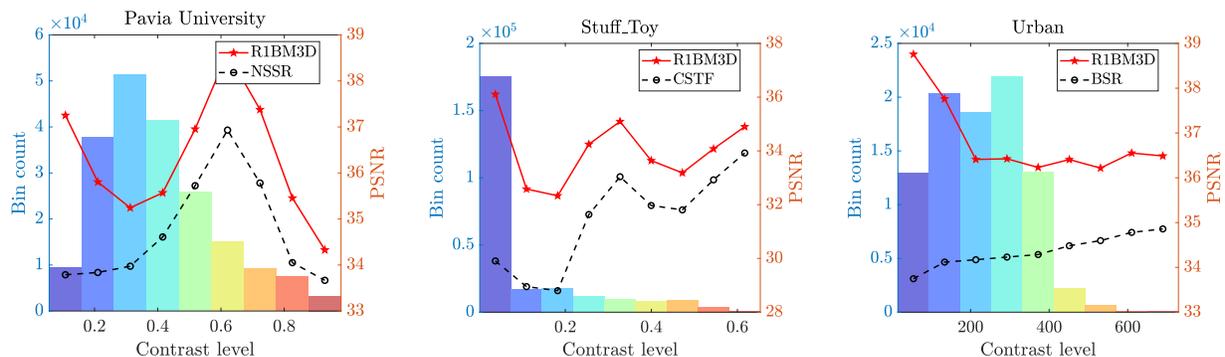


Figure 22. Analysis of fusion quality vs block contrast.

**4.5.3. Parameters selection.** Due to the very different spatial and spectral sampling conditions characterizing each dataset, the fusion algorithms all benefit from separate tuning of their key parameters. The comparison methods are tuned by trial and error as follows in order to maximize the PSNR of the fused image separately for each dataset. For NSSR the number of dictionary atoms, the number of iterations, and the regularization terms are set to  $K=256$ ,  $iter=11$ ,  $\eta_1=1e^{-4}$ , and  $\eta_2=8e^{-3}$  for Pavia University;  $K=280$ ,  $iter=12$ ,  $\eta_1=1e^{-1}$ , and  $\eta_2=3.3e^{-4}$  for Stuff\_Toy; and  $K=300$ ,  $iter=12$ ,  $\eta_1=3.5e^{-5}$ , and  $\eta_2=1.7$  for Urban dataset. For the CSTF the number of dictionary atoms of the spatial-spectral modes are set to  $n_w=256$ ,  $n_h=256$  and  $n_s=11$

for Pavia University;  $n_w = 150$ ,  $n_h = 150$  and  $n_s = 12$  for Stuff\_Toy; and  $n_w = 250$ ,  $n_h = 250$  and  $n_s = 5$  for Urban dataset. For the CMS the size of the overlapped local full band patches, the number of clusters, and the interval between patches are set to  $\sqrt{q} = 2$ ,  $K = 250$ , and  $d = 1$  for Pavia University;  $\sqrt{q} = 8$ ,  $K = 500$ , and  $d = 2$  for Stuff\_Toy; and  $\sqrt{q} = 2$ ,  $K = 300$ , and  $d = 1$  for Urban dataset.

R1BM3D has three key parameters,  $\lambda$  controlling the implicit regularization term,  $\rho$  controlling the dual variable weights, and  $\mu$  controlling the weight of the fidelity term to the MSI. In the experiments, I fix  $\lambda = \sigma^2 \rho$ , where  $\sigma$  stands for the MSI effective noise standard deviation, which is assumed by the multichannel BM3D for denoising. Then,  $\rho$  and  $\mu$  are tuned through simulations for each dataset. Figure 23 shows the global PSNR as function of  $\rho$  and  $\mu$ . From the simulations, I set  $\mu = 7$  and  $\rho = 0.16$  for Pavia University;  $\mu = 0.6$  and  $\rho = 0.02$  for Stuff\_Toy; and  $\mu = 1.65$  and  $\rho = 0.009$  for Urban dataset. In general, the selection of  $\rho$  affects the convergence speed. Very small values will produce a fast convergence but will reduce the importance of the dual terms, affecting the quality. Meanwhile, very high values of  $\rho$  will produce slow convergence. Also, notice that  $\mu$  is commonly higher than  $\rho$ , so that the fidelity to the MSI is more relevant to obtain good quality.

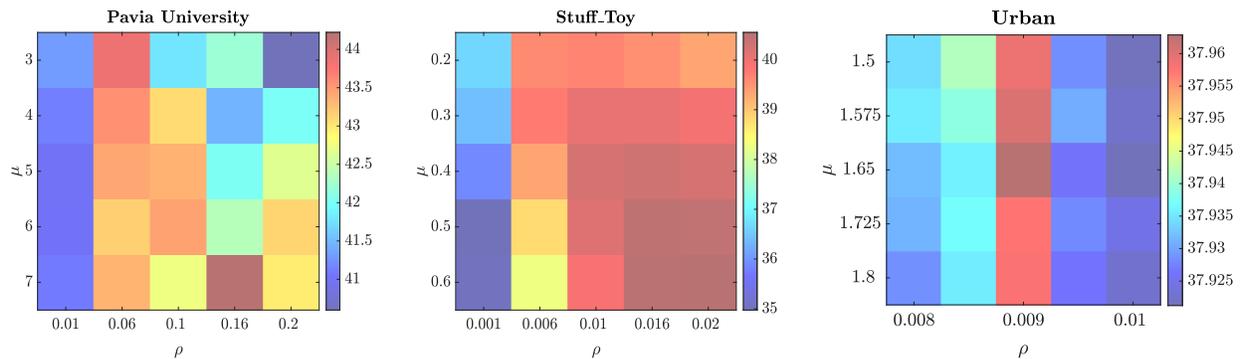


Figure 23. Parameters Analysis for each dataset.

**4.5.4. Rank analysis.** The number of PCs that are preserved at each iteration is analyzed. This number is directly related to the underlying HRI rank, which is expected to be low. Figure 24 shows the nuclear norm, the number of preserved PCs, and the global PSNR as the iterations progress. Analysis of the number of preserved PCs during the iterations of R1BM3D. The plots show the behaviour of the number of preserved PCs that are considered significant, the nuclear-norm of the estimated HRI, and the obtained global PSNR across the iterations. After some iterations, the number of PCs converges to a small number for the three datasets. This behavior was expected because of the low-rank property. Furthermore, it can be observed a counter-relation between the nuclear norm and the global PSNR during the first few iterations, in which a high value in the nuclear norm implies low quality. These results indicate that the information removed during the first iterations effectively corresponds to noise. It can be observed that the number of preserved PCs converges to a relatively small number, inducing a low-rank structure. Notice how the nuclear norm fluctuates opposite to the PSNR; this is most noticeable during the first stages of recovery when the filtering suppresses spurious and erroneous structures from the HRI estimate.

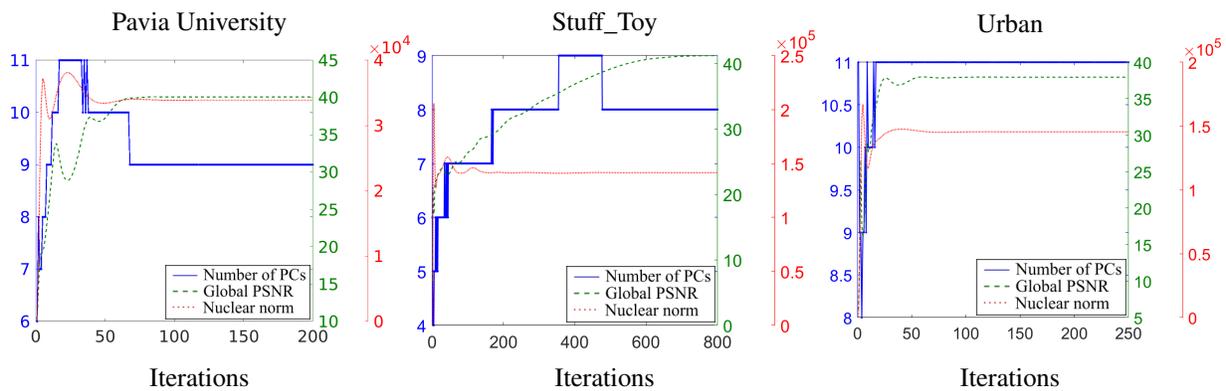


Figure 24. Analysis of the number of preserved PCs during the iterations of R1BM3D.

#### 4.6. Discussion

The presented fusion approach takes advantage of a rank-one similarity prior, under which the dissimilarity maps of high-contrast reference blocks are assumed to be proportional to each other across the spectral domain. Nonetheless, the analysis reported in Fig. 22 shows that the greatest improvement comes from low-contrast blocks instead of from high-contrast blocks, which is perhaps unexpected. These results could indicate that the strategy of finding the dissimilarity maps in the first PC can improve the recovery of low-contrast blocks. Since low-contrast areas are more impacted by noise, their reconstruction can be challenging for the comparison methods. Using the rank-one prior in R1BM3D helps by guiding the block-matching under heavy noise, and therefore the quality is improved.

In connection to this work, it is interesting to mention the hyperspectral denoising method FastHyDe presented in (Zhuang and Bioucas-Dias, 2018). While FastHyDe is also based on a low-rank approximation of the HRI and self-similarity via BM3D, the combination of these principles with the rank-one similarity prior proposed here is stronger and more efficient than the one adopted by FastHyDe. Specifically, in FastHyDe it is assumed that each PC is internally self-similar and the denoising is applied separately on each PCs implying that the block-matching is operated independently on each PC, which leads to increased computation and can also be impaired by the low SNR of some PCs. R1BM3D is instead a multichannel approach, where through the rank-one self-similarity prior, the block-matching is operated only once, on the PC with the highest SNR, and reuse the matching positions for all PCs; this leads to computational savings and substantial

benefit for content with low SNR and low-contrast.

#### **4.7. Conclusions**

This chapter considered the joint of various low-rank regularizations in order to improve the HSI-MSI fusion recovery quality. Unlike previous chapters using explicit regularizers, the proposed R1BM3D method combined the low-rank priors through an implicit regularization function. Notably, I introduced an alternative low-rank prior, dubbed rank-one similarity prior, as an intrinsic HRI characteristic under which the structural similarities of small blocks are shared across all spectral bands. I conducted an extensive empirical study over several real-world objects and remote-sensing and satellite datasets, finding support for such rank-one similarity prior. Furthermore, I developed an effective algorithm that uses a non-local patch-based denoiser to promote the proposed low-rank priors. Finally, the experiments showed that R1BM3D can significantly improve the HSI-MSI fusion, with the most remarkable improvement observed in the recovery of the spectral information at low-contrast locations.

## 5. Mixture-Net: Low-Rank Deep Image Prior Inspired by Mixture Models for Spectral Image Recovery

*Part of this chapter has been adapted from the conferences (Gelvez et al., 2021a,b) [© [2021] IEEE. Reprinted] (published) and journal Mixture-Net: Low-Rank Deep Image Prior Inspired by Mixture Models for Spectral Image Recovery Submitted to IEEE Trans. on Image Processing.*

This chapter presents a strategy to incorporate the low-rank prior without requiring a defined regularization function. The proposed approach, dubbed Mixture-Net, aims to learn implicitly the SI low-rankness through a non-data-driven deep neural network, whose architecture is inspired on mixture models. Section 5.1 describes the related works, where the hand-crafted regularizer design and DL data-driven limitations, motivate the proposal with the following main contributions.

1. Mixture-Net: An interpretable network architecture inspired by linear and non-linear low-rank mixture models, where the learned weights and features can be interpreted as the SI abundances and endmembers. (Section 5.2.2).
2. A scheme of multiple deep-blocks, whose loss functions contain particular model-based regularizers to improve the SI learned spatial-spectral correlations (Section 5.2.3).
3. A non-data-driven DIP-based approach for SI denoising, HSI-SR, and CSI reconstruction (Section 5.2).
4. A significant improvement for SI recovery in terms of recovery quality while reducing the processing complexity through the non-data-driven approach (Section 5.3).

### 5.1. Spectral image recovery related works

Using prior information is crucial in current SI recovery methods to effectively solve the ill-posed inverse problems, through *greedy* algorithms, *model-based* optimization, or DL approaches.

Greedy methods include the prior information by following a sequence of reasonable steps seeking locally optimal solutions (Blumensath et al., 2012; Maggioni et al., 2012). Meanwhile, model-based optimization methods consider the prior information by designing hand-crafted regularizers to reduce the feasible search set, finding thus globally optimal solutions (Mullah et al., 2020; Xu et al., 2020; Li et al., 2020a; Gelvez et al., 2017). For instance, the low-rank regularizer promotes a low-dimensional subspace prior by modeling the SI as a low-rank structure (Kanatsoulis et al., 2018; Bacca et al., 2019; Dian et al., 2018). Nonetheless, hand-crafted regularizers are often insufficient to handle the wide variety of spectral information.

Data-driven DL methods learn the prior information by training a black-box non-linear mapping from a SI dataset. The DL black-box nature has been tackled by making systematic connections between model-based iterative algorithms and deep neural networks (Wang et al., 2020b; Ramírez et al., 2021; Vu et al., 2020). Nonetheless, DL methods are still impractical because of the expensive acquisition cost of several SI datasets with high spectral resolution.

DIP methods overcame the DL data dependency, thus showing that a single generator network is sufficient to capture the low-level spectral statistics, recovering the SI in a non-data-driven manner (Ulyanov et al., 2018; Bacca et al., 2021; Zhang et al., 2019). Nonetheless, the DIP networks also behave as a black-box, where the architecture disregards the prior information. Re-

cently, authors in (Miao et al., 2021) and (Gelvez et al., 2021a) proposed a *proper* neural network architecture to incorporate the LMM as prior information to solve SI denoising and CSI reconstruction, respectively. However, the SI non-linear interactions and the benefits of using an interpretable network have not been considered.

This chapter tackles aforementioned limitation through a non-data driven DIP-based approach that incorporates the low-rankness in an interpretable architecture.

## 5.2. Proposed spectral image recovery method

The proposal follows the mathematical notation and formulation of the SI problems defined in Section 1.3, and the low-rank mixture models for SIs defined in Section 1.4.2.

The proposed SI recovery method aims to include the prior information implicitly in the architecture of a deep model that generates the SI measurements by minimizing

$$\hat{\theta} \in \underset{\theta}{\operatorname{argmin}} \mathcal{L}(\mathbf{y}, \Phi \mathcal{M}_{\theta}(\mathbf{f}^0)), \quad (45)$$

where  $\hat{\mathbf{f}} := \mathcal{M}_{\hat{\theta}}(\mathbf{f}^0)$  denotes the recovered SI,  $\mathcal{M}_{\theta}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n$  denotes the deep generative network with  $\theta$  adjustable weights,  $\mathcal{L}(\cdot) : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$  denotes a customized loss function, and  $\mathbf{f}^0 \in \mathbb{R}^n$  denotes the model input. Notice that (45) only requires the measurements  $\mathbf{y}$ , the acquisition operator  $\Phi$ , and the input  $\mathbf{f}^0$ , i.e., the proposed method is non-data-driven.

Then, the proposed SI recovery method consists of three components schematized in Fig. 25. (i) The input structure is based on tensor decomposition. (ii) The Mixture-Net interpretable architecture generates the recovered SI, the learned features interpreted as the abundances, and

the adjusted weights interpreted as the endmembers. (iii) The customized loss function includes the forward model, the losses of each deep-block, and the mixture model physical constraints. Specifically, (i) the input  $\mathbf{f}_0$  is learned as a tensor decomposition to impose a low-rank structure that passes through the network. (ii) Mixture-Net is composed by a sequence of multiple interpretable deep-blocks  $\mathcal{M}_{\theta^k}(\cdot)$ . The  $k^{th}$  interpretable deep-block contains an abundance block-layer  $\mathcal{A}_{\theta}^k(\mathbf{f}^{k-1})$ , consisting of a CNN to learn the spatial correlations, an endmember block-layer  $\mathcal{E}_{\mathbf{E}}(\bar{\mathbf{a}}^k)$  performing the matrix product according to the LMM, and a non-linearity block-layer  $\mathcal{N}_{\theta^k}(\mathcal{E}_{\mathbf{E}}(\bar{\mathbf{a}}^k))$  containing a CNN to learn the non-linear mapping operator of the NLMM. (iii) The customized loss function considers the forward operator  $\Phi$  and the sum of the single losses employed at each interpretable deep-block including the abundance constraint

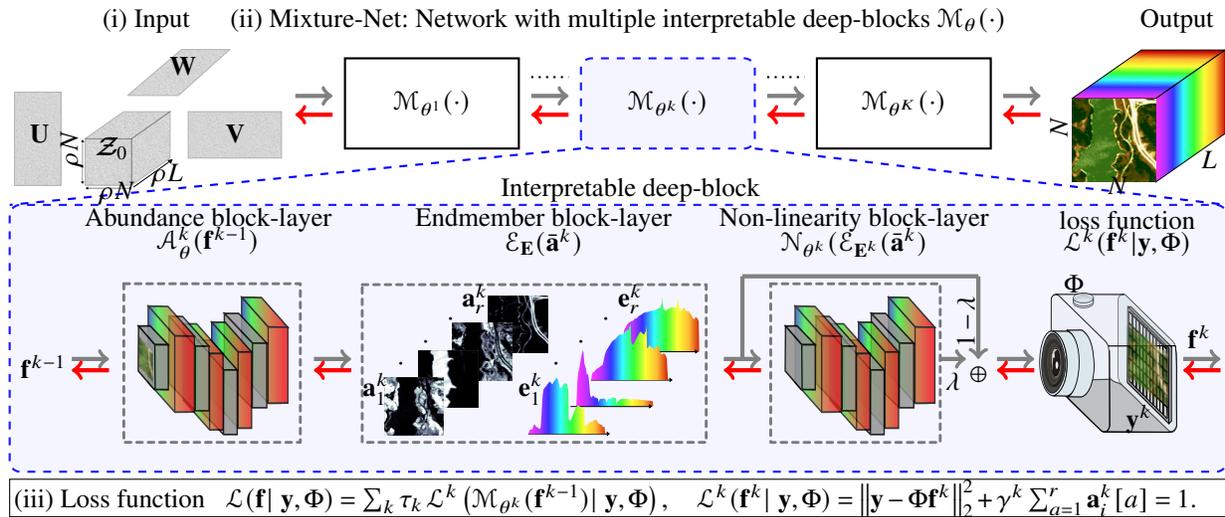


Figure 25. Proposed SI recovery method scheme.

**5.2.1. Input structure component.** The input component determines the input structure for the network  $\mathcal{M}_{\theta}(\cdot)$ , considering the *fixed* and *learned* approaches.

**Fixed Input.** The input is fixed in the sense that its entries do not change during the learning process. Thus, the input is chosen as a fixed array  $\mathbf{f}^0 \in \mathbb{R}^n$  generated as random noise, constant values, SI rough estimation, or so on.

**Learned Input.** The input is computed as an adjustable variable from a blind representation by solving the problem

$$\{\hat{\theta}, \hat{\mathbf{f}}^0\} \in \underset{\theta, \mathbf{f}^0 \in \mathbb{R}^n}{\operatorname{argmin}} \mathcal{L}(\mathbf{y}, \Phi \mathcal{M}_\theta(\mathbf{f}^0)). \quad (46)$$

Authors in (Bacca et al., 2021) suggested imposing a low-dimensional structure over the input  $\mathbf{f}^0$  to force obtaining a low-dimensional output. Therefore, a low-rank structure is imposed by learning the input according to the Tucker decomposition as  $\mathbf{f}^0 = \operatorname{vec}(\mathcal{Z})$ ,  $\mathcal{Z} = \mathcal{Z}_0 \times_1 \mathbf{U} \times_2 \mathbf{V} \times_3 \mathbf{W}$ , where the variables  $\mathcal{Z}_0, \mathbf{U}, \mathbf{V}, \mathbf{W}$  are fitted by minimizing (46). The Tucker decomposition maintains the SI 3D structure and guarantees low-dimensionality given that  $\mathcal{Z}_0 \in \mathbb{R}^{N_\rho \times N_\rho \times L_\rho}$  stands for a 3D low-rank tensor, where  $N_\rho < N$  and  $L_\rho < L$ , where  $\rho$  is a scale factor.

**5.2.2. Mixture-Net: Network architecture component.** To capture the SI prior information implicitly, Mixture-Net is composed of a sequence of  $K$  interpretable deep-blocks inspired by low-rank mixture models, containing an abundance, an endmember, and a non-linearity block-layer.

**Abundance block-layer.** The first block-layer consists of a CNN that filters the input with the size of the reference SI to obtain an output whose structure and dimensions should match for being interpreted as the abundances. The abundance block-layer of the  $k$ -th interpretable

deep-block can be expressed as

$$\mathbf{a}^k = \mathcal{A}_\theta^k(\mathbf{f}^{k-1}), \quad (47)$$

where  $\mathcal{A}_\theta^k(\cdot) : \mathbb{R}^{N^2L} \rightarrow \mathbb{R}^{N^2r}$  connotes the CNN with  $r$  being a tunable hyper-parameter related to the SI rank. This block-layer incorporates the non-negativity and sum-to-one physical constraints described in Section 1.4.1 by using the *sigmoid* function as the activation of the last layer and including the following regularization term in the loss function

$$R(\bar{\mathbf{a}}^k) = \sum_{a=1}^r \mathbf{a}_i^k[a] = 1, \quad \mathbf{a}_i[a] \geq 0, \forall i. \quad (48)$$

**Endmember block-layer.** The second block-layer consists of an operator that performs the matrix multiplication between the learned features in the abundance block-layer  $\mathbf{a}^k$  and the adjusted model weights in the endmember block-layer  $\mathbf{E}^k$ , whose dimensions should match to be interpreted as the endmembers. The endmember block-layer can be expressed as the linear component of the low-rank mixture model as

$$\mathbf{L}^k = \mathcal{E}_{\mathbf{E}^k}(\bar{\mathbf{a}}^k) = (\mathbf{E}^k \otimes \mathbf{I}_{N^2})\bar{\mathbf{a}}^k, \quad (49)$$

where  $\mathcal{E}_{\mathbf{E}^k}(\cdot) : \mathbb{R}^{N^2r} \rightarrow \mathbb{R}^{N^2L}$  models the fully connected endmember block-layer. This block-layer includes the non-negativity constraint described in Section 1.4.1 by projecting the estimated endmembers into  $\mathbb{R}_+$  at each gradient step.

**Non-Linearity block-layer.** The third block-layer consists of a convolutional operator determining the non-linear transformation applied to the linear component obtained in the endmember block-layer  $\mathbf{L}^k$ . Therefore, the block-layer learns the non-linear interactions providing an output whose structure and dimension should match to be interpreted as the recovered SI. The third block-layer can be expressed as the non-linear component of the low-rank mixture model as

$$\mathbf{NL}^k = \mathcal{N}_{\theta^k}(\mathcal{E}_{\mathbf{E}^k}(\mathbf{a}^k)), \quad (50)$$

where  $\mathcal{N}_{\theta}^k(\cdot) : \mathbb{R}^{N^2r} \rightarrow \mathbb{R}^{N^2L}$  connotes a CNN that determines the non-linear transformation following the NLMM.

In this manner, the  $k^{th}$  interpretable deep-block estimates the recovered SI, generating three outputs given by

$$\begin{aligned} \mathcal{M}_{\theta^k}(\mathbf{f}^{k-1}) &= \mathbf{f}^k = (1 - \lambda)\mathbf{L}^k + \lambda\mathbf{NL}^k, \\ \mathbf{L}^k &= \mathcal{E}_{\mathbf{E}^k}(\mathcal{A}_{\theta}^k(\mathbf{f}^{k-1})), \\ \mathbf{NL}^k &= \mathcal{N}_{\theta^k}(\mathcal{E}_{\mathbf{E}^k}(\mathcal{A}_{\theta}^k(\mathbf{f}^{k-1}))). \end{aligned} \quad (51)$$

for  $k = 1, \dots, K$  interpretable deep-blocks. Notice that  $\mathbf{L}^k$  results from multiplying the adjusted weights  $\mathbf{E}^k$  with the learned features  $\mathbf{a}^k$ , such that  $\mathbf{L}^k$  can be interpreted as the SI LMM. Similarly,  $\mathbf{NL}^k$  results from applying a non-linear transformation  $\mathcal{N}_{\theta^k}(\cdot)$  to the linear term  $\mathbf{L}^k$ , such that  $\mathbf{NL}^k$  can be interpreted as the NLMM. Therefore, each estimated  $\mathbf{f}^k$  results from an affine combination of a linear and a non-linear term regulated by the parameter  $0 \leq \lambda \leq 1$ .

**5.2.3. Customized loss function component.** The chosen loss function for each SI recovery problem is decisive for the learning effectiveness. Therefore, a customized loss function is presented that considers the forward model, the independent deep-blocks, and the physical constraints. First, the forward sensing operator  $\Phi$  is applied at each deep-block output to predict the measurements. Subsequently, the model weights  $\theta$  are adjusted by minimizing a loss measuring the difference between the observed and predicted measurements, and containing the mixture model physical constraints. The customized loss function is composed of the sum of the independent losses at each deep-block as follows

$$\begin{aligned} \{\theta^*\} &\in \underset{\theta}{\operatorname{argmin}} \sum_k \tau_k \mathcal{L}^k(\mathbf{f}^k | \mathbf{y}, \Phi), \\ \mathcal{L}^k(\mathbf{f}^k | \mathbf{y}, \Phi) &= F(\mathbf{f}^k | \mathbf{y}) + \gamma^k \sum_{a=1}^r \mathbf{a}_i^k[a] = 1, \end{aligned} \quad (52)$$

where,  $\mathbf{f}^k := \mathcal{M}_{\theta^k}(\mathbf{f}^{k-1})$ ,  $\mathcal{M}_{\theta^k}(\cdot) : \mathbb{R}^{N^2L} \rightarrow \mathbb{R}^{N^2L}$  stands for the  $k^{\text{th}}$  interpretable deep-block,  $\tau_k > 0$  denotes the  $k^{\text{th}}$  loss function relative weight in the sum, and  $\gamma_k > 0$  denotes the  $k^{\text{th}}$  regularization parameter that controls the trade-off between the fidelity-data  $F(\mathbf{f}^k | \mathbf{y})$  and the abundance constraint. The output at each deep-block could be interpreted as the recovered SI; therefore, two possible strategies are studied for defining the recovered SI (i) using the last deep-block output, i.e.,  $\hat{\mathbf{f}} := \mathbf{f}^K$  and (ii) using the average between the last two deep-blocks outputs, i.e.,  $\hat{\mathbf{f}} := (\mathbf{f}^K + \mathbf{f}^{K-1})/2$ .

### 5.3. Simulations and Results

The experiments to evaluate the Mixture-Net performance were conducted over the Pavia University, Pavia Center, CAVE, KAIST, ARAD, and Jasper-Ridge publicly available spectral datasets

described in Appendix 1.

Mixture-Net was tested in three SI recovery tasks, SI denoising, HSI-SR, and CSI reconstruction. For HSI-SR and CSI reconstruction, the  $\ell_2$ -norm was used as the fidelity term, given by  $F(\mathbf{f}^k | \mathbf{y}) = \|\mathbf{y} - \Phi \mathbf{f}^k\|_2^2$ . For SI denoising, the SURE loss in (Nguyen et al., 2020) was used given by

$$F(\mathbf{f}^k | \mathbf{y}) = \|\mathbf{y} - \Phi \mathbf{f}^k\|_2^2 - \sigma^2 + \frac{2\sigma}{N^2 L} \text{div}_{\mathbf{y}}(\Phi \mathbf{f}^k), \quad (53)$$

where  $\text{div}_{\mathbf{y}}(\Phi \mathbf{f}^k)$  is the divergence of  $\Phi \mathbf{f}^k := \Phi \mathcal{M}_{\theta^k}(\mathbf{f}^0)$  computed with the Monte-Carlo SURE strategy in (Ramani et al., 2008) as

$$\text{div}_{\mathbf{y}}(\Phi \mathbf{f}^k) \approx \mathbf{b}^T \left( \frac{\Phi(\mathcal{M}_{\theta^k}(\mathbf{f}^0 + \epsilon)) - \Phi \mathcal{M}_{\theta^k}(\mathbf{f}^0)}{\epsilon} \right). \quad (54)$$

$\mathbf{b} \in \mathbb{R}^{N^2 L}$  is an i.i.d. Gaussian distribution with zero mean and unit variance, and  $\epsilon$  is a small value of order  $1 \times 10^{-5}$ .

The quality improvement is quantified through the spectral angle mapper (SAM), the root mean squared error (RMSE), the dimensionless global relative error of synthesis (ERGAS), the peak signal-to-noise ratio (PSNR), and the structural similarity (SSIM) metrics defined in Appendix 2. All experiments were run on an Intel Xeon W-3223 with 64GB of memory, and an NVIDIA

RTX 3090 GPU with 24GB of memory<sup>2</sup>.

**5.3.1. Characterization of Mixture-Net.** The first experiment illustrates the effect of varying each one of the components and hyper-parameters affecting the Mixture-Net. (i) The input strategy, imposing various structures over the network input. (ii) The abundance block-layer scheme, involving the number of layers and filters. (iii) The non-linearity block-layer, varying the number of endmembers. (iv) The number  $K$  of employed deep-blocks. For this purpose, the HSI-SR task is solved over the Pavia Center dataset.

*Input structure.* I evaluated five strategies imposing a particular structure over the Mixture-Net input  $\mathbf{f}^0$ . The former four belong to the *fixed* approach in Section 5.2.1 as follows. A *Constant* input, using a tensor with all values equal to 0.5, i.e.,  $\mathcal{Z} \in \{0.5\}^{N \times N \times L}$ . A *Random* input, generating a random tensor with a normal Gaussian distribution. A *Mesh-grid* input, initializing the tensor by the method in (Ulyanov et al., 2018). A *Estimated* input, using a rough image estimation given by  $\text{vec}(\mathcal{Z}) = \Phi^T \mathbf{y}$ . The fifth is the *learned* strategy learning a low-dimensional Tucker Decomposition from random noise using  $\rho = 0.4$  as described in Section 5.2.1.

Furthermore, the input  $\mathbf{f}^0$  considers a perturbation at each iteration of the learning process, emulating external noise. This strategy has been shown to improve the results quality (Ulyanov et al., 2018). The perturbation is given by

$$\mathbf{f}^{(0)} = \text{vec}(\mathcal{Z}) + \beta \boldsymbol{\eta}, \quad (55)$$

---

<sup>2</sup> The source code is publicly available in Mixture-Net code

where  $\text{vec}(\cdot)$  denotes a vectorization operator,  $\beta \in \mathbb{R}_+$  is a parameter controlling the perturbation level, and  $\boldsymbol{\eta}_{i,j} \sim \mathcal{N}(0, 1)$  stands for Gaussian additive noise.

Figure 26 shows the Mixture-Net quality results when using the five strategies imposing a structure in the input of Mixture-Net. The box-plot of the quantitative performance is measured in terms of the (*left*) PSNR (purple) / SAM (green) and (*right*) SSIM (yellow) / RMSE (red) metrics. The level of perturbation was varied across the values  $[0, 1e^{-2}, 3e^{-2}, 5e^{-2}, 8e^{-2}, 1e^{-1}]$ . The Abundance block-layer is a CNN with six layers using thirty two filters, and I set the hyper-parameters as  $lr = 5e^{-3}$ ,  $\lambda = 0$  (only the LMM), and  $r = 8$ . It can be observed that the random input is far away from being a good input in a DIP architecture, where even the mesh-grid or rough estimation strategies improves the obtained quality. The proposed *Learned* input emerges as the best strategy, exhibiting a variance no greater than 0.47 [dBs] of PSNR and 0.27 degrees of SAM.

**Abundance-block-layer.** The abundance block-layer tuning evaluates three architectures to learn the spatial features leading to the estimated abundance maps. The *Convolutional* refers to a sequence of 2D convolutional layers with padding as explained in (Goodfellow et al., 2016). The *Auto-encoder* refers to a sequence of 2D convolutional layers, where the first half increase at the double at each feature, i.e.,  $[\ell, 2\ell, \dots, 2\nu\ell]$ , and the remaining half decrease  $[2\nu\ell, (2\nu - 1)\ell, \dots, \ell]$  as presented in (Choi et al., 2017a). The *ResNet* refers to a sequence of residual neural layers, where the first and the last convolutional layers are concatenated as described in (Bacca et al., 2021). Figure 27 shows Quantitative performance in terms of PSNR (*left*) and SAM (*right*) metrics varying the number of layers and the learned features per layer for three architectures to estimate the abundances. The performance is shown varying the corresponding

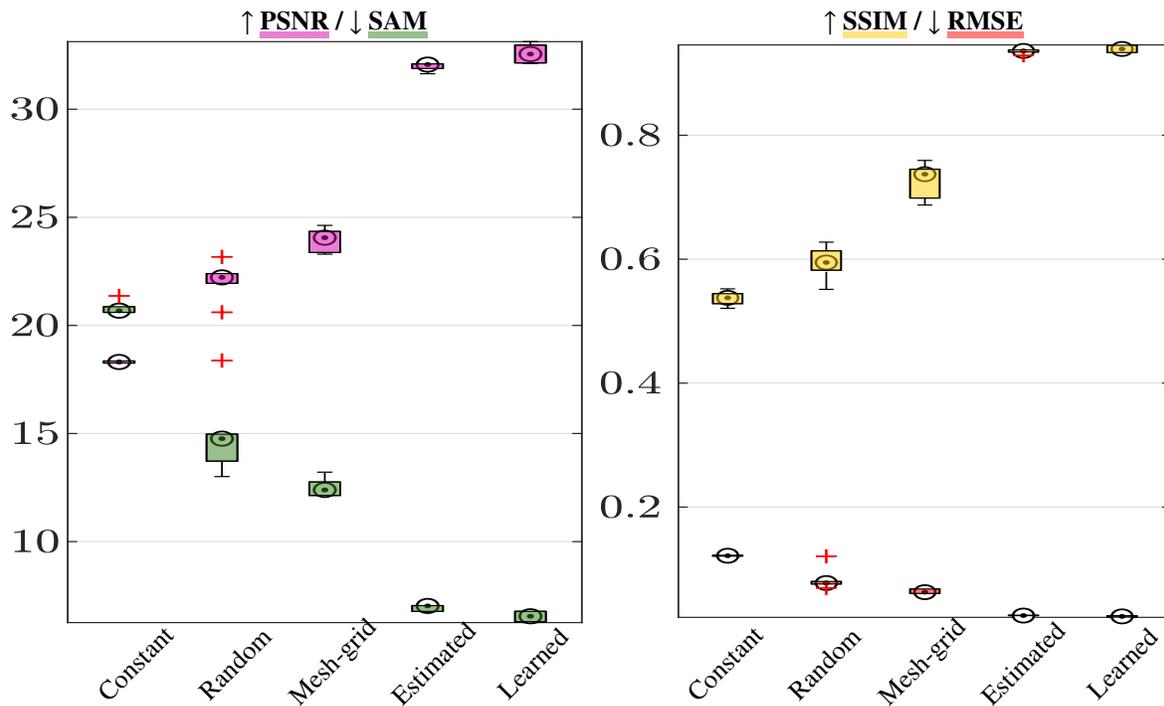


Figure 26. Input structure quantitative performance box-plot.

hyper-parameters: the number of layers varies in the range  $[1, 2, 3, 4, 5]$  for the *Convolutional* and *Res-Net* architectures, and in the range  $[3, 5, 7]$  for the *Auto-encoder*; and the number of features per layer varies in the range  $[8, 16, 32, 64]$ . The best learning-rate was found to be  $lr = 1e^{-3}$ , and the noise perturbation of the input was set to be  $\beta = 0.7$ . It can be observed that the *Convolutional* architecture provides the highest quality, showing the most stable behavior with minimal variances when using thirty two features or more.

**Non-linearity block-layer.** This experiment evaluates the effect of introducing the non-linearity block-layer in the architecture varying the rank. The rank can be addressed as a hyper-parameter related to the number of different materials in the SI. Figure 28 shows a quantitative comparison between Mixture-Net with just the LMM ( $\lambda = 0$  in (51)) and Mixture-Net when the

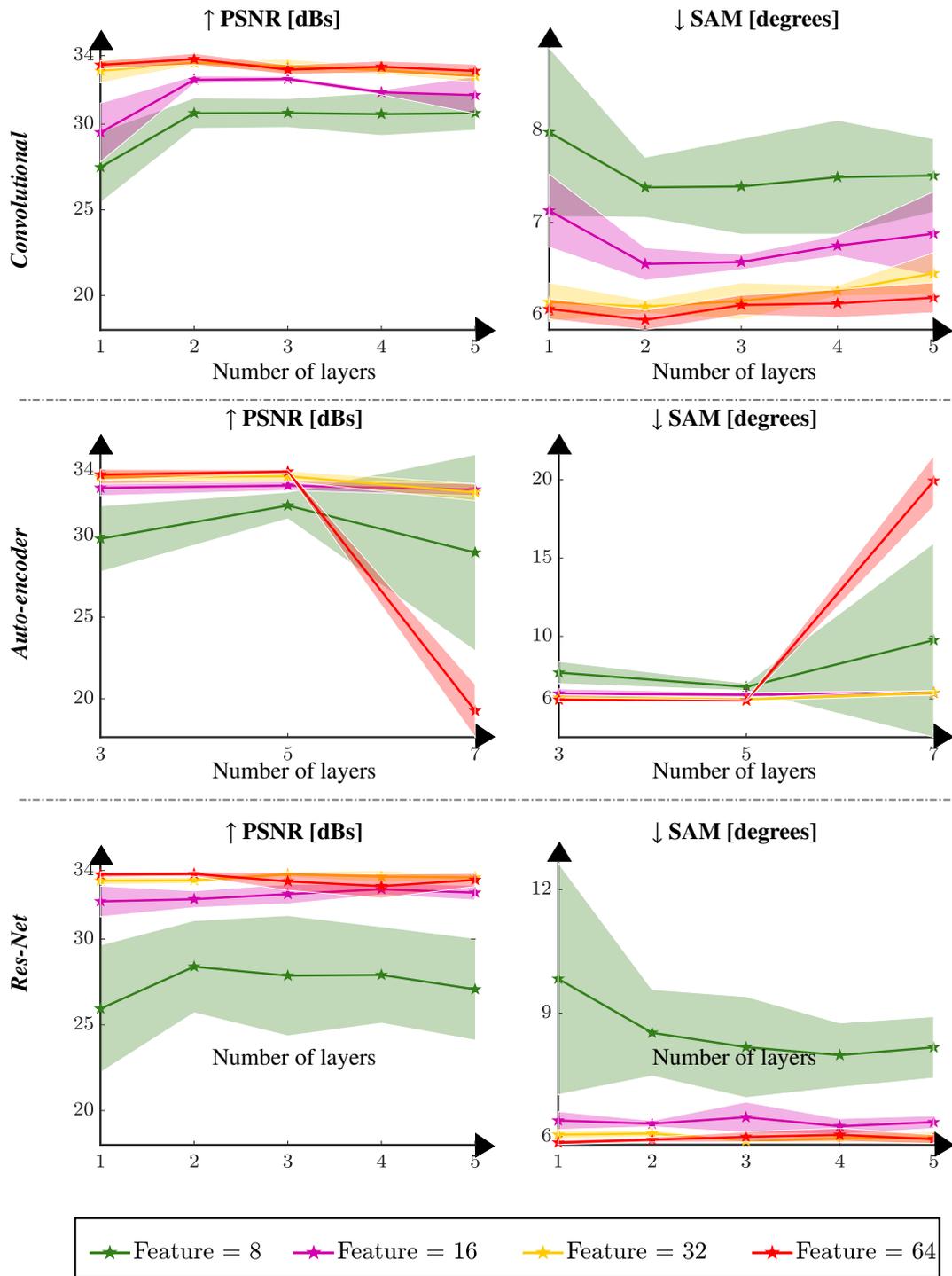


Figure 27. Architectures quantitative performance.

non-linearity block-layer is included, i.e.,  $\lambda > 0$ . Specifically, the performance is measured in terms of the PSNR (*left*) and SAM (*right*) metrics when using the linear block-layer, i.e.,  $\lambda = 0$  in (51), and when including the non-linearity block-layer with  $\lambda = 0.7$ . The rank value varies in the range [3, 4, 5, 6, 7, 8, 9, 10, 15, 20]. The non-linearity block-layer is composed of two consecutive spatial-spectral networks presented in (Wang et al., 2019a), where the best learning rate was found to be  $lr = 1e^{-3}$ . It can be observed a clear improvement in the quality, especially for small values of the rank. This result indicates that using many endmembers yields to compensate for the non-linear relations in the underlying scene. However, applications such as material identification should be aware of just a few ground truth endmembers that interact in non-linear ways to form the scene, instead of using a large number of endmembers that do not match the spectral response of any material (Bioucas-Dias et al., 2012).

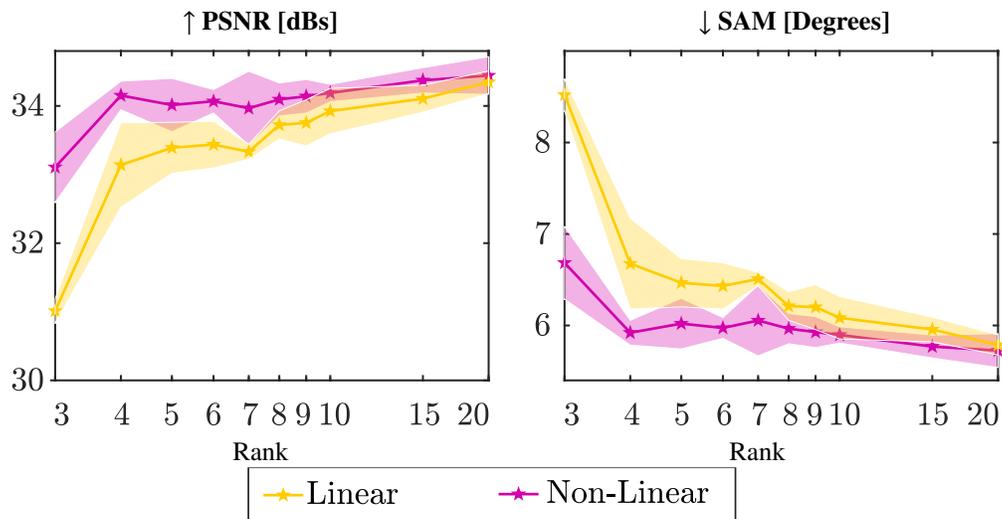


Figure 28. Non-linearity block-layer performance.

***Number of interpretable deep-blocks.*** This experiment evaluates the effect of concatenating multiple interpretable deep-blocks, where the non-linearity block-layer is included only at the last deep-block. Each deep-block in the sequence provides an output  $\mathbf{f}^k$  that could be interpreted as the recovered image. Therefore, Fig. 29 plots the quantitative comparison among the output quality at each deep-block when concatenating two, three, four, and five deep-blocks. For instance, the experiment of five deep-blocks (Block-5) provides six recovered images, five given by the linear representation obtained at each deep-block, and one given by the non-linear representation obtained at the fifth deep-block. Quantitative performance in terms of PSNR when using two, three, four and five deep-blocks, i.e.,  $K = [2, 3, 4, 5]$ . The dark lines indicate the mean value, and the width of the light boxes indicates the variance of running five experiments. The non-linearity block-layer is composed of two consecutive spatial-spectral networks presented in (Wang et al., 2019a), where the best learning rate was found to be  $lr = 1e - 3$ . The labels are shared among the plots with exception to the label of the non-linearity block shown at each box. It can be observed that using two and three deep-blocks improves the quality. However, from using four deep-blocks, the quality decreases because the neural network begins getting over-fitting to the data-fidelity term. Therefore, I concatenate two deep-blocks that provide a slight result variance as shown in Fig. 29.

**5.3.2. Mixture-Net Performance for Spectral Image Recovery.** After determining the hyper-parameters and Mixture-Net optimal structure, the second experiment compares the recovery performance against state-of-the-art methods when solving denoising, single HSI-SR, and CSI reconstruction.

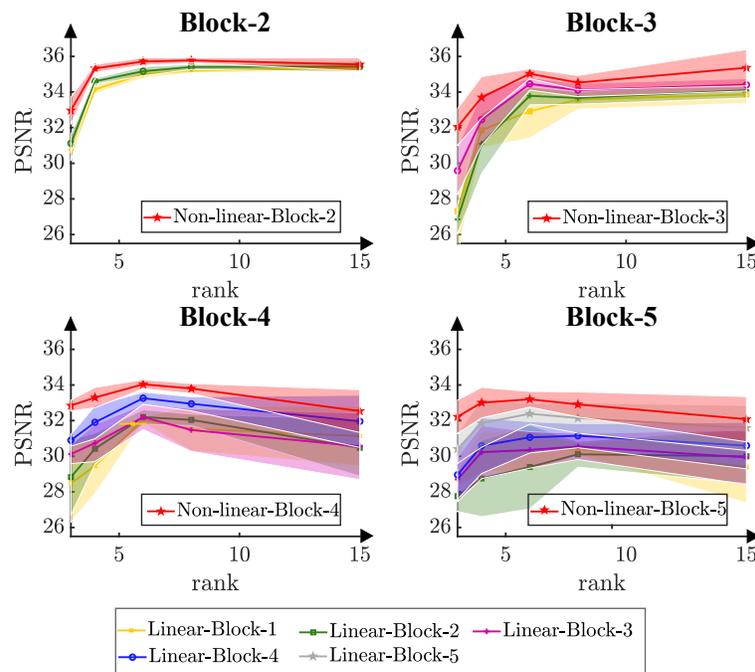


Figure 29. Multiple deep-blocks quantitative performance.

**Denoising.** This experiment evaluates the Mixture-Net performance for SI denoising in the presence of Gaussian additive noise over the Pavia University dataset. The SURE loss is used to avoid over-fitting in the learning process without an early stop strategy. The performance was evaluated against the block-matching and 3D filtering (Dabov et al., 2007), the first order spectral roughness penalty for denoising (Rasti et al., 2013), the hyperspectral restoration method (Rasti et al., 2017), the hyperspectral-DIP (Sidorov and Yngve Hardeberg, 2019), the SURE-CNN (Choi et al., 2017a), and the unsupervised disentangled spatio-spectral deep priors (DS2DP) method (Miao et al., 2021).

Table 14 presents the denoising quantitative results regarding PSNR and SSIM metrics, where Mixture-Net obtains a significant improvement, especially in the SSIM metric, indicating

that the recovered SI intrinsic structures with Mixture-Net are improved over the state-of-the-art methods as observed in Fig. 30. Visual RGB mapping of the denoised Pavia University for a noise level of  $\sigma = 50/255$ . Notice in the zoomed version that Mixture-Net suppresses the Gaussian noise in the smooth regions maintaining the shapes and structures. A qualitative comparison of the spectral quality is shown at the bottom for two spatial pixels **P1** and **P2**.

Table 14

*Denoising Quantitative Results for Pavia University*

Noise	Metric	BM3D	FORPDN	HyRes	DIP	SURE CNN	DS2DP	Mixture -Net
$\frac{100}{255}$	$\uparrow$ PSNR	29.14	26.03	28.41	26.47	<u>29.62</u>	27.54	<b>30.95</b>
	$\downarrow$ SSIM	0.7536	0.5971	0.7379	0.6834	<u>0.8019</u>	0.7187	<b>0.8686</b>
$\frac{50}{255}$	$\uparrow$ PSNR	32.97	30.44	31.78	30.69	<u>33.29</u>	32.12	<b>34.47</b>
	$\downarrow$ SSIM	0.8814	0.7989	0.8549	0.8460	<u>0.9048</u>	0.8963	<b>0.9424</b>
$\frac{25}{255}$	$\uparrow$ PSNR	<u>36.48</u>	34.34	35.35	34.48	36.09	35.55	<b>36.99</b>
	$\downarrow$ SSIM	0.9416	0.9063	0.9271	0.9165	0.9452	<u>0.9511</u>	<b>0.9651</b>

**Single Hyperspectral Super-Resolution.** This experiment employs the Pavia Center and Stuff-Toys datasets for two spatial downsampling factors,  $d = 4$  and  $d = 8$ . The quality was compared against several state-of-the-art methods, including the bicubic interpolation; four deep single gray/RGB image SR methods, EDSR (Lim et al., 2017), RCAN (Zhang et al., 2018c), and SAN (Dai et al., 2019); four data-driven DL single HSI-SR methods, 3DCNN (Mei et al., 2017), GDRRN (Li et al., 2018b), and SSPSR (Jiang et al., 2020); and the non-data-driven Hyperspectral-DIP method (Sidorov and Yngve Hardeberg, 2019). The optimization uses the Adam algorithm (Kingma and Ba, 2014) with a learning rate of  $1e^{-3}$  and all hyper-parameters  $\gamma^k$  equal to 0.5. A different

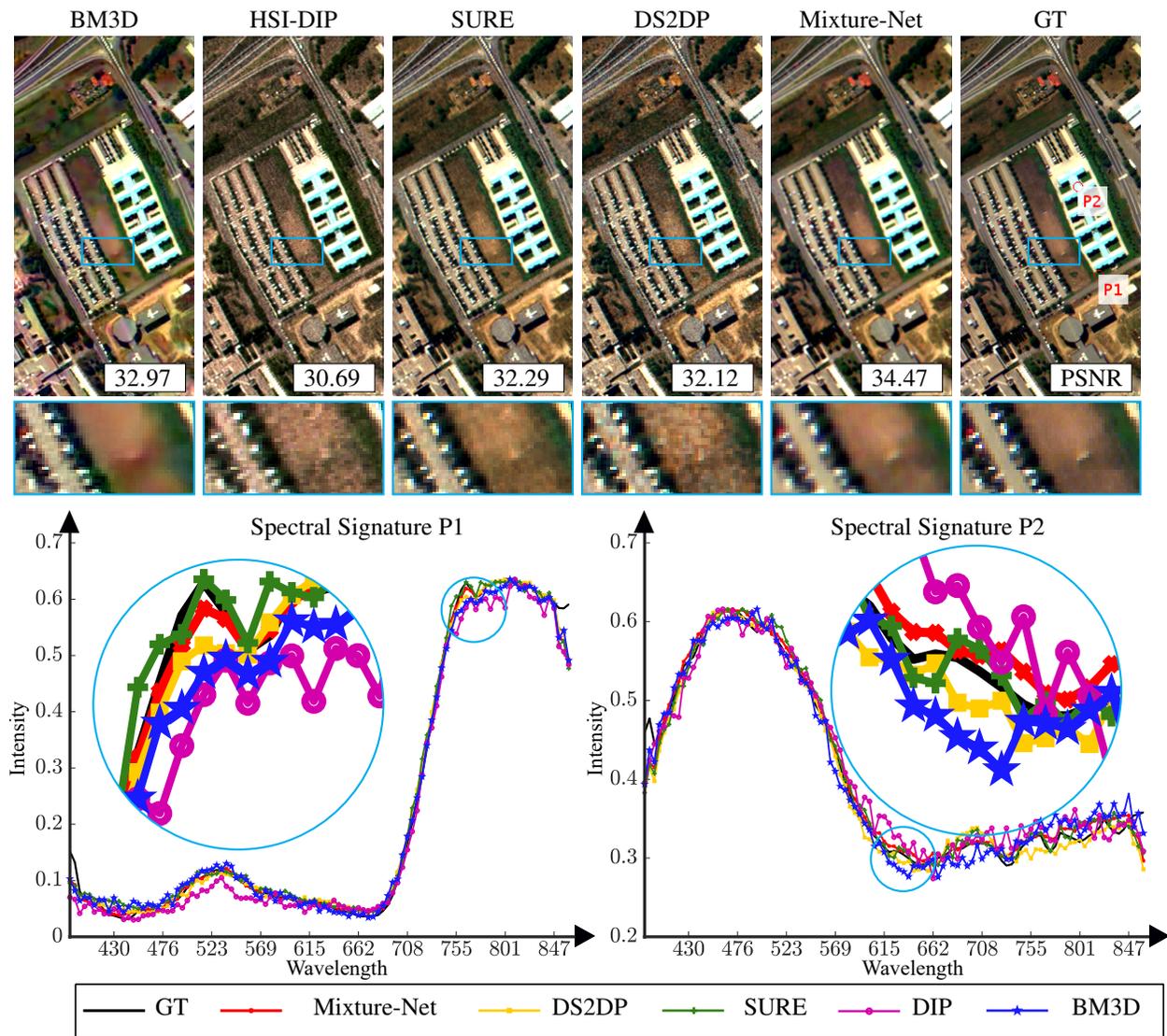


Figure 30. Visual RGB mapping of the denoised Pavia University.

value of the rank  $r$  in (48) was set for each dataset. The number of deep-blocks,  $K$ , were found for each dataset through cross-validation. Precisely, I established  $r = 6$  and  $K = 3$ , and  $r = 12$  and  $K = 4$ , for Pavia Center and Stuff-Toys, respectively.

Tables 15 and 16 compare the quantitative results for the single HSI-SR over the evaluated methods. It can be seen that the non-data-driven Mixture-Net outperforms or achieves competitive

quality even against the data-driven methods such as the SSPSR for both datasets, both downsampling factors, and all evaluated metrics. The spectral quality measured with the SAM metric results remarkably improved for the Pavia Center dataset. Thus, the intrinsic low-rank prior is more substantial when a higher number of correlated spectral bands are considered. Figure 31 shows a visual comparison of the RGB mapping of some super-resolved images for the Pavia Center dataset for spatial downsampling factor  $d = 4$  (*top*) and  $d = 8$  (*bottom*), where the proposed method improves the spatial quality, particularly for the highest down-sampling factor  $d = 8$ .

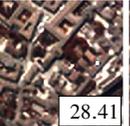
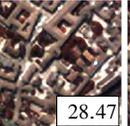
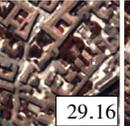
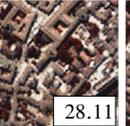
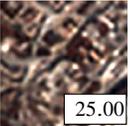
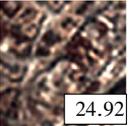
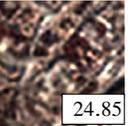
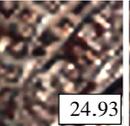
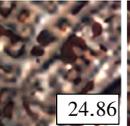
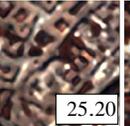
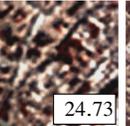
	Ground truth	EDSR	RCAN	SAN	3DCNN	GDRRN	SSPSR	DIP	Mixture-Net
$d = 4$									
	PSNR	28.80	28.82	28.86	28.41	28.47	29.16	28.11	<b>29.91</b>
$d = 8$									
	PSNR	25.00	24.92	24.85	24.93	24.86	25.20	24.73	<b>26.90</b>

Figure 31. RGB representation of the reconstructed composite Pavia Center dataset

**Compressive Spectral Imaging Reconstruction.** This experiment compares the effectiveness of Mixture-Net for CSI reconstruction against data-driven and non-data-driven state-of-the-art methods, including the non-data-driven Plug-and-Play (PnP) (Yuan et al., 2020) and TL-DIP (Bacca et al., 2021) methods, and the data-driven Deep Non-local Unrolling (DNU) (Wang et al., 2020b) and Autoencoder (AE) (Choi et al., 2017a) methods. This experiment employs the KAIST and the ARAD datasets, where the DNU and AE methods were trained with the publicly available codes using a training dataset of 27 and 450 scenes for the KAIST and ARAD datasets,

Table 15

*Single Hyperspectral Super-Resolution Quantitative Results for Pavia Center*

Method	$d$	SAM↓	RMSE↓	ERGAS↓	PSNR↑	SSIM ↑
Bicubic	4	6.1399	0.0437	6.8814	27.5874	0.6961
EDSR	4	5.8657	0.0379	6.0199	28.7981	0.7722
RCAN	4	5.9785	0.0376	6.0485	28.8165	0.7719
SAN	4	5.9590	0.0374	5.9903	28.8554	0.7740
3DCNN	4	5.8669	0.0396	6.2665	28.4114	0.7501
GDRRN	4	5.4750	0.0393	6.2264	28.4726	0.7530
SSPSR	4	<u>5.4612</u>	<u>0.0362</u>	<b>5.8014</b>	<u>29.1581</u>	<u>0.7903</u>
DIP	4	6.2665	0.0410	6.4845	28.1061	0.7365
Mixture-Net	4	<b>4.2120</b>	<b>0.0352</b>	<u>5.8084</u>	<b>29.914</b>	<b>0.8396</b>
Bicubic	8	7.8478	0.0630	4.8280	24.5972	0.4725
EDSR	8	7.8594	0.05983	4.6359	25.0041	0.5130
RCAN	8	7.9992	0.0604	4.6930	24.9183	0.5086
SAN	8	8.0371	0.0609	4.7646	24.8485	0.5054
3DCNN	8	7.6878	0.0605	4.6469	24.9336	0.5038
GDRRN	8	7.3531	0.0607	4.6220	24.8648	0.5014
SSPSR	8	<u>7.3312</u>	<u>0.0586</u>	<u>4.5266</u>	<u>25.1985</u>	<u>0.5365</u>
DIP	8	7.9281	0.0618	4.7366	24.7252	0.4963
Mixture-Net	8	<b>6.7855</b>	<b>0.0485</b>	<b>4.0015</b>	<b>26.9041</b>	<b>0.7148</b>

respectively.

Figure 32 shows a visual representation of the estimations with the reconstruction quality measured in PSNR and SSIM using the non-data-driven PnP and DIP methods, and the training-data dependent DNU, and AE methods against the interpretable method Mixture-Net for the KAIST and ARAD datasets. It can be observed that Mixture-Net improves the quality of reconstruction by

Table 16

*Single Hyperspectral Super-Resolution Quantitative Results for Stuff-Toys*

Method	$d$	SAM↓	RMSE↓	ERGAS↓	PSNR↑	SSIM ↑
Bicubic	4	4.1759	0.0212	5.2719	34.7214	0.9277
EDSR	4	3.5499	0.0149	3.5921	38.1575	0.9522
RCAN	4	3.6050	0.0142	3.4178	38.7585	0.9530
SAN	4	3.5951	0.0143	3.4200	38.7188	0.9531
3DCNN	4	3.3463	0.0154	3.7042	37.9759	0.9522
GDRRN	4	<u>3.4143</u>	0.0145	3.5086	38.4507	0.9538
SSPSR	4	<b>3.1846</b>	0.0138	3.3384	<u>39.0892</u>	0.9553
DIP	4	8.4935	<u>0.0124</u>	<u>2.5358</u>	38.1329	<u>0.9631</u>
Mixture-Net	4	5.3285	<b>0.0110</b>	<b>2.1997</b>	<b>39.1640</b>	<b>0.9821</b>
Bicubic	8	5.8962	0.0346	4.2175	30.2056	0.8526
EDSR	8	5.6865	0.0279	3.3903	32.4072	0.8842
RCAN	8	5.9771	0.0268	3.1781	32.9544	0.8884
SAN	8	5.8683	0.0267	3.1437	33.0012	0.8888
3DCNN	8	<u>5.0948</u>	0.0292	3.5536	31.9691	0.8863
GDRRN	8	<u>5.3597</u>	0.0280	3.3460	32.5763	0.8890
SSPSR	8	<b>4.4874</b>	0.0257	3.0419	<u>33.4340</u>	0.9010
DIP	8	8.3342	<u>0.0231</u>	<u>2.3697</u>	32.7324	<u>0.9291</u>
Mixture-Net	8	5.5027	<b>0.0208</b>	<b>2.1061</b>	<b>33.6270</b>	<b>0.9432</b>

up to 7 [dB] and obtains the highest SSIM outperforming even the methods employing data during the training step. Figure 32, at right, shows the spectral reconstruction of a random spatial location, illustrating that Mixture-Net improves the spectral response estimation.

**5.3.3. High-level Tasks Experiments.** This experiment aims to evaluate the potential application of Mixture-Net to perform high-level tasks such as unmixing and material identi-

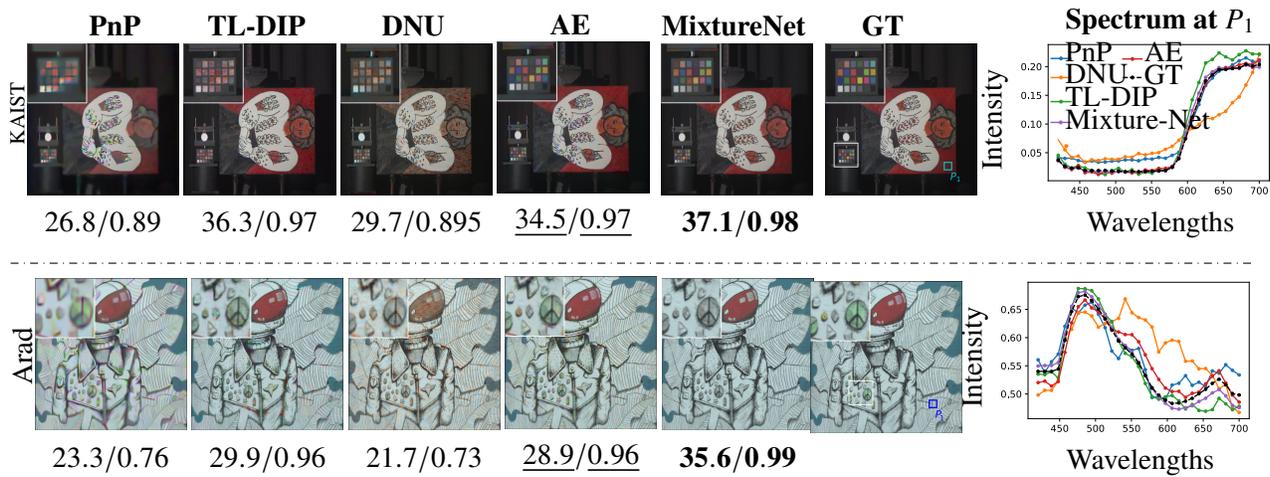


Figure 32. Recovery quality comparison.

fication without requiring to run complementary routines. For this, Mixture-Net takes advantage of the structure inspired by mixture models, implicitly providing two outputs interpreted as the corresponding abundances and endmembers.

A single shot of the Dual Disperser Coded Aperture Snapshot Spectral Imaging (DD-CASSI) (Gehm et al., 2007) and the Jasper dataset were employed for the unmixing experiment. Figure 33 shows the unmixing high-level task application resulting from the interpretable weights and features of the proposed deep model. It can be seen a false RGB representation of Jasper's ground truth and Mixture-Net estimation, and the interpreted endmember. and abundance matrices. Notice that the learned features in the abundance block-layer converge to a rough estimation of the abundances. Similarly, the adjusted weights in the endmember block-layer converge to a rough estimation of the endmembers. It demonstrates that Mixture-Net could be used for the unmixing problem at any additional cost.

A single shot of the DD-CASSI is employed and the Fake image containing a real and a

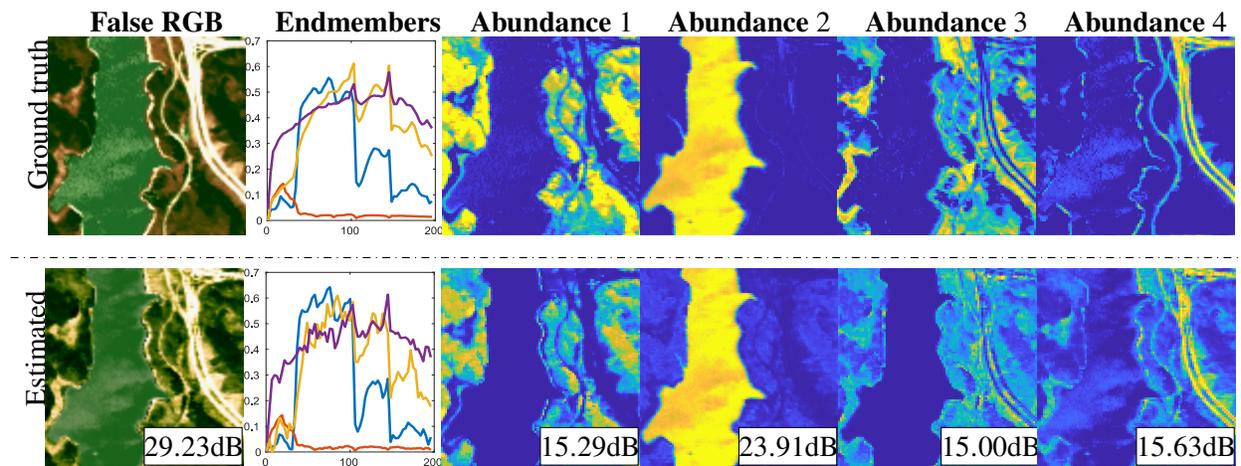


Figure 33. Unmixing application.

fake plastic fruit with similar shape and color for the material identification experiment. The image is spatially sub-sampled to  $256 \times 256$  pixels. Figure 34 shows the (*left*) ground truth and a false RGB mapping of the Mixture-Net estimation with  $r = 15$ . At the top-right, the figure shows two of the obtained features that can be interpreted as abundances. A thresholding is applied over the abundances and show the binary maps at the bottom-right to determine if the materials of both objects are the same. The color checker help shows that the obtained abundance can determine that the materials are different, identifying the fake object without running additional routines.

#### 5.4. Conclusions

Unlike the convex optimization framework employed in previous chapters, this chapter addressed the low-rank incorporation for SI recovery under a DL framework. Specifically, I presented a non-data-driven SI recovery method based on DIP, where the low-rank mixture models inspired the network dubbed Mixture-Net. Beyond previous literature methods, the introduced Mixture-Net implicitly incorporated the low-rankness prior knowledge in the network architecture, addressing

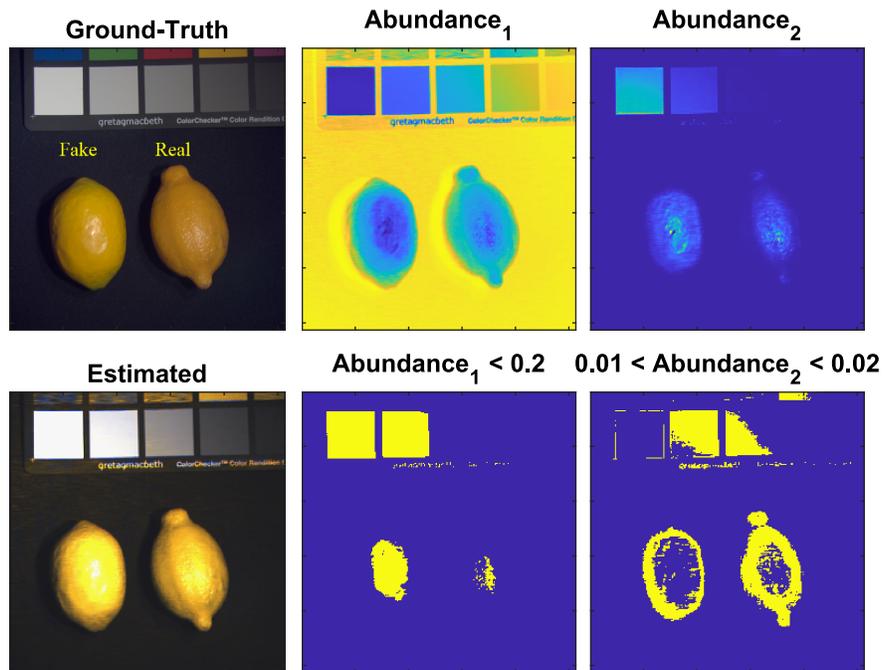


Figure 34. Material identification application

the black-box nature of standard DL. The proposal was divided into three components, the input, imposing a low-dimensional structure through the learning of a Tucker decomposition; the interpretable Mixture-Net, following a sequence of multiple deep-blocks to estimate the abundances, the endmembers, and the SI with non-linear relations; and the custom loss function that considers a regularization term related to physical constraints. The non-data-driven Mixture-Net effectiveness was demonstrated over three main SI recovery problems: SI denoising, single HSI-SR, and CSI reconstruction in terms of different metrics, outperforming even data-driven methods requiring the training of a vast amount of data. Along with the experiments, I noted that imposing a low-dimensional structure over the input improves the quality of the recovered image. On the other hand, the non-linearity block-layer drastically improved the obtained quality by considering the

non-linear relationships between the endmembers. The extension of Mixture-Net to address any other SI recovery task is suggested for future works. Remark that Mixture-Net supports the use of different loss functions, where the interpretable advantage enabled to perform further high-level tasks as linear unmixing and material identification without using additional routines.

## Conclusions

The objective of this dissertation was to develop approaches based on low-rank regularization to recover a SI from a set of random spectral observations. Therefore, the thesis was composed of three main fronts relating to current problems found in the literature: (i) the analysis of the low-rank property exhibited by natural SIs across the spatial and spectral dimensions, (ii) the formulation of regularization strategies to effectively take advantage of the SI low-rankness, and (iii) the design of practical algorithms based on the low-rank regularization to improve the SI recovery quality.

The low-rank prior was analyzed through theoretical premises and empirical studies, where the low-rankness was found to have a different connotation along the spatial and spectral dimensions. Fulfilling the first specific objective, *“to analyze the low-rank properties of SIs to validate the assumption that a low-rank regularization could improve the performance of the spectral inverse problem from encoded random projections,”* Section 2.2 analyzed the low-rankness as a global property cause of the high redundancy of a few spectral signatures repeated at several spatial locations. Section 3.2 analyzed the low-rankness as a non-local property across the spatial dimension cause of structural self-similarities. Beyond the spatial and spectral dimensions, Section 4.2 analyzed the low-rankness across a proposed similarity domain, assuming that the sorting of similar blocks can be shared among the spectral bands. The introduced rank-one similarity prior was validated through an extensive empirical study over several real and synthetic spectral datasets, suggesting that the rank-one similarity prior is strongest over hyperspectral remote sensing datasets. The analysis results supported the inclusion of the low-rank prior to address ill-posed SI inverse problems.

The low-rank prior was mainly regularized through hand-crafted regularization functions and model-based premises incorporated in traditional convex optimization formulations. Fulfilling the second and third specific objectives “*To design two regularization terms to measure the SI low-rank characteristic*” and “*To formulate an optimization problem which includes the low-rank regularization terms to solve the spectral inverse problem from encoded random projections,*” Section 2.3 incorporated the global spectral low-rank prior through the nuclear norm. Section 3.3 incorporated the global spectral low-rank prior through the LMM and the non-local low-rank prior through a proposed hand-crafted regularization term that promoted self-similarities between small extracted patches. Differently, Section 4.3 combined the global spectral low-rank, non-local self-similarities, and the rank-one similarity prior through an implicit regularization term. Contrary to the traditional regularization with an explicit or implicit function, Section 5.2 incorporated the global spectral low-rank prior into the architecture of a proposed neural network, providing a new strategy to consider prior information into general DL-based approaches.

Fulfilling the fourth specific objective, “*To design a numerical algorithm to solve the proposed spectral inverse optimization problem,*” the low-rank regularization was incorporated in four developed practical algorithms covering the convex optimization with global optimums and DL frameworks to solve ill-posed SI recovery problem. Sections 2.4, 3.4, and 5.3 tackled inverse problems from the compressive acquisition approach, providing a competitive alternative to traditional sparse and transform-based approaches. Notice that, fulfilling the fifth specific objective, “*To verify the developed algorithm by using two state-of-the-art random encoder projectors,*” the SI was recovered using CASSI and SSCSI compressed random spectral projections. Further, Sections 4.4

and 5.3 addressed the HSI-MSI fusion and single HSI-SR inverse problems, providing strategies to alleviate the current technology limitation in the acquisition of a high-spatial-spectral resolution image.

In perspective, there is still room for improvement when applying the low-rank prior to solve SI recovery problems. Specifically, future works can focus on: (i) to determine in which dimension the low-rank property results more substantial for spectral imaging analysis. This research could significantly facilitate the selection of the regularization term promoting the low-rankness, alleviating the formulation of complex models combining several low-rank priors. It could be expected to find that the low-rankness has a different relevance according to the nature of the SI, i.e., depending on the spatial-spectral resolution and the intrinsic structures in the present objects. (ii) to extend the use of the novel rank-one similarity prior. This dissertation specifically used the rank-one similarity prior to solve the HSI-MSI fusion problem. Nonetheless, the concept of the proposed prior could be extended to many more applications involving SIs. It could be expected that the rank-one similarity prior will improve the obtained quality, especially in applications facing limitations at recovering the spectral details at low-contrast regions, for instance, SI denoising. (iii) to extrapolate the idea of implicitly incorporating prior information into a neural network architecture to promote SI prior information. This dissertation specifically addressed the global spectral low-rank prior by proposing two layers representing the theoretical endmembers and abundances. Therefore, it could be expected that the non-local self-similarities and rank-one similarity prior can also be incorporated through a combination of proper layers in a neural network.

### References

- Afonso, M. V., Bioucas-Dias, J. M., and Figueiredo, M. A. (2010). An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Transactions on Image Processing*, 20(3):681–695.
- Afonso, M. V., Bioucas-Dias, J. M., and Figueiredo, M. A. (2011). An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Transactions on Image Processing*, 20(3):681–695.
- Akhtar, N., Shafait, F., and Mian, A. (2014). Sparse spatio-spectral representation for hyperspectral image super-resolution. In *European conference on computer vision*, pages 63–78. Springer.
- Akhtar, N., Shafait, F., and Mian, A. (2015). Bayesian sparse representation for hyperspectral image super resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3631–3640.
- Akhtar, N., Shafait, F., and Mian, A. (2016). Hierarchical beta process with gaussian process prior for hyperspectral image super resolution. In *European Conference on Computer Vision*, pages 103–120. Springer.
- Altmann, Y., Dobigeon, N., and Tourneret, J.-Y. (2014). Unsupervised post-nonlinear unmixing of hyperspectral images using a hamiltonian monte carlo algorithm. *IEEE Transactions on Image Processing*, 23(6):2663–2675.

- Amigo, J. M. and Santos, C. (2020). Preprocessing of hyperspectral and multispectral images. In *Data Handling in Science and Technology*, volume 32, pages 37–53. Elsevier.
- Arad, B., Timofte, R., Ben-Shahar, O., Lin, Y.-T., and Finlayson, G. D. (2020). Ntire 2020 challenge on spectral reconstruction from an rgb image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 446–447.
- Arguello, H. and Arce, G. R. (2014). Colored coded aperture design by concentration of measure in compressive spectral imaging. *IEEE Transactions on Image Processing*, 23(4):1896–1908.
- Bacca, J., Correa, C. V., and Arguello, H. (2019). Noniterative hyperspectral image reconstruction from compressive fused measurements. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(4):1231–1239.
- Bacca, J., Fonseca, Y., and Arguello, H. (2021). Compressive spectral image reconstruction using deep prior and low-rank tensor representation. *arXiv preprint arXiv:2101.07424*.
- Barajas, C., Ramirez, J.-M., and Arguello, H. (2020). Convolutional sparse coding framework for compressive spectral imaging. *Journal of Visual Communication and Image Representation*, 66:102690.
- Bioucas-Dias, J. M. and Figueiredo, M. A. (2007). A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transactions on Image processing*, 16(12):2992–3004.

- Bioucas-Dias, J. M. and Nascimento, J. M. (2008). Hyperspectral subspace identification. *IEEE Transactions on Geoscience and Remote Sensing*, 46(8):2435–2445.
- Bioucas-Dias, J. M., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N., and Chanussot, J. (2013). Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and remote sensing magazine*, 1(2):6–36.
- Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., and Chanussot, J. (2012). Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE journal of selected topics in applied earth observations and remote sensing*, 5(2):354–379.
- Blumensath, T., Davies, M. E., Rilling, G., Eldar, Y., and Kutyniok, G. (2012). Greedy algorithms for compressed sensing.
- Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122.
- Brady, D. J. (2009). *Optical Imaging and Spectroscopy*. John Wiley Sons, Inc, Hoboken, NJ, 1 edition.
- Cai, J.-F., Candès, E. J., and Shen, Z. (2010). A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982.

- Candes, E. J. and Recht, B. (2008). Exact low-rank matrix completion via convex optimization. In *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, pages 806–812.
- Candès, E. J. and Wakin, M. B. (2008). An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30.
- Cao, C., Yu, J., Zhou, C., Hu, K., Xiao, F., and Gao, X. (2019a). Hyperspectral image denoising via subspace-based nonlocal low-rank and sparse factorization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(3):973–988.
- Cao, X., Xu, Z., and Meng, D. (2019b). Spectral-spatial hyperspectral image classification via robust low-rank feature extraction and markov random field. *Remote Sensing*, 11(13):1565.
- Chan, S. H., Wang, X., and Elgendy, O. A. (2016). Plug-and-play admm for image restoration: Fixed-point convergence and applications. *IEEE Transactions on Computational Imaging*, 3(1):84–98.
- Chang, Y., Yan, L., Chen, B., Zhong, S., and Tian, Y. (2020a). Hyperspectral image restoration: Where does the low-rank property exist. *IEEE Transactions on Geoscience and Remote Sensing*.
- Chang, Y., Yan, L., Zhao, X.-L., Fang, H., Zhang, Z., and Zhong, S. (2020b). Weighted low-rank tensor recovery for hyperspectral image restoration. *IEEE transactions on cybernetics*, 50(11):4558–4572.

- Chen, Y., Cao, X., Zhao, Q., Meng, D., and Xu, Z. (2017). Denoising hyperspectral image with non-iid noise structure. *IEEE transactions on cybernetics*, 48(3):1054–1066.
- Chen, Y., He, W., Yokoya, N., and Huang, T.-Z. (2019). Hyperspectral image restoration using weighted group sparsity-regularized low-rank tensor decomposition. *IEEE transactions on cybernetics*, 50(8):3556–3570.
- Chen, Y., Wang, S., and Zhou, Y. (2018). Tensor nuclear norm-based low-rank approximation with total variation regularization. *IEEE Journal of Selected Topics in Signal Processing*, 12(6):1364–1377.
- Chen, Z., Pu, H., Wang, B., and Jiang, G. M. (2014). Fusion of hyperspectral and multispectral images: A novel framework based on generalization of pan-sharpening methods. *IEEE Geoscience and Remote Sensing Letters*, 11(8):1418–1422.
- Choi, I., Jeon, D. S., Nam, G., Gutierrez, D., and Kim, M. H. (2017a). High-quality hyperspectral reconstruction using a spectral prior. *ACM Transactions on Graphics (TOG)*, 36(6):1–13.
- Choi, I., Jeon, D. S., Nam, G., Gutierrez, D., and Kim, M. H. (2017b). High-quality hyperspectral reconstruction using a spectral prior. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia 2017)*, 36(6):218:1–13.
- Cohen, J. E. and Gillis, N. (2017). Spectral unmixing with multiple dictionaries. *IEEE Geoscience and Remote Sensing Letters*, 15(2):187–191.

- Correa, C. V., Hinojosa, C., Arce, G. R., and Arguello, H. (2016). Multiple snapshot colored compressive spectral imager. *Optical Engineering*, 56(4):041309.
- Correa-Pugliese, C. V., Galvis-Carreño, D. F., and Arguello-Fuentes, H. (2016). Sparse representations of dynamic scenes for compressive spectral video sensing. *Dyna*, 83(195):42–51.
- Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K. (2007). Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095.
- Dai, T., Cai, J., Zhang, Y., Xia, S.-T., and Zhang, L. (2019). Second-order attention network for single image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11065–11074.
- Dian, R. and Li, S. (2019). Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization. *IEEE Transactions on Image Processing*, 28(10):5135–5146.
- Dian, R., Li, S., Fang, L., and Bioucas-Dias, J. (2018). Hyperspectral image super-resolution via local low-rank and sparse representations. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 4003–4006. IEEE.
- Dian, R., Li, S., Sun, B., and Guo, A. (2021). Recent advances and new guidelines on hyperspectral and multispectral image fusion. *Information Fusion*, 69:40–51.
- Dong, W., Fu, F., Shi, G., Cao, X., Wu, J., Li, G., and Li, X. (2016). Hyperspectral image super-

- resolution via non-negative structured sparse representation. *IEEE Transactions on Image Processing*, 25(5):2337–2352.
- Fan, S., Li, C., Huang, W., and Chen, L. (2018). Data fusion of two hyperspectral imaging systems with complementary spectral sensing ranges for blueberry bruising detection. *Sensors*, 18(12):4463.
- Fang, L., Zhuo, H., and Li, S. (2018). Super-resolution of hyperspectral image via superpixel-based sparse representation. *Neurocomputing*, 273:171–177.
- Figueiredo, M., Nowak, R., and Wright, S. (2007a). Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):586–597.
- Figueiredo, M. A., Nowak, R. D., and Wright, S. J. (2007b). Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of selected topics in signal processing*, 1(4):586–597.
- Fu, Y., Zheng, Y., Sato, I., and Sato, Y. (2016). Exploiting spectral-spatial correlation for coded hyperspectral image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3727–3736.
- Gao, L., Hong, D., Yao, J., Zhang, B., Gamba, P., and Chanussot, J. (2020). Spectral superresolution of multispectral imagery with joint sparse and low-rank learning. *IEEE Transactions on Geoscience and Remote Sensing*, 59(3):2269–2280.

- Gehm, M., John, R., Brady, D., Willett, R., and Schulz, T. (2007). Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics express*, 15(21):14013–14027.
- Gelvez, T. and Arguello, H. (2018). Spectral image fusion from compressive projections using total-variation and low-rank regularizations. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 1985–1989. IEEE.
- Gelvez, T. and Arguello, H. (2020). Nonlocal low-rank abundance prior for compressive spectral image fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 59(1):415–425.
- Gelvez, T., Bacca, J., and Arguello, H. (2021a). Interpretable deep image prior method inspired in linear mixture model for compressed spectral image recovery. In *2021 IEEE International Conference on Image Processing (ICIP)*, pages 1934–1938. IEEE.
- Gelvez, T., Bacca, J., and Arguello, H. (2021b). Interpretable multiple loss functions in a low-rank deep image prior based method for single hyperspectral image super-resolution. In *2021 29th European Signal Processing Conference (EUSIPCO)*, pages 696–700. IEEE.
- Gelvez, T., Rueda, H., and Arguello, H. (2017). Joint sparse and low rank recovery algorithm for compressive hyperspectral imaging. *Applied optics*, 56(24):6785–6795.
- Gélvez, T. C., Rueda, H. F., and Arguello, H. (2016). Simultaneously sparse and low-rank hyperspectral image recovery from coded aperture compressive measurements via convex optimization. In *SPIE Defense+ Security*, pages 98401J–98401J. International Society for Optics and Photonics.

- Gelvez-Barrera, T. (2021). Hyperspectral datasets ONLINE. <https://github.com/TatianaGelvez/Datasets.git>.
- Gogna, A., Shukla, A., Agarwal, H., and Majumdar, A. (2014). Split bregman algorithms for sparse/joint-sparse and low-rank signal recovery: Application in compressive hyperspectral imaging. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 1302–1306. IEEE.
- Golbabaee, M. and Vandergheynst, P. (2012). Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2741–2744. Ieee.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.
- Guilloteau, C., Oberlin, T., Berné, O., and Dobigeon, N. (2020). Hyperspectral and multispectral image fusion under spectrally varying spatial blurs—application to high dimensional infrared astronomical imaging. *IEEE Transactions on Computational Imaging*, 6:1362–1374.
- Han, X.-H., Shi, B., and Zheng, Y. (2018). Self-similarity constrained sparse representation for hyperspectral image super-resolution. *IEEE Transactions on Image Processing*, 27(11):5625–5637.
- Hasanlou, M. and Seydi, S. T. (2018). Hyperspectral change detection: An experimental comparative study. *International journal of remote sensing*, 39(20):7029–7083.

- He, W., Yao, Q., Li, C., Yokoya, N., Zhao, Q., Zhang, H., and Zhang, L. (2020). Non-local meets global: An iterative paradigm for hyperspectral image restoration. *arXiv preprint arXiv:2010.12921*.
- Heylen, R., Parente, M., and Gader, P. (2014). A review of nonlinear hyperspectral unmixing methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):1844–1868.
- Hinojosa, C., Bacca, J., and Arguello, H. (2018). Coded aperture design for compressive spectral subspace clustering. *IEEE Journal of Selected Topics in Signal Processing*, 12(6):1589–1600.
- Jia, S. and Qian, Y. (2009). Constrained nonnegative matrix factorization for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 47(1):161–173.
- Jia, Y., Feng, Y., and Wang, Z. (2015). Reconstructing hyperspectral images from compressive sensors via exploiting multiple priors. *Spectroscopy Letters*, 48(1):22–26.
- Jian, W., Chunxia, Q., Xiufei, Z., Ke, Y., and Ping, R. (2020). A multi-source image fusion algorithm based on gradient regularized convolution sparse representation. *Journal of Systems Engineering and Electronics*, 31(3):447–459.
- Jiang, J., Sun, H., Liu, X., and Ma, J. (2020). Learning spatial-spectral prior for super-resolution of hyperspectral imagery. *IEEE Transactions on Computational Imaging*, 6:1082–1096.
- Junjun-Jiang (2020). Hyperspectral image super resolution benchmark ONLINE. <https://github.com/junjun-jiang/Hyperspectral-Image-Super-Resolution-Benchmark>.

- Kanatsoulis, C. I., Fu, X., Sidiropoulos, N. D., and Ma, W.-K. (2018). Hyperspectral super-resolution: A coupled tensor factorization approach. *IEEE Transactions on Signal Processing*, 66(24):6503–6517.
- Kervrann, C. and Boulanger, J. (2006). Unsupervised patch-based image regularization and representation. In *European Conference on Computer Vision*, pages 555–567. Springer.
- Khan, M. J., Khan, H. S., Yousaf, A., Khurshid, K., and Abbas, A. (2018). Modern trends in hyperspectral image analysis: a review. *IEEE Access*, 6:14118–14129.
- Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- León-López, K. M. and Fuentes, H. A. (2020). Online tensor sparsifying transform based on temporal superpixels from compressive spectral video measurements. *IEEE Transactions on Image Processing*, 29:5953–5963.
- Li, F., Wang, J., Lan, R., Liu, Z., and Luo, X. (2019a). Hyperspectral image classification using multi-feature fusion. *Optics & Laser Technology*, 110:176–183.
- Li, J., Liu, X., Yuan, Q., Shen, H., and Zhang, L. (2019b). Antinoise hyperspectral image fusion by mining tensor low-multilinear-rank and variational properties. *IEEE Transactions on Geoscience and Remote Sensing*, 57(10):7832–7848.
- Li, P., Chen, W., and Ng, M. K. (2020a). Compressive total variation for image reconstruction and restoration. *Computers & Mathematics with Applications*, 80(5):874–893.

- Li, S., Dian, R., Fang, L., and Bioucas-Dias, J. M. (2018a). Fusing hyperspectral and multispectral images via coupled sparse tensor factorization. *IEEE Transactions on Image Processing*, 27(8):4118–4130.
- Li, X., Yuan, Y., and Wang, Q. (2020b). Hyperspectral and multispectral image fusion via nonlocal low-rank tensor approximation and sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*.
- Li, Y., Zhang, L., Dingl, C., Wei, W., and Zhang, Y. (2018b). Single hyperspectral image super-resolution with grouped deep recursive residual network. In *2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM)*, pages 1–4. IEEE.
- Lim, B., Son, S., Kim, H., Nah, S., and Mu Lee, K. (2017). Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 136–144.
- Lin, B., Tao, X., Xu, M., Dong, L., and Lu, J. (2017). Bayesian hyperspectral and multispectral image fusions via double matrix factorization. *IEEE Transactions on Geoscience and Remote Sensing*, 55(10):5666–5678.
- Lin, X., Liu, Y., Wu, J., and Dai, Q. (2014). Spatial-spectral encoded compressive hyperspectral imaging. *ACM Transactions on Graphics (TOG)*, 33(6):233.
- Liu, H., Sun, P., Du, Q., Wu, Z., and Wei, Z. (2018). Hyperspectral image restoration based on

- low-rank recovery with a local neighborhood weighted spectral–spatial total variation model. *IEEE Transactions on Geoscience and Remote Sensing*, 57(3):1409–1422.
- Liu, J., Wu, Z., Xiao, L., Sun, J., and Yan, H. (2020). A truncated matrix decomposition for hyperspectral image super-resolution. *IEEE Transactions on Image Processing*, 29:8028–8042.
- Lloyd, S. (1982). Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137.
- Lu, B., He, Y., and Dao, P. D. (2019). Comparing the performance of multispectral and hyperspectral images for estimating vegetation properties. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(6):1784–1797.
- M Graña, MA Veganzons, B. A. (2008). Hyperspectral remote sensing scenes ONLINE. [http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral\\_Remote\\_Sensing\\_Scenes](http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes).
- Ma, F., Yang, F., and Wang, Y. (2020). Low-rank tensor decomposition with smooth and sparse regularization for hyperspectral and multispectral data fusion. *IEEE Access*, 8:129842–129856.
- Maggioni, M., Katkovnik, V., Egiazarian, K., and Foi, A. (2012). Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE transactions on image processing*, 22(1):119–133.
- Mäkinen, Y., Azzari, L., and Foi, A. (2020). Collaborative filtering of correlated noise: Exact transform-domain variance for improved shrinkage and patch matching. *IEEE Transactions on Image Processing*, 29:8339–8354.

- Mei, S., Hou, J., Chen, J., Chau, L.-P., and Du, Q. (2018). Simultaneous spatial and spectral low-rank representation of hyperspectral images for classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(5):2872–2886.
- Mei, S., Yuan, X., Ji, J., Zhang, Y., Wan, S., and Du, Q. (2017). Hyperspectral image spatial super-resolution via 3d full convolutional neural network. *Remote Sensing*, 9(11):1139.
- Miao, Y.-C., Zhao, X.-L., Fu, X., Wang, J.-L., and Zheng, Y.-B. (2021). Hyperspectral denoising using unsupervised disentangled spatio-spectral deep priors. *arXiv preprint arXiv:2102.12310*.
- Mullah, H. U., Deka, B., and Prasad, A. (2020). Fast multi-spectral image super-resolution via sparse representation. *IET Image Processing*, 14(12):2833–2844.
- Nascimento, J. M. and Dias, J. M. (2005). Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE transactions on Geoscience and Remote Sensing*, 43(4):898–910.
- Nguyen, H. V., Ulfarsson, M. O., and Sveinsson, J. R. (2020). Sure based convolutional neural networks for hyperspectral image denoising. In *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*, pages 1784–1787. IEEE.
- Nocedal, J. and Wright, S. (2006). *Numerical Optimization*. Springer Series in Operations Research and Financial. Springer New York.
- Ongie, G. and Jacob, M. (2017). A fast algorithm for convolutional structured low-rank matrix recovery. *IEEE transactions on computational imaging*, 3(4):535–550.

- Ongie, G., Jalal, A., Baraniuk, C. A. M. R. G., Dimakis, A. G., and Willett, R. (2020). Deep learning techniques for inverse problems in imaging. *IEEE Journal on Selected Areas in Information Theory*.
- Oymak, S., Jalali, A., Fazel, M., Eldar, Y. C., and Hassibi, B. (2015). Simultaneously structured models with application to sparse and low-rank matrices. *IEEE Transactions on Information Theory*, 61(5):2886–2908.
- Piella, G. (2003). A general framework for multiresolution image fusion: from pixels to regions. *Information fusion*, 4(4):259–280.
- Ramani, S., Blu, T., and Unser, M. (2008). Monte-carlo sure: A black-box optimization of regularization parameters for general denoising algorithms. *IEEE Transactions on image processing*, 17(9):1540–1554.
- Ramírez, J. M., Torre, J. I. M., and Fuentes, H. A. (2021). Ladmm-net: An unrolled deep network for spectral image fusion from compressive data. *arXiv preprint arXiv:2103.00940*.
- Rasti, B., Sveinsson, J. R., Ulfarsson, M. O., and Benediktsson, J. A. (2013). Hyperspectral image denoising using first order spectral roughness penalty in wavelet domain. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):2458–2467.
- Rasti, B., Ulfarsson, M. O., and Ghamisi, P. (2017). Automatic hyperspectral image restoration using sparse and low-rank modeling. *IEEE Geoscience and Remote Sensing Letters*, 14(12):2335–2339.

- Rueda, H., Arguello, H., and Arce, G. R. (2016). Compressive spectral testbed imaging system based on thin-film color-patterned filter arrays. *Applied Optics*, 55(33):9584–9593.
- Sanchez, K., Hinojosa, C., and Arguello, H. (2019). Supervised spatio-spectral classification of fused images using superpixels. *Applied optics*, 58(7):B9–B18.
- Sidorov, O. and Yngve Hardeberg, J. (2019). Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 0–0.
- Simoes, M., Bioucas-Dias, J., Almeida, L. B., and Chanussot, J. (2014). A convex formulation for hyperspectral image superresolution via subspace-based regularization. *IEEE Transactions on Geoscience and Remote Sensing*, 53(6):3373–3388.
- Song, S., Zhou, H., Yang, Y., and Song, J. (2019). Hyperspectral anomaly detection via convolutional neural network and low rank with density-based clustering. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(9):3637–3649.
- Sui, L., Li, L., Li, J., Chen, N., and Jiao, Y. (2019). Fusion of hyperspectral and multispectral images based on a bayesian nonparametric approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(4):1205–1218.
- Teodoro, A., Bioucas-Dias, J., and Figueiredo, M. (2017). Sharpening hyperspectral images using plug-and-play priors. In *International Conference on Latent Variable Analysis and Signal Separation*, pages 392–402.

- Tseng, P. (2001). Convergence of a block coordinate descent method for nondifferentiable minimization. *Journal of optimization theory and applications*, 109(3):475–494.
- Tu, X., Shen, X., Fu, P., Wang, T., Sun, Q., and Ji, Z. (2020). Discriminant sub-dictionary learning with adaptive multiscale superpixel representation for hyperspectral image classification. *Neurocomputing*, 409:131–145.
- Ulyanov, D., Vedaldi, A., and Lempitsky, V. (2018). Deep image prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9446–9454.
- Vaish, A., Gautam, S., and Kumar, M. (2019). A wavelet based approach for simultaneous compression and encryption of fused images. *Journal of King Saud University-Computer and Information Sciences*, 31(2):208–217.
- Vargas, E., Arguello, H., and Tourneret, J. (2019). Spectral image fusion from compressive measurements using spectral unmixing and a sparse representation of abundance maps. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7):5043–5053.
- Vargas, E., Arguello, H., and Tourneret, J.-Y. (2017). Spectral image fusion from compressive measurements using spectral unmixing. In *Proc. of 7th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'17)*, pages 10–13.
- Vargas, E., Espitia, Ó., Arguello, H., and Tourneret, J.-Y. (2018a). Spectral image fusion from compressive measurements. *IEEE Transactions on Image Processing*.

- Vargas, E., Pinilla, S., and Arguello, H. (2018b). A fast endmember estimation algorithm from compressive measurements. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 2210–2214. IEEE.
- Vargas, H., Camacho, A., and Arguello, H. (2019). Spectral unmixing approach in hyperspectral remote sensing: a tool for oil palm mapping. *TecnoLógicas*, 22(45):131–145.
- Vu, H., Cheung, G., and Eldar, Y. C. (2020). Unrolling of deep graph total variation for image denoising. *arXiv preprint arXiv:2010.11290*.
- Wagadarikar, A., John, R., Willett, R., and Brady, D. (2008). Single disperser design for coded aperture snapshot spectral imaging. *Appl. Opt.*, 47(10):B44–B51.
- Wang, K., Wang, Y., Zhao, X.-L., Chan, J. C.-W., Xu, Z., and Meng, D. (2020a). Hyperspectral and multispectral image fusion via nonlocal low-rank tensor decomposition and spectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*.
- Wang, L., Sun, C., Fu, Y., Kim, M. H., and Huang, H. (2019a). Hyperspectral image reconstruction using a deep spatial-spectral prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8032–8041.
- Wang, L., Sun, C., Zhang, M., Fu, Y., and Huang, H. (2020b). Dnu: Deep non-local unrolling for computational spectral imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1661–1671.

- Wang, M., Zhao, M., Chen, J., and Rahardja, S. (2019b). Nonlinear unmixing of hyperspectral data via deep autoencoder networks. *IEEE Geoscience and Remote Sensing Letters*, 16(9):1467–1471.
- Wang, Y., Yin, W., and Zeng, J. (2019c). Global convergence of admm in nonconvex nonsmooth optimization. *Journal of Scientific Computing*, 78(1):29–63.
- Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612.
- Wei, Q., Bioucas-Dias, J., Dobigeon, N., and Tourneret, J.-Y. (2015). Hyperspectral and multispectral image fusion based on a sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*, 53(7):3658–3668.
- Wei, Q., Bioucas-Dias, J., Dobigeon, N., Tourneret, J.-Y., Chen, M., and Godsill, S. (2016a). Multiband image fusion based on spectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12):7236–7249.
- Wei, Q., Dobigeon, N., Tourneret, J.-Y., Bioucas-Dias, J., and Godsill, S. (2016b). R-fuse: Robust fast fusion of multiband images based on solving a sylvester equation. *IEEE signal processing letters*, 23(11):1632–1636.
- Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., and Ma, Y. (2008). Robust face recognition

- via sparse representation. *IEEE transactions on pattern analysis and machine intelligence*, 31(2):210–227.
- Wright, S. J., Nowak, R. D., and Figueiredo, M. A. (2009). Sparse reconstruction by separable approximation. *IEEE Transactions on Signal Processing*, 57(7):2479–2493.
- Wu, R., Ma, W.-K., Fu, X., and Li, Q. (2020a). Hyperspectral super-resolution via global-local low-rank matrix estimation. *IEEE Transactions on Geoscience and Remote Sensing*.
- Wu, R., Wai, H.-T., and Ma, W.-K. (2020b). Hybrid inexact bcd for coupled structured matrix factorization in hyperspectral super-resolution. *IEEE Transactions on Signal Processing*, 68:1728–1743.
- Xu, T., Huang, T.-Z., Deng, L.-J., Zhao, X.-L., and Huang, J. (2020). Hyperspectral image super-resolution using unidirectional total variation with tucker decomposition. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:4381–4398.
- Xu, Y., Wu, Z., Chanussot, J., and Wei, Z. (2019). Nonlocal patch tensor sparse representation for hyperspectral image super-resolution. *IEEE Transactions on Image Processing*, 28(6):3034–3047.
- Xue, J., Leung, Y., and Fung, T. (2017a). A bayesian data fusion approach to spatio-temporal fusion of remotely sensed images. *Remote Sensing*, 9(12):1310.
- Xue, J., Zhao, Y., Liao, W., and Chan, J. C.-W. (2019). Nonlocal tensor sparse representation and

- low-rank regularization for hyperspectral image compressive sensing reconstruction. *Remote Sensing*, 11(2):193.
- Xue, J., Zhao, Y., Liao, W., and Kong, S. G. (2017b). Joint spatial and spectral low-rank regularization for hyperspectral image denoising. *IEEE Transactions on Geoscience and Remote Sensing*, 56(4):1940–1958.
- Yasuma, F., Mitsunaga, T., Iso, D., and Nayar, S. (2008). Generalized Assorted Pixel Camera: Post-Capture Control of Resolution, Dynamic Range and Spectrum. Technical report, Department of Computer Science, Columbia University.
- Yi, C., Zhao, Y.-q., Chan, J. C.-W., and Kong, S. G. (2020). Joint spatial-spectral resolution enhancement of multispectral images with spectral matrix factorization and spatial sparsity constraints. *Remote Sensing*, 12(6):993.
- Yokoya, N., Grohnfeldt, C., and Chanussot, J. (2017). Hyperspectral and multispectral data fusion: A comparative review of the recent literature. *IEEE Geoscience and Remote Sensing Magazine*, 5(2):29–56.
- Yuan, X., Liu, Y., Suo, J., and Dai, Q. (2020). Plug-and-play algorithms for large-scale snapshot compressive imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1447–1457.
- Zeng, H., Xie, X., Cui, H., Yin, H., and Ning, J. (2020). Hyperspectral image restoration via

- global l 1-2 spatial–spectral total variation regularized local low-rank tensor recovery. *IEEE Transactions on Geoscience and Remote Sensing*, 59(4):3309–3325.
- Zhang, H., He, W., Zhang, L., Shen, H., and Yuan, Q. (2014). Hyperspectral image restoration using low-rank matrix recovery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8):4729–4743.
- Zhang, H., Liu, L., He, W., and Zhang, L. (2019). Hyperspectral image denoising with total variation regularization and nonlocal low-rank tensor decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, 58(5):3071–3084.
- Zhang, L., Wei, W., Bai, C., Gao, Y., and Zhang, Y. (2018a). Exploiting clustering manifold structure for hyperspectral imagery super-resolution. *IEEE Transactions on Image Processing*, 27(12):5969–5982.
- Zhang, X., Yuan, X., and Carin, L. (2018b). Nonlocal low-rank tensor factor analysis for image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8232–8241.
- Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., and Fu, Y. (2018c). Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European conference on computer vision (ECCV)*, pages 286–301.
- Zhao, Y.-Q. and Yang, J. (2015). Hyperspectral image denoising via sparse representation and low-rank constraint. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1):296–308.

- Zhou, X., Zhang, Y., Zhang, J., and Shi, S. (2020). Alternating direction iterative nonnegative matrix factorization unmixing for multispectral and hyperspectral data fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:5223–5232.
- Zhou, Y., Feng, L., Hou, C., and Kung, S.-Y. (2017). Hyperspectral and multispectral image fusion based on local low rank and coupled spectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 55(10):5997–6009.
- Zhu, F., Wang, Y., Fan, B., Xiang, S., Meng, G., and Pan, C. (2014a). Spectral unmixing via data-guided sparsity. *IEEE Transactions on Image Processing*, 23(12):5412–5427.
- Zhu, F., Wang, Y., Fan, B., Xiang, S., Meng, G., and Pan, C. (2014b). Spectral unmixing via data-guided sparsity. *IEEE Transactions on Image Processing*, 23(12):5412–5427.
- Zhuang, L. and Bioucas-Dias, J. M. (2018). Fast hyperspectral image denoising and inpainting based on low-rank and sparse representations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(3):730–742.

## **Appendices**

## Appendix A. Spectral Image Datasets

The developed SI recovery methods and low-rank analysis are validated along various publicly available datasets covering daily objects, remote sensing, and satellite sources, and two hyperspectral datasets acquired at the Optics Laboratory from the Universidad Industrial de Santander.

**ARAD hyperspectral database.** Set of 510 images<sup>3</sup> of  $512 \times 512 \times 31$  spatial-spectral resolution taken with the Specim IQ mobile hyperspectral camera spanning the  $0.4 - 0.70 \mu\text{m}$  spectral range in intervals of  $0.01 \mu\text{m}$  Arad et al. (2020). A fake RGB visual representation of four images from the ARAD dataset are depicted in Fig. 35

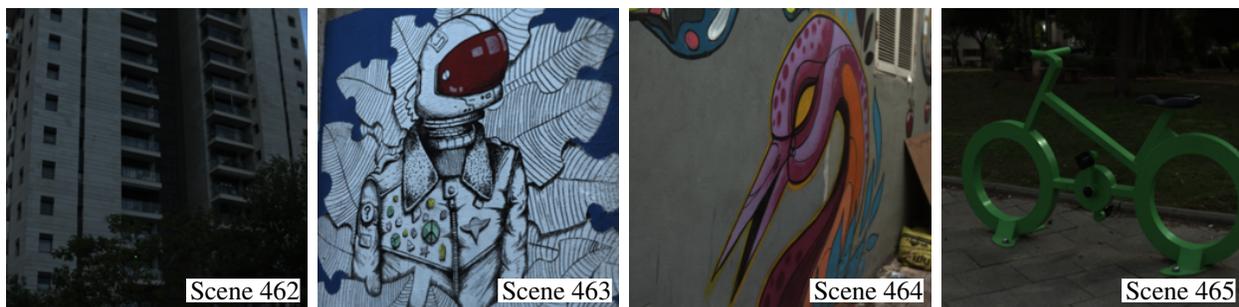


Figure 35. RGB mapping of the ARAD database.

**CAVE multispectral database.** Subset of six images from the standard CAVE dataset<sup>4</sup> of  $512 \times 512 \times 31$  spatial-spectral resolution taken with the Cooled CCD camera spanning the  $0.4 - 0.7 \mu\text{m}$  spectral range in intervals of  $0.01 \mu\text{m}$ . A fake RGB visual representation of four images from the CAVE dataset are depicted in Fig. 36

---

<sup>3</sup> Available in <https://competitions.codalab.org/competitions/22226>.

<sup>4</sup> Available in <https://www.cs.columbia.edu/CAVE/databases/multispectral/>. Accessed: 20-Nov-2020



Figure 36. RGB mapping of the CAVE database.

**KAIST hyperspectral database.** Set of 30 images<sup>5</sup> of  $512 \times 512 \times 31$  spatial-spectral resolution taken with the Pointgrey Grasshopper 9.1MP Monochromatic (GS3-U3-91S6M-C) camera spanning the  $0.4 - 0.72 \mu\text{m}$  spectral range in intervals of  $0.01 \mu\text{m}$  Choi et al. (2017b). A fake RGB visual representation of four images from the KAIST dataset are depicted in Fig. 37

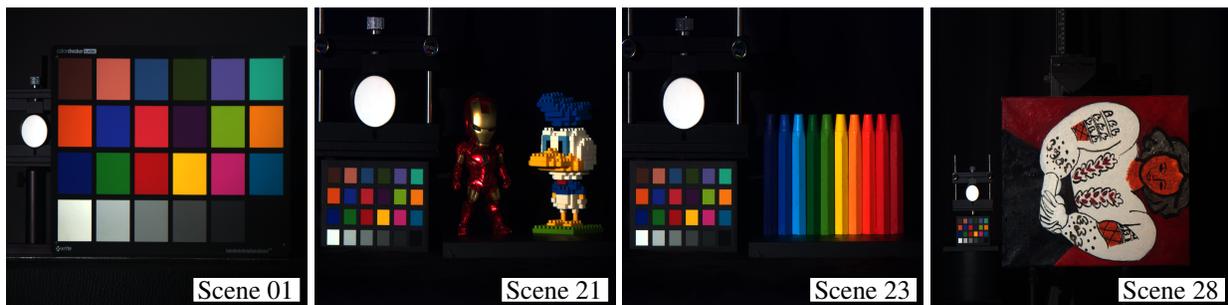


Figure 37. RGB mapping of the KAIST database.

**Botswana.** The Botswana dataset of  $1476 \times 256 \times 145$  spatial-spectral resolution taken with the Hyperion sensor spanning the  $0.4 - 2.5 \mu\text{m}$  spectral range in intervals of  $0.01 \mu\text{m}$ , contains fourteen identified classes representing the land cover types in seasonal swamps, occasional swamps, and drier woodlands.

---

<sup>5</sup> Available in KAIST Dataset. Accessed: 15-Jan-2022.

**China.** The China dataset belongs to a farmland near the city of Yuncheng Jiangsu province in China, which was acquired on 3 May 2006 and 23 April 2007. This scene is mainly a combination of soil, river, tree, building, road, and agricultural field Hasanlou and Seydi (2018).

**Cuprite.** The Cuprite dataset of  $250 \times 190 \times 224$  spatial-spectral resolution taken with AVIRIS spanning the  $0.4 - 2.5 \mu\text{m}$  spectral range in intervals of  $0.01 \mu\text{m}$ , contains information about twelve different minerals. After removing the (1 – 2, 221 – 224, 104 – 113, and 148 – 167) water absorption and noisy bands, a subset of size  $250 \times 190 \times 188$  is employed.

**Jasper Ridge.** The Jasper Ridge dataset of  $512 \times 614 \times 224$  spatial-spectral resolution spanning the ( $0.38 - 2.5 \mu\text{m}$  spectral range in intervals of  $0.0946 \mu\text{m}$  Zhu et al. (2014a), contains information about four materials: Road, Soil, Water, and Tree. After removing the 1 – 3, 108 – 112, 154 – 166 and 220 – 224 noisy bands, a subset of size  $100 \times 100 \times 198$  aligned at the  $(105, 269)^{\text{th}}$  spatial location in the original image is employed.

**KSC.** The Kennedy Space Center dataset (KSC) of  $512 \times 614 \times 224$  spatial-spectral resolution taken with AVIRIS spanning the  $0.4 - 2.5 \mu\text{m}$  spectral range in intervals of  $0.01 \mu\text{m}$  contains information of thirteen classes representing various land cover types. After removing water absorption and noisy bands, a subset of size  $512 \times 614 \times 176$  is employed.

**Pavia Center.** Semi-synthetic Pavia Center dataset of  $1096 \times 1096 \times 102$  spatial-spectral resolution taken with the ROSIS sensor spanning the  $0.43 - 0.86 \mu\text{m}$  spectral range. Following the setup in Jiang et al. (2020) a subset aligned at the top-left of the (1, 670) spatial location of size  $256 \times 256 \times 102$  is employed.

**Pavia University.** Semi-synthetic Pavia University dataset of  $610 \times 340 \times 103$  spatial-spectral resolution taken with ROSIS spanning the  $(0.43 - 0.86) \mu\text{m}$  spectral range. Following the setup in Wang et al. (2020a) removing the water absorption and noisy bands, the top-left subset of size  $256 \times 256 \times 93$  is employed.

**Samson.** The Samson dataset of  $952 \times 952 \times 156$  spatial-spectral resolution spanning the  $0.4 - 0.889 \mu\text{m}$  spectral range contains information about three materials: soil, tree, and water Zhu et al. (2014a). A subset of size  $95 \times 95 \times 156$  aligned at the  $(252, 332)^{\text{th}}$  spatial location in the original image is employed to reduce the computational complexity.

**Urban.** The Urban dataset of  $307 \times 307 \times 210$  spatial-spectral resolution spanning the  $0.4 - 2.5 \mu\text{m}$  spectral range is presented in three versions containing information about four, five, or six materials Zhu et al. (2014a). After removing the 1 – 4, 76, 87, 101 – 111, 136 – 153, and 198 – 210 noisy bands, a subset of size  $307 \times 307 \times 162$  is employed.

**USA.** The USA dataset belongs to an irrigated agricultural field of Hermiston city in Umatilla County, Oregon, OR, the USA, which was acquired with Hyperion sensors on May 1, 2004, and May 8, 2007, respectively. The land cover types are soil, irrigated fields, river, building, type of cultivated land and grassland Hasanlou and Seydi (2018).

A fake RGB and resized visual representation of the above mentioned datasets are depicted in Fig. 38. The datasets are publicly available in <sup>6</sup>.

---

<sup>6</sup> Available in [http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral\\_Remote\\_Sensing\\_Scenes#Pavia\\_Centre\\_and\\_University](http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes#Pavia_Centre_and_University) and <https://rslab.ut.ac.ir/data> Accessed: 15-Jan-2022.

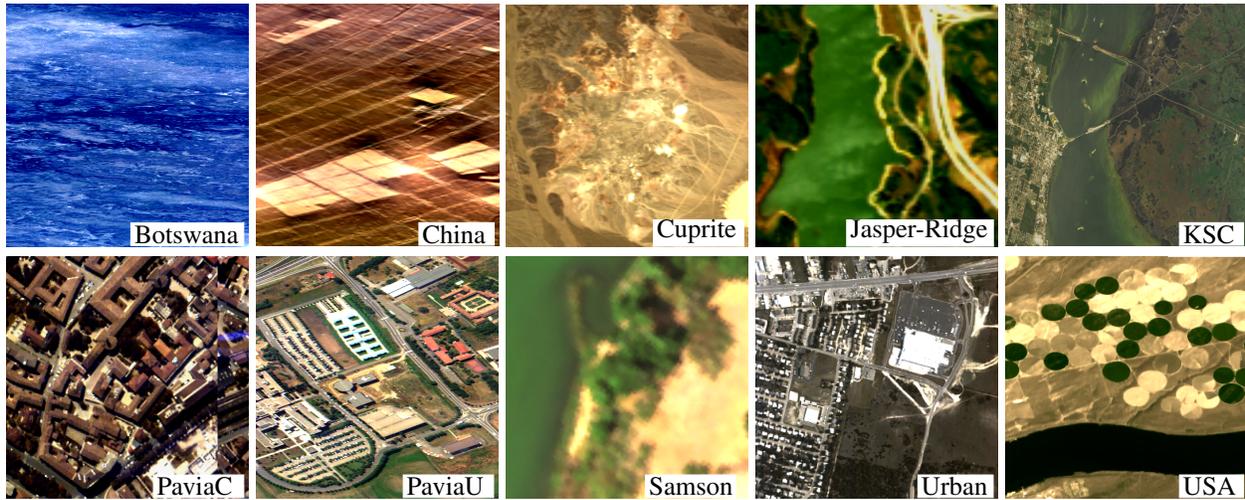


Figure 38. RGB mapping of the remote sensing datasets.

## Appendix B. Quantitative Quality Metrics

The developed methods are evaluated in terms of the following quantitative quality metrics.

**Global Peak-signal-to-noise-ratio.** The peak signal to noise ratio (PSNR), measured in decibels [dB], is defined as the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation, so that, a higher value indicates superior quality of fusion. The global PSNR is calculated as

$$\text{PSNR}(\mathbf{f}, \hat{\mathbf{f}}) = 10 \log_{10} \left( \frac{\max(\mathbf{f})^2}{\text{MSE}(\mathbf{f}, \hat{\mathbf{f}})} \right), \quad (56)$$

where  $\max(\mathbf{f})$  is the maximum possible pixel value of  $\mathbf{f}$ , and MSE stands for the mean square error between the original signal  $\mathbf{f}$  and its estimation  $\hat{\mathbf{f}}$ .

**Spatial peak-signal-to-noise-ratio.** The spatial PSNR is calculated as the average between the  $\text{PSNR}^\ell$ , for  $\ell = 1, \dots, L$  spectral bands. The PSNR for the  $\ell^{\text{th}}$  spectral band  $\mathbf{f}^\ell$  is given

by

$$\text{PSNR}^\ell = 10 \log_{10} \left( \frac{\max(\mathbf{f}^\ell)^2}{\text{MSE}(\mathbf{f}^\ell, \hat{\mathbf{f}}^\ell)} \right). \quad (57)$$

**Spectral peak-signal-to-noise-ratio.** The spectral PSNR is calculated as the average between the  $\text{PSNR}_i$  for  $i = 1, \dots, N^2$ , spatial locations. The PSNR at the  $i^{\text{th}}$  spatial location is given by

$$\text{PSNR}_i = 10 \log_{10} \left( \frac{\max(\mathbf{f}_i)^2}{\text{MSE}(\mathbf{f}_i, \hat{\mathbf{f}}_i)} \right), \quad (58)$$

**Structural Similarity.** The structural similarity SSIM metric allows to measure the local similarity between two images Wang et al. (2004). Unlike the PSNR metric, SSIM considers the image degradation as the perceived change in the structural information, hence, it is considered a perception-based model. The SSIM between sections  $X$  and  $Y$  of an image is calculated as,

$$\text{SSIM}(X, Y) = \frac{(2\mu_X\mu_Y + c_1)(2\text{cov}(X, Y) + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)}, \quad (59)$$

where  $\mu_X$  and  $\mu_Y$  are the average of  $X$  and  $Y$ ,  $\sigma_X^2$  and  $\sigma_Y^2$  are the variance of  $X$  and  $Y$ ,  $\text{cov}(X, Y)$  is the covariance of  $X$  and  $Y$ , and  $c_1$  and  $c_2$  are stabilization variables.

**Universal image quality index.** The universal image quality index (UIQI) measures the amount of transformation of relevant data from reference image into fused image. The closer to 1 indicates that the reference and fused image are more similar. The UIQI is calculated as

$$\text{UIQI}(\mathbf{f}, \hat{\mathbf{f}}) = \frac{4\text{cov}(\mathbf{f}, \hat{\mathbf{f}})\mu_f\mu_{\hat{f}}}{(\sigma_f^2 + \sigma_{\hat{f}}^2)(\mu_f^2 + \mu_{\hat{f}}^2)}. \quad (60)$$

**Spectral angle mapper.** The SAM determines the spectral similarity between two spectra by calculating the angle between two pixels viewed as vectors in a space with dimensionality equal to the number of bands. Smaller angles represent closer matches to the reference spectrum. The global SAM between  $\mathbf{f}$  and  $\hat{\mathbf{f}}$  is calculated as

$$\text{SAM}(\mathbf{f}, \hat{\mathbf{f}}) = \frac{1}{n_m} \sum_{\ell=1}^{n_m} \arccos \left( \frac{\langle \mathbf{z}_\ell, \hat{\mathbf{f}}_\ell \rangle}{\|\mathbf{z}_\ell\|_2 \|\hat{\mathbf{f}}_\ell\|_2} \right). \quad (61)$$

**Relative dimensionless global error:.** The ERGAS computes the quality of the fused image in terms of the normalized average error of each band of processed image. Increase in the value of ERGAS indicates distortion in the fused image, lower value of ERGAS indicates that the fused image is similar to the reference image. The ERGAS is calculated as

$$\text{ERGAS}(\mathbf{f}, \hat{\mathbf{f}}) = 100 * d_h \sqrt{\frac{1}{L} \sum_{\lambda=1}^L \left( \frac{\text{RMSE}(\mathbf{f}^\lambda, \hat{\mathbf{f}}^\lambda)}{\mu_\lambda} \right)^2}, \quad (62)$$

where  $\text{RMSE}(\mathbf{z}^\lambda, \hat{\mathbf{f}}^\lambda)$  denotes the root MSE between the  $\lambda^{\text{th}}$  spectral band of the image and its estimation, and  $\mu_\lambda$  denotes the mean of the  $\lambda^{\text{th}}$  spectral band of the HSI.