



***Università degli Studi di Salerno***

Dipartimento di Ingegneria dell'Informazione, Ingegneria Elettrica e  
Matematica Applicata  
Dottorato di Ricerca in Ingegneria dell'Informazione  
XII Ciclo – Nuova Serie

TESI DI DOTTORATO

# **Multispectral and Hyperspectral Pansharpening: A Critical Examination and New Developments**

CANDIDATO: **GEMINE VIVONE**

TUTOR: **PROF. ROCCO RESTAINO**

COORDINATORE: **PROF. ANGELO MARCELLI**

Anno Accademico 2012 – 2013

*To my dearest Love*

*Ilaria*



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>An Approach to Data Fusion: Pansharpening</b>	<b>7</b>
2.1	Introduction to Pansharpening . . . . .	8
2.2	A Discussion of Assessed Pansharpening Methods . . . . .	11
2.2.1	Notation . . . . .	12
2.2.2	Component Substitution . . . . .	13
	Intensity-Hue-Saturation . . . . .	16
	Principal Component Analysis . . . . .	17
	Gram-Schmidt Orthogonalization . . . . .	18
	Adaptive CS . . . . .	19
2.2.3	Multiresolution Analysis . . . . .	21
	Low Pass Filtering . . . . .	23
	Pyramidal Decompositions . . . . .	24
2.3	Quality Assessment of Fusion Products . . . . .	27
2.3.1	Reduced Scale Assessment . . . . .	28
	Quality Indexes . . . . .	29
2.3.2	Full Scale Validation . . . . .	31
<b>3</b>	<b>Pansharpening Algorithms: A Critical Comparison</b>	<b>33</b>
3.1	The Pansharpening Toolbox . . . . .	34
3.2	Datasets . . . . .	35
3.3	Algorithms . . . . .	37
3.4	Experimental Results . . . . .	38
3.4.1	Wald Protocol . . . . .	38
3.4.2	Full Scale Validation . . . . .	48

3.4.3	Summary and Discussion . . . . .	55
<b>4</b>	<b>The Contrast and Error-based Injection Models</b>	<b>57</b>
4.1	MTF-Preserving Injection Methods . . . . .	58
4.2	Details Extraction Methods . . . . .	62
4.3	Experimental Results . . . . .	63
<b>5</b>	<b>Pansharpening Based on Semiblind Deconvolution</b>	<b>67</b>
5.1	Motivations and Contributions . . . . .	68
5.1.1	Notation . . . . .	71
5.2	Semiblind Deconvolution . . . . .	71
5.3	Pansharpening Method Based on Filter Estimation	75
5.4	Experimental Results . . . . .	77
5.4.1	Wald Protocol . . . . .	79
5.4.2	Full Scale Validation . . . . .	89
<b>6</b>	<b>Hyperspectral Pansharpening</b>	<b>95</b>
6.1	Classical Approaches . . . . .	96
6.2	Pansharpening Based on Non-Linear PCA . . . . .	101
6.2.1	Non-Linear PCA . . . . .	101
6.2.2	Proposed Approach . . . . .	105
6.2.3	Experimental Results . . . . .	107
	Pavia Dataset . . . . .	110
	CHRIS+QB Dataset . . . . .	117
<b>7</b>	<b>Conclusions</b>	<b>123</b>
	<b>Bibliography</b>	<b>128</b>

# List of Figures

2.1	Flowchart presenting the steps of a generic pansharpening process based on the CS approach. . . .	15
2.2	Flowchart of a generic pansharpening algorithm belonging to the MRA family. . . . .	23
3.1	China Dataset: (a) Reference Image; (b) PAN; (c) EXP; (d) PCA; (e) IHS; (f) Brovey; (g) BDSD; (h) GS; (i) GSA; (j) PRACS; (k) HPF; (l) SFIM; (m) Indusion; (n) ATWT; (o) AWLP; (p) ATWT M2. .	42
3.2	China Dataset: (a) Reference Image; (b) ATWT M3; (c) GLP; (d) MTF HPM PP; (e) MTF HPM; (f) GLP-CBD. . . . .	43
3.3	Rome Dataset: (a) Reference Image; (b) PAN; (c) EXP; (d) PCA; (e) IHS; (f) Brovey; (g) BDSD; (h) GS; (i) GSA; (j) PRACS; (k) HPF; (l) SFIM; (m) Indusion; (n) ATWT; (o) AWLP; (p) ATWT M2. .	44
3.4	Rome Dataset: (a) Reference Image; (b) ATWT M3; (c) GLP; (d) MTF HPM PP; (e) MTF HPM; (f) GLP-CBD. . . . .	45
3.5	Pléiades dataset: (a) Reference Image; (b) PAN; (c) EXP; (d) PCA; (e) IHS; (f) Brovey; (g) BDSD; (h) GS; (i) GSA; (j) PRACS; (k) HPF; (l) SFIM; (m) Indusion; (n) ATWT; (o) AWLP; (p) ATWT M2. .	46
3.6	Pléiades dataset: (a) Reference Image; (b) ATWT M3; (c) GLP; (d) MTF HPM PP; (e) MTF HPM; (f) GLP-CBD. . . . .	47

3.7	Toulouse Dataset: (a) PAN; (b) EXP; (c) PCA; (d) IHS; (e) Brovey; (f) BDSD; (g) GS; (h) GSA; (i) PRACS; (j) HPF; (k) SFIM; (l) Indusion. . . . .	50
3.8	Toulouse Dataset: (a) PAN; (b) EXP; (c) ATWT; (d) AWLP; (e) ATWT M2; (f) ATWT M3; (g) GLP; (h) MTF HPM PP; (i) MTF HPM; (j) GLP-CBD. . . . .	51
3.9	Rio Dataset: (a) PAN; (b) EXP; (c) PCA; (d) IHS; (e) Brovey; (f) BDSD; (g) GS; (h) GSA; (i) PRACS; (j) HPF; (k) SFIM; (l) Indusion. . . . .	53
3.10	Rio Dataset: (a) PAN; (b) EXP; (c) ATWT; (d) AWLP; (e) ATWT M2; (f) ATWT M3; (g) GLP; (h) MTF HPM PP; (i) MTF HPM; (j) GLP-CBD. . . . .	54
4.1	Example of estimation of the MTF at a given frequency $\nu$ according to Eq. (4.4). (a) Periodic pattern (specimen) at frequency $\nu$ ; (b) resulting image as acquired by the optical system; (c) estimation of the MTF at frequency $\nu$ . . . . .	59
4.2	<i>China dataset</i> false color representation (NIR, Red and Green): (a) Full resolution MS image (reference); (b) PAN image; (c) IHS-ERR ( <i>IHS</i> ); (d) IHS-CON ( <i>Brovey</i> ); (e) MTF-ERR ( <i>GLP</i> ); (f) MTF-CON ( <i>MTF HPM</i> ); (g) ATWT-ERR ( <i>ATWT</i> ); (h) ATWT-CON. . . . .	65
4.3	<i>India dataset</i> false color representation (NIR, Red and Green): (a) Full resolution MS image (reference); (b) PAN image; (c) IHS-ERR ( <i>IHS</i> ); (d) IHS-CON ( <i>Brovey</i> ); (e) MTF-ERR ( <i>GLP</i> ); (f) MTF-CON ( <i>MTF HPM</i> ); (g) ATWT-ERR ( <i>ATWT</i> ); (h) ATWT-CON. . . . .	66
5.1	Error in angle over (a) $\lambda$ and (b) $\mu$ variations for the proposed approach with support size equal to 25, using the model in Eq.(5.1) without additional noise. . . . .	81

5.2	Error in angle Vs. the support size using the model in Eq.(5.1) without additional noise. . . . .	81
5.3	Error in angle over (a) $\lambda$ and (b) $\mu$ variations for the proposed approach with support size equal to 25, using the model in Eq.(5.1) with Gaussian noise. The SNR is equal to 45 db. . . . .	82
5.4	Error in angle Vs. the support size using the model in Eq.(5.1) with Gaussian noise. The SNR is equal to 45 db. . . . .	83
5.5	Proposed method's estimated blur function, when $\mathbf{\hat{M}}$ is obtained with (a) the bicubic interpolator or (b) the half band polynomial with 23 coefficient interpolator. . . . .	84
5.6	From the top to the bottom: (a)-(d) Expected blurs; (e)-(h) Estimated blurs by the <i>FE MS</i> method; (i)-(l) Estimated blurs by the <i>FE</i> approach. The columns are ordered by wavelengths ( <i>i.e.</i> , Blue, Green, Red and NIR). . . . .	85
5.7	IKONOS sensor: Relative spectral responses. . . . .	85
5.8	(a) MS full resolution (4 m) image (used as reference image); (b) <i>EXP</i> , (c) <i>MTF</i> , (d) <i>FE</i> methods. . . . .	88
5.9	The results in terms of <i>SAM</i> and <i>SCC</i> indexes of the proposed method obtained by varying the support sizes. . . . .	89
5.10	From the top to the bottom: (a)-(d) Expected blurs; (e)-(h) Estimated blurs by the <i>FE</i> approach. The columns are ordered by wavelengths ( <i>i.e.</i> , Blue, Green, Red and NIR). . . . .	91
5.11	Full scale fused results: (a) <i>EXP</i> , (b) <i>PAN</i> , (c) <i>MTF</i> and (d) <i>FE</i> images. . . . .	92
5.12	Differences in absolute value between the details achieved by the <i>FE</i> and <i>MTF</i> methods over the bands, ordered by wavelengths ( <i>i.e.</i> Blue, Green, Red, NIR). . . . .	93

5.13	A small area in the <i>Indianapolis dataset</i> : (a) MS image; (b) PAN image; (c) RGB fusion result using the <i>FE</i> approach; Details extracted as $\frac{\mathbf{p}}{\mathbf{p}_{LP}}$ using (d) <i>FE</i> and (e)-(h) <i>MTF</i> (different details for each spectral band) approaches. . . . .	94
6.1	<i>Hyp+ALI</i> dataset: (a) MS; (b) PAN; (c) PCA; (d) GS; (e) GSA; (f) SFIM; (g) GLP; (h) MTF HPM. .	98
6.2	<i>CHRIS+QB</i> dataset: (a) MS; (b) PAN; (c) PCA; (d) GS; (e) GSA; (f) SFIM; (g) GLP; (h) MTF HPM.	98
6.3	Networks implementing (a) mapping $\mathbf{z}$ and (b) demapping $\phi$ functions. . . . .	102
6.4	Complete scheme of the proposed hybrid approach.	107
6.5	The 4 NLPCs and the PAN image for the <i>Pavia dataset</i> . . . . .	110
6.6	RGB images obtained by combining bands 70, 50 and 20 for the <i>Pavia dataset</i> : (a) Reference image; (b) $\widetilde{\mathbf{HS}}$ ; Fused products achieved by means of (c) NLPCA; (d) PCA; Hybrid-NLPCA: (e) HPF; (f) SFIM; (g) GLP; (h) MTF HPM; (i) ATWT; (j) AWLP; Directly (all bands) applied to HS and PAN images: (k) HPF; (l) SFIM. The fused images are obtained using the NLPCs 2 and 3. . . . .	114
6.7	RGB images obtained by combining bands 70, 50 and 20 for the <i>Pavia dataset</i> : (a) Reference image; Fused products obtained by directly applying to HS and PAN images the following methods: (b) GLP; (c) MTF HPM; (d) ATWT; (e) AWLP. . . . .	115
6.8	The 3 NLPCs and the PAN image for the <i>CHRIS+QB dataset</i> . . . . .	117

6.9	RGB images obtained by combining bands 11, 9 and 7 for the <i>CHRIS+QB dataset</i> : (a) PAN; (b) $\widetilde{\mathbf{HS}}$ ; Fused products by the means of (c) NLPCA; (d) PCA; Hybrid-NLPCA: (e) HPF; (f) SFIM; (g) GLP; (h) MTF HPM; (i) ATWT; (j) AWLP; Directly (all bands) applied to HS and PAN images: (k) HPF; (l) SFIM; (m) GLP; (n) MTF HPM; (o) ATWT; (p) AWLP. The NLPCA-based fused images are obtained using the third NLPC. . . . .	119
-----	---	-----





# List of Tables

2.1	List of the main acronyms. . . . .	12
2.2	List of the main symbols. . . . .	12
2.3	Values of the weights in Eqs. (2.2) and (2.1) for different CS-based pansharpening techniques. . . .	16
2.4	Employed MRA scheme and weights in Eq. (2.14) for different MRA-based pansharpening techniques.	24
3.1	China dataset: Quantitative results. . . . .	41
3.2	Rome dataset: Quantitative results. . . . .	43
3.3	Pléiades dataset: Quantitative results. . . . .	45
3.4	Toulouse dataset: Quantitative results. . . . .	49
3.5	Rio dataset: Quantitative results. . . . .	52
4.1	<i>China dataset</i> : Comparison between error-based (ERR) and contrast-based (CON) injection models.	64
4.2	<i>India dataset</i> : Comparison between error-based (ERR) and contrast-based (CON) injection models.	64
5.1	MTF gains at Nyquist cutoff frequency. . . . .	80
5.2	Performance evaluation for the <i>MTF</i> and <i>FE</i> meth- ods, when $\widetilde{\mathbf{M}}$ is obtained with the bicubic interpo- lator. . . . .	86
5.3	Performance evaluation for the <i>MTF</i> , <i>FE MS</i> and <i>FE</i> methods, when $\widetilde{\mathbf{M}}$ is obtained with the 23 co- efficient interpolator. . . . .	86
5.4	Performance evaluation for the <i>EXP</i> , <i>SFIM</i> , <i>ATWT</i> and <i>MTF</i> methods, when $\widetilde{\mathbf{M}}$ is obtained with the 23 coefficient interpolator. . . . .	87

5.5	Execution times for the compared methods on the <i>China dataset</i> . . . . .	87
5.6	Full scale performance evaluation using SAM and SCC indexes for <i>EXP</i> , <i>SFIM</i> , <i>MTF</i> and <i>FE</i> methods.	90
6.1	Full scale analysis: <i>Hyp+ALI</i> dataset with perfectly coregistrated images. . . . .	99
6.2	Full scale analysis: <i>Hyp+ALI</i> dataset with 60 meters misalignment (2 HS pixels). . . . .	100
6.3	Full scale analysis: <i>CHRIS+QB</i> dataset with perfectly coregistrated images. . . . .	101
6.4	Correlation coefficients between PAN image and each NLPC on the <i>Pavia dataset</i> . . . . .	110
6.5	SAM quality indexes (measured in degrees) obtained with different fusion approaches applied to the <i>Pavia dataset</i> . Pure MRA are indicated by “All bands”. CS approaches are labeled as NLPCA and PCA, while the others are hybrid approaches based on NLPCA projection and the indicated MRA methodology. The symbol “/” indicates an unavailable value. . . . .	112
6.6	ERGAS quality indexes obtained with the different fusion approaches applied to the <i>Pavia dataset</i> . Pure MRA are indicated by “All bands”. CS approaches are labeled as NLPCA and PCA, while the others are hybrid approaches based on NLPCA projection and the indicated MRA methodology. The symbol “/” indicates an unavailable value. . . . .	113
6.7	SCC quality indexes obtained with the different fusion approaches applied to the <i>Pavia dataset</i> . Pure MRA are indicated by “All bands”. CS approaches are labeled as NLPCA and PCA, while the others are hybrid approaches based on NLPCA projection and the indicated MRA methodology. The symbol “/” indicates an unavailable value. . . . .	113

6.8	Computational time (expressed in seconds) evaluated for the different methods applied to the <i>Pavia dataset</i> . Pure MRA are indicated by “All bands”. CS approaches are labeled as NLPCA and PCA, while, the others are hybrid approaches based on NLPCA projection and the indicated MRA methodology. The symbol “/” indicates an unavailable value. . . . .	116
6.9	Correlation coefficients between PAN image and each NLPC obtained by the <i>CHRIS+QB dataset</i> . .	117
6.10	SAM (expressed in degrees) and SCC quality indexes over the different fusion approaches applied to the <i>CHRIS+QB dataset</i> . The symbol “/” indicates an unavailable value. . . . .	118
6.11	Computational time (expressed in seconds) evaluated for the different methods applied to the <i>CHRIS+QB dataset</i> . The symbol “/” indicates an unavailable value. . . . .	120



# Acknowledgments

This research activity would not have been possible without the support of the University of Salerno and the Grenoble Institute of Technology. I am heartily thankful to my Advisor, Rocco Restaino, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. I would like to express my thankfulness to the Prof. Maurizio Longo for his guide and availability. Furthermore, special thanks go to the Prof. Jocelyn Chanussot, who hosted me at the GIPSA-Lab at the Grenoble Institute of Technology, for his guide, support and devotion to research. Few words are not surely enough to demonstrate the gratitude to Mauro Dalla Mura for his support and helpful discussions. Deepest gratitude is also due to other people who with their knowledge and skills helped the success of my research activities. I would like to mention them here in a pure alphabetically order: Paolo Addesso, Prof. Luciano Alparone, Prof. José Bioucas-Dias, Prof. Andrea Garzelli, Roberto Conte, Giorgio Licciardi, Antonino Maltese, Miguel Simões and Prof. Lucien Wald. Without each of you this thesis would never find its way.

I would also like to thank the Prof. José Bioucas-Dias and the Prof. Antonio Plaza for their availability to provide me a review of my dissertation in order to fulfill to the European Label's requirements.

The new experiences allow the growing of each one. For this reason, I would like to thank all the guys at the GIPSA-Lab and the NATO Undersea Research Centre (NURC), where I spent about one year of my Ph.D. period. Special thanks go to Paolo

Braca for the supervision and assistance during the three months at NURC and my lab mates Alessandro Chiancone at the GIPSA-Lab to share his spare time with me in France.

I am glad to have met Mario Di Mauro, Prof. Maurizio Guida, Prof. Stefano Marano, Vincenzo Matta, Rita Montone, Prof. Vincenzo Pierro and Fabio Postiglione at the University of Salerno, where I was hosted in the last three years.

Thanks to my girlfriend Ilaria for her understanding and endless love, my father Giuseppe, my mother Luciana and my sister Antonia for their constant love and support through the duration of my studies.

Lastly, I would like to thank all those who have been with me over the last years and those who are no longer with us.

# Chapter 1

## Introduction

Remote sensing consists in measuring some characteristics of an object from a distance. Many devices have been proposed in the scientific literature to carry out this task, in the attempt of reproducing the human ability in collecting information through the remote sense organs (hearing, smell and eyes). A key example of remote sensing is the Earth Observation from sensors mounted on satellites that is a crucial aspect of space programs.

The first satellite used for Earth observation was Explorer VII, which was launched in October 1959 with the purpose of measuring the “Earth’s radiation budget”, namely the amount of heat reflected by the Earth. It has been followed by thousands of satellites, many of which are still working. They collect information in a wide range of frequency, which is widely exploited in a large number of military and civil applications.

Due to the availability of a large number of different sensors and the subsequent huge amount data collected, the idea of obtaining improved products by means of fusion algorithms is becoming more intriguing. *Data fusion* is often exploited for indicating the process of integrating multiple data and knowledge related to the same real-world scene into a consistent, accurate, and useful representation. This term is very generic and it includes different levels of fusion. In this dissertation, we will focus on the *low level data fusion*, which consists in combining several sources of raw data. In

this field, one of the most relevant scientific application, which is demonstrated by both the organization of a contest by the IEEE Data Fusion Committee in 2006 and the huge number of papers published in the recent years in this research area, is surely the *Pansharpening*.

*Pansharpening* refers to the fusion of a panchromatic image (a single band that covers the visible and near infrared spectrum) and a multispectral/hyperspectral image (tens/hundreds bands) acquired on the same area. This can be seen as a particular problem of data fusion since one would aim at combining in a single product the spatial details resolved by the panchromatic image, but not present in the multispectral/hyperspectral image, and the multiple spectral bands of the other image characterized by a greater spectral content. The relevance of these techniques is mainly due to the necessity of obtaining a data with high spatial and spectral resolutions for visual inspections. Very popular software, such as Google Earth, exploit these algorithms to provide enhanced products to final users. In the latest years, the possibility of applying them as preprocessing for algorithms of object and change detection, linear unmixing and classification has been discussed and tested. The fusion of these data compensates the unavailability of images with high resolutions in both domains, which are precluded by physical limits of remote sensor realizations and by the necessity of alleviating problems, such as, the data storage on-board and the transmission to ground stations.

In this dissertation, we will report the details of the relevant aspects related to this problem both from a methodological and application oriented point of view. Firstly, an overview of the wide existing literature will be presented, together with the classification of the approaches into two main families. More in detail, the pansharpening algorithms are usually divided into two categories: Component substitution and multiresolution analysis. The former is based on the projection into a new space, the substitution of the most similar component with the panchromatic image and the projection back into the original space. The main idea in the second approach is the extraction of the spatial details of



the panchromatic image and their injection into the multispectral/hyperspectral data exploiting a proper injection model. Under the hypothesis of linear transformation and the substitution of only one component with the panchromatic image, even the component substitution family can be represented as an extraction detail phase followed by a proper injection rule (see Chapter 2 for further details). Therefore, a generic classical pansharpening algorithm can be divided into two phases: Detail extraction and detail injection.

By focusing on the former, many powerful pansharpening approaches exploit the functional relation between the panchromatic and multispectral images. To this purpose, the modulation transfer function of the higher spectral resolution sensor is typically used, being easily approximated as a Gaussian filter whose analytic expression is fully specified by the sensor gain at the Nyquist frequency. However this characterization is often inadequate in practice, or, in some cases, the gains at the Nyquist frequency could be unavailable. In Chapter 5, we will present an algorithm for estimating the relation between panchromatic and multispectral images directly from the available data, through the solution of an optimization problem. Moreover, a critical analysis and comparison among the state-of-art extraction detail approaches will be provided to justify the advantages.

With reference to the other main step of a pansharpening algorithm, i.e., the injection of the previously extracted details, we will analyze two of the most popular models. The main contributions are: 1) The analysis of the injection methods for pansharpening in the light of physical considerations regarding the point spread function and 2) the derivation of the constraint on the modulation transfer functions in terms of image local contrast. The findings presented in Chapter 4 can be related to rather more general image fusion contributions, which already highlighted the superiority of methods based on contrast pyramids. Focusing on the pansharpening applications, the same considerations justify the very appealing visual features of the high pass modulation method.

The pansharpening problem is usually related to the fusion of

multispectral and panchromatic images. Few attempts for fusing hyperspectral images are present in the literature. In this dissertation, we firstly investigate the use of classical pansharpening approaches in the hyperspectral case. We will also analyze the specific features of hyperspectral data fusion with respect to the multispectral image case. It is worth to note that the satellite platforms, which are able to acquire both panchromatic and hyperspectral data, are rare. Therefore, procedures as co-registration and intercalibration of the acquired data have to be performed as preliminary step. The hyperspectral sharpening differs from the classical multispectral one mainly because of the large number of bands, implying a significantly increase of the computational burden and their narrowness, which reduces the similarities among the acquired data and results in the raise of the spectral distortions of the final products. To overcome the computational issue, the application of techniques, which are able to reduce the dimensionality of the hyperspectral data, are surely advisable. In the literature, one of the most used methodology to reach this goal is the Principal Component Analysis (PCA). In this thesis, we propose to use the non-linear PCA, which exploits auto-associative neural networks, and compare its performances with the classical PCA. Moreover, in order to reduce the spectral distortions, hybrid approaches are suitable. In fact, after the reduction step, we only inject the spatial details of the panchromatic image into one or more selected components instead of substituting it.

Finally, an experimental activity requires a proper implementation of the validation procedures, quality indexes, algorithms for the comparison and preprocessing procedures. For these reasons, after a collaboration among the GIPSA-Lab (Grenoble Institute of Technology), the universities of Florence, Siena, Salerno and the MINES ParisTech, a MATLAB Toolbox has been developed, which allows, starting from multispectral and panchromatic images, the comparison of several algorithms belonging to the state-of-art; it exploits the two main validation procedures (i.e., at reduced and full scale) and the commonly used quality indexes (see Chapter 2 for details). This work represents a step towards the

standardization of some processes and, after its distribution on the net, it will help all the researches who want to propose a new pansharpening algorithm, making easier the improvements in this scientific area. Another goal of this work is the critical comparison of the existing pansharpening methods, which has not yet been sufficiently investigated. The contest launched by the IEEE Data Fusion Committee in 2006 has made a first step tackling this issue, since it performed an explicit comparison among the principal existing methods assessed on the same datasets, with the same metrics and in the same working conditions. The current work moves in the same direction, allowing for a wide critical analysis of several state-of-art pansharpening approaches and to the comparison of the main validation protocols for the performance assessment.

This dissertation is organized as follows. In Chapter 2, the state-of-art approaches and their classification will be discussed. In Chapter 3, the critical comparison of many pansharpening techniques, exploiting the two principal validation procedures will be shown. In Chapter 4, the two main injection models will be presented and the advantages of the high pass modulation injection model will be pointed out both physically and numerically. In Chapter 5, a technique to properly estimate the filter to spatially degrade the panchromatic image to the same resolution of the multispectral data will be described and analyzed. In Chapter 6, the hyperspectral pansharpening problem will be discussed and analyzed by means of classical approaches, and, moreover, a novel hybrid approach based on non-linear PCA and multiresolution methodologies will be exploited to overcome the issues related to the application of pansharpening to hyperspectral data. Finally, in Chapter 7, some final remarks and indications for further developments of this work will be presented.



## Chapter 2

# An Approach to Data Fusion: Pansharpening

In this chapter, the fusion between a high spatial resolution and reduced spectral content image with another one characterized by complementary features is analyzed. This raw data fusion approach is usually called *Pansharpening*. An introduction to the methodologies proposed in the literature is provided to the reader and a more detailed overview of the main techniques exploited in the recent years to face this problem is also carried out. More in detail, many pansharpening techniques belonging to the Component Substitution (CS) and to the MultiResolution Analysis (MRA) families are described. Furthermore, the two main assessment procedures (i.e., at full and reduced scale) are shown and commented.

The remainder of the chapter is as follows. Sec. 2.1 shows an introduction to the problem and a review of the methodologies proposed in the literature. While, Sec. 2.2 presents a review of the CS- and MRA-based approaches, providing a detailed description of the main algorithms belonging to these families. Finally, Sec. 2.3 is devoted to the description of the two main assessment procedures operating at reduced and full scale by introducing the corresponding quality indexes.

## 2.1 Introduction to Pansharpening

Pansharpening refers to the fusion of a *PANchromatic* (PAN) and a *MultiSpectral* (MS) image acquired simultaneously on the same area. This can be seen as a particular problem of data fusion since one would aim at combining in a single product the spatial details resolved by the PAN (but not present in the MS) and the spectral bands of the MS image. With respect to the general problem of multisensor fusion, the pansharpening could not require the challenging phase of spatial coregistration since the images are typically acquired simultaneously, being the PAN and MS sensors both mounted on the same platform [1]. Nowadays, PAN and MS images can be obtained in bundle by several commercial optical satellites such as IKONOS, OrbView, Landsat, Spot, Quickbird and WorldView. The spatial resolution is even below half meter for the PAN (thus representing the commercial product with highest spatial resolution) and the spectral resolution can be up to eight bands, captured in the visible and near infrared domain, for the MS. The fusion of the PAN and MS images constitutes the sole possibility for having in the same product the highest resolution in both the spatial and spectral domains, since physical constraints preclude to achieve this result by using a single device. The demand for pansharpened data is continuously growing due to the increasing diffusion of commercial products based on spatially enhanced images, as for example Google Earth and Microsoft Virtual Earth. Furthermore, pansharpening represents an important preliminary step for many remote sensing applications, such as change detection [2], object recognition [3] and visual image interpretation [4].

The interest of the image processing community in pansharpening is evident by reviewing the recent technical literature; detailed surveys of pansharpening algorithms can be found in [1,5,6]. We have chosen to focus on some of the most popular algorithms belonging to the two main classical approaches: The *Component Substitution* and the *MultiResolution Analysis*. The former approach relies on the substitution of a component (e.g., obtained

with a transformation of the data) of the MS image representation with the PAN image. This family comprises algorithms such as *Intensity-Hue-Saturation* (IHS) [7, 8], *Principal Component Analysis* (PCA) [9, 10] and *Gram-Schmidt* (GS) [11]. The MRA approach is based on the injection of spatial details, which are typically obtained by the details through a multiscale decomposition of the PAN image, into the MS data. The spatial details can be extracted by several approaches, ranging from very simple methods based on a single low pass filter [8, 12] to more complex decompositions based on *Laplacian* [13] pyramids, or *wavelet/contourlet* operators [10, 14].

These two classes of methods will be widely discussed in Sec. 2.2. Some other methods have been proposed in the literature that cannot be classified within the two above-mentioned families. Some algorithms use the Bayesian paradigm for carrying out the data fusion task [15]. The difficulty in finding a suitable statistical model to jointly characterize the available MS and PAN images [16] has strongly limited its use for pansharpening. However many contributions based on Bayesian estimation theory have been presented in the recent literature. They are based on regularized solutions of the ill-posed problem consisting in the high resolution image reconstruction from coarse measurements. Some proposals are based on *Total Variation* penalization terms [17] and many others rely upon recent developments in sparse signal representation, or *Compressive Sensing* theory, which was proposed in [18, 19]. Among the latter, it is possible to refer to the seminal works [20, 21] that introduced this approach, whereas more recent improvements have been achieved through the application of super-resolution techniques, which are already largely diffused in many image processing and computer vision applications [22]. In particular, methods belonging to this family are [23–25].

Pansharpening have been also proposed for fusing panchromatic and HyperSpectral (HS) data [26, 27]. Clearly, this task cannot be addressed by employing conventional methods due to the particular issues that have to be faced (e.g., non simultaneous acquisition, co-registration of the data and different spatial cover-

age and resolution [1, 28]). Furthermore, it should pay attention to the computational problem when fusing the images. The use of very simple fusion techniques could be advisable in order to alleviate it. Accordingly, component substitution approaches are surely preferable, but, they can generate spectral distortions, which are, in general, more present in the hyperspectral case. Therefore, the use of hybrid approaches (i.e., the application of MRA methodologies on some components obtained by component substitution approaches) could lead to advantages in both directions. In [27], the use of non linear Principal Component Analysis [29] and Indusion [30] is proposed to solve the above-mentioned problems. In [26], state-of-art pansharpening approaches are tested on a scenario acquired by the Hyperion and ALI sensors and the necessity of very simple but suitable injection models to reduce both the computational burden and distortions is pointed out. Moreover, in [31] the use of a model that properly minimizes the spatial distortion index of the Quality with No Reference (QNR) is exploited for hyperspectral pansharpening. Other methodologies in this literature are based on variational approaches to enforce spectral coherence in the fusion process [32], unmixing methods combined with pansharpening [33] and Bayesian approaches based on [34], whose effectiveness in this field has been already demonstrated [35].

Due to the lack of objective reference images, universal measures for quantifying the enrichment introduced by pansharpening cannot be explicitly formulated, being strongly dependent on the intended application. Accordingly, a common practice is the reference to ideal criteria, among which the most credited dates back to Wald et al. [36]. The latter formalizes the required features of the fused product by pointing out two specific requirements, which are also known as the *consistency* and the *synthesis* properties. The first, which is more affordable in the practice, involves the reversibility of the pansharpening process; it states that the original MS image should be obtainable by simply degrading the pansharpened image. On the other hand, the synthesis property points out the effective purpose of the pansharpening procedure.



It states that the final product has to reproduce the characteristics of the original MS image at a higher resolution. Since the tackled process deals with MS data, this condition entails that both the features of each single channel and the mutual relationships among bands have to be preserved, justifying the original articulation in two distinguished statements.

The definition of a technique that fulfills the constraints defined in the protocol is still an open problem [5, 37] and relates closely to the general discussion regarding image quality assessment [38] and image fusion [39, 40]. Moreover, there are additional issues linked to the lack of an universally accepted index matching the human capability in evaluating the difference of two images (for example the Mean Square Error (MSE) has been proven to be inadequate for this task [41], which has given rise to the definition of many other indexes for assessing the image quality (some are presented in Sect. 2.3.1)). In addition, the unavailability of a reference high resolution MS image precludes the evaluation of the results regardless the chosen quality index.

In order to face these aspects and perform a quantitative evaluation of the results, two main solutions have been proposed. The first relies on the reduction of the spatial resolution of both the original MS and PAN images and then the original MS image is used as reference for the evaluation of the results [42]. In this strategy the invariance among scales of the fusion procedures is assumed [36]. However, this hypothesis is not always verified in the practice, especially for very high resolution urban area data [39]. The second employs indexes that do not require the availability of the reference image [39, 40]. Clearly in this case, the results are heavily dependent on the definition of such indexes.

## 2.2 A Discussion of Assessed Pansharpening Methods

The goal of this section is the presentation and analysis of some widely used methods proposed in the technical literature, which

can be considered as state-of-art algorithms for pansharpening. They can be grouped into two main classes: Component Substitution methods and techniques based on MultiResolution Analysis. Before describing these two approaches, the notation and conventions used in this thesis are introduced.

### 2.2.1 Notation

Acronym	Description
MS	MultiSpectral image
PAN	PANchromatic image
HS	HyperSpectral image
HRMS	High Resolution MS image
LRMS	Low Resolution MS image
CS	Component Substitution
MRA	MultiResolution Analysis

**Table 2.1** List of the main acronyms.

Symbol	Description
HRMS	Reference image
MS	MS image
$\widetilde{\text{MS}}$	Upsampled MS image
P	PAN image
$\widehat{\text{MS}}$	Pansharpened image
$R$	Spatial resolution ratio between MS and PAN
$N$	Number of MS bands

**Table 2.2** List of the main symbols.

In Tabs. 2.1 and 2.2, the lists of the main acronyms and symbols are shown, together with their brief description. Other acronyms and symbols will be defined in this thesis, as necessary.

The notation followed in this thesis (except for some cases where it will be properly reported) will be as follows. We indi-

cate the vectors with the bold lowercase (e.g.,  $\mathbf{x}$ ) and the bidimensional and three dimensional arrays with the bold uppercase (e.g.,  $\mathbf{X}$ ). Moreover, a generic MS image  $\mathbf{X} = \{\mathbf{X}_k\}_{k=1,\dots,N}$  is a three dimensional array, composed by  $N$  bands indexed by the subscript  $k = 1, \dots, N$ ; accordingly,  $\mathbf{X}_k$  indicates the  $k$ -th band of  $\mathbf{X}$ . A generic PAN image is a bidimensional matrix indicated as  $\mathbf{Y}$ . Furthermore, the  $i$ -th element of a generic vector  $\mathbf{x}$  will be indicated as  $x_i$ .

### 2.2.2 Component Substitution

This family is based on the projection of the MS image into another space, spanned by decorrelated components, which permits to handle the redundancy of the data. The main assumption behind this approach is that the transformation separates the spatial structure from the spectral information in different components [1]. Subsequently the transformed MS image can be enhanced by substituting part or the entire panchromatic image, to one or more components representing the spatial details. Accordingly, the greater the correlation between the PAN image and the replaced component(s), the less distortion will be introduced by the approach [1]. To this purpose, histogram matching of the selected component(s) and the panchromatic image is performed before the substitution. Finally, the pansharpening process is completed through the inverse transformation that brings the data back to the original space.

This approach is global (i.e., it operates in the same way on the whole image) leading to advantages and drawbacks. In greater details, techniques belonging to this family usually yield in a high fidelity in rendering the spatial details in the final product [43] and they are in general fast and easy to implement. On the other side, they are not able to account for local dissimilarities between panchromatic and multispectral image which can produce significant spectral distortions [1, 5].

In some works, the denomination CS is often substituted by *Projection - Substitution* in order to underline the main steps of

the algorithms into this family. A new look at these methodologies was proposed by Tu et al. [44] and then analyzed in other subsequent works [45]. In fact, under the hypothesis of a linear transformation applied to the MS data and if the substitution involves only a single component in the new transformed domain, it is simple to see that the fusion process can be obtained without any forward and backward transformation of the data but only by applying a proper injection scheme [45]. This observation allows a faster implementation of these methods. Starting from this consideration, a general formulation of the CS fusion scheme is given by

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + g_k (\mathbf{P} - \mathbf{I}_L), \quad (2.1)$$

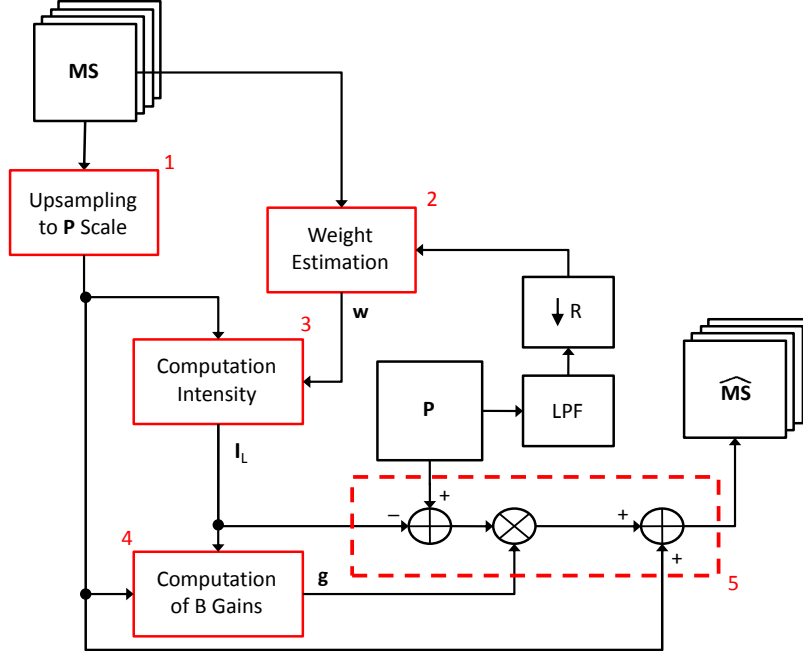
in which the subscript  $k$  indicates the  $k$ -th band,  $\mathbf{g} = [g_1, \dots, g_k, \dots, g_N]$  is the vector of the *injection gains*, while,  $\mathbf{I}_L$  is defined as follows:

$$\mathbf{I}_L = \sum_{i=1}^N w_i \widetilde{\mathbf{MS}}_i, \quad (2.2)$$

where the weights  $\mathbf{w} = [w_1, \dots, w_i, \dots, w_N]$  measure the spectral overlap among the spectral bands and the panchromatic image  $\mathbf{P}$  [1].

Eq. (2.1) testifies that the fusion process can be carried out without the explicit computation of the spectral transformation [43, 46]. It is worth to note that another family could be found in the literature [1, 5]. In fact the name of *Relative Spectral Contribution (RSC)* approaches is used for methods in which the substituted component  $\mathbf{I}_L$  is defined as in Eq. (2.2). However, under the two previously stated hypotheses, this family collapses into the CS one, as defined by Eq. (2.1). For this reason, in this thesis, we use a classification of pansharpening methods based on two main families, as already done in [6].

In Fig. 2.1 a flowchart which describes a generic fusion process into the CS family is shown. Specifically, it is possible to notice the presence of blocks aimed to: 1) Upsample the MS image to reach the panchromatic resolution; 2) Estimate the weight vector  $\mathbf{w}$  from



**Figure 2.1** Flowchart presenting the steps of a generic pansharpening process based on the CS approach.

the original MS image and a low pass version of the PAN image; the latter is obtained through a proper Low Pass Filter (LPF) and downsampled with a factor  $R$ ; 3) Calculate the intensity component by Eq. (2.2); 4) Compute the band dependent injection gains  $\{g_k\}_{k=1,\dots,N}$ ; 5) Inject the extracted details according to Eq. (2.1).

The CS family includes many popular pansharpening approaches, such as the *Intensity-Hue-Saturation* (IHS) [46], *Principal Component Analysis* (PCA) [8, 46] and *Gram-Schmidt* (GS) [43] methods, each based on different projections of the MS image. Since there is not an unique transform for extracting the component most suited for substitution, and the choice of the appropriate spectral transformation for each dataset ends up in a challenging task, other methods are based on the estimation of the optimal component and are known as *Adaptive CS* [43, 47].

In the following, a more detailed description of the main CS methods is presented and summarized in Tab. 2.3, in which the

Method	$w_{i,k}$	$g_k$
<b>IHS</b>	$1/N$	1
<b>Brovey</b>	$1/N$	$\frac{\widetilde{\mathbf{MS}}_k}{\mathbf{I}_L}$
<b>PCA</b>	$x$	$y$
<b>GS</b>	$1/N$	$\frac{\text{cov}(\mathbf{I}_L, \widetilde{\mathbf{MS}}_k)}{\text{var}(\mathbf{I}_L)}$
<b>GSA</b>	$\widehat{w}_i$	$\frac{\text{cov}(\mathbf{I}_L, \widetilde{\mathbf{MS}}_k)}{\text{var}(\mathbf{I}_L)}$
<b>BDS</b>	$\widehat{w}_{i,k}$	$\widehat{g}_k$
<b>PRACS</b>	$\widehat{w}_{i,k}$	Eqs. (2.12)-(2.13)

**Table 2.3** Values of the weights in Eqs. (2.2) and (2.1) for different CS-based pansharpening techniques.

specific coefficients needed for the implementation based on Eqs. (2.2) and (2.1) are reported.

### Intensity-Hue-Saturation

The *Intensity-Hue-Saturation* (IHS) pansharpening method [7, 8] exploits the transformation into the IHS color space that mimics the human visual system in processing the intensity (I), hue (H) and saturation (S) information. The IHS transform can be only applied to RGB true color images, leading to a major limitation for processing MS images. In [46] the authors have generalized the concept of IHS to images with more than three bands. Furthermore, they have shown that the IHS fusion approach can be reformulated, for  $k = 1, \dots, N$ , as follows:

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \mathbf{P} - \mathbf{I}_L, \quad (2.3)$$

where  $\mathbf{I}_L$  follows from Eq. (2.2) with the coefficients  $\{w_i\}_{k=1,\dots,N}$  all equal to  $1/N$ . Eq. (2.3) fits the general paradigm described by Eq. (2.1), with all the gains  $\{g_k\}_{k=1,\dots,N}$  equal to one (see Tab. 2.3). This approach is also called *Fast IHS* because it avoids the sequential computation of the direct transformation, substitution and the final backward step.

Moreover, by replacing the detail injection rule in Eq. (2.3) with a multiplicative scheme, the *Brovey* pansharpening approach

[48] is obtained. Indeed in this technique the fused image is defined, for  $k = 1, \dots, N$ , as:

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k \cdot \frac{\mathbf{P}}{\mathbf{I}_L}. \quad (2.4)$$

Since Eq. (2.4) can be recast as

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{I}_L} \cdot (\mathbf{P} - \mathbf{I}_L), \quad (2.5)$$

it is easy to identify the model (2.1), achievable by setting, for  $k = 1, \dots, N$ ,

$$g_k = \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{I}_L}. \quad (2.6)$$

This specific parameter choice, together with the values of the weight vector  $\mathbf{w}$ , is reported in Tab. 2.3.

### Principal Component Analysis

*Principal Component Analysis* (PCA) (also known as *Karhunen-Loève transform* or *Hotelling transform*) is another spectral transformation widely employed for pansharpening applications [8, 49]. It is achieved through a roto-translation of the original space (i.e., linear transformation) that yields decorrelated axes called Principal Components (PCs). PCs are in general ordered according to the variance of the corresponding data projectons; the variance can be thus used to quantify the informative content of each PC.

More specifically, the hypothesis underlying its application to pansharpening is that the spatial information, that is shared by all the channels, is concentrated in the first component, while the spectral information, that is a peculiar characteristic of each single band, is accounted for by the other components. Actually, the spatial information is mapped to the first component, to an extent proportional to the correlation among the MS channels [5].

Again, the whole fused process can be described by the general formulation stated by Eq. (2.1), where the  $\mathbf{w}$  and  $\mathbf{g}$  coefficient

vectors are derived by the PCA procedure on the MS image; consequently, no specific expression is provided in Tab. 2.3 for  $\mathbf{w}$  and  $\mathbf{g}$ , since they are dependent on the particular processed dataset.

### Gram-Schmidt Orthogonalization

The *Gram-Schmidt* (GS) orthogonalization transformation is used for a powerful pansharpening approach, which was firstly implemented in a patent by Kodak [11]. The GS transformation is a common technique used in linear algebra and multivariate statistics to orthogonalize the spectral content of data.

In the fusion process, the mean of each band is subtracted from each pixel, before the orthogonalization procedure. The latter starts by using, as the first basis vector, a synthetic low resolution panchromatic image  $\mathbf{I}_L$  at the same scale of the MS image, which is successively replaced, in the fusion step, by the full resolution PAN. Therefore, GS constitutes a more general method than PCA, which can be obtained by using the first principal component as the low resolution panchromatic image [50]. The representation of the MS image is then carried out by constructing a complete orthogonal basis. The pansharpening procedure is thus completed by substituting the first component with the histogram-matched PAN image and inverting the transformation.

Again, the latter process is achieved according to Eq.(2.1), by using, for  $k = 1, \dots, N$ , the gains [43]:

$$g_k = \frac{\text{cov}(\widetilde{\mathbf{MS}}_k, \mathbf{I}_L)}{\text{var}(\mathbf{I}_L)}, \quad (2.7)$$

where  $\text{cov}(\mathbf{X}, \mathbf{Y})$  indicates the covariance between two images  $\mathbf{X}$  and  $\mathbf{Y}$ , and  $\text{var}(\mathbf{X})$  is the variance of  $\mathbf{X}$  (see also Tab. 2.3).

Diverse algorithms, characterized by different definition of the weights in Eq.(2.2), are achieved by changing the generation method of  $\mathbf{I}_L$ . The simplest way to obtain this low-resolution panchromatic image consists in simply averaging the MS bands (i.e., in setting  $w_i = 1/N$ , for all  $i = 1, \dots, N$ ); this modality is simply called *GS* [11]. In [43] the authors proposed an enhanced



version, called *GS Adaptive* (GSA), in which  $\mathbf{I}_L$  is generated by a weighted average of the MS bands. The required weights are estimated by minimizing the Mean Square Error (MSE), with respect to a filtered and downsampled version of the panchromatic image. This implementation of the GS algorithm could be also classified in the Adaptive CS category discussed in the next section. A last widespread method for generating the low resolution panchromatic image  $\mathbf{I}_L$  entails the application of a low pass filter to the original PAN image and is named *GS mode 2* [11], leading to a hybrid approach which is sometimes included in the MRA family.

### Adaptive CS

This category incorporates data-dependent self-adaptive approaches that are briefly described in the following and whose parameters are reported in Tab. 2.3.

In this category, we can find the *Band-Dependent Spatial-Detail* (BDSD) algorithm [51], which uses the generic formulation reported in Eq. (2.1) and points at estimating separately for each band both  $\mathbf{w}$  and  $\mathbf{g}$  according to the Minimum MSE (MMSE) criterion. By defining the coefficients

$$\gamma_{k,i} = \begin{cases} g_k, & \text{if } i = N + 1, \\ -g_k w_i, & \text{otherwise,} \end{cases} \quad (2.8)$$

Eq. (2.1) can be rewritten in the compact matricial form

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \mathbf{H}\boldsymbol{\gamma}_k \quad (2.9)$$

in which  $\mathbf{H} = [\widetilde{\mathbf{MS}}_1, \dots, \widetilde{\mathbf{MS}}_N, \mathbf{P}]$ ,  $\boldsymbol{\gamma}_k = [\gamma_{k,1}, \dots, \gamma_{k,N+1}]^T$  and all the images are reorganized by columns. Optimal MSE estimation of weights vector  $\boldsymbol{\gamma}_k$  would encompass the use of the unavailable target image  $\widetilde{\mathbf{MS}}_k$  and is thus performed at a reduced scale. Consequently, the solution is found as

$$\boldsymbol{\gamma}_b = (\mathbf{H}_d^T \mathbf{H}_d)^{-1} \mathbf{H}_d^T \left( \widetilde{\mathbf{MS}}_k - \widetilde{\mathbf{MS}}_k^{LP} \right), \quad (2.10)$$

where  $\mathbf{H}_d$  is the reduced scale version of  $\mathbf{H}$  and  $\widetilde{\mathbf{MS}}_k^{LP}$  is a version of  $\widetilde{\mathbf{MS}}_k$  obtained by filtering via the Modulation Transfer Function (MTF) of the MS sensor.

Another algorithm belonging to this category has been proposed in [45]. In this case, only the weight vector  $\mathbf{w}$  is adaptively set by minimizing the MSE between  $\mathbf{I}_L$  and a filtered and down-sampled version of the panchromatic image. Moreover, the authors exploit in this paper the concept of partial replacement and, for this reason, we will call this method *Partial Replacement Adaptive CS* (PRACS), from now on. Specifically, they construct each component of the high spatial resolution image as a convex combination of PAN and MS images with a proper weight, in order to meet the spectral characteristics of the individual MS bands. Therefore, for  $k = 1, \dots, N$ , the new, band-dependent, panchromatic image  $\mathbf{P}_k^{new}$  is achieved as:

$$\mathbf{P}_k^{new} = CC(\mathbf{I}_L, \widetilde{\mathbf{MS}}_k) \cdot \mathbf{P} + [1 - CC(\mathbf{I}_L, \widetilde{\mathbf{MS}}_k)] \cdot \widetilde{\mathbf{MS}}_k, \quad (2.11)$$

where  $CC(\mathbf{X}, \mathbf{Y})$  is the correlation coefficient between images  $\mathbf{X}$  and  $\mathbf{Y}$ . Finally, the weights  $\{g_k\}_{k=1, \dots, N}$  are obtained according to

$$g_k = \beta \cdot CC(\mathbf{I}_{L,k}^{new}, \widetilde{\mathbf{MS}}_k) \frac{std(\widetilde{\mathbf{MS}}_k)}{\sum_{i=1}^N std(\widetilde{\mathbf{MS}}_i)} L_k, \quad k = 1, \dots, N, \quad (2.12)$$

which is given by the product of: 1) a parameter  $\beta$  that normalizes the high frequencies, so that they lie in the right dynamic range; 2) a correlation coefficient that adjusts the relative magnitude of the high-frequency information and has the purpose of minimizing the global dissimilarity between the final low spatial resolution image  $\mathbf{I}_{L,k}^{new}$  achieved by spatially degrading  $\mathbf{P}_k^{new}$  and each MS band; 3) a coefficient that takes into account the spectral distortion due to the differences in standard deviation among the MS bands; 4) an adaptive factor  $L_k$  defined according to

$$L_k = 1 - \left| 1 - CC(\mathbf{I}_L, \widetilde{\mathbf{MS}}_k) \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{I}_{L,k}^{new}} \right| \quad (2.13)$$

and aimed to remove the local spectral instability error between the synthetic component image and the MS band. Due to the partial replacement, both the flowchart showing the CS fusion procedure in Fig. 2.1 and Eq. (2.1) must be slightly modified to incorporate it. In particular, in Eq. (2.1) the role of the original panchromatic image  $\mathbf{P}$  is played by  $\mathbf{P}_k^{new}$ .

### 2.2.3 Multiresolution Analysis

In the second class of pansharpening methods the contribution of the PAN image to the fused product is achieved by calculating the difference between  $\mathbf{P}$  and its low pass version  $\mathbf{P}_L$ . Namely the HRMS image is defined as

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + g_k (\mathbf{P} - \mathbf{P}_L), \quad k = 1, \dots, N. \quad (2.14)$$

In [52] this paradigm has been denoted as *Amélioration de la Résolution Spatiale par Injection de Structures (ARSIS)* underlining that the purposes of these methods are the preservation of the whole content of the LRMS image and the addition of further information, obtained from the PAN image through spatial filtering.

According to definition (2.14) the different approaches belonging to this class are uniquely characterized by the algorithm employed for obtaining the image  $\mathbf{P}_L$  and by the combination coefficients  $\{g_k\}_{k=1, \dots, N}$ .

In a very general setting,  $\mathbf{P}_L$  is achieved through an iterative decomposition scheme (called multiresolution analysis) aiming at constructing a sequence of signals with successively reduced information (or *pyramid*), through the repeated application of the analysis operators. The type of decomposition constitutes a first distinguishing feature of the various approaches and can range from very simple methods based on a single level decomposition to more complex techniques relying upon a multilevel analysis.

In addition, the mathematical expression of  $\mathbf{g}$  assumes different forms in the literature, among which two options are the most

commonly reported. By setting all the coefficients equal to 1 the *additive* injection scheme is achieved:

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + (\mathbf{P} - \mathbf{P}_L), \quad k = 1, \dots, N. \quad (2.15)$$

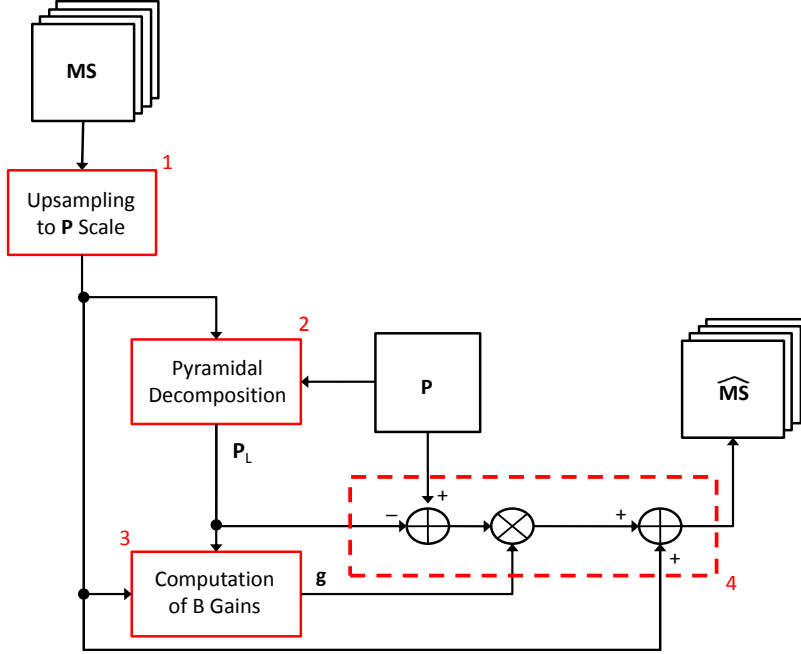
The *High Pass Modulation* (HPM) method (also called *multiplicative* injection scheme) is the other well-known paradigm, defined through the expression

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{P}_L} (\mathbf{P} - \mathbf{P}_L), \quad k = 1, \dots, N, \quad (2.16)$$

in which the details are weighted by the ratio of the MS and the PAN low pass version  $\mathbf{P}_L$ , with the aim of reproducing, in the fused image, the local intensity contrast of the PAN [53]. Furthermore, if a unique low pass image  $\mathbf{P}_L$  is used for all the MS bands, this algorithm reduces the spectral distortion with respect to  $\mathbf{MS}$ , which can be quantified by the *Spectral Angle Mapper* (SAM) (see Sect. 2.3.1), and thus is an element of the *Spectral Distortion Minimization* (SDM) class [54].

The general scheme of MRA fusion methods is reported in Fig. 2.2. The flowchart is slightly simpler than for CS (Fig. 2.1). In fact, as one can notice, it lacks the block aimed at estimating the weight vector  $\mathbf{w}$ , needed for the construction of the equivalent panchromatic image. Accordingly, the required blocks are devoted to: 1) Upsample the MS image to reach the panchromatic scale; 2) Perform the pyramidal decomposition yielding the low pass version  $\mathbf{P}_L$  of the PAN image; 3) Compute the band dependent injection gains  $\{g_k\}_{k=1,\dots,N}$ ; 4) Inject the extracted details according to Eq. (2.14).

Tab. 2.4 reports a brief summary of the implemented MRA-based approaches, evidencing the paradigm used for calculating the injection coefficients required by Eq. (2.14) and the details extraction scheme. More implementation details are provided in the following pages.



**Figure 2.2** Flowchart of a generic pansharpening algorithm belonging to the MRA family.

### Low Pass Filtering

The most direct implementation of the ARSIS concept consists in applying a single Linear Time-Invariant (LTI) Low Pass Filter (LPF)  $h_{LP}$  to the PAN image  $\mathbf{P}$  for obtaining  $\mathbf{P}_L$ , thus resulting in the following formula

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + g_k (\mathbf{P} - \mathbf{P} * h_{LP}), \quad k = 1, \dots, N, \quad (2.17)$$

in which  $*$  denotes the convolution operator.

This approach can be performed by employing several different low pass filters  $h_{LP}$ , among which the most diffused use Box, Gaussian and Laplacian masks [5]. Furthermore, the simple additive injection scheme (Eq.(2.15)) or more elaborated alternatives can be exploited for injecting the spatial details.

Among the possible couples of filters and coefficients, we chose the simplest scheme achievable by using the Box mask (i.e., mask

Method	MRA	$g_k$
<b>HPF</b>	Box Filter	1
<b>SFIM</b>	Box Filter	$\frac{\mathbf{MS}_k}{\mathbf{P}_L}$
<b>Indusion</b>	Decimated WT	1
<b>ATWT</b>	ATWT	1
<b>AWLP</b>	ATWT	$\frac{\mathbf{MS}_k}{\frac{1}{N} \sum_{i=1}^N \mathbf{MS}_i}$
<b>ATWT M2</b>	ATWT + Detail Equal. [52]	1
<b>ATWT M3</b>	ATWT + Detail Equal. [52]	1
<b>GLP</b>	Gaussian Pyramid	1
<b>MTF HPM</b>	Gaussian Pyramid	$\frac{\mathbf{MS}_k}{\mathbf{P}_L}$
<b>MTF HPM PP</b>	Gaussian Pyramid	$\frac{\mathbf{MS}_k}{\mathbf{P}_L}$
<b>GLP-CBD</b>	Gaussian Pyramid	$\frac{\text{cov}(\mathbf{P}_L, \mathbf{MS}_k)}{\text{var}(\mathbf{P}_L)}$

**Table 2.4** Employed MRA scheme and weights in Eq. (2.14) for different MRA-based pansharpening techniques.

with uniform weights, implementing an average filter) and additive injection, which leads to the pansharpening algorithm known as *High Pass Filtering* (HPF) method [8, 12].

The corresponding methodology employing the HPM injection scheme has been proposed in [55] and is named *Smoothing Filter-based Intensity Modulation* (SFIM).

### Pyramidal Decompositions

The resolution reduction needed to obtain the low pass signal  $\mathbf{P}_L$  at the original MS scale can be performed in one or more steps, namely by employing a single low pass filter with cutoff frequency equal to  $1/R$  and decimating by  $R$ , or by multiple fractional steps. This second method, that includes the first as a particular case, is commonly referred to as pyramidal decomposition and dates back to the seminal work of Burt and Adelson [13] which utilized Gaussian LPFs to carry out the analysis steps. The corresponding differential representation, achieved by calculating the differences between the Gaussian pyramid levels, is named *Laplacian pyramid*

and has later been proven to be very valuable for pansharpening purposes [56]. Indeed the Gaussian filters can be tuned to closely match the sensor MTF. This allows to extract from the PAN image the details, which due to the coarser spatial resolution are not seen by the MS sensor [57]. Since the Gaussian mask is defined by a single parameter (i.e., its standard deviation), its frequency response is fully specified by fixing it. To this aim, the value of the amplitude response at the Nyquist frequency is used, being it commonly provided by the manufacturer as a sensor specification, or obtained through on-orbit measurements. However it is useful to keep in mind that components aging can induce a significant incertitude of this parameter, also in the more favorable case of on-orbit estimation.

Both the additive (Eq. (2.15)) and the multiplicative injection scheme (Eq. (2.16)) [58] have been used in the literature and implemented in this work. They will be referred to as GLP [57] and MTF HPM [58], respectively. Following the indications of the authors of [58], we also tested the MTF HPM method followed by a PostProcessing (PP) phase aimed at correcting the noise generated in the presence of discontinuities. For the latter approach, whose details can be found in the cited reference, we use here the acronym MTF HPM PP.

As a further example of Gaussian pyramids exploiting an analysis filter matched with the MS sensor MTF we considered an algorithm that relies upon the optimization of the injection coefficients by least square fitting. In particular, it is based on Eq.(2.14), in which the coefficients are calculated as:

$$g_k = \frac{\text{cov}(\widetilde{\mathbf{MS}}_k, \mathbf{P}_L)}{\text{var}(\mathbf{P}_L)}, \quad (2.18)$$

where, in general,  $\mathbf{P}_L$  depends on the  $k$ -th band. This injection rule, also used by GS approaches [43], leads to a powerful algorithm. It is commonly known as *GLP with Context Based Decision* (GLP-CBD) [42], since the injection coefficient can be locally optimized by patching the image in overlapping zones. Actually, the consequent performance improvement is paid in terms of a

significant increase in the execution time.

In the same field, very appreciable results have been achieved by employing wavelet and contourlet decomposition pyramids [10, 14]. By focusing of the widely diffused wavelet decompositions, the decimated Mallat's method has been initially used for this purpose [52, 59]. As an example of this class, we selected for the comparison the more recent *Indusion* method [30], which profits from multiple equalization steps to achieve remarkable results.

However, the undecimated “à trous” method [60] has soon emerged as a very effective method [61]. Indeed, even if non-orthogonality (which implies that a wavelet plane could retain information for a neighbouring plane) could compromise the spectral quality of the fused product [61], its beneficial characteristics such as the shift-invariance property [62] and the capability of being easily matched to the sensor MTF [57] produce accurate pansharpened image. We adopt here a widely used implementation of the “à trous” filter based on the sequential application (thanks to the separability property), in the vertical and horizontal directions, of 1-D filters [63]

$$h = [1 \quad 4 \quad 6 \quad 4 \quad 1], \quad (2.19)$$

that derives from the choice of a  $B_3$  cubic spline as scaling function [64].

In this case, here denoted as *ATWT*, choosing the additive injection formula (2.15) for achieving the final product leads to the formulation as in [60]. Furthermore, alternative implementations using the *Model 2* and *Model 3* [52] have been added to the comparison and named *ATWT M2* and *ATWT M3*, respectively. They consist in equalizing the first two moments of the details extracted by the PAN images, before injection. In both cases the mean and the standard deviation of the PAN and MS wavelet coefficients at the original MS scale are used. Whilst Model 2 employs a deterministic relation among the corresponding quantities, Model 3 is optimized by least square fitting.

All the previous wavelet methods are based on the choice of unitary injection coefficients  $\{g_k\}_{k=1,\dots,N}$ . However, some further improvements can be achieved by injecting the details using the



HPM paradigm (2.16) [53]. As an example of Wavelet-based method employing a different choice we implemented the *Additive Wavelet Luminance Proportional* (AWLP) [65] that uses the more general fusion formula reported in Eq. (2.14), with the injection coefficients defined as

$$g_k = \frac{\widetilde{\text{MS}}_k}{\frac{1}{N} \sum_{i=1}^N \widetilde{\text{MS}}_i}, \quad k = 1, \dots, N. \quad (2.20)$$

## 2.3 Quality Assessment of Fusion Products

As in most data fusion problems, the absence of a reference image is the main limitation for the evaluation of the results. In the context of pansharpening this lack prevents the direct application of the Wald criteria. Thus, two assessment procedures have been proposed in order to circumvent this problem. The first one considers the images at a spatial resolution lower than the original (induced artificially) and uses the original MS image as a reference. Although this procedure allows for a precise evaluation of the results with established indexes, there might be mismatches between the performances obtained at reduced scale and the quality (interpreted by visual inspection) for the fusion product at the original scale [4]. Indeed, the performances are intrinsically related to the way the resolution degradation is performed, especially in the case of pansharpening methods exploiting spatial filters [6].

The second approach for validation uses quality indexes that do not require a reference image, but operate on the relationships among the original images and the pansharpened product. This approach is appealing since it operates directly on the data at the native scale but is biased by the definition of the indexes.

Due to the sub-optimality of both the quantitative evaluation procedures, a qualitative evaluation of the results through visual inspection is still advisable. This is especially useful for appreciat-

ing local spectral distortions and precision in rendering the spatial details in the fused images [5].

### 2.3.1 Reduced Scale Assessment

The procedure operating at reduced scale is mainly based on the Wald protocol [36] that is composed by the following three requirements:

1. Any fused synthetic image  $\widehat{\mathbf{MS}}_k$ , where  $k$  ranges from 1 to the number of available channels  $N$ , once degraded to its original resolution, should be as identical as possible to the original image  $\mathbf{MS}_k$ .
2. Any fused synthetic image  $\widehat{\mathbf{MS}}_k$  should be as identical as possible to the image  $\mathbf{HRMS}_k$  that the corresponding sensor would observe with the highest resolution.
3. The multispectral set of synthetic images  $\widehat{\mathbf{MS}} = \{\widehat{\mathbf{MS}}_k\}_{k=1,\dots,N}$  should be as identical as possible to the multispectral set of images  $\mathbf{HRMS} = \{\mathbf{HRMS}_k\}_{k=1,\dots,N}$  that the corresponding sensor would observe with the highest resolution.

Considering the images at reduced scale allows one to easily check the synthesis property of the Wald protocol, expressed by the second and third statements, since all the required quantities, namely the fusing images, obtained by degrading the available MS and PAN images, and the reference, represented by the original MS image, are made accessible in this way.

More in detail, the degradation of the resolution is obtained by applying to the available MS and PAN images a low pass filter and a decimation operator characterized by a sampling factor equal to the resolution ratio between the two images. Let us denote the reduced scale multispectral and panchromatic images by  $\mathbf{MS}^*$  and  $\mathbf{P}^*$ , respectively.

Clearly, in this validation protocol, the choice of the filter is crucial. In general, the filter is defined for ensuring the *consistency* (defined by the first Wald's statement) of the pansharpening process. Since the pansharpened image (that here should match as close as possible the original image  $\mathbf{MS}$ ), once degraded to its original resolution, should be identical to the original multispectral image (whose part is acted by  $\mathbf{MS}^*$ ), it comes natural that the resolution reduction has to be performed by employing a filter simulating the transfer function of the remote sensor. In other terms the degradation filter has to match the MTF of the sensor [57]. In addition, the filter used for obtaining the panchromatic image  $\mathbf{P}^*$  has to be designed in order to preserve the details that would have been seen if the image were acquired at the reduced resolution. Accordingly, it is commonly degraded through an ideal filter [57].

### Quality Indexes

Several indexes have been proposed for evaluating spatial and spectral distortion of the fused product with respect to an available reference image. According to the Wald's procedure both scalar (i.e., measurements on a single spectral band) and vector (which take into account of all the spectral bands) similarity indexes are required. The most widely used ones are briefly described in the following.

The vector measures are useful to quantify the spectral distortion. A simple index that has assumed a key role in the technical literature is the *Spectral Angle Mapper* (SAM) [66], which consists in calculating the angle between the corresponding pixels of the fused and reference images in the space defined by considering each spectral band as axis. Let  $\mathbf{I}_{\{n\}} = [I_{1,\{n\}}, \dots, I_{N,\{n\}}]$  be a pixel vector of the MS image  $\mathbf{I}$  with  $N$  bands, the SAM between  $\mathbf{I}_{\{i\}}$  and  $\mathbf{J}_{\{i\}}$  is defined as:

$$SAM(\mathbf{I}_{\{i\}}, \mathbf{J}_{\{i\}}) = \arccos \left( \frac{\langle \mathbf{I}_{\{i\}}, \mathbf{J}_{\{i\}} \rangle}{\|\mathbf{I}_{\{i\}}\| \|\mathbf{J}_{\{i\}}\|} \right) \quad (2.21)$$

in which  $\langle \cdot, \cdot \rangle$  denotes the scalar product (or inner product) and  $\|\cdot\|$  the vector  $\ell_2$  norm. The global value of SAM for the whole image is obtained by averaging the single measures over all the pixels. The optimal value of the SAM index is 0.

Other indexes account for radiometric distortions. One of the most popular quantity is the *Root Mean Square Error* (*RMSE*), which is defined as:

$$RMSE(\mathbf{I}, \mathbf{J}) = \sqrt{E[(\mathbf{I} - \mathbf{J})^2]}. \quad (2.22)$$

The ideal value of the *RMSE* is zero, and is achieved if and only if  $\mathbf{I} = \mathbf{J}$ ; however its use in visual applications has been strongly criticized [41].

A more credited global index is the *Erreur Relative Globale Adimensionnelle de Synthèse* (*ERGAS*) that was properly proposed for pansharpening in [52] and is defined as:

$$ERGAS = \frac{100}{R} \sqrt{\frac{1}{N} \sum_{k=1}^N \left( \frac{RMSE(\mathbf{I}_k, \mathbf{J}_k)}{\mu(\mathbf{I}_k)} \right)^2}, \quad (2.23)$$

where the *RMSE* is defined as in Eq.(2.22) and  $\mu$  represents the mean of the image. Since the *ERGAS* is composed by a sum of *RMSE* values, its optimal value is 0.

Another index which was developed to overcome some limitations of the *RMSE* is the *Universal Image Quality Index* (*UIQI*) or *Q-index*, proposed by Wang and Bovik [67]. Its physical interpretation becomes straightforward by writing its expression in the form:

$$Q(\mathbf{I}, \mathbf{J}) = \frac{\sigma_{\mathbf{I}\mathbf{J}}}{\sigma_{\mathbf{I}}\sigma_{\mathbf{J}}} \frac{2\bar{\mathbf{I}}\bar{\mathbf{J}}}{(\bar{\mathbf{I}})^2 + (\bar{\mathbf{J}})^2} \frac{2\sigma_{\mathbf{I}}\sigma_{\mathbf{J}}}{(\sigma_{\mathbf{I}}^2 + \sigma_{\mathbf{J}}^2)}, \quad (2.24)$$

that comprises an evaluation of the correlation coefficient, the difference between the mean luminances and the similarity of the contrasts. The *Q-index* varies in the range  $[-1, 1]$ , with 1 denoting the best quality.

In addition, an extension of the *Q-index* for accounting also spectral distortions has been proposed in [68]. Specifically, the *Q*

vector index can be used with datasets composed by four spectral bands and is based on modeling each pixel  $\mathbf{I}_{\{i\}}$  as a *quaternion*

$$\mathbf{I}_{\{i\}} = \mathbf{I}_{\{i\},1} + \mathbf{i}\mathbf{I}_{\{i\},2} + \mathbf{j}\mathbf{I}_{\{i\},3} + \mathbf{k}\mathbf{I}_{\{i\},4}. \quad (2.25)$$

The index is calculated by Eq.(2.24), after replacing the vector operations with the corresponding quaternion operations. A generalization of this index, allowing for the evaluation of images with a number of spectral bands different from four, is presented in [69] and is called  $Q\mathcal{Q}^n$ -index.

### 2.3.2 Full Scale Validation

In order to perform the quality evaluation at the original resolution of the data, the *Quality w/No Reference* (QNR) index [39] was proposed. The QNR index is defined as:

$$QNR = (1 - D_\lambda)^\alpha (1 - D_S)^\beta. \quad (2.26)$$

It is thus composed by the product, weighted through the coefficients  $\alpha$  and  $\beta$ , of two separate values  $D_\lambda$  and  $D_S$ , which quantify the spectral and the spatial distortion, respectively. Higher is the QNR index and better should be the quality of the fused product. The maximum theoretical value of this index is 1 when both  $D_\lambda$  and  $D_S$  are equal to 0.

The spectral distortion is estimated by

$$D_\lambda = \sqrt[p]{\frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \left| Q(\mathbf{MS}_i, \mathbf{MS}_j) - Q(\widehat{\mathbf{MS}}_i, \widehat{\mathbf{MS}}_j) \right|^p}. \quad (2.27)$$

This formulation aims at producing a synthetic image with the same spectral features of the original MS image. Accordingly, the relations among the MS bands have to be preserved during the enhancement procedure. The *Q-index* is used to calculate the dissimilarities between couples of bands and the parameter  $p$  is typically set to one [39].

The spatial distortion is calculated by

$$D_S = \sqrt[q]{\frac{1}{N} \sum_{i=1}^N \left| Q(\widehat{\mathbf{MS}}_i, \mathbf{P}) - Q(\mathbf{MS}_i, \mathbf{P}_{LP}) \right|^q} \quad (2.28)$$

where  $\mathbf{P}_{LP}$  is a low resolution panchromatic image at the same scale of the MS image and  $q$  is usually set to one [39]. From a practical point of view, the perfect alignment between the upsampled version of the MS and the PAN images should be assured to avoid the loss of meaning for this quality index.

## Chapter 3

# Pansharpening Algorithms: A Critical Comparison

In the last decades many pansharpening algorithms have been presented in the literature. However, the lack of universally recognized evaluation criteria, image datasets for benchmarking and public implementations of the algorithms make difficult to achieve a thorough evaluation and comparison of the different pansharpening techniques. We try to fill this gap with this work by providing a critical description and extended comparison of some of the main state-of-art pansharpening methods. More in detail, several pansharpening techniques belonging to the Component Substitution or MultiResolution Analysis families are analyzed. These techniques are evaluated both at reduced and full scale. Five datasets acquired by different satellites (i.e., IKONOS, WorldView-2, Pléiades) allow a detailed comparison of the algorithms, characterization of their performances and consistency between the two main validation procedures. In addition, the implementation of all the pansharpening techniques considered in this chapter and the framework used for running the experiments are collected in a MATLAB Toolbox that will be provided to the Community.

The remainder of the chapter is as follows. In Sec. 3.1 a brief introduction to the pansharpening MATLAB Toolbox is done,

while, Sec. 3.2 and Sec. 3.3 describe the exploited datasets and the algorithms, respectively. Finally, in Sec. 3.4 an extended experimental phase is shown.

## 3.1 The Pansharpening Toolbox

The Pansharpening Toolbox has been developed thanks to international collaborations among different universities and research centers between Italy and France. In greater details, the Universities of Florence and Siena have been involved by means of Luciano Alparone and Andrea Garzelli, respectively. Furthermore, we want to thank Lucien Wald (MINES ParisTech, Paris), Jocelyn Chanussot, Mauro Dalla Mura and Giorgio Licciardi (GIPSA-Lab, Grenoble) for the support in developing the above-mentioned Toolbox.

The Toolbox has been developed in MATLAB with the aim of aiding the new developments in this topic thanks to an easier comparison with state-of-art pansharpening approaches by comparing them on the same datasets in the same conditions (for instance, with same preprocessing, if required) in order to understand the real powerful of each approach by analyzing the most relevant characteristics of the different algorithms (e.g., the filters exploited in the extraction detail phase, the injection model and so forth). As far as we know, this is the first time that so many state-of-art pansharpening approaches are compared on same datasets under same conditions. Moreover, after its spread on the net, it can be used to facilitate the standardization of the procedures for evaluating the fused products, which represent critical points in this research field.

The Toolbox allows two kinds of validations for the final products: At reduced and full scale. These procedures are well described in the previous chapter together with the indexes exploited for quantifying the goodness of the fused images. It is possible to run default test cases, which exploit different datasets acquired by different sensors. The sensors involved are: IKONOS, WorldView-



2 and Pléiades. These sensors acquire both panchromatic and multispectral data except for the Pléiades sensor in which case the panchromatic has been simulated (further details on the scenes acquired and the characteristics of the exploited sensors can be found in the next section). Furthermore, it is very simple to incorporate new test cases and run the proposed algorithms.

The Toolbox is very flexible and it is simple to add a new pansharpening algorithm by exploiting pre-processing functions, as the interpolation of the MS image to the dimension of the PAN image and the validation procedures and indexes to properly quantify the goodness of the new developed approach. The standard outputs of a pansharpening algorithm into the software chain are: The execution time of the method, the values of the quality indexes (which depend on the validation procedure exploited) and the fused image (in this case, even an eps printable image can be directly provided to the user).

## 3.2 Datasets

Five datasets acquired by different sensors were considered. Their characteristics are detailed in the following.

**China dataset** The employed dataset<sup>1</sup> represents a mountainous and vegetated area of the China-Sichuan region. The dataset has a size equal to  $300 \times 300$  pixels. It is acquired by the IKONOS sensor which works in the visible and near infrared spectrum range; the multispectral sensor is characterized by four bands (Blue, Green, Red and NIR) and also a panchromatic channel is available. The resolution cell is  $4 \times 4$  meters for the multispectral bands and  $1 \times 1$  meters for the panchromatic channel. The resolution ratio  $R$  is therefore equal to 4.

**Toulouse dataset** This dataset represents an urban area of the city of Toulouse (France). The size is equal to  $512 \times 512$  pixels. The

---

<sup>1</sup>The dataset is available at <http://glcf.umiacs.umd.edu>

sensor used to acquire the images is IKONOS, and, therefore, it has the same spectral and spatial characteristics as in the previous case.

**Rome dataset** This represents an urban area of the city of Rome (Italy). The dimension of the image is  $300 \times 300$  pixels. The dataset has been acquired by the WorldView-2 sensor which provides a high resolution panchromatic channel and eight MS bands. Four standard colors (red, green, blue, and near-infrared 1) and four new bands (coastal, yellow, red edge, and near-infrared 2) are acquired. The *Ground Sample Distance* (GSD) for the panchromatic is 0.46 meters at Nadir and 0.52 meters at  $20^\circ$  off-nadir, while the GSD for the Multispectral sensor is 1.8 meters at Nadir and 2.4 meters GSD at  $20^\circ$  off-nadir. It is worth noting that images are resampled to 0.5 meters due to the US Government policies. The resolution ratio  $R$  is again equal to 4.

**Rio dataset** This represents an urban area of the city of Rio de Janeiro (Brazil). The dimension of the image is  $256 \times 256$  pixels. The dataset has been acquired by the WorldView-2 sensor, and, therefore, it has the same instrumental characteristics as in the previous case.

**Pléiades dataset** The *Pléiades dataset*, collected by an aerial platform, was used for the 2006 contest [42] and was provided by CNES, the French space agency. This represents an urban area of Toulouse (France) and it has a size equal to  $1024 \times 1024$  pixels. The resolution of the four MS bands is 60 cm, but the dataset misses the availability of the panchromatic image, since the corresponding sensor was under development. Accordingly, the high-resolution panchromatic data was simulated by the following procedure. The green and red channels were averaged and the result was filtered with a system characterized by the nominal Modulation Transfer Function (MTF) of the panchromatic sensor. After the resampling to 80 cm, thermal noise was added and the final simulated image was achieved by inverse filtering and wavelet

denoising. As a consequence, also the low resolution MS image were simulated according to the Wald protocol, namely by MTF-filtering and decimation. A resolution ratio  $R$  of 4 was chosen.

### 3.3 Algorithms

The algorithms used for the comparison are listed below. It is worth to underline that for MRA approaches a preprocessing step is often required. Hence, the matching of the two first moments between the panchromatic image and each MS band is performed before applying the fusion algorithm [5]. Moreover, when no authors' specifications are provided, the interpolation of the original multispectral image  $\mathbf{MS}$  for generating the upsampled image  $\widetilde{\mathbf{MS}}$  is carried out by using a polynomial kernel with 23 coefficients [70]. A more detailed description of the methods can be found in the relative references or in Sec. 2.2:

- *EXP*: MS image interpolation, using a polynomial kernel with 23 coefficients [70]
- *PCA*: *Principal Component Analysis* [8]
- *IHS*: *Fast Intensity-Hue-Saturation* image fusion [46]
- *Brovey*: Brovey Transform [48]
- *BDSD*: *Band-Dependent Spatial-Detail* [51]
- *GS*: *Gram Schmidt* [11]
- *GSA*: *Gram Schmidt Adaptive* [43]
- *PRACS*: *Partial Replacement Adaptive Component Substitution* [45]
- *HPF*: Box filter and *High Pass Filter* [8]
- *SFIM*: Box filter and *High Pass Modulation*, also called *Smoothing Filter-based Intensity Modulation* [12, 55]

- *Indusion*: Decimated Wavelet Transform using *additive* injection model [30]
- *ATWT*: *Additive A Trous Wavelet Transform* [60]
- *AWLP*: *Additive Wavelet Luminance Proportional* [65]
- *ATWT M2*: A Trous Wavelet Transform using the Model 2 proposed in [52]
- *ATWT M3*: A Trous Wavelet Transform using the Model 3 proposed in [52]
- *GLP*: *Generalized Laplacian Pyramid* (GLP) [70] with MTF-matched filter [57]
- *MTF HPM*: Gaussian MTF-matched filter [57] with *multiplicative* injection model [58]
- *MTF HPM PP*: Gaussian MTF-matched filter [57] with *multiplicative* injection model and Post-Processing [58]
- *GLP-CBD*: *Generalized Laplacian Pyramid* (GLP) [70] with MTF-matched filter [57] and Gram Schmidt based injection model [42]

## 3.4 Experimental Results

This section is devoted to describe a set of experiments (by exploiting several state-of-art pansharpening approaches), which have been carried out on very high resolution optical data. The results have been validated at both full and reduced scale following the procedures described in Chapter 2.

### 3.4.1 Wald Protocol

Three datasets are considered for evaluating the results at reduced scale: The China, Rome and Pléiades datasets. In Figs. 3.1 - 3.6,

the fusion results of a small area are shown by using an RGB representation. The quality indexes used are the SAM (in degree) to measure the spectral distortion, the ERGAS and the  $Q2^n$  (i.e.,  $Q4$  for four band datasets and  $Q8$  for the WorldView-2 dataset). The quantitative results obtained are tabulated in Tabs. 3.1, 3.2 and 3.3.

A first analysis of the results can be done per family of pan-sharpened algorithms. Among the CS approaches, the adaptive methods, namely the *BDSD* and *PRACS*, achieve the best performances with a good visual appearance of the final products and a reduced spectral distortion with respect to the other members of this family. They are followed by the *GS*, *PCA* and *IHS* approaches, ranked in this order for most of the datasets and the quality indexes. However, for the *China dataset* the *PCA* gives slightly better results than the *GS* approaches. More in general, *PCA* and *IHS*-based approaches have higher spectral distortion, as measured by the *SAM* value and confirmed by the visual inspection.

A more detailed analysis can be done within the algorithms based on Gram-Schmidt Orthogonalization. In fact, *GSA* shows its superiority with respect to the *GS*, thanks to its adaptive weights for building the equivalent panchromatic starting from the MS image.

Considering the techniques belonging to the MRA family, the best accuracy is obtained by the *Gaussian MTF matched filters*, followed by the “à trous” *wavelet transform*. Such good performance can be attributed to the similarity of the frequency response of the filters with that of the MS sensor MTF [57]. On the contrary, the application of the simple *Box filter* leads to modest performances. This is caused by the spatial artifacts due to the presence of ripples in the obscure band in the representation of this filter in the frequency domain and the blur caused by its dissimilarity with respect to the transfer function of the acquisition device. Even in the case of *Indusion*, severe artifacts due to the aliasing introduced by decimation can be noticed (see Figs. 3.1, 3.3 and 3.5).

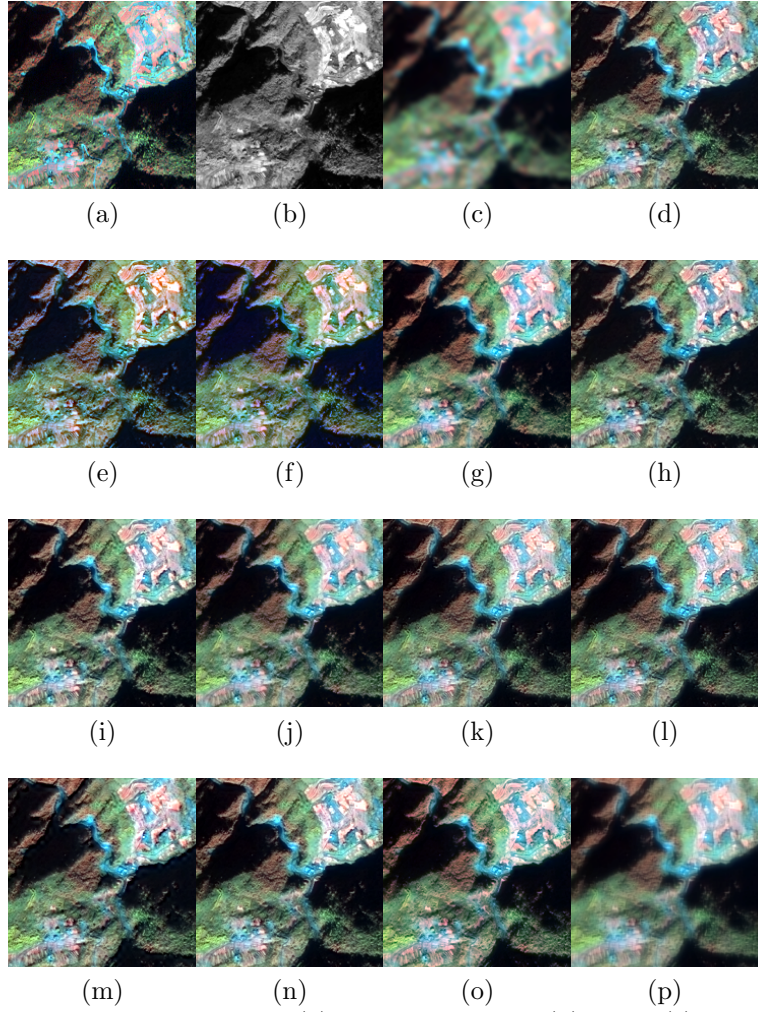
Regarding to the injection methodologies, the *HPM* injection model turns out to be better suited than the *additive* injection model. This can be explained by the flexibility of the *HPM* in setting the local weights that govern the detail injection and its close relationship with the image local contrast (see Chapter 4 for further details).

*ATWT M2* and *ATWT M3* [52] perform poorly both from a numerical and visual (with a more blurred result) points of view when compared to the other approaches that use the same detail extraction filters but different injection methodologies. Finally, it is worth to underline that the results of the *MTF HPM PP* method are, in general, poorer than the ones obtained by *MTF HPM* with no gain due to the *PP* phase and drawbacks provided by the iterative application of *MTF*-based filters and non-idealities in the upsampling step.

Finally, we compare the *CS* and *MRA* families. The former shows a higher spectral distortion, but a better visual appearance of the final product, whereas the latter has opposite characteristics, namely higher spatial distortion and a superior spectral consistency, as shown by the quality indexes and the visual inspection. Accordingly, the overall results are comparable, especially when considering images of four bands. Approaches as *BDSD* and *PRACS* often obtain comparable results with respect to the best performing *MRA* algorithms and in some cases they represent the best choice. On the contrary, in the case of the *Rome dataset* acquired by *WorldView-2* (i.e., an eight band dataset), in general, the gap among algorithms within the two families slightly increases with a greater appeal of *MRA*-based ones. In the *CS* family, only the adaptable approaches (i.e., *GSA*, *BDSD* and *PRACS*) attain reasonable performances due to the capability of well-estimating all the weights to generate the intensity component.

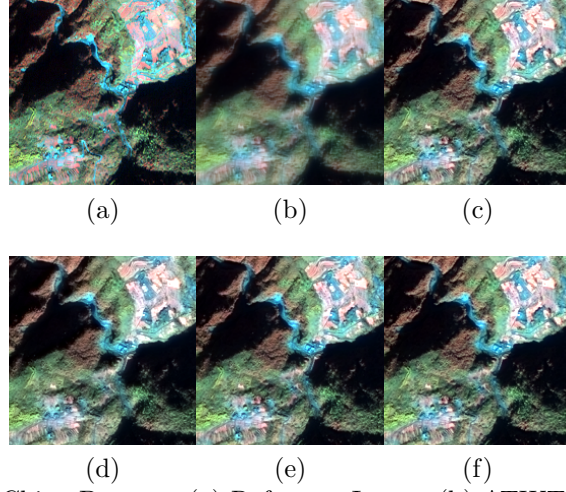
	Q4	SAM(°)	ERGAS
Reference	1	0	0
EXP	0.7398	4.4263	3.8471
PCA	0.8578	3.5433	2.6715
IHS	0.7308	4.9892	3.5766
Brovey	0.7314	4.4263	3.1722
BDSD	<b>0.8869</b>	<b>2.9123</b>	<b>2.4124</b>
GS	0.8500	3.5304	2.7982
GSA	0.8756	2.9889	2.5521
PRACS	0.8793	3.1514	2.5745
HPF	0.8555	3.4817	2.8243
SFIM	0.8582	3.4193	2.7941
Indusion	0.8043	3.9059	3.2846
ATWT	0.8791	3.0786	2.5178
AWLP	0.8830	2.9424	2.4073
ATWT M2	0.8021	4.0493	3.2930
ATWT M3	0.8198	4.3388	3.3357
GLP	0.8787	3.0387	2.5106
MTF HPM PP	0.8643	3.3030	2.7540
MTF HPM	0.8819	3.0041	2.4624
GLP-CBD	0.8780	2.9673	2.5067

Table 3.1 China dataset: Quantitative results.



**Figure 3.1** China Dataset: (a) Reference Image; (b) PAN; (c) EXP; (d) PCA; (e) IHS; (f) Brovey; (g) BDSD; (h) GS; (i) GSA; (j) PRACS; (k) HPF; (l) SFIM; (m) Indusion; (n) ATWT; (o) AWLP; (p) ATWT M2.

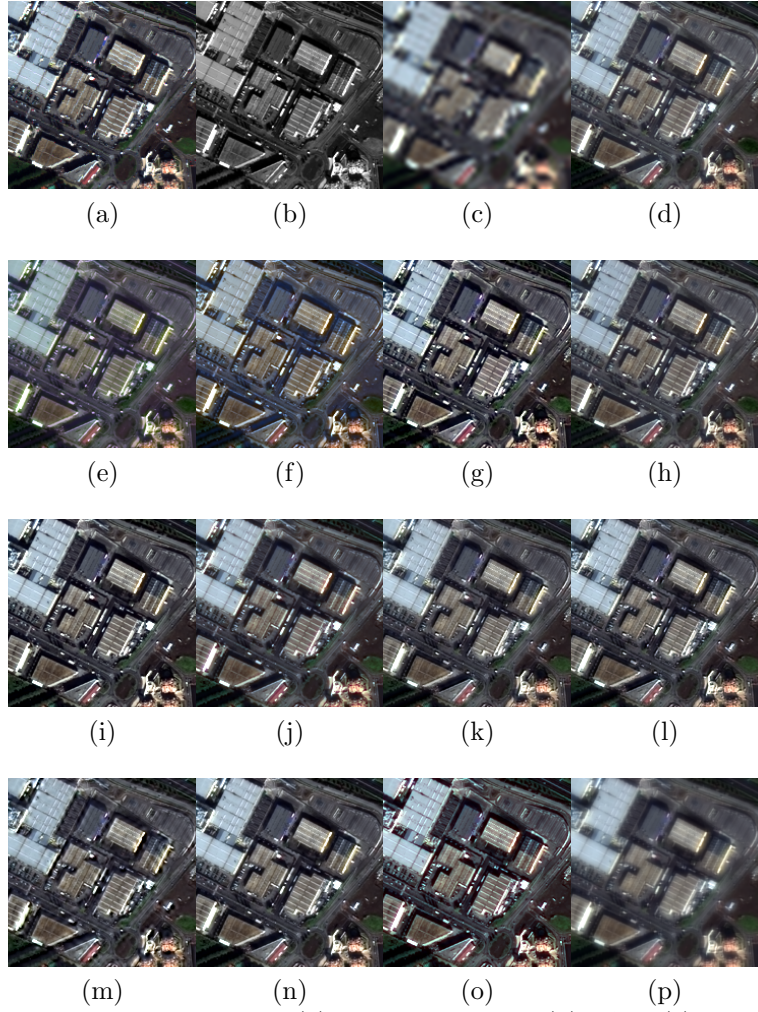




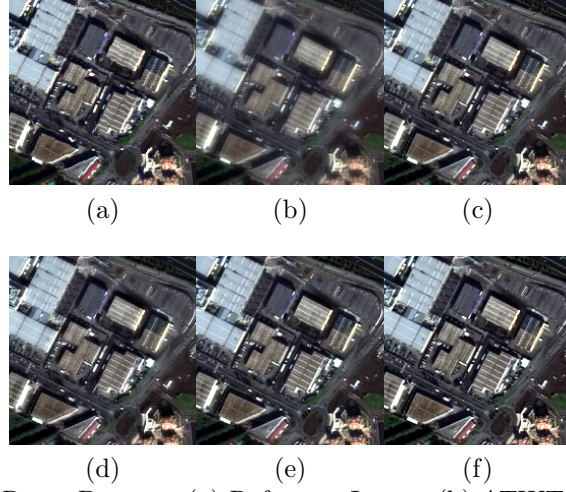
**Figure 3.2** China Dataset: (a) Reference Image; (b) ATWT M3; (c) GLP; (d) MTF HPM PP; (e) MTF HPM; (f) GLP-CBD.

	Q8	SAM( $^{\circ}$ )	ERGAS
<b>Reference</b>	1	0	0
<b>EXP</b>	0.7248	4.9263	5.4171
<b>PCA</b>	0.8169	5.2153	4.4128
<b>IHS</b>	0.7439	5.1455	4.1691
<b>Brovey</b>	0.7487	4.9263	4.1407
<b>BDSD</b>	0.8762	4.8717	3.8619
<b>GS</b>	0.8335	4.8592	4.0144
<b>GSA</b>	0.8907	4.1415	3.4062
<b>PRACS</b>	0.8878	4.6678	3.6768
<b>HPF</b>	0.8715	4.4523	3.8594
<b>SFIM</b>	0.8758	4.2457	3.7591
<b>Indusion</b>	0.8030	5.1415	4.8864
<b>ATWT</b>	0.9013	4.1117	3.3237
<b>AWLP</b>	0.9011	4.5146	3.3572
<b>ATWT M2</b>	0.7969	5.0277	4.6487
<b>ATWT M3</b>	0.8379	5.1042	4.3684
<b>GLP</b>	0.9016	4.0957	3.2982
<b>MTF HPM PP</b>	0.8900	4.3736	3.4777
<b>MTF HPM</b>	<b>0.9092</b>	<b>3.8871</b>	<b>3.1005</b>
<b>GLP-CBD</b>	0.8940	4.1125	3.3479

**Table 3.2** Rome dataset: Quantitative results.



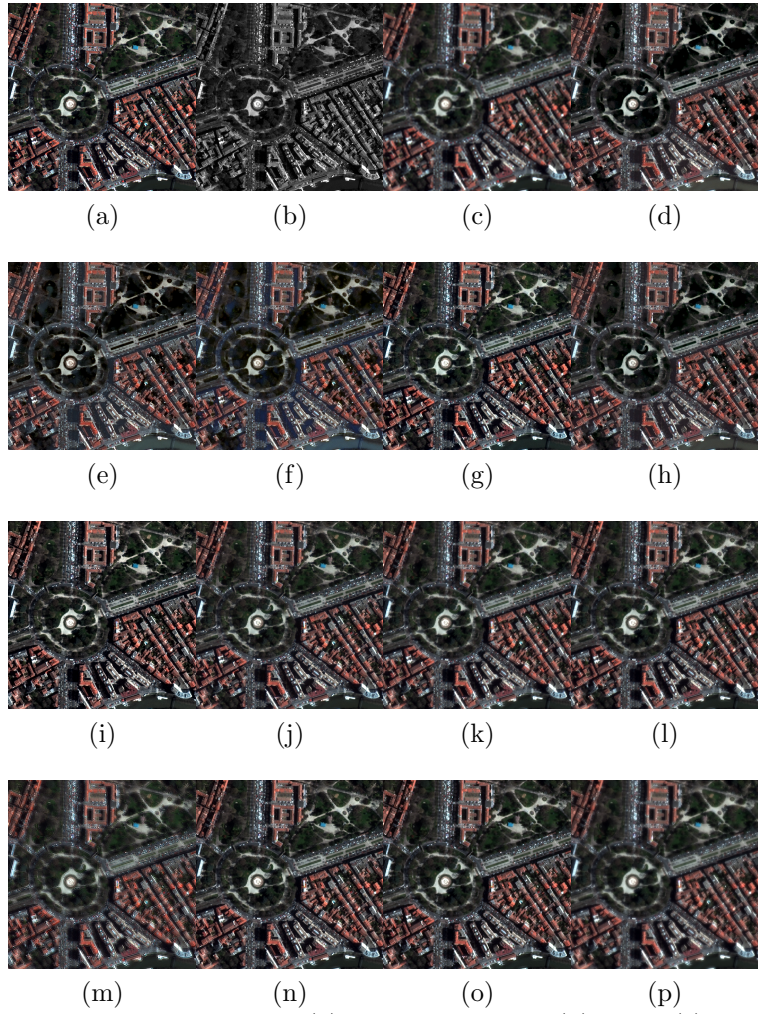
**Figure 3.3** Rome Dataset: (a) Reference Image; (b) PAN; (c) EXP; (d) PCA; (e) IHS; (f) Brovey; (g) BDSD; (h) GS; (i) GSA; (j) PRACS; (k) HPF; (l) SFIM; (m) Indusion; (n) ATWT; (o) AWLP; (p) ATWT M2.



**Figure 3.4** Rome Dataset: (a) Reference Image; (b) ATWT M3; (c) GLP; (d) MTF HPM PP; (e) MTF HPM; (f) GLP-CBD.

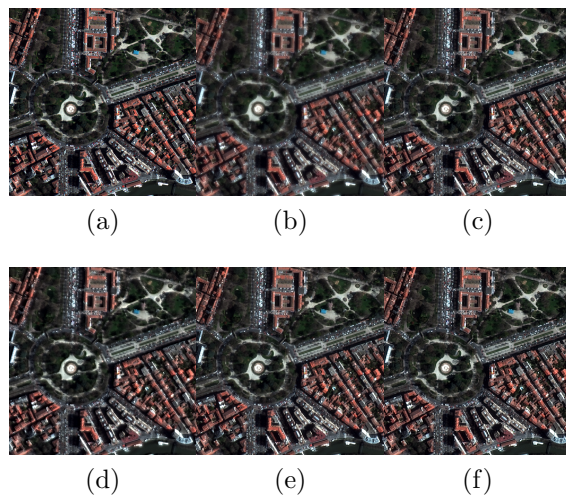
	Q4	SAM( $^{\circ}$ )	ERGAS
<b>Reference</b>	1	0	0
<b>EXP</b>	0.7782	4.6742	6.0826
<b>PCA</b>	0.8122	6.1435	6.0028
<b>IHS</b>	0.8381	5.1788	5.0549
<b>Brovey</b>	0.8411	4.6742	5.1506
<b>BDSD</b>	<b>0.9650</b>	<b>4.0749</b>	<b>2.7811</b>
<b>GS</b>	0.8448	5.3304	5.0468
<b>GSA</b>	0.9572	4.4449	3.0013
<b>PRACS</b>	0.9325	4.5157	3.6523
<b>HPF</b>	0.9041	4.4343	4.2562
<b>SFIM</b>	0.9007	4.3599	4.2923
<b>Indusion</b>	0.8251	5.8322	5.6366
<b>ATWT</b>	0.9479	4.1401	3.3157
<b>AWLP</b>	0.9426	4.3356	3.5219
<b>ATWT M2</b>	0.8300	5.1505	5.4053
<b>ATWT M3</b>	0.8612	5.2915	5.0569
<b>GLP</b>	0.9485	4.1529	3.3193
<b>MTF HPM PP</b>	0.8786	5.2343	5.1103
<b>MTF HPM</b>	0.9488	4.1541	3.2741
<b>GLP-CBD</b>	0.9549	4.4459	3.1521

**Table 3.3** Pléiades dataset: Quantitative results.



**Figure 3.5** Pléiades dataset: (a) Reference Image; (b) PAN; (c) EXP; (d) PCA; (e) IHS; (f) Brovey; (g) BDSD; (h) GS; (i) GSA; (j) PRACS; (k) HPF; (l) SFIM; (m) Indusion; (n) ATWT; (o) AWLP; (p) ATWT M2.





**Figure 3.6** Pléiades dataset: (a) Reference Image; (b) ATWT M3; (c) GLP; (d) MTF HPM PP; (e) MTF HPM; (f) GLP-CBD.

### 3.4.2 Full Scale Validation

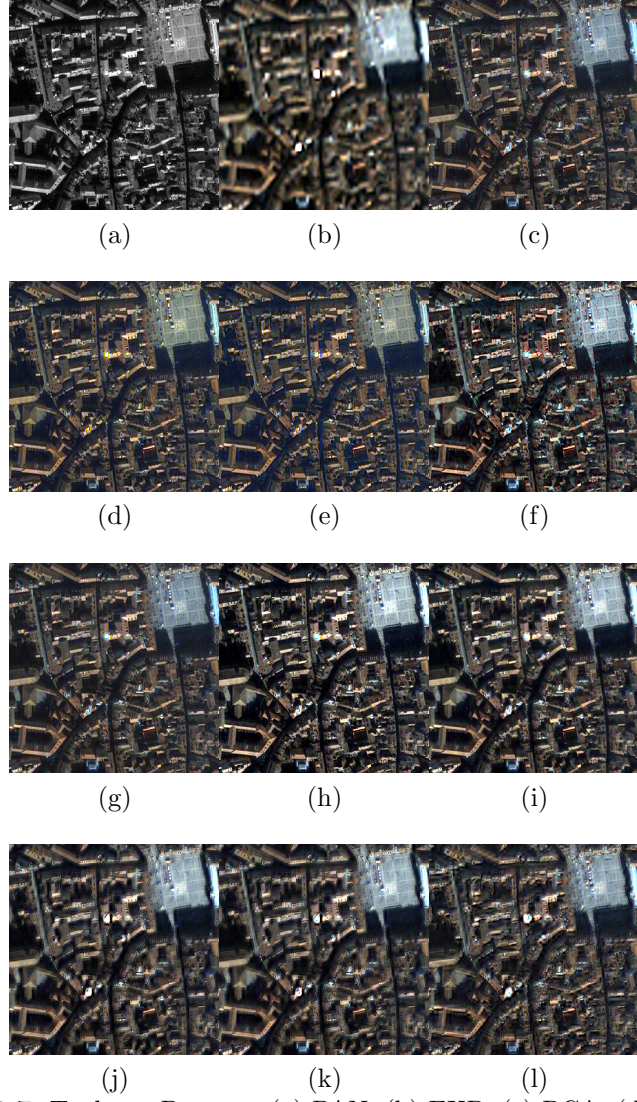
The evaluation of the algorithms at their original scale was performed by employing the Toulouse and the Rio datasets (since Pléiades has a simulated PAN). In Figs. 3.7 - 3.10, the fusion results of a small area are shown by using an RGB representation. Tabs. 3.4 and 3.5 report the values of the  $QNR$ ,  $D_\lambda$  and  $D_S$ . Many results obtained in the analysis performed at reduced scale are in line with those obtained at full scale. In particular, the algorithms within the *MRA* category performing the best are those based on *Gaussian* and “à trous”. The superiority of *HPM*-based injection scheme evidenced at reduced scale is still confirmed here. Within the *CS* family, the best techniques in terms of obtained accuracy are the *adaptive* algorithms followed by the other *CS* approaches. As for the reduced scale, the difference in accuracy between the two main families is more evident in the case of the eight band *Rome dataset*. Again, the algorithms based on *IHS* show a higher spectral distortion confirmed by the values of  $D_\lambda$ .

It is worth to point out that the *PCA* and *IHS*-based approaches perform very well from a quantitative point of view, but these results are not fully confirmed by the visual analysis since the pansharpened images are more spectral distorted if compared to the results of other approaches.

The obtained results give rise to some general considerations. Indeed, the validation at full scale allows to avoid some problems and hypotheses related to the reduce scale protocol, but the values of the indexes are less reliable (since discrepancies between the quantitative values and the visual appearance of the results can be noticed). In more detail, the comparison of algorithms in terms of  $QNR$  index is accurate within each category. However, it is difficult to quantify the existing differences from a visual point of view among algorithms belonging to different families.

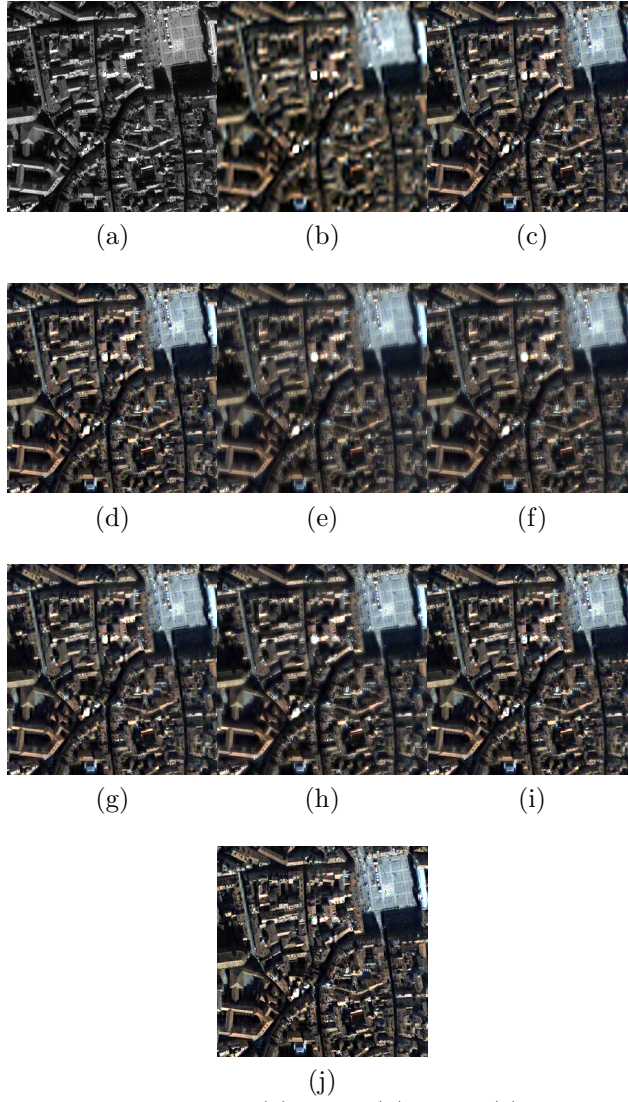
	$D_\lambda$	$D_s$	QNR
<b>EXP</b>	<b>0</b>	0.1975	0.8025
PCA	0.0114	0.0991	0.8906
IHS	0.0367	0.0591	0.9063
Brovey	0.0164	0.0579	0.9266
BDSB	0.0079	0.0404	<b>0.9520</b>
GS	0.0111	0.0912	0.8987
GSA	0.0364	0.0555	0.9101
PRACS	0.0137	0.0612	0.9259
HPF	0.0279	0.0616	0.9122
SFIM	0.0255	0.0606	0.9155
Indusion	0.0274	0.0795	0.8952
ATWT	0.0348	0.0428	0.9239
AWLP	0.0412	0.0392	0.9212
ATWT M2	0.0303	0.0975	0.8752
ATWT M3	0.0544	0.0725	0.8771
GLP	0.0384	0.0406	0.9225
MTF HPM PP	0.0440	0.0407	0.9171
MTF HPM	0.0338	<b>0.0383</b>	0.9292
GLP-CBD	0.0401	0.0484	0.9134

Table 3.4 Toulouse dataset: Quantitative results.



**Figure 3.7** Toulouse Dataset: (a) PAN; (b) EXP; (c) PCA; (d) IHS; (e) Brovey; (f) BDSD; (g) GS; (h) GSA; (i) PRACS; (j) HPF; (k) SFIM; (l) Indusion.

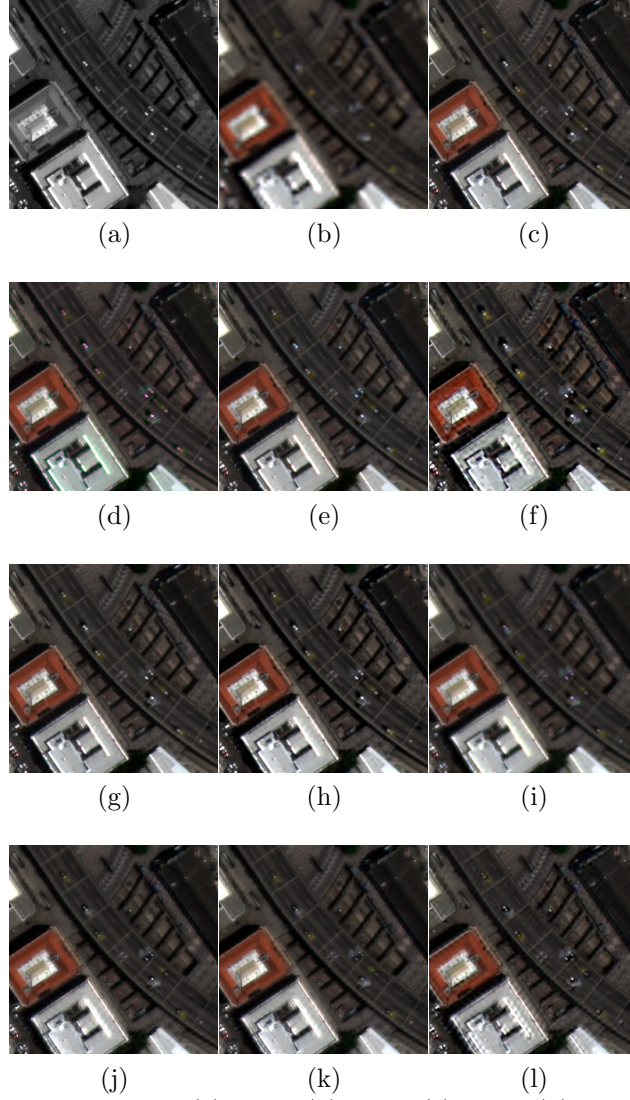




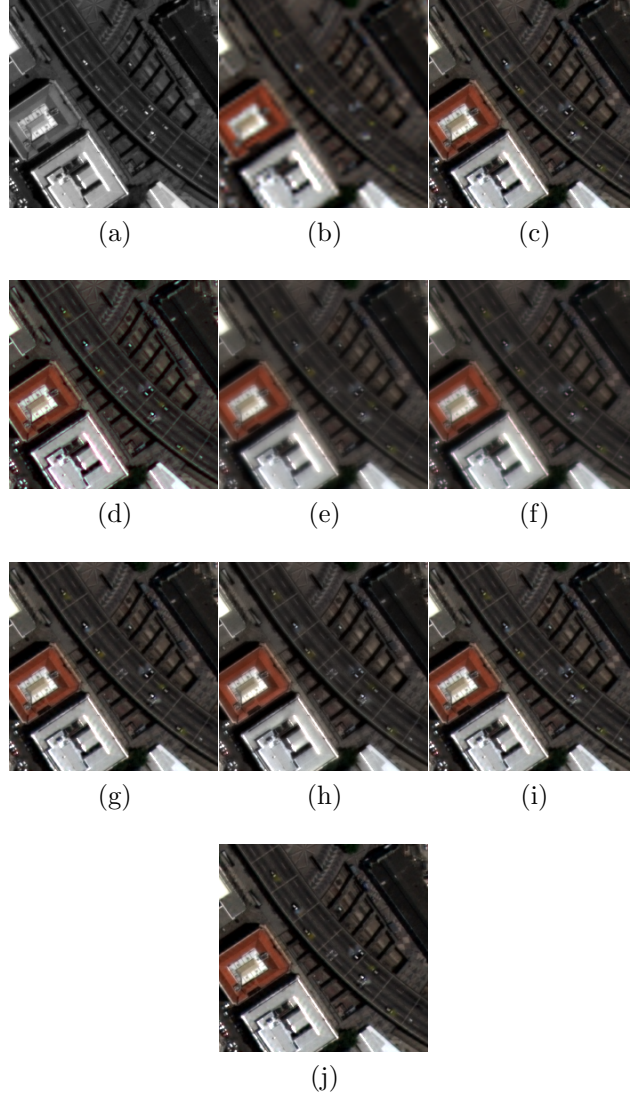
**Figure 3.8** Toulouse Dataset: (a) PAN; (b) EXP; (c) ATWT; (d) AWLP; (e) ATWT M2; (f) ATWT M3; (g) GLP; (h) MTF HPM PP; (i) MTF HPM; (j) GLP-CBD.

	$D_\lambda$	$D_s$	QNR
<b>EXP</b>	<b>0</b>	0.0940	0.9060
<b>PCA</b>	0.0265	0.0582	0.9168
<b>IHS</b>	0.0454	0.0560	0.9012
<b>Brovey</b>	0.0326	0.0615	0.9080
<b>BDSD</b>	0.0272	<b>0.0347</b>	<b>0.9390</b>
<b>GS</b>	0.0261	0.0579	0.9176
<b>GSA</b>	0.0369	0.0552	0.9100
<b>PRACS</b>	0.0158	0.0647	0.9205
<b>HPF</b>	0.0228	0.0619	0.9167
<b>SFIM</b>	0.0213	0.0631	0.9170
<b>Indusion</b>	0.0258	0.0600	0.9157
<b>ATWT</b>	0.0361	0.0474	0.9182
<b>AWLP</b>	0.0399	0.0568	0.9056
<b>ATWT M2</b>	0.0246	0.0906	0.8870
<b>ATWT M3</b>	0.0405	0.0713	0.8911
<b>GLP</b>	0.0359	0.0421	0.9235
<b>MTF HPM PP</b>	0.0525	0.0533	0.8970
<b>MTF HPM</b>	0.0348	0.0423	0.9244
<b>GLP-CBD</b>	0.0375	0.0416	0.9225

Table 3.5 Rio dataset: Quantitative results.



**Figure 3.9** Rio Dataset: (a) PAN; (b) EXP; (c) PCA; (d) IHS; (e) Brovey; (f) BDSD; (g) GS; (h) GSA; (i) PRACS; (j) HPF; (k) SFIM; (l) Indusion.



**Figure 3.10** Rio Dataset: (a) PAN; (b) EXP; (c) ATWT; (d) AWLP; (e) ATWT M2; (f) ATWT M3; (g) GLP; (h) MTF HPM PP; (i) MTF HPM; (j) GLP-CBD.

### 3.4.3 Summary and Discussion

The obtained results give evidence that the good visual appearance and the spectral content preservation represent the main salient features of the *CS* and *MRA* methods, respectively. Accordingly, approaches of the first class aimed at improving the spectral quality and those of the second class properly designed for enhancing the spatial properties obtain the best results. Indeed, very interesting performances are attained by *adaptive CS* approaches (in particular on four band datasets), with the reduction of the spectral distortion, and by some *MRA* algorithms, which benefit of a proper extraction detail phase. In particular the match of the low-pass filter with the sensor's MTF allows to significantly reduce the classical blur of the *MRA* final products.

Another interesting consideration arises from the performed tests. Specifically, for the case of eight band datasets, the performance degradations of some *CS* approaches are clear from both numerical and visual points of view. This is mainly due to the greater spectral distortion. For this reason, *MRA* approaches are, in general, advisable to the growing of the band number.

Furthermore, it is worth underlining that *CS* approaches are often preferable with respect to the *MRA* ones thanks to their robustness to aliasing and misregistration errors. In fact, by considering that they employ an aliasing-free version of the PAN image, it is simple to derive that the fusion rule adopted by the *CS* algorithms is able to compensate the aliasing present in the MS image. On the contrary, the aliasing patterns are visible when the *MRA* fusion rule is adopted [71]. However, such problem can be strongly reduced by employing *MRA* decimated approaches whose analysis filter is matched with the MTF of the MS sensor, as obtainable by properly designing Laplacian pyramids [72]. On the other hand, temporal misalignments suggest the use of *MRA* methodologies with respect to *CS* ones [6]. This kind of robustness turns out to be very helpful when the fusing images are acquired with a time delay, as, for example, when they are provided from sensors mounted on different remote sensing platforms.

From a computational point of view, the *CS* approaches are surely preferable to the *MRA* ones, since the filtering phase significantly slows down the algorithms. According to our tests, the former ones are about ten times faster than the latter ones.

The analysis carried out in this work allowed us to confirm some features of the validation procedures: The reduced scale protocol leads to a very accurate evaluation of the quality indexes, but the scale invariance hypothesis is hardly verified in practice. Furthermore, the operation for generating a synthetic version of the MS image at reduced scale introduces a strong bias in the analyzed algorithms, privileging those employing a similar procedure for extracting the spatial details. On the contrary, the full scale validation gets quite accurate results with respect to algorithms of the same family, whereas being less reliable when the comparison includes algorithms belonging to different classes. This validation approach, which exploits the *QNR* index, has the advantage of validating the products at the original scale, thus avoiding any hypothesis on the behavior at different scales. However, due to the less precision of the quality indexes (since a reference is missing), the results of this analysis can be affected by some mismatches between the quantitative results and the visual appearance of the fused image (e.g., as for the *IHS* method).

## Chapter 4

# The Contrast and Error-based Injection Models

In this chapter, we focus on two of the most popular injection methodologies for panshaping applications. The appealing features of the *multiplicative* scheme are shown and supported by experimental results. In the following, an analysis in the light of physical considerations with respect to the *Point Spread Function* (PSF) of the above-mentioned injection model is provided. Furthermore, the derivation of the constraint on the Modulation Transfer Functions (MTFs) in terms of local contrast is shown. It is worth to be pointed out that the findings in this chapter are mainly based on some considerations, which already underline the superiority of the contrast pyramids with respect to the error ones [73]. Furthermore, by focusing on pansharpening applications, in [58], the authors already point out the goodness of the *High Pass Modulation* (HPM) method from an experimental point of view. The analysis and the relative results presented in this chapter have been already published by the author in [53].

The rest of the chapter is organized as follows. Sec. 4.1 presents the importance of the MTF in pansharpening and highlights its link with the image contrast. In Sec. 4.2, the methods used for

the extraction of details are described and the results of the comparison are reported in Sec. 4.3.

## 4.1 MTF-Preserving Injection Methods

Essentially, the sensors can be approximated as a non-ideal optical systems whose effect can be modeled as the output of a linear spatially invariant system. In these terms, the acquisition system can be fully characterized by means of its spatial impulse response, or PSF [12]. In this thesis we leverage on this aspect and we explicitly express the pansharpening objective of having the same spatial resolution for both the  $\mathbf{P}$  and the enriched  $\mathbf{MS}$  by imposing the following equality:

$$\text{PSF}_{\widehat{\mathbf{MS}}} = \text{PSF}_{\mathbf{P}}. \quad (4.1)$$

We recall that  $\mathcal{F}\{\text{PSF}\}$ , the frequency representation of the PSF, is defined as the product of the MTF and the Phase Transfer Function (PTF), which are respectively the modulus and the phase of the PSF in the Fourier domain. The equality in Eq. (4.1) can be approximated by matching the MTFs in the frequency domain, since the modulus of the PSF constitutes the main contribution to the overall response of the acquisition system [57].

The goal of this work is to explore the possibility of obtaining the desired spatial properties in  $\widehat{\mathbf{MS}}$  by mimicking the response of the PAN sensor. More in detail we focus on the MTF and propose its use for driving the design of the pansharpening algorithms. Starting from Eq. (4.1) we impose for all bands  $k = 1, \dots, N$  of the  $\mathbf{MS}$  the equality

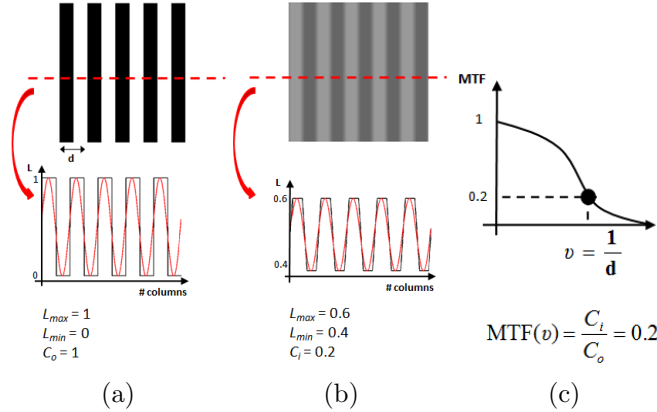
$$\text{MTF}_{\widehat{\mathbf{MS}}_k}(\nu) = \text{MTF}_{\mathbf{P}}(\nu), \quad (4.2)$$

where  $\nu$  refers to the spatial frequency.

The MTF is defined as

$$\text{MTF}(\nu) = \left. \frac{C_i}{C_o} \right|_{o=s(\nu)}, \quad (4.3)$$





**Figure 4.1** Example of estimation of the MTF at a given frequency  $\nu$  according to Eq. (4.4). (a) Periodic pattern (specimen) at frequency  $\nu$ ; (b) resulting image as acquired by the optical system; (c) estimation of the MTF at frequency  $\nu$ .

where  $C_o$  is the contrast of the target object  $o$  and  $C_i$  the contrast measured from the image of the object acquired by the optical system, both typically normalized by the modulation at zero frequency. Let us recall that according to Michelson [74], the contrast  $C$ , denoted also as *modulation*, is

$$C = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}, \quad (4.4)$$

where  $L_{max}$  and  $L_{min}$  are the maximum and minimum luminance (or equivalently radiance when considering radiometric measures) in the pattern, respectively.

Figure 4.1 shows an example of the practical estimation of the MTF (Fig. 4.1(c)) at a given frequency  $\nu$ , performed by computing the contrast according to Eq. (4.4) on the object (Fig. 4.1(a)) and acquired image (Fig. 4.1(b)), by employing a binary square wave with fundamental period  $d = \nu^{-1}$ .

By considering several specimens  $o(\nu)$  with differing spatial frequencies  $\nu$ , it can be shown that the image modulation varies as a function of the spatial frequency. Consequently Eq. (4.2) can

be rewritten as

$$\frac{C_{\widehat{\mathbf{MS}}}}{C_o} = \frac{C_{\mathbf{P}}}{C_o}, \quad (4.5)$$

that implies the relation between the contrast of  $\mathbf{P}$  and  $\widehat{\mathbf{MS}}$  images to be

$$C_{\widehat{\mathbf{MS}}} = C_{\mathbf{P}}. \quad (4.6)$$

When dealing with real acquisitions, the imaged scenes are typically composed by small objects on a background of uniform luminance [12]. In that case the peak luminance is that of the object to resolve, say  $L$ , and the mean luminance coincides with that of the background, say  $L_b$ . In this scenario, the contrast can be computed by *Weber's formula* [75] as

$$C = \frac{L - L_b}{L_b} = \frac{L}{L_b} - 1. \quad (4.7)$$

This transforms the contrast equality (4.6) into the relationship, holding for all bands  $k = 1, \dots, N$ ,

$$\frac{\widehat{\mathbf{MS}}_k}{\widehat{\mathbf{MS}}_k^L} - \mathbf{I} = \frac{\mathbf{P}}{\mathbf{P}_L} - \mathbf{I}, \quad (4.8)$$

with  $\mathbf{I}$  the identity matrix,  $\widehat{\mathbf{MS}}^L$  and  $\mathbf{P}_L$  the background luminance of  $\widehat{\mathbf{MS}}$  and  $\mathbf{P}$ , respectively. The latter has been denoted by  $\mathbf{P}_L$  since it can be computed by degrading the original image with a low pass filter, which might be different for each band  $k$ .

According to the Wald protocol, the low pass version of  $\widehat{\mathbf{MS}}$  is constituted by the original MS image and thus Eq. (4.2) turns into the well known HPM formulation [58]:

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k \cdot \frac{\mathbf{P}}{\mathbf{P}_L}, \quad (4.9)$$

in which  $\widetilde{\mathbf{MS}}$ , an upsampled version of  $\mathbf{MS}$ , is used for guaranteeing the compatibility of the image sizes. Eq. (4.9) characterizes the fusion methods employing Ratio Of Low-Pass decompositions

(ROLP), whose superiority with respect to Difference Of Low-Pass (DOLP) approaches (even called *additive*; see Sec. 2.2.3) that are based on the formula

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + (\mathbf{P} - \mathbf{P}_L) \quad (4.10)$$

has been proven in [73].

By defining the details  $\mathbf{D}$  of the PAN image as

$$\mathbf{D} = \mathbf{P} - \mathbf{P}_L, \quad (4.11)$$

Eq. (4.9) can be also rewritten as

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{P}_L} \mathbf{D}. \quad (4.12)$$

This highlights the inclusion of the HPM method into the general injection model

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \alpha_k \mathbf{D}, \quad (4.13)$$

where  $\alpha_k$ , for  $k = 1, \dots, N$ , are the modulation coefficients that weigh the contribution of  $\mathbf{P}$ .

Noticeably, the model (4.13) can be designed to satisfy a very interesting property: if the low pass image  $\mathbf{P}_L$  used in the calculation of  $\alpha_k$  does not depend on  $k$ , this algorithm belongs to the *Spectral Distortion Minimization* (SDM) class, or in other words the pansharpened image  $\widehat{\mathbf{MS}}$  exhibits the same spectral distortion of  $\widetilde{\mathbf{MS}}$  [54].

However, this property that turns out to be useful in showing the preservation of a crucial spectral feature of  $\widetilde{\mathbf{MS}}$ , does not necessarily imply an acceptable quality of the fusion product and has thus to be considered as a secondary requirement for the algorithm design. Indeed, increasing the quantity of the injected details typically implies the increase of spectral distortion [46], that on the contrary could be trivially eliminated by neglecting any contribution of  $\mathbf{P}$ .

## 4.2 Details Extraction Methods

In the following the experimental results performed on different data sets and with different types of detail extraction algorithms will be presented to show the very appealing features of pansharp-ening methods based on Eq. (4.9) with respect to approaches employing formula (4.10).

Two different methods for estimating the required lowpass version  $\mathbf{P}_L$  of the PAN image will be considered. One is based on the combination of the available MS images and the other on a proper degradation of the original PAN data. The former is exploited into the CS family and, in this case,  $\mathbf{P}_L$  is often indicated with the symbol  $\mathbf{I}_L$ . The building of the  $\mathbf{P}_L$  via the degradation step is instead a peculiarity of the MRA family. Further details can be found in Chapter 2.

The two detail injection algorithms can be combined with many different methods for achieving the low pass image  $\mathbf{P}_L$ .

The simplest algorithm into the CS family is the *IHS* (properly equalized to improve the results) [46] in which the PAN image is obtained by averaging all multispectral bands, or in other terms by adopting the assumption (hardly verified in the practice) that the MS bands equally contribute to the PAN image (see Sec. 2.2.2, for more details). Furthermore the IHS method employs Eq. (4.10) for the injection of details [44]. The corresponding contrast-based method using Eq. (4.9) is named *Brovey* [44].

Good results, for the other family, can be achieved by considering a system with complementary frequency responses with respect to the MTF of the sensor, whose shape can be safely approximated by a Gaussian function. This technique, referred as *MTF method*, was firstly proposed in [57] and successively efficiently implemented through the HPM details injection scheme (4.9) in [58]. A more sophisticated way for building the  $\mathbf{P}_L$  image consists in utilizing a MRA [52], which aim at separating the informative content at the various spatial scales. In particular we chose for the comparison of the two injection methods a Wavelet implementation based on the *à trous filters*, that allows to design

the analysis operator so that the equivalent low pass filter matches the sensor MTF [76]. In the same reference an HPM implementation, satisfying the SDM requirement has been utilized, while an A Trous Wavelet Trasform (ATWT) technique employing an *additive* model for detail injection was previously proposed in [60].

### 4.3 Experimental Results

Two of the most popular sensors for pansharpening applications are IKONOS and QuickBird. They work in the visible and near infrared spectrum range with the MS sensor that is characterized by four bands (Blue, Green, Red and Near InfraRed (NIR)) and they also have a PAN channel. The spatial resolution of IKONOS is  $4 \times 4$  m for the MS bands and  $1 \times 1$  m for the PAN. QuickBird has an asymmetric pixel. Indeed, the resolution cell for the multispectral bands is  $2.44 \times 2.88$  m, while for the PAN channel is  $0.61 \times 0.72$  m. Thus, the spatial resolution ratio is 4. The experimental results have been conducted on two real datasets<sup>1</sup>. In particular, we selected a scene of the *China dataset* (see Sec. 3.2 for details) in the case of IKONOS and a region of India (called *India dataset*, Fig. 4.3) acquired by the Quickbird sensor.

A quantitative analysis is performed by considering a reference image according to the Wald protocol [36] (see Sec. 2.3.1 for details), which results accurate in evaluating injection approaches. The low resolution MS bands (four times lower than that of the PAN image) are obtained by applying a low-pass filtering and decimation [42]. The frequency response of the low-pass filter is designed to match the MTF of each spectral channel of the sensor and the panchromatic channel is degraded by means of an ideal low-pass filtering [57, 58].

The assessment of the fused products (with respect to the reference image) is carried out by calculating different indexes [42]. The ERGAS [52] and the Q4 [68] are exploited to evaluate the global quality of images.

---

<sup>1</sup>Available at <http://glcf.umiaccs.umd.edu>.

	Injection	Q4	SAM(°)	ERGAS	SCC
<b>IHS</b>	ERR	0.8512	3.5836	2.8178	0.9069
	CON	<b>0.8528</b>	<b>3.4985</b>	<b>2.7788</b>	<b>0.9102</b>
<b>MTF</b>	ERR	0.8763	3.2425	2.6207	0.9022
	CON	<b>0.8792</b>	<b>3.1872</b>	<b>2.5801</b>	<b>0.9059</b>
<b>ATWT</b>	ERR	0.8767	3.1973	2.6058	0.9042
	CON	<b>0.8800</b>	<b>3.1272</b>	<b>2.5600</b>	<b>0.9083</b>

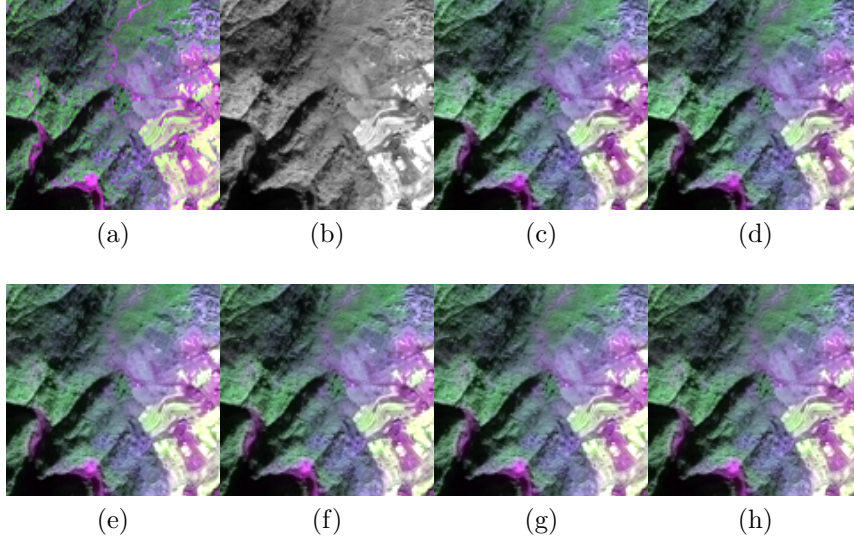
**Table 4.1** *China dataset*: Comparison between error-based (ERR) and contrast-based (CON) injection models.

	Injection	Q4	SAM(°)	ERGAS	SCC
<b>IHS</b>	ERR	0.7450	3.2976	2.1561	0.9074
	CON	<b>0.7484</b>	<b>3.1214</b>	<b>2.1064</b>	<b>0.9162</b>
<b>MTF</b>	ERR	0.8373	2.7817	1.7109	0.9097
	CON	<b>0.8451</b>	<b>2.7212</b>	<b>1.6581</b>	<b>0.9158</b>
<b>ATWT</b>	ERR	0.8338	2.8112	1.7262	0.9091
	CON	<b>0.8421</b>	<b>2.7507</b>	<b>1.6705</b>	<b>0.9154</b>

**Table 4.2** *India dataset*: Comparison between error-based (ERR) and contrast-based (CON) injection models.

The spectral and spatial distortions are assessed by the *Spectral Angle Mapper* (SAM) and the *Spatial Correlation Coefficient* (SCC) [65], respectively. We recall that the optimal values of the indexes are one for Q4 and SCC and zero for ERGAS and SAM.

The numerical results achieved by contrast-based (CON) and error-based (ERR) injection methods (Sec. 4.2) on the *China* and *India datasets* are reported in Tabs. 4.1 and 4.2. The superiority of the CON model is evident in all the considered scenarios and to a different extent for all kinds of algorithms. All the quality indexes are improved by this choice, but a finer analysis reveals that the use of (4.9) allows to reduce some typical drawbacks of CS and MRA details extraction methodologies. Indeed the CS methods are known to be affected by a more significant spectral distortion that is in part compensated by the CON method, as testified by the enhancement achieved by the SAM index. On the other side

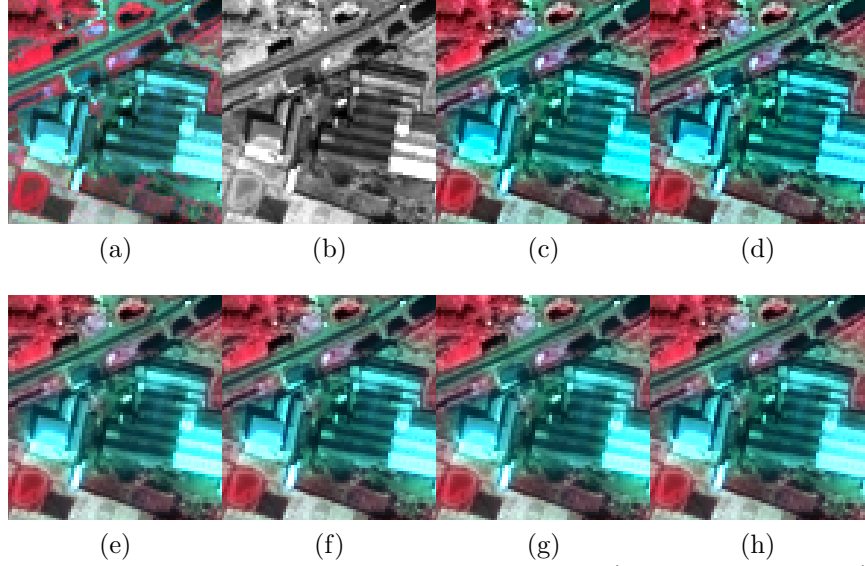


**Figure 4.2** *China dataset* false color representation (NIR, Red and Green): (a) Full resolution MS image (reference); (b) PAN image; (c) IHS-ERR (*IHS*); (d) IHS-CON (*Brovey*); (e) MTF-ERR (*GLP*); (f) MTF-CON (*MTF HPM*); (g) ATWT-ERR (*ATWT*); (h) ATWT-CON.

the MRA methods benefit from the use of rule (4.9) especially in terms of SCC index, indicating a substantial contribution in terms of spatial details.

Visual analysis can be performed by Figs. 4.2 and 4.3 in which two magnified parts of the considered dataset have been shown in false colors (NIR, Red and Green). In both cases the inspection corroborates the considerations derived during the analysis of quantitative parameters. In particular, the superior spectral quality of images achieved by the CON paradigm in the *IHS* approach (Figs. 4.2(d), 4.3(d)) with respect to those attained by the same CS method with ERR injection rule (Figs. 4.2(c), 4.3(c)) can be observed by comparing the fused products with the reference images reported in Figs. 4.2(a) and 4.3(a). The presence of additional details in images obtained by the CON rule can be noted in the corresponding images of Fig. 4.3.

Finally, it is worth noticing that among the algorithms used for the current analysis the ATWT and MTF methods compare



**Figure 4.3** *India dataset* false color representation (NIR, Red and Green):  
 (a) Full resolution MS image (reference); (b) PAN image; (c) IHS-ERR (*IHS*); (d) IHS-CON (*Brovey*); (e) MTF-ERR (*GLP*); (f) MTF-CON (*MTF HPM*); (g) ATWT-ERR (*ATWT*); (h) ATWT-CON.

favorably in all the test cases. Indeed, these approaches utilizing a filter designed to match the sensor MTF take advantage with respect to the other competitors from a more proper details extraction phase.



## Chapter 5

# Pansharpening Based on Semiblind Deconvolution

Classical approaches to pansharpening are based on two main steps: Extraction of details from the PANchromatic (PAN) image and their injection into the MultiSpectral (MS) data. The former is often performed exploiting *a priori* knowledge about the MS sensor instrumentation. Indeed, the Modulation Transfer Function (MTF) is typically used and defined starting from the Gaussian hypothesis and by characterizing its analytic expression by the means of the sensor gain at the Nyquist frequency. However this characterization could be often inadequate. For this reason, in this chapter we provide a formulation of a fast optimization problem trying to identify the relationship between PAN and MS images directly from the available data. The effectiveness of the approach is validated both on a reduced scale dataset achieved by degrading images acquired by the IKONOS sensor and on full scale data consisting of images collected by the QuickBird sensor. In the first case the proposed method achieves performances very similar to the algorithm relying upon the full knowledge of the degrading filter. In the second it is shown to compare favorably to several very credited approaches used in the current literature.

The rest of the chapter is organized as follows. Sec. 5.1 underlines the motivations and the contributions related to this work.

In Sec. 5.2, the formulation of the estimation problem and its solution are provided; In Sec. 5.3, we detail the employed algorithm for spatially enhancing the MS image, which is based on a well-known injection model, while, in Sec. 5.4, the two validation procedures are fully described and exploited to assess the performances of the proposed method with respect to state-of-art methods.

## 5.1 Motivations and Contributions

This chapter is focused on a particular issue that is crucial in many pansharpening approaches: The characterization of the relationship between the MS and the PAN image through the MTF of the MS sensor. We recall that the MTF is defined as the absolute value of the Fourier transform of the sensor's *Point Spread Function*<sup>1</sup>. In this regard, it is supposed that the less detailed MS image is achievable as the convolution of the PAN image by a filter matched to the MTF. This assumption relies upon the hypothesis of ideality of the frequency response of the PAN sensor and of the interpolation applied to the MS for upsampling. The well established pansharpening technique in [57] leverages on this model. In more details, this technique employs a classical scheme consisting of two parts: *Detail extraction*, in which the missing spatial content is extracted from the PAN image and *Detail injection* in which such information is introduced into the MS image. Specifically, in [57] the suitability of using a High Pass Filter matched to the MS sensor's MTF for the first step is justified by the purpose of restoring all the spatial details not resolved by the MS. The relationship under study is also required by the pansharpening algorithms that demand a version of the PAN image at the resolution of the original MS image, as, for example, the *Band-Dependent Spatial-Detail* [51], or the *Gram-Schmidt mode 2* [11].

---

<sup>1</sup>The Point Spread Function is the inverse Fourier Transform of the Optical Transfer Function, which is the product between the Phase Transfer Function (PTF) and the MTF.

As a further motivating example, we cite the approach recently proposed in [23] and based on Compressive Sensing theory [19], which exploits the PAN-MS relationship for ensuring the consistency of the high resolution dictionary with the low resolution available data.

The design of a filter matching the MTF is not straightforward. Gaussian-shaped filters are usually employed for approximating the shape of the real (and unknown) MTF of the MS [58, 70]. The effectiveness of this model relies upon the fact that typically the MTF closely follows the shape of a Gaussian bell in many optical acquisition systems [70]. This family is defined by a single parameter (i.e., the standard deviation) that fully specifies the filter response. Usually the value of the standard deviation of the filter is determined from the gain at the Nyquist frequency, which is often provided by the sensor's manufacturer (derived by construction specifications or, more properly, by on-board measurements) [57, 58]. "A trous" wavelet transforms have been also used for modeling the MTF [60]. Indeed, it was noted that when these filters are designed for matching the sensor specifications they resemble a Gaussian shape filter [57].

Unfortunately, in many cases the information available does not lead to a specification of the sensor MTF with the desirable accuracy. A first issue concerns the modeling of the MTF by a Gaussian-shaped filter, which is sometimes unconfirmed in practice. Even when it constitutes a reasonable approximation, its specification through the MTF gain at Nyquist frequency is not always possible, being it unavailable for many sensors. Even in the most favorable case, the gain at the Nyquist frequency provided by the sensor manufacturer is typically the one measured after the launch of the sensor. However, due to the aging of the optical and electronics payload of the sensor, it may have changed, thus losing the precision of this datum. A final concern regards the symmetry hypothesis of the MTF; in fact for several sensors the MTF shape varies between the along-track and the across-track directions and thus it cannot be fully characterized through the single value of the standard deviation.

We propose to *estimate* the relationship between the PAN and MS images directly from the data in order to overcome the above-mentioned limitations. To the best of our knowledge, methods that propose to estimate the filter for detail extraction from PAN by employing the same images are absent in the pansharpening literature. The comparison between the proposed approach and those based on the determination of the MTF from the sensor knowledge is thus a very intriguing topic to investigate.

In order to perform the filter estimation, we assume that each MS band is a spatially degraded version of the PAN image, all blurred by the same filter, and we model the estimation as a problem of *blind image deblurring*. We recall that blind image deblurring, a classical challenge in image processing, is the inverse problem of recovering sharp images from degraded ones by both estimating the blur affecting the image and the target high resolution (deblurred) image. In general, this problem is hard to solve since it is ill-posed (i.e., there are multiple pairs of estimated image and blurring operator that can produce the degraded image) and the blurring operators are usually very ill-conditioned in practice. For these reasons, in order to obtain a well-defined solution, additional hypotheses (or *a priori* information) about the original image and/or the blurring operator are needed. The problem can thus be formulated into the Bayesian framework or as a regularization problem, both leading to a similar optimization problem form (please refer to [77, 78] and to the more recent paper [79] for a review of the topic). Actually, in this chapter only the estimation of the blurring operator is required (i.e., the filter approximating the MTF of the MS sensor), thus rendering the solution significantly simpler.

Starting from the motivations presented earlier, the objectives of this chapter can be pointed out as: 1) Provide an estimation procedure for the degradation filter based only on the acquired images, with the objective of obtaining an optimal system for details extraction from the PAN image; 2) Compare this approach with the *state-of-art* pansharpening methods and in particular with those exploiting the knowledge of the sensor MTF. Correspondingly, two

different validation procedures are utilized. The first one aims at evidencing the capabilities of the estimation procedure to properly approximate the actual response of the imaging systems; the second one points out the advantages of extracting details through the filters derived according to the proposed procedure, instead of those employed in classical approaches.

### 5.1.1 Notation

We use bold lowercase to denote vectors (e.g.,  $\mathbf{x}$ ,  $\mathbf{y}$ ) and bold uppercase to denote matrices (e.g.,  $\mathbf{H}$ ,  $\mathbf{MS}$ ).  $\mathbf{X}^T$  indicates the Hermitian transpose of  $\mathbf{X}$ . A monochromatic image is represented by lexicographic ordering its pixels, namely by stacking either their rows or their columns into a vector. In particular, vector  $\mathbf{p} \in \mathbb{R}^r$  indicates the observed panchromatic image, where  $r$  represents the number of pixels. Accordingly, MS images are organized as a matrix in which each row corresponds to a spectral band. The observed MS image is thus denoted as  $\mathbf{MS} \in \mathbb{R}^{N \times q}$ , in which  $N$  is the number of bands and  $q < r$  is the number of pixels in each MS channel. Note that, with this notation, each column contains the spectrum of a given pixel.

## 5.2 Semiblind Deconvolution

In this section, we detail the mathematical formulation of the general problem of filter estimation. The objective is to infer the relationship between the blurred and the sharp images, which, in this context, are represented by the MS and the PAN data, respectively.

We restrict the analysis of this work to the blurring process through a linear shift-invariant system in the presence of additive noise that is described, for a given monochromatic (sharp) image  $x$ , by the model

$$y = x * h + n \quad (5.1)$$

in which  $y$ ,  $h$  and  $n$  are the observed degraded image, the blur-

ring filter, and the observation noise, respectively and  $*$  denotes the convolution operator. This equation can be also expressed in matrix-vector notation by

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{n}, \quad (5.2)$$

where  $\mathbf{h}$ ,  $\mathbf{y}$ , and  $\mathbf{n}$  are the lexicographic ordering of  $h$ ,  $y$ , and  $n$ , all belonging to  $\mathbb{R}^q$  and  $\mathbf{X} \in \mathbb{R}^{q \times q}$  is a matrix operator which is constructed by properly arranging the elements of  $\mathbf{x}$  [80]. In the case of the two-dimensional convolution with periodic boundary conditions,  $\mathbf{X}$  is a Block Circulant with Circulant Blocks (BCCB) matrix. The latter can be spectrally decomposed as  $\mathbf{X} = \mathbf{F}^T \mathbf{\Lambda} \mathbf{F}$ , where  $\mathbf{F}$  and  $\mathbf{F}^T = \mathbf{F}^{-1}$  are the DFT (Discrete Fourier Transform) and the inverse DFT matrices, respectively, and  $\mathbf{\Lambda}$  is a diagonal matrix. This decomposition has two main advantages. First, by using Fast Fourier Transforms (FFTs), computing matrix-vector multiplications with  $\mathbf{F}^T$  and  $\mathbf{F}$  can be done without explicitly constructing them; this implies significant computational savings, since the cost of the FFT algorithm is  $O(M \log_2 M)$  for power-of-two length vectors, where  $M$  is the length of the vector. Second, the diagonal matrix  $\mathbf{\Lambda}$  is easily invertible, which is very convenient, as it will be seen later. Furthermore, we restrict the estimation of the blur vector  $\mathbf{h}$  to a given nonempty convex set  $\mathcal{H}$  of  $\mathbb{R}^q$ , that is used to impose a finite support to  $\mathbf{h}$  (namely to limit the number of nonzero values of  $\mathbf{h}$ ).

The simplest and most intuitive approach to apply the deconvolution problem to pansharpening consists in searching for a relationship between each MS band and the PAN image. In other words, the estimation problem is separated into  $N$  independent parts, where  $N$  represents the number of the MS spectral bands. The intrinsic suboptimal nature of this method, which will be referred to as *FE MS* (*Filter Estimation with Multi Spectral optimization*), entails the lack of correlation among the MS bands, leading to worse performances (as shown in Sec. 5.4). In particular this incoherence can cause the wrong estimation of some filters, as it is the case, for example, of the blur filter relative to the blue channel of the IKONOS sensor (see Fig. 5.6). For this

reason, other more sophisticated approaches have to be devised to solve the problem. In the following, we discuss in more detail our proposed method in light of the previous considerations.

A different method, aimed at preserving the coherence among details of the MS and PAN images, is the object of our current proposal and hence will be simply denoted as *FE* (*Filter Estimation*). Our approach consists in applying Eq. (5.1) to model the relation between the original PAN image and its low resolution version, which is constructed as a linear combination of the MS bands, as it is common for Relative Spectral Contribution (RSC) methods [1]. The equivalent panchromatic image  $\mathbf{p}_e$  can be written as:

$$\mathbf{p}_e = \widetilde{\mathbf{MS}}^T \boldsymbol{\alpha}, \quad (5.3)$$

where  $\widetilde{\mathbf{MS}}$  is a version of the MS image upsampled to the PAN resolution and the vector  $\boldsymbol{\alpha} \equiv [\alpha_1, \dots, \alpha_L]^T$  contains the weighting coefficients, accounting for the spectral overlap of MS channels and PAN image [1].

More specifically, the filter estimation problem is formulated as:

$$\begin{aligned} & \underset{\mathbf{h}, \boldsymbol{\alpha}}{\text{minimize}} \quad \|\mathbf{p}_e - \mathbf{P}_C \mathbf{h}\|^2 + \lambda \|\mathbf{h}\|^2 + \mu (\|\mathbf{D}_v \mathbf{h}\|^2 + \|\mathbf{D}_h \mathbf{h}\|^2) \\ & \text{subject to} \quad \mathbf{h}^T \mathbf{1} = 1, \mathbf{h} \in \mathcal{H}. \end{aligned} \quad (5.4)$$

The first term is known as the *data-fitting term*, which imposes that the blurred version of the panchromatic image must be close to the equivalent panchromatic. Note that  $\mathbf{P}_C$  is defined such that  $\mathbf{P}_C \mathbf{h}$  represents the linear convolution in matrix form between the panchromatic image  $\mathbf{p}$  and the blur  $\mathbf{h}$ , as in (5.2). The second and third addends act as *regularization terms* (in the sense of Tikhonov) aimed at dealing with the ill-posedness of the inverse problem [81]. These terms can be also seen as an *a priori* information under a Bayesian framework [78]. Here  $\mathbf{D}_h, \mathbf{D}_v \in \mathbb{R}^{q \times q}$ , which are BCCB matrices, stand for the first-order finite differences operator in the horizontal and vertical directions, respectively, and can be also interpreted as the convolution between  $\mathbf{h}$  and the derivative filters  $\mathbf{d}_h \in \mathbb{R}^q$  and  $\mathbf{d}_v \in \mathbb{R}^q$ . The constraints

induce the normalization of the blur (i.e.,  $\mathbf{h}^T \mathbf{1} = 1$ ) and its finite support (i.e.,  $\mathbf{h} \in \mathcal{H}$ ), respectively. The selection of the squared  $\ell_2$  norm  $\|\cdot\|^2$  allows for a closed form solution, which can be computed efficiently in the frequency domain. The same choice proves favorable also in the regularization terms, matching some desirable physical conditions. Specifically, its use in the second addend forces the obtained solution to have limited energy. This makes sense since the blur degrading effect is usually confined to a small region. Analogously, when acting on the filter finite differences, it assures smooth transitions among the filter values. This is desirable, since blurs experienced in the practice tend to be smooth and have Gaussian-like shapes. The two regularization terms are properly weighted by coefficients  $\lambda$  and  $\mu$ , which can be used as input parameters to the algorithm.

The solution of the optimization problem follows an alternated minimization scheme, where each variable is minimized separately. First the coefficients  $\boldsymbol{\alpha}$  are estimated, and, subsequently, the blur is optimized; this sequence is repeated for every iteration. The initialization of the algorithm is performed by starting from the low pass filter commonly used for the “à trous” implementation of the wavelet decomposition, matched with the typical MTF shape in the Visible Near InfraRed (VNIR) spectrum [57,60]. Note that we do not impose the optimization constraints directly. In fact, the estimated filter is normalized after each iteration and values outside a given window (typically being very close to zero) are then discarded. This allows us to find an approximate solution to the optimization problem in a simple way. For each iteration, the first step consists in computing  $\boldsymbol{\alpha}$ , given the current value of  $\mathbf{h}$ . This corresponds to find the solution of a simple least squares problem:

$$\widetilde{\mathbf{M}}\widetilde{\mathbf{M}}^T \boldsymbol{\alpha}^{\text{sol}} = \widetilde{\mathbf{M}}\mathbf{P}_C \mathbf{h}. \quad (5.5)$$

In the second step, the computation of  $\mathbf{h}$  given  $\boldsymbol{\alpha}$  involves the solution of the optimization problem described by Eq. (5.4) with respect to  $\mathbf{h}$ . Being the cost function quadratic, it has a global



minimum achieved when

$$[\mathbf{P}_C^T \mathbf{P}_C + \lambda \mathbf{I} + \mu \mathbf{D}_v^T \mathbf{D}_v + \mu \mathbf{D}_h^T \mathbf{D}_h] \mathbf{h}^{\text{sol}} = \mathbf{P}_C^H \mathbf{p}_e. \quad (5.6)$$

As seen before, the BCCB matrices  $\mathbf{P}_C$ ,  $\mathbf{D}_h$  and  $\mathbf{D}_v$  are diagonalized by the 2-D DFT matrix,  $\mathcal{F}\{\cdot\}$ . This accelerates the computation of the solution:

$$\mathbf{h}^{\text{sol}} = \mathcal{F}^{-1} \left\{ \circ \frac{\mathcal{F}\{\mathbf{p}\}^* \circ \mathcal{F}\{\mathbf{p}_e\}}{\mathcal{F}\{\mathbf{p}\}^* \circ \mathcal{F}\{\mathbf{p}\} + \lambda + \mu(\mathcal{F}\{\mathbf{d}_h\}^* \circ \mathcal{F}\{\mathbf{d}_h\} + \mathcal{F}\{\mathbf{d}_v\}^* \circ \mathcal{F}\{\mathbf{d}_v\})} \circ \right\}, \quad (5.7)$$

where  $\mathcal{F}^{-1}\{\cdot\}$  denotes the inverse Fourier transform,  $\mathcal{F}\{\cdot\}^*$  is the complex conjugate of the Fourier transform,  $\circ$  is the component-wise multiplication, and  $\circ - \circ$  is the component-wise division. This equation involves a diagonal inversion, which has cost  $O(m)$ , but is dominated by the FFTs — with cost  $O(m \log m)$ ; note that  $\mathcal{F}\{\mathbf{p}\}$ ,  $\mathcal{F}\{\mathbf{d}_h\}$ , and  $\mathcal{F}\{\mathbf{d}_v\}$  can be computed in advance. In order to take advantage of the properties of the FFT, it is necessary to work under periodic boundary conditions, as discussed earlier. However, when dealing with real-world images, this assumption is usually too strong, since it is highly improbable that an image's external borders (unobserved) are repeated periodically. Processing these images as they are usually leads to development of undesirable artifacts and, in order to reduce them, a pre-processing step is usually taken, namely by blurring the borders of the images. This allows the discontinuities to be smoothed out [82].

The followed procedure can be summarized as in Algorithm 1.

The required stopping criterion can be imposed by considering the relative changes between the estimated variables or even by imposing a fixed number of iterations.

### 5.3 Pansharpening Method Based on Filter Estimation

One of the main issues for pansharpening algorithms is how to properly inject into the MS the spatial details extracted from the PAN in order to obtain the required spatial enhancement.

---

**Algorithm 1:** Filter Estimation Algorithm

---

**Data:** The observed PAN and MS data,  $\mathbf{p}$ ,  $\mathbf{MS}$ , respectively; the regularizer weights  $\lambda$  and  $\mu$ ; an initial filter estimate  $\mathbf{h}^0$ , dimension of the filter's support, threshold value.

**Result:** Degradation filter  $\mathbf{h}^j$ .

**begin**

    Upsample the MS image  $\mathbf{MS}$

**while** *some stopping criterion is satisfied* **do**

        Set  $\boldsymbol{\alpha}^j$  as indicated in (5.5)

        Set  $\mathbf{h}^j$  as indicated in (5.7)

$\mathbf{h}^j$ : Normalize, threshold and use a mask to ignore values outside the filter's support

---

Many approaches have been proposed in the literature and a widespread classification distinguishes between local and global approaches [50]. An example of the former is the Context-Based Decision (CBD) algorithm [70] which relies upon the partitioning of the fusing images in blocks. The details extracted by the PAN image are injected, if needed, after an equalization phase based on the relationship between the standard deviations calculated within the corresponding blocks of the two images. Furthermore, details injection is performed only if the correlation between the lowpass version of the PAN image and each of the expanded MS bands is greater than a given threshold [70]. The injection schemes belonging to the second class are based on global rules, as it is the case of the simple additive injection model [60]. In this method, a histogram matching procedure between the PAN image and each MS band is advisable before fusing data from different sensors. We want to stress that the objective of this work is the study of the extraction detail phase and the comparison of several approaches under the same conditions. To reach this scope, we selected a well know injection model based on the concept of modulation (often called *High Pass Modulation* (HPM)). In Chap-

ter 4, it has been shown that the HPM scheme is related to the local contrast of an image and in general outperforms the *additive* injection approaches. Moreover, some powerful fusing procedures are also based on the HPM, as in [58].

More in detail, let us define  $\mathbf{P}^{hm} \in \mathbb{R}^{N \times r}$  as the PAN image ( $\mathbf{p}$ ) after histogram matching with the  $N$  MS spectral bands ( $\mathbf{MS}$ ). For the  $k$ -th band, with  $k \in [1, \dots, N]$ ,  $\mathbf{P}_k^{hm}$  is given by:

$$\mathbf{P}_k^{hm} = [\mathbf{p} - \mu_{\mathbf{p}}] \cdot \frac{\sigma_{\mathbf{MS}_k}}{\sigma_{\mathbf{p}}} + \mu_{\mathbf{MS}_k}, \quad (5.8)$$

with  $\mu$  and  $\sigma$  denoting the mean and standard deviation, respectively, estimated on the PAN and  $k$ -th MS band. The HPM injection model defines the fused product  $\widehat{\mathbf{MS}} \in \mathbb{R}^{N \times r}$  (in the lexicographic ordering) according to the formula:

$$\widehat{\mathbf{MS}} = \widetilde{\mathbf{MS}} \circ \left( \circ \frac{\mathbf{P}^{hm}}{\mathbf{P}_{LP}^{hm}} \circ \right), \quad (5.9)$$

with  $\circ$  the component-wise multiplication, and  $\circ - \circ$  the component-wise division. The  $\mathbf{P}_{LP}^{hm} \in \mathbb{R}^{N \times r}$  is the low resolution version of  $\mathbf{P}^{hm}$  at the same scale of the MS image, obtained by filtering  $\mathbf{P}^{hm}$  with the spatial Low Pass Filter (LPF)  $\mathbf{h}$ .

Accordingly, the fusion procedure exploiting the HPM model can be summarized as in Algorithm 2.

In the following section some different types of blur filters  $\mathbf{h}$  will be used to generate  $\mathbf{P}_{LP}^{hm}$  from  $\mathbf{P}^{hm}$  and they will be compared along with the proposed method presented in Sec. 5.2.

## 5.4 Experimental Results

In order to validate the proposed method and point out its advantages with respect to filters conventionally used in state of the art techniques for detail extraction, an accurate experimental analysis has been performed by considering two datasets. The first one is the *China dataset* (see Sec. 3.2) acquired by IKONOS, while, the

**Algorithm 2:** Pansharpening Algorithm

---

**Data:** The observed PAN and MS data,  $\mathbf{p}$ ,  $\mathbf{MS}$ , respectively; the degradation filter  $\mathbf{h}$ .

**Result:** Pansharpened image  $\widehat{\mathbf{MS}}$ .

**begin**

Upsample the MS image to get  $\widetilde{\mathbf{MS}}$  with the same dimensions of  $\mathbf{p}$

Equalize  $\mathbf{p}$  with each band of  $\mathbf{MS}$  using Eq. (5.8) to get  $\mathbf{P}^{hm}$

Convolve  $\mathbf{P}^{hm}$  with the degradation filter  $\mathbf{h}$  to get  $\mathbf{P}_{LP}^{hm}$

Use Eq. (5.9) to obtain the final product  $\widehat{\mathbf{MS}}$

---

second dataset is acquired by QuickBird (see Sec. 4.3 for its spectral and spatial characteristics) and it represents an urban area of Indianapolis, and is named, for this reason, *Indianapolis dataset*.

All the algorithms considered for the comparison are based on the HPM injection model presented in Sec. 5.3. Thus the pansharpening algorithms differ only by the detail extraction phase, which is defined by the filter used to generate the equivalent panchromatic with same resolution of the MS image (i.e.,  $\mathbf{P}_{LP}$  in Eq.(5.9)). More in detail, *SFIM* [55] is based on the *Box* filter for extracting details; *ATWT* is an “à trous” wavelet method [60] matched with a typical MTF shape in the visible and NIR spectrum [57], while *MTF* employs a *Gaussian* filter [58, 70], designed to match the sensor MTF by exploiting the knowledge of the gain at the Nyquist frequency [57]. For the sake of comparison, *EXP* is also taken into account in the evaluation being the reference image with no detail injection. In this case, only the interpolation of the original MS image is performed to reach the same image size of the PAN. The interpolation algorithm exploits the polynomial kernel with 23 coefficients as in [70]. This interpolation method is also used to generate the  $\widetilde{\mathbf{MS}}$  in Eq.(5.9) for all the algorithms.

Two validation strategies are considered:

- The first one is performed by following the Wald protocol.

A reduced scale MS image is generated by applying a Gaussian MTF matched filter to the original MS image that is also used as a reference for validating the fused products. The main objective of this test is to show that the proposed method is able to properly approximate the LPF used for degrading the images in the simulation phase.

- The other experiment is performed at full scale with no degradation. In this case no reference image is available and thus the spectral quality of the fused image has to be evaluated against the original MS image, while the spatial quality is assessed by comparing the details of the final product with those of the original PAN image. Moreover, in this case, a visual analysis is a mandatory step to appreciate the quality of the pansharpened images.

#### 5.4.1 Wald Protocol

In the literature, there are two well-known techniques to validate pansharpening products. One of the most used (see the pansharpening contest organized by the IEEE Data Fusion Committee in 2006 [42]) is derived from the Wald Protocol and aims at verifying the *synthesis* property [36]. In this case, reduced scale MS and PAN images are simulated; the two images are then fused and the pansharpened image is compared with the original image, used as a reference. A problem implied by this validation procedure regards the quantification of the similarity between the two MS images, or, in another words, the choice of the quality indexes to use in order to verify the spectral and spatial consistency of the fused product with the given reference image. We will exploit in this chapter: The *Spectral Angle Mapper* (SAM) (in degree) that evaluates the spectral distortion, the *Erreur Relative Globale Adimensionnelle de Synthèse* (ERGAS) [52] that is able to measure both the spectral and radiometric distortions and the Q4 index [68], which represents a vectorial extension of the  $Q$  [67], relevant to four band datasets.

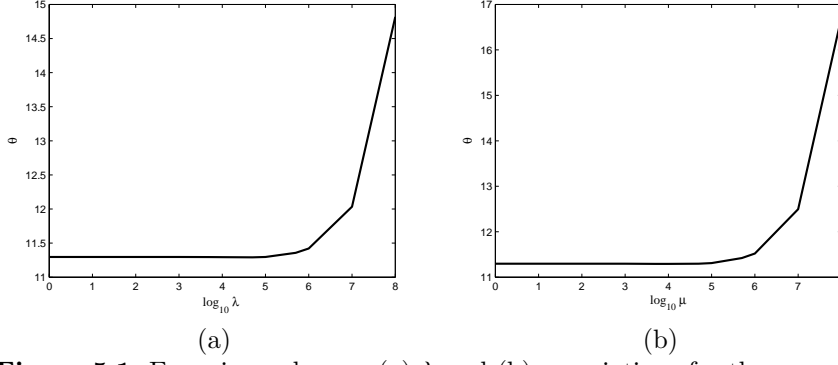
Sensor	Blue	Green	Red	NIR
IKONOS	0.27	0.28	0.29	0.28
QuickBird	0.34	0.32	0.30	0.22

**Table 5.1** MTF gains at Nyquist cutoff frequency.

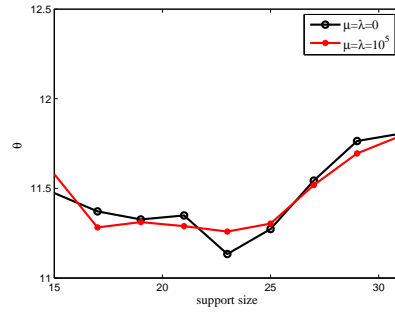
The last question that arises in this protocol is related to the procedure for simulating the low resolution MS and PAN images. Proper low pass spatial filters have to be applied to the two images. Usually, in the case of the MS image, the filter is matched with the MTF shape of the MS sensor, designed by exploiting the hypothesis of Gaussian shape and the knowledge of the gains at Nyquist frequency of the MS sensor (see Tab. 5.1 for values of sensors involved in this analysis) [57]. On the other hand, an ideal filter is applied in this phase to the PAN image [57].

Obviously, this procedure is very accurate, but the hypothesis of invariance between scales, on which it is founded, is not always fulfilled in practice. Moreover, a strong bias is introduced in the comparison among the algorithms, since a specific known filter is applied to degrade the initial images. In fact, in the detail extraction stage of a pansharpening algorithm, the best way to extract details should be obtained matching the filter used to simulate the products for the validation.

In this validation scenario preliminary experiments aim at evaluating the role of the parameters characterizing the proposed method for the estimation of the filter. The first analysis is carried out by varying the  $\lambda$  and  $\mu$  coefficients, fixing the support dimensions to a value equal to the double of the supposed real support size (i.e., 12 pixels, which as detailed afterwards can be a reasonable size for the filter support). A large range of values are used for the coefficients of the regularization terms in a model of the system (i.e., Eq. (5.1) without any noise). The performances are measured by using the error in angle (in degrees) between the filter used in the simulation phase and the one estimated through the proposed optimization procedure illustrated by Algorithm 2. The results are shown in Fig. 5.1. From the experiments it can be



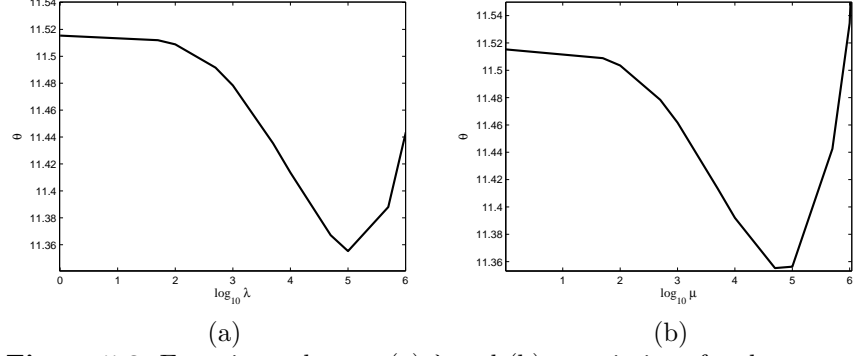
**Figure 5.1** Error in angle over (a)  $\lambda$  and (b)  $\mu$  variations for the proposed approach with support size equal to 25, using the model in Eq.(5.1) without additional noise.



**Figure 5.2** Error in angle Vs. the support size using the model in Eq.(5.1) without additional noise.

stated that the proposed method shows no particular sensitivity with respect to the values of the regularization coefficients.

Subsequently, an analysis for evaluating the robustness of the technique to different filter support sizes is performed. A reasonable choice for the support dimension is considered to be at least  $3 \cdot R$  assuming a blur with Gaussian shape (i.e., for a  $R = 4$  the support size is at least equal to  $3 \cdot R = 12$ ). The performances corresponding to different sizes of the filter support are computed fixing  $\lambda = \mu = 0$  and  $\lambda = \mu = 10^5$  and are reported in Fig. 5.2 showing that this rule of thumb for setting the size of the support is reasonable. By investigating more in details the obtained re-



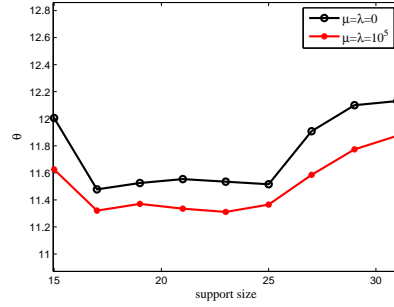
**Figure 5.3** Error in angle over (a)  $\lambda$  and (b)  $\mu$  variations for the proposed approach with support size equal to 25, using the model in Eq.(5.1) with Gaussian noise. The SNR is equal to 45 db.

sults, errors in angle are comparable, but the robustness over high support dimensions values is better for higher values of the parameters. For this reason and for the greater robustness to noise of this configuration, we choose the latter setting. Quantifying the robustness to noise can be useful, especially in real scenarios where the Signal to Noise Ratio (SNR) could be lower than in a simulated case. In order to further corroborating this choice, we show in Figs. 5.3 and 5.4 the angle error curves vs.  $\lambda$  and  $\mu$  coefficients and support dimensions obtained by using i.i.d. white Gaussian noise in the degradation model described by Eq. (5.1). We set the SNR value to 45 db that is well-suited for simulating real scene acquisitions by MS sensors. The effectiveness of employing higher values for the regularization coefficients to face the additional noise contribution is evident from the results.

Regarding the stopping criterion, a fixed number of iterations is chosen. The estimation in Eq. (5.5) is not very sensitive to the variation of the degradation filter applied to the PAN image, leading to a fast convergence of the proposed iterative approach for the filter estimation. A couple of iterations are usually enough to guarantee the convergence. The maximum number of iterations is set to 10 in order to ensure the stop of the algorithm.

Another interesting analysis is carried out by varying the LPF

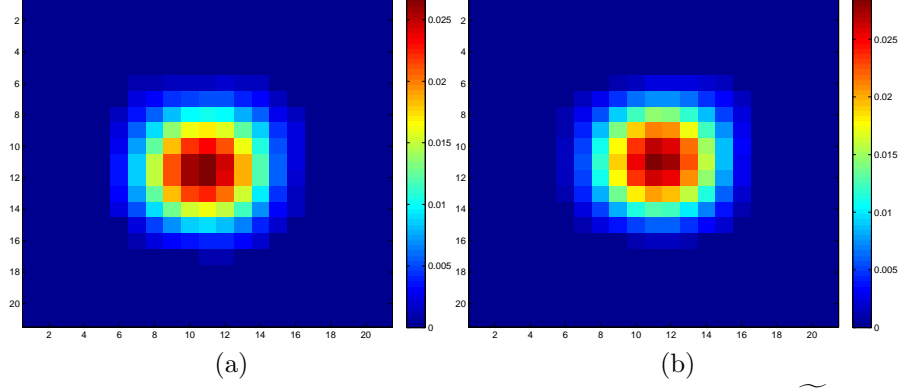




**Figure 5.4** Error in angle Vs. the support size using the model in Eq.(5.1) with Gaussian noise. The SNR is equal to 45 db.

needed by the interpolation procedure that yields the upsampled version **MS** of the MS image. In fact, the estimation problem in Eq. (5.5) takes into account both the actual degradation and the one introduced by the non-ideality of the LPF used in the upsampling step. In Fig. 5.5, we can see the different blurs estimated by the algorithm with different interpolation methods. It is worth noticing that, when the bicubic interpolator (i.e., the farthest one from ideal) is used, the estimated blur is more significant than the one achieved when a half band polynomial with 23 coefficients (i.e., closer to the ideal filter) is exploited. This is due to the mentioned necessary compensation of the non-ideality (particularly the attenuation introduced by the non-ideal LPF) in the estimation phase. The final results reported in Tab. 5.2 show that the proposed approach overcomes the method based on the same filter used for simulating the degraded images, which should obtain the best performances. On the contrary, by using an almost ideal LPF in the upsampling phase, the filter estimation procedure tends to approximate the system applied during the simulation, as confirmed by the very similar values achieved by MTF and FE methods (see Tab. 5.3).

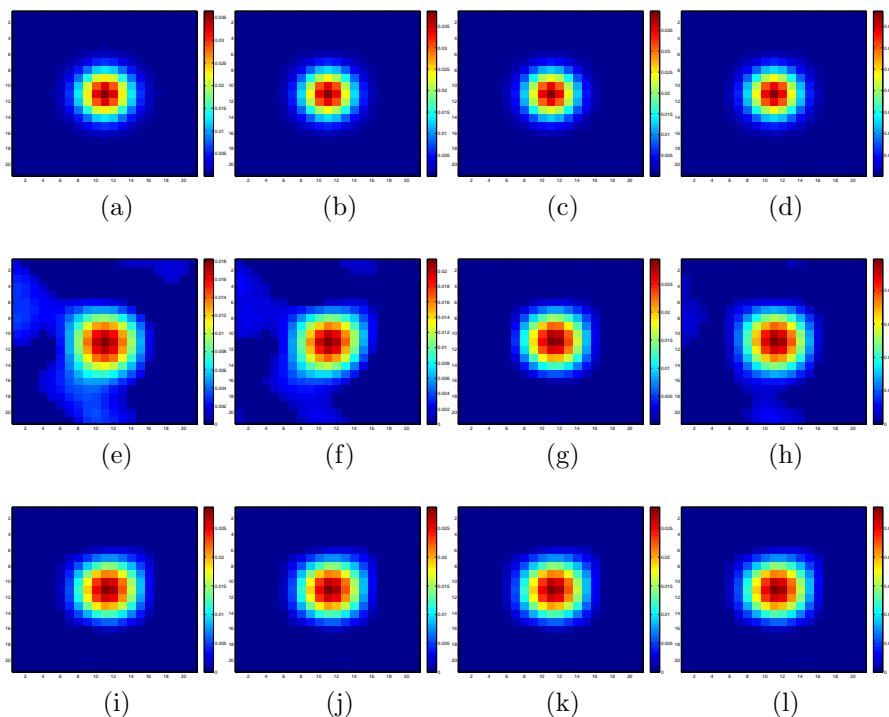
In Fig. 5.6 a visual comparison between the actual MTF-based blurs and the estimated ones is shown. The first row reports the ideal shapes of the degradation filters and the second contains the results obtained by applying the estimation procedure performed



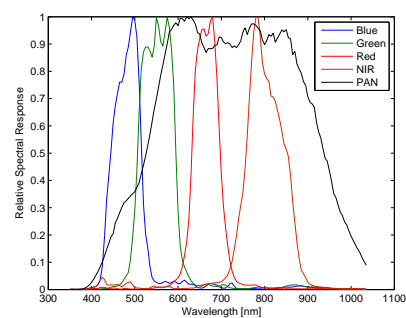
**Figure 5.5** Proposed method's estimated blur function, when  $\widetilde{\mathbf{M}}$  is obtained with (a) the bicubic interpolator or (b) the half band polynomial with 23 coefficient interpolator.

band by band (*FE MS* approach). Finally, the third row reports the filter achieved by examining the relation between the PAN image and its equivalent low pass version, which constitutes the proposed method of this paper (*FE* method). In the first case a different filter is obtained for each band. This is the simplest and more intuitive choice, but it obtains poor results due to the incoherence between a MS band and the PAN image (in particular, the Blue and Green bands) [1]. In fact, the coherence between a MS channel and the PAN image increases with the contribution that the band gives to the PAN image (see the relative spectral responses for the sensor in Fig. 5.7); accordingly a better estimation of the degradation filter can be achieved. In Tabs. 5.2 and 5.3, the results obtained by the *FE MS*, the *FE* and the *MTF* methods are shown. The advantages of the *FE* method with respect to the *FE MS* approach are evident in terms of all the performance indexes. Furthermore it can be underlined that the proposed method constitutes a good approximation of the optimal *MTF* approach.

In conclusion, the comparison with other very popular detail extraction filters is reported in Tab. 5.4. Naturally, the best results are obtained by the *MTF* approach (see a visual comparison between the *MTF* method and the proposal in Fig. 5.8). Indeed, the



**Figure 5.6** From the top to the bottom: (a)-(d) Expected blurs; (e)-(h) Estimated blurs by the *FE MS* method; (i)-(l) Estimated blurs by the *FE* approach. The columns are ordered by wavelengths (i.e., Blue, Green, Red and NIR).



**Figure 5.7** IKONOS sensor: Relative spectral responses.

Algorithm	$Q4$	$SAM(^{\circ})$	ERGAS
$MTF$	0.8744	3.1595	2.5908
$FE\ MS$	0.8674	3.1990	2.6888
$FE$	<b>0.8758</b>	<b>3.0990</b>	<b>2.5745</b>

**Table 5.2** Performance evaluation for the  $MTF$  and  $FE$  methods, when  $\widetilde{\mathbf{M}}$  is obtained with the bicubic interpolator.

Algorithm	$Q4$	$SAM(^{\circ})$	ERGAS
$MTF$	<b>0.8826</b>	2.9925	<b>2.4573</b>
$FE\ MS$	0.8774	3.0679	2.5828
$FE$	0.8825	<b>2.9807</b>	2.4623

**Table 5.3** Performance evaluation for the  $MTF$ ,  $FE\ MS$  and  $FE$  methods, when  $\widetilde{\mathbf{M}}$  is obtained with the 23 coefficient interpolator.

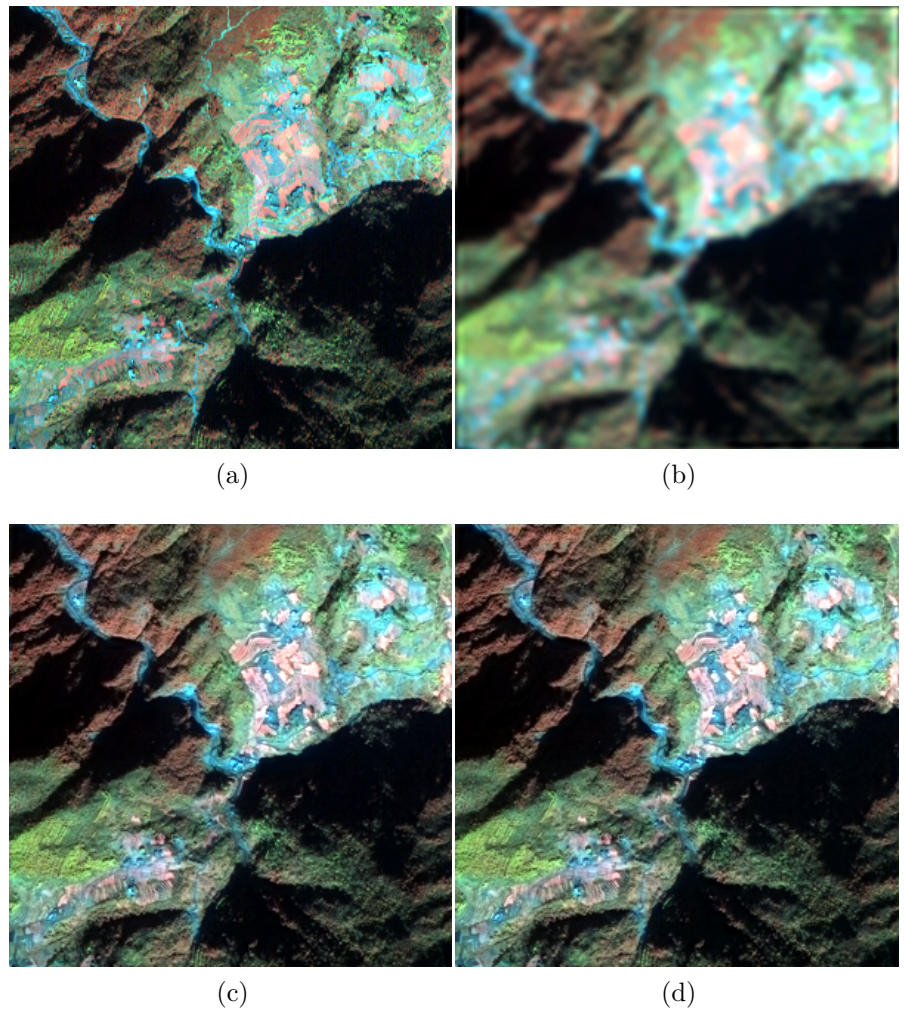
performances of the Gaussian approaches are slightly better than wavelet-based ones and, in general, they are preferred in the detail extraction phase [72]. Poorer performances are instead shown by the  $SFIM$  method, which is based on the  $Box$  filter. They are mainly due to the ripple in the obscure band and the shape in the frequency domain that does not well approximate the MTF of the MS sensor. Moreover, in Tab. 5.5, the execution times are shown with reference to the *China dataset*, using an image with dimensions equal to  $300 \times 300$  and a processor AMD Athlon 1.6 GHz. The interpolation phase with the half band polynomial with 23 coefficient interpolator requires 0.3 seconds. The fastest approach is  $SFIM$  followed by  $MTF$ ,  $FE$ , which takes more time due to the filter estimation step, and  $ATWT$ . However, all the approaches obtain good results in terms of this index, also thanks to the chosen simple but efficient injection rule.

Algorithm	$Q4$	SAM( $^{\circ}$ )	ERGAS
<i>EXP</i>	0.7423	4.4100	3.8337
<i>SFIM</i>	0.8597	3.4037	2.7823
<i>ATWT</i>	<b>0.8829</b>	3.0308	2.4659
<i>MTF</i>	0.8826	<b>2.9925</b>	<b>2.4573</b>

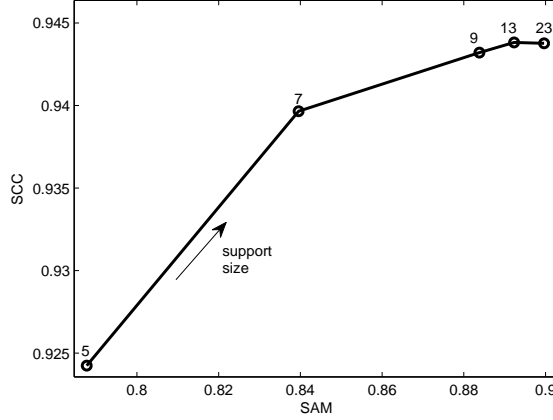
**Table 5.4** Performance evaluation for the *EXP*, *SFIM*, *ATWT* and *MTF* methods, when  $\widetilde{\mathbf{M}}$  is obtained with the 23 coefficient interpolator.

Algorithms	<i>EXP</i>	<i>SFIM</i>	<i>ATWT</i>	<i>MTF</i>	<i>FE</i>
Execution Times [sec.]	0.30	0.53	2.05	1.04	1.74

**Table 5.5** Execution times for the compared methods on the *China dataset*.



**Figure 5.8** (a) MS full resolution (4 m) image (used as reference image);  
(b) *EXP*, (c) *MTF*, (d) *FE* methods.



**Figure 5.9** The results in terms of *SAM* and *SCC* indexes of the proposed method obtained by varying the support sizes.

### 5.4.2 Full Scale Validation

The second test case is performed at full scale. Because of the unavailability of the reference image, the Spectral Angle Mapper (SAM) [66] between the MS image and the fused product has been calculated, together with the Spatial Correlation Coefficient (SCC) [65] between the details of the PAN image and the ones of the fused product, are used in order to evaluate the performance of the algorithm. This test is performed on the *Indianapolis dataset*.

Fig. 5.9 shows the performances of the proposed method achieved by varying the support dimension and fixing  $\mu = \lambda = 10^5$ . A good compromise between spectral consistency and spatial enhancement is provided by a support dimension equal to 13. In Tab. 5.6 the results of the *MTF* and *FE* algorithms are shown together with a comparison with state of the art pansharpening algorithms based on different detail extraction phases. The proposed approach obtains the best *SCC*, which indicates a good extraction phase, and a reasonable *SAM* value that evidences a good spectral consistency.

The most straightforward consequence of the reference image unavailability concerns the numerical assessment of fused prod-

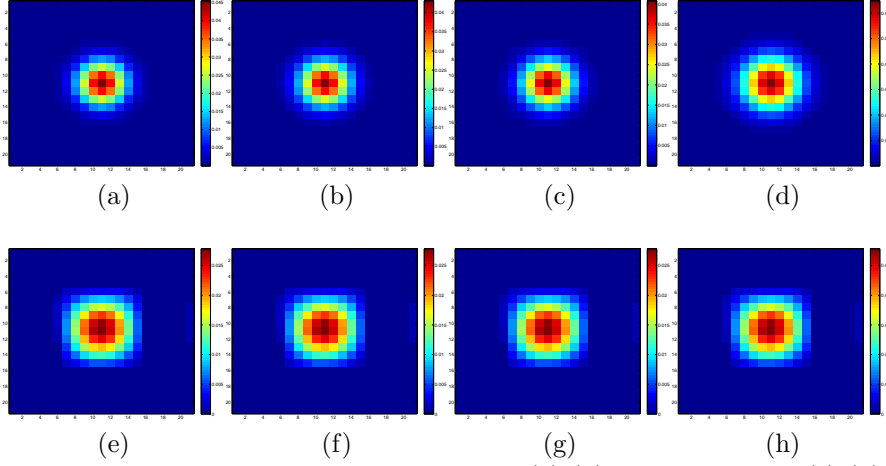
Algorithm	SAM( $^{\circ}$ )	SCC
<i>EXP</i>	<b>0</b>	0.4864
<i>SFIM</i>	0.6137	0.8664
<i>ATWT</i>	0.7904	0.9351
<i>MTF</i>	0.9772	0.9366
<i>FE</i>	0.8923	<b>0.9438</b>

**Table 5.6** Full scale performance evaluation using SAM and SCC indexes for *EXP*, *SFIM*, *MTF* and *FE* methods.

ucts quality. Thus, visual inspection constitutes the key step to understand the effectiveness of a pansharpening algorithm. Nevertheless, as further information, we quantify the performances of the considered approaches through the SAM and the SCC indexes. The former is a measure of the spectral distortion with respect to the original MS product, but its value has to be considered *cum grano salis*. Indeed, since the details are characterized by peculiar spectral features, a small SAM could eventually indicate a limited quantity of injected information [1]. Accordingly, the SAM has to be analyzed in conjunction with the SCC index that, on the other side, quantifies the amount of spatial details. It is worth to underline that other approaches are often employed in the literature to evaluate the fused products at full scale (e.g., the Quality with No Reference index [39]) but the applicability to MRA approaches could be compromised by the presence of severe aliasing effects [71], as it is the case of QuickBird data.

Fig. 5.10 shows for every band the comparison between the blurs derived by the *MTF* method and those estimated by the proposed technique. It is worth to note the asymmetric estimated shape (due to the blur introduced by the platform motion) which represents the main difference between the two filters. Moreover, in Fig. 5.11 the fused results achieved by the *MTF* and *FE* approaches are depicted by using a RGB representation, together with the MS and PAN starting images. Fig. 5.12 points out the differences in absolute value over the bands among the details extracted by the above-mentioned two methods. They are partic-

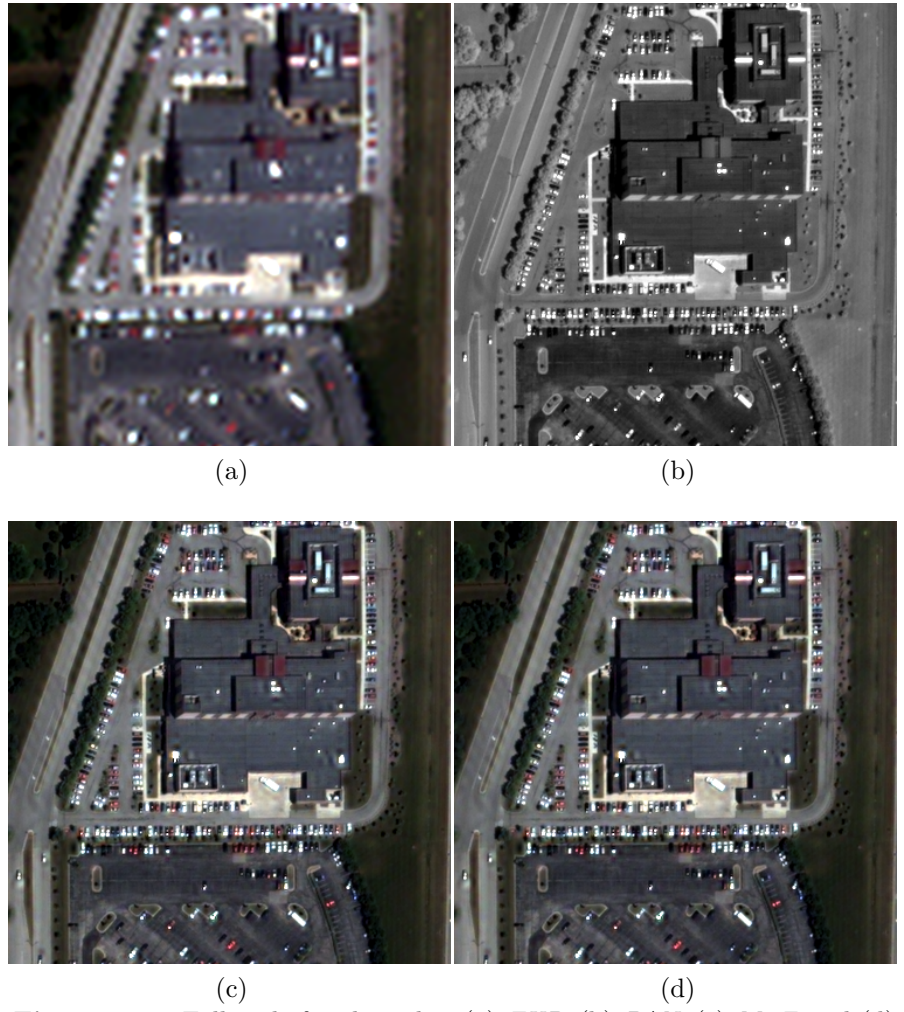




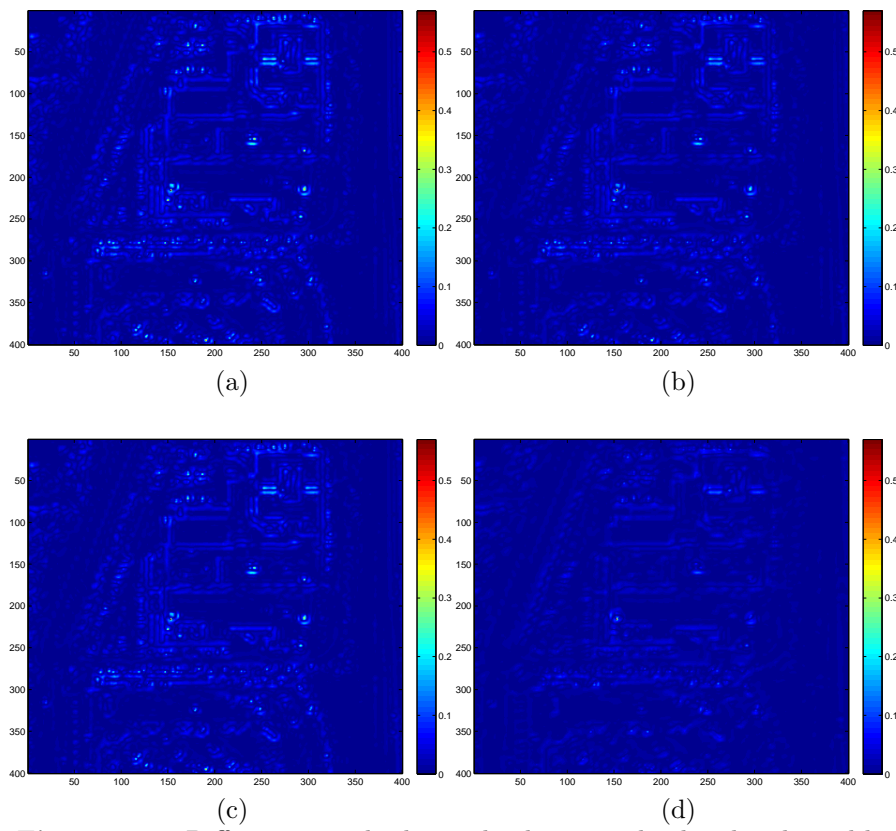
**Figure 5.10** From the top to the bottom: (a)-(d) Expected blurs; (e)-(h) Estimated blurs by the *FE* approach. The columns are ordered by wavelengths (*i.e.*, Blue, Green, Red and NIR).

ularly evident on the edges of small structures (such as paths and cars) and on the larger ones (such as buildings). The main differences can be highlighted in the comparison between the MTF for the blue band and the proposed method, while the results are closer for the NIR band. This similarity is mainly due to the fact that the NIR band is the one that most significantly contributes to the generation of the equivalent PAN image used for estimation.

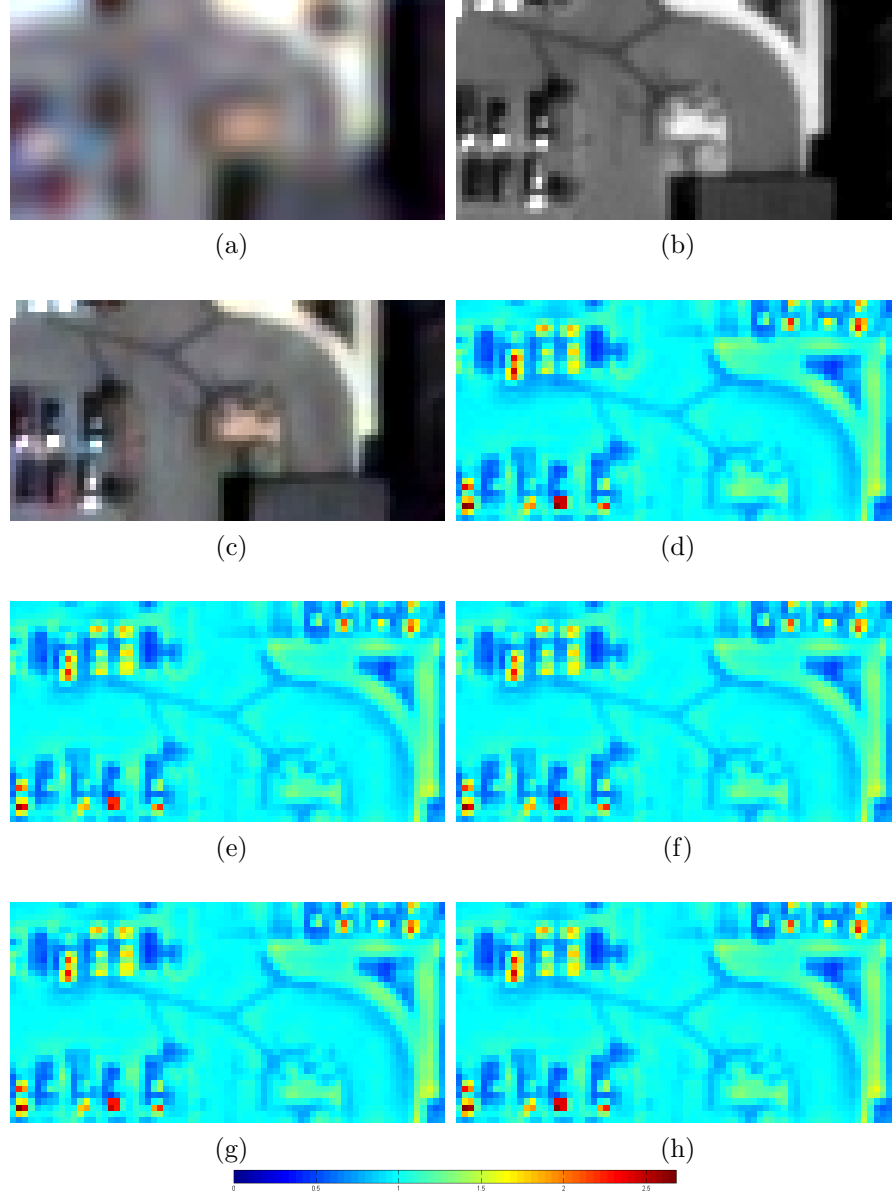
To help a visual inspection, a small area of the *Indianapolis dataset* is shown in Fig. 5.13. Since the injection model is the same, only the details for both the algorithms are reported. An unique image is presented for the proposed method, while four images, one for each band, are shown for the *MTF* approach. The greater evidence in the details of the proposed method compared to the ones of the *MTF* approach is clear, in particular, for some zones as the path and cars present in the area under study.



**Figure 5.11** Full scale fused results: (a) *EXP*, (b) *PAN*, (c) *MTF* and (d) *FE* images.



**Figure 5.12** Differences in absolute value between the details achieved by the *FE* and *MTF* methods over the bands, ordered by wavelengths (*i.e.* Blue, Green, Red, NIR).



**Figure 5.13** A small area in the *Indianapolis* dataset: (a) MS image; (b) PAN image; (c) RGB fusion result using the *FE* approach; Details extracted as  $\frac{\mathbf{P}}{\mathbf{P}_{LP}}$  using (d) *FE* and (e)-(h) *MTF* (different details for each spectral band) approaches.

## Chapter 6

# Hyperspectral Pansharpening

HyperSpectral (HS) sensors have a higher spectral resolution but a lower spatial resolution with respect to multispectral or panchromatic acquisitions. In order to increase their utilization and interpretation, hyperspectral images with both high spectral and spatial resolutions are desired. This goal can be achieved by means of pansharpening algorithms. Unfortunately, many pansharpening methods proposed in the literature have been properly developed for multispectral images but they have been never tested on hyperspectral data. In fact, several problems can arise when they are applied to the enhancement of hyperspectral images (e.g., the introduction of significant spectral distortions and the increasing of the computational effort due to the huge number of bands to fuse). For these reasons, in this chapter we firstly perform a critical comparison among classical approaches on hyperspectral data. Then, in order to alleviate the computational burden, we propose an approach based on: 1) Projection of the HS image into a lower dimensional feature space by means of Non-Linear Principal Component Analysis (NLPCA), 2) histogram matching and substitution of the panchromatic image into one or more nonlinear components and 3) projection back of the spatially enhanced components into the original spectral space. The comparison with

the Principal Component Analysis (PCA) is performed to demonstrate the effectiveness of the proposal. Furthermore, in order to reduce the spectral distortion, a hybrid approach is exploited by fusing through a MRA algorithm the most correlated nonlinear components with the panchromatic image instead of substituting them.

The rest of the chapter is as follows. In Sec. 6.1 classical approaches into the Component Substitution (CS) and MultiResolution Analysis (MRA) families are applied to hyperspectral data and their advantages and drawbacks are pointed out through an accurate at full scale experimental phase, while, Sec. 6.2 describes the use of the non-linear principal component analysis in hyperspectral sharpening and promotes the application of hybrid approaches to reduce both the spectral distortion and the computational burden.

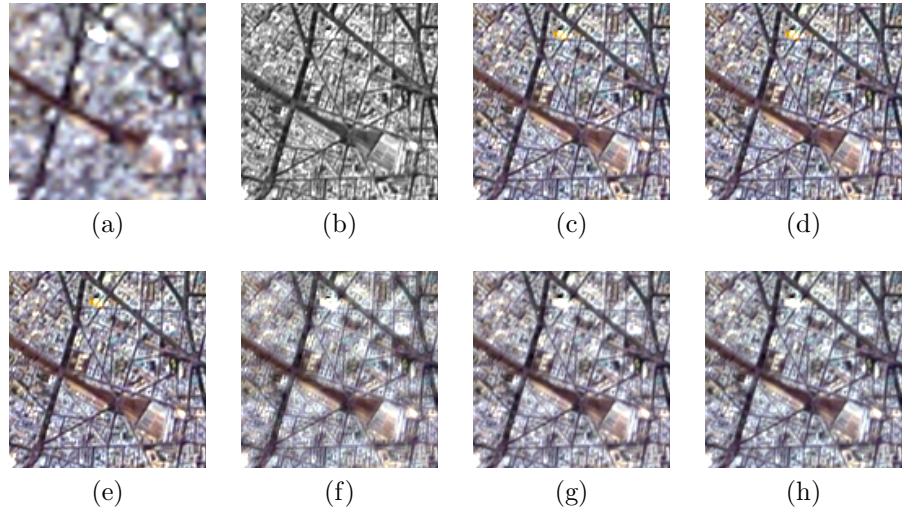
## 6.1 Classical Approaches

Pansharpening methods are generally divided into Component Substitution and MultiResolution Analysis techniques (see Chapter 2). *Principal Component Analysis* (PCA) and *Gram Schmidt* (GS) are surely very popular CS-based image fusion methods. All these methods rely upon the assumption that a component of the transformation tends to comprise the spatial information that is common to all the bands and thus can be substituted with a histogram matched version of the PAN image to achieve the HS enhanced data. MRA methods are based on the decomposition of each image by iterative applications of a given operator into a sequence of signals (or *pyramid*) with decreasingly informative content. Many approaches based on this concept have been proposed in the literature, employing *Wavelets* or *Gaussian pyramids*. On one hand, CS methods obtain fused images with high spatial quality but affected by spectral distortions. On the other hand, images obtained through MRA techniques in general are not very sharp but are more spectrally consistent.

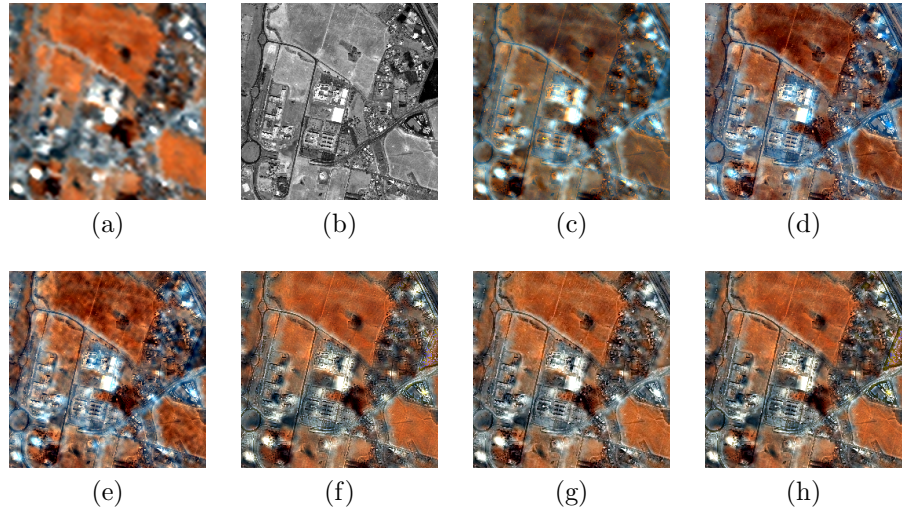
Here we compare several classical algorithms (see also Chapter 2 for details) applied to hyperspectral data: *Principal Component Analysis* [8], *Gram Schmidt* [11], *Gram Schmidt Adaptive* (GSA) [43], *Smoothing Filter-based Intensity Modulation* (SFIM) [55], *Generalized Laplacian Pyramid* with MTF-matched filter (GLP) [57], *Gaussian MTF-matched filter* with *High Pass Modulation* (MTF HPM) injection model [58].

The experiments have been conducted on two different datasets. The first one represents a small part of the center of Paris and is collected by the Hyperion sensor, which is capable of resolving 220 spectral bands (from 0.4 to 2.5  $\mu\text{m}$ ) with a 30-meter resolution, and by the panchromatic ALI sensor, which has a spatial resolution of 10 meters and a spectral coverage from 0.48  $\mu\text{m}$  to 0.69  $\mu\text{m}$ . Both sensors are mounted on the same platform, thus alleviating image co-registration problems. In the experiment, the sole bands which overlap the spectral range of the panchromatic channel are used (i.e., from band 14 to 33). The second scene represents a part of the city of Rome observed by the CHRIS-Proba sensor. In this case, no panchromatic channel is present on-board and the fusion is performed with a PAN image acquired by QuickBird (see Sec. 4.3 for its spectral and spatial characteristics). CHRIS-Proba has a spatial resolution of 17 meters and 18 bands with a variable spectral resolution, increasing from 1.25 nm at 415 nm to 11.25 nm at 1050 nm. Since the panchromatic channel of QuickBird has a spatial resolution of 0.6 m, it has been properly reduced by using almost ideal low pass filters to obtain a more suitable scale ratio of 4. We will name *Hyp+ALI* the first dataset and *CHRIS+QB* the second one.

We perform the validation of the fused products at full scale using the Quality with No Reference (QNR) index. The QNR consists of two parts: The spectral ( $D_\lambda$ ) and spatial ( $D_S$ ) distortions. The former, which estimates the spectral fidelity of the data, is time consuming when the number of bands to fuse increases. For this reason, we substitute it with the very popular Spectral Angle Mapper (SAM). We recall that the best values for  $D_S$  and SAM are zero.



**Figure 6.1** *Hyp+ALI* dataset: (a) MS; (b) PAN; (c) PCA; (d) GS; (e) GSA; (f) SFIM; (g) GLP; (h) MTF HPM.



**Figure 6.2** *CHRIS+QB* dataset: (a) MS; (b) PAN; (c) PCA; (d) GS; (e) GSA; (f) SFIM; (g) GLP; (h) MTF HPM.



The fusion results for the *Hyp+ALI* dataset are shown in Fig. 6.1 and the corresponding values of the numerical indexes are reported in Tab. 6.1. More in detail, in the MRA family the advantages of Gaussian MTF-matched filters with respect to the Box one (used by *SFIM*) are evident. In fact, the former are preferable mainly due to an increased amount of extracted details (as the  $D_S$  index values testify) and the reduction of artifacts (see Fig. 6.1). By comparing the injection strategies, the HPM injection model shows its superiority with respect to detail addition, as it is also the case for MS data fusion [53]. Into the CS family, the results are very similar. The behavior of the *GSA* algorithm is slightly penalized by the error committed in the estimation of the channel weights to construct the equivalent PAN image. Indeed, the performances should be very close to the ones obtained by the *GS* method, because the spectral responses of the Hyperion channels are very tight and all with the same bandwidth (as implicitly assumed by *GS*). The comparison between the two classes of pansharpening methods shows the performance advantages of the MRA approaches thanks to the reduction of the spectral distortion, which becomes increasingly relevant with the growth of the band number. On the other hand, the reduced computational burden and the robustness to misregistration errors (very common for images acquired by sensors on different platforms) suggest the use of CS algorithms. This last characteristic has been further investigated in this work, by introducing a shift in the grid of

Algorithm	SAM( $^\circ$ )	$D_S$
<b>PCA</b>	0.7928	0.1054
<b>GS</b>	0.7913	0.1053
<b>GSA</b>	0.8008	0.1062
<b>SFIM</b>	<b>0.4352</b>	0.1666
<b>GLP</b>	0.6650	0.0619
<b>MTF HPM</b>	0.6554	<b>0.0593</b>

**Table 6.1** Full scale analysis: *Hyp+ALI* dataset with perfectly coregistrated images.

the PAN image and the results are reported in Tab. 6.2. The CS algorithms obtain lower  $D_S$  values, indicating the ability to compensate the misregistration errors. The slight increase of all the SAM values is due to the differences of details among the bands, as well analyzed and justified in [71].

The results concerning the *CHRIS+ALI* dataset are in line with the previous test case (see Tab. 6.3 and Fig. 6.2). In general, higher values of SAM and  $D_S$  are reported, since the images were acquired by different sensors, resulting in different fields of view, temporal incoherence and so forth. Slight advantages in performances of CS methods can be evidenced, thanks to the greater robustness to misregistration errors. Since the CHRIS-Proba sensor has a significant variation of the spectral resolution over the bands, the *GSA* approach is advisable with respect to the *GS*. Indeed, we report that, by limiting the analysis to a small number of bands with higher bandwidths, the results of the *GS* and *GSA* becomes more comparable, as expected.

Finally, it is worth to note that the Box filter used by SFIM seems to work better than the MTF matched Gaussian one because of the preliminary spatial degradation of PAN image, carried out to obtain a reasonable scale ratio. From a visual point of view, the temporal coherency of the MRA approaches compared to the CS ones is evident, in particular, at the burned area, which is only present in the PAN image (see the right side of Fig. 6.2). Accordingly, the correct behavior consists in discarding this feature as for

Algorithm	SAM(°)	$D_S$
<b>PCA</b>	1.3194	0.0597
<b>GS</b>	1.3159	<b>0.0595</b>
<b>GSA</b>	0.8088	0.2525
<b>SFIM</b>	<b>0.4380</b>	0.5462
<b>GLP</b>	0.6895	0.3561
<b>MTF HPM</b>	0.6590	0.3664

**Table 6.2** Full scale analysis: *Hyp+ALI* dataset with 60 meters misalignment (2 HS pixels).

MRA approaches.

## 6.2 Pansharpening Based on Non-Linear PCA

In this section we want to detail the use of NLPCA [29] for hyperspectral pansharpening. In particular, we briefly describe the NLPCA approach by focusing on its implementation through Artificial Neural Networks. We subsequently describe the proposed algorithm aimed at mitigating the main drawbacks in fusing HS data, such as, the spectral distortion and the computational burden. Finally, the experimental results are shown.

### 6.2.1 Non-Linear PCA

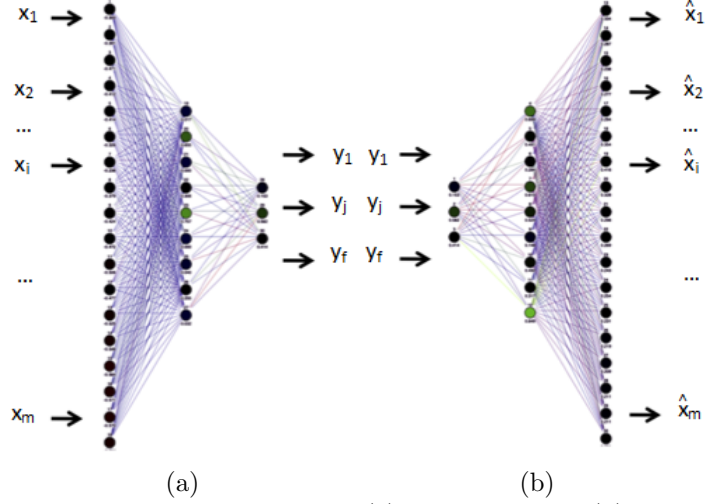
Let us consider a set of  $n$  observations of dimensionality  $m$ , which are denoted as  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_m]$ , where  $\mathbf{x}_i$  is the column vector composed by the  $i$ -th components of the  $n$  observations. In the NLPCA approach,  $\mathbf{X}$  is mapped into a lower dimensionality feature space through a set of nonlinear functions

$$\mathbf{Y} = \mathbf{z}(\mathbf{X}). \quad (6.1)$$

In this equation the  $n \times f$  resulting matrix  $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_j, \dots, \mathbf{y}_f]$  contains the Non-Linear Principal Components (NLPCs) and  $\mathbf{z} = [z_1, \dots, z_f]$  is the set of  $f < m$  nonlinear functions.

Algorithm	SAM(°)	$D_S$
<b>PCA</b>	10.6386	<b>0.2904</b>
<b>GS</b>	5.4846	0.4768
<b>GSA</b>	4.9747	0.3866
<b>SFIM</b>	<b>3.0762</b>	0.3097
<b>GLP</b>	4.2274	0.3904
<b>MTF HPM</b>	3.2355	0.3654

**Table 6.3** Full scale analysis: *CHRIS+QB* dataset with perfectly coregistered images.



**Figure 6.3** Networks implementing (a) mapping  $\mathbf{z}$  and (b) demapping  $\phi$  functions.

The inverse transformation, which aims at reconstructing the  $m$ -dimensional data from its  $f$ -dimensional representation, can be performed by a set of  $m$  nonlinear transformations  $\phi = [\phi_1, \dots, \phi_m]$ :

$$\hat{\mathbf{X}} = \phi(\mathbf{Y}) \quad (6.2)$$

where the  $n \times m$  matrix  $\hat{\mathbf{X}}$  indicates the reconstruction of  $\mathbf{X}$  (see also Fig. 6.3).

Analogously to PCA, the loss of information between the original and reconstructed data, measured by the error matrix  $\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}}$ , can be minimized by properly selecting the functions  $\mathbf{z}$  and  $\phi$ .

A widely employed method for implementing the latter transformations, relies upon the use of Artificial Neural Networks (ANNs), whose power is the consequence of the result that any nonlinear function can be approximated by a superposition of a set of transformations  $\sigma(x)$  that are continuous and monotonically increasing functions, with  $\sigma(x) \rightarrow 1$  as  $x \rightarrow +\infty$  and  $\sigma(x) \rightarrow 0$  as  $x \rightarrow -\infty$  [83]. This property is often called *universal fitting* and its hypotheses are fulfilled by sigmoidal functions [84]. Specifically,

the approximation of any nonlinear function can be obtained by an ANN using one single layer of nodes with sigmoidal activation functions and two layers of weighted connections [84]. In practical implementations, sigmoidal nonlinearities are often included in the nodes of the output layer so that the network produces outputs in a fixed and finite range. The ability of the ANN to fit arbitrary nonlinear functions depends on the presence of a hidden layer with nonlinear nodes. Without the hidden layer (or with linear nodes in the hidden layer), the network is only capable of producing linear combinations of the inputs, given linear nodes in the output layer. A network lacking a hidden layer but including sigmoidal nonlinearities in the output layer is only capable of generating multivariable sigmoidal functions.

Starting from these considerations, the NLPCA can be implemented by two ANNs playing the role of the nonlinear functions  $\mathbf{z}$  and  $\phi$  [29]. The ANN approximating  $\mathbf{z}$ , called *coding subnet*, is composed by an input layer with  $m$  nodes followed by the hidden layer (often called *mapping layer*) with  $M_1 > f$  nodes and sigmoidal transfer functions (to verify the *universal fitting* property). The output layer of this subnet contains  $f < m$  nodes and for this reason it is often called *bottleneck*. The second ANN (also called *decoding subnet*) approximates the  $\phi$  function. Its input layer has  $f$  nodes followed by the hidden layer (often called *demapping layer*) with  $M_2 > f$  nodes and sigmoidal transfer functions (aimed again to verify the *universal fitting* property). The output layer yields the reconstructed data and thus contains  $m$  nodes. The nodes of the output layers can be linear or sigmoidal. ANNs require supervised training that consists in tuning the network in order to obtain a desired output. However, only the input of the coding network and the output of the decoding network are known, while the outputs of the coding subnet, and hence the inputs of the decoding subnet, are unknown. Therefore, direct supervised training of the two networks is unfeasible. To overcome this problem, one can observe that by combining in series the two ANNs, or, equivalently, by defining a composite function  $\psi(\cdot) = (\mathbf{z} \circ \phi)(\cdot) = \phi(\mathbf{z}(\cdot))$  which links the original data  $\mathbf{X}$  with

their reconstructed version  $\hat{\mathbf{X}}$ :

$$\hat{\mathbf{X}} = \psi(\mathbf{X}), \quad (6.3)$$

the combined network can be trained to produce the identity mapping. In other terms, the parameters of the network implementing  $\psi$  are optimized so that the reconstructed outputs match the inputs as closely as possible. The training aimed at learning the identity mapping has been called *self-supervised backpropagation* or *autoassociation* [85], thus suggesting the name of AutoAssociative NNs (AANNs). The training of AANNs is an iterative process that is completed when the sum of squared errors is minimized:

$$E = \sum_{p=1}^n \sum_{i=1}^m (\hat{\mathbf{x}}_i - \mathbf{x}_i)_p^2 \quad (6.4)$$

or, equivalently, when the matrix error norm  $\|\mathbf{E}\|$  is minimized (as in the case of PCA).

Once the complete AANN  $\psi$  is trained, it is possible to use the *coding subnet* to project the original data into a lower dimensional space. Thus, the  $f$  NLPCs can be obtained from the output layer of the *bottleneck*. The obtained NLPCs can be subsequently used as input to the *decoding subnet* in order to obtain the reconstructed data.

As it can be seen, the main difference between PCA and NLPCA is that the latter employs nonlinear relations between variables, while PCA is based on linear ones. As a consequence, NLPCA has the relevant advantage to retain most information, that is typically quantified by the variance of the data, in fewer components with respect to PCA.

One of the main difficulties in designing the AANN relies in the selection of the correct number of nodes that minimizes the loss of information produced in the three hidden layers, and, in particular, in the *bottleneck*. Being the AANN designed to minimize the reconstruction error, the best NN topology can be retrieved by using a simple grid search algorithm that varies recursively the number of nodes of the hidden layers and selects the topology presenting the smallest error. However, without a well-chosen

starting point, this approach can be extremely time consuming and the probability of locally optimal solutions due to data overfitting turns out to be very high. To mitigate this problem a further condition on the network complexity is forced; since the outputs impose  $n \cdot m$  constraints, a more regular solution can be achieved by imposing that the number of the adjustable parameters (weights and biases for all network connections and nodes, respectively) must be much less than this number. Because the number of network weights is  $(M_1 + M_2)(m + f + 1) + m + f$  the following inequality has to be fulfilled:

$$M_1 + M_2 \ll \frac{m(n - 1) - f}{m + f + 1}. \quad (6.5)$$

Since the aim of a dimensionality reduction method is to obtain a representation of the original signal into a lower dimensional space, i.e.,  $f \ll m, n$ , Eq. (6.5) can be easily approximated as:

$$M_1 + M_2 \ll n. \quad (6.6)$$

Furthermore, by assuming a balanced structure of the AANN, i.e.,  $M_1$  and  $M_2$  should have the same dimensions ( $M_1 = M_2 = M$ ), a widely used design condition is the following:

$$2M \ll n. \quad (6.7)$$

It is worth to note that Eq. (6.7) is effective only if the number of mapping/demapping nodes  $M$  is greater then the number of nodes in the bottleneck layer  $f$ . Otherwise, there will not be enough data to effectively extract  $f$  NLPCs. Moreover, since the output has simply the task of replicating the input, no *a priori* knowledge is needed for the implementation of the learning phase. This implies that the AANN training can be performed in a fully automatic way and that all the image pixels can be used for this step, as it has been done in this work.

### 6.2.2 Proposed Approach

To fully describe the proposed pansharpening algorithm, we firstly recall the used notation. We indicate the HS low spatial resolution

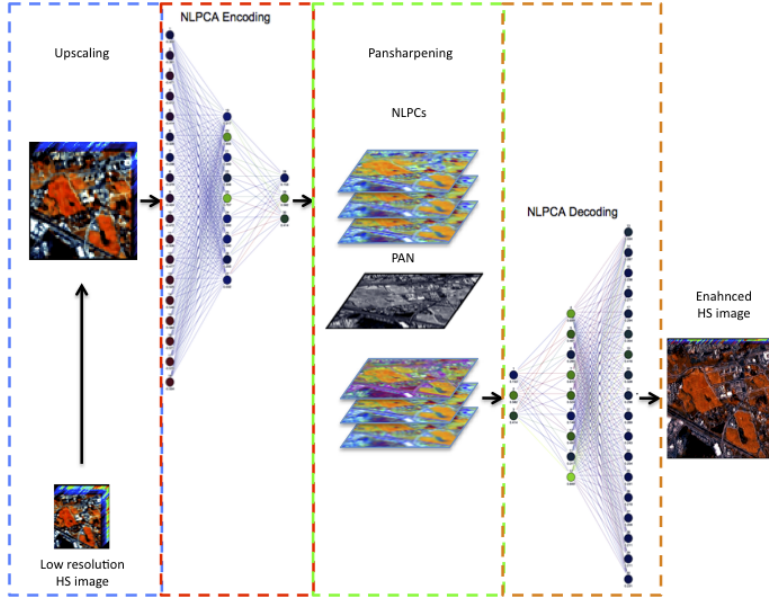
image as  $\mathbf{HS}$  and define the vectorized form of the panchromatic image  $\mathbf{P}$  as  $\mathbf{p} = \text{vec}(\mathbf{P})$ , where  $\text{vec}$  is the operator that transforms the original image in its vectorized form (i.e., arranges in a column vector the pixels of the image, extracted by columns). Then the interpolated version of  $\mathbf{HS}$  to the resolution of  $\mathbf{p}$  is denoted as  $\widetilde{\mathbf{HS}} = [\widetilde{\mathbf{hs}}_1, \dots, \widetilde{\mathbf{hs}}_k, \dots, \widetilde{\mathbf{hs}}_N]$  where  $\widetilde{\mathbf{hs}}_k$  is the vectorized form of the  $k$ -th band of the interpolated HS image and  $N$  the band number.

The first step of the approach is to train the AANN by using the whole dataset  $\widetilde{\mathbf{HS}}$  in order to obtain the functions  $\mathbf{z}$  and  $\phi$  for the coding and decoding phases, respectively. Subsequently we exploit Eq. (6.1) to get the set of  $f$  NLPCs, called  $\mathbf{NLPC} = [\mathbf{nlpc}_1, \dots, \mathbf{nlpc}_f]$ , from now on. In Eq. (6.1), the role of  $\mathbf{X}$  is played by  $\widetilde{\mathbf{HS}}$ .

Since the NLPCs tend to represent different characteristics of the spectra, the injection of the same amount of PAN spatial details into each component would not lead to an acceptable result. In order to overcome this problem, only the NLPCs for which the correlation coefficients with the PAN image are higher than a given threshold  $\theta$  are selected for the sharpening. The threshold  $\theta$  is chosen equal to  $0.9 \cdot \rho_{max}$ , where,  $\rho_{max}$  represents the maximum value of the correlations between each NLPC and the panchromatic image. The selected NLPCs are then substituted with a histogram-matched version of the PAN image. At the end, both the histogram-matched PAN images and the NLPCs that do not exceed the threshold are used in Eq. (6.2) to obtain the final pansharpened product  $\widehat{\mathbf{HS}}$ . This CS algorithm will be named *NLPCA*, from now on.

Furthermore, in order to reduce the relevant spectral distortion, a modification of the above-mentioned algorithm is proposed. Indeed, a MRA fusion approach is exploited instead of the classical substitution of the selected components with a histogram matched version of the PAN image. Therefore, only the details of the latter are injected into the NLPC to get the fused components. Several MRA approaches have been tested and they will be described in Sec. 6.2.3. The complete NLPCA-based hybrid processing pro-





**Figure 6.4** Complete scheme of the proposed hybrid approach.

posed in this chapter is resumed in Algorithm 3 and the schematic representation of the fusion process is depicted in Fig. 6.4.

### 6.2.3 Experimental Results

The experimental results are conducted on two different HS datasets. Firstly, we exploit the *Pavia dataset*, a 103 bands hyperspectral image acquired by the ROSIS airborne sensor. In this case, no panchromatic image is available and it is simulated by averaging the values of its bands. This choice has the relevant advantage that the resulting pansharpened image does not suffer from distortions caused by misregistration between the HS and PAN. Moreover, both synthetic PAN and HS images have the same spectral coverage and temporal coherence, thus avoiding the insurgence of further spectral distortions. Another important remark concerning these kinds of data is the linear correlations among the spectral bands [86]. This suggests that, since the NLPCA approach is able to deal with both linear and nonlinear correlations,

---

**Algorithm 3:** NLPCA Hybrid Pansharpening Algorithm
 

---

**Data:** The upsampled HS image  $\widetilde{\mathbf{HS}}$ , the PAN image  $\mathbf{p}$  and the threshold  $\theta$ .

**Result:** Fused HS image  $\widehat{\mathbf{HS}}$ .

**begin**

    Train the AANN using  $\widetilde{\mathbf{HS}}$  to get the functions  $\mathbf{z}$  and  $\phi$

    Use Eq. (6.1) with  $\mathbf{X} = \widetilde{\mathbf{HS}}$  to get

$\mathbf{NLPC} = [\mathbf{nlpc}_1, \dots, \mathbf{nlpc}_f]$

$\mathcal{H} = \emptyset$

**for**  $i \in \{1, \dots, f\}$  **do**

        Compute the correlation  $\rho$  between  $\mathbf{p}$  and  $\mathbf{nlpc}_i$  to

        get  $\rho_i = \rho(\mathbf{p}, \mathbf{nlpc}_i)$

**if**  $\rho_i > \theta$  **then**

            Inject the spatial information of  $\mathbf{p}$  into the  $\mathbf{nlpc}_i$   
             to get  $\widehat{\mathbf{nlpc}}_i$

$\mathcal{H} = \mathcal{H} \cup \{\widehat{\mathbf{nlpc}}_i\}$

**else**

$\mathcal{H} = \mathcal{H} \cup \{\mathbf{nlpc}_i\}$

    Use Eq. (6.2) with  $\mathbf{Y} = \mathcal{H}$  to get  $\widehat{\mathbf{HS}}$

---

the results obtained by NLPCA and PCA (which only takes into account of linear correlations) should be very close.

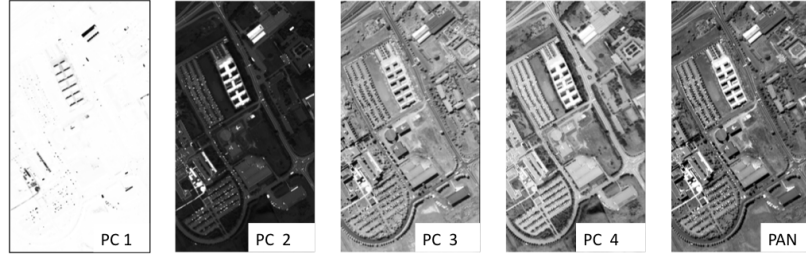
The second test is based on the *CHRIS+QB dataset* (already described in Sec. 6.1). In this case the correlations among bands have been shown to be nonlinear [87], thus suggesting the exploitation of nonlinear approaches to improve the results.

In both cases, the resolution ratio  $R$  is chosen to be 4 because it is typical for pansharpening applications and offers the best tradeoff between spatial enhancement and spectral distortion [30].

The algorithms for the comparison are listed below:

- *PCA*: Substitution of the first principal component with the PAN image [8]
- *HPF*: *Box filter* and *additive* injection model [5]
- *SFIM*: *Box filter* and *HPM* injection model, even called Smoothing Filter-based Intensity Modulation [55]
- *GLP*: *Generalized Laplacian Pyramid* with MTF adjustment [57]
- *MTF HPM*: Gaussian pyramid with MTF adjustment [57] and *HPM* injection model [58]
- *ATWT*: A Trous Wavelet with *additive* injection model [60]
- *AWLP*: *Additive Wavelet Luminance Proportional* [65]

The quantitative analysis is performed both at reduced and full scale. In the former case, the Spectral Angle Mapper (SAM), the Relative Global Error of Synthesis (ERGAS) and the Spatial Correlation Coefficient (SCC) [5] are used to evaluate the performances. At full scale, only SAM and SCC are used for quantifying the spectral and spatial distortions, respectively. Best values for SAM and ERGAS are 0, while, 1 for SCC.



**Figure 6.5** The 4 NLPCs and the PAN image for the *Pavia dataset*.

	NLPC 1	NLPC 2	NLPC 3	NLPC 4
PAN	-0.0536	0.9064	0.9477	0.2870

**Table 6.4** Correlation coefficients between PAN image and each NLPC on the *Pavia dataset*.

### Pavia Dataset

To demonstrate the effectiveness of the NLPCA, we evaluate the performances exploiting the validation procedure at reduced scale on the *Pavia dataset*.

According to the proposed method, the first step consists in projecting the HS image into a nonlinear feature space by means of NLPCA. The *Pavia dataset* consists of 103 bands, hence a grid search to optimize the number of nodes in the hidden layers of the nonlinear AANN has been carried out. The best topology is found to have 103 inputs/output nodes, 50 nodes in both the outer hidden layers and 4 nodes in the bottleneck layer, corresponding to 4 NLPCs. The PAN image is compared to each of the 4 NLPCs in order to detect the most similar components. As it can be noted by observing Fig. 6.5, from a qualitative point of view, component 3 presents characteristics very similar to the PAN image. Even by analyzing the correlation values between each component and PAN image, we can see that component 3 reaches the highest correlation, as reported in Tab. 6.4.

These values suggest that components 3 and 2 present similar characteristics with the PAN image. For this reason, in these experiments we perform different tests by fusing the information of the PAN image not only with the component 3, but also with

the other components. In greater details, four different configurations are tested. They consist in combining the PAN image with component 3, components 2 and 3, components 2, 3 and 4 and all the components, respectively. In each test, the selected components are then spatially enhanced and substituted in the NLPCs dataset. This new feature dataset is then projected back to the original spectral domain by using the decoding subnet. In each test, the fusion process is carried out by employing different well-known techniques and the final fused images are evaluated in terms of the cited quality indexes.

Tabs. 6.5, 6.6 and 6.7 show the SAM, ERGAS and SCC quality indexes associated to the different approaches and to the yardsticks consisting in applying the different techniques directly to the HS images (indicated by the label “*All bands*”, from now on). As expected, among the results obtained with the proposed approach, the best tradeoff between spectral and spatial quality is obtained by fusing the PAN image with components 2 and 3, suggesting that the use of the sole component 3 is not sufficient, even if the latter is the most correlated with the PAN image. Furthermore, the fusion of the components 4 and 1 do not improve the quality of the enhanced images, but, in most cases, tends to increase the distortions. This choice also supports the rule of thumb proposed in the previous section to select the components to fuse. By a comparison of these results with those obtained by utilizing the same fusion techniques directly to the spectral bands, it can be noted that the use of the NLPCA pre-processing improves the spectral quality of the enhanced images, though preserving the spatial coherency.

Other important remarks come from the comparison of the results obtained by using the linear (i.e., PCA) and the non-linear PCA (i.e., NLPCA), respectively. The two methods produce enhanced images having almost the same overall quality, with a slightly larger spectral distortion of the *NLPCA* compared to *PCA*, but with a better radiometric distortion, as measured by the ERGAS index. Among the fusion techniques, *GLP*, applied to components 2 and 3, achieves the best results in terms of spectral

NLPCs	3	2, 3	2, 3, 4	1, 2, 3, 4	All bands
<b>NLPCA</b>	7.0120	6.5856	9.9686	9.5204	/
<b>PCA</b>	/	/	/	/	<b>6.6061</b>
<b>HPF</b>	6.8376	6.6364	6.9274	6.8534	6.8021
<b>SFIM</b>	6.8380	6.6978	7.0706	6.8425	11.1472
<b>GLP</b>	6.7972	<b>6.4943</b>	<b>6.8159</b>	6.6975	6.8451
<b>MTF HPM</b>	6.7369	6.5685	6.9588	6.7327	8.9873
<b>ATWT</b>	<b>6.7157</b>	6.5087	6.8447	<b>6.6953</b>	6.8832
<b>AWLP</b>	<b>6.7157</b>	6.5087	6.8447	6.6957	6.7253

**Table 6.5** SAM quality indexes (measured in degrees) obtained with different fusion approaches applied to the *Pavia dataset*. Pure MRA are indicated by “All bands”. CS approaches are labeled as NLPCA and PCA, while the others are hybrid approaches based on NLPCA projection and the indicated MRA methodology. The symbol “/” indicates an unavailable value.

quality, while *AWLP* applied to NLPCs 2 and 3 results in the higher spatial consistency value. After a more general evaluation, *ATWT*, *AWLP*, *MTF HPM* and *GLP* applied to NLPCs 2 and 3 present the best tradeoff between spatial enhancement and spectral distortion.

From a qualitatively point of view, the enhanced images obtained by combining the PAN image with NLPCs 2 and 3 according to the different fusion paradigms appear to be very sharp and spectrally consistent with the reference image, as illustrated by Figs. 6.6 and 6.7.

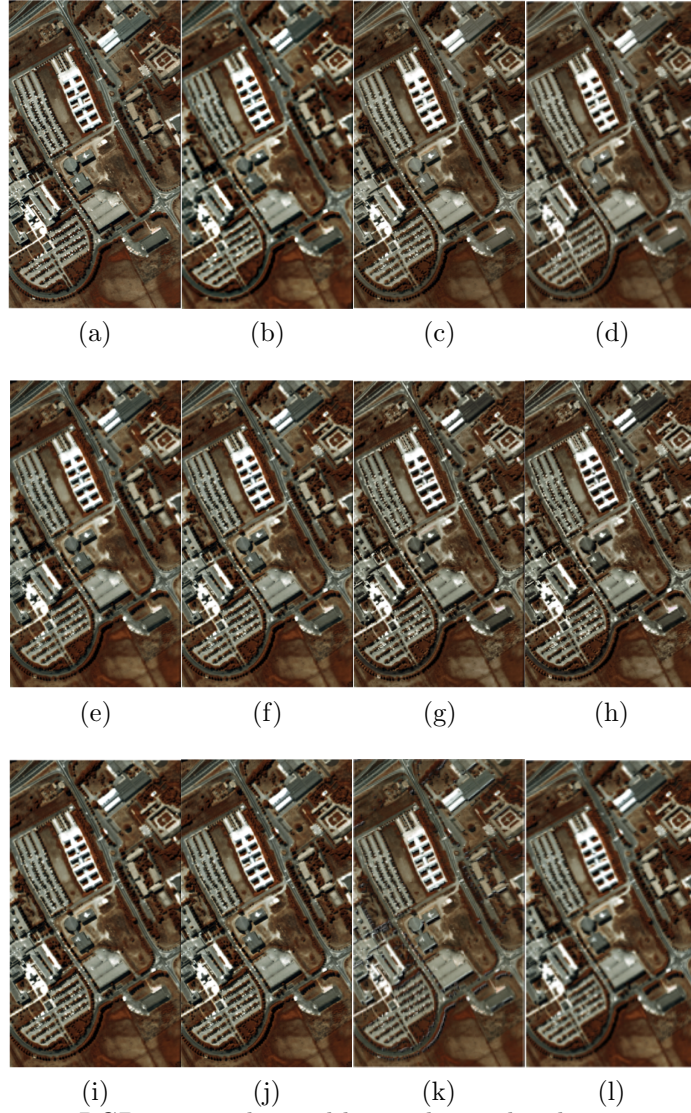
A further analysis is carried out in order to evaluate the computational burden of the different methods. Tab. 6.8 reports the required times, expressed in seconds, for each technique to produce the enhanced image. As it can be clearly seen, the use of NLPCA improves the computational efficiency of the fusion process by reducing the band number to fuse. This is mainly evident for MRA approaches that have higher order filters to apply to the images to fuse (e.g., Gaussian-based ones).

NLPCs	3	2, 3	2, 3, 4	1, 2, 3, 4	All bands
<b>NLPCA</b>	<b>7.0540</b>	6.9151	16.0414	15.6749	/
<b>PCA</b>	/	/	/	/	8.8601
<b>HPF</b>	7.1860	7.1711	8.0134	8.0237	6.8725
<b>SFIM</b>	7.1459	7.1140	8.3495	8.2264	7.3360
<b>GLP</b>	7.1406	<b>6.0995</b>	<b>7.8815</b>	<b>7.8650</b>	<b>6.5513</b>
<b>MTF HPM</b>	7.1325	6.0783	8.4154	8.2797	7.2173
<b>ATWT</b>	7.1346	6.1002	7.9359	7.8999	6.6527
<b>AWLP</b>	7.1346	6.1007	7.9359	7.8997	6.7529

**Table 6.6** ERGAS quality indexes obtained with the different fusion approaches applied to the *Pavia dataset*. Pure MRA are indicated by “All bands”. CS approaches are labeled as NLPCA and PCA, while the others are hybrid approaches based on NLPCA projection and the indicated MRA methodology. The symbol “/” indicates an unavailable value.

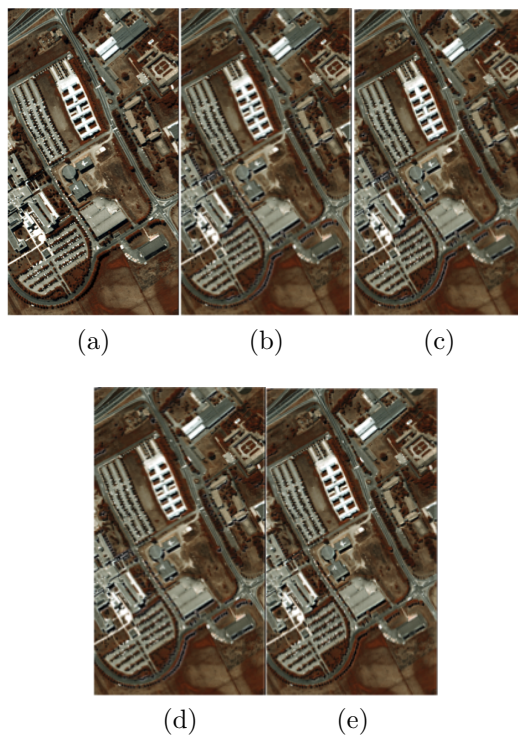
NLPCs	3	2, 3	2, 3, 4	1, 2, 3, 4	All bands
<b>NLPCA</b>	<b>0.8364</b>	<b>0.8669</b>	0.7565	0.7288	/
<b>PCA</b>	/	/	/	/	0.7737
<b>HPF</b>	0.7911	0.8104	0.7876	0.7859	0.8467
<b>SFIM</b>	0.7910	0.8061	0.7898	0.7903	0.6928
<b>GLP</b>	0.7972	0.8178	0.7982	0.7978	0.8520
<b>MTF HPM</b>	0.7964	0.8140	0.8031	<b>0.8056</b>	0.7926
<b>ATWT</b>	0.7985	0.8199	0.7989	0.8013	0.8525
<b>AWLP</b>	0.7985	0.8565	<b>0.8589</b>	0.8012	<b>0.8597</b>

**Table 6.7** SCC quality indexes obtained with the different fusion approaches applied to the *Pavia dataset*. Pure MRA are indicated by “All bands”. CS approaches are labeled as NLPCA and PCA, while the others are hybrid approaches based on NLPCA projection and the indicated MRA methodology. The symbol “/” indicates an unavailable value.



**Figure 6.6** RGB images obtained by combining bands 70, 50 and 20 for the *Pavia dataset*: (a) Reference image; (b)  $\widetilde{\mathbf{HS}}$ ; Fused products achieved by means of (c) NLPCA; (d) PCA; Hybrid-NLPCA: (e) HPF; (f) SFIM; (g) GLP; (h) MTF HPM; (i) ATWT; (j) AWLP; Directly (all bands) applied to HS and PAN images: (k) HPF; (l) SFIM. The fused images are obtained using the NLPCs 2 and 3.

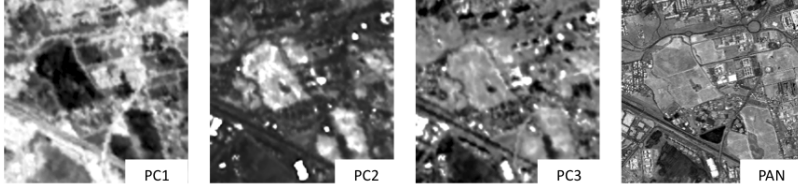




**Figure 6.7** RGB images obtained by combining bands 70, 50 and 20 for the *Pavia dataset*: (a) Reference image; Fused products obtained by directly applying to HS and PAN images the following methods: (b) GLP; (c) MTF HPM; (d) ATWT; (e) AWLP.

NLPCs	3	2, 3	2, 3, 4	1, 2, 3, 4	All bands
<b>NLPCA</b>	5.11	5.15	5.15	5.20	5.25
<b>PCA</b>	/	/	/	/	5.00
<b>HPF</b>	5.10	5.14	5.17	5.25	6.25
<b>SFIM</b>	5.11	5.13	5.18	5.21	6.25
<b>GLP</b>	6.33	7.37	8.86	11.59	104.99
<b>MTF HPM</b>	6.33	7.37	8.86	11.59	104.99
<b>ATWT</b>	5.29	5.55	5.80	6.28	32.81
<b>AWLP</b>	5.29	5.56	5.82	6.33	32.85

**Table 6.8** Computational time (expressed in seconds) evaluated for the different methods applied to the *Pavia dataset*. Pure MRA are indicated by “All bands”. CS approaches are labeled as NLPCA and PCA, while, the others are hybrid approaches based on NLPCA projection and the indicated MRA methodology. The symbol “/” indicates an unavailable value.



**Figure 6.8** The 3 NLPCs and the PAN image for the *CHRIS+QB* dataset.

	NLPC 1	NLPC 2	NLPC 3
PAN	-0.0906	0.2340	0.3951

**Table 6.9** Correlation coefficients between PAN image and each NLPC obtained by the *CHRIS+QB* dataset.

### CHRIS+QB Dataset

In this second experiment, we apply the NLPCA to the *CHRIS+QB* dataset. The CHRIS-Proba image presents 18 bands and a spatial resolution of 20 meters, while the PAN image has a resolution of about 0.6 meters (degraded with an almost ideal filter [57] to 5 meters to obtain a more suitable spatial resolution ratio of 4). As in the previous experiment, in order to project the HS image into the nonlinear feature space, all the pixels of the upsampled CHRIS-Proba image are used to train the AANN. The best AANN topology configuration is found to have 18 inputs/output nodes, 9 nodes in both the outer hidden layers and 3 nodes in the bottleneck layer, resulting in 3 NLPCs. Being the two images acquired in different dates, with different atmospheric conditions and with different angles of view, high correlation values between the NLPCs and the PAN image are not expected, as confirmed by Fig. 6.8 and Tab. 6.9. Indeed, in this case, only one component (the third one) presents an appreciable correlation with the PAN image and it is the unique component that overcomes the threshold  $\theta$  for the fusion.

Even in this case, several other pansharpening approaches are compared in order to assess the performances of the proposal. After the fusion, all the components are projected back to the original spectral domain by using the demapping subnet of the

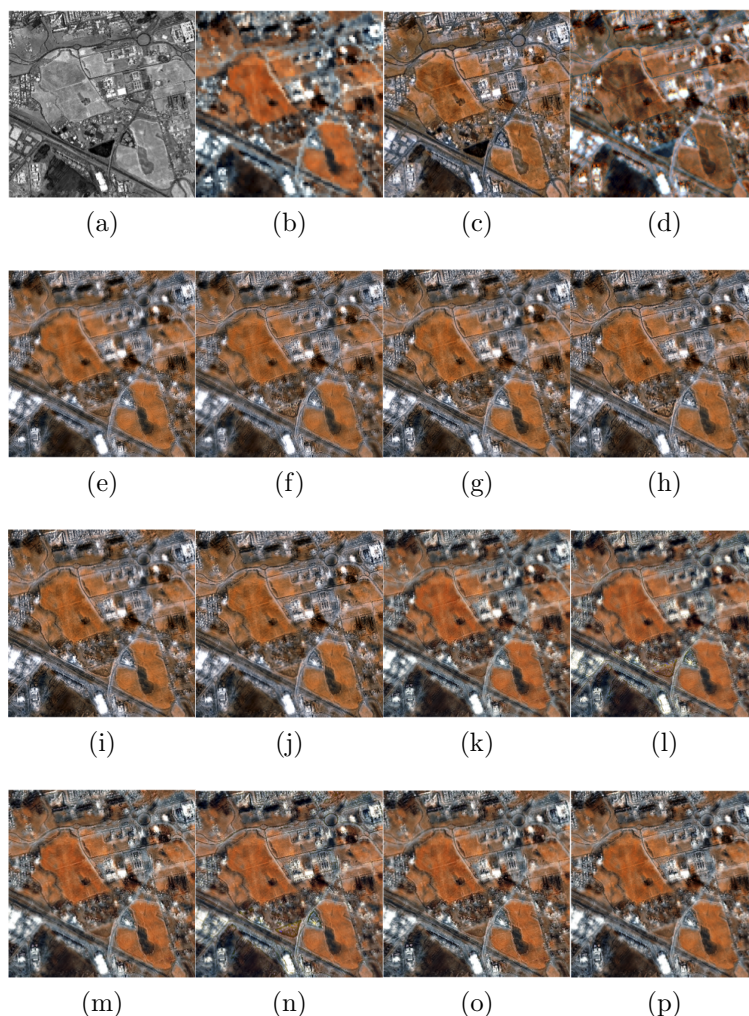
	SAM		SCC	
	Hybrid-NLPCA	All bands	Hybrid-NLPCA	All bands
<b>NLPCA</b>	/	3.6093	/	<b>0.8893</b>
<b>PCA</b>	/	9.9237	/	0.6624
<b>HPF</b>	<b>3.1074</b>	3.8072	0.7241	0.7249
<b>SFIM</b>	3.2007	3.5930	0.7019	0.7067
<b>GLP</b>	3.1597	3.9273	0.7973	0.7949
<b>MTF HPM</b>	3.2477	3.8603	0.7817	0.7934
<b>ATWT</b>	3.1674	4.2287	0.8029	0.8001
<b>AWLP</b>	3.1674	4.1565	0.8029	0.7944

**Table 6.10** SAM (expressed in degrees) and SCC quality indexes over the different fusion approaches applied to the *CHRIS+QB dataset*. The symbol “/” indicates an unavailable value.

AANN. For the sake of comparison, also the results obtained with the same techniques, directly applied to the HS images (labeled again as “*All bands*”) are considered.

In this case, a full scale validation is performed. For this reason, we only consider SAM, as the spectral quality index, evaluated between the fused product and the original HS image, and SCC, evaluated between the details of the outcome of the fused algorithm and the PAN image, for the spatial quality. Even if the latter is not advisable as a general evaluation index [1], it is possible to use it as a parameter quantifying the ability of the different techniques to extract spatial details and, thus, assess the sharpness of the final product.

Tab. 6.10 shows the SAM and SCC quality indexes associated to the different approaches. It can be shown that, in most cases, independently from the chosen extraction detail technique, the projection into the non-linear feature space permits to achieve better results, in terms of spectral and spatial quality, if compared with the same fusion method applied directly to the HS image. In particular, the enhanced images obtained with a NLPCA pre-processing present a lower spectral distortion with a comparable spatial consistency. A direct comparison of the *NLPCA* method with the *PCA* demonstrates that, in this test case, non-linear contributions have to be taken into account. Indeed, the NLPCA permits to enhance the spatial information of the hyperspectral



**Figure 6.9** RGB images obtained by combining bands 11, 9 and 7 for the *CHRIS+QB* dataset: (a) PAN; (b)  $\widetilde{\mathbf{HS}}$ ; Fused products by the means of (c) NLPCA; (d) PCA; Hybrid-NLPCA: (e) HPF; (f) SFIM; (g) GLP; (h) MTF HPM; (i) ATWT; (j) AWLP; Directly (all bands) applied to HS and PAN images: (k) HPF; (l) SFIM; (m) GLP; (n) MTF HPM; (o) ATWT; (p) AWLP. The NLPCA-based fused images are obtained using the third NLPC.

	Hybrid-NLPCA	All bands
<b>NLPCA</b>	0.16	/
<b>PCA</b>	/	0.13
<b>HPF</b>	0.22	2.69
<b>SFIM</b>	0.22	2.69
<b>GLP</b>	0.95	28.26
<b>MTF HPM</b>	0.95	28.25
<b>ATWT</b>	0.38	5.81
<b>AWLP</b>	0.38	5.83

**Table 6.11** Computational time (expressed in seconds) evaluated for the different methods applied to the *CHRIS+QB dataset*. The symbol “/” indicates an unavailable value.

image without adding relevant distortions. This is also evident by qualitatively analyzing the results in Fig. 6.9. In particular, it can be noted that, while the image obtained following the *PCA* approach presents evident spectral distortion, the one obtained by substituting the third NLPC with the PAN image is spatially well-defined and presents spectral characteristics very close to the original image. In this case, the best spectral quality is achieved by the *HPF* method applied to the third NLPC, while the best spatial consistency is obtained by directly substituting the PAN image to NLPC 3. However, on a more general analysis, the best tradeoff between spectral and spatial qualities is again achieved by *ATWT*, *AWLP* and *GLP* methods. At the end, it is worth to underline that the hybrid approaches mitigate the problem of the temporal coherence into the CS family. In fact, the cascade of a CS and a MRA approach tends to have a behavior similar to a traditional MRA algorithm with a greater temporal coherence with respect to the product obtained by exploiting only the CS algorithm (see Sec. 6.1). This is evident in Fig. 6.9, where the burned area (depicted only in the PAN image) is still shown in the *NLPCA* (or, equivalently, *PCA*) fused image but not in the hybrid fusion products.

Tab. 6.11 shows the required time, expressed in seconds, for each technique to produce the fused image. Again, the use of

the NLPCA improves the computational efficiency of the image enhancement process. The overall time to perform the NLPCA transformation has been measured to be 0.16 seconds.





# Chapter 7

## Conclusions

In recent years, pansharpening has received a great attention into the Data Fusion Community, as can be seen by the number of different techniques proposed for accomplishing this task. In this dissertation, a classification of the methods presented in the literature into two main families (i.e., Component Substitution and MultiResolution Analysis) has been presented in Chapter 2 and some widely used and popular algorithms belonging to these families have been described. Due to the differences in spectral and spatial characteristics of the products to fuse, the validation of the final outcome represents a difficult task. For this reason, Chapter 2 has also focused its attention on the validation methods for the pansharpening products. The two approaches used in the literature for the validation, i.e., analysis at reduced and full scale, have been considered. The former follows the Wald protocol to verify the fulfillment of the synthesis property. To this aim, the original high spectral resolution images are used as references, while the pansharpening algorithms are applied to the images after reducing their spatial resolution. Due to the availability of an objective final product, the indexes quantifying the similarity of the fused and the reference products can be employed. Typically, indexes taking into account both radiometric and spectral distortions are preferable in this phase. On the contrary, the full scale validation procedure avoids to reduce the resolution of input images. Ac-

cordingly, since no reference image can be considered, appropriate indexes (without reference) have to be designed and exploited to verify the accuracy of the sharpened product.

In Chapter 3 a comparison among several pansharpening algorithms presented in the literature has been performed. The different behaviors of the algorithms varying the validation procedures, scenario and satellite for acquiring the scene have been pointed out. The peculiarities of the two classes of pansharpening algorithms have been evidenced during the assessment phase. Specifically, the appeal, from a visual point of view, of component substitution methods has been shown by highlighting the absence of aliasing artifacts. The latter property, together with the robustness of these methods against the errors induced by the misregistration of the available images, has supported their large use also featured by a low computational burden. On the other side, the best overall performances are often achieved by multiresolution analysis approaches, which are characterized by a more exact reproduction of the spectral characteristics. Moreover, they can be easily designed to match the multispectral sensor properties in the phase of spatial detail extraction at the aim of decreasing the spatial distortion. Furthermore, this class of algorithms is currently drawing increasing attention due to their temporal coherence, useful when dealing with multiplatform data. The experiments presented in this chapter have been performed by a MATLAB implementation of the algorithms, quality indexes and validation procedures. We have chosen to make the developed MATLAB Toolbox available to the Community in order to allow a fair and easy comparison of some of the most widely used state-of-art algorithms. We also hope that this Toolbox can foster the development (and validation through a benchmark with established algorithms) of novel techniques for pansharpening.

In Chapter 4 we have shown that, considering the physics of acquisition systems, the injection of the high spatial resolution details of the panchromatic image into low resolution multispectral images can be accurately achieved by imposing the equality between the modulation transfer functions of the panchromatic

image and the estimated high spatial resolution multispectral image. This approach constitutes a step towards the construction of a synthetic multispectral image with the same spatial resolution of the panchromatic image fulfilling the Wald protocol. The method has been applied to pansharpening methods both based on component substitution and on multiresolution analysis. The experimental analysis confirmed the capability of the proposed high pass modulation details injection paradigm to improve the quality of the fused product with respect to the error based injection schemes.

A pansharpening algorithm can be often divided into two main phases: The extraction details from the panchromatic image and their injection into the image with higher spectral resolution. The former is often carried out by filters matching the modulation transfer function of the multispectral sensor in order to maximize the amount of extracted details. Gaussian filters matched with the sensor modulation transfer function represent the state-of-art for this step. In this case, *a priori* knowledge on the sensor characteristics are exploited to specify the filter. The unavailability (or the inaccuracy) of this information results in a strong limitation for the above-mentioned approach. In addition, being filters matched with the modulation transfer function, they represent an approximation of the real (and unknown) response of the acquisition devices, and, thus, the extracted details can be incorrect due to this residual mismatch. To overcome this problem, in Chapter 5 we have proposed a procedure for estimating the filter that models the blur between the multispectral and the panchromatic image by only using the available images and no additional information. The proposed technique has been compared to state-of-art pansharpening techniques extracting the details through filters. The experimental results are carried out by exploiting the two different validation procedures: at reduced scale and at full image scale. The former validation underlines the capability of the proposed method of correctly approximating the unknown blur filter and its robustness with respect to the tuning of its free parameters. The second protocol points out the advantages of the proposed

method with respect to the state-of-art pansharpening techniques both from a numerical and visual analysis.

Finally, in Chapter 6 we have presented a study based on the fusion of hyperspectral data and panchromatic images. Firstly, a critical comparison of classical approaches belonging to the component substitution and multiresolution analysis families has been performed. In greater details, the fused approaches have been applied to two different datasets. The first experiment has been carried out by exploiting images acquired by the same platform. This has permitted to mitigate co-registration effects in order to focus the analysis only on the distortions produced by the fusion itself. On the other hand, the second experiment has been performed by fusing products acquired by different platforms. The spectral angle mapper and the spatial distortion index have been considered to quantitatively evaluate the quality of the enhanced images, while the qualitative analysis has been carried out by visual inspecting the final results.

Furthermore, in Chapter 6, a novel hybrid pansharpening approach based on dimensionality reduction and multiresolution analysis methodologies has been proposed for hyperspectral pansharpening. Specifically, we have proposed the use of the non-linear principal component analysis (exploiting auto-associative neural networks) for representing the hyperspectral data into a lower non-linear feature space; several filters for extracting details and injection strategies have been considered instead of directly substituting the panchromatic image into the more correlated non-linear components. The proposed approach has been applied to both synthetic and real images. It has been demonstrated both the superiority of the non-linear approach with respect to the linear principal component analysis and that the hybrid approaches, based on non-linear principal component analysis for dimensionality reduction purposes, represent a good tradeoff between the reduction of the spectral distortion and the computation burden for hyperspectral sharpening.

The main research topics arising from this thesis will focus on:  
 1) The extension of the Toolbox with the introduction of different

pansharpening methodologies, such as, the ones based on Bayesian approaches, Compressive Sensing and Total Variation techniques; 2) The reformulation of the semiblind deconvolution problem to take into account of the possible filter diversities among spectral bands helpful, when we deal with old sensors as QuickBird or, in future applications, for hyperspectral sensors since they could be characterized by strongly band-dependent modulation transfer functions because of a wider spectral range covered; 3) Characterization of the non-linear principal component analysis inside the component substitution family in order to point out the advantages and drawbacks even in typical cases of fusing panchromatic and multispectral data.



# Bibliography

- [1] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, “Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics,” *IEEE Trans. Geosci. and Remote Sens.*, vol. 46, no. 5, pp. 1301–1312, May 2008.
- [2] C. Souza Jr, L. Firestone, L. M. Silva, and D. Roberts, “Mapping forest degradation in the eastern Amazon from SPOT 4 through spectral mixture models,” *Remote Sens. Environ.*, vol. 87, no. 4, pp. 494–506, Nov. 2003.
- [3] A. Mohammadzadeh, A. Tavakoli, V. Zoj, and J. Mohammad, “Road extraction based on fuzzy logic and mathematical morphology from pan-sharpened IKONOS images,” *Photogramm. Rec.*, vol. 21, no. 113, pp. 44–60, Feb. 2006.
- [4] F. Laporterie-Déjean, H. de Boissezon, G. Flouzat, and M.-J. Lefèvre-Fonollosa, “Thematic and statistical evaluations of five panchromatic/multispectral fusion methods on simulated PLEIADES-HR images,” *Inform. Fusion*, vol. 6, no. 3, pp. 193–212, Sep. 2005.
- [5] I. Amro, J. Mateos, M. Vega, R. Molina, and A. Katsaggelos, “A survey of classical methods and new trends in pansharpening of multispectral images,” *EURASIP Journal on Advances in Signal Processing*, no. 79, pp. 1–22, Sep. 2011.
- [6] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, “Twenty-five years of pansharpening: A critical review and

- new developments,” in *Signal and Image Processing for Remote Sensing, 2nd Ed.*, C. Chen, Ed. CRC Press, 2012, pp. 533–548.
- [7] W. Carper, T. Lillesand, and R. Kiefer, “The use of Intensity-Hue-Saturation transformations for merging SPOT panchromatic and multispectral image data,” *Photogramm. Eng. Remote Sens.*, vol. 56, no. 3, pp. 459–467, Apr. 1990.
- [8] P. Chavez, S. Sides, and J. Anderson, “Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic,” *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 295–303, Mar. 1991.
- [9] P. Chavez and A. Kwarteng, “Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis,” *Photogramm. Eng. Remote Sens.*, vol. 55, no. 3, pp. 339–348, Mar. 1989.
- [10] V. Shah, N. Younan, and R. King, “An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets,” *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1323–1335, May 2008.
- [11] C. Laben and B. Brower, “Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening, Eastman Kodak Company,” Patent US 6 011 875, 2000.
- [12] R. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing, 3rd Ed.* Elsevier, 2007.
- [13] P. Burt and E. Adelson, “The Laplacian pyramid as a compact image code,” *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, Apr. 1983.
- [14] K. Amolins, Y. Zhang, and P. Dare, “Wavelet based image fusion techniques - An introduction, review and comparison,” *ISPRS Journal Photogramm.*, vol. 62, no. 4, pp. 249–263, Sep. 2007.



- [15] J. C. Price, "Combining panchromatic and multispectral imagery from dual resolution satellite instruments," *Remote Sens. Environ.*, vol. 21, no. 2, pp. 119–128, Mar. 1987.
- [16] D. Fasbender, "Bayesian data fusion for adaptable image pan-sharpening," *IEEE Trans. Geosci. and Remote Sens.*, vol. 46, no. 6, pp. 1847–1857, Jun. 2008.
- [17] F. Palsson, J. Sveinsson, and M. Ulfarsson, "A new pansharpening algorithm based on Total Variation," *IEEE Geosci. and Remote Sens. Letters*, vol. 11, no. 1, pp. 318–322, Jan. 2014.
- [18] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [19] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [20] S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. and Remote Sens.*, vol. 49, no. 2, pp. 738–746, Feb. 2011.
- [21] X. Zhu and R. Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Trans. Geosci. and Remote Sens.*, vol. 51, no. 5, pp. 2827–2836, May 2013.
- [22] S. Park, M. Park, and M. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.
- [23] S. Li, H. Yin, and L. Fang, "Remote sensing image fusion via sparse representations over learned dictionaries," *IEEE Trans. Geosci. and Remote Sens.*, vol. 51, no. 9, pp. 4779–4789, Sep. 2013.

- [24] S. Yang, F. Sun, M. Wang, Z. Liu, and L. Jiao, "Novel super resolution restoration of remote sensing images based on compressive sensing and example patches-aided dictionary learning," in *International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping (M2RSM)*, 2011, pp. 1–6.
- [25] J. Pan, Z. and Yu, H. Huang, S. Hu, A. Zhang, H. Ma, and W. Sun, "Super-resolution based on compressive sensing and structural self-similarity for remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4864–4876, Sep. 2013.
- [26] A. Garzelli, L. Capobianco, L. Alparone, B. Aiazzi, and S. Baronti, "Hyperspectral pansharpening based on modulation of pixel spectra," in *2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2010*, 2010, pp. 1–4.
- [27] G. Licciardi, M. M. Khan, J. Chanussot, A. Montanvert, L. Condat, and C. Jutten, "Fusion of hyperspectral and panchromatic images using multiresolution analysis and nonlinear PCA band reduction," *EURASIP Journal on Advances in Signal Processing*, no. 1, pp. 1–17, Sep. 2012.
- [28] M. Ehlers, S. Klonus, P. J. Astrand, and P. Rosso, "Multi-sensor image fusion for pansharpening in remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 1, no. 1, pp. 25–45, Feb. 2010.
- [29] M. A. Kramer, "Nonlinear principal component analysis using autoassociative neural networks," *AIChE J.*, vol. 37, pp. 233–243, Feb. 1991.
- [30] M. M. Khan, J. Chanussot, L. Condat, and A. Montanvert, "Indusion: Fusion of multispectral and panchromatic images using the induction scaling technique," *IEEE Geosci. and Remote Sens. Letters*, vol. 5, no. 1, pp. 98–102, Jan. 2008.

- [31] M. M. Khan, J. Chanussot, and L. Alparone, "Pansharpening of hyperspectral images using spatial distortion optimization," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, 2009, pp. 2853–2856.
- [32] M. Moeller, T. Wittman, and A. Bertozzi, "A variational approach to hyperspectral image fusion," in *SPIE Defense, Security, and Sensing*, 2009, pp. 73 341E–73 341E.
- [33] G. A. Licciardi, A. Villa, M. M. Khan, and J. Chanussot, "Image fusion and spectral unmixing of hyperspectral images for spatial improvement of classification maps," in *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, 2012, pp. 7290–7293.
- [34] R. C. Hardie, M. T. Eismann, and G. L. Wilson, "MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1174–1184, Sep. 2004.
- [35] Y. Zhang, S. De Backer, and P. Scheunders, "Noise-resistant wavelet-based bayesian fusion of multispectral and hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3834–3843, Nov. 2009.
- [36] L. Wald, T. Ranchin, and Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogram. Eng. Remote Sensing*, vol. 63, no. 6, pp. 691–699, Jun. 1997.
- [37] Q. Du, N. Younan, R. King, and V. Shah, "On the performance evaluation of pan-sharpening techniques," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 518–522, Oct. 2007.
- [38] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

- [39] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, "Multispectral and panchromatic data fusion assessment without reference," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200, Feb. 2008.
- [40] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. of International Conference on Image Processing (ICIP)*, vol. 2, 2003, pp. III – 173–1585.
- [41] Z. Wang and A. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.
- [42] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.
- [43] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS+Pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.
- [44] T. Tu, S.-C. Su, H.-S. Shyu, and P.-S. Huang, "A new look at IHS-like image fusion methods," *Inform. Fusion*, vol. 2, no. 3, pp. 177–186, Sep. 2001.
- [45] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 295–309, Jan. 2011.
- [46] T. Tu, P. Huang, C. Hung, and C. Chang, "A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery," *IEEE Geosci. and Remote Sens. Letters*, vol. 1, no. 4, pp. 309–312, Oct. 2004.

- [47] A. Garzelli and F. Nencini, "Fusion of panchromatic and multispectral images by genetic algorithms," in *Proc. of IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2006, pp. 3810–3813.
- [48] A. Gillespie, A. Kahle, and R. Walker, "Color enhancement of highly correlated images. I. Decorrelation and HSI contrast stretches," *Remote Sens. Environ.*, vol. 20, no. 3, pp. 209–235, Dec. 1986.
- [49] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2005.
- [50] B. Aiazzi, F. Baronti, S. and Lotti, and M. Selva, "A comparison between global and context-adaptive pansharpening of multispectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 302–306, Apr. 2009.
- [51] A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE pan sharpening of very high resolution multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 228–236, Jan. 2008.
- [52] T. Ranchin and L. Wald, "Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation," *Photogram. Eng. Remote Sensing*, vol. 66, no. 1, pp. 49–61, Jan. 2000.
- [53] G. Vivone, R. Restaino, M. Dalla Mura, G. Licciardi, and J. Chanussot, "Contrast and error-based fusion schemes for multispectral image pansharpening," *IEEE Geosci. and Remote Sens. Letters*, vol. 11, no. 5, pp. 930–934, May 2014.
- [54] B. Aiazzi, L. Alparone, S. Baronti, I. Pippi, and M. Selva, "Generalized Laplacian pyramid-based fusion of MS + P image data with spectral distortion minimisation," *ISPRS Internat. Archives Photogramm. Remote Sensing*, vol. 34, no. 3A-W3, pp. 3–6, 2002.

- [55] J. Liu, "Smoothing filter based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, 2000.
- [56] L. Alparone, B. Aiazzi, S. Baronti, and A. Garzelli, "Sharpening of very high resolution images with spectral distortion minimization," in *Proc. of IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 1, 2003, pp. 458–460.
- [57] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and Pan imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.
- [58] J. Lee and C. Lee, "Fast and efficient panchromatic sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 155–163, Jan. 2010.
- [59] N. Garguet-Duport, J. Girel, J.-M. Chassery, and G. Patou, "The use of multiresolution analysis and wavelets transform for merging SPOT panchromatic and multispectral image data," *Photogram. Eng. Remote Sensing*, vol. 62, no. 9, pp. 1057–1066, Sep. 1996.
- [60] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1204–1211, May 1999.
- [61] M. González-Audcana, X. Otazu, O. Fors, and A. Seco, "Comparison between Mallat's and the 'à trous' discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images," *Int. J. Remote Sens.*, vol. 26, no. 3, pp. 595–614, Feb. 2005.
- [62] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Prentice Hall, 1995.

- [63] J.-L. Starck, J. Fadili, and F. Murtagh, "The undecimated wavelet decomposition and its reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 297–309, Feb. 2007.
- [64] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, 2nd Ed. Wellesley Cambridge Press, 1996.
- [65] X. Otazu, M. Gonzalez-Audicana, O. Fors, and J. Nunez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, Oct. 2005.
- [66] R. Yuhas, A. Goetz, and J. Boardman, "Discrimination among semi-arid landscape endmembers using the Spectral Angle Mapper (SAM) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*, 1992, pp. 147–149.
- [67] Z. Wang and A. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
- [68] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.
- [69] A. Garzelli and F. Nencini, "Hypercomplex quality assessment of multi/hyperspectral images," *IEEE Geosci. and Remote Sens. Letters*, vol. 6, no. 4, pp. 662–665, Oct. 2009.
- [70] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, Oct. 2002.

- [71] S. Baronti, B. Aiazzi, M. Selva, A. Garzelli, and L. Alparone, "A theoretical analysis of the effects of aliasing and misregistration on pansharpened imagery," *J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 446–453, Jun. 2011.
- [72] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "Advantages of Laplacian pyramids over "à trous" wavelet transforms for pansharpening of multispectral images," in *SPIE Remote Sensing*. International Society for Optics and Photonics, 2012, pp. 853 704–853 704.
- [73] A. Toet, "Image fusion by a ratio of low pass pyramid," *Pattern Recogn. Lett.*, vol. 9, pp. 245–253, May 1989.
- [74] A. Michelson, *Studies in Optics*. University of Chicago Press, 1927.
- [75] E. Peli, "Contrast in complex images," *J. Opt. Soc. Amer. A*, vol. 7, no. 10, pp. 2032–2039, Oct. 1990.
- [76] A. Garzelli and F. Nencini, "Interband structure modeling for pan-sharpening of very high-resolution multispectral images," *Inform. Fusion*, vol. 6, no. 3, pp. 213–224, Sep. 2005.
- [77] D. Kundur, "Blind image deconvolution," *IEEE Signal Processing Mag.*, vol. 13, no. 3, pp. 43–64, May 1996.
- [78] P. Campisi and K. Egiazarian, *Blind image deconvolution: Theory and applications*. CRC Press, 2007.
- [79] A. Levin, Y. Weiss, F. Durand, and W. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, no. 2, 2009, pp. 1964–1971.
- [80] A. K. Jain, *Fundamentals of digital image processing*. Prentice-Hall Englewood Cliffs, 1989.
- [81] A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*. Winston and Sons, 1977.



- [82] S. Reeves, "Fast image restoration without boundary artifacts," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1448–1453, Oct. 2005.
- [83] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Control Signals*, vol. 2, no. 4, pp. 303–314, 1989.
- [84] C. Bishop, "Neural networks for pattern recognition," *Oxford Univ. Press, London, U.K.*, 1995.
- [85] G. W. Cottrell, P. Munro, and D. Zipser, "Learning internal representations from gray-scale images: An example of extensional programming," *Proc. Conf. of the Cognitive Sci. SOC.*, 1987.
- [86] G. Licciardi, P. Marpu, J. Benediktsson, and J. Chanussot, "Extended morphological profiles using auto-associative neural networks for hyperspectral data classification," in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2011 3rd Workshop on*, 2011, pp. 1–4.
- [87] G. A. Licciardi and F. D. Frate, "Pixel unmixing in hyperspectral data by means of neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4163–4172, Oct. 2011.