UNIVERSIDAD DE GRANADA

E.T.S. DE INGENIERÍAS INFORMÁTICA Y TELECOMUNICACIÓN

Universidad de Granada

Departamento de Ciencias de la Computación
e Inteligencia Artificial

FUSIÓN DE IMÁGENES MULTIESPECTRALES MEDIANTE MÉTODOS MULTIESCALA Y SUPER-RESOLUCIÓN

TESIS DOCTORAL

Israa Amro

Granada, enero de 2011

Editor: Editorial de la Universidad de Granada
Autor: Israa Amro
D.L.: GR 1979-2011
ISBN: 978-84-694-1317-3

Universidad de Granada

FUSIÓN DE IMÁGENES MULTIESPECTRALES MEDIANTE MÉTODOS MULTIESCALA Y SUPER-RESOLUCIÓN

MEMORIA QUE PRESENTA
ISRAA AMRO
PARA OPTAR AL GRADO DE DOCTOR EN INFORMÁTICA
ENERO 2011

DIRECTOR
JAVIER MATEOS DELGADO

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN
E INTELIGENCIA ARTIFICIAL

E.T.S. INGENIERÍAS INFORMÁTICA Y TELECOMUNICACIÓN   UNIVERSIDAD DE GRANADA

La memoria titulada **Fusión de imágenes multiespectrales mediante métodos multiescala y super-resolución**, que presenta Dña. Israa Amro para optar al grado de DOCTOR, ha sido realizada en el Departamento de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada bajo la dirección del Doctor D. Javier Mateos Delgado.

Granada, enero de 2011

La doctoranda	El director
Israa Amro	Javier Mateos Delgado

ACKNOWLEDGEMENT

بِسمِ اللّٰهِ الرَّحمٰنِ الرَّحِيمِ

I would like to thank all people who have helped and inspired me during my doctoral study. First and foremost I offer my sincerest gratitude to my supervisor, Prof. Javier Mateos, who has supported me throughout my thesis with his patience and knowledge. I attribute the level of my PhD degree to his encouragement and effort and without him this thesis, too, would not have been completed or written. In addition, he was always accessible and willing to help his students with their research. As a result, research life became smooth and rewarding for me. One simply could not wish for a better or friendlier supervisor. I was delighted to interact with Profs. Rafael Molina and Aggelos Katsaggelos by attending their classes and having them as my tutors. Their insight to Bayesian methods is second to none. Besides, they set an example of a world-class researchers for their rigor and passion on research. My sincere thanks also goes to Prof. Miguel Vega, for his detailed and constructive comments, and for his important support throughout this thesis. I warmly thank Prof. Nicolas Pérez de la Blanca for his valuable advice and friendly help, who gave me the opportunity to know this University from the first day I talked to him from Palestine. Where would I be without my family? My parents deserve special mention for their inseparable support and prayers. My Father, in the first place is the person who put the fundament of my learning character, showing me the joy of intellectual pursuit ever since I was a child. My Mother, is the one who sincerely raised me with her caring and gently love. All my brothers, brother in law, sister, and sisters in law, thanks for being supportive and caring siblings. Words fail me to express my appreciation to my fiancé whose dedication, love and persistent confidence in me, has taken the load off my shoulder. All my Spanish friends, especially Miriam Lopez and all Palestinians in Granada, thanks :) for everything. I would like to thank everybody who was important to the successful realization of thesis, as well as expressing my apology that I could not mention personally one by one. Last but not least, thanks be to Allah for my life through all tests in the past three years. You have made my life more bountiful. May your name be exalted, honored, and glorified.

FUSIÓN DE IMÁGENES MULTIESPECTRALES MEDIANTE MÉTODOS MULTIESCALA Y SUPER-RESOLUCIÓN

ISRAA AMRO

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

General Introduction

1.1 Overview

Nowadays, a huge quantity of satellite images is available from many different Earth observation platforms. Earth observation satellites, with physical and technological constraints, provide two different classes of images: a panchromatic image (PAN) with high spatial and low spectral resolutions, allowing to distinguish features spatially but not spectrally, and a multispectral image (MS) with high spectral and low spatial resolutions, that is, with the number of spectral bands needed to distinguish features spectrally but not spatially. The fusion of a high spatial resolution PAN image and a high spectral resolution MS image is an important issue for many remote sensing and mapping applications. For many applications, that combination of data from multiple sensors provides more comprehensive information [1]. The pixel level fusion technique used to increase the spatial resolution of the MS image while simultaneously retaining its spectral information is called pansharpening [2]. Pansharpening provides feature enhancement and improved classification as information from two different images is used effectively [3].

The benefit of merged images has been demonstrated in many practical applications, especially for vegetation, land use, precision farming, urban studies [4], feature detection, change monitoring and land cover classification, which often demand the highest spatial and spectral resolution for the best accomplishment of their objectives [5]. Several works have demonstrated the usefulness of fused products offering high spectral and spatial resolutions at the same time in various environmental applications [6, 7]. Pansharpened products are becoming very popular, (see for example Google Earth [8]), and data providers are offering higher and higher amounts of them at lower and lower costs [7]. In this disertation we will study classical and state-of-the-art pansharpening techniques and propose new techniques based on multiscale methods and super-resolution.

1.2 Background

1.2.1 Remote Sensing and Sensors Images

Remote sensing is defined, for our purposes, as the measurement of object properties on the Earth's surface using data acquired from aircraft and satellites [9]. It is therefore an attempt to measure something at a distance, rather than in situ. Since we are not in direct contact with the object of interest, we must rely on propagated signals of some sort, for example optical, acoustical, or microwave. Remote sensing data consists of discrete, (point measurements or a profile along a flight path), measurements over a two-dimensional spatial grid, i.e., images. Remote sensing systems, particularly those deployed on satellites, provide a repetitive and consistent view of the Earth that is invaluable to monitoring short-term and long-term changes and the impact of human activities. Some of the important applications of remote sensing technology are [9]:

1. Environmental assessment and monitoring (urban growth, hazardous waste).
2. Global change detection and monitoring (atmospheric ozone depletion, deforestation, global warming).
3. Agriculture (crop condition, yield prediction, soil erosion).
4. Nonrenewable resource exploration (minerals, oil, natural gas).
5. Renewable natural resources (wetlands, soils, forests, oceans).
6. Meteorology (atmosphere dynamics, weather prediction).
7. Mapping (topography, land use, civil engineering)
8. Military surveillance and reconnaissance (strategic policy, tactical assessment).
9. News media (illustrations, analysis).

To meet the needs of different data users, many remote sensing systems have been developed, offering a wide range of spatial, spectral, and temporal parameters. Some users may require frequent, repetitive coverage with relatively low spatial resolution (meteorology). Others may desire the highest possible spatial resolution with repeat coverage only infrequently (mapping); while some users need both high spatial resolution and frequent coverage, plus rapid image delivery (military surveillance). Properly calibrated remote sensing data can be used to initialize and validate large computer models, such as Global Climate Models (GCMs), that attempt to simulate and predict the Earth's environment. In this case, high spatial resolution may be undesirable because computational requirements, but accurate and consistent sensor calibration over time and space essential [9]. Examples of remote sensing satellite series in operations are: Landsat [10], SPOT [11], GMS, etc. Remote sensing images are recorded in digital forms and then processed by the computers to produce images for interpretation purposes. Images are available in two forms: photographic film form and digital form. Digital image consists of discrete picture elements called pixels. Associated with each pixel is a number represented as DN (Digital Number), that depicts the average radiance of relatively small area within a scene. The range of DN values being normally 0 to 255. The size of this area effects the reproduction of details within the scene. As the pixel size is reduced more scene detail is preserved in digital representation [12].

The modern era of Earth remote sensing from satellites began when the Landsat Multispectral Scanner System (MSS) provided, for the first time, in 1972 a consistent set of synoptic, high resolution Earth images to the world scientific community. The characteristics of this new sensor were multiple spectral bands, with reasonably high spatial resolution for the time (80m), large area (185km by 185km), and repeating (every 18 days) coverage. Moreover, the MSS provided general purpose satellite image data directly in digital form [9]. An excellent history and discussion of the motivations for the Landsat program and data processing are provided in [13].

Since 1972, there have been four additional MSS systems, two Thematic Mapper (TM) systems, and the Enhanced Thematic Mapper Plus (ETM+) in the Landsat series. There have also been five higher resolution French SPOT systems [11], several lower resolution AVHRR and GOES systems, and NASA's [14] sensor suites on the Earth Observing System (EOS) [15] Terra and Aqua satellites, as well as a wide variety of other multispectral sensors on aircraft and satellites. The Advanced Visible/InfraRed Imaging Spectrometer (AVIRIS) [16] and the HyMap are airborne sensors that produce hundreds of images of the same area on the ground in spectral bands about 10nm wide over the solar reflective portion of the spectrum from 400 to 2400 nm. The Hyperion was on NASA's Earth- Observing-1 (EO-1) satellite as the first civilian hyper-spectral satellite system. Although it has relatively fewer spectral bands, the European Space Agency's MEdium Resolution Imaging Spectrometer (MERIS) [17] is also an imaging spectrometer. The separation of spectral bands in these systems is achieved with a continuously dispersive optical element, such as a grating or prism. The MODerate Imaging Spectroradiometer (MODIS) [18], a discrete filter-based system, on Terra and Aqua provides images in 36 spectral bands over the range 0.4 to 14 μm. Such sensors have provided large improvements in the quantity and quality of information that can be gathered about the Earth's surface and near environment. Commercial development in the late 1990s of high performance orbital sensors, with resolutions of 0.5 to 1m in panchromatic mode and 2.5 to 4m in multispectral mode, has opened new commercial markets and public service opportunities for satellite imagery, such as real estate marketing, design of cellular telephone and wireless Personal Communications System (PCS) coverage areas (which depend on topography and building structures), urban and transportation planning, and natural and man-made disaster mapping and management. These systems also have value for military intelligence and environmental remote sensing applications. The first generation includes the IKONOS [19], QuickBird [20], and OrbView [21] sensors, and further development is expected, particularly toward higher resolution capabilities, subject to legal regulations and security concerns of various countries [9]. A depiction of some of these optical remote sensing systems in a performance space defined by two key sensor parameters, the number of spectral bands and the Ground-projected Sample Interval (GSI), is shown in Figure 1.1 [22].

In this work, the experiments will focus in different sensors images such as Landsat 7 [23], SPOT 5 [11] and QuickBird [20] sensor images. In 1999, the US Congress authorized the procurement, launch and operation of a new Landsat satellite. This new system, Landsat 7, it is the latest in a series of Earth observation satellites. Landsat 7 have a unique and essential role in the realm of Earth observing satellites in orbit. No other system will match Landsat's combination of synoptic coverage, high spatial resolution, spectral range and radiometric calibration. The primary new features on Landsat 7 are:

1. A panchromatic band with 15m spatial resolution.

2. On board, full aperture, 5% absolute radiometric calibration.

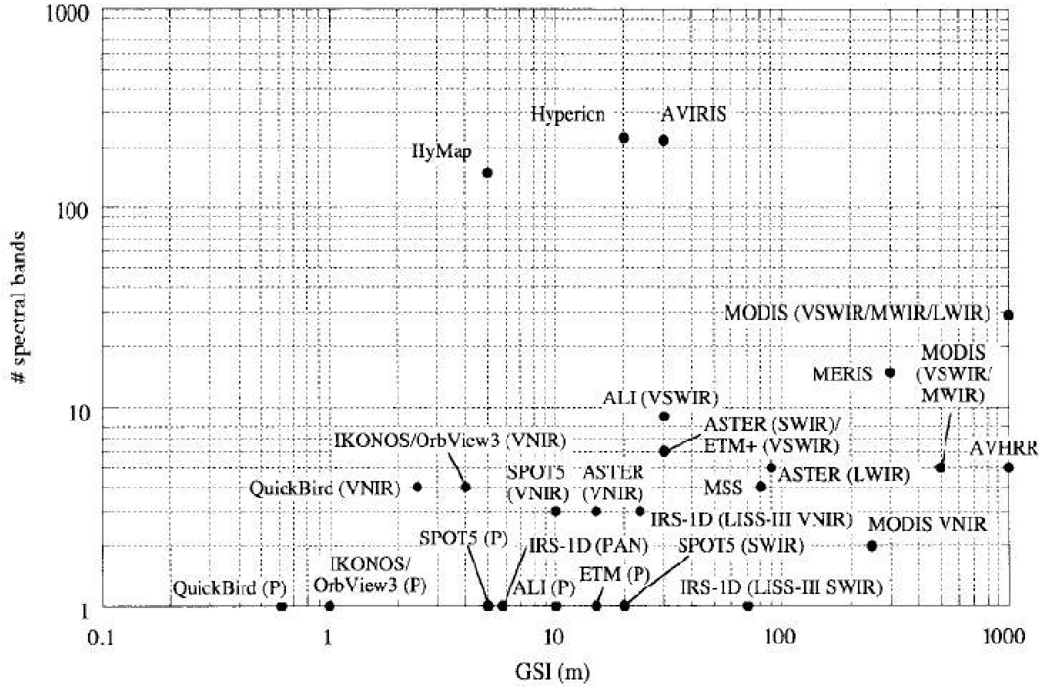


Figure 1.1: *A plot of some remote-sensing systems in a two-dimensional parameter space. V = Visible, NIR = Near InfraRed, LWIR = Long Wave IR, MWIR = Mid Wave IR, SWIR = Short Wave IR, and P = Panchromatic.*

Band	**Spectral Range** (microns)	**Spectral Response**	**Ground Resolution** (m)
1	0.45 to 0.515	Blue-Green	30
2	0.525 to 0.605	Green	30
3	0.63 to 0.690	Red	30
4	0.75 to 0.90	Near IR	30
5	1.55 to 1.75	Mid-IR	30
6	10.40 to 12.5	Thermal-IR	60
7	2.09 to 2.35	Mid-IR	30
PAN	0.52 to 0.90	visible	15

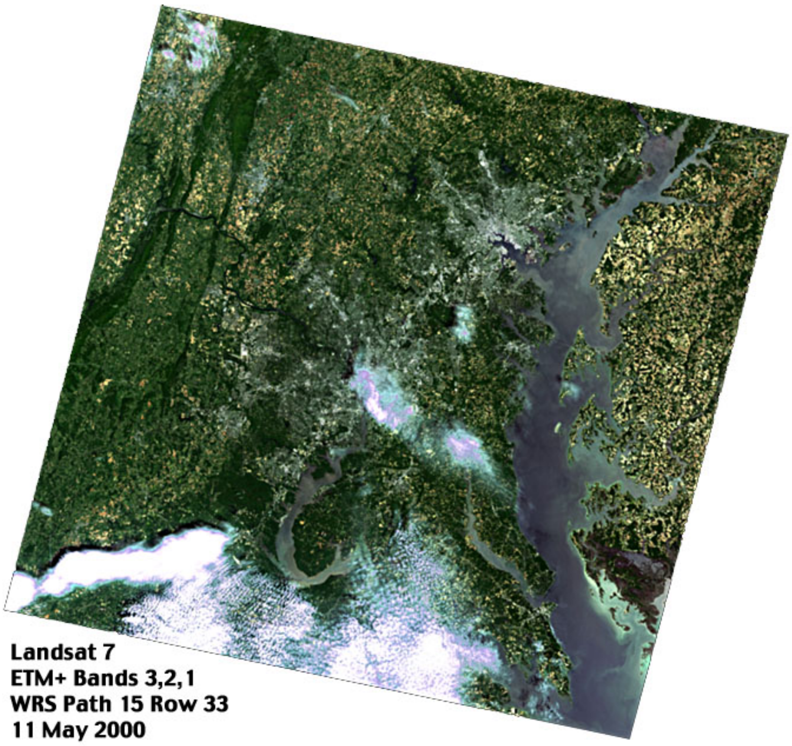
Table 1.1: Landsat 7 Bands

3. A thermal InfraRed (IR) channel with 60m spatial resolution.

The Landsat 7 satellite uses an instrument that collects seven images at once. Each image shows a specific section of the electromagnetic spectrum, called a band. Landsat 7 has seven different bands. The seven bands of Landsat 7 [14], and their characteristics described in Table 1.1.

Figure 1.2 shows two Landsat 7 pictures of the exact same location. The picture on the above is a “true color” image, this means that the picture shows objects in the same colors that our eyes would normally see. The picture on the below is a “false color” image, this means that the colors have been assigned to three different wavelengths that our eyes might not normally see.

The SPOT 5 Earth observation satellite was successfully placed into orbit by an Ariane 4 from the Guiana Space Centre in Kourou during the night of 3 to 4 May 2002. Compared to its predecessors, SPOT 5 offers greatly enhanced capabilities, which provide additional cost-effective imaging solutions. Thanks to SPOT-5’s improved 5 m and 2.5 m resolution and wide imaging swath, which covers 60 x 60 km or 60 km x 120 km in twin-instrument mode, the SPOT 5 satellite provides an ideal balance between high resolution and wide-area coverage. The coverage offered by SPOT 5 is a key asset for applications such as medium-scale mapping (at 1:25 000 and 1:10 000 locally), urban and rural planning, oil and gas exploration, and natural disaster management. SPOT 5’s other key feature is the unprecedented acquisition capability of the on-board HRS stereo viewing instrument, which can cover vast areas in a single pass.



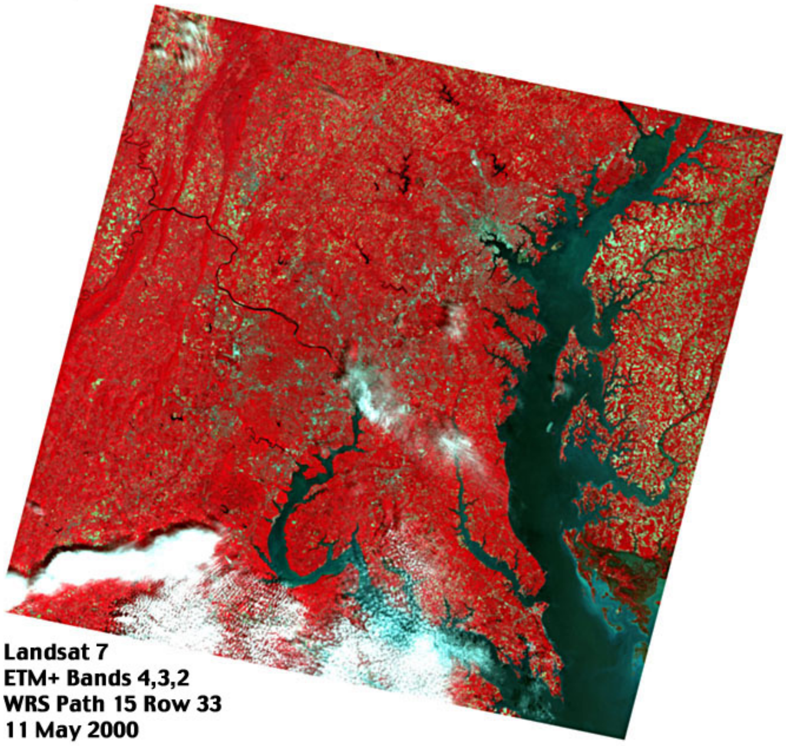


Figure 1.2: *”True Color” and “False Color” Landsat 7 Images*

Stereo pair imagery is vital for applications that call for 3D terrain modeling and computer environments, such as flight simulator databases, pipeline corridors, and mobile phone network planning [11].

For SPOT5 imaging dataset, the MS image pixel resolution of 10 m, while the PAN image pixel resolution of 5 m and 2.5 m. The MS image consists of four bands from the visible and infrared region corresponding to green (0.50-0.59 μm), red (0.61-0.68 μm), Near IR (0.78-0.89 μm), Mid IR(1.58-1.75 μm), while the PAN image consists of a single band covering the visible and NIR (0.48-0.71 μm).

QuickBird is a high resolution satellite owned and operated by DigitalGlobe. Using a state-of-the-art BGIS 2000 sensor, QuickBird collects image data to 0.61 m pixel resolution degree of detail. This satellite is an excellent source of environmental data useful for analyses of changes in land usage, agricultural and forest climates. QuickBird's imaging capabilities can be applied to a host of industries, including Oil and Gas Exploration and Production (E&P), Engineering and Construction and environmental studies [20].

For QuickBird imaging dataset, the MS image pixel resolution of 2.44 m, while the PAN image pixel resolution of 61 cm to 72 cm . The MS image consists of four bands from the visible and infrared region corresponding to blue (0.45-0.52 μm), green (0.52-0.60 μm), red (0.63-0.69 μm), near IR (0.76-0.90 μm), while the PAN image consists of a single band covering the all MS bands (0.45-0.90 μm).

1.2.2 Spatial and Spectral Resolution

The spatial resolution of a remote sensing imaging system is expressed as the area of the ground captured by one pixel and affects the reproduction of details within the scene. As the pixel size is reduced, more scene details are preserved in the digital representation [24]. The instantaneous field of view (IFOV) is the ground area sensed by the sensor at a given instant in time (see Figure 1.4). The spatial resolution depends on the IFOV. For a given number of pixels, the finer the IFOV is, the higher the spatial resolution. Spatial resolution is also viewed as the clarity of the high frequency detail information available in an image. Spatial resolution in remote sensing is usually expressed in meters or feet, which represents the length of the side of the area covered by a pixel. Figure 1.3 shows three images of the same ground area but with different spatial resolutions. The image at $5m$ depicted in Figure. 1.3a was captured by the SPOT 5 satellite while the other two images, at $10m$ and $20m$, are simulated from the first


(a) (b) (c)

Figure 1.3: *Images of the same area with different spatial resolutions (a) Spatial resolution =* $5m$*, (b) Spatial resolution =* $10m$*, (c) Spatial resolution =* $20m$*.*

image. As one can see from the images, the detail information becomes clearer as the spatial resolution increases from $20m$ to $5m$.

Spectral resolution is the width of the band, within the electromagnetic spectrum, that can be captured by a sensor which is capable of separating light reflected from the Earth into discrete spectral bands. The narrower the spectral bandwidth is, the higher the spectral resolution. If the platform captures a few spectral bands, typically 4 to 7, they are referred to as multispectral data, while if the number of spectral bands is measured in hundreds or thousands, they are referred to as hyperspectral (HS) data [25]. A color image is a very simple example of a multispectral image that contains three bands. Together with the MS or HS image, satellites usually provide a panchromatic image. This is an image that contains one wide band of reflectance data representative of a wide range of bands and wavelengths such as visible or thermal infrared, that is, it integrates the chromatic information therefore the name is “pan” chromatic. A PAN image of the visible bands is a combination of red, green and blue data into a single measure of reflectance. Modern multispectral scanners also generally include some radiation at slightly longer wavelengths than red light, called ”near infrared” radiation [26].

1.2.3 Image Fusion and Pansharpening

Remote sensor systems are designed within often competing constraints, among the most important being the trade-off between IFOV and signal-to-noise ratio (SNR). Since multispectral, and to a greater extent hyperspectral, sensors have reduced spectral bandwidths compared to

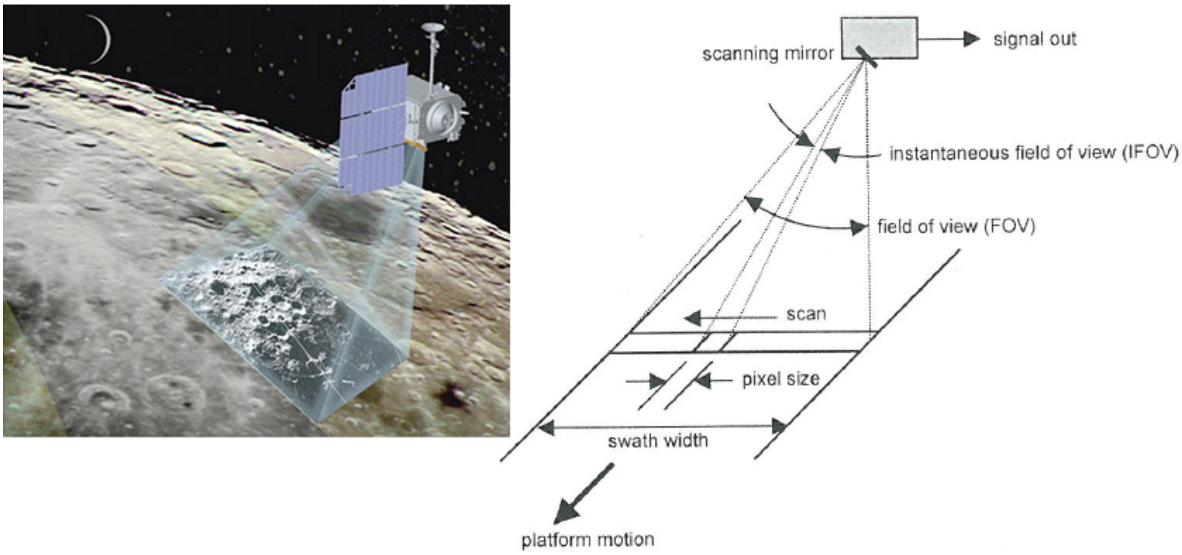


Figure 1.4: *Remote Sensing Capturing System*

panchromatic sensors, they typically have a larger IFOV in order to collect more photons and maintain image SNR. Many sensors such as SPOT, ETM+, IKONOS, OrbView, and QuickBird have a set of multispectral bands and a co-registered higher spatial resolution panchromatic band. With appropriate algorithms it is possible to combine these data and produce multispectral imagery with higherspatial resolution. This concept is known as multispectral or multisensor merging, fusion, or sharpening (of the lower-resolution image) [9].

Data fusion is a process dealing with data and information from multiple sources to achieve refined/improved information for decision making. A general definition of image fusion "is the combination of two or more different images to form a new image by using a certain algorithm" [2]. Image fusion is performed at three different processing levels according to the stage at which the fusion takes place (see Figure 1.5 for graphical illustration of those levels) [3]:

1. Pixel level: Fusion at the lowest processing level referring to the merging of measured physical parameters.

2. Feature Level: This requires the extraction of objects recognized in the various data sources, e.g., using segmentation procedures. Features correspond to characteristics extracted from the initial images which are depending on their environment such as extent, shape and neighborhood.

3. Decision (interpretation) Level: Represents a method that uses value-added data where

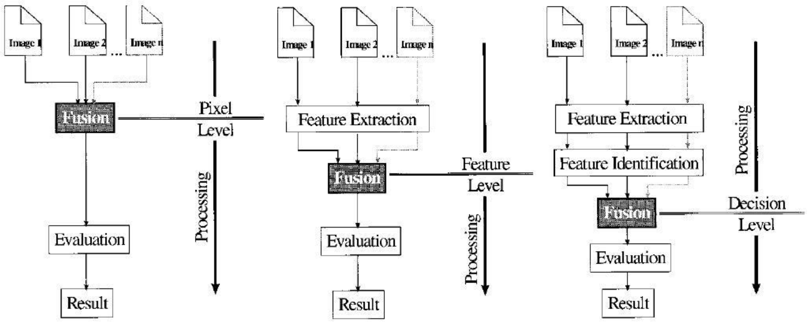


Figure 1.5: *Processing levels of image fusion*

the input images are processed individually for information extraction. The obtained information is then combined applying decision rules to reinforce common interpretation and resolve differences and furnish a better understanding of the observed objects

Pansharpening, can consequently be defined as, a pixel level fusion technique used to increase the spatial resolution of the MS image [3]. Pansharpening is shorthand for Panchromatic Sharpening meaning the use of a PAN (single band) image to sharpen an MS image. In this sense, to sharpen means to increase the spatial resolution of an MS image. Thus, pansharpening techniques increase the spatial resolution while simultaneously preserving the spectral information in the MS image, giving the best of the two worlds: high spectral resolution and high spatial resolution [25]. Some of the applications of pansharpening include improving geometric correction, enhancing certain features not visible in either of the single data alone, change detection using temporal data sets, and enhancing classification [27].

1.3 Objectives and Outline

The primary contribution of this thesis is the development of three new pansharpening methods that give us an estimation of the high spatial and spectral resolution MS image, that an ideal remote sensor will provide us.

The content of this dissertation is divided in the following chapters. In Chapter 2, we study the different classical and state-of-the-art methods proposed in the literature to tackle the pansharpening problem, giving a clear classification of the methods and a description of their main characteristics. The objective of this chapter is to know the advantages and drawbacks for each approach in order to propose, in the following chapters, methods that improve the quality of the pansharpened image. In Chapter 3 we will study the sources of degradation affecting to the ideal high spatial and spectral resolution multispectral image, and explain the low resolutiom multispectral and panchromatic image formation models. Once we know how these images are formed, we will discuss Bayesian Formulation in order to provide solutions to the remote sensing pansharpening problem, and the possible Bayesian solutions to perform the image reconstruction inference and the parameter estimation simultaneously. Chapter 4 explains briefly the contourlet transform and study different pansharpening methods that depend on both of wavelet and contourlet transforms, and casts the popular wavelets-based algorithm called WiSPeR [5] into the contourlets domain. Part of this chapter was published as a work in the "8th International Workshop on Information Optics (WIO'09)" [28] and as a chapter of the book Advances in Information Optics and Photonics [29]. Chapter 5 proposes a new pansharpening method that combines the super-resolution technique with non-subsampled contourlet transform in order to obtain a method that efficiently preserves the texture and contours information of the PAN image while improving all the bands of the image, even those that are not covered by the PAN image. Part of this chapter was published in the 2010 International Conference on Image Processing (ICIP 2010) [30] and it was selected as finalist of the Huawei Best Student Paper Award. In Chapter 6 we propose a Bayesian fusion method based on non-subsampled contourlet transform that comprises, as particular cases, popular contourlet domain methods such as, substitution, addition and some other mathematical models. Part of this chapter was published in the 2010 European Signal Processing Conference (EUSIPCO-2010) [31]. Finally, Chapter 7 concludes this dissertation and future research topics are outlined.

Chapter 2

Classical methods and new trends in pansharpening of multispectral images [1]

2.1 Introduction

During the past years an enormous amount of pansharpening techniques have been developed and, in order to choose the one that better serves to the user needs, there are some points, mentioned by Pohl [3], that have to be considered. In the first place, the objective or application of the pansharpened image can help in defining the necessary spectral and spatial resolution. For instance, some users may require frequent, repetitive coverage with relatively low spatial resolution (i.e. meteorology applications), others may desire the highest possible spatial resolution (i.e. mapping), while other users may need both high spatial resolution and frequent coverage, plus rapid image delivery (i.e. military surveillance).

Then, the data which are more useful to meet the needs of the pansharpening applications, like the sensor, the satellite coverage, atmospheric constraints such as cloud cover, sun angle, etc. have to be selected. We are mostly interested in sensors that can capture simultaneously a PAN channel with high spatial resolution and some MS channels with high spectral resolution like SPOT 5, Landsat 7 and QuickBird satellites. In some cases, PAN and MS images captured by different satellites sensors at different dates for the same scene can be used for some appli-

---

[1] Part of this chapter has been submitted to the IEEE Journal of Selected Topics in Signal Processing, 2010 [32].

cations [27], like in the case of fusing different MS SPOT 5 images captured at different times with one PAN IKONOS image [33], which can be considered as a multi-sensor, multi-temporal and multi-resolution pansharpening case.

We also have to take into account the need for data pre-processing, like registration, upsampling and histogram matching, as well as, the selection of a pansharpening technique that makes the combination of the data most successful. Finally, evaluation criteria are needed to specify which is the most successful pansharpening approach.

In this chapter we examine the classical and state-of-the-art pansharpening methods described in the literature giving a clear classification of the methods and a description of their main characteristics.

This chapter is organized as follows. In section 2.2 data pre-processing techniques are described. In section 2.3 a classification of the pansharpening methods is presented, with a description of the methods related to each category and of some examples. In section 2.4 we analyze how the quality of the pansharpened images can be assessed both visually and quantitatively and examine the different quality measures proposed for that purpose and finally section 2.5 concludes the chapter.

2.2 Pre-processing

Remote sensors acquire raw data which need to be processed in order to convert it to images. The grid of pixels that constitutes a digital image is determined by a combination of scanning in the cross-track direction (orthogonal to the motion of the sensor platform) and by the platform motion along the in-track direction. A pixel is created whenever the sensor system electronically samples the continuous data stream provided by the scanning [9]. The image data recorded by sensors and aircrafts can contain errors in geometry and measured brightness value of the pixels (which are referred to as radiometric errors) [34]. The relative motion of the platform, the non-idealities in the sensors themselves and the curvature of the Earth, can lead to geometric errors of varying degrees of severity. The radiometric errors can result from the instrumentation used to record the data, the wavelength dependence of solar radiation and the effect of the atmosphere. For many applications using these images, it is necessary to make corrections in geometry and brightness before the data are used. By using correction techniques (see [9, 34]), an image can be registered to a map coordinate system and therefore have its pixels addressable in terms of map coordinates rather than pixel and line numbers, a process often referred to as geocoding.

The Earth Observing System Data and Information System (EOSDIS) receives “raw” data from all spacecrafts and processes it to remove telemetry errors, eliminate communication artifacts, and create *Level 0* Standard Data Products that represent raw science data as measured by the instruments. Other levels of remote sensing data processing were defined in [35] by the NASA Earth Science program. In *Level 1A* the reconstructed, unprocessed instrument data at full resolution, time-referenced and annotated with ancillary information (including radiometric and geometric calibration coefficients and georeferencing parameters) are computed and appended, but not applied to *Level 0* data (i.e., *Level 0* can be fully recovered from *Level 1A*). Some instruments have *Level 1B* data products, where the resulting data from *Level 1A* are processed to sensor units. At *Level 2*, the geographical variables are derived (e.g. Ocean wave height, soil moisture, ice concentration) at the same resolution and location as *Level 1* data. *Level 3* maps the variables on uniform space-time grids usually with some completeness and consistency, and finally, *Level 4* gives the results from the analysis of the previous levels data. For many applications, *Level 1* data is the most fundamental data records with significant scientific utility, and it is the foundation upon which all subsequent data sets are produced.

Once the image data has already been processed in one of the standard levels previously described, and in order to apply pansharpening techniques, the images are pre-processed to accommodate the pansharpening algorithm requirements. This pre-processing may include registration, re-sampling and histogram matching of the MS and PAN images. Let us now study these processes in detail.

2.2.1 Image registration

Many applications of remote sensing image data require two or more scenes of the same geographical region, acquired at different dates or from different sensors, in order to be processed together. In this case, the role of image registration is to make the pixels in the two images precisely coincide to the same points on the ground [9]. Two images can be registered to each other by registering each to a map coordinate base separately, or one image can be chosen as a master to which the other is to be registered [34]. However, due to the different physical characteristics of the different sensors, the problem of registration is more complex than registration of images from the same type of sensors [36] and have also to face problems like features present in one image that might appear only partially in the other image or do not appear at all. Contrast reversal in some image regions, multiple intensity values in one image that need to be mapped to a single intensity value in the other or considerably dissimilar images of the same scene produced by the image sensor when configured with different imaging parameters are also problems to be solved by the registration techniques.

Many image registration methods have been proposed in the literature. They can be classified into two categories: area-based methods and feature-based methods. Examples of area-based methods, which deal with the images without attempting to detect common object, include Fourier methods, cross-correlation methods and mutual information methods [37]. Since the gray-level values of the images to be matched may be quite different and taking into account that the joint intensity probability is not always maximum when two images are spatially aligned for any two different modality of images, area-based techniques are not well adapted to the problem of multisensor image registration [36]. Feature-based methods, which extract and match the common structures (features) from two images, have been shown to be more suitable for this task. Example methods in this category include methods using spatial relations, those based on invariant descriptors, relaxation, and pyramidal and wavelet image decompositions, among others [37].

2.2.2 Image up-sampling and interpolation

When the registered remote sensing image is too coarse and does not meet the required resolution, up-sampling may be needed to obtain a higher resolution version of the image [38]. The up-sampling process may involve interpolation which allocates intensity values into a new generated pixel. The interpolation itself is usually performed via convolution of the image with an interpolation kernel [39]. In order to reduce the computational cost, preferably separable interpolants have been considered [37]. Many interpolants for various applications have been proposed in the literature [40]. A brief discussion of interpolation methods used for up-sampling purposes is provided in [37]. In [38] the authors describe methods that are suitable to interpolate remote sensing data and prevent the degradation of the images quality. These methods include *nearest neighbor interpolation* which only considers the closest pixel to the interpolated point, thus it requires the least processing time of all interpolation algorithms; *bilinear interpolation* which creates the new pixel in the target image from a weighted average of its 4 nearest neighboring pixels in the source image; *interpolation with smoothing filter* which produces a weighted average of the pixels contained in the area spanned by the filter mask. This process produces an image with smooth transitions in gray level; *interpolation with sharpening filter* where the sharpening filter is usually used in order to enhance details that have been blurred or to highlight fine details in an image. A sharpening filter produces a sharp image by enhancing the edges of the image. However, this filter has some undesirable side effects, causing aliasing in the output image; and *interpolation with unsharp masking* where unsharp masking is a process to sharpen images by subtracting a blurred version of an image from the image itself. Note that interpolation does not increase the high frequency detail information in the image but it is needed to match the number of pixels of images with different spatial resolutions.

2.2.3 Histogram matching

Some pansharpening algorithms assume that the spectral characteristics of the PAN image match those of each band of the MS image or match those of a transformed image based on the MS image. Unfortunately this is not usually the case [34] and those pansharpening methods are prone to produce spectral distortions.

Matching the histograms of the PAN image and MS bands will minimize the variation of the brightness value during the fusion process, which may help to reduce the spectral distortion in the pansharpened image. Furthermore, there are histogram matching techniques for general applications as the ones described, for instance in [34] and [39], that are used also in remote sensing. A technique for histogram matching in remote sensing fusion was presented in [41]. This technique minimizes the modification of the spectral information of the fused high resolution multispectral (HRMS) image with respect to the original low resolution multispectral (LRMS) image. This method modifies the value of the PAN image at each pixel (i,j) as

$$Stretched_{PAN}(i,j) = (PAN(i,j) - \mu_{PAN})\frac{\sigma_b}{\sigma_{PAN}} + \mu_b, \tag{2.1}$$

where μ_{PAN} and μ_b are the mean of the PAN and MS image band b, respectively, and σ_{PAN} and σ_b are the standard deviation of the PAN and MS image band b, respectively. This technique ensures that the mean and standard deviation of PAN image and MS bands are within the same range, thus reducing the chromatic difference between both images.

2.3 Pansharpening categories

Once the remote sensing images are pre-processed in order to satisfy the pansharpening method requirements, the pansharpening process is performed. The literature shows a large collection of these pansharpening methods developed over the last two decades as well as a large number of terms used to express the idea of image fusion. In 1980, Wong *et al.* [42] proposed a technique for the integration of Landsat MSS and Seasat synthetic aperture radar (SAR) images based on the modulation of the intensity of each pixel of the MSS channels with the value of the corresponding pixel of the SAR image, hence named Intensity Modulation (IM) integration method. Other scientists evaluated multi-sensor image data in the context of co-registered [43], resolution enhancement [44] or coincident [45] data analysis.

After the launch of the French SPOT satellite system in February of 1986, the civilian remote sensing sector was provided with the capability of applying high resolution MS imagery to a range of land use and land cover analyses. Cliche *et al.* [46] who worked with SPOT simulation data prior to the satellite's launch, showed that simulated 10-m resolution color images can be produced by modulating each SPOT MS (XS) band with PAN data individually, using three different IM methods. Welch *et al.* [47] used the term "merge" instead of 'integration' and proposed merging of SPOT PAN and XS data using the Intensity-Hue-Saturation (IHS) transformation, a method previously proposed by Haydn *et al.* [48] to merge Landsat MSS with Return Beam Vidicon (RBV) data and Landsat MSS with Heat Capacity Mapping Mission data. In 1988, Chavez *et al.* [49] used SPOT panchromatic data to "sharpen" Landsat TM images by high pass filtering (HPF) the SPOT PAN data before merging it with the TM data. A review of the so called classical methods, which include IHS, HPF, Brovey Transform (BT) [50] and Principal Component Substitution (PCS) [51,52], among others, can be found in [3].

In 1987, Price [53] developed a fusion technique based on the statistical properties of remote sensing images, for the combination of the two different spatial resolutions of the HRV-SPOT sensor. Besides the Price method, the literature shows other pansharpening methods based on the statistical properties of the images, such as spatially adaptive methods [54] and Bayesian-based methods [55,56].

More recently, multi-resolution analysis employing the generalized Laplacian pyramid (GLP) [57,58], the discrete wavelet transform [59,60] and the contourlet transform [61–63] has been used in pansharpening using the basic idea of extracting the spatial detail information from the PAN image not present in the low resolution MS image, to inject it into the later.

Image fusion methods have been classified in several ways. Schowengerdt [9] classified them into spectral domain, spatial domain, and scale space techniques. Ranchin and Wald [64] classified them into three groups: projection and substitution methods, relative spectral contribution methods, and those relevant to the ARSIS concept (from its French acronym *"Amélioration de la Résolution Spatiale par Injection de Structures"* which means "Enhancement of the spatial resolution by structure injections"). It was found that many of the existing image fusion methods, such as the HPF and additive wavelet transform (AWT) methods, can be accommodated within the ARSIS concept [65], but Tu *et al.* [4] found that the PCS, BT, and AWT methods could be also considered as IHS-like image fusion methods. Meanwhile Bretschneider *et al.* [66] classified IHS and PCA methods as transformation based methods, in a classification that also includes more categories such as addition and multiplication fusion, filter fusion (which includes HPF method), fusion based on inter-band relations, wavelet decomposition fusion and further fusion methods (based on statistical properties). Fusion methods that involve linear forward and backward transformation, had been classified by Sheftigara [67] as component substitution methods.

Although it is not possible to find a universal classification, in this work we classify the pansharpening methods into the following categories according to the main technique they use:

1. Component Substitution family, which includes IHS, PCS and GramSchmidt (GS), because all these methods utilize, usually, a linear transformation and substitution for some components in the transformed domain.

2. Relative Spectral Contribution family, which includes BT, IM and P+XS, where a linear combination of the spectral bands, instead of substitution, is applied.

3. High-Frequency Injection family, which includes HPF and HPM, where these two methods involve extracting high frequencies by subtracting a low pass filtering PAN image from the original one.

4. Methods based on the statistics of the image, which include Price and spatially adaptive methods, Bayesian-based and super resolution methods.

5. Multiresolution family including generalized Laplacian pyramid, wavelet and contourlet

methods and those methods using multiresolution analysis combined with methods from other categories.

2.3.1 Component Substitution (CS) family

The component substitution methods start by upsampling the low resolution MS image to the size of the PAN image. Then the MS image is transformed into a set of components, using usually a linear transform of the MS bands. The CS methods work by substituting a component of the (transformed) MS image, C_l, with a component, C_h, from the PAN image. These methods are physically meaningful only when these two components, C_l and C_h, contain almost the same spectral information. In other word, the C_l component should contain all the redundant information of the MS and PAN data but C_h should contain more spatial information. An improper construction of the C_l component tends to introduce high spectral distortion. The general algorithm for the CS sharpening techniques is summarized in Algorithm 2.1.

Algorithm 2.1 *Component substitution pansharpening*

```
1. Upsample the MS image to the size of the PAN image.
2. Forward transform the MS image to the desired components.
3. Match the histogram of the PAN image with the $C_l$ component to be substituted.
4. Replace the $C_l$ component with the histogram-matched PAN image.
5. Backward transform the components to obtain the pansharpened image.
```

The CS family includes many popular pansharpening methods, such as the IHS, PCS, Gram-Schmidt (GS) methods [67, 68], where each of them involves a different transformation of the MS image. CS techniques are attractive because they are fast and easy to implement and they allow users expectations to be fulfilled most of the time since they provide pansharpened images with good visual/geometrical quality in most cases [69]. However, the results obtained by thees methods highly depend on the correlation between the bands, and since the same transformation is applied to the whole image, it does not take into account local dissimilarities between PAN and MS images [27, 70].

Since the spectral transformation approaches are data dependent and a single type of transformation does not always obtain the optimal component required for substitution or transformation, it would be difficult to choose the appropriate spectral transformation method for various datasets. In order to alleviate this problem, recent methods incorporate statistical tests or weighted measures to adaptively select an optimal component for substitution and transformation. This results in a new approach known as *adaptive component substitution* [71–73].

The ***Intensity-Hue-Saturation (IHS)*** pansharpening method [49,74] is one of the classical techniques included in this family and it uses the IHS color space, which is often chosen due to the tendency of the visual cognitive system of human beings to treat the intensity (I), hue (H) and saturation (S) components as roughly orthogonal perceptual axes. IHS transformation originally was applied to RGB true color, but in the remote sensing applications and for display purposes only, arbitrary bands are assigned to RGB channel to produce false color composites [1]. The ability of IHS transformation to separate effectively spatial information (band I) and spectral information (bands H and S) [39] makes it very applicable in pansharpening methods. There are different models of IHS transformation, differing in the method used to compute the intensity value. Smith's hexacone and triangular models are two of the most widely used ones [25]. A full description of the IHS transformation and the algorithms to compute it can be found in [75]. An example of pansharpened image using IHS method is shown in Figure. 2.1(b).

The major limitation of this technique is that only three bands are involved. A new methodology, proposed by Silva *et al.* [76], reduces the dimensionality limitation by using multiple input images with a Generalized IHS transformation. In any case, since the spectral response of I, as synthesized from the MS bands, does not generally have the same radiometry as the histogram-matched PAN [69], when the fusion result is displayed in color composition, large spectral distortion may appear as color changes. In order to minimize the spectral distortion in IHS pansharpening, Rahmani *et al.* [71] proposed a new adaptive IHS method in which the intensity band approximates the PAN image as closely as possible. Therefore, this adaptive IHS method determines the coefficients α_i that best approximate

$$PAN = \sum_i \alpha_i MS_i, \tag{2.2}$$

subject to the physical constraint of non-negativity of the coefficients α_i. Note that, although this method will reduce the spectral distortion, local dissimilarities between MS and PAN images might be present and not be treated by this method [27].

Another method in the CS family is the ***principal component substitution (PCS)*** method which depends on the PCA mathematical transformation. The principal component analysis (PCA), also known as the Karhunen-Loéve transform or the Hotelling transform [77], is widely used in signal processing, statistics, and many other areas. This transformation generates a new set of images, referred to as components or axes, which are linear combinations of the original images, and allows the user to generate a new set of rotated orthogonal axes where the new image components are not correlated. The largest amount of the variance is mapped to the first component, with decreasing variance going to each of the following components. The sum of the variances in all the components is equal to the total variance present in the original input images. PCA and the calculation of the transformation matrices, can be performed following the steps specified in [39]. Theoretically, the first principal component, $PC1$, collects the information that is common to all bands used as input data to the PCA, i.e., the spatial information, while the spectral information that is specific to each band is captured in the other principal components [51,60]. This makes PCS an adequate technique when merging MS and PAN images. PCS is similar to the IHS method, with the main advantage that an arbitrary number of bands can be considered. However, some spatial information may not be mapped to the first component, depending on the degree of correlation and spectral contrast that exist among the MS bands [51], resulting in the same problems the IHS methods have. To overcome this drawback, Shah *et al.* [72] proposed a new Adaptive PCS-based pansharpening method that determines, using cross-correlation, the appropriate PC component that is to be substituted by the PAN image. By replacing the selected component with the highly correlated high spatial resolution PAN component, adaptive PCS method will produce better results than traditional ones [72].

A widespread CS technique is the ***Gram-Schmidt (GS)*** spectral sharpening. This method was invented by Laben and Brover in 1998 and patented by Eastman Kodak [78]. The GS transformation, as described in [79], is a common technique used in linear algebra and multivariate statistics. GS is used to orthogonalize matrix data or bands of a digital image removing redundant (i.e., correlated) information that is contained in multiple bands. If there were perfect correlation between input bands, the GS orthogonalization process would produce a final band with all its elements equal to zero. For its use in pansharpening, GS transformation had been modified [78]. In the modified process, the mean of each band is subtracted from each pixel in the band before the orthogonalization is performed to produce a more accurate outcome.

In GS-based pansharpening, a lower resolution PAN band needs to be simulated and used as the first band of the input to the GS transformation, together to the MS image. Two methods are used in [78] to simulate this band; in the first method, the LRMS bands are combined into a single lower resolution PAN (LR PAN) as the weighted mean of MS image. These weights depend on the spectral response of the MS bands and high resolution PAN (HR PAN) image and on the optical transmittance of the PAN band. The second method simulates the LR PAN image by blurring and sub-sampling the observed PAN image. The major difference in results, mostly noticeable in a true color display, is that the first method exhibits outstanding spatial quality, but spectral distortions may occur. This distortion is due to the fact that the average of the MS spectral bands is not likely to have the same radiometry as the PAN image. The second method is unaffected by spectral distortion but generally suffers from a lower sharpness and spatial enhancement. This is due to the injection mechanism of high-pass details taken from PAN, which is embedded into the inverse GS transformation, carried out by using the full resolution PAN, while the forward transformation uses the low resolution approximation of PAN obtained by resampling the decimated PAN image provided by the user. In order to avoid this drawback, Aiazzi *et al.* [73] proposed an Enhanced GS method, where the LR PAN is generated by a weighted average of the MS bands and the weights are estimated to minimize the MMSE with the downsampled PAN.

2.3.2 Relative Spectral Contribution (RSC) family

The RSC family can be considered as a variant of the CS pansharpening family, when a linear combination of the spectral bands, instead of substitution, is applied.

Let PAN^h be the high spatial resolution PAN image, MS_b^l the b low resolution MS image band, h the original spatial resolution of PAN, and l the original spatial resolution of MS_b ($l < h$), while MS_b^h is the image MS_b^l re-sampled at resolution h. RSC works only on the spectral bands MS_b^l lying within the spectral range of the PAN^h image. The synthetic (pansharpened) bands $HRMS_b^h$ are given at each pixel (i,j) by

$$HRMS_b^h(i,j) = \frac{MS_b^h(i,j)PAN^h(i,j)}{\sum_b MS_b^h(i,j)}, \tag{2.3}$$

where $b = 1,2,\cdots,B$ and V is the number of MS bands. The process flow diagram of the RSC sharpening techniques is shown in Algorithm 2.2. This family does not tell what to do when MS_b^l lies outside the spectral range of PAN^h. In Eq. (2.3) there is an influence of the other spectral bands on the assessment of $HRMS_b^h$, thus causing a spectral distortion. Furthermore the method does not preserve the original spectral content once the pansharpened images $HRMS_b^h$ are brought back to the original low spatial resolution [64]. These methods include the Brovey transform (BT) [50], the P+XS [80, 81] and the Intensity modulation (IM) method [82].

```
Algorithm 2.2 Relative spectral contribution pansharpening
  1. Upsample the MS image to the size of the PAN image.
  2. Match the histogram of the PAN image with each MS band.
  3. Obtain the pansharpened image by applying Eq. (2.3).
```

The Brovey Transform (BT), named after its author, is a simple method to merge data from different sensors based on the chromaticity transform [50], with the limitation that only three bands are involved [1, 60]. A pansharpened image using the BT method is shown in Figure 2.1(c).

The Brovey transform provides excellent contrast in the image domain but greatly distorts the spectral characteristics [83]. The Brovey sharpened image is not suitable for pixel-based classification as the pixel values are changed drastically [25]. A variation of the BT method subtracts the intensity of the MS image from the PAN image before applying Eq. (2.3) [1]. Although, the first BT method injects more spatial details, the second one preserves better the spectral details.

The concept of ***intensity modulation (IM)*** was originally proposed by Wong *et al.* [42] in 1980 for integrating Landsat MSS and Seasat SAR images. Later, this method was used by Cliche *et al.* [46] for enhancing the spatial resolution of three-bands SPOT MS (XS) images. As a method in the relative spectral contribution family, we can derive IM from Eq. (2.3), by replacing the sum of all MS bands, by the intensity component of the IHS transformation [84]. Note that the use of the IHS transformation limits to three the number of bands utilized by this method. The intensity modulation may cause color distortion if the spectral range of the intensity replacement (or modulation) image is different from the spectral range covered by the three bands used in the color composition [85]. In the literature, different versions based on the IM concept, have been used [46, 84, 85].

2.3.3 High-frequency injection family

The high-frequency injection family methods were first proposed by Schowengerdt [86], working on full-resolution and spatially compressed Landsat MSS data. He demonstrated the use of a high-resolution band to "sharpen" or edge-enhance lower resolution bands having the same approximate wavelength characteristics. Some years later, Chavez [87] proposed a project whose primary objective was to extract the spectral information from the Landsat TM and combine (inject) it with the spatial information from a data set having much higher spatial resolution. To extract the details from the high resolution data set he used a high pass filter in order to "enhance the high frequency/spatial information but, more important, suppress the low frequency/spectral information in the higher resolution image" [49]. This was necessary so that simple addition of the images did not distort the spectral balance of the combined product.

An useful concept for understanding spatial filtering is that any image is made of spatial components at different kernel sizes. Suppose we process an image in such a way that the value at each output pixel is the average of a small neighborhood of input pixels, a box filter. The result is a Low-Pass (LP) blurred version of the original image that will be noted as LP. Subtracting this image from the original one, produces High-Pass (HP) image, that represents the difference between each original pixel and the average of its neighborhood. This relation can be written as,

$$image(i,j) = LP(i,j) + HP(i,j), \tag{2.4}$$

which is valid for any neighborhood size (scale). As the neighborhood size is increased, the LP image hides successively larger and larger structures, while the HP image picks up the smaller structures lost in the LP image (see Eq. (2.4)) [9].

The idea behind this type of spatial domain fusion is to transfer the high frequency content of the PAN image to the MS images by applying spatial filtering techniques [88]. However, the size of the filter kernels can not be arbitrary because it has to reflect the radiometric normalization between the two images. Chavez *et al.* [52] suggested that the best kernel size is approximately twice the size of the ratio of the spatial resolutions of the sensors, which produce edge-enhanced synthetic images with the least spectral distortion and edge noises. According to [89] pansharpening methods based on injecting high-frequency components into re-sampled versions of the MS data have demonstrated a superior performance, and compared with many other pansharpening methods such as the methods in the CS family. Several variations of high-frequency injection pansharpening methods have been proposed as High-Pass Filtering Pansharpening and High Pass Modulation.

As we have already mentioned, the main idea of the ***High-Pass Filtering (HPF)*** pansharpening method is to extract from the PAN image the high-frequency information, to later add or inject it into the MS image previously expanded to match the PAN pixel size. This spatial information extraction is performed by applying a low-pass spatial filter to the PAN image,

$$filtered_{PAN} = h_0 * PAN, \tag{2.5}$$

where h_0 is a low-pass filter and $*$ the convolution operator. The spatial information injection is performed adding, pixel by pixel, the filtered image that results from subtracting $filtered_{PAN}$ from the original PAN image, to the MS one [49, 90]. There are many different filters that can be used, like Box filter, Gaussian, Laplacian, and so on. Recently, the use of the modulated transfer function (MTF) of the sensor as the low-pass filter has been proposed in [91]. The MTF is the amplitude spectrum of the system point spread function (PSF) [92]. In [91], the HP image is also multiplied by a weight selected to maximize the Quality Not requiring a Reference (QNR) criterion proposed in the paper.

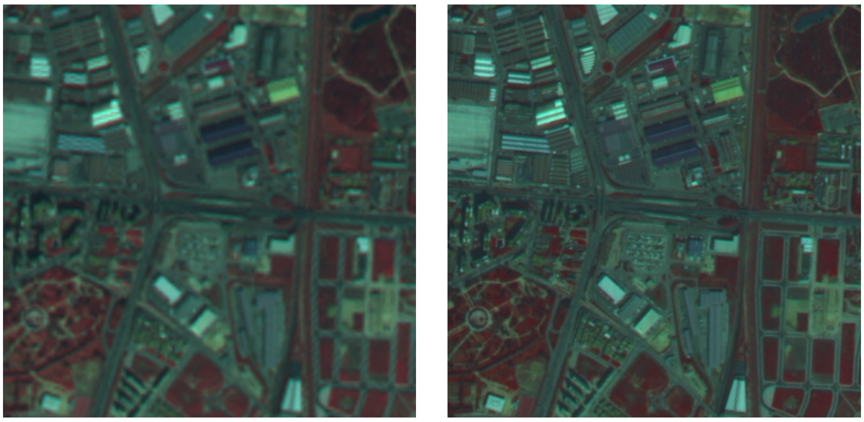
As expected, HPF images present low spectral distortion. However, the ripple in the frequency response will have some negative impact [1]. The HPF method could be considered the predecessor of an extended group of image pansharpening procedures based on the same principle: to extract spatial detail information from the PAN image not present in the MS image and inject it into the latter in a multiresolution framework. This principle is known as the *ARSIS concept* [64].

In the ***High Pass Modulation (HPM)***, also known as ***High Frequency Modulation (HFM)*** algorithm [9], the HR PAN image is multiplied by each band of the LRMS image and normalized by a low-pass filtered version of the PAN image to estimate the enhanced MS image bands. The principle of HPM is to transfer the high-frequency information of the PAN image to the LRMS band b ($LRMS_b$) with a modulation coefficient k_b which equals the ratio between the LRMS and the low-pass filtered version of the PAN image [1]. Thus, the algorithm assumes that each pixel of the enhanced (sharpened) MS image in band b is simply proportional to the corresponding higher-resolution image at each pixel. This constant of proportionality is a spatially-variable gain factor, calculated by,

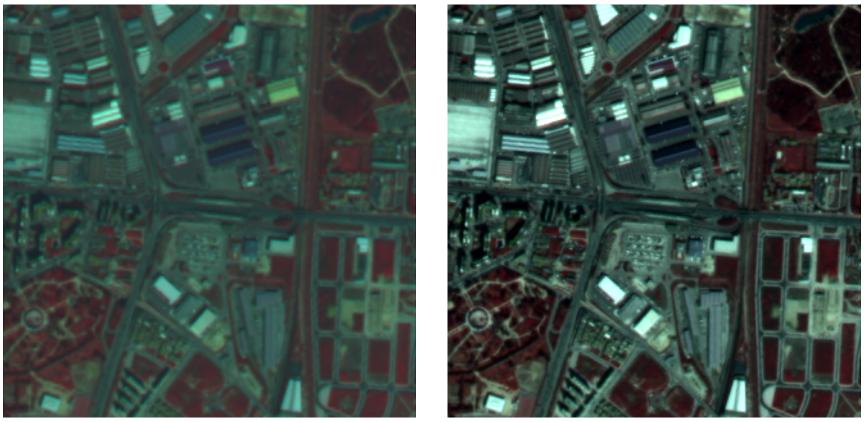
$$k_b(i,j) = \frac{LRMS_b(i,j)}{filtered_{PAN}(i,j)}, \tag{2.6}$$

where $filtered_{PAN}$ is a low-pass filtered version of PAN image (see Eq. (2.5)) [9]. According to [1] when the low-pass filter is chosen as in the HPF method, the HPM method will give slightly better performance than HPF because the color of the pixels is not biased toward gray.

The process flow diagram of the HFI sharpening techniques is shown in Algorithm 2.3. Also a pansharpened image using the HPM method is shown in Figure 2.1(d).



(a) Original LRMS image (b) IHS



(c) BT (d) HPF

Figure 2.1: *Results of some classical pansharpening methods.*

Algorithm 2.3 *High-frequency injection pansharpening*

1. *Upsample the MS image to the size of the PAN image.*

2. *Apply a low pass filter on the PAN image using Eq. (2.5).*

3. *Calculate the high frequency image by subtracting the filtered PAN from the original PAN.*

4. *Obtain the pansharpened image by adding the high frequency image to each band of the MS image (modulated by the factor k_i in Eq. (2.6) in the case of HPM).*

2.3.4 Methods based on the statistics of the image

The methods based on the statistics of the image include a set of methods that exploit the statistical characteristics of the MS and PAN images in the pansharpening process. The first known method in this family was proposed by Price [53] to combine PAN and MS imagery from dual resolution satellite instruments based on the substantial redundancy existing in the PAN data and the local correlation between the PAN and MS images. It is based on two previous works where Price compared the information content of data from satellites imagery [93] and did a statistical analysis for this data [94] in order to compute their redundancy. Later, the method was improved by Price [95] by computing the local statistics of the images and by Park *et al.* [54] in the so called spatially adaptive algorithm.

Price's method [95] uses the statistical relationship between each band of the LRMS image and HR images to sharpen the former. It models the relationship between the pixels of each band of the HRMS z_b, the PAN image x and the corresponding band of the LRMS image y_b linearly as

$$z_b - \hat{y_b} = \hat{a}(x - \hat{x}), \tag{2.7}$$

where $\hat{y_b}$ is the band b of the LRMS image y upsampled to the size of the HRMS image by pixel replication, $\hat{x}$ represents the panchromatic image downsampled to the size of the MS image by averaging the pixels of x in the area covered by the pixels of y and upsampling again to its original size by pixel replication, and $\hat{a}$ is a matrix defined as the upsampling, by pixel replication, of a weight matrix a whose elements are calculated from a window 3×3 of each LR image pixel.

Price's algorithm succeeds in preserving the lower resolution radiometry in the fusion process, but sometimes it produces blocking artifact because it uses the same weight for all the HR pixels corresponding to one LR pixel. If the HR and LR images have little correlation, the blocking artifacts will be severe. A pansharpened image using Price's method proposed in [96] is shown in Figure 2.3 (a).

The ***spatially adaptive algorithm*** [54] starts from Price's method [95], but with a more general and improved mathematical model.

It features adaptive insertion of information according to the local correlation between the two images preventing spectral distortion as much as possible and sharpening the MS images simultaneously. This algorithm has also the advantage that a number of high resolution images, not only one PAN image, can be utilized as references of high frequency information, which is not the case for most methods [54].

Besides those methods, most of the papers in this family have used the Bayesian framework to model the knowledge about the images and estimate the pansharpened image. Since the work of Mascarenhas [55], a number of pansharpening methods have been proposed using the Bayesian framework (see [97,98] for instance).

Bayesian methods model the degradation suffered by the original HRMS image, z as the conditional probability distribution of the observed LRMS image, y, and the PAN image, x, given the original z, called the likelihood and denoted as $p(y,x|z)$. They take into account the available prior knowledge about the expected characteristics of the pansharpened image, modeled in the so called prior distribution $p(z)$, to determine the posterior probability distribution $p(z|y,x)$ by using Bayes law,

$$p(z|y,x) = \frac{p(y,x|z)p(z)}{p(y,x)}, \tag{2.8}$$

where $p(y,x)$ is the joint probability distribution. Inference is performed from the posterior distribution to draw estimates of the HRMS image, z.

The main advantage of the Bayesian approach is to place the problem of pansharpening into a clear probabilistic framework [98], although assigning suitable distributions for the conditional and prior distributions and the selection of an inference method are critical points that lead to different Bayesian-based pansharpening methods.

As prior distribution, Fasbender *et al.* [98] assumes a noninformative prior, $p(z) \propto 1$ which gives equal probability to all possible solutions, that is, no solution is preferred as no clear information on the HRMS image is available. This prior has also been used by Hardie *et al.* [99]. In [55] the prior information is carried out by an interpolation operator and its covariance matrix; both will be used as the mean vector and the covariance matrix, respectively, for a Bayesian synthesis process. In [100] the prior knowledge about the smoothness of the object luminosity distribution within each band makes it possible to model the distribution of z using a Simultaneous Autorregressive Model (SAR) as

$$p(z) = \prod_{b=1}^{B} p(z_b) \propto \prod_{b=1}^{B} exp\left\{-\frac{1}{2}\alpha_b \|Cz_b\|^2\right\} \tag{2.9}$$

where C denotes the Laplacian operator and $1/\alpha_b$ is the variance of the Gaussian distribution of z_b, $b = 1,\dots,B$, with B being the number of bands in the MS image. More advanced models try to incorporate a smoothness constrain while preserving the edges in the image. Those models include adaptive SAR model [56], Total Variation (TV) [96], Markov Random Fields (MRF) [101] based models or Stochastic Mixing Models (SMM) [102]. Note that the described models do not take into account the correlation existing between the MS bands. In [103] the authors propose a TV prior model to take into account spatial pixel relationships and a SAR model to enforce similarity between the pixels in the same position in the different bands.

It is usual to model the LRMS and PAN images as degraded versions of the HRMS image by two different processes; one modeling the LRMS image and usually described as

$$y = g_s(z) + n_s, \tag{2.10}$$

where $g_s(z)$ represents a function that relates z with y and n_s represents the noise of the LRMS image, and a second one that models how the PAN image is obtained from the HRMS image which is written as

$$y = g_p(z) + n_p, \tag{2.11}$$

where $g_p(z)$ represents a function that relates z with x and n_p represents the noise of the PAN image. Note that, since the success of the pansharpening algorithm will be limited by the accuracy of those models, the physics of the sensor should be considered. In particular the MTF of the sensor and the sensor's spectral response should be taken into account.

The conditional distribution of the observed images given the original one, $p(y,x|z)$, is usually defined as

$$p(y,x|z) = p(y|z)p(x|z) \tag{2.12}$$

by considering that the observed LRMS image and the PAN image are independent given the HRMS image. This allows an easier formulation of the degradation models. However, Fasbender *et al.* [98] took into account that y and x may carry information of quite different quality about z and defined $p(y,x|z) = p(y|z)^{2(1-w)}p(x|z)^{2w}$, where the parameter $w \in [0,1]$ can be interpreted as the weight to be given to the panchromatic information at the expense of the MS information. Note that $w = 0.5$ leads back to Eq. (2.12) while a value of zero or one means that we are discarding the PAN or the MS image, respectively.

Different models have been proposed for the conditional distributions $p(y|z)$ and $p(x|z)$. The simpler model is to assume that $g_s(z) = z$, so that y will be then $y = z + n_s$ [98] where $n_s \sim N(0, \Sigma_s)$. Note that in this case, y has the same resolution as z so an interpolation method has to be used to obtain y from the observed MS image. However, most of the authors consider the relation $y = Hx + n_s$ where H is a matrix representing the blurring, usually represented by its MTF, the sensor integration function, and the spatial subsampling and n_s is the capture noise, assumed to be Gaussian with zero mean and variance $1/\beta$, leading to the distribution

$$p(y|z) \propto \exp\left\{-\frac{1}{2}\beta\|y - Hx\|^2\right\}. \tag{2.13}$$

This model has been extensively used (see [101, 102, 104]) and it is the base for the so called super-resolution based methods [105] as the ones described, for instance, in [56,96]. The degradation model in [55] can be also written in this way. A pansharpened image using the super-resolution method proposed in [96] is shown in Figure 2.3 (b).

On the other hand, $g_p(z)$ has been defined as a linear-regression model linking the MS pixels to the PAN ones, as estimated from both observed images, so that $g_p(z) = a + \sum_{b=1}^{B} \lambda_b z_b$, where a and λ_b, $b = 1, 2, \ldots, B$, are the regression parameters. Note that this model is used by IHS, PCA, and Brovey methods to relate the PAN and HRMS images. Mateos *et al.* [106] (and also [56,96,101] for instance) used a special case for $g_p(z)$, where $a = 0$ and $\lambda_b \geq 0$, $b = 1, 2, \ldots, B$, are known quantities that can be obtained from the sensor spectral characteristics (see Figure 2.2 for the Landsat 7 ETM+ spectral response) that represent the contribution of each MS band to the PAN image. In all those papers the noise n_p is assumed to be Gaussian with zero mean and covariance matrix C_p and hence,

$$p(x|z) \propto \exp\left\{-\frac{1}{2}((x - g_p(z))^t C_p^{-1}(x - g_p(z))\right\}. \tag{2.14}$$

Finally, we want to mention that a similar approach has been used in the problem of hyperspectral (HS) resolution enhancement in which a HS image is sharpened by a higher resolution MS or PAN image. In this context, Eismann and Hardie [102, 104] and other authors later (see for instance [107]) proposed to use the model $x = S^t z + n$, where z is the HR original HS image, x is a HRMS or PAN image used to sharpen a LRHS image, S is the spectral response matrix and n is assumed to be a spatially independent zero-mean Gaussian noise with covariance matrix C. The spectral response matrix is a sparse matrix that contains in each column the spectral response of a MS band of x. Note that in the case of pansharpening, the image x has only one band and the matrix S will be a column vector with components λ_b as in the model proposed in [106].

Once the prior and conditional distributions have been defined, Bayesian inference is performed to find an estimate of the original HRMS image. Different methods have been used in the literature to carry out the inference, depending on the form of the chosen distributions. Maximum likelihood (ML) [98], linear minimum mean square error (LMMSE) [107], Maximum a Posteriori (MAP) [99], the Variational approach [56, 96] or Simulated Annealing [101] are some of the techniques used.

Although all approaches already mentioned use the hypothesis of Gaussian additive noise for mathematical convenience, in practice, remote sensing imaginary noise shows non-Gaussian

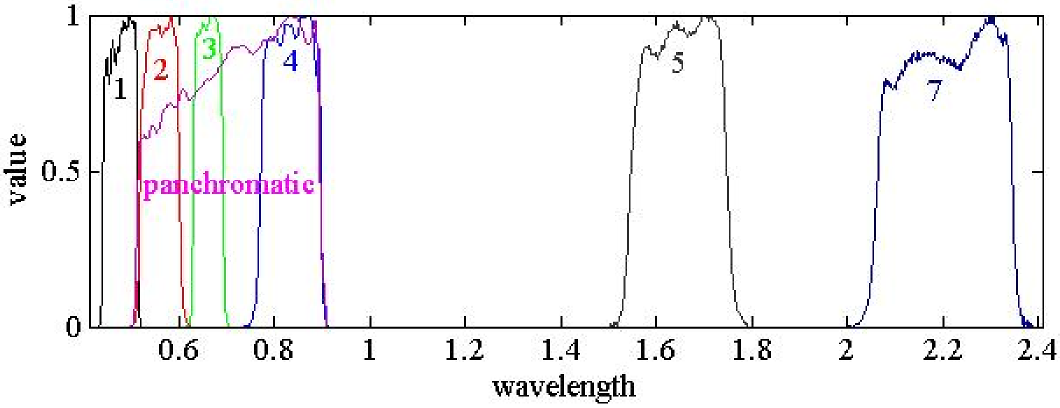
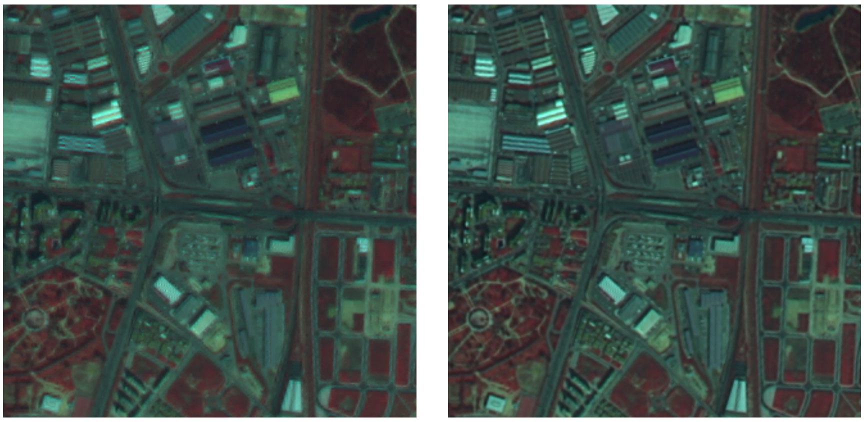


Figure 2.2: *Landsat 7 ETM+ band spectral response*



(a) Price (b) Super-resolution [96]

Figure 2.3: *Results of some statistical pansharpening methods*

characteristics [108]. In some applications, such as astronomical image restoration, Poisson noise is usually used, or a shaping filter [109] may be used in order to transform non-Gaussian noise into Gaussian. Recently, Niu *et al.* [108] proposed the use of a mixture of Gaussian (MoG) noise for multisensor fusion problems.

2.3.5 Multiresolution Family

In order to extract or modify the spatial information in remote sensing images, spatial transforms represent also a very interesting tool. Some of these transforms use only local image information (i.e., within a relatively small neighborhood of a given pixel), such as convolution, while other use frequency content, such as the Fourier transform. Beside these two extreme transformations, there is a need for a data representation allowing the access to spatial information over a wide range of scales from local to global [9]. This increasingly important category of scale-space filters utilize multi-scale decomposition techniques such as Laplacian pyramids [110], wavelet transform [59], contourlet transform [111] and curvelets transform. These techniques are used in pansharpening methods to decompose MS and PAN images with different levels in order to derive spatial details that are imported into finer scales of the MS images, highlight the relationship between PAN and MS images in coarser scales and enhance spatial details [112]. This is the idea behind the methods based on the successful ARSIS (from French “Amèlioration de la Résolution by Structure Injection”) concept [64].

We will now describe each of the above multiresolution methods and their different types in detail.

Multiresolution analysis based on the ***Laplacian pyramid (LP),*** originally proposed in [110], is a bandpass image decomposition derived from the Gaussian pyramid (GP) which is a multiresolution (multiscale) image representation obtained through a recursive reductions of the image. LP is an oversampled transform, that the image is decomposed into nearly disjointed bandpass channels in the spatial frequency domain, without losing the spatial connectivity of its edges [113]. Figure 2.4 shows the concept of GP and its relation to LP. The Generalized Laplacian Pyramid (GLP) is an extension of the LP where a factor different of two is used [114]. An attractive characteristic of the GLP is that the lowpass reduction filter used to analyze the PAN image may be designed to match the MTF of the band into which the details extracted will be injected. The benefit is that the restoration of the spatial frequency content of the MS band is embedded into the enhancement procedure of the band itself, instead of being accomplished ahead of time.

The steps for merging Landsat images using this GLP are described in Algorithm 2.4, where different injection methods can be used with GLP [58, 115]. In this context, injection means adding the details from the GLP to each MS band weighted by the coefficients obtained by the injection method. The Spectral Distortion Minimizing (SDM) injection model is both a spatially and spectrally varying model where the injected details at a pixel position must be parallel to the re-sampled MS vector at the same resolution. At the same time the details are weighted to minimize the radiometric distortion measured as the Spectral Angle Mapper (SAM). In Context-Based Decision (CBD) injection model, the weights are calculated locally between the MS band resampled to the scale of the PAN image and a approximation of the

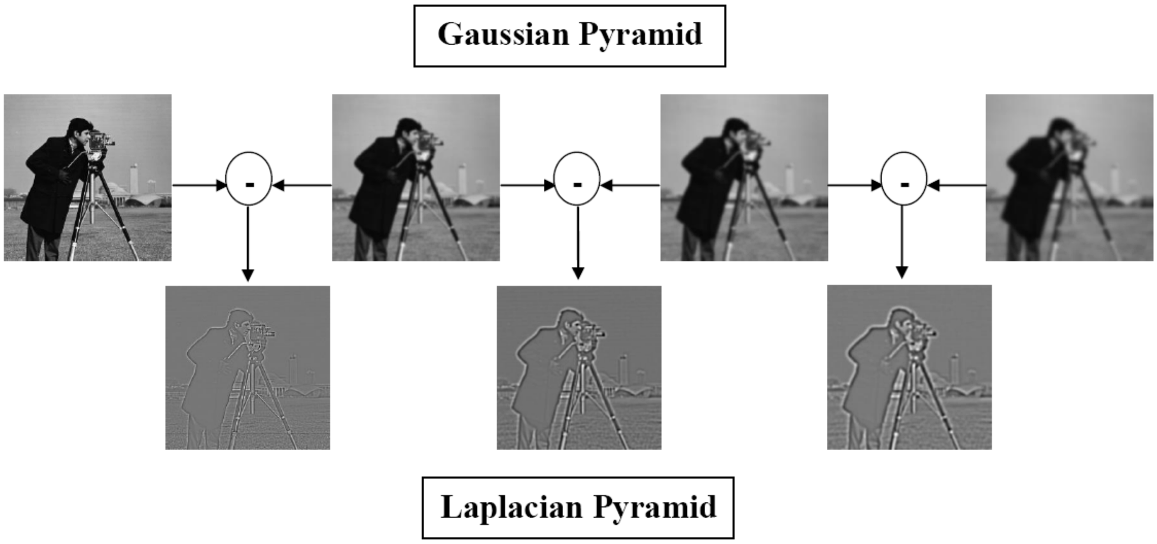


Figure 2.4: *Laplacian Pyramid created from Gaussian pyramid by subtraction*

PAN image at the resolution of the MS bands and only injects details if the local correlation coefficient between those images, calculated on a window of size $N \times N$, is larger than a given threshold. The CBD model is uniquely defined by the set of thresholds generally different for each band, and by the window size N depending on the spatial resolutions and scale ratio of the images to be merged, as well as on the landscape characteristics to avoid loss of local sensitivity [58]. Pansharpened images using wavelet/contourlet-based methods are shown in Figure 2.5 (a).

Algorithm 2.4 *General Laplacian Pyramid-based pansharpening*

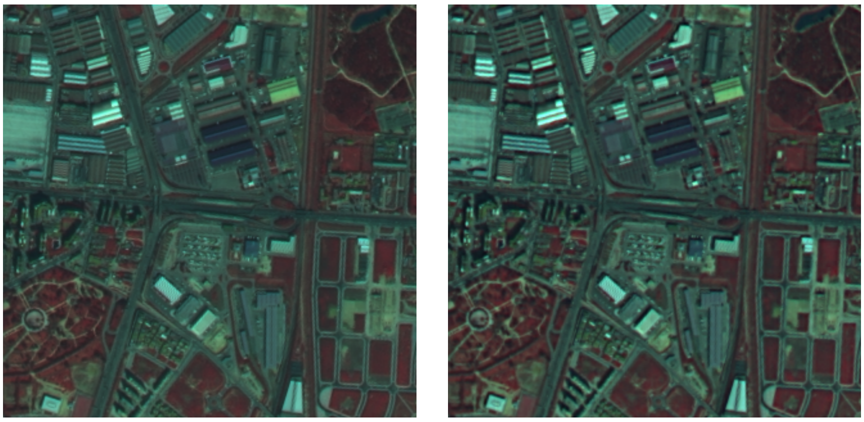
```
1. Upsample each MS band to the size of the PAN image.
2. Apply GLP on the PAN image.
3. According to the injection model, select the weights for the details from GLP at each level.
4. Obtain the pansharpened image by adding the details from the GLP to each MS band weighted by the coefficients obtained in the previous step.
```

The Ranchin-Wald-Mangolini (RWM) injection model [58], unlike the SDM and CBD models, is calculated on bandpass details instead of approximations. In the RWM model the details of the MS band are modeled using a linear model of the coefficients of the PAN image where the scale and offset are space and spectral-varying coefficients, respectively, and their calculations are based on modeling the detail coefficients as having a non-zero mean.

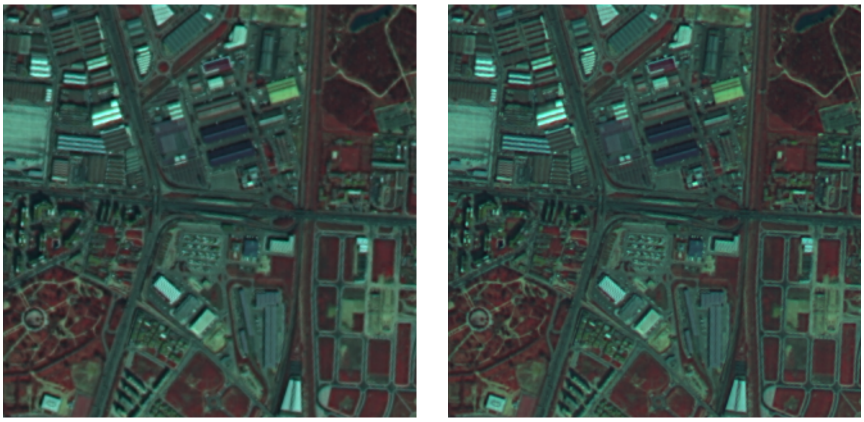
Another popular category of multiresolution pansharpening methods is the one based on ***Wavelet and Contourlet***. The wavelet transform provides a framework to decompose images into a number of new images, each of them with a decreasing degree of resolution, and to separate the spatial detail information of the image between two successive resolution degrees [116]. The discrete approach of the wavelet transform, named Discrete Wavelet Transform (DWT), can be performed using several different approaches, probably the most popular ones for image pansharpening being Mallat's [60, 64, 117, 118] and the "a' trous" [65, 119, 120] algorithms. Each one has its particular mathematical properties and leads to different image decompositions. The first is an orthogonal, dyadic, non-symmetric, decimated, non-redundant DWT algorithm, while "a'trous" is a non-orthogonal, shift-invariant, dyadic, symmetric, undecimated, redundant DWT algorithm [116]. Redundant wavelet decomposition, as well as GLP, has an attractive characteristic: the lowpass reduction filter used to analyze the PAN image may be easily designed to match the MTF of the band to be enhanced. If the filters are correctly chosen, the extracted high spatial frequency components from the PAN image are properly retained, thus resulting in a greater spatial enhancement.

Contourlets provide a new representation system for image analysis [111]. The contourlet transform is so called because of its ability to capture and link the point of discontinuities to form a linear structure (contours). The two-stage process used to derive the contourlet coefficients involves a multiscale transform and a local directional transform. The point of discontinuities and multiscale transformation is obtained via the Laplacian pyramid. The local directional filter bank is used to group these wavelet-like coefficients to obtain a smooth contour. Contourlets provide $2l$ directions at each scale, where l is the number of required orientation. This flexibility of having different numbers of direction at each scale makes contourlets different from other available multiscale and directional image representations [72]. Similarly to wavelets, contourlets also have different implementations of the subsampled [111] and non-subsampled [61] transforms.

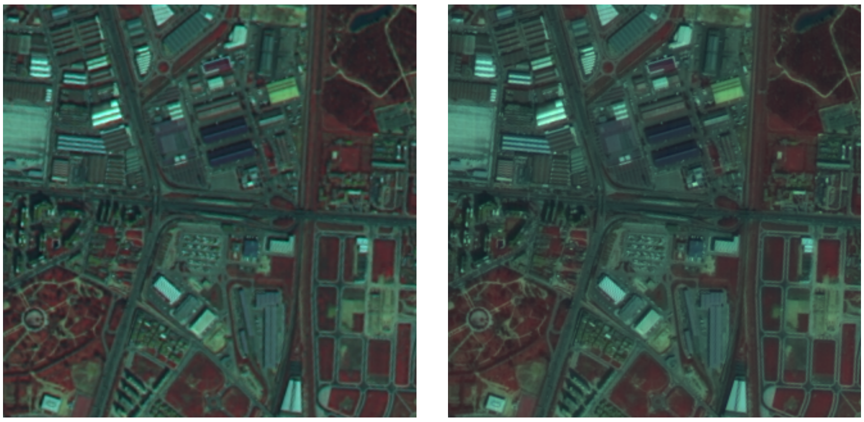
A number of pansharpening methods using the wavelet and, more recently, the contourlet transform have been proposed. In general, all the transform based pansharpening methods use the algorithm described in Algorithm 2.5. In the wavelet/contourlet-based approach, the MS and PAN images need to be decomposed multiple times in step 1 of the algorithm in Algorithm 2.5.



(a) GLP-SDM (b) Additive wavelet



(c) Wavelet Additive IHS (d) Additive contourlet



(e) Contourlet Additive IHS (f) Method in [31]

Figure 2.5: *Results of some multiresolution pansharpening methods*

Algorithm 2.5 *Wavelets / contourlets based pansharpening*

```
1. Forward transform the PAN and MS images using a sub-band and directional decomposition such as the subsampled or non-subsampled wavelet or contourlet transform.
2. Apply a fusion rule onto the transform coefficients.
3. Obtain the pansharpened image by performing the inverse transform.
```

Preliminary studies have shown that the quality of the pansharpened images produced by the wavelet-based techniques is a function of the number of decomposition levels. If few decomposition levels are applied, the spatial quality of the pansharpened images is less satisfactory. If an excessive number of levels is applied, the spectral similarity between the original MS and the pansharpened images decreases. Pradhan *et al.* [121] attempt in their work to determine the optimal number of decomposition levels for the wavelet-based pansharpening, producing the optimal spatial and spectral quality.

The fusion rules in step 2 of the algorithm comprise, for instance, substituting the original MS coefficient bands by the coefficients of the PAN image or adding the coefficients of the PAN to the coefficients of the original MS bands weighted sometimes by a factor related with the contribution of the PAN image to each MS band. It results in the different wavelet and contourlet-based pansharpening methods that will be described next.

The ***Additive Wavelet/Contourlet*** method for fusing MS and PAN images uses the wavelet [116] /contourlet [62] transform in steps 1 and 3 in Algorithm 2.5 and, for the fusion rule in step 2, it adds the detail bands of the MS image to those corresponding of the PAN image, having previously matched the MS histogram to one of the PAN image.

The ***Substitutive Wavelet/Contourlet*** methods are quite similar to the additive ones but, instead of adding the information of the PAN image to each band of the MS image, these methods simply replace the MS detail bands with the details obtained from the PAN image (see [119] for wavelet and [122] for contourlet decomposition).

A number of hybrid methods have been developed to attempt to combine the best aspects of classical methods and wavelet and contourlet transforms. Research has mainly focused on incorporating the IHS, PCA, BT into wavelet and contourlet methods.

As we have seen, some of the most popular image pansharpening methods are based on the IHS transformation. The main drawback of these methods is the high distortion of the original spectral information present in the resulting MS images. To avoid this problem, the

IHS transformation is followed by the additive wavelet or countourlet method in the so called ***wavelet*** [116] and ***contourlet*** [123, 124] ***additive IHS*** pansharpening. If the IHS transform is followed by the substitutive wavelet method, the ***wavelet substitutive IHS*** [125] pansharpening method is obtained.

Similarly to the IHS wavelet/contourlet methods, the ***PCA wavelet [90, 116] /contourlet [72]*** methods are based on applying substitutive wavelet/contourlet methods to the first principal component (PC1) instead of applying it to the bands of the MS image. Adaptive PCA has also been applied in combination with contourlets [72].

The ***WiSpeR*** [5] method can be considered as a generalization of different wavelet-based image fusion methods. It uses a modification of the non-subsampled additive wavelet algorithm where the contribution from the PAN image to each of the fused bands depends on a factor generated both from the sensor spectral response and the physical properties of the observed object. A new contourlet panshapening method named ***CiSper*** was proposed in [63] that, similarly to WiSper, weights the contribution of the PAN image to each MS band but it uses a different method to calculate these weights and uses the non-subsampling contourlet transform instead of the wavelet transform. In order to take advantage of multiresolution analysis, the use of ***pansharpening based on the statistics of the image on the wavelet/contourlet domain*** has been suggested [31, 126]. Pansharpened images using wavelet and contourlet-based methods are shown in Figure 2.5 (b-f).

Some authors [59,60] state that multi-sensor image fusion is a trade-off between the spectral information from an MS sensor and the spatial information from a PAN sensor and wavelet transform fusion methods easily control this trade-off. The trade-off idea, however, is just a convenient simplification, as discussed in [27], and ideal fusion methods must be able to simultaneously reach both spectral and spatial quality, and not one at the expense of the other. To do so, the physics of the capture process have to be taken into account and the methods have to adapt to the local properties of the images.

2.4 Quality Assessment

In the previous section, a number of different pansharpening algorithms have been described to produce images with both high spatial and spectral resolutions. The suitability of these images for various applications depends on the spectral and spatial quality of the pansharpened images.

Besides visual analysis, there is a need to quantitatively assess the quality of different pansharpened images. Quantitative assessment is not easy as the images to be compared are at different spatial and spectral resolutions. Wald *et al.* [89] formulated that the pansharpened image should have the following properties:

1. Any pansharpened image once downsampled to its original spatial resolution should be as similar as possible to the original image.

2. Any pansharpened image should be as similar as possible to the image that a corresponding sensor would observe with the same high spatial resolution.

3. The MS set of pansharpened images should be as similar as possible to the MS set of images that a corresponding sensor would observe with the same high spatial resolution.

These three properties have been reduced to two properties: consistency property and synthesis property [127]. The consistency property is the same as the first property and the synthesis property combines the second and third properties defined by [89]. The synthesis property emphasizes the synthesis at an actual higher spatial and spectral resolution.

Note that the reference image for the pansharpening process is the MS image at the resolution of the PAN image. Since this image is not available, Wald *et al.* [89] proposed a protocol for quality assessment and several quantitative measures for testing the three properties. The consistency property is verified by down-sampling the fused image from the higher spatial resolution h to their original spatial resolution l using suitable filters. To verify the synthesis properties, the original PAN at resolution h and MS at resolution l are down-sampled to their lower resolutions l and v respectively. Then, PAN at resolution l and MS at resolution v are fused to obtain fused MS at resolution l that can be then compared with the original MS image. The quality assessed at resolution l is assumed to be close to the quality at resolution h. This reduces the problem of reference images. However, we cannot predict the quality at higher resolution from the quality of lower resolution [128]. Recently a set of methods has been proposed to assess the quality of the pansharpened without the requirement of a reference image. Those methods aim at providing reliable quality measures at full scale following Wald's protocol.

2.4.1 Visual Analysis

Visual analysis is needed to check if the objective of pansharpening has been met. The general visual quality measures are the global image quality (geometric shape, size of objects), the spatial details and the local contrast. Some visual quality parameters for testing the properties are [127]: (1) *spectral preservation of features in each MS band*, where the appearance of the objects in the pansharpened images are analyzed in each band based on the appearance of the same objects in the original MS images; (2) *multispectral synthesis in pansharpened images*, where different color composites of the fused images are analyzed and compared with that of original images to verify that MS characteristics of objects at higher spatial resolution is similar to that in the original images; and (3) *synthesis of images close to actual images at high resolution as defined by the synthesis property of pansharpened images*, that cannot be directly verified but can be analyzed from our knowledge of the spectra of objects present in the lower spatial resolutions.

2.4.2 Quantitative Analysis

A set of measures has been proposed to quantitatively asses the spectral and spatial quality of the images. In this section we will present the measures more commonly used for this purpose.

Spectral Quality Assessment: To measure the spectral distortion due to the pansharpening process, each merged image is compared to the reference MS image, using one or more of the following quantitative indicators:

1. *Spectral Angle Mapper (SAM):* SAM denotes the absolute value of the angle between two vectors, whose elements are the values of the pixels of the HRMS image and the MS image at each image location. A SAM value equal to zero denotes the absence of spectral distortion, but radiometric distortion may be present (the two pixels vectors are parallel but have different lengths). SAM is measured either in degrees or radians and is usually averaged over the whole image to yield a global measurement of spectral distortion [129].

2. *Relative-shift mean (RM):* The RM [130] of each band of the fused image helps to visualize the change in the histogram of fused image and is defined in [130] as the percentage of variation between the mean of the reference image and the pansharpened image.

3. *Correlation coefficient (CC):* The CC between each band of the reference and the pansharpened image indicates the spectral integrity of pansharpened image [83]. However,

CC is insensitive to a constant gain and bias between two images and does not allow for subtle discrimination of possible pansharpening artifacts [1]. CC should be as close to 1 as possible.

4. *Root mean square error (RMSE):* The RMSE between each band of the reference and the pansharpened image measures the changes in radiance of the pixel values [89]. RMSE is a very good indicator of the spectral quality when it considered along homogeneous regions in the image [130]. RMSE should be as close to 0 as possible.

5. *Structure Similarity Index (SSIM):* SSIM [131] is a perceptual measure that combines several factors related to the way humans perceive the quality of the images. Beside luminosity and contrast distortions, structure distortion is considered in SSIM index and calculated locally in 8×8 square windows. The value varies between -1 and 1. Values close to 1 show the highest correspondence with the original images.

   The *Universal Image Quality Index (UIQI)* proposed in [132], can be considered a special case for SSIM index.

While these parameters only evaluate the difference in spectral information between each band of the merged and the reference image, in order to estimate the global spectral quality of the merged images the following parameters are used:

1. *Erreur relative globale adimensionnelle de synthése (ERGAS) index:* whose English translation is relative dimensionless global error in fusion [133], is a global quality index sensitive to mean shifting and dynamic range change [134]. The lower the ERGAS value, specially a value lower than the number of bands, the higher the spectral quality of the merged images.

2. *Mean SSIM (MSSIM) index and the average quality index (Q_{avg}):* These indices [5, 131] are used to evaluate the overall image SSIM and UIQI quality, by averaging these measures. The higher, closer to one, the value the higher the spectral and radiometric quality of the merged images.

3. Another global measure, *Q4*, proposed in [135] depends on the individual UIQI of each band, but also on spectral distortion, embodied by the spectral angle SAM. The problem of this index is that it may not be extended to images with a number of bands greater than four.

Spatial Quality Assessment: To assess the spatial quality of a pansharpened image, its spatial detail information must be compared to the that present in the reference HR MS image. Just a few quantitative measures have been found in the literature to evaluate the spatial quality of merged images. Zhou [60] proposed the following procedure to estimate the spatial quality of the merged images: to compare the spatial information present in each band of these images with the spatial information present in the PAN image. First, the images to be compared is extracted using a Laplacian filter. Second, the correlation between these two filtered images is calculated thus obtaining the spatial correlation coefficient (SCC). However, the use of the PAN image as a reference is incorrect as demonstrated in [27, 136] and the HR MS image has to be used, as done by Otazu *et al.* in [5]. A high SCC indicates that many of the spatial detail information of one of the images is present in the other one. The SCC ideal value of each band of the merged image is 1.

Recently, a new spatial quality measure was suggested in [121], related to quantitative edge analysis. The authors claim that a good pansharpening technique should retain all the edges present in the PAN image in the sharpened image [121]. Thus, a Sobel edge operator is applied on the image in order to calculate its edges, and then compared with the edges of the PAN image. However, the concept behind this index is false since the reference image is not the PAN but the HRMS [136].

Additionally, some spectral quality measures have been adapted to spatial quality assessment. Pradhan *et al.* [121] suggested the use of structural information in SSIM measure between panchromatic and pansharpened images as a spatial quality measure. Lillo-Saavedra *et al.* [137] proposed to use the spatial ERGAS index, that includes in its definition the spatial RMSE calculated between each fused spectral band and the image obtained by adjusting the histogram of the original PAN image to the histogram of the corresponding band of the fused MS image. Again, since the PAN image is used as reference those measures should not be used [136].

Although the objective of this chapter is not the comparison of the different methods, as an example for these quality measures, Table 2.1 shows the figures of merit of some of the pansharpened images already presented in this chapter. We have highlighted the two best values for each measure. From the obtained results it is clear that IHS and HPF, depicted in Figs. 2(b) and 2(d), respectively, produce higher spectral distortions, with lower values of SSIM and MSSIM and high of ERGAS, the BT (Figure 2(c)) and Price (Figure 4(a)) methods produce lower spectral distortion but less spatial information is introduced. Methods based in Multiresolution approaches, GLP (Figure 6(a)), AW (Figure 6(b)) and the method described in [31] (Figure 6(f)), provide the best results with lower spectral distortion and high SCC values.

2.4.3 Quality Assessment without a Reference

Quantitative quality of data fusion methods can be provided when using reference images, usually obtained by degrading all available data to a coarser resolution and carrying out fusion from such data.

A set of global indices capable of measuring the quality of pansharpened MS images and working at full scale without performing any preliminary degradation of the data has been recently proposed.

The Quality with No Reference (QNR) index [138] comprises two indices, one pertaining to spectral distortion and the other to spatial distortion. As proposed in [138] and [139] the two distortions may be combined to yield a unique global quality measure. While the QNR measure proposed in [138] is based on the UIQI index, the one proposed in [139] is based on the measure of the Mutual Information (MI) between the different images. The spectral distortion index defined in [138] can be derived from the difference of inter-band UIQI values calculated among all the fused MS bands and among all the LR MS bands. The spatial distortion index defined in [138] is based on differences between the UIQI of each band of the HRMS image and the PAN image and the UIQI of each band of the LRMS image with a low resolution version of the PAN image. This LR image is obtained by filtering the PAN image with a lowpass filter with normalized frequency cutoff at the resolution ratio between MS and PAN images, followed by decimation. The QNR index is obtained by the combination of the spectral and spatial distortion indices into one single measure varying from zero to one. The maximum value is only obtained when there is no spectral and spatial distortion between the images. The main advantage of the proposed index is that, in spite of lack of a reference data set, the global quality of a fused image can be assessed at the full scale of the PAN image.

The QNR method proposed in [139] is based on the MI measure instead of UIQI. The mutual information between resampled original and fused MS bands is used to measure the spectral quality, while the mutual information between the PAN image and the fused bands yields a measure of spatial quality.

Another protocol was proposed by Khan *et al.* [91] to assess spectral and spatial quality at full scale. For assessing spectral quality, the MTF of each spectral channel is used to low-pass filter the HR MS image. This filtered image, once it has been decimated, will give a degraded low-resolution MS image. For comparing the degraded and original low-resolution MS images, the Q4 index [135] is used. Note that the MTF filters for each sensor are different and the exact filter response is not usually provided by the instrument manufacturers. However, the filter gain at Nyquist cutoff frequency may be derived from on-orbit measurements. Using this information and assuming that the frequency response of each filter is approximately Gaussian shaped, MTF filters for each sensor of each satellite can be estimated. To assess the spatial quality of the fused image, the high-pass complement of the MTF filters is used to extract the high-frequency information from the MS images at both high (fused) and low (original) resolutions. In addition, the PAN image is downscaled to the resolution of the original MS image and high-frequency information is extracted from high- and low-resolution PAN images. The UIQI value is calculated between the details of the each MS band and the details of the PAN image at the two resolutions. The average of the absolute differences in the UIQI values across scale of each band produces the spatial index.

2.5 Conclusion

In this chapter we have provided a complete overview of the different methods proposed in the literature to tackle the pansharpening prolblem and classified them into different categories according to the main technique they use. As previously described in sections 2.3.1 and 2.3.2, the classical CS and RSC methods provide pansharpened images of adequate quality for some applications but usually they introduce high spectral distortion. Their results highly depend on the correlation between each spectral band and the PAN image. A clear trend in the CS family, as we explained in section 2.3.1, is to use transformations of the MS image so that the transformed image mimics the PAN image. In this sense, a linear combination of the MS image is usually used by weighting each MS band with weights obtained either from the spectral response of the sensor or by minimizing, in the MMSE sense, the difference between the PAN and this linear combination. By using this techniques the spectral distortion can be significantly reduced. Another already mentioned important research area is the local analysis of the images, producing methods that inject structures in the pansharpened image depending on the local properties of the image.

The High-frequency injection family, described in section 2.3.3, can be considered the predecessor of the methods based on the ARSIS concept. HFI methods low-pass filter the image using different filters. As we have seen, the use the MTF of the sensor as the low-pass filter is preferable since, in our opinion, introducing sensor characteristics into the fusion rule will make the method more realistic.

The described methods based on the statistics of the image provide a flexible and powerful way to model the image capture system as well as incorporating the knowledge available about the HR MS image. Those methods allow, as explained in section 2.3.4, to accurately model the relationship between the HR MS image and the MS and PAN images incorporating the physics of the sensors (MTF, spectral response, or noise properties, for instance) and the conditions in which the images were taken. Still the models used are not very sophisticated thus presenting an open research area in this family.

The multiresolution analysis, as already mentioned, is one of the most successful approaches for the pansharpening problem. Most of those techniques have been previously classified into techniques relevant to the ARSIS concept. Decomposing the images at different resolution levels allows to inject the details of the PAN image into the MS one depending on the context. From the methods described in section 2.3.5, we can see that the generalized Laplacian pyramid and redundant wavelet and contourlet transforms are the most popular multiresolution techniques applied to this fusion problem. From our point of view, the combination of multiresolution analysis with techniques that take into account the physics of the capture process will provide prominent methods in the near future.

Finally, we want to stress, again, the importance of a good protocol to assess the quality of the pansharpened image. In this sense, Wald's protocol, described in section 2.4, is the most suitable assessment algorithm if no reference image is available. Besides visual inspection, numerical indices give a way to rank different methods and give an idea of their performance. Recent advances in full scale quality measures as the ones presented in section 2.4.3 set the trend for new measures specific for pansharpening that have to be considered.

Method	**Band**	**SCC**	**SSIM**	**MSSIM**	**ERGAS**
IHS [1]	B1	0.91	0.81	0.853	5.34
	B2	**0.96**	0.89		
	B3	**0.98**	0.86		
BT [50]	B1	0.63	**0.98**	**0.98**	1.76
	B2	0.80	**0.98**		
	B3	0.83	**0.98**		
HPF [49]	B1	**0.97**	0.53	0.627	16.14
	B2	0.95	0.70		
	B3	**0.97**	0.65		
	B4	**0.97**	0.63		
Price [95]	B1	0.32	**0.94**	0.927	**0.78**
	B2	0.79	0.91		
	B3	0.83	0.92		
	B4	0.44	**0.94**		
GLP [58]	B1	**0.96**	0.74	0.742	3.92
	B2	**0.98**	0.82		
	B3	**0.98**	0.81		
	B4	**0.97**	0.60		
AW [116]	B1	0.91	0.85	0.815	3.22
	B2	0.96	0.83		
	B3	**0.96**	0.85		
	B4	**0.95**	0.73		
Method in [31]	B1	**0.96**	0.90	**0.93**	**1.74**
	B2	**0.96**	**0.93**		
	B3	**0.97**	**0.97**		
	B4	0.92	**0.92**		

Table 2.1: Numerical results on the presented pansharpened images.

Chapter 3

Bayesian Formulation for Remote Sensing Image Pansharpening

3.1 Introduction

In this chapter we will discuss the Bayesian Formulation in order to provide solutions to the remote sensing pansharpening problem. Within the Bayesian Paradigm, pansharpening is understood as the process that generates the HRMS image (the image that the sensor would capture in the ideal conditions) from a set of degraded image observations (LRMS and PAN). This paradigm will allow us to perform the HRMS image reconstruction and the estimation of the associated parameters using a solid and flexible framework. To follow the Bayesian formulation, we will describe briefly the elements required for pansharpening problem, such as the conditional and prior probability distributions. Also there is a need to explain the low resolution and high resolution image formation models. After defining the elements of the Bayesian paradigm we will discuss the possible Bayesian solutions to perform the image reconstruction inference and the parameter estimation simultaneously.

The structure of this chapter as following: First in Section 3.2 we are going to study the different degradation sources that can take place in the image formation process and the image formation process for the problem of HRMS image reconstruction . Finally, in section 3.3 the Bayesian paradigm is described.

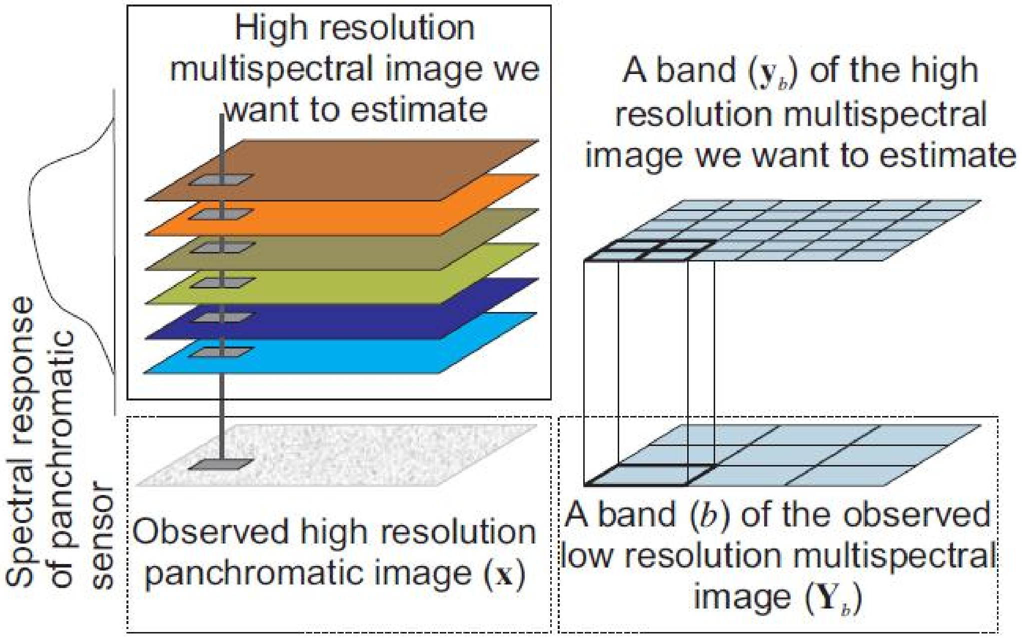


Figure 3.1: *Acquisition model*

3.2 Sources of Image Degradation

As we already mentioned in section 1.2, satellites usually provides us with two different images of the same scene. Let us assume that y, the unknown HRMS image of the scene we would have observed under ideal conditions with a HR sensor, has B bands, y_b, $b = 1,...,B$, each one centered on a narrow spectral band, of size $p = m \times n$. Satellites provide a LRMS image Y with the same spectral size that the HRMS image, that is, it has B bands Y_b, $b = 1,...,B$, but with lower spatial quality, that is, each band b has a size of $P = M \times N$ pixels, with $M < m$ and $N < n$. The sensor also provides us with a panchromatic image x of the size $p = m \times n$ of the HRMS image, that contains reflectance data in a single band that covers a wide area of the spectrum, that is it lacks of the spectral quality of the HRMS image.

For instance, the Landsat 7 satellite (http://landsat.gsfc.nasa.gov/), equipped with the ETM+ sensor, allows for the capture of a MS image with six bands (three bands on the visible spectrum plus three bands on the infrared) with a resolution of 30 meters per pixel, a thermal band with a resolution of 60 meters per pixel and a panchromatic band (covering a large zone on the visible spectrum and the near infrared), with a resolution of 15 meters per pixel. Figure 3.1 illustrates the acquisition model.

Having two images, we have to model the degradation model for both the observed LRMS and PAN images. Both of them are effected by blur and noise, but for PAN image the blur is removed in early re-processing stages [34], while this is not the case for LRMS image which also may be affected by interband component blur [140]. Also, the LRMS image is affected by a decimation process, since its resolution is lower than the one of the PAN image. In order to study the Bayesian reconstruction of HRMS image, it is necessary before understanding how the images in general, and the LRMS and PAN images in particular, are degraded. There are numerous situations in which an imaging system can contribute in the degradation of the quality of acquired images. We must remember that we are capturing a three-dimensional scene that resulted on a two-dimensional scene by using the capturing system. Since this capturing system is never perfect, both deterministic and stochastic (random) distortions may be introduced in general.

3.2.1 Deterministic Degradations

Deterministic degradations is related to the mode of image acquisitions, to possible defects of the imaging system (such as blur created by incorrect lens adjustment or by motion) or another phenomena such as atmospheric turbulence [141]. This degradation which introduced by the imaging process may be very complex for the following reasons [142]:

1. The 2-D imaging system may not be able to capture all 3-D scene, such as 3-D rotation of an object and 3-D geometry effects.
2. The complexity of the transfer function of the imaging system due to; diffraction effects in optical system, system aberrations, atmospheric turbulence, motion blurs and defocused systems. Usually these imperfections can be considered to be a bandwidth reduction (or blurring) of the captured scene. Furthermore, the severity of the bandwidth reduction may vary as a function of the image coordinates (spatially varying blur).
3. All kinds of nonlinearities due to the response of the sensor may further degrade the captured image.

Since modeling 3-D degradations is virtually impossible in an image reconstruction context, here we restrict the discussion on 2-D degradations. However, we can not discard the nonlinear response of the sensor, which sometimes can be approximated by a linear response (e.g. the logarithmic behavior of a photographic medium becomes linear when the image is of low contrast. If such a linearization is not possible, there are two ways to handle the nonlinearity [142]; mathematically by incorporating the nonlinear response into the image formation model, but such a model yields complex reconstruction algorithm, or by more practical approach which assumes that it is allowed to apply the inverse sensor response to the observed image before any further processing is done. Although the last way has several faults, it has been shown to be unexpectedly successful in the practice in related fields as image restoration [143, 144]. For remote sensing images, in this dissertation we are assuming that nonlinearities introduced by the sensor are removed in the pre-processing steps.

From the above discussion we learn that in many cases of practical interest it is useful to restrict the modeling of the image formation system to the usually dominant effect of blurring. The entire process then becomes a linear system characterized by a 2-D Point Spread Function (PSF). The observed image is given by the following 2-D superposition (or Fredholm) integral of the first kind [39, 142]:

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q(x,y;s,t) f(s,t) ds dt \tag{3.1}$$

where $f(x,y)$ and $g(x,y)$ are the original image and the observed image respectively, and they represent real-valued intensity distributions and take nonnegative values only. As a result, $q(x,y;s,t)$ is real-valued and nonnegative also.

Unfortunately, the model in Eq. (3.1) is not very useful for image identification and reconstruction purposes, because the complexity implied by the possibility of having a different PSF $q(x,y;s,t)$ at each coordinate (s,t) of the image is unacceptable from a computational viewpoint. Furthermore, it is rather unrealistic to assume that one might be able to estimate a different PSF for each location in the image simply because of the lack of sufficient information for the estimation procedures. Therefore, we have to assume that the PSF of the image formation process is stationary or spatially invariant over the image (or at least a significant portion of it), yielding to [39, 142],

$$\begin{aligned} g(x,y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q(x-s,y-t) f(s,t) ds dt \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q(s,t) f(x-s,y-t) ds dt \\ &= q(x,y) * f(x,y), \end{aligned} \tag{3.2}$$

where $*$ is used to denote 2-D convolution.

In order to perform digital image reconstruction, the discrete equivalent of Eq. (3.2) is required. The blurred image $g(x,y)$ is sampled on a 2-D regularly spaced lattice after appropriately restricting its bandwidth. The discrete observed image $g(i,j)$ is then given by [142]:

$$g(i,j) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} q(m,n)f(i-m,j-n), \tag{3.3}$$

where M and N are the numbers of row and columns of the image, respectively.

A convenient shorthand notation of Eq. (3.3) can be obtained by lexicographic ordering of the image raws and stacking the data into a vector, yielding [39, 142]

$$g = Qf, \tag{3.4}$$

where f and g are the lexicographically ordered vectors of size $MN \times 1$. If a circular convolution is assumed in Eq. (3.3), the blurring matrix Q (of size $MN \times MN$) has a block-circulant structure. The advantage of having a circular convolution in Eq. (3.3) is that the eigenvalues and eigenvectors of Q can be computed easily, and are in fact given by the coefficients of the discrete Fourier transform of $q(m,n)$ and the discrete Fourier basis functions, respectively. Therefore an alternative formulation of Eq. (3.3) and Eq. (3.4) is the frequency domain model

$$\mathscr{F}_g(u,v) = \mathscr{F}_q(u,v)\mathscr{F}_f(u,v), \tag{3.5}$$

where $\mathscr{F}$ means denotes the Fourier domain samples and u,v the discrete vertical and horizontal frequency variables.

In [39, 142] there are reviews of the most common PSF functions encountered in the problem of image reconstruction and restoration. A PSF function cannot take arbitrary values. In the observation model in Eq. (3.1) it was observed that the original and captured image are nonnegative real-valued quantities because of the physics of the underlying image formation process. As a consequence, the PSF function needs to be nonnegative and real-valued also [39, 142]. Moreover, the imperfections in an image formation system normally act as passive operations on the data (i.e. they don't absorb or generate energy). Consequently, all energy arising from a specific point in the original image should be preserved, that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q(s,t)\,ds\,dt = 1.0, \tag{3.6}$$

while a discrete PSF is constrained to satisfy,

$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} q(m,n) = 1.0. \tag{3.7}$$

Recently, the use of the modulated transfer function (MTF) of the sensor as the low-pass filter has been proposed in [91]. The MTF is the amplitude spectrum of the system point spread function (PSF) [92]. The MTF of a real imaging system is generally bell shaped, and its magnitude value at the cutoff Nyquist frequency is far lower than 0.5, to prevent aliasing. Furthermore, the MTFs of the MS sensors may be significantly different from one another in terms of decay rate, and especially are different from that of the PAN sensor. Hence models empirically optimized at a coarser scale on data degraded by means of digital filters that are close to be ideal, may yield little enhancement when are utilized at the finer scale [145]. Blur can also be spectral (see [140]), but in this dissertation we are not going to consider this case.

We can consider the satellite multispectral capture system as an image formation system composed of B image sensors, each one consisting of $M \times N$ sensing elements, where the size of each sensing element is $T \times T$. Each one of the sensors is equipped with a filter that only allows to capture light in a narrow spectral band. Then, the LRMS image will be composed of B spectral bands, where each one is a $M \times N$ discrete image with two-dimensional rectangularly sampled base interval $T \times T$.

On the other hand, the panchromatic capture system can be considered as a single image sensor consisting of $m \times n$ sensing elements, where the size of each sensing element is $T/L \times T/L$ where L is the ratio between the size of the pixels of the panchromatic and LRMS images. The panchromatic image sensor is able to capture light in a wide spectral band ranging from the blue to the near infrared. We can assume that both the multispectral and panchromatic image sensors are perfectly aligned so that they capture exactly the same area although with different resolutions. Note that the HRMS image grid is the same that the one of the panchromatic image.

With those considerations, we can write the deterministic part of the process to obtain Y_b, a low resolution image band of size $(M \times N) \times 1$, from y_b, the high resolution image band of size $(m \times n) \times 1$. Let us assume that C is the $(m \times n) \times (m \times n)$ integration matrix, that represents the way in which a set of pixels in the high resolution image affects each pixel in the low resolution image. C may have different forms; in [146, 147] it is associated with the blur function,

$$C(x,y) = C_1(x) \otimes C_1^t(y), \tag{3.8}$$

where $x = -m, \ldots, m$ and $y = -n, \ldots, n$, $\otimes$ denote the Kronecker product operator and t is the transpose operator, with

$$C_1(u) = \begin{cases} \frac{1}{2L} & u = -\frac{L}{2} \\ \frac{1}{L} & |u| < \frac{L}{2} \\ \frac{1}{2L} & u = \frac{L}{2} \\ 0 & otherwise. \end{cases} \tag{3.9}$$

In [100], C_1 is defined as

$$C_1(u) = \begin{cases} \frac{1}{L} & |u| < \frac{L}{2} \\ \frac{1}{L} & u = \frac{L}{2} \\ 0 & otherwise. \end{cases} \tag{3.10}$$

In [91] the integration function C is considered together to the blurring function into the sensor MTF and not as a separate function.

Let D_1 and D_2 to be the 1-D horizontal and vertical down-sampling matrices defined by,

$$D_1 = I_M \otimes e^t, \tag{3.11}$$

$$D_2 = I_N \otimes e^t, \tag{3.12}$$

where I_i is the $(i \times i), i = M, N$ identity matrix, e is the $(L \times 1)$ unit vector whose nonzero element is in the first position. Then, the 2-D downsampling matrix defined as

$$D = D_1 \otimes D_2, \tag{3.13}$$

is a $(m \times n) \times (M \times N)$ matrix that performs the downsampling of each HRMS image band to the LRMS image band.

As we have already commented, the panchromatic image has the same spatial resolution that the HRMS image but it lacks of its spectral resolution. In [100] this panchromatic image is modeled by spectrally averaging the unknown high resolution images y_b. The spectral averaging is controlled by a set of weights $\lambda_b \geq 0$, $b = 1, 2, \ldots, B$, that are known quantities that can be obtained from the sensor spectral characteristics or manually fixed. Note that, usually, the panchromatic image x does not depend on all the multispectral image bands but on a subset of them, i. e., some of the λ_b's are equal to zero. For example, for Landsat ETM+ images, the panchromatic image only covers the region from the end of band 1 to the end of band 4 and, so, the rest of the bands have no influence on the x. Figure 2.2 shows the spectral response covered by the observed low resolution and panchromatic Landsat 7 ETM+ bands (except the thermal band).

3.2.2 Stochastic Degradations

Beside the deterministic distortions described in the last sections, captured images are invariably degraded by stochastic degradations, usually referred to as *observation noise* or *adquisition noise*. It may originate from the image formation process, the transmission medium, the capturing process, quantization of the data or any combination of these [142]. Here, we will model the noise contributions as an additive zero-mean white Gaussian noise process with covariance matrix $\sigma_n^2 I$, which is statistically uncorrelated with the images,

$$n \sim N(0, \sigma_n^2 I), \tag{3.14}$$

where I is the identity matrix. This is a simplification since some types of noise are not uncorrelated with the input and may even be non-additive. This simplification nonetheless leads to reconstruction methods which can be applied to a wide class of problems.

Denoting the noise contribution to each band b of LRMS image by n_b, the complete image formation model becomes, for each one of the LRMS image bands, Y_b, written as matrix-vector [100]

$$Y_b = H y_b + n_b, \tag{3.15}$$

where $H = DCQ$ is a $(m \times n) \times (M \times N)$ that synthesizes the blurring, integration and downsampling processes, and the image formation model for the panchromatic image, x, becomes,

$$x = \sum_{b=1}^{B} \lambda_b y_b + n, \tag{3.16}$$

where n denotes the acquisition noise.

3.3 Hierarchical Bayesian Paradigm

The hierarchical Bayesian paradigm has been used in many areas related to image analysis. Some examples are the construction of classification trees [148], construction [149, 150] or the likelihood refinement network [151] and neural networks with back-propagation [152, 153]. This paradigm has also been applied to interpolation problems [154, 155], reconstruction of tomographic images [156, 157], reconstruction of compressed images [158] and restoration problems [143, 144], even when the blur which presents in the image is partially known [159].

The Bayesian modeling of a problem is based on a statistical approach and it is related to the decision theory in the presence of statistical knowledge that could highlight some uncertainties involved in the problem decision. Classical statistics is directed towards the use of the information obtained from the data of the statistical research to make inferences about the unknown data. The decision theory, on the other hand, attempts to combine information about this data with other relevant aspects of the problem to make better decisions.

The approach known as Bayesian inference can be distinguished from other statistical approaches for the use of prior information on the images that we are working on. In the case of fusion of a LRMS image and a PAN image we are dealing with, this information usually contains a priori spatial and spectral information about the properties of the image, for instance smoothness of the luminosities in the image. This approach has wide application in the field of the Astronomy image restoration [160–163] and remote sensing pansharpening (see, for example, [55, 56, 96, 98, 100, 101, 103], and references therein).

3.3.1 Bayesian Paradigm Description

To follow the Bayesian paradigm in general, we distinguish between $f = y$, the original high resolution image we would have observed under ideal conditions, and $g = \{Y, x\}$, the set of the observed low resolution MS image and the PAN image, as described in section 3.2. Our objective is to reconstruct f, the original image, from g, the set of the observed images. The Bayesian approach start with a priori distribution, $p(f|\alpha)$, the image model, over the possible reconstructions where information on the expected structure within an image is incorporated. This approach also specifies the degradation model from the original, f, to the observed images, g, by the conditional distribution $p(g|f,\beta)$. Both the prior and the conditional distributions depend on a number of parameters, α and β, or vectors of hyper-parameters, that must be set or estimated from the data. The hierarchical Bayesian paradigm combines the information from the observed image data obtained from the capture system with a priori information to construct the posterior distribution of f given g,

$$p(f|g,\alpha,\beta) = \frac{p(\alpha,\beta,g,f)}{P(f)}, \tag{3.17}$$

from which we can take decisions and develop inferences [143].

In the hierarchical approach to image reconstruction we have at least two stages. In the first stage, we have knowledge about the structural form of the degradation process and local characteristics of the original image. These degradation and image models depend on unknown hyperparameters β and α, respectively. In the second stage, the hierarchical Bayesian approach defines a hyperprior on β and α, where information about these hyperparameters is included.

Although in some cases it would be possible to know relations between the hyperparameters, we shall study here the model where the global probability is defined as [144]

$$p(\alpha,\beta,f,g) = p(\alpha)p(\beta)p(f|\alpha)p(g|f,\beta). \tag{3.18}$$

3.3.2 Hierarchical Bayesian Analysis

Once we have defined the elements needed to perform the analysis based on the hierarchical Bayesian paradigm $p(\alpha,\beta,f,g)$, we can perform the analysis in a set of different ways; the evidence and maximum a posteriori (MAP) analysis.

Evidence Analysis

In this approach $\hat{\alpha}$ and $\hat{\beta}$ are first selected as [144],

$$\begin{aligned}\hat{\alpha},\hat{\beta} &= \arg\max_{\alpha,\beta} p(\alpha,\beta|g),\\ &= \arg\max_{\alpha,\beta} \int_y p(\alpha,\beta,f,g)df,\end{aligned} \tag{3.19}$$

then, $y_{(\hat{\alpha},\hat{\beta})}$, the restored image is defined as,

$$f_{(\hat{\alpha},\hat{\beta})} = \arg\max_f p(f|g,\hat{\alpha},\hat{\beta}). \tag{3.20}$$

MAP Analysis

This analysis, suggested in [143, 152], used to estimate the image and the hyperparameters simultaneously by integrating $p(f|g,\hat{\alpha},\hat{\beta})$ over the hyperparameters, α and β, to obtain the real likelihood and maximize this likelihood in f. The process to estimate the image f and the hyperparameters α,β starting by the image estimation as,

$$\hat{f} = \arg\max_f p(f,g), \tag{3.21}$$

$$= \arg\max_f \int_\alpha \int_\beta p(\alpha,\beta,f,g)d\alpha d\beta, \tag{3.22}$$

then estimating the hyperparameters as,

$$\hat{\alpha},\hat{\beta} = \arg\max_{\alpha,\beta} p(\alpha,\beta|\hat{f},g). \tag{3.23}$$

Note that we are not really concerned with the estimation of α and β in the MAP analysis and that the above equations can be understood as intermediate steps to calculate $\hat{f}$.

In [143,144] there are a study in depth of both approaches where it is found that the evidence approach provides better results than the MAP in the related problem of restoration of images. So we will follow the evidence approach when possible.

The Variational Approach to the Bayesian Analysis

The inference procedures aim at optimizing a given function and not at obtaining posterior distributions that can be analyzed or simulated to obtain additional information about the quality of the estimates. Instead of having a distribution over all possible values of the parameters and the image, the above inference procedures choose a specific set of values. This means that we have neglected many other interpretations of the data. If the posterior is sharply peaked, other values of the hyperparameters and the image will have a much lower posterior probability but, if the posterior is broad, choosing a unique value will neglect many other choices of them with similar posterior probabilities.

Variational methods provide a way to approximate the posterior $p(\alpha,\beta,g|f)$ by a simpler distribution, from which it is easier, to extract observations. See the very interesting [164, 165], books [166, 167] and book chapter [168] for a comprehensive introduction to variational methods.

The last few years have seen a growing interest in the application of variational methods [164,168] to inference problems. These methods attempt to approximate posterior distributions with the use of the Kullback-Leibler cross-entropy [169]. Application of variational methods to Bayesian inference problems include graphical models and neural networks [168], independent component analysis [164], mixtures of factor analyzers, linear dynamic systems, hidden Markov models [165], support vector machines [170] and blind deconvolution problems (see [171,172]).

As previously stated, inference on (α,β,f) should be based on $p(\alpha,\beta,f|g)$. In some cases $p(\alpha,\beta,f|g)$ can not be found in closed form, since

$$p(f) = \int\int\int p(\alpha,\beta,g,f)dg\,d\alpha\,d\beta \tag{3.24}$$

cannot be calculated analytically. The variational methods approximate this distribution by the distribution $q(\alpha,\beta,f)$, which is similar to the true posterior but for which it is tractable to perform the marginalization required for the inference.

The variational criterion used to find $q(\alpha,\beta,f)$ is the minimization of the Kullback-Leibler divergence, given by [169, 173],

$$\begin{aligned} C_{KL}(q(\alpha,\beta,f)||p(\alpha,\beta,f|g)) &= \int_{\alpha,\beta}\int_f q(\alpha,\beta,f)\log\left(\frac{q(\alpha,\beta,f)}{p(\alpha,\beta,f|g)}\right)d\alpha d\beta df \\ &= \int_{\alpha,\beta}\int_f q(\alpha,\beta,f)\log\left(\frac{q(\alpha,\beta,f)}{p(\alpha,\beta,f,g)}\right)d\alpha d\beta df + const, \end{aligned} \tag{3.25}$$

which is always non negative and equal to zero only when $q(\alpha,\beta,f) = p(\alpha,\beta,f|g)$. We choose to approximate the posterior distribution $p(\alpha,\beta,f|g)$ by the distribution

$$q(\alpha,\beta,f) = q(\alpha)q(\beta)q(f) = q(\alpha,\beta)q(f), \tag{3.26}$$

where $q(f)$, $q(\alpha)$ and $q(\beta)$ denote distributions on f,α and β, respectively. We now proceed to find the best of these distributions in the divergence sense. Using Eq. (3.26) we have in Eq. (3.25),

$$\begin{aligned} C_{KL}(q(\alpha,\beta,f)||p(\alpha,\beta,f|g)) &= \int_{\alpha,\beta} q(\alpha,\beta)\left(\int_f q(f)\log\left(\frac{q(\alpha,\beta)q(f)}{p(\alpha,\beta,f,g)}\right)df\right)d\alpha d\beta + const \\ &= \int_f q(f)\left(\int_{\alpha,\beta} q(\alpha,\beta)\log\left(\frac{q(\alpha,\beta)q(f)}{p(\alpha,\beta,f,g)}\right)d\alpha d\beta\right)df + const. \end{aligned}$$

Now, given $\hat{q}(\alpha,\beta)$, an estimate for $q(\alpha,\beta)$, we can obtain as estimate of $q(f)$ by solving,

$$\hat{q}(f) = \arg\min_{q(f)} C_{KL}((\hat{q}(\alpha,\beta)q(f))||p(\alpha,\beta,f|g)), \tag{3.27}$$

and given $\hat{q}(f)$, an estimate of $q(\alpha,\beta)$, we can obtain an estimate of $q(\alpha,\beta)$ by,

$$\hat{q}(\alpha,\beta) = \arg\min_{q(\alpha,\beta)} C_{KL}((q(\alpha,\beta)\hat{q}(f))||p(\alpha,\beta,f|g)). \tag{3.28}$$

Chapter 4

Multispectral image pansharpening based on the contourlet transform [1]

4.1 Introduction

In the literature many pansharpening methods have been proposed for combining PAN with MS image (see Chapter 2 for a detailed review). Among them, methods such as IHS [49] and PCS [51, 60] provide superior visual high-resolution multispectral images but ignore the requirement of high-quality synthesis of spectral information. More recently, an underlying multi-resolution analysis employing the discrete wavelet transform has been used in image fusion. Properties, such as multiresolution, localization, critical sampling, and limited directionality (horizontal, vertical, and diagonal directions) have made the wavelet transform a popular choice for feature extraction, image denoising, and pansharpening. However, wavelets fail to capture the smoothness along the contours [174]. The contourlet transform, an alternative multiresolution approach, provides an efficient directional representation and also efficient in capturing intrinsic geometrical structures of the natural image along the smooth contours [111]. Remote sensing images have presence of natural and man-made objects, e.g., rivers, roads, coastal areas, buildings, etc. which indicate higher geometrical content. Thus, the transformations taking in consideration the geometric structure along with other properties of wavelet transformation will be more useful for pansharpening.

---

[1] Part of this chapter was published as a work in the "8th International Workshop on Information Optics (WIO'09)" [28] and as a chapter of the book Advances in Information Optics and Photonics [29].

In most of the proposed methods no explicit physical information about the detection system has been taken into account. However, a new technique proposed in [5], WiSpeR, was used to define a wavelet-based fusion method which does incorporate information from the spectral response of the sensor in each band of the low-resolution MS image (LRMS). This prior knowledge is employed in the transformation model which injects spatial detail into the LRMS image. In this chapter we will briefly describe the wavelet based pansharpening algorithms and the contourlet based algorithms sharing the same ideas, propose a new method that cast the WiSpeR method defined in [5] using the contourlet transform and we compare it with some of the most popular methods for pansharpening described in the literature.

The chapter is organized as follows. Section 4.2 provides a short explanation of the contourlet transform. Section 4.3 describes the contourlet based pansharpening and the proposed algorithm. Experimental results and comparisons are presented in Section 4.4 for different datasets and, finally, Section 4.5 concludes the chapter.

4.2 Contourlet transform

Contourlets provide a new system representation for image analysis. The contourlet transform is so called because of its ability to capture and link the point of discontinuities to form a linear structure (contours). In general, the ideal image representations have to have the following features: multiresolution, localization, critical sampling, directionality, and anisotropy [111]. Among these features, the first three are successfully provided by separable wavelets, and the result of contourlet transform offers a high degree of directionality and anisotropy. With a rich set of basis functions, contourlets represent a smooth contour with fewer coefficients compared with discrete wavelets. As the resolution becomes finer, the limitation of discrete wavelets is that it needs many fine ”dots” to capture the contour. However, contourlets effectively explore the smoothness of the contour by different elongated shapes and in a variety of directions following the contour [111].

The two-stage process used to derive the contourlet coefficients involves a multiscale transform and a local directional transform. The point of discontinuities and multiscale transformation is obtained via the Laplacian pyramid. The local directional filter bank is used to group these wavelet-like coefficients to obtain a smooth contour. Contourlets provide $2l$ directions at each scale, where l is the number of required orientation. This flexibility of having different

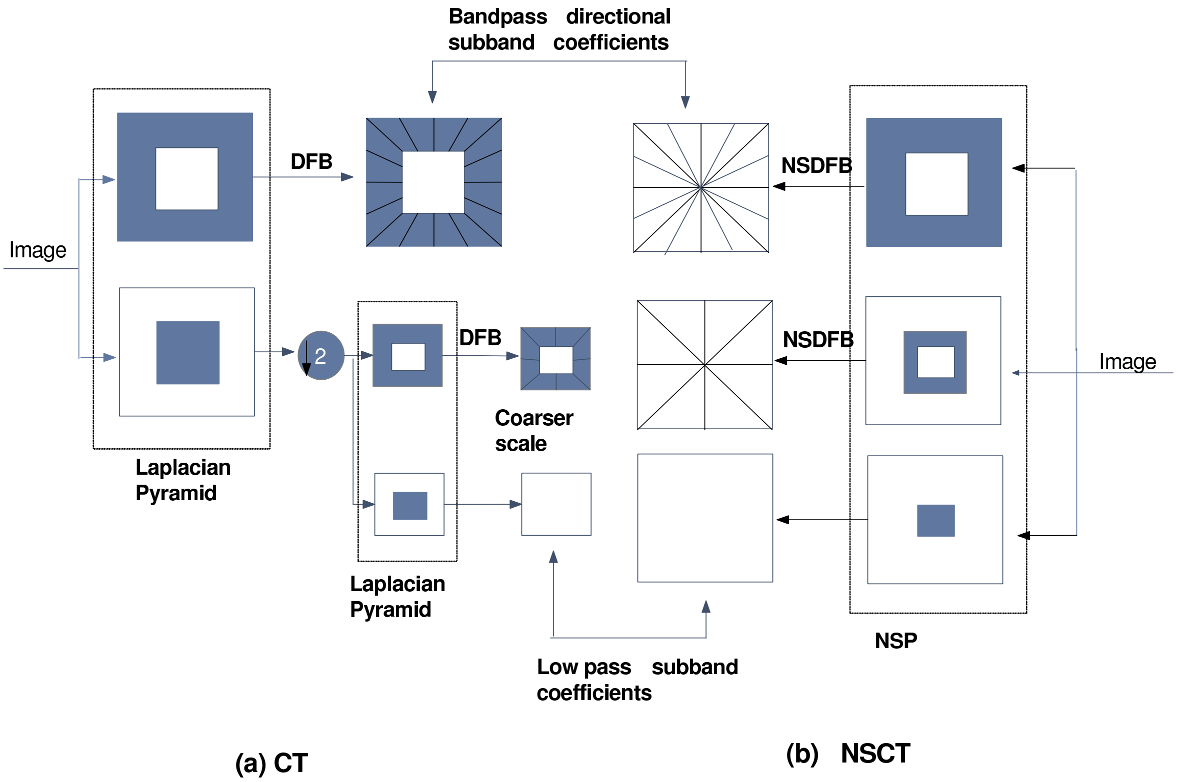


Figure 4.1: *Discrete contourlet transform: a) Subsampled Contourlet Transform. b) Non-subsampled Contourlet Transform*

numbers of direction at each scale makes contourlets different from other available multiscale and directional image representation [72]. Similarly to wavelets, both subsampled and non-subsampled transforms have been defined for contourlets:

1. *Discrete Contourlet Transform (CT)*: The discrete CT is developed in the discrete domain using the fast iterated nonseparable filter banks having an order of N operations for N-pixel images. The transformation stage includes two filter banks: the Laplacian pyramid to generate multiscale decomposition and the directional filter bank (DFB) to reveal directional details at each decomposition level [111] as illustrated in Figure 4.1(a). Similarly to the discrete wavelet transform, the discrete contourlet transform is also shift variant.

2. *Nonsubsampled Contourlet Transform (NSCT)*: The NSCT provides a complete shift-invariant and multiscale representation, similar to the redundant wavelet transform [61],

with a fast implementation. The building block of the NSCT is the 2-D two-channel nonsubsampled filter banks (NSFBs). The NSCT is also obtained via a two stage non shift-invariant process [61] as depicted in Figure 4.1(b). The first part achieves the multiscale property, via the nonsubsampled pyramid (NSP) subband decomposition, while the second part provides directionality information using nonsubsampled directional filter bank (NSDFB). Both stages of the NSCT are constructed to be invertible in order to have an overall invertible system.

4.3 Wavelet and Contourlet-based pansharpening

A number of pansharpening methods using the wavelet and, more recently, the contourlet transform has been proposed. In general, all the transform based fusion methods consist of three stages. The first stage provides a sub-band and directional decomposition by the application of the subsampled or non-subsampled wavelet or contourlet transform to the PAN and MS images. It is followed by the application of various fusion rules onto the transform coefficients. These fusion rules usually comprise, for instance, substituting the original MS coefficient bands by the coefficients of the PAN image or adding the coefficients of the PAN to the coefficients of the original MS bands weighted sometimes, as for the method we propose in this chapter, by a factor related with the contribution of the PAN image to each MS band. The fusion schemes ends with the inverse transform.

This chapter summarizes the most important wavelet-based pansharpening methods, and compares them with the existing contourlet-based ones and the new proposed contourlet method. Since contourlet-based and wavelet-based methods share same stages for similar methods, except the transform type (contourlet or wavelet), let us to describe them together.

4.3.1 Additive Wavelet/Contourlet

The steps for fusing MS and PAN images using the additive wavelet [116] /contourlet [62] method are:

1. Register the LRMS image to the same size as the PAN image in order to be superimposed.

2. For each band of the MS image taken into account, generate a new panchromatic image

which histogram match that of the MS image using, for instance, [41]

$$PAN_k = \left(PAN - \mu_{PAN}\right)\frac{\sigma_{b_k}}{\sigma_{PAN}} + \mu_{b_k}, \tag{4.1}$$

where μ_{PAN}, μ_{b_k} are the mean of the PAN and the MS band b_k, respectively, $k \in B$, and B is the set of bands we are interested in. σ_{PAN} and σ_{b_k} are the standard deviation of PAN and MS band b_k, respectively.

3. Apply the wavelet/contourlet transform to each histogram-matched panchromatic images. Repeat the same transform to each MS band.

$$PAN_k = \bigcup_{i=1}^{n} wPAN_k^i \cup PAN_k^r, \tag{4.2}$$

$$b_k = \bigcup_{i=1}^{n} wb_k^i \cup b_k^r, \ \forall k \in B \tag{4.3}$$

where $wPAN_k^i$ and wb_k^i are the wavelet/contourlet coefficients for PAN and MS bands, respectively, PAN_k^r and wb_k^r are the residual (low pass filtered version of original) images of PAN and MS bands, respectively, n is the wavelet/contourlet resolution levels, usually $n = 2$ or 3. The $\cup$ operator means the composition operator that merges the different wavelet/contourlet bands since each band may have different resolution. Note that in the non subsampled case, this operator means just adding the different bands.

4. Introduce the details of the panchromatic image into each MS band adding the wavelet/contourlet coefficients of the panchromatic image to those of the MS image

$$b_k^{coef} = \bigcup_{i=1}^{n} (wPAN_k^i + wb_k^i), \tag{4.4}$$

where b_k^{coef} is the new wavelet/contourlet coefficients of the MS band k.

5. Apply the inverse wavelet/contourlet transform to each MS transformed band

$$b_k^{new} = b_k^{coef} \cup b_k^r \tag{4.5}$$

to obtain b_k^{new}, the pansharpened MS band k, $k \in B$. Note that since for the undecimated case, $\sum_{i=1}^{n} wb_k{}^i + b_k^r = b_k$, we don't need to decompose the MS image and we can add b_k to the corresponding PAN coefficients.

4.3.2 Substitutive wavelet/contourlet

The substitute wavelet/contourlet methods are quite similar to the additive ones but, instead of adding the information of the panchromatic image to each band of the MS image, the pansharpening method simply replaces the MS detail bands with the details obtained by the panchromatic image following these steps for wavelet [119] and contourlet [122] reconstruction:

1. Perform the first 3 steps of the algorithm described in section 4.3.1.
2. For each MS band k, perform the inverse wavelet/contourlet transform to the transformed image formed by the wavelet/contourlet coeficient planes of the histogram-matched PAN image $wPAN_k^i, i = 1...n$, and the residual band of the MS image k, that is,

$$b_k^{new} = \bigcup_{i=1}^{n} wPAN_k^i \cup b_k^r. \tag{4.6}$$

4.3.3 IHS wavelet/contourlet

One of the most popular image pansharpening methods are those based on the IHS transformation. The main drawback of these methods is the high distortion of the original spectral information that the resulting MS images present. To avoid this problem, the IHS transformation is followed by the wavelet or countourlet transform to take advantage of the multiresolution property of this transform. Another disadvantage of the IHS based method, that cannot be solved by these transforms, is that they can only work with three bands due to the IHS transformation.

Additive IHS

In order to perform the wavelet [116] and contourlet [123, 175] additive IHS pansharpening, the following steps are followed:

1. Register the LRMS image to the same size as the PAN image in order to be superimposed.
2. Apply the IHS transform to the RGB composition of three MS image, using,

$$\begin{pmatrix} I \\ v1 \\ v2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & 0 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}, \tag{4.7}$$

$$H = \tan^{-1}[v2/v1], S = \sqrt{v1^2 + v2^2}.$$

3. Perform histogram matching between the panchromatic image and the intensity component of the IHS image using Eq. (4.1) to obtain PAN_I, the PAN image histogram-matched to the I band.

4. Apply wavelet/contourlet decomposition algorithm to the I band of the IHS image and to the 'histogram-matched' PAN one using,

$$I = \bigcup_{i=1}^{n} wI^i \cup I^r, \tag{4.8}$$

$$PAN_I = \bigcup_{i=1}^{n} wPAN_I^i \cup PAN_I^r. \tag{4.9}$$

5. Generate the wavelet/contourlets coefficients of the pansharpened intensity image as the sum of the wavelet/contourlet coefficients of the initial intensity and the histogram-matched PAN image,

$$I^{coef} = \bigcup_{i=1}^{n} (wPAN_I^i + wI^i). \tag{4.10}$$

6. Apply the inverse wavelet/contourlet transform to reconstruct new intensity image I^{new},

$$I^{new} = I^{coef} \cup I^r. \tag{4.11}$$

   Note that, as already happened in the additive wavelet case, since for the undecimated case $\sum_{i=1}^{n} wI^i + I^r = I$, we do not need to decompose the I image.

7. Insert the spatial information of the panchromatic image into the MS one, by applying the inverse IHS transform,

$$\begin{pmatrix} R^{new} \\ G^{new} \\ B^{new} \end{pmatrix} = \begin{pmatrix} 1 & \frac{-1}{\sqrt{6}} & \frac{3}{\sqrt{6}} \\ 1 & \frac{-1}{\sqrt{6}} & \frac{-3}{\sqrt{6}} \\ 1 & \frac{2}{\sqrt{6}} & 0 \end{pmatrix} \begin{pmatrix} I^{new} \\ v1 \\ v2 \end{pmatrix}. \tag{4.12}$$

Substitutive IHS

In order to perform the wavelet [125] and contourlet substitutive IHS pansharpening, the following steps are followed:

1. Perform the first 4 steps of the section 4.3.3.

2. Perform the inverse wavelet/contourlet transform to the wavelet/contourlet image formed by substituting the wavelet/contourlet coefficient planes of the intensity image with the corresponding wavelet/contourlet planes of the histogram-matched PAN image,

$$I^{new} = \bigcup_{i=1}^{n} wPAN_I^i \cup I^r. \tag{4.13}$$

3. Apply the inverse IHS transform using Eq. (4.12).

4.3.4 PCA wavelet/contourlet

The PCA-based method has been popularly used for spectral transformation because the first principal component (PC1) consists of the most variance, making it a suitable choice to replace the PAN component. Like IHS, the main drawback of this method is the high distortion of the original spectral information that the resulting MS images may present. To overcome this problem, Gonzalez et al. [90] proposed a pansharpening method based the PCA and wavelets methods where only the details of PC1 are replaced by the details of the PAN image.

The steps we need to pansharpen an image using the PCA wavelet [116] and contourlet [72] methods are the following:

1. Register the LRMS image to the same size as the PAN image in order to be superimposed.
2. Apply the PCA transformation to the MS image to obtain the PC1 image.
3. Histogram match the PAN image to the PC1 image.
4. Apply a subsampled or non-subsampled wavelet or contourlet transformation to the PC1 image and the histogram matched PAN image.
5. Replace the detail wavelet or contourlet coefficients of PC1 with the detail wavelet or contourlet coefficients of the histogram-matched PAN image.
6. Perform inverse wavelet/contourlet transformation and inverse PCA transformation to obtain a PAN image.

4.3.5 WiSper/ CiSper

The WiSpeR method can be understood as a generalization of different wavelet-based image fusion methods [5]. It uses a modification of the non-subsampled additive wavelet algorithm where the contribution from the PAN image to each of fused bands depends on a factor generated both from the sensor spectral response and physical properties of the observed object.

The steps for merging MS and PAN images using WiSper method are as follows:

1. Register the LRMS image to the same size as the PAN image in order to be superimposed.

2. Generate new panchromatic images, whose histograms match those of each band of the MS image, using Eq. (4.1).

3. Perform the n undecimated wavelet planes transform only on the panchromatic images, using Eq. (4.2).

4. Calculate the spectral factor λ_k, related to b_k bands [5], where k is the band number.

5. Add the wavelet planes of the panchromatic decomposition to each band of the MS dataset, as the following:

   $$b_k^{new} = b_k + \lambda_k \sum_{i=1}^{n} wPAN_k^i, \tag{4.14}$$

   where b_k^{new} is the fused band k, $k \in B$, and B is the number of bands we take into account.

We proposed a new contourlet panshapening method, named CiSper, that, similarly to WiSpeR, depends on a spectral factor to determine the amount of spatial detail of the PAN image that has to be injected into each MS band but it uses the non-subsampling contourlet transform with necessary filters, and n resolution levels and m directions in each level. In order to apply these methods, we need to know the contribution of each band to the panchromatic image, that is, the values of $\lambda_k, k = 1,2,...,B$ in Eq. 4.14. These values can be obtained from the spectral response of the imagery sensor (see Figure 2.2 for the Landsat 7 ETM+ spectral response). Note that the panchromatic image covers a region of wavelengths from almost the end of band 1 to the end of band 4 and that the sensor sensibility is not constant over the whole range. Taking into account these considerations, we obtain values for $\lambda_k, k = 1,2,3,4$ by summing up the spectral response of the panchromatic sensor weighted by the response of the sensor for each multispectral band. Following [5],λ_k takes into the account the residual image of the used decomposition transform. The obtained values are then normalized so that their sum equals one, thus producing the values of λ_k displayed on Table 4.1 for the Landsat 7 ETM+ imagery.

λ_1	λ_2	λ_3	λ_4
0.2568	0.2366	0.2021	0.5263

Table 4.1: Estimated values for $\lambda_k, k = 1,2,3,4$ for the Landsat 7 ETM+ imagery.

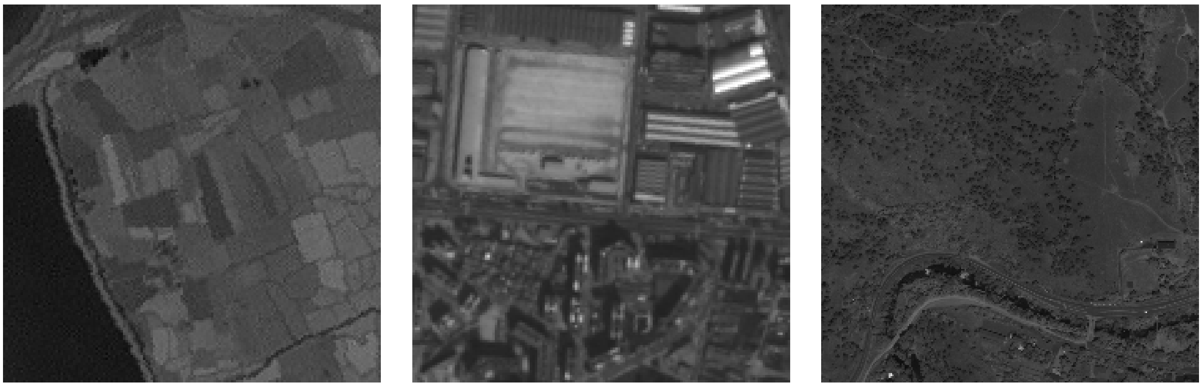

(a) Landsat 7 ETM+ (b) SPOT 5 (c) QuickBird

Figure 4.2: *Panchromatic Imagery of the Dataset*

4.4 Experimental Results

The contourlet-based and wavelet-based methods mentioned in section 4.3 are evaluated by performing pansharpening on dataset acquired by SPOT5, Landsat 7 ETM+ and QuickBird satellites. The MS and PAN images are co-registered for each dataset. The panchromatic images, for both SPOT5, Landsat 7 ETM+ and QuickBird datasets, are shown in Figure 4.2.

In the Landsat 7 ETM+ Dataset, we chose a region of interest of the MS image of 256 by 256 pixels with a pixel resolution of 28.5 m, and a region of interest of the PAN image of 512 by 512 pixels, with a pixel resolution of 14.25 m. The MS image consists of the following six bands from the visible and infrared (IR) region: blue (0.45-0.515 μm), green (0.525-0.605 μm), red (0.63-0.690 μm), Near IR (0.75-0.90 μm), Mid IR (1.55-1.75 μm), and Mid IR (2.09-2.35 μm), while the PAN image consists of a single band covering the visible and Near IR (0.52-0.90 μm). The scene, depicted in Figure 4.3(a), was acquired over The Netherlands on May 13, 2000.

The MS image in SPOT5 dataset covers a region of interest of 80 by 80 pixels with a pixel resolution of 10 m, while the PAN image is 160 by 160 pixels with a pixel resolution of 5 m. The MS image consists of four bands from the visible and infrared region corresponding to green (0.50-0.59 μm), red (0.61-0.68 μm), Near IR (0.78-0.89 μm), Mid IR(1.58-1.75 μm), while the PAN image consists of a single band covering the visible and NIR (0.48-0.71 μm). The

(a) MS (R-G-B) (b) ANW (c) ANC

(d) I-ANW (e) I-ANC (f) PCA-NW

(g) PCA-NC (h) WiSpeR (i) CiSpeR

Figure 4.3: *(a) Low resolution image formed from the R-G-B bands of the MS LandSat image. (b)–(i) Pansharpened images using the methods under study.*

scene, depicted in Figure 4.4(a), was acquired over Sevilla (Spain) on February 15, 2003. The $\lambda_k, k = 1,2,3,4$ values for the SPOT 5 imagery are displayed on Table 4.2.

λ_1	λ_2	λ_3	λ_4
0.44	0.56	0	0

Table 4.2: Estimated values for $\lambda_k, k = 1,2,3,4$ for the SPOT 5 imagery.

The MS image in QuickBird dataset covers a region of interest of 60 by 60 pixels with pixel resolution of 2.44 m, while the PAN image is 240 by 240 pixels with a pixel resolution of 61 cm to 72 cm. The $\lambda_k, k = 1,2,3,4$ values for the QuickBird imagery are displayed on Table 4.3.

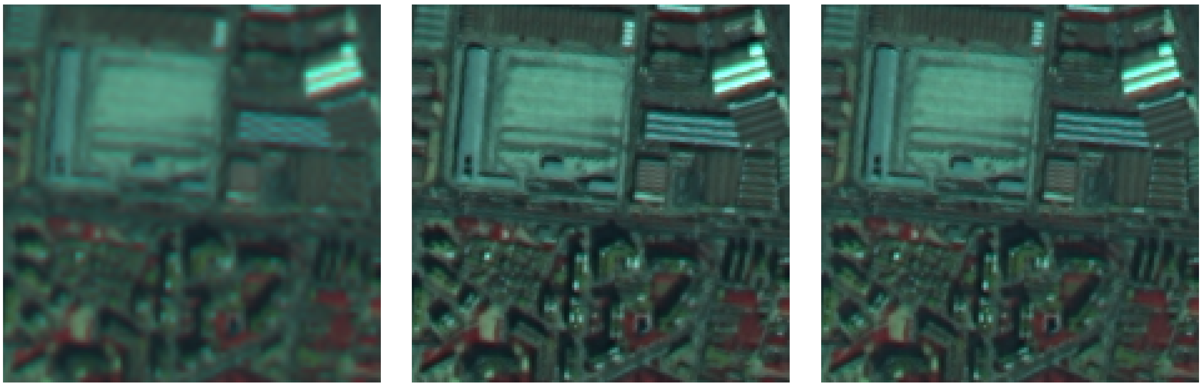
λ_1	λ_2	λ_3	λ_4
0.1472	0.2512	0.1639	0.4378

Table 4.3: Estimated values for $\lambda_k, k = 1,2,3,4$ for the QuickBird imagery.

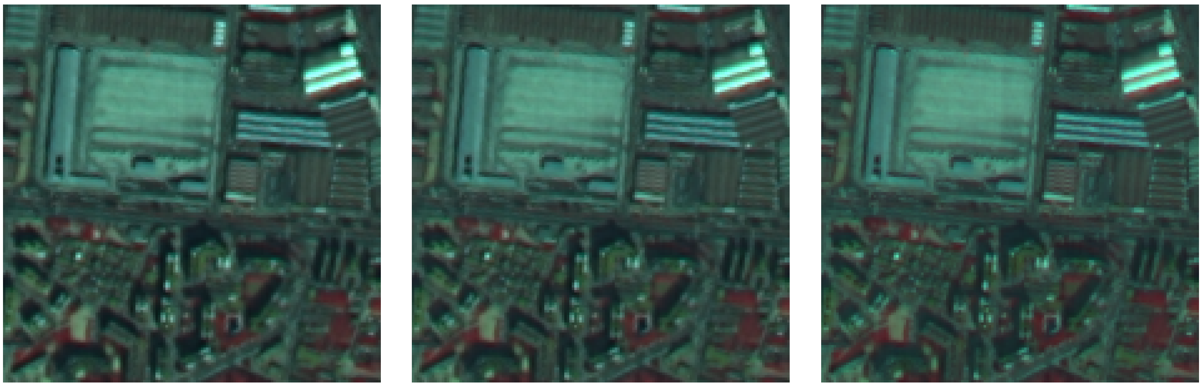
Pansharpening results are evaluated visually and numerically using some well known global quality indexes. These indexes are explained before in Chapter 2, here we are using the COR for assessing spatial quality and, UIQI and ERGAS for assessing spectral quality.

Since more than twenty methods for pansharpening have been presented, we compare in this chapter the pansharpening methods from the discussed above that, from our point of view, are the most significant. In our preliminary experiments we have realized that the non-subsampled decompositions always provide better results than their subsampled counterpart so we are going to center on the non-subsampled approaches. For the Landsat 7 and SPOT 5 experiments we used 3 levels of non-subsampled wavelet/countourlet decomposition, while for the QuickBird experiment we found that there is a need to more decomposition levels, since the spatial resolution ratio between QuickBird MS and PAN images is 1:4. Also, we realized that, usually, the additive methods performs better than the substitute ones so, in this chapter, we are going to compare the following eight methods: Additive Non-subsampled Wavelet (ANW), IHS Additive Non-subsampled Wavelet (I-ANW), PCA Non-subsampled Wavelet (PCA-NW), Additive Non-subsampled Contourlet (ANC), IHS Additive Non-subsampled Contourlet (I-ANC), PCA Non-subsampled Contourlet (PCA-NC), WiSpeR, and CiSpeR.

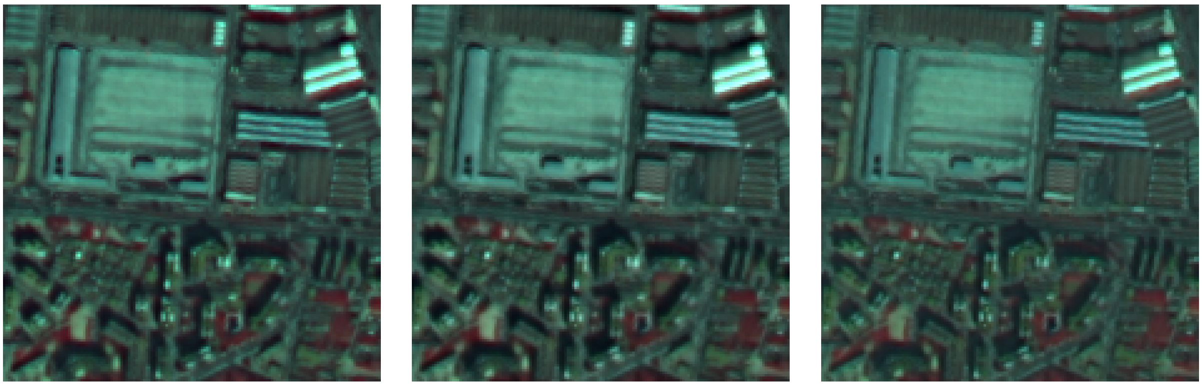
The pansharpened images resulted from using the wavelet/contourlet methods under study are presented in Figure 4.3 for Landsat 7 dataset, in Figure 4.4 for SPOT 5 dataset and in



(a) MS(G-R-NIR) (b) ANW (c) ANC

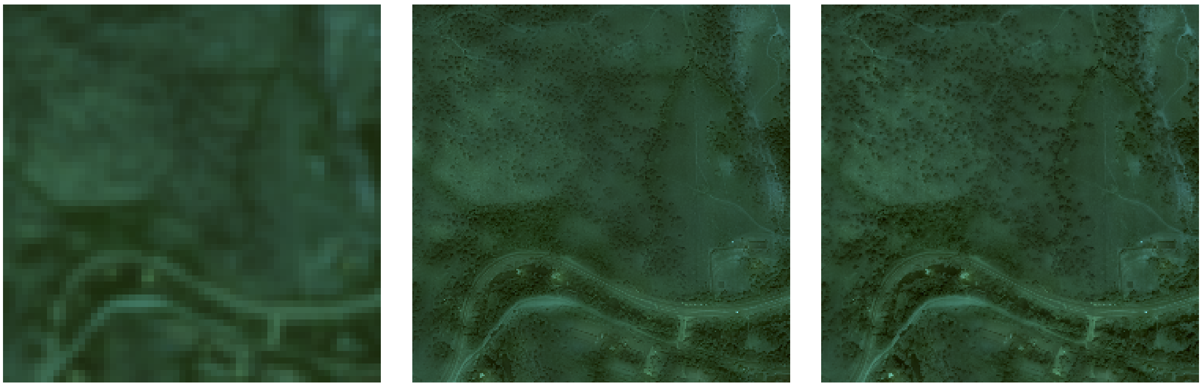


(d) I-ANW (e) I-ANC (f) PCA-NW

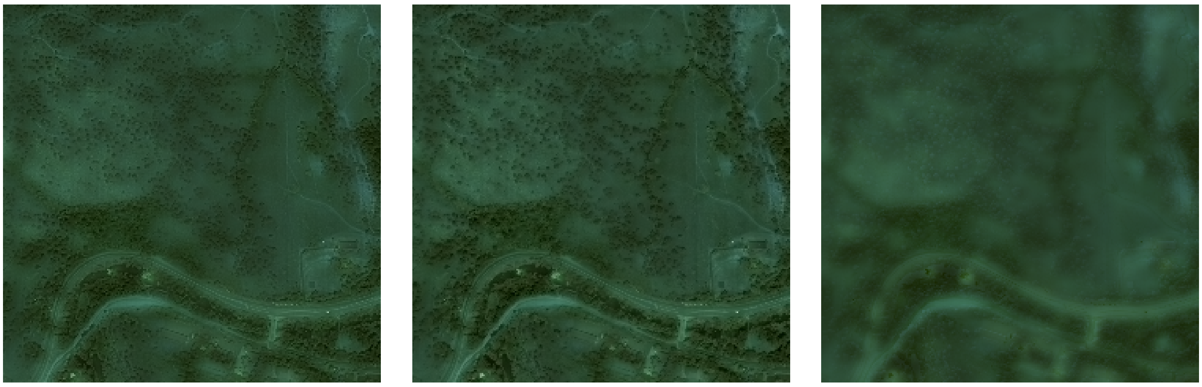


(g) PCA-NC (h) WiSpeR (i) CiSpeR

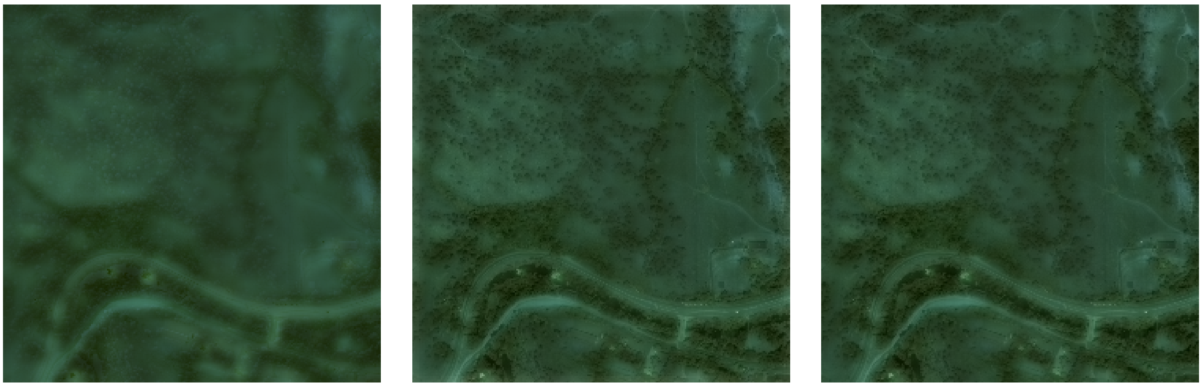
Figure 4.4: *(a) Low resolution image formed from the G-R-NIR bands of the MS SPOT image. (b)–(i) Pansharpened images using the methods under study.*



(a) MS(G-R-NIR) (b) ANW (c) ANC



(d) I-ANW (e) I-ANC (f) PCA-NW



(g) PCA-NC (h) WiSpeR (i) CiSpeR

Figure 4.5: *(a) Low resolution image formed from the R-G-B bands of the MS QuickBird image. (b)–(i) Pansharpened images using the methods under study.*

Figure 4.5 for the QuickBird dataset. From the pansharpened images we observe that the contourlet-based and the new pansharpening method not only enhance the spatial resolution, but also preserve the spectral information of the original MS image better than wavelet ones. Although the IHS-based fusion methods improve the spatial resolution, they induce some color distortion, so these methods cannot be used to preserve the spectral information effectively. The additive contoulet method preserves the spectral information to some extent; however it contains little spatial information which is manifested as blur on the fused image. PCA-based resulted image with Landsat and QuickBird does not differ much from the original MS image, while with SPOT dataset it improves the spatial resolution and preserves the spectral information of the original MS image. The proposed method, CiSpeR, enhances the images spatially while accurately preserves the spectral information for all imagery datasets.

Measure	**Band**	**ANW**	**ANC**	**I-ANW**	**I-ANC**	**PCA-NW**	**PCA-NC**	**WiSpeR**	**CiSpeR**
COR	b1	0.81	0.86	**0.87**	**0.90**	0.33	0.33	**0.87**	**0.90**
I.V. =1	b2	0.85	0.88	0.89	**0.91**	0.48	0.47	0.88	**0.90**
	b3	**0.90**	**0.90**	0.84	**0.85**	0.41	0.40	0.79	0.80
	b4	**0.95**	**0.95**	-	-	**0.93**	0.91	0.84	0.85
UIQI	b1	0.77	0.84	0.70	0.79	**0.94**	**0.95**	0.57	0.79
I.V. =1	b2	0.81	0.87	0.76	0.83	**0.95**	**0.96**	0.66	0.85
	b3	0.80	0.87	0.88	0.91	**0.96**	**0.97**	0.89	0.93
	b4	0.83	0.86	-	-	0.77	**0.88**	**0.88**	**0.93**
ERGAS	-	5.08	3.51	4.25	2.77	3.25	**2.08**	4.72	**2.50**
I. = Lowest	-	-	-	-	-	-	-	-	-

Table 4.4: Landsat 7 Quantative Analysis

The visual inspection of the images of Figure 4.3, Figure 4.4 and Figure 4.5 agree also with the quantative analysis results. Table 4.4 shows the quantitative results for Landsat 7 imagery. The highlighted values in the table present the two highest ideal values for each measure. It is clear that the contourlet-based methods, have better results than wavelet ones except for the PCA based approaches where the non-subsampled wavelet-based method obtain slighter better values for the COR coefficient. These results are expected since the contourlet is known to have a better representation for directional information, and the nonsubsampled version provides a shift-invariant representation. The figures of merit in this table also indicate that some

Measure	**Band**	**ANW**	**ANC**	**I-ANW**	**I-ANC**	**PCA-NW**	**PCA-NC**	**WiSpeR**	**CiSpeR**
COR	b1	0.90	0.92	**0.93**	**0.94**	0.83	0.85	0.89	0.92
I.V. =1	b2	0.95	**0.96**	0.91	0.93	**0.98**	0.95	0.92	0.94
	b3	0.94	**0.95**	0.93	**0.95**	**0.97**	**0.95**	0.61	0.61
	b4	0.93	**0.95**	-	-	**0.95**	**0.96**	0.45	0.45
UIQI	b1	0.86	0.91	0.83	0.88	**0.95**	**0.95**	0.80	0.86
I.V. =1	b2	0.91	0.94	**0.95**	**0.97**	0.93	**0.95**	0.86	0.91
	b3	0.91	**0.95**	0.92	**0.95**	0.93	**0.95**	0.98	**0.98**
	b4	0.82	0.88	-	-	0.88	0.89	0.99	**0.99**
ERGAS	-	5.36	3.52	5.40	3.30	**3.25**	**3.03**	5.27	4.3
I. = Lowest	-	-	-	-	-	-	-	-	

Table 4.5: SPOT 5 Quantative Analysis

approaches, like PCA, are spectrally efficient but they did not add many spatial details, while the proposed CiSpeR approach achieves consistent results, spectrally and spatially, providing one of the highest COR values with one of the highest UIQI values and a very low ERGAS value.

Table 4.5 shows the evaluation results with the same quality indexes for SPOT 5 imagery. SPOT 5 quantative analysis again shows that contourlet-based methods provide better results compared to all other wavelet-based methods. Here all the contourlet-based methods performs very well, obtaining a high COR and high UIQI with low ERGAS, probably due to the high resolution of the images, with only 5 meter per pixel, that allows a very good representation of the spatial structures using the contourlet trasnform. For the SPOT image we can see that PCA give better results than PCA with Landsat. This may be due to the correlation between PC1 and MS bands. While, for the Landsat 7 image, PC1 is very similar to band 4, in SPOT 5 it takes information from the four bands. Again, CiSpeR almost achieves close to ideal values in all the spatial and spectral measures in the first 2 bands, where small details were injected to the bands 3 and 4, since these bands are not covered with the panchromatic band.

Table 4.6 shows the evaluation results with the same quality indexes for QuickBird imagery. QuickBird quantative analysis shows that PCA contourlet-based method and Cisper provide better results compared to the PCA wavelet-based method and Wisper, while for additive and IHS based approaches we can see that contourlet-based methods injected more details than

Measure	**Band**	**ANW**	**ANC**	**I-ANW**	**I-ANC**	**PCA-NW**	**PCA-NC**	**WiSpeR**	**CiSpeR**
COR	b1	0.97	**0.96**	**0.97**	**0.97**	0.83	0.85	0.88	0.93
I.V. =1	b2	0.96	**0.97**	0.95	0.9627	**0.98**	0.95	0.86	0.90
	b3	**0.97**	**0.97**	**0.96**	**0.97**	**0.97**	0.95	0.90	0.94
	b4	**0.88**	**0.96**	-	-	**0.96**	**0.96**	0.76	0.81
UIQI	b1	**0.81**	0.92	0.96	0.82	0.96	**0.98**	0.91	0.92
I.V. =1	b2	0.94	0.93	**0.99**	0.94	**0.98**	**0.98**	0.92	0.95
	b3	0.94	0.92	**0.99**	0.95	0.96	0.97	0.94	0.97
	b4	**0.95**	0.86	-	-	0.78	0.86	**0.96**	**0.96**
ERGAS	-	2.67	2.80	**1.23**	2.92	2.38	**1.68**	3.17	2.03
I. = Lowest	-	-	-	-	-	-	-	-	

Table 4.6: QuickBird Quantative Analysis

wavelet but we can notice that the non-subsampled wavelet-based methods obtain better values for ERGAS and UIQI. For the Quickbird image, all the contourlets-based methods obtain higher COR values than the wavelets-based methods. That means that spatial structures are better represented using the contourlet transform than using the wavelets transform. Here all the contourlet-based methods performs very well, obtaining a high COR that allows a very good representation of the spatial structures using the contourlet transform, while the spectral details are preserved. For the QuickBird image we can see that PCA-based methods have low COR value in band 1. Regarding to CiSper, we can see that it obtain slighter worse values than ANW for some bands, but it achieves almost ideal values in all the spatial and spectral measures. This may be due to the spatial resolution ratio between QuickBird MS and PAN images that is 4, while for SPOT5 and Landsat is only 2.

4.5 Conclusions

In this chapter, wavelet and contourlet based pansharpening approaches have been compared and their efficiency to merge Landsat 7, SPOT 5 and Quickbird images has been evaluated by means of visual and quantative analysis.

Different image pansharpening methods based on the undecimated wavelet and contourlet transform (Additive, IHS and PCA) have been experimentally compared. Also the new proposed method CiSper was compared with these methods and with WiSper. In all methods, contourlet-based pansharpened images present, visually and numerically, better results than those obtained by wavelet for both Landsat and SPOT imagery extracting spatial information from the PAN image missing in the MS image, without modifying its spectral information content.

CiSper obtain a very low ERGAS value, smaller than 3, in all the imagery, and values very close to the ideal in the other measures. It is a consistent approach that works well spatially and spectrally with different imagery dataset.

Chapter 5

Bayesian super-resolution pansharpening using contourlets [1]

5.1 Introduction

In the previous chapter, the contourlet transform and some of the pansharpening methods based on it have been briefly discussed. In recent years, contourlet-based algorithms, specifically non-subsampled contourlet transform (NSCT) [61] based ones, are becoming popular [62]. The main problems of the classical methods based on contourlet are the control of the noise in the images and their dependency on an initial interpolation but they efficiently preserve texture and contour information.

Recently, a new super-resolution (SR) approach has been proposed [96]. Within the Bayesian formulation, this method incorporates prior knowledge on the expected characteristics of MS images using a Total Variation prior, and considers the sensor characteristics to model the observation process of both PAN and MS images. However, the MS bands which are not covered by PAN image cannot be improved properly using this method and some color bleeding may appear [176].

In this chapter we propose a new pansharpening method that combines the super-resolution technique presented in [96] with non-subsampled contourlet transform in order to obtain a method that efficiently preserves the texture and contour information of the PAN image while

---

[1] Part of this chapter was published in the International Conference on Image Processing 2010 [30]. This paper has been selected as finalist of the Huawei Best Student Paper Award.

improving all the bands of the image, even those that are not covered by the PAN image. This chapter is organized as follows. In section 5.2 the Bayesian SR using contourlet approach is described and the used notation introduced. Section 5.3 describes the Bayesian SR formulation and the actual parameters hyperpriors, image prior, and observation models used in this chapter. Section 5.4 describes the variational approach to distribution approximation for Bayesian SR pansharpening using contourlets and how inference is performed. Experimental results and comparison are presented in section 5.5 for synthetic, SPOT5 and QuickBird images and finally, section 5.6 concludes the chapter.

5.2 Problem formulation

Let us assume that y, the unknown high resolution MS image we would have observed under ideal conditions, has B bands, y_b, $b = 1,...,B$, each one centered on a narrow spectral band, of size $p = m \times n$, that is, $y = [y_1^t, y_2^t, ..., y_B^t]^t$, where each band of this image is expressed as a column vector by lexicographically ordering the pixels in the band, and t denotes the transpose of a vector or matrix. The observed low resolution MS image Y has B bands Y_b, $b = 1,...,B$, each of size $P = M \times N$ pixels, with $M < m$ and $N < n$. These images are also stacked into the vector $Y = [Y_1^t, Y_2^t, ..., Y_B^t]^t$, where each band of this image is also expressed as a column vector by lexicographically ordering the pixels in the band. The sensor also provides us with a panchromatic image x of size $p = m \times n$, that contains reflectance data in a single band that covers a wide area of the spectrum.

Using the contourlet transform, these images can also be expressed in NSCT domain, as

$$x = \sum_{l=1}^{L} \sum_{d=1}^{D} C^{ld} x + C^r x, \qquad y_b = \sum_{l=1}^{L} \sum_{d=1}^{D} C^{ld} y_b + C^r y_b, \tag{5.1}$$

where $C^{ld} x$ and $C^{ld} y_b$ are the NSCT coefficients for the PAN and MS bands, respectively, at a level of scale decomposition l and direction d, and $C^r x$ and $C^r y_b$ are the residual (low pass filtered version of original) image of PAN and MS bands, respectively. For the sake of simplicity, we will use in this chapter the notation $\sum_j C_j$ instead of $\sum_{l=1}^{L} \sum_{d=1}^{D} C^{ld}$, where $j \in l,d$ the set of the levels and directions.

Following [96], we define the relationship between high resolution MS image and its low resolution counterpart as

$$Y_b = H y_b + n_b, \tag{5.2}$$

where the $P \times p$ degradation matrix H combines the subsampling, integration and blur present in the image and n_b is the capture noise assumed to be independent white Gaussian of known variance β_b^{-1}.

The panchromatic image contains the details of the high resolution MS image but lacks of its spectral information. Following the wavelet and contourlet based pansharpening methods, see [7, 62], the PAN image is considered here as a combination of the high frequency details of the high resolution MS bands, plus a residual, low pass filtered image, which have an unknown relationship with the MS bands. The relationship between the PAN and the high resolution MS images is then defined in this chapter as

$$x = \sum_j C_j x + C^r x = \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j y_b + \frac{1}{B}\sum_{b=1}^{B} C^r y_b + v, \tag{5.3}$$

and assuming $\frac{1}{B}\sum_{b=1}^{B} C^r y_b = C^r x$, we have

$$x_d = \sum_j C_j x = \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j y_b + v, \tag{5.4}$$

where x_d contains the details of the PAN image, obtained using the contourlet transform, which is able to effectively extract the details of an image and v is the noise that is assumed to be Gaussian with zero mean and known variance γ^{-1}. Note that this model doesn't take into account the residuals of the NSCT, in a similar way as the additive wavelet or contourlet methods do [62]. So, we are assuming that the objects structure is present in the MS bands but, since the PAN image doesn't cover all the spectral range of the multispectral image, its intensities will not necessarily coincide with the MS bands intensities and, hence, we are not forcing this similarity.

5.3 Bayesian Formulation

The Bayesian formulation of the high resolution multispectral reconstruction problem requires the definition of the joint distribution $P(\Omega, y, Y, x)$ of the PAN high resolution observation x, the low resolution multispectral Y, the unknown high resolution multispectral y and the rest of the hyper-parameters Ω. The posterior distribution of the unknowns given the observed LRMS and PAN, $p(\Omega, y|Y, x)$ has to be calculated and used to estimate the HR image y and the parameters. To calculate this joint distribution, we utilize the hierarchical Bayesian paradigm, which have been described in Chapter 3.

In the hierarchical approach to our HRMS image reconstruction problem we have two stages; in the first stage, knowledge about the structural form of LRMS and PAN image observation noise and the structural behavior of the HRMS image is used in forming $p(Y,x|y,\Omega)$ and $p(y|\Omega)$, respectively. These noise and image models depend on the unknown hyper-parameters Ω. In the second stage, a hyper-prior on the hyper-parameters is defined, allowing the incorporation of information about the hyper-parameters into the process. So, following the Bayesian paradim we can define the joint distribution on the observation, hyper-parameters and HRMS image, $p(\Omega,y,Y,x)$ as

$$p(\Omega,y,Y,x) = p(\Omega)p(y|\Omega)p(Y,x|y,\Omega), \tag{5.5}$$

and the inference will based on $p(\Omega,y|Y,x)$. Let us now define these distributions:

5.3.1 Hyper-priors, Priors and observation models used in Bayesian SR HRMS image reconstruction

First Stage: Prior Model on the MS image

In this chapter we choose a prior model based on the Total Variation (TV) [177]. The idea behind this model is to consider the image as a set of relatively smooth objects or regions separated by strong edges. This knowledge is common in practically all the satellite and natural images. This model enforces smoothness within the objects in each image band while preserving their edges [176]. Following [96], we assume that there is no correlation between the different HR bands, hence defining

$$p(y|\Omega) = \prod_{b=1}^{B} p(y_b|\alpha_b) \propto \prod_{b=1}^{B} \alpha_b^{p/2} \exp\{-\alpha_b TV(y_b)\}, \tag{5.6}$$

with

$$TV(y_b) = \sum_{i=1}^{p} \sqrt{(\Delta_i^h(y_b))^2 + (\Delta_i^v(y_b))^2} \tag{5.7}$$

where the operators Δ_i^h and Δ_i^v correspond, respectively, to the horizontal and vertical first order differences at pixel i, that is,

$$\Delta_i^h(y) = y_i - y_{l(i)}, \tag{5.8}$$

$$\Delta_i^v(y) = y_i - y_{a(i)}, \tag{5.9}$$

where $l(i)$ is the nearest neighbors of i to the left, $a(i)$ is the nearest neighbors of i to the above and α_b is the model parameter of the band b. Note that the partition function of the TV image prior has been approximated using the approach in [178].

First Stage: Observation Model of the LRMS and PAN images

Since the observed PAN and LRMS images are independent given the real HRMS image to be estimated, we can write

$$p(Y,x|y,\Omega) = p(Y|y,\Omega)p(x|y,\Omega). \tag{5.10}$$

Each band of the LRMS image, Y_b, is related to its corresponding HR image band by the degradation model defined in Eq. (5.2). Using this degradation, the distribution of the observed Y given y, and the set of parameters Ω, is defined by

$$p(Y|y,\Omega) = \prod_{b=1}^{B} p(Y_b|y_b,\beta_b) \propto \beta^{P/2} \prod_{b=1}^{B} \exp\left\{-\frac{1}{2}\beta_b \|Y_b - Hy_b\|^2\right\}. \tag{5.11}$$

Using the degradation model defined in Eq. (5.4), the distribution of the details of the PAN image x_d, given y and a set of parameters Ω is given by,

$$p(x|y,\Omega) = p(x_d|y,\gamma) \propto \gamma^{p/2} \exp\left\{-\frac{1}{2}\gamma \left\| x_d - \frac{1}{B}\sum_{b=1}^{B}\sum_{j} C_j y_b \right\|^2\right\}. \tag{5.12}$$

Second Stage: Hyper-prior on the hyper-parameters

The set of hyper-parameters, Ω, will be formed by the prior model hyper-parameter $\alpha_1,\ldots,\alpha_B$, $\beta_1,\ldots,\beta_B$ and γ. Hence, let $\Omega = (\alpha_1,\ldots,\alpha_B,\beta_1,\ldots,\beta_B,\gamma)$. A large part of Bayesian literature is devoted to find hyper-prior distribution $p(\Omega)$ for which $p(\Omega,y|Y,x_d)$ can be calculated in a straightforward way or can be approximated. These are so called conjugate priors [179]. Conjugate priors have the intuitive feature of allowing one to begin with a certain functional form for the prior and end up with a posterior of the same functional form, but with the parameters updated by the sample information. Taking this consideration about conjugate priors into account, we will assume that each of the hyper-parameters, $w \in \Omega$, has a hyper-prior the gamma distribution,

$$p(w|a_w,c_w) = \Gamma(w|a_w,c_w), \tag{5.13}$$

defined by,

$$\Gamma(w|a_w,c_w) = \frac{(c_w)^{a_w}}{\Gamma(a_w)} w^{a_w-1} \exp[-c_w w], \tag{5.14}$$

where $w > 0$ denotes a hyper-parameter, and the two parameters $a_w > 0$ and $c_w > 0$ are respectively, the shape and the inverse scale parameters of the distribution, that are assumed to be known,we will discuss on their calculation in section 5.5. The gamma distribution has the following mean, mode and variance,

$$E[w] = \frac{a_w}{c_w}, \tag{5.15}$$

$$mode[w] = \frac{a_w - 1}{c_w}, \tag{5.16}$$

$$var[w] = \frac{a_w}{(c_w)^2}. \tag{5.17}$$

Note that the mean and the mode do not coincide. Using gamma distribution as hyper-priors for the hyper-parameters allows us to incorporate, in a straightforward manner, prior knowledge about the expected value of the hyper-parameters and also about the confidence on such expected value.

We will then use the following distribution as the hyper-prior on the hyper-parameters,

$$\begin{aligned} p(\Omega) &= p(\alpha_1, \ldots, \alpha_B, \beta_1, \ldots, \beta_B, \gamma) \\ &= \prod_{b=1}^{B} p(\alpha_b) \cdot \prod_{b=1}^{B} p(\beta_b) \cdot p(\gamma), \end{aligned} \tag{5.18}$$

where the hyper-prior for each hyper-parameter $w \in \Omega$ is the gamma distribution defined in Eq. (5.14).

Finally, combining the first and second stages of the hierarchical Bayesian approach, we have the joint distribution,

$$\begin{aligned} p(\Omega, y, Y, x) &= p(\Omega) p(y|\Omega) p(Y|y, \Omega) p(x|y, \Omega) \\ &= \prod_{b=1}^{B} p(\alpha_b) \cdot \prod_{b=1}^{B} p(\beta_b) \cdot p(\gamma) \cdot \prod_{b=1}^{B} p(y_b|\alpha_b) \cdot \prod_{b=1}^{B} \left(p(Y_b|y_b, \beta_b) \cdot p(x_d|y_b, \gamma) \right), \end{aligned} \tag{5.19}$$

where $p(y_b|\alpha_b)$, $p(Y_b|y_b, \beta_b)$ and $p(x_d|y_b, \gamma)$ are given in Eqs. (5.6), (5.11) and (5.12), respectively.

5.4 Bayesian Inference and variational approximation of the posterior distribution for SR reconstruction of MS image

For our selection of hyper-parameters in the previous section, the set of all unknowns is given by,

$$(\Omega, y) = (\alpha_1, \ldots, \alpha_B, \beta_1, \ldots, \beta_B, \gamma, y). \tag{5.20}$$

As already known, the Bayesian paradigm dictates that inference on (Ω, y) should be based on,

$$p(\Omega, y|Y, x) = \frac{p(\Omega, y, Y, x)}{p(Y, x)}, \tag{5.21}$$

once $p(\Omega, y, Y, x)$ has been calculated, y can be integrated out to obtain $p(\Omega|Y, x)$. This distribution is then used to simulate or select the value of the hyper-parameters. If a point estimate, $\hat{\Omega} = (\hat{\alpha_1}, \ldots, \hat{\alpha_B}, \hat{\beta_1}, \ldots, \hat{\beta_B}, \hat{\gamma})$ is required, then the mode or the mean of this posterior distribution can be used. Finally, a point estimate of the original HRMS image $\hat{y}$ can be obtained by maximizing $p(y|Y, x, \hat{\Omega})$. Alternatively the mean value of this posterior distribution can be selected as the estimate of the MS image.

From the above discussion it is clear that in order to perform inference we need to either calculate or approximate the posterior distribution $p(\Omega, y|Y, x)$. However, $p(\Omega, y|Y, x)$ can not be found in closed form, since

$$p(Y, x) = \int\int p(\Omega, y|Y, x) dy d\Omega \tag{5.22}$$

can not be calculated analytically. Thus, we will apply variational methods to approximate this distribution by the distribution $q(\Omega, y)$. We utilize a mean field approximation for the posterior distributions of Ω and y [164], so that these posterior distributions are assumed to be independent given the observation, that is, we are going to approximate $q(\Omega, y)$ by $q(\Omega)q(y)$. We will later show that particular selections of the distributions $q(\Omega)$ and $q(y)$ lead to hyper-parameters and image point estimates provided by the evidence and empirical analysis described later in this section. Notice, however, that unless the distributions $q(\Omega)$ and $q(y)$ are degenerate, the variational approximation provides us with additional information that goes beyond simple point estimates.

The variational criterion used to find $q(\Omega, y)$ is the minimization of the Kullback-Leibler divergence [173], given by

$$\begin{aligned} & C_{KL}(q(\Omega, y)||p(\Omega, y|Y, x_d)) \\ & \quad = \int_\Omega \int_y q(\Omega, y) \log\left(\frac{q(\Omega, y)}{p(\Omega, y|Y, x_d)}\right) d\Omega dy \qquad (5.23) \\ & \quad = \int_\Omega \int_y q(\Omega, y) \log\left(\frac{q(\Omega, y)}{p(\Omega, y, Y, x_d)}\right) d\Omega dy + const, \qquad (5.24) \end{aligned}$$

which is always non-negative and equal to zero only when $q(\Omega, y) = p(\Omega, y|Y, x_d)$.

As we already mentioned, we choose to approximate the posterior distribution $p(\Omega, y|Y, x_d)$ by the distribution,

$$q(\Omega, y) = q(\Omega)q(y), \tag{5.25}$$

where $q(y)$ and $q(\Omega)$ denote distributions on y and Ω, respectively, and then, we can rewrite Eq. (5.24) as ,

$$\begin{aligned} &C_{KL}(q(\Omega, y)||p(\Omega, y|Y, x_d)) \\ &\quad= \int_\Omega q(\Omega)\left(\int_y q(y)\log\left(\frac{q(\Omega)q(y)}{p(\Omega, y, Y, x_d)}\right)dy\right)d\Omega + const \\ &\quad= \int_y q(y)\left(\int_\Omega q(\Omega)\log\left(\frac{q(\Omega)q(y)}{p(\Omega, y, Y, x_d)}\right)d\Omega\right)dy + const. \end{aligned} \tag{5.26}$$

Due to the form of the TV prior, the above integral is difficult to evaluate. However, following [96], we can majorize the TV prior by a function which renders the integral easier to calculate.

Let us consider the following inequality, also used in [180], which states that, for any $w \geq 0$ and $z > 0$

$$\sqrt{wz} \leq \frac{w+z}{2} \Rightarrow \sqrt{w} \leq \frac{w+z}{2\sqrt{z}}. \tag{5.27}$$

Using this inequality in Eq. (5.27) with $w = (\Delta_i^h(y_b))^2 + \Delta_i^v(y_b))^2$ and $z = u_b(i)$, we define the following functional

$$M(\alpha_b, y_b, u_b) = \alpha_b^{p/2} \times \exp\left[-\frac{\alpha_b}{2}\sum_{i=1}^{p}\frac{(\Delta_i^h(y_b))^2 + \Delta_i^v(y_b))^2 + u_b(i)}{\sqrt{u_b(i)}}\right]. \tag{5.28}$$

where $u_b \in (R^+)^p$ is a p-dimensional vector with components $u_b(i), i = 1, \ldots, p$, that need to be calculated and have, as we will show later, an intuitive interpretation related to the unkown images y_b.

Comparing Eq. (5.28) with Eq. (5.27), we obtain

$$p(y_b|\alpha_b) \geq c.M(\alpha_b, y_b, u_b). \tag{5.29}$$

Inequality (5.29) leads to the following lower bound for the joint probability distribution,

$$\begin{aligned} p(\Omega, y, Y, x_d) &\geq c.p(\Omega)\prod_{b=1}^{B} M(\alpha_b, y_b, u_b)p(Y|y, \beta)p(x_d|y, \gamma) \\ &= F(\Omega, y, Y, x_d, u), \end{aligned} \tag{5.30}$$

where $u = [u_1^t, u_2^t, \ldots, u_B^t]^t$.

Hence, by defining,

$$\tilde{\mathcal{M}}(q(\Omega,y),u)=\int_{\Omega}\int_{y}q(\Omega,y)\times\log\left(\frac{q(\Omega,y)}{F(\Omega,y,Y,x_d,u)}\right)d\Omega dy, \tag{5.31}$$

and utilizing inequality (5.30) we obtain,

$$\mathcal{M}(q(\Omega,y))\le\min_{u}\tilde{\mathcal{M}}(q(\Omega,y),u). \tag{5.32}$$

Therefore, by finding a sequence of distributions $\{q^k(\Omega,y)\}$ that monotonically decrease $\tilde{\mathcal{M}}(q(\Omega,y),u)$ for a fixed u, a sequence of an ever decreasing upper bound of $C_{KL}(q^k(\Omega,y)||p(\Omega,y|Y,x_d))$ is also obtained due to Eq. (5.23).

However, also minimizing $\mathcal{M}(q(\Omega,y),u)$ with respect to u, generates a sequence of vectors $\{u^k\}$ that tightens the upper bound for each distribution $\{q^k(\Omega,y)\}$. Therefore, the two sequences $\{q^k(\Omega,y)\}$ and $\{u^k\}$ are coupled. We developed an iterative algorithm, presented as Algorithm 5.1, to find such sequence.

Inequality (5.27) provides a local quadratic approximation to the TV prior. Had a fixed u^0 with same elements been used, a global conditional auto-regression model approximating the TV prior would have been obtained. Clearly, the procedure which updates u will provide a tighter upper bound for $\mathcal{M}(q^k(\Omega,y))$, since we are using $\min_u\tilde{\mathcal{M}}(q^k(\Omega,y),u)$ instead of $\tilde{\mathcal{M}}(q^k(\Omega,y),u^0)$.

Finally, we note that the process to find the best posterior distribution approximation of the image in combination with u is a very natural extension of the majorization-minimization approach to function optimization [181] and that local majorization has also been applied to variational logistic regression, as well as, to the inference of its parameters.

The following algorithm can, therefore, be used for calculating the approximating posteriors $q(\Omega,y)=q(\Omega)q(y)$.

Algorithm 5.1 *Posterior parameter and image distributions estimation in TV restoration using* $q(\Omega,y)=q(\Omega)q(y)$

Given $u^1 \in (R^+)^p$ and $q^1(\Omega)$, an initial estimate of the distribution $q(\Omega)$, *for* $k=1,2,...$ until a stopping criterion is met.

1. *Find*

$$q^k(y)=\arg\min_{q(y)}\int_\Omega\int_y q^k(\Omega)q(y)\times\log\left(\frac{q^k(\Omega)q(y)}{F(\Omega,y,Y,x_d,u^k)}\right)d\Omega dy. \tag{5.33}$$

2. *Find*

$$u^{k+1}=\arg\min_{u}\int_\Omega\int_y q^k(\Omega)q^k(y)\times\log\left(\frac{q^k(\Omega)q^k(y_i^{ld})}{F(\Omega,y,Y,x_d,u)}\right)d\Omega dy. \tag{5.34}$$

3. *Find*

$$q^{k+1}(\Omega)=\arg\min_{q(\Omega)}\int_\Omega\int_y q(\Omega)q^k(y)\times\log\left(\frac{q(\Omega)q^k(y)}{F(\Omega,y,Y,x_d,u^{k+1})}\right)d\Omega dy. \tag{5.35}$$

Set $q(\Omega)=lim_{k\to\infty}q^k(\Omega), q(y)=lim_{k\to\infty}q^k(y)$.

The convergence of the parameters defining the distributions $q^k(y)$ and $q^{k+1}(\Omega)$ can be used as stopping criterion for the above iterations. In order to simplify such criterion, the condition $\left\|E[y]_{q^k(y)}-E[y]_{q^{k-1}(y)}\right\|^2/\left\|E[y]_{q^{k-1}(y)}\right\|^2<\epsilon$, where ϵ is a prescribed bound, can also be used for terminating Algorithm 5.1. Note that this is a convergence criterion over the MS image but it normally also implies convergence on the posterior hyper-parameter distribution, since its convergence is required for the convergence of the posterior distribution of the image.

Let us now further develop each of the steps of the above algorithm. Assume that at the $k-th$ iteration step of Algorithm 5.1 we have,

$$E_{q^k(\Omega)}[\alpha_b]=\alpha_b^k, b=1,\dots,B \tag{5.36}$$

$$E_{q^k(\Omega)}[\beta_b]=\beta_b^k, b=1,\dots,B \tag{5.37}$$

$$E_{q^k(\Omega)}[\gamma]=\gamma^k. \tag{5.38}$$

To calculate $q^k(y)$, we observe the differentiating the integral on the right-hand side of Eq. (5.33) with respect to $q(y)$ and setting it equal to zero, we obtain,

$$q^k(y)\propto\exp\left\{E_{q^k(\Omega)}\left[\ln F(\Omega,y,Y,x_d,u^k)\right]\right\}, \tag{5.39}$$

and so,

$$q^k(y)\propto\exp\left[\sum_{b=1}^{B}\left(\alpha_b^k\mathrm{TV}(y_b)-\frac{\beta_b^k}{2}\|Y_b-Hy_b\|^2\right)-\frac{\gamma^k}{2}\left\|x_d-\frac{1}{B}\sum_{b=1}^{B}\sum_j C_j y_b\right\|^2\right]. \tag{5.40}$$

Thus we have that $q^k(y)$ is an $p-$dimensional Gaussian distribution with parameters,

$$q^k(y) = \mathcal{N}(y|E_{q^k(y)}[y], cov_{q^k(y)}[y]), \tag{5.41}$$

with

$$cov_{q^k(y)}[y] = \mathscr{A}^{-1}(u^k), \tag{5.42}$$

and

$$E_{q^k(y)}[y] = cov_{q^k(y)}[y]\phi^k, \tag{5.43}$$

where ϕ^k is the $(B \times p) \times 1$ vector,

$$\phi^k = (diag(\beta^k) \otimes H^t)Y + \frac{\gamma^k}{B}(1_B \otimes \sum_j C_j^t x_d), \tag{5.44}$$

where 1_B is the column vector of size $1 \times B$ with all its elements equal to one, and

$$\mathscr{A}(u^k) = \begin{pmatrix} \alpha_1^k \varsigma(u_1^k) & 0_p & \cdots & 0_p \\ 0_p & \alpha_2^k \varsigma(u_2^k) & \cdots & 0_p \\ \vdots & \vdots & \ddots & \vdots \\ 0_p & 0_p & \cdots & \alpha_B^k \varsigma(u_B^k) \end{pmatrix} \tag{5.45}$$

$$+diag(\beta^k) \otimes H^t H + \frac{\gamma^k}{B^2}(1_{B\times B} \otimes \sum_{j,k} C_j^t C_k),$$

where $1_{B\times B}$ is a $B \times B$ matrix with all its elements equal to one, I_B is a $B \times B$ identity matrix, $\otimes$ is the Kronecker product, $\beta = (\beta_1, \beta_2, ..., \beta_B)^t$ and

$$\varsigma(u_b^k) = (\Delta^h)^t W(u_b^k)(\Delta^h) + (\Delta^v)^t W(u_b^k)(\Delta^v), \tag{5.46}$$

for $b = 1, ..., B$, where Δ^h and Δ^v represent $p \times p$ convolution matrices associated with the first order horizontal and vertical differences, respectively, and

$$W(u_b^k) = diag\left(u_b^k(i)^{-\frac{1}{2}}\right), \tag{5.47}$$

is a $p \times p$ diagonal matrix, for $i = 1, ..., p$. This is a spatial adaptivity matrix since it controls the amount of smoothing at each pixel location depending on the strength of the intensity variation at that pixel, as expressed by the horizontal and vertical intensity gradient [96].

To calculate u^{k+1}, we have from Eq. (5.34) that,

$$u_b^{k+1} = \arg\min_{u_b} \sum_{i=1}^{p} \frac{E_{q^k(y_b)}\left[(\Delta_i^h(y_b))^2 + (\Delta_i^v(y_b))^2 + u_b(i)\right]}{\sqrt{u_b(i)}}, \tag{5.48}$$

and consequently,

$$u_b^{k+1}(i) = E_{q^k(y_b)}\left[(\Delta_i^h(y_b))^2 + (\Delta_i^v(y_b))^2\right], i = 1, \dots, p. \tag{5.49}$$

Notice that $q^k(\Omega)$ is not required in calculating u^{k+1}. It is clear from Eq. (5.49) that the vector u^{k+1} is a function of the spatial first order differences of the unknown image y under the distribution $q^k(y)$ and represents the local spatial activity of y. Therefore, matrix $W(u^k)$ in Eq. (5.47) can be interpreted as the spatial adaptivity matrix, since it controls the amount of smoothing at each pixel location depending on the strength of the intensity variation at that pixel, as expressed by the horizontal and vertical intensity gradients. That is, for the pixels with high spatial activity the corresponding entries of $W(u^k)$ are very small or zero, which means that no smoothness is enforced, while for the pixels in a flat region the corresponding entries of $W(u^k)$ are very large, which means that smoothness is enforced. This matrix $W(u^k)$ has also been referred to as the visibility matrix, since it describes the masking property of the human visual system, according to which noise is not visible in high spatial activity regions (its high frequencies are masked by the edges), while it is visible in the low spatial frequency (flat) regions. The visibility matrix and its complementary matrix $I - W(u^k)$ have been also used in iterative image restoration [178].

Once we know $q^k(y)$ and u^{k+1}, the next step is to calculate $q^{k+1}(\Omega)$. By differentiating the integral on the right-hand side of Eq. (5.35) with respect to $q(\Omega)$ and setting it equal to zero, we obtain that,

$$q^{k+1}(\Omega) \propto \exp\left\{E_{q^k(y)}\left[\ln F(\Omega, y, Y, x_d, u^{k+1})\right]\right\}, \tag{5.50}$$

and thus,

$$q^{k+1}(\Omega) = q^{k+1}(\alpha_1), \dots, q^{k+1}(\alpha_B) q^{k+1}(\beta_1), \dots, q^{k+1}(\beta_B) q^{k+1}(\gamma), \tag{5.51}$$

which produces,

$$q^{k+1}(\Omega) = q^{k+1}(\gamma) \prod_{b=1}^{B} \left(q^{k+1}(\alpha_b) q^{k+1}(\beta_b)\right), \tag{5.52}$$

where

$$
\begin{aligned}
q^{k+1}(\alpha_b) &= \alpha_b^{a_{\alpha_b}} \exp\left\{-c_{\alpha_b}\alpha_b\right\} \alpha_b^{\frac{p}{2}} \exp\left\{-\alpha_b E_{q^k(y)}[TV(y_b)]\right\} \\
&= \alpha_b^{(a_{\alpha_b}+\frac{p}{2})} \exp\left\{\alpha_b\left(-c_{\alpha_b} - E_{q^k(y)}[TV(y_b)]\right)\right\}, && (5.53) \\
q^{k+1}(\beta_b) &= \beta_b^{a_{\beta_b}} \exp\left\{-c_{\beta_b}\beta_b\right\} \beta_b^{\frac{p}{2}} \exp\left\{-\frac{1}{2}\beta_b E_{q^k(y)}\left[\|Y_b - Hy_b\|^2\right]\right\} \\
&= \beta_b^{(a_{\beta_b}+\frac{p}{2})} \exp\left\{\beta_b\left(-c_{\beta_b} - \frac{1}{2}E_{q^k(y)}\left[\|Y_b - Hy_b\|^2\right]\right)\right\}, && (5.54) \\
q^{k+1}(\gamma) &= \gamma^{a_\gamma} \exp\left\{-c_\gamma\gamma\right\} \gamma^{\frac{p}{2}} \exp\left\{-\frac{1}{2}\gamma E_{q^k(y)}\left[\left\|x_d - \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j y_b\right\|^2\right]\right\} \\
&= \gamma^{(a_\gamma+\frac{p}{2})} \exp\left\{\gamma\left(-c_\gamma - \frac{1}{2}E_{q^k(y)}\left[\left\|x_d - \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j y_b\right\|^2\right]\right)\right\}. && (5.55)
\end{aligned}
$$

where

$$E_{q^k(y)}[TV(y_b)] = \sum_{i=1}^{p} \sqrt{u_b^{k+1}(i)}, \tag{5.56}$$

$$
\begin{aligned}
E_{q^k(y)}\left[\|Y_b - Hy_b\|^2\right] &= \left\|Y_b - HE_{q^k(y)}[y_b]\right\|^2 && (5.57) \\
&\quad + trace\left(H^t H cov_{q^k(y)}[y_b]\right), \\
E_{q^k(y)}\left[\left\|x_d - \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j y_b\right\|^2\right] &= \left\|x_d - \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j E_{q^k(y)}[y_b]\right\|^2 && (5.58) \\
&\quad + trace\left(\frac{1}{B^2}(1_{B\times B} \otimes \sum_{j,k} C_j^t C_k) cov_{q^k(y)}[y]\right).
\end{aligned}
$$

From previous discussion, we can see that $cov_{q^k(y)}[y]$ is explicitly needed to calculate these quantities. However, since the calculation of $cov_{q^k(y)}[y]$ is very intense, we propose the following approximation of the covariance matrix [178]. We first approximate $W(u_b^k)$ using

$$W(u_b^k) \approx z(u_b^k) I, \tag{5.59}$$

where $z(u_b^k)$ is calculated as the mean value of the diagonal values in $W(u_b^k)$, that is

$$z(u_b^k) = \frac{1}{p}\sum_{i=1}^{p} \frac{1}{\sqrt{u_b^k(i)}}. \tag{5.60}$$

We then approximate $cov_{q^k(y)}$ using

$$
\begin{aligned}
cov_{q^k(y)}[y_b] &\approx (E_{q^k(\alpha)}[\alpha_b] z(u_b^k)(\Delta^h)^t(\Delta^h) + E_{q^k(\alpha)}[\alpha_b] z(u_b^k)(\Delta^v)^t(\Delta^v) \\
&\quad + E_{q^k(\beta)}[\beta_b] H^t H + E_{q^k(\gamma)}[\gamma] \frac{1}{B^2}\sum_{j,k} C_j^t C_k)^{-1}, && (5.61) \\
&= COV^{-1}, && (5.62)
\end{aligned}
$$

where $u_b^{k+1}(i)$ was defined in Eq. (5.49), and its calculation is carried out by expanding the right-hand side of Eq. (5.49) thus obtaining,

$$\begin{aligned} E_{q^k(y)}\left[(\Delta_i^h(y_b))^2+(\Delta_i^v(y_b))^2\right] &= \left(\Delta_i^h\left(E_{q^k(y)}[y_b]\right)\right)^2+\left(\Delta_i^v\left(E_{q^k(y)}[y_b]\right)\right)^2 \\ &+E_{q^k(y)}\left[\left(\Delta_i^h(y_b-E_{q^k(y)}[y_b])\right)^2\right]+E_{q^k(y)}\left[\left(\Delta_i^v(y_b-E_{q^k(y)}[y_b])\right)^2\right]. \end{aligned} \tag{5.63}$$

Note that the matrix COV is a block circulant matrix with circulant blocks (BCCB); thus computing its inverse can be performed if Fourier domain, which is very efficient.

Using this approximation, the last two terms in Eq. (5.63) can be expressed as,

$$\begin{aligned} &E_{q^k(y_b)}\left[\left(\Delta_i^h(y_b-E_{q^k(y_b)}[y_b])\right)^2\right]+E_{q^k(y_b)}\left[\left(\Delta_i^v(y_b-E_{q^k(y_b)}[y_b])\right)^2\right] \\ &\approx \frac{1}{p}trace\left[COV_b^{-1}\times((\Delta^h)^t(\Delta^h)+(\Delta^v)^t(\Delta^v))\right]. \end{aligned} \tag{5.64}$$

From the definition of the gamma distribution in Eq. (5.14), previous distributions have the following means,

$$E[\alpha_b]_{q^k(\Omega)}=\frac{a_{\alpha_b}+\frac{p}{2}}{c_{\alpha_b}+E_{q^k(y_b)}[TV(y_b)]}, b=1,\dots,B \tag{5.65}$$

$$E[\beta_b]_{q^k(\Omega)}=\frac{a_{\beta_b}+\frac{P}{2}}{c_{\beta_b}+\frac{1}{2}E_{q^k(y_b)}\left[\|Y_b-Hy_b\|^2\right]}, b=1,\dots,B \tag{5.66}$$

$$E[\gamma]_{q^k(\Omega)}=\frac{a_\gamma+\frac{p}{2}}{c_\gamma+\frac{1}{2}E_{q^k(y_b)}\left[\left\|x_d-\frac{1}{B}\sum_{b=1}^{B}\sum_j C_j y_b\right\|^2\right]}, \tag{5.67}$$

We can rewrite those equations having into account that the mean of the prior distribution on the parameters are $\overline{\alpha_b}=a_{\alpha_b}/c_{\alpha_b}, \overline{\beta_b}=a_{\beta_b}/c_{\beta_b}, \overline{\gamma}=a_\gamma/c_\gamma$ and

$$\lambda_{\alpha_b}=\frac{a_{\alpha_b}}{a_{\alpha_b}+p/2}, \lambda_{\beta_b}=\frac{a_{\beta_b}}{a_{\beta_b}+p/2}, \lambda_\gamma=\frac{a_\gamma}{a_\gamma+p/2}, \tag{5.68}$$

we can rewrite the above means as follows,

$$
\begin{aligned}
(E[\alpha_b]_{q^k(\Omega)})^{-1} &= \frac{\lambda_{\alpha_b}}{\overline{\alpha_b}} + (1-\lambda_{\alpha_b})\frac{\sum_{i=1}^{p}\sqrt{u_b^{k+1}(i)}}{p/2} \\
&= \frac{\lambda_{\alpha_b}}{\overline{\alpha_b}} + (1-\lambda_{\alpha_b})\times \\
&\frac{2\sum_{i=1}^{p}\sqrt{\left(\Delta_i^h\left(E_{q^k(y)}[y_b]\right)\right)^2 + \left(\Delta_i^v\left(E_{q^k(y)}[y_b]\right)\right)^2 + \frac{1}{p}trace\left[COV_b^{-1}\times((\Delta^h)^t(\Delta^h)+(\Delta^v)^t(\Delta^v))\right]}}{p},
\end{aligned}
\tag{5.69}
$$

$$
\begin{aligned}
(E[\beta_b]_{q^k(\Omega)})^{-1} &= \frac{\lambda_{\beta_b}}{\overline{\beta_b}} + (1-\lambda_{\beta_b})\frac{E_{q^k(y)}\left[\|Y_b - Hy_b\|^2\right]}{P} \\
&= \frac{\lambda_{\beta_b}}{\overline{\beta_b}} + (1-\lambda_{\beta_b})\times\frac{\left\|Y_b - HE_{q^k(y)}[y_b]\right\|^2 + trace\left(H^tHCOV_b^{-1}\right)}{P},
\end{aligned}
\tag{5.70}
$$

$$
\begin{aligned}
(E[\gamma]_{q^k(\Omega)})^{-1} &= \frac{\lambda_\gamma}{\overline{\gamma}} + (1-\lambda_\gamma)\frac{E_{q^k(y)}\left[\left\|x_d - \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j y_b\right\|^2\right]}{p} \\
&= \frac{\lambda_\gamma}{\overline{\gamma}} + (1-\lambda_\gamma)\frac{\left\|x_d - \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j E_{q^k(y)}[y_b]\right\|^2 + trace\left(\frac{1}{B^2}(1_{B\times B}\otimes\sum_{j,k}C_j^tC_k)COV^{-1}\right)}{p}.
\end{aligned}
\tag{5.71}
$$

These mean values are then used to recalculate the distributions of y in Algorithm 5.1.

Equation (5.68) indicates that λ_{α_b}, λ_{β_b} and λ_γ are taking values in the interval $[0,1)$, and so, they can be understood as normalized confidence parameters. As can be seen from Eqs. (5.69),(5.70) and (5.71) the inverse of the means of the hyper-priors are calculated as convex combinations of their initial values and their maximum likelihood (ML) estimates. These ML estimates have been derived before either empirically or by using regularization formulations [182, 183]. According to Eqs. (5.69),(5.70) and (5.71) when they are equal to zero, no confidence is placed on the initial values of the hyper-parameters and ML estimates are used, making the observation fully responsible of the parameters estimation, while when they are asymptotically equal to one, the prior knowledge of the mean is fully enforced (i.e., no estimation of the hyper-parameters is performed).

5.5 Experimental results

In order to test the proposed method, we used both a synthetic color image and a real SPOT5 and QuickBird images. We compared the proposed SR using contourlets method with the SR

Measure	**Band**	**NSCT in [62]**	**SR in [96]**	**Proposed**
COR	R	0.91	0.84	**0.97**
	G	0.91	**0.98**	0.97
	B	0.90	0.62	**0.95**
SSIM	R	0.79	0.90	**0.96**
	G	0.81	0.94	**0.96**
	B	0.81	0.85	**0.95**
PSNR	R	26.75	32.68	**37.25**
	G	27.17	35.50	**37.51**
	B	27.65	30.15	**36.13**
ERGAS	-	5.76	3.12	**1.85**

Table 5.1: Synthetic Image Quantative Results

method in [96] and the additive NSCT method [62]. To assess the spatial improvement of the pansharpened images we use the correlation of the high frequency components (COR) [25] which takes values between zero and one (the higher the value the better the quality of the pansharpened image). Spectral fidelity was assessed by means of the peak signal-to-noise ratio (PSNR), the Structural Similarity Index Measure (SSIM) [131], an index ranging from -1 to $+1$ with $+1$ corresponding to exactly equal images, and the *erreur relative globale adimensionnelle de synthése* (ERGAS) [89] index, a global criterion for what the lower the value, specially a value lower than the number of bands in the image, the higher the quality of the pansharpened image.

Our first experiment focuses on understanding the behaviour of the proposed model under controlled conditions. To this end, we used synthetic multispectral observations, obtained from the color image, displayed in Figure 5.1(a), by convolving it with mask $0.25 \times 1_{2\times 2}$ to simulate sensor integration, and then downsampling it by a factor of two by discarding every other pixel in each direction and adding zero mean Gaussian noise with variance 16. For the PAN image we used the luminance of the original color image and zero mean Gaussian noise of variance 9 was added. The observed PAN image and MS image, scaled to the size of the PAN image for displaying purposes, are shown in Figure 5.1(b) and (c), respectively.

The proposed algorithm was run until the criterion $\left\|E[y]_{q^k(y)} - E[y]_{q^{k-1}(y)}\right\|^2 / \left\|E[y]_{q^{k-1}(y)}\right\|^2 <$

10^{-4} was satisfied, which typically is reached within 5 iterations. The values of parameters were automatically estimated using the following method. The initial distribution on the parameters, $q^1(\Omega)$, in Algorithm 5.1 was estimated based on the observed MS and PAN images as, $\alpha_b^1 = p/2TV(y_b)$, $\beta_b^1 = P/\|Y_b - Hy_b\|^2$, for all $b = 1,\ldots,B$ and $\gamma^1 = p\left\|x_d - \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j y_b\right\|^2$, assuming that $q^0(y)$ is a degenerate distribution on the bicubic interpolation of the observed MS image, that is, we used the observation to initialize the hyperparameters. The initial value of u is calculated using Eq. (5.49) from the bicubic interpolation of the observed MS image. Note that the algorithms are initialized automatically without any manual input. The resulted images corresponding to the reconstruction of the synthetic image using the NSCT method in [62], the SR method in [96], and the proposed method are displayed in Figure 5.1(d)-(f), respectively, and Table 5.1 shows the corresponding quantitative results. The highlighted value in the table presents the highest value for each measure. The proposed method provides better results for each measure except for the COR of band 2 (green) where a very high value, similar to the one obtained by the method in [96] is obtained. The COR values reflect that all methods are able to incorporate the details of the PAN image into the pansharpened one, although the SR method in [96], see Figure 5.1(e), introduced less details in the band 3 (blue) since the blue band contributes only a 10% to the PAN image and more into the band 2 since it has the highest contribution, a 60%, which is reflected as a greenish color near the edges of the image. The NSCT method in [62] incorporates details in all the bands but produces a noisy image, see Figure 5.1(d). The proposed method (Figure 5.1(f)) is able to incorporate detail in all the bands while controlling the noise. The spectral fidelity measures show that the proposed method performs better than the competing method, which is also clear from the image in Figure 5.1(f), producing an image that is not as noisy as the NSCT method in [62] (Figure 5.1(d)) and preserves better the colors than the SR method in [96] (Figure 5.1(e)), while better controlling the noise. It is remarkable the high SSIM and low ERGAS values which reflect the high quality of the resulting images. Note also that the PSNR for the proposed method is about 10dB higher than NSCT method in [62] and from 2 to almost 6 dB higher than for the SR method in [96].

We next examine the effect of the introduction of additional information about the unknown hyper parameters through the use of the confidence parameters λ_{α_b}, λ_{β_b} and λ_γ on the performance of the algorithm. As we have already explained before, in the case of $\lambda_{\alpha_b} = \lambda_{\beta_b} = \lambda_\gamma = 0$, no information about the hyper parameters is available, and the observed image is responsible for the estimation of the hyperparameters and the image. In our experiments, we provided the observed MS and PAN synthetic images to the algorithm and run the algorithm while varying the confidence parameters λ_{α_b}, λ_{β_b} and λ_γ from 0 to 1 in 0.1 intervals, and setting $\overline{\alpha_b} = \alpha_b^1$, $\overline{\beta_b} = \beta_b^1$ and $\overline{\gamma} = \gamma^1$. Table 5.2 shows the ERGAS, the mean COR and mean SSIM using the proposed algorithm for selected values of the confidence parameters. The confidence values are selected to demonstrate the behavior of the proposed algorithm in the following cases:

1. when full information about the HRMS image and noise variance is available,
2. when no information is provided, i.e., the observation is fully responsible for the reconstruction,
3. when some information about the image prior parameter α is provided,
4. when some information about the noise variance is provided.

The evolution of ERGAS, mean COR and mean SSIM for the full set of confidence parameters are depicted in the Figures 5.2-5.4.

From those figures and the numerical values in Table 5.2 we can see that very good results are obtained even in the absence of prior knowledgement on the value of the parameters although introducing prior knowledgement on them improves the performance. Also we can notice that the knowledgement extracted from the observed images is not very reliable. Setting λ_α, λ_β or λ_γ to one decreases the performance. However it helps to guide the estimation of the parameters increasing COR and SSIM and decreasing ERGAS when values around 0.9 are selected. The algorithm is able to estimate very precisely the value of alpha and so, introducing knowledgement on this values does not improve significantly the performance of the algorithm while introducing knowledgement of the value of β and γ makes the method to converge in less iterations and produces better reconstructions. Best values are obtained when $\lambda_{\alpha_b} = 0.9$, $\lambda_{\beta_b} = 0.9$ and $\lambda_\gamma = 0.0$. The image correspond to $\lambda_{\alpha_b} = \lambda_{\beta_b} = \lambda_{gamma} = 0$ and the one correspond to the best reconstruction ($\lambda_{\alpha_b} = 0.9$, $\lambda_{\beta_b} = 0.9$ and $\lambda_\gamma = 0.0$) are shown in Figure 5.1(g) and (h), respectively.

In a second experiment, the method was tested on real SPOT5 dataset, where the MS image covers a region of interest of 80 by 80 pixels with pixel resolution of 10 m, while the PAN image is 160 by 160 pixels with a pixel resolution of 5 m. The MS image consists of four bands from the visible and infrared region corresponding to green (b1: 0.50-0.59 μm), red (b2: 0.61-0.68 μm), Near IR (b3: 0.78-0.89 μm), Mid IR(b4: 1.58-1.75 μm), while the PAN image consists of

λ_α	λ_β	λ_γ	**ERGAS**	**mean COR**	**mean SSIM**
0.0	0.0	0.0	2.7324	0.9967	0.9613
0.0	0.0	0.9	2.5266	0.9960	0.9615
0.0	0.0	1.0	2.5062	0.9959	0.9615
0.0	0.9	0.0	2.2480	0.9950	0.9635
0.0	1.0	0.0	2.7406	0.9967	0.9576
0.0	0.9	0.9	2.3048	0.9958	0.9618
0.0	1.0	1.0	2.2767	0.9940	0.9621
0.9	0.0	0.0	2.2717	0.9921	0.9688
0.9	0.0	0.9	2.2775	0.9937	0.9654
0.9	0.9	0.0	2.2514	0.9919	0.9693
0.9	0.9	0.9	2.3039	0.9937	0.9663
0.9	1.0	0.0	2.6248	0.9694	0.9646
0.9	0.0	1.0	2.2831	0.9937	0.9655
0.9	1.0	1.0	2.2717	0.9923	0.9671
1.0	0.0	0.0	3.3060	0.9622	0.9548
1.0	0.9	0.9	3.2486	0.9828	0.9657
1.0	1.0	1.0	3.1305	0.9831	0.9674

Table 5.2: ERGAS, Mean COR and Mean SSIM for synthetic HRMS image for selected values of λ_α, λ_β and λ_γ .

a single band covering the visible and NIR (0.48-0.71 μm). Figure 5.5(a) shows a region of the G-R-IR color image representing bands 1 to 3 of the MS image. Its corresponding PAN image is depicted in Figure 5.5(b). The resulting images after running the proposed method when no prior information about the parameters value is introduced, $\lambda_{\alpha_b} = \lambda_{\beta_b} = \lambda_{gamma} = 0$, are depicted in Figure 5.6(e), this figure also shows the resulting images for bicubic interpolation, CiSper in [29], NSCT method in [62] and SR method in [96]. Following the results, for the synthetic image, we also run the method with introducing prior knowledgement on the value of the parameters. We used $\overline{\alpha_b} = \alpha_b^1$, $\overline{\beta_b} = \beta_b^1$ and $\overline{\gamma} = \gamma^1$, and $\lambda_{\alpha_b} = 0.9$, $\lambda_{\beta_b} = 0.9$ and $\lambda_\gamma = 0.0$, which were the best values for the synthetic image. The resulting image is depicted in Figure 5.6(f). Table 5.3 shows the quantitative results corresponding to these images. The highlighted value in the table presents the highest value for each measure. The quality measure values estimated from the observed image is slightly better than the ones obtained with best values for the synthetic image, which indicates that incorporating this prior information in the SPOT 5 case makes not helping the method for better reconstruction for the HRMS image.

Visual inspections of the resulting images, displayed in Figures 5.6(a)-(f), reveals close conclusions to the obtained for the synthetic image. The proposed method provides better details in all the bands, but band 1 and band 4 seems to be spectrally modified, lower values for SSIM and PSNR, nevertheless we can not notice the visual impact of it on the resulted image. Figure 5.6(a) depicts the bicubic interpolation of the MS image and although the colors are perfect, no detail is appreciated. Figure 5.6(b) show the result of CiSper method in [29], presents the best spectral values but lowest COR values, that is, it preserves spectral properties but does not incorporate much of the details of the PAN image. This is clear in the image which is not as sharp as the one obtained with the proposed method. Also, from PSNR values, it is clear that for band 3 and 4 in SPOT 5, these bands are not changed by the method. The NSCT method in [62] (Fig.5.6(c)) provides a detailed image but quite noisy, the SR method in [96] provides good details for bands 1 and 2, see Figure 5.6(d), but not for bands 3 and 4 since the PAN image does not cover those bands, this is why the blue color in Figure 5.6(d), seems to be vanished. The proposed method in Figure 5.6(e)-(f) preserves better the colors, incorporates more details from the PAN image into the pansharpened image and controls the noise in the images.

In a third experiment, the method was tested on real QuickBird dataset, where the MS image covers a region of interest of 60 by 60 pixels with pixel resolution of 2.44 m, while the PAN image is 240 by 240 pixels with a pixel resolution of 61 cm to 72 cm. Figure 5.7(a) shows a region of the RGB color image representing bands 1 to 3 of the MS image. The resulting images after running the proposed method when no prior information about the parameters value is introduced, $\lambda_{\alpha_b} = \lambda_{\beta_b} = \lambda_\gamma = 0$, are depicted in Figure 5.8(e), this figure also shows the resulting images for bicubic interpolation, CiSper in [29], NSCT method in [62] and SR method in [96]. Following the results, for the synthetic image, we also run the method with introducing prior knowledgement on the value of the parameters. We used $\overline{\alpha_b} = \alpha_b^1$, $\overline{\beta_b} = \beta_b^1$ and $\overline{\gamma} = \gamma^1$, and $\lambda_{\alpha_b} = 0.9$, $\lambda_{\beta_b} = 0.9$ and $\lambda_\gamma = 0.0$, which were the best values for the synthetic image. The resulting image is depicted in Figure 5.8(f). Table 5.4 shows the quantitative results corresponding to these images. The highlighted value in the table presents the highest value for each measure. Again, the quality measure values estimated from the observed image is

Measure	**Band**	**NSCT [62]**	**SR [96]**	CiSper [29]	**Proposed(No Prior)**	**Proposed(Prior)**
COR	b1	0.91	0.87	0.92	**0.99**	**0.99**
	b2	0.95	0.95	0.95	**0.99**	**0.99**
	b3	0.95	0.96	0.95	**0.99**	**0.99**
	b4	0.94	0.78	0.96	**0.98**	**0.98**
SSIM	b1	0.79	0.74	**0.84**	0.59	0.50
	b2	0.75	0.78	0.86	**0.87**	0.80
	b3	0.76	0.78	0.84	**0.85**	0.77
	b4	0.67	**0.89**	0.81	0.64	0.56
PSNR	b1	29.40	27.23	**29.77**	28.23	27.58
	b2	24.76	26.36	26.20	**29.07**	28.24
	b3	26.47	27.10	∞	27.98	**28.17**
	b4	24.77	**27.00**	∞	25.35	23.98
ERGAS	-	6.36	6.02	**3.03**	6.12	6.74

Table 5.3: SPOT 5 Image Quantative Results

slightly better than the ones obtained with best values for the synthetic image for the COR measure, although it preserving better the spectral quality with slightly higher values of SSIM and PSNR.

Visual inspections of the resulting images, displayed in Figures 5.8(a)-(f), reveals similar conclusions to the obtained for the synthetic and SPOT 5 images. The proposed method provides better details in all the bands. Figure 5.8(a) depicts the bicubic interpolation of the MS image, as the SPOT 5 case it preserves the image colors but no detail is appreciated. CiSper method in [29] (Figure 5.8(b)) again present better spectral details but less details in the band 3 and band 4. The NSCT method in [62] (Figure 5.8(c)) provides again a detailed image but quite noisy, the SR method in [96] provides good details except for band 1, but has the problem of color bleeding (see Figure 5.8(d)). The proposed method in Figure 5.8(e)-(f) preserves better the colors, incorporates more details from the PAN image into the pansharpened image and controls the noise in the images.

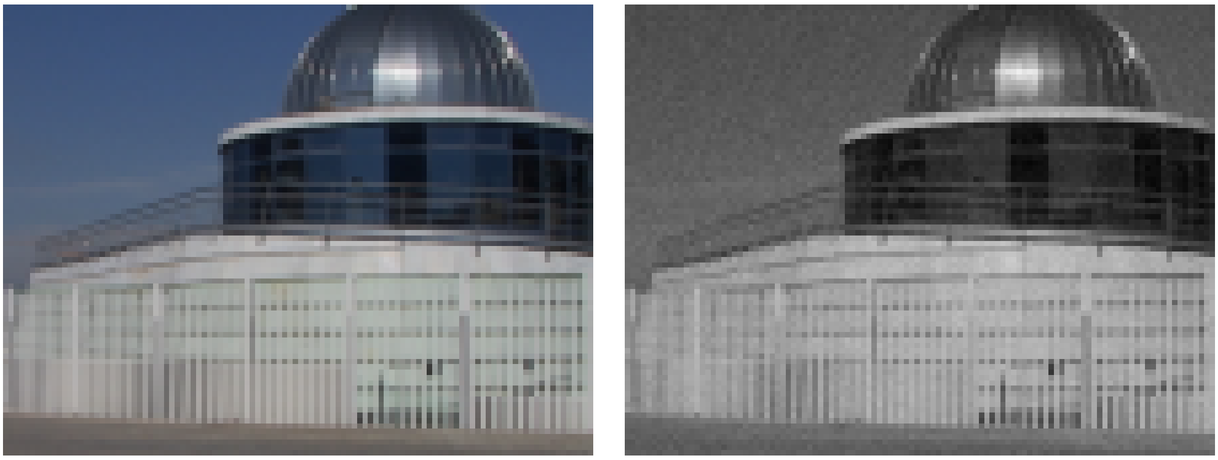
Measure	**Band**	**NSCT [62]**	**SR [96]**	CiSper [29]	**Proposed (no Prior)**	**Proposed (Prior)**
COR	b1	0.96	0.39	0.92	**0.99**	0.98
	b2	0.97	0.87	0.94	**0.99**	0.98
	b3	0.97	0.86	0.61	**0.99**	0.98
	b4	**0.96**	0.88	0.45	0.95	94
SSIM	b1	0.80	**0.96**	0.75	0.51	0.52
	b2	0.61	0.58	**0.78**	0.52	0.54
	b3	0.58	0.62	**0.79**	0.52	0.53
	b4	0.39	0.44	**0.81**	0.55	0.56
PSNR	b1	**34.68**	31.44	30.28	27.77	27.85
	b2	29.21	20.75	**30.12**	27.08	27.13
	b3	28.37	21.97	**30.44**	26.85	26.89
	b4	23.43	18.51	**29.34**	25.89	25.83
ERGAS	-	2.80	14.76	**4.30**	5.93	6.02

Table 5.4: QuickBird Image Quantitative Results

5.6 Conclusions

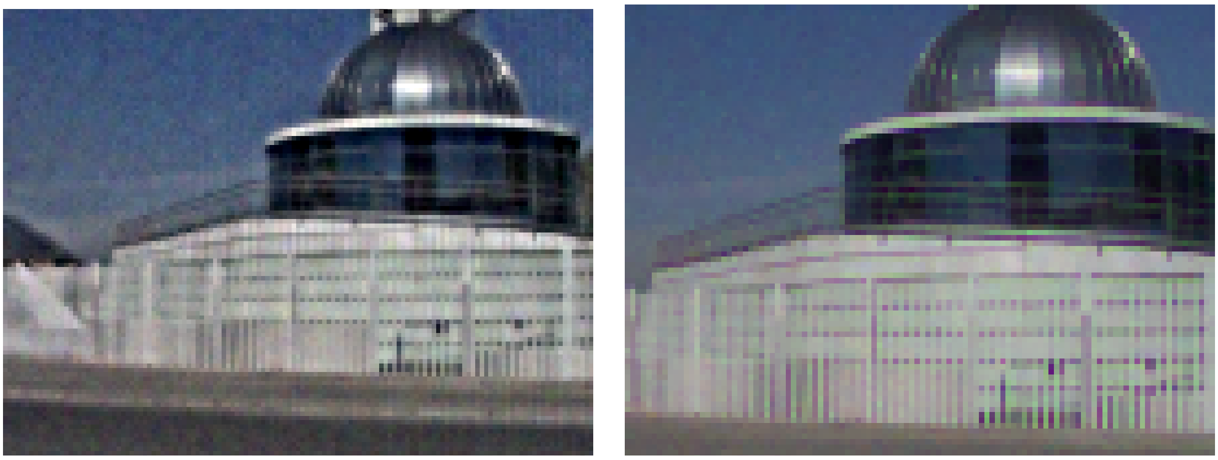
In this chapter, a new pansharpening method based on super-resolution reconstruction and non subsampled contourlet transform has been presented. The residual and coefficients bands were reconstructed independently, and the relationship between the contourlet coefficients has been examined. This relationship used by the algorithm and improved its performance. The method also estimate simultaneously the HRMS image and the unknown hyperparameters. The proposed method preserves the spectral properties of MS image while incorporating the high frequencies from the panchromatic image and controlling the noise in the image.

The efficiency of pansharpening methods has been evaluated by means of visual and quantitative analysis, for synthetic and real data. Based on the presented experiments, the proposed method does significantly outperform NSCT-based and TV-based super-resolution methods and the methods have been proposed in previous chapters.



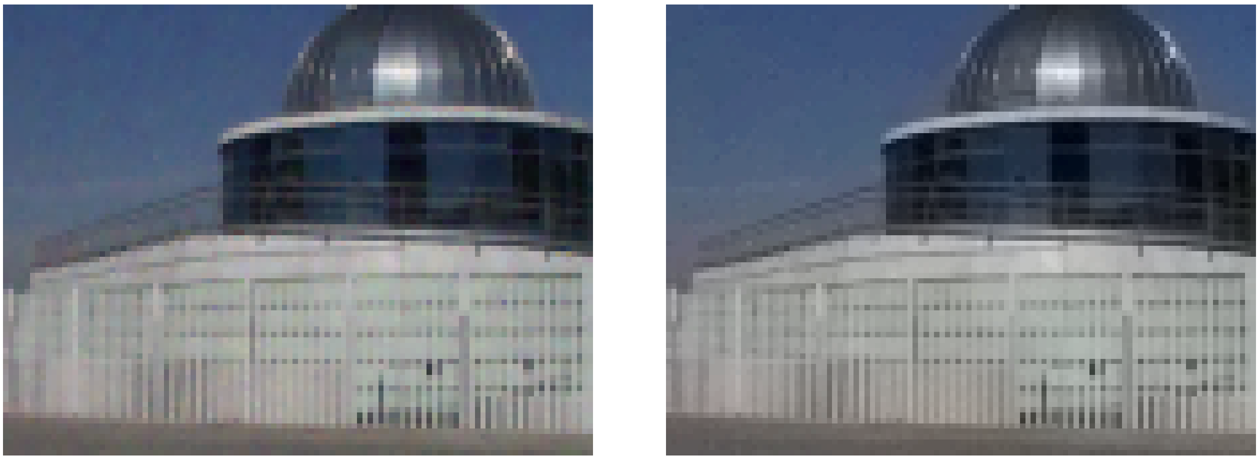
(a) Original image (b) Observed PAN image

(c) Observed MS image (d) Bicubic Interpolation



(e) NSCT method in [62] (f) SR method in [96]

Figure 5.1: *Results for the synthetic image*



(g) proposed method ($\lambda_{\alpha_b} = \lambda_{\beta_b} = \lambda_\gamma = 0$)

(h) proposed method ($\lambda_{\alpha_b} = 0.9$, $\lambda_{\beta_b} = 0.9$ and $\lambda_\gamma = 0.0$)

Figure 5.1: *Results for the synthetic image (contd.)*

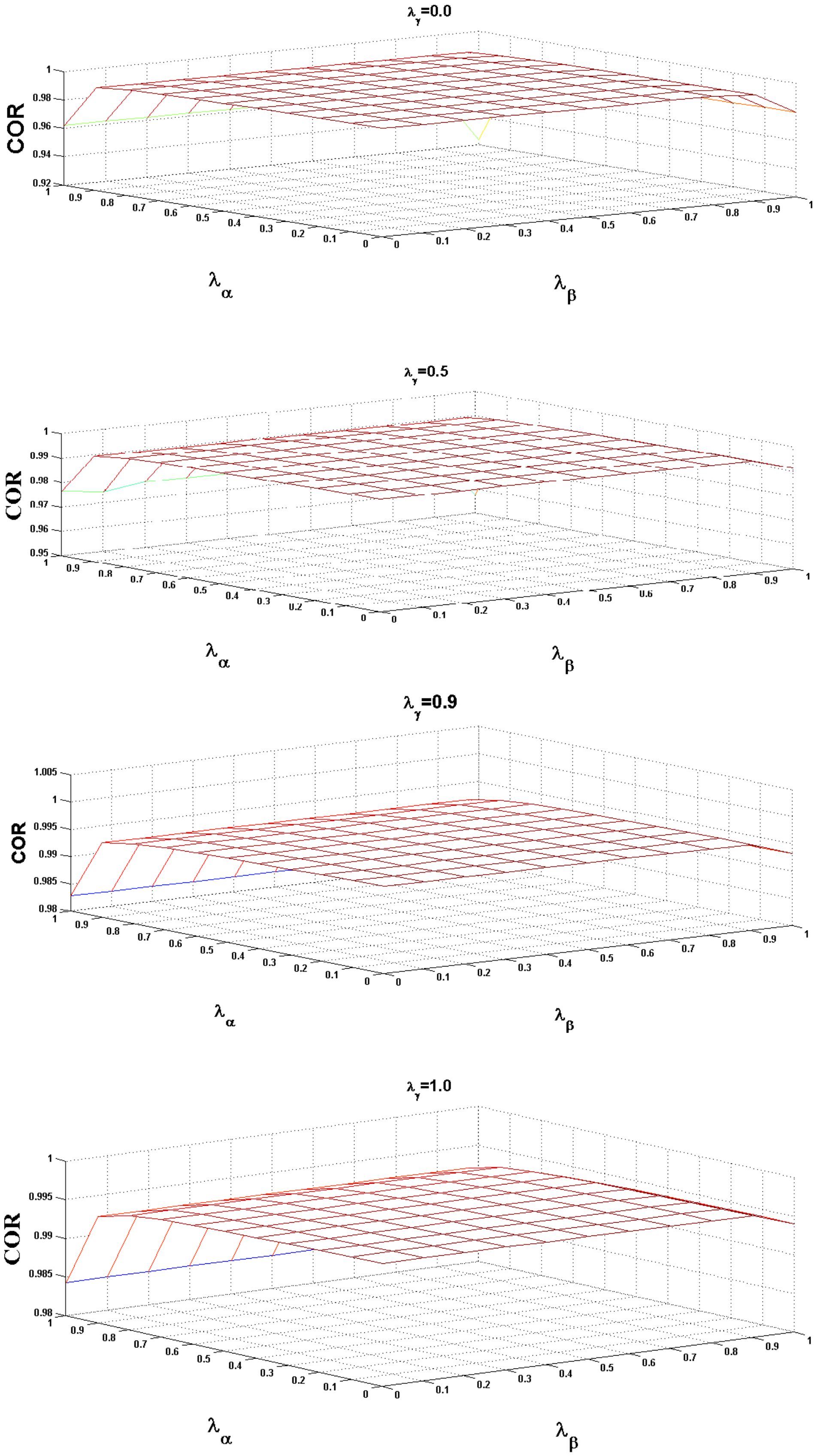


Figure 5.2: *Mean COR evolution for the synthetic HRMS image with different values of* λ_γ

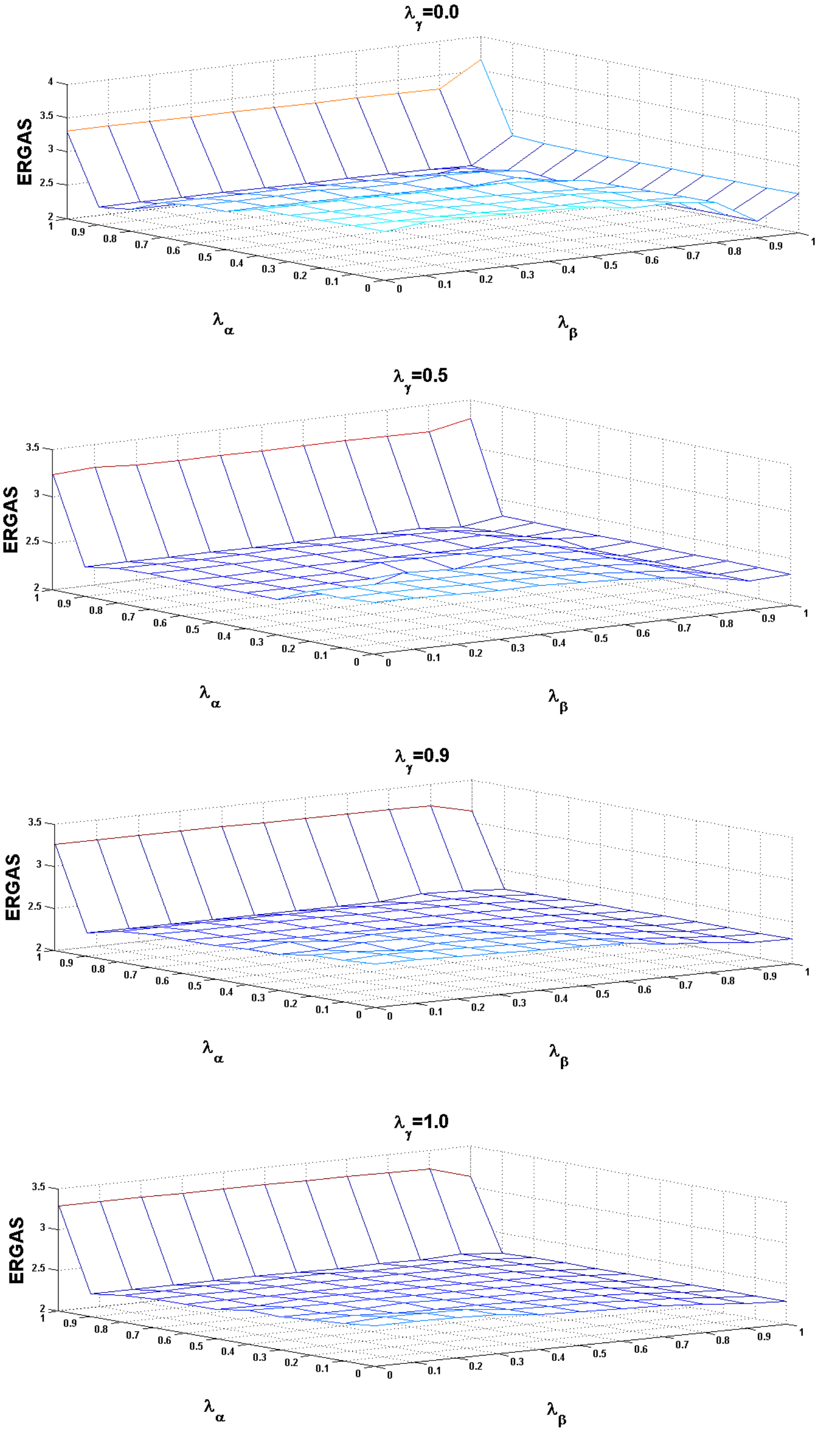


Figure 5.3: *ERGAS evolution for the synthetic HRMS image with different values of* λ_β

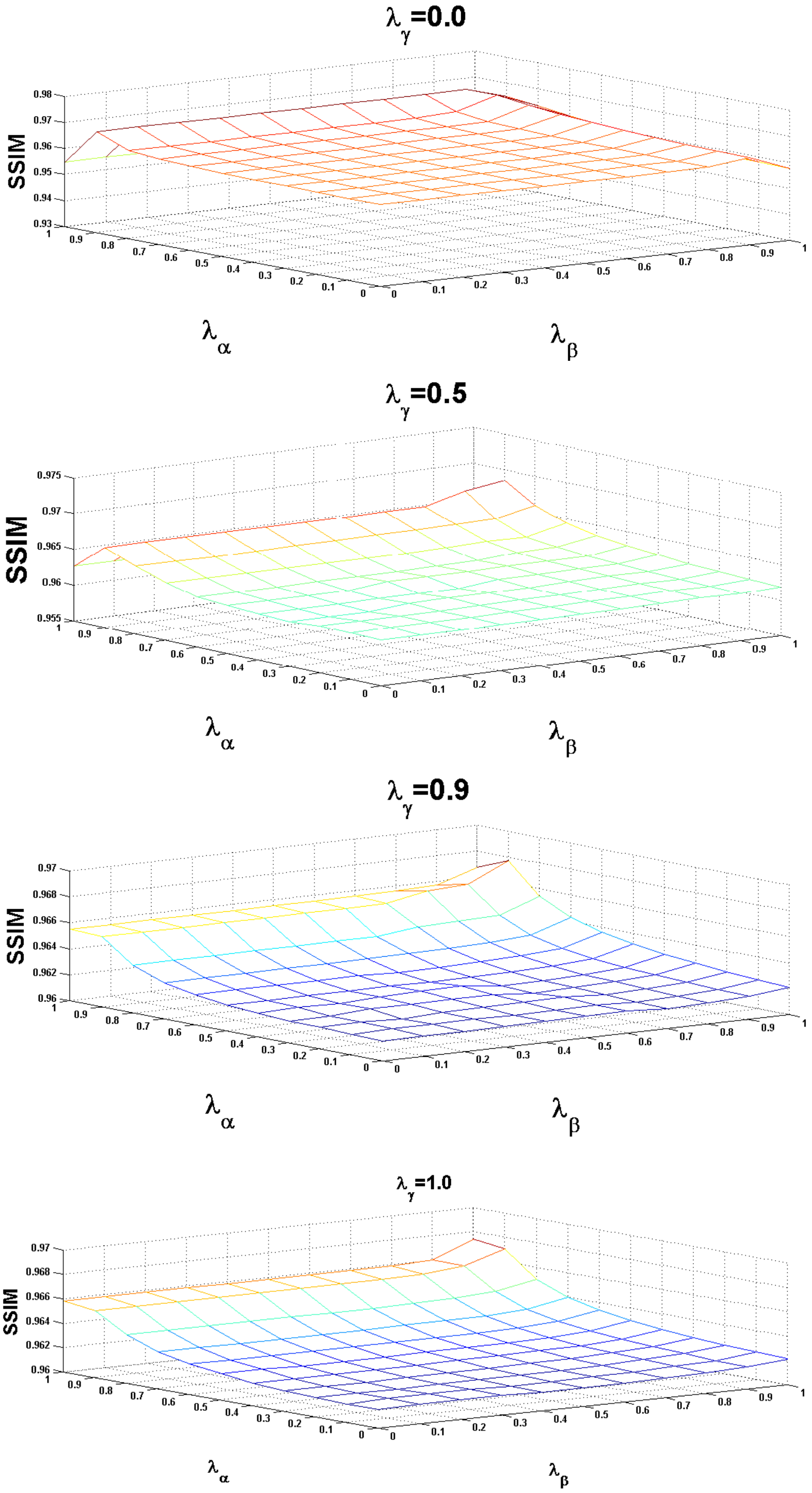
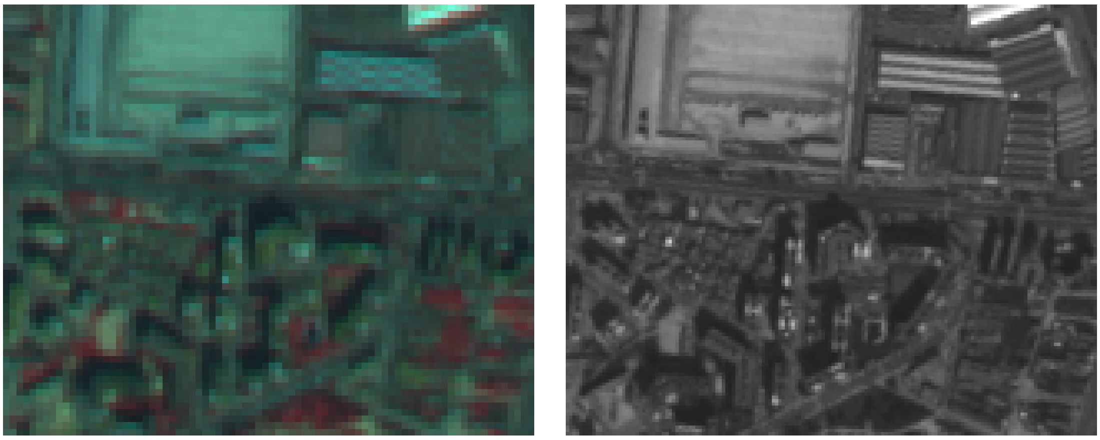
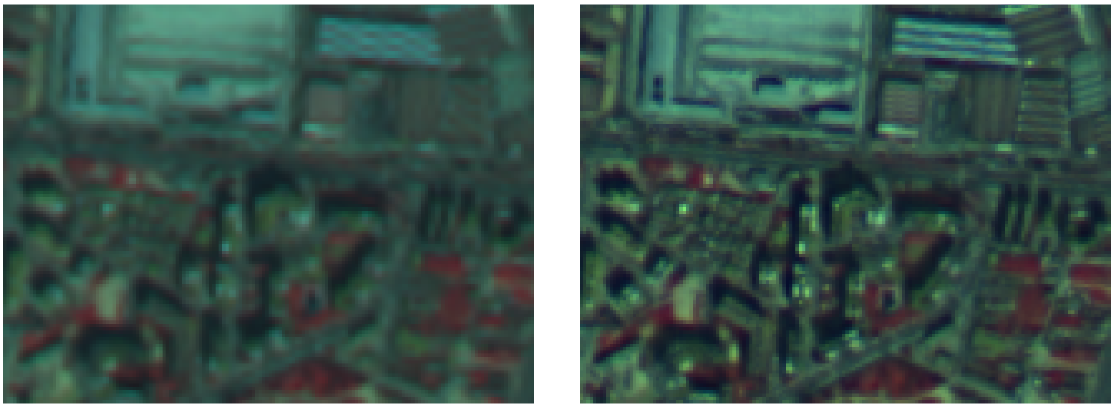


Figure 5.4: *Mean SSIM evolution for the synthetic HRMS image with different values of* λ_γ

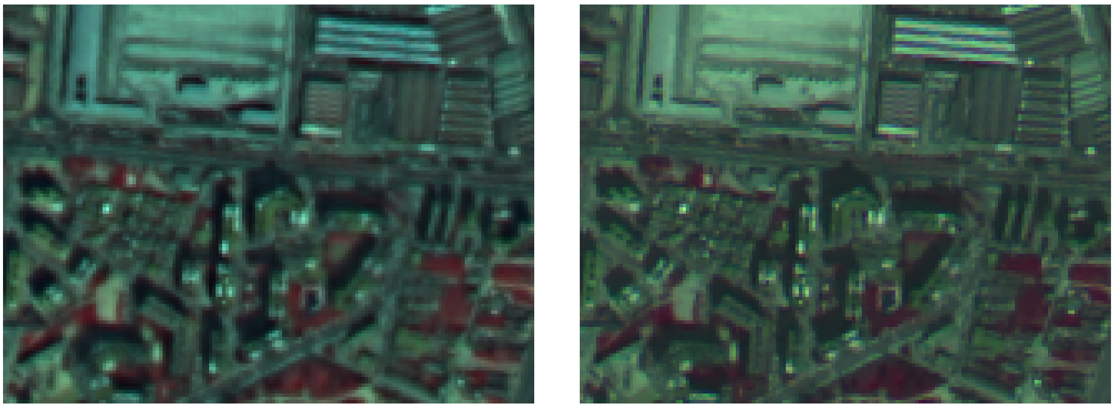


(a) Observed MS image (G-R-IR) (b) Observed PAN image

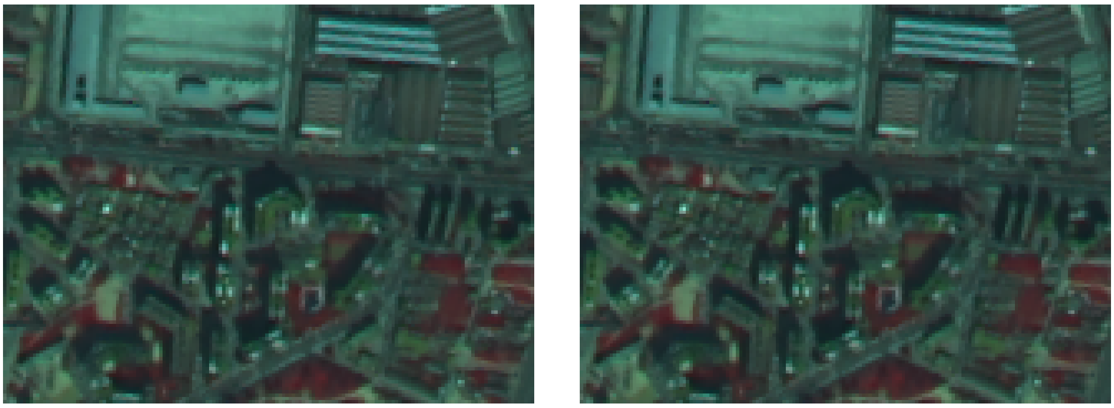
Figure 5.5: *Observed MS and PAN images of SPOT 5 sensor.*



(a) Bicubic interpolation (b) CiSper in [29]

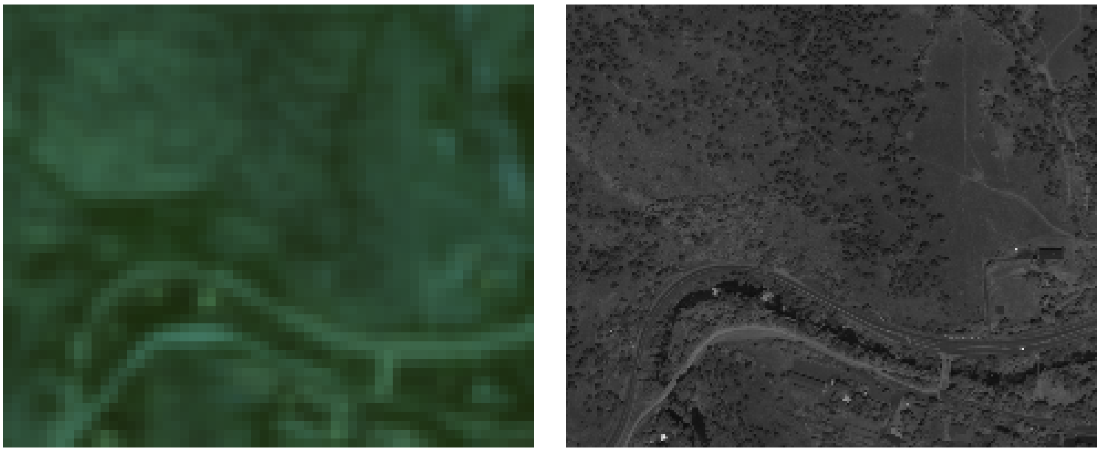


(c) NSCT method in [62] (d) SR method in [96]



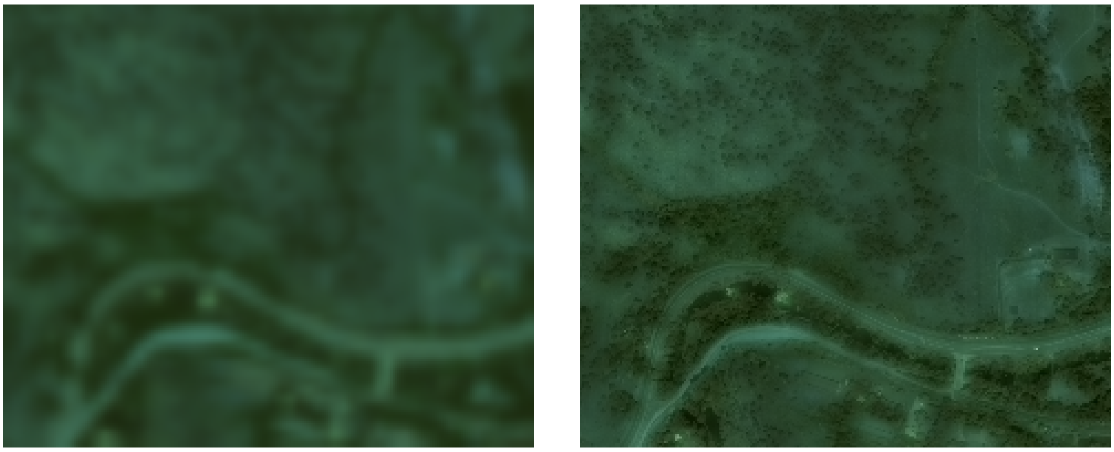
(e) proposed method ($\lambda_{\alpha_b} = \lambda_{\beta_b} = \lambda_\gamma = 0$) (f) Proposed($\lambda_{\alpha_b} = 0.9$, $\lambda_{\beta_b} = 0.9$ and $\lambda_\gamma = 0.0$)

Figure 5.6: *Results for the SPOT5 image*

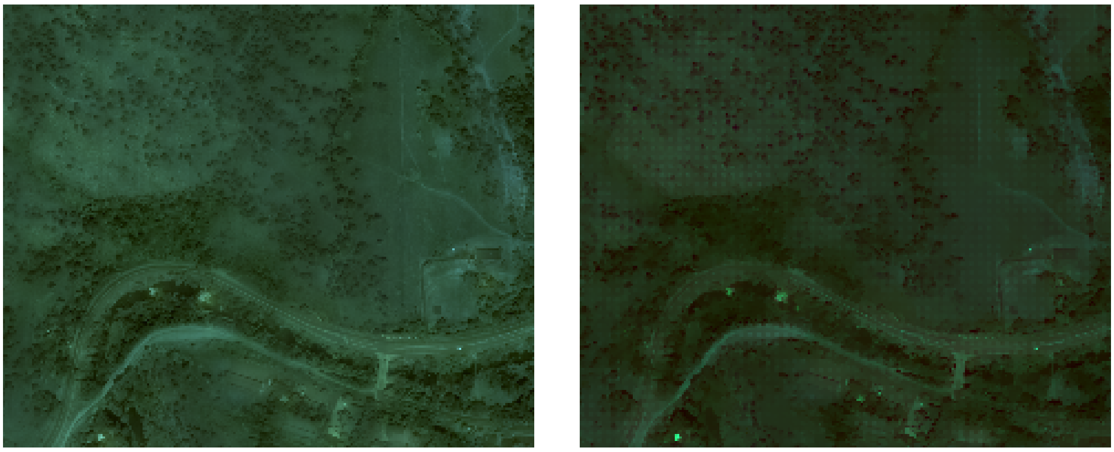


(a) Observed MS image (R-G-B) (b) Observed PAN image

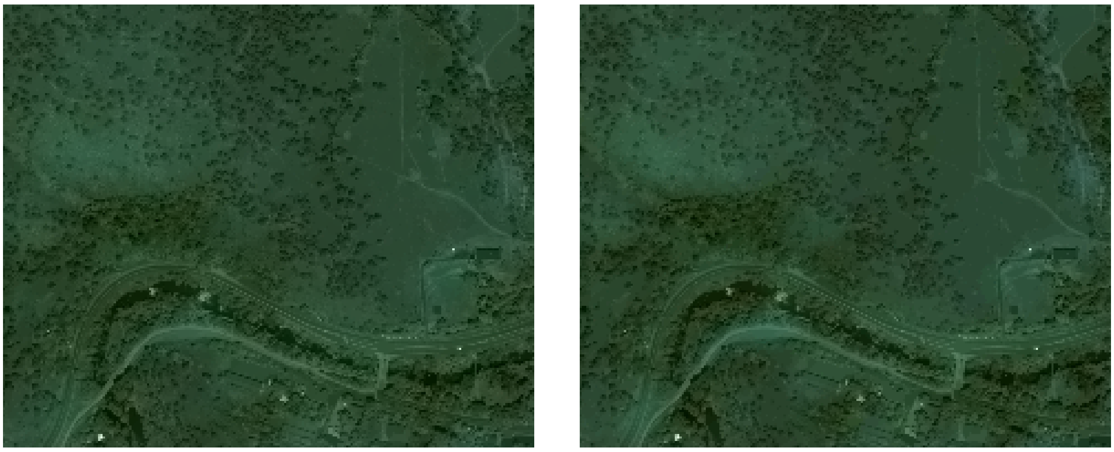
Figure 5.7: *Observed MS and PAN images of QuickBird sensor.*



(a) Bicubic interpolation (b) CiSper in [29]



(c) NSCT method in [62] (d) SR method in [96]



(e) proposed method ($\lambda_{\alpha_b} = \lambda_{\beta_b} = \lambda_\gamma = 0$) (f) Proposed($\lambda_{\alpha_b} = 0.9$, $\lambda_{\beta_b} = 0.9$ and $\lambda_\gamma = 0.0$)

Figure 5.8: *Results for the QuickBird image*

Chapter 6

General Contourlet Pansharpening Method using Bayesian Inference [1]

6.1 Introduction

In chapters 2 and 4, we revised the pansharpening methods based on NSCT and showed that they can use several schemes for the injection of the image details: substitution, which is the simplest one, addition, and other more complex mathematical models. Regardless of the injection scheme, NSCT based pansharpening methods start with images that have the same resolution so a preprocessing step is needed to upsample the MS image to the size of the PAN image. In this chapter we propose a Bayesian fusion method based on NSCT that comprises, as particular cases, substitution, addition and some other mathematical models.

This chapter is organized as follows. In section 6.2 the general algorithm for NSCT pansharpening, using different injection methods, is described and the used notation is introduced. Section 6.3 explains the Bayesian modeling for both residual and coefficients bands. The inference of the high resolution MS image residual and coefficients are explained in section 6.4. Experimental results and comparison with other methods are presented in 6.5 for synthetic, SPOT5 and QuickBird images and, finally, section 6.6 concludes the chapter.

---

[1] Part of this chapter was published in the 2010 European Signal Processing Conference (EUSIPCO-2010) [31]

6.2 General Pansharpening algorithm

Contourlet pansharpening methods are based on the ability of the NSCT transform for obtaining the high frequencies image details at different scales and different directions. Based on this ability, NSCT-based pansharpening methods intend to extract the detail information from the PAN image for injecting them into the MS image. The main drawback of NSCT-based pansharpening is the spectral distortion that it may produce. Pansharpening based on multiresolution decompositions such as NSCT, can be carried out in different ways, that following Amolins *et al.* [184] can be classified as:

A. *Substitution Model:* It involves completely replacing the MS image details, extracted with the NSCT transform, with those of the PAN image.

B. *Additive Model:* Add the NSCT details information of PAN image directly to the MS image bands, or to the NSCT details information of MS image.

C. *Mathematical-based Model:* It is more sophisticated than the above models since it applies a mathematical model to the details information in both PAN and MS images and then use the model to weight the information of both images in order to control noise and color bleeding effects.

All those models start with the observed low resolution MS image, Y, with B bands, $Y_b, b = 1,\dots,B$, each of size $P = M \times N$ pixels, and the PAN image, x, of size $p = m \times n$, with $M < m$ and $N < n$, which contains reflectance data in a single band that covers a wide area of the spectrum. Based on those observations, they find an estimation of y, the high resolution multispectral (HRMS) image, with B bands, $y_b, b = 1,\dots,B$, each of size $p = m \times n$ pixels. A general algorithm for pansharpening based on NSCT is summarized in Algorithm 6.1, to obtain an estimation of y, $\hat{y}$, from x and Y.

Algorithm 6.1 *NSCT pansharpening algorithm of x and $\{Y_b\}$ into $\{\hat{y}_b\}$*

---

1. Upsample each band of the MS image, Y_b, to the size of the PAN, x, and register them obtaining s_b, $b = 1,\dots,B$.

2. Apply NSCT decomposition on the PAN image x and registered MS image $\{s_b\}$,

$$x = C^r x + \sum_{l=1}^{L}\sum_{d=1}^{D} C^{ld} x = x^r + \sum_{l=1}^{L}\sum_{d=1}^{D} x^{ld}, \tag{6.1}$$

$$s_b = C^r s_b + \sum_{l=1}^{L}\sum_{d=1}^{D} C^{ld} s_b = s_b^r + \sum_{l=1}^{L}\sum_{d=1}^{D} s_b^{ld}, \quad b = 1,\dots,B, \tag{6.2}$$

where we are using the superscript r to denote the residual (low pass filtered version) NSCT coefficients band and the superscript ld to refer to the detail bands, with $l = 1,\dots,L$, representing the scale and $d = 1,\dots,D$, representing the direction for each coefficient band. For the sake of simplicity, for a given image, say o, we will use here the notation o^r and o^{ld} instead of $C^r o$ and $C^{ld} o$, respectively.

3. Merge the details of PAN $\{x^{ld}\}$ and MS $\{s_b^{ld}\}$ images getting $\{\hat{y}_b^{ld}\}$, keeping the residual image unchanged,

$$\hat{y}_b^{ld} = a^{ld} x^{ld} + b^{ld} s_b^{ld}, \tag{6.3}$$

$$\hat{y}_b^r = s_b^r. \tag{6.4}$$

4. Apply the inverse NSCT to merge the MS band coefficients $\{\hat{y}_b^r\}$, $\{\hat{y}_b^{ld}\}$, getting $\{\hat{y}_b\}$,

$$\hat{y}_b = \hat{y}_b^r + \sum_{l=1}^{L}\sum_{d=1}^{D} \hat{y}_b^{ld}, b = 1,\dots,B. \tag{6.5}$$

Note that for $b^{ld} = 0$, we get the substitution model, and for $a^{ld} = b^{ld} = 1, \forall l = 1,\dots,L$ and $d = 1,\dots,D$ we have the additive one, while using different a^{ld} and b^{ld} values we will get different weighted models proposed in the literature. In this chapter we propose to modify the merging strategy in step 3 of Algorithm 6.1 by using Bayesian inference as a mathematical way to estimate the details coefficients of the HRMS image from those of the PAN and MS images and to reconstruct HRMS residual image. The parameters incorporated in our model are estimated at each level of decomposition and direction for each band, providing a sound way to control the noise, preventing color bleeding and generalizing all previous models.

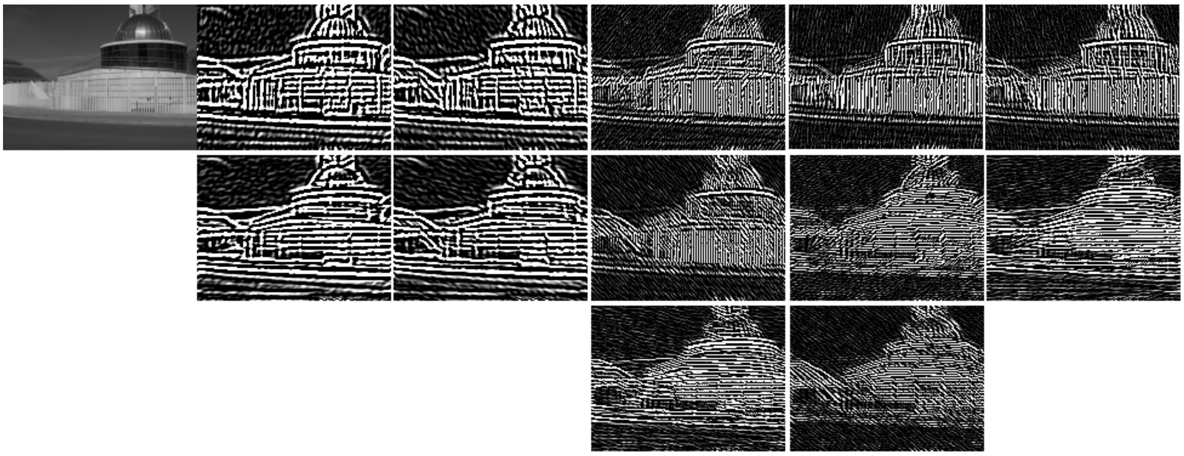


Figure 6.1: *NSCT contourlet transform of the first band of synthetic image. The image is decomposed into two scale levels, which are then decomposed into four and eight directional subbands.*

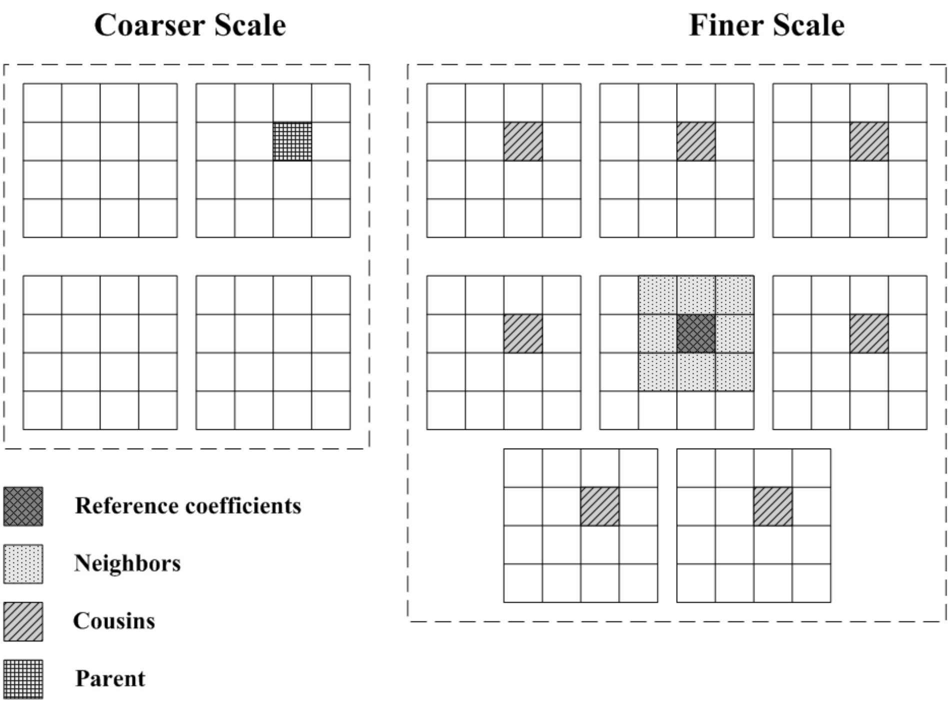
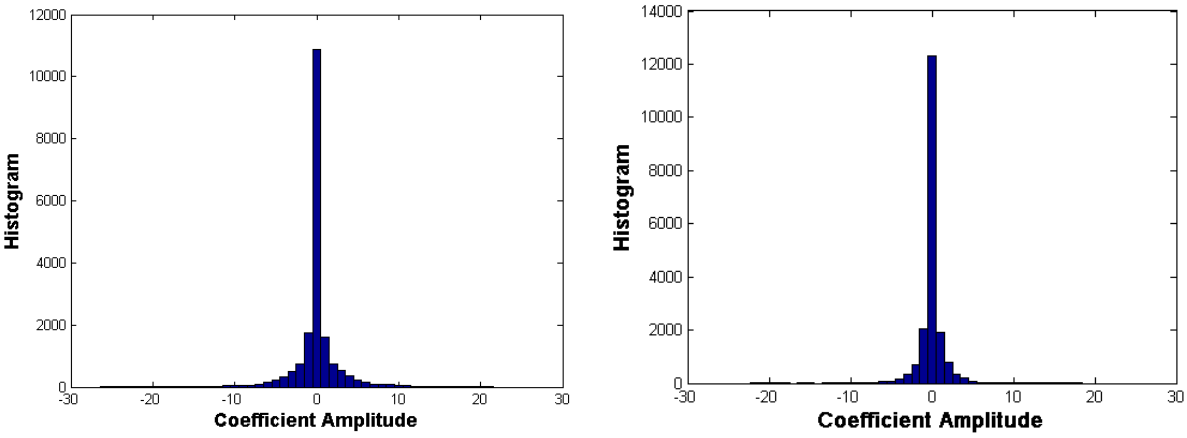


Figure 6.2: *NSCT coefficient relationships.*



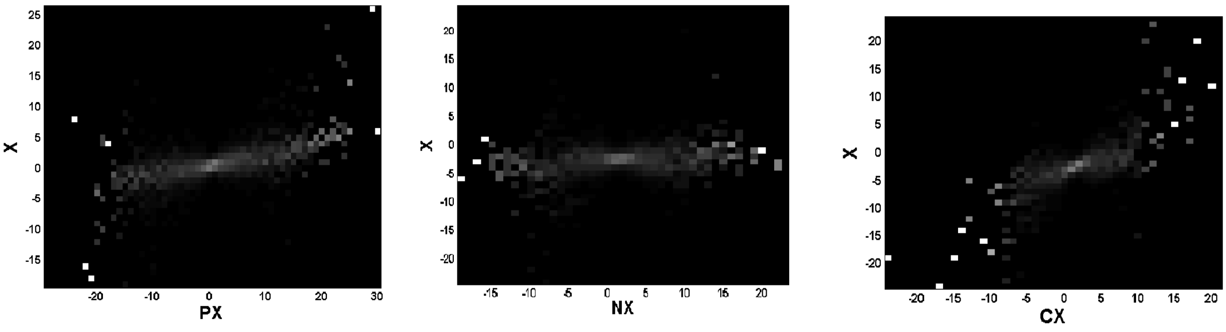
(a) (b)

Figure 6.3: *Marginal statistics of two finest subbands of the first band of synthetic image. The kurtosis of the two distributions are measured at (a) 15.58 and (b) 25.23.*

6.3 Bayesian Formulation

In order to model the relationships between the images in the contourlet domain, we have to study first the dependency of the contourlet coefficients (see Figure 6.1). As an important topic in the sparse representation of images, coefficients characteristics of subsampled contourlet have been studied in [185]. There are three relationships in contourlet coefficients, which are shown in Figure 6.2. The reference coefficient has eight neighbors (NX) in the same subband, parent (PX) at the same spatial location in the immediately coarser scale and cousins (CX) at the same scale and spatial location but in different directional subbands. Figure 6.3 plots the histogram of two finest coefficient subbands in Figure 6.1. These distributions exhibit a sharp peak at zero amplitude and heavy tails to both sides of the peak. This implies that the contourlet transform is sparse, as the majority of coefficients are close to zero. The kurtosis of the two shown distributions are 15.58 and 25.23, which are much higher than the kurtosis of 3 for Gaussian distributions. Thus, the subband marginal distributions of natural images in the contourlet domain are highly non-Gaussian. This marginal statistics only describe the individual behaviors of transform coefficients without accounting for their dependencies.

The authors in [185] suggested the study of the conditional distributions of contourlet coefficients, conditioned on their parents, neighbors and cousins, in order to know the relationship



(a) (b) (c)

Figure 6.4: *Conditional distribution of a finest subband of the first band of synthetic image, conditioned on (a) parent$p(X|PX)$, (b) neighbor $p(X|NX)$ and (c) cousin $p(X|CX)$*

between these coefficients for the subsampled contourlet transform. Figure 6.4 shows those conditional distributions for the non-subsampled contourlet transform coefficients. We notice that all of these conditional distributions exhibit a "bow-tie" shape where the variance of the coefficients is related to the magnitude of the conditioned coefficient. Moreover, even though coefficients are correlated due to the over completeness of the non-subsampled contourlet transform, they are approximately decorrelated since conditional expectation $E[X|.] \approx 0$. Therefore, we conclude that contourlet coefficients of natural images are approximately uncorrelated yet dependent on each other. These dependencies, however are local. Figure 6.5 shows the conditional distribution of contourlet coefficients conditioned on distant relatives and neighbors of the first band of the synthetic image. We observe that these conditional distributions are approximately invariant to the conditioned value, indicating independence.

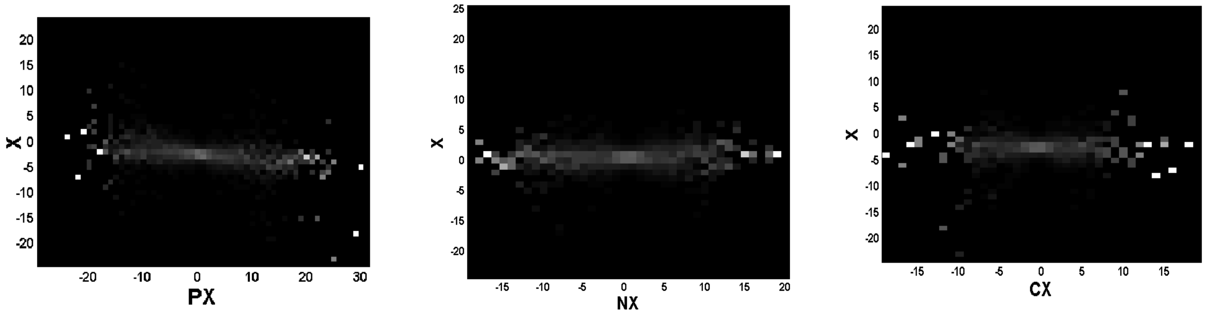
In this chapter we are going to relate the HRMS image and the upsampled MS image following the model in [186]. Starting from modeling the relation between y_b and s_b, as

$$s_b = y_b + n_b, \tag{6.6}$$

with n_b being the capture noise assumed to be Gaussian with zero mean and covariance matrix $\sigma_b^2 I$. Applying the NSCT to both sides of Eq. (6.6) and since the NSCT decomposition is just a convolution, we can write

$$s_b^r + \sum_{l=1}^{L}\sum_{d=1}^{D} s_b^{ld} = C^r(y_b + n_b) + \sum_{l=1}^{L}\sum_{d=1}^{D} C^{ld}(y_b + n_b) = y_b^r + n_b^r + \sum_{l=1}^{L}\sum_{d=1}^{D} y_b^{ld} + n_b^{ld}. \tag{6.7}$$

Assuming that the noise is separable, that is, it decomposes in the same way as the image does,



(a) (b) (c)

Figure 6.5: *Distribution of a finest subband of of the first band of synthetic image conditioned on (a) ancestors, (b) neighbors, and (c) cousins, all at distances of three coefficients away.*

and therefore low frequencies of the noise affect only to the low frequencies of the MS band, and the high frequencies in each direction affect only to their corresponding high frequencies in each direction, we can write that

$$s_b^r = y_b^r + n_b^r, \tag{6.8}$$

$$s_b^{ld} = y_b^{ld} + n_{s_b}^{ld}, \tag{6.9}$$

In those equations, the noise of the residual band, n_b^r will follow a Gaussian distribution of the form

$$n_b^r \sim \mathcal{N}(0, (\sigma_b^2 C^r (C^r)^t)), \tag{6.10}$$

and we are going to approximate the covariance matrix $\sigma_b^2 C^r (C^r)^t$ by $(\beta_b^r)^{-1} I$, where β_b^r, the inverse of the variance of the noise of the residual band of the MS band b, is a parameter that have to be estimated. We are going to use the same approximation for the noise of the detail bands, n_b^{ld}, approximating it by a Gaussian distribution with zero mean and covariance matrix $(\beta_b^{ld})^{-1} I$, with β_b^{ld} the inverse of the noise of the detail band at level l and direction d of the MS band b.

Since the PAN image contains the details of the high resolution MS image but lacks of its spectral information, and the MS image have the spectral information of the HRMS images, the relationship between the HRMS band coefficients and the PAN image could be written,

$$x^{ld} = y_b^{ld} + n_x^{ld}, \tag{6.11}$$

where n_x^{ld} is the noise of the coefficients bands at each NSCT decomposition level, l, and direction, d, for PAN image, that, as we already did for the MS image, it is approximated by a Gaussian distribution with zero mean and covariance matrix $(\gamma_b^{ld})^{-1}I$.

Note that, with this modeling, we have decoupled each one of the band of the contourlet transform and since we have demonstrated that bands are uncorrelated,, we can do the estimation of each band independently of the other bands. Note also that the residual band of the contourlet transform have different characteristics that the coefficients bands and so, we are going to propose different prior models for the residual band and the coefficient bands and different optimization procedures will be applied to the residual and the coefficients bands.

The study in [185] concludes that the contourlet coefficients are non-Gaussian but conditionally Gaussian. In other words, the contourlet coefficients of natural images may be accurately modeled, for instance, by mixtures of Gaussian distributions whose variances depend on their generalized neighborhood coefficients. While some authors used, l1 or p-norm based priors, with $p \leq 2$, on the coefficients (see [187], for instance), in this dissertation we chose the TV prior that penalizes differences between neighbor coefficients. The reason is that we observed that the contourlets coefficients in the detail bands are comprised of smooth regions with a few strong edges and this can be conveniently modeled by the TV prior on the coefficient imposing smoothness on the regions while preserving the edges. Figure 6.6 shows the distribution of the difference between neighbor coefficients for a given band. The red line in Figure 6.6 depicts a TV prior fitted to this distribution. From the figure it is clear that the TV prior adjust almost perfectly to the distribution of the difference of neighbor coefficients and so, it is possible to use the TV as a prior for the contourlet coefficient bands.

Let us now explain in detail the Bayesian modeling and inference for the residual and coefficients bands.

6.3.1 Bayesian Modeling for the Residual Image

The Bayesian formulation of the residual high resolution multispectral image reconstruction problem requires the definition of the joint distribution, $p(\Omega_b^r, y_b^r, s_b^r)$, of the residual of the upsampled multispectral observation, s_b^r, the unknown residual for high resolution multispectral image y_b^r, and the hyperparameters Ω_b^r. Then, the posterior distribution of the unknowns given the residual of the upsampled MS band $p(\Omega_b^r, y_b^r | s_b^r)$ has to be calculated and used to estimate the residual for the high resolution image band y_b. To model the joint distribution, we

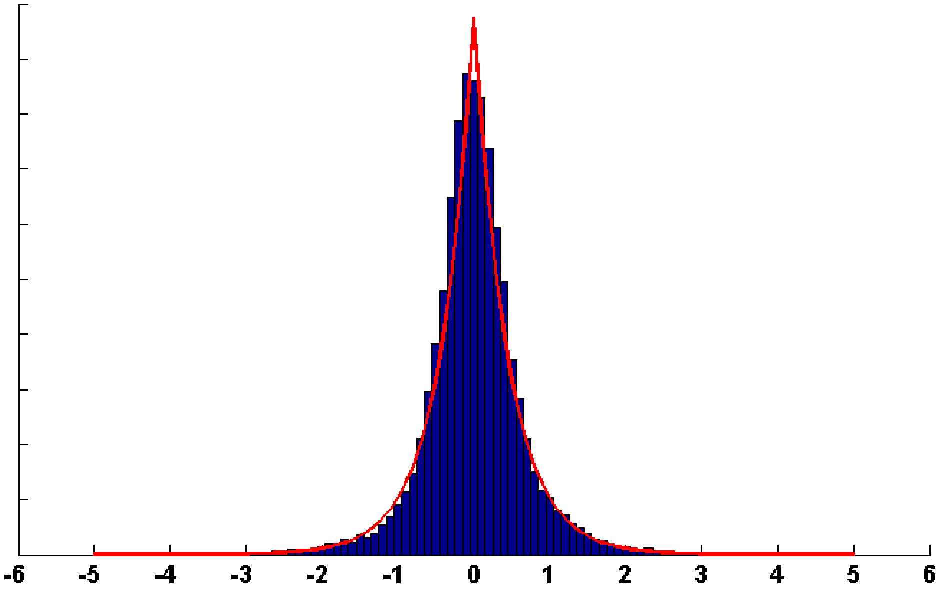


Figure 6.6: *Marginal distribution of the difference between neighbor coefficients and the distribution of TV prior.*

utilize the hierarchical Bayesian paradigm.

As already described in the previous chapters, in the hierarchical approach to our high resolution residual image reconstruction problem we have two stages. In the first stage, knowledge about the structural form of the low resolution observation noise and the structural behavior of the high resolution multispectral image is used in forming $p(s_b^r|y_b^r,\Omega_b^r)$ and $p(y_b^r|\Omega_b^r)$, respectively. These noise and image models depend on the unknown hyperparameters Ω_b^r. In the second stage a hyperprior on the hyperparameters is defined, thus allowing the incorporation of information about these hyperparameters into the process. We note here that each of the two above mentioned conditional distributions will depend only on a subset of Ω_b^r, but we use this more general notation until we precisely describe the hyperparameters.

For Ω_b^r, y_b^r and s_b^r the following distribution is defined

$$p(\Omega_b^r,y_b^r,s_b^r) = p(\Omega_b^r)p(y_b^r|\Omega_b^r)p(s_b^r|y_b^r,\Omega_b^r), \tag{6.12}$$

and inference is based on $p(\Omega_b^r,y_b^r|s_b^r)$. Let us now describe the form of these distributions

First Stage: Prior model on residual of the HRMS image

Bayesian methods start with the definition of a prior model where we incorporate the expected characteristics of the original NSCT coefficients for the residual band. Since the residual band is a smoothed version of the original HRMS band, for y_b^r we choose a quadratic prior, the Simultaneously Autoregressive (SAR) prior, given by

$$p(y_b^r|\Omega_b^r) = p(y_b^r|\alpha_b^r) \propto (\alpha_b^r)^{(p-1)/2} \exp\left\{-\frac{1}{2}\alpha_b^r \left\|Q y_b^r\right\|^2\right\}, \tag{6.13}$$

where Q denotes the Laplacian operator, and α_b^r is the model parameter that control the degree of smoothness of the residual MS band b.

First Stage: Observation model of the low resolution MS residual image

From the observation model of the MS image in Eq. (6.8) we have the following conditional probability distribution

$$p(s_b^r|y_b^r,\Omega_b^r) = p(s_b^r|y_b^r,\beta_b^r) \propto (\beta_b^r)^{p/2} \exp\left\{-\frac{1}{2}\beta_b^r \left\|s_b^r - y_b^r\right\|^2\right\}, \tag{6.14}$$

Second Stage: Hyperprior on the hyperparameters

The set of hyperparameters, Ω_b^r, will be formed by the prior model hyperparameter α_b^r and the noise hyperparameter β_b^r. Hence, $\Omega_b^r = (\alpha_b^r, \beta_b^r)$. As we already justified in Chapter 5, we will use as hyperprior for each of the hyperparameters, $w \in \Omega_b^r$, the gamma distribution

$$p(w) = \Gamma(w|a_w, c_w) = \frac{(c_w)^{a_w}}{\Gamma(a_w)} w^{a_w-1} \exp[-c_w w] \tag{6.15}$$

where $w > 0$ denotes a hyperparameter, and the two parameters $a_w > 0$ and $c_w > 0$ are respectively, the shape and the inverse scale parameters of the distribution, that are assumed to be known. We will then use the following distribution as the hyperprior on the hyperparameters

$$p(\Omega_b^r) = p(\alpha_b^r, \beta_b^r) = p(\alpha_b^r)p(\beta_b^r), \tag{6.16}$$

where the hyperprior for each hyperparameter $w \in \Omega_b^r$, is given by Eq. (6.15).

Finally, combining the first and second stage of the problem modeling we have the global distribution in Eq. (6.12).

6.3.2 Bayesian Modeling for the Coefficient Bands

Following the same steps we already performed for the residual image, when modeling the coefficient bands into the Bayesian framework we have two stages. In the first stage, knowledge about the structural form noise coefficients and the structural behavior of the high resolution multispectral image coefficients is used in forming $p(s_b^{ld}, x^{ld} | y_b^{ld}, \Omega_b^{ld})$ and $p(y_b^{ld} | \Omega_b^{ld})$, respectively. These noise and image models depend on the unknown hyperparameters Ω_b^{ld}. In the second stage a hyperprior on the hyperparameters is defined, thus allowing the incorporation of information about these hyperparameters into the process.

For Ω_b^{ld}, y_b^{ld}, x^{ld} and s_b^{ld} the following joint distribution is defined

$$p(\Omega_b^{ld}, y_b^{ld}, x^{ld}, s_b^{ld}) = p(\Omega_b^{ld}) p(y_b^{ld} | \Omega_b^{ld}) p(s_b^{ld} | y_b^{ld}, \Omega_b^{ld}) p(x^{ld} | y_b^{ld}, \Omega_b^{ld}), \tag{6.17}$$

and inference is based on $p(\Omega_b^{ld}, y_b^{ld} | s_b^{ld}, x^{ld})$.

First Stage: Prior model on MS Coefficient image

For the coefficient band, we choose a prior model based on the Total Variation (TV) [177]. As we mentioned before, the idea behind this model is to consider the image as a set of relatively smooth objects or regions separated by strong edges, such as the coefficients of the NSCT. This model enforces smoothness within the objects in each coefficient band while preserving their edges. Following [96], we assume that there is no correlation between the different coefficient bands, hence defining

$$p(y_b^{ld} | \alpha_b^{ld}) \propto (\alpha_b^{ld})^{p/2} \exp\left\{ -\alpha_b^{ld} \mathrm{TV}(y_b^{ld}) \right\}, \tag{6.18}$$

with $\mathrm{TV}(y_b^{ld}) = \sum_{i=1}^{p} \sqrt{(\Delta_i^h(y_b^{ld}))^2 + (\Delta_i^v(y_b^{ld}))^2}$ where $\Delta_i^h(y_b^{ld})$ and $\Delta_i^v(y_b^{ld})$ represent the horizontal and vertical first order differences at pixel i, respectively, and α_b^{ld} is the model parameter of the MS band b coefficients at level l and direction d.

First Stage: Observation model of the low resolution MS and Panchromatic coefficient image

From the observation model of the MS and panchromatic image in Eqs. (6.9) and (6.11) respectively, we have the following probability distribution

$$p(s_b^{ld}|y_b^{ld}) \propto (\beta_b^{ld})^{p/2} \exp\left\{-\frac{1}{2}\beta_b^{ld} \left\|s_b^{ld} - y_b^{ld}\right\|^2\right\}, \tag{6.19}$$

$$p(x^{ld}|y_b^{ld}) \propto (\gamma_b^{ld})^{p/2} \exp\left\{-\frac{1}{2}\gamma_b^{ld} \left\|x^{ld} - y_b^{ld}\right\|^2\right\}, \tag{6.20}$$

Second Stage: Hyperprior on the hyperparameters

From the distributions defined in Eqs. (6.18), (6.19) and (6.20) we have that the set of hyperparameters is $\Omega_b^{ld} = (\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld})$, and as in the residual case we will use gamma distribution as a hyperprior resulting in.

$$p(\Omega_b^{ld}) = p(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}) = p(\alpha_b^{ld})p(\beta_b^{ld})p(\gamma_b^{ld}), \tag{6.21}$$

where the hyperprior for each hyperparameter $w \in \Omega_b^{ld}$, is given by Eq. (6.15).

6.4 Bayesian Inference

In this chapter, we have defined different prior distributions for the residual and coefficient bands. For the HRMS residual image, all used distributions are quadratics, and so in this case the evidence approach can be used to perform the inference. However for the HRMS coefficient bands, the evidence can not be used due to the TV prior, and the variational approach will be used to perform the inference. Let us now study in detail how the inference is carried out for the residual and coefficients images.

6.4.1 Bayesian Inference for the residual image

The set of all unknowns is given by $(\Omega_b^r, y_b^r) = (\alpha_b^r, \beta_b^r, y_b^r)$. The Bayesian paradigm dictates that inference on (Ω_b^r, y_b^r) should be based on

$$p(\alpha_b^r, \beta_b^r, y_b^r|s_b^r) = \frac{p(\alpha_b^r, \beta_b^r, y_b^r, s_b^r)}{p(s_b^r)} \propto p(\alpha_b^r, \beta_b^r, y_b^r, s_b^r) \tag{6.22}$$

We are going to use here the evidence approach in order to do this inference (see section 3.3.2). To do the estimation of α_b^r, β_b^r and y_b^r, we have to follow two steps:

1. *Step I: Estimation of the parameters*, $\hat{\alpha}_b^r$ and $\hat{\beta}_b^r$ are first selected as

$$\hat{\alpha}_b^r, \hat{\beta}_b^r = \arg\max_{\alpha_b^r, \beta_b^r} \log p(s_b^r | \alpha_b^r, \beta_b^r), \tag{6.23}$$

   where

$$p(s_b^r | \alpha_b^r, \beta_b^r) = \int_{y_b^r} p(y_b^r | \alpha_b^r) p(s_b^r | y_b^r, \beta_b^r) d y_b^r. \tag{6.24}$$

2. *Step II: Estimation of the original multispectral residual image,* Once the parameters α_b^r and β_b^r have been estimated, the estimation of the original MS residual image, $\hat{y}_b^r$, is selected as the image satisfying

$$\hat{y}_b^r = \underset{y_b^r}{\operatorname{argmax}} \, p(y_b^r | s_b^r, \hat{\alpha}_b^r, \hat{\beta}_b^r) = \underset{y_b^r}{\operatorname{argmin}} [-2 \log p(y_b^r | s_b^r, \hat{\alpha}_b^r, \hat{\beta}_b^r),] \tag{6.25}$$

   which produces,

$$\hat{y}_b^r = \underset{y_b^r}{\operatorname{argmin}} \left[\hat{\alpha}_b^r \, \| Q y_b^r \|^2 + \hat{\beta}_b^r \, \| s_b^r - y_b^r \|^2 \right]. \tag{6.26}$$

   Differentiating with respect to y_b^r, and setting the result equal to zero, we obtain

$$\hat{\alpha}_b^r Q^t Q y_b^r - \hat{\beta}_b^r s_b^r + \hat{\beta}_b^r y_b^r = 0 \tag{6.27}$$

   and so, an estimate of y_b^r, $\hat{y}_b^r$, can be obtained as

$$\hat{y}_b^r = COV^{-1}(\alpha_b^r, \beta_b^r) \hat{\beta}_b^r s_b^r \tag{6.28}$$

   where

$$COV(\alpha_b^r, \beta_b^r) = \hat{\beta}_b^r I_p + \hat{\alpha}_b^r Q^t Q. \tag{6.29}$$

Note that we are using maximum likelihood for estimating the parameters and the *Maximum a Posterior* (MAP) for estimating $\hat{y}_b^r$. Furthermore, the steps I and II are performed within the hierarchical Bayesian approach iteratively in order to perform the image and parameter estimation simultaneously, by including priors on the unknown parameters α_b^r and β_b^r.

Let us examine the parameter estimation process in detail. Fixing α_b^r and β_b^r and expanding the function

$$M(y_b^r, s_b^r | \alpha_b^r, \beta_b^r) = \alpha_b^r \, \| Q y_b^r \|^2 + \beta_b^r \, \| s_b^r - y_b^r \|^2 \tag{6.30}$$

around $\hat{y}_b^r$, we have

$$M(y_b^r, s_b^r | \alpha_b^r, \beta_b^r) = M(\hat{y}_b^r, s_b^r | \alpha_b^r, \beta_b^r) + \frac{1}{2} (y_b^r - \hat{y}_b^r)^t COV(\alpha_b^r, \beta_b^r)(y_b^r - \hat{y}_b^r). \tag{6.31}$$

Therefore

$$p(s_b^r|\alpha_b^r,\beta_b^r) = \frac{1}{p(\alpha_b^r)p(\beta_b^r)} \exp\left\{-\frac{1}{2}M(\hat{y}_b^r, s_b^r|\alpha_b^r,\beta_b^r)\right\} \\ \times \int_{y_b^r} \exp\left\{-\frac{1}{2}(y_b^r-\hat{y}_b^r)^t COV(\alpha_b^r,\beta_b^r)(y_b^r-\hat{y}_b^r)\right\} dy_b^r. \tag{6.32}$$

We then have that two times the likelihood of α_b^r and β_b^r is equal to

$$2\mathscr{L}_{s_b^r}(\alpha_b^r,\beta_b^r) = -\alpha_b^r \left\|Qy_b^r\right\|^2 - \beta_b^r \left\|s_b^r - y_b^r\right\|^2 - \log\det(COV(\alpha_b^r,\beta_b^r)) - 2\log p(\alpha_b^r) - 2\log p(\beta_b^r) + const. \tag{6.33}$$

We now differentiate $-2\mathscr{L}_{s_b^r}(\alpha_b^r,\beta_b^r)$ with respect to α_b^r and β_b^r to find the conditions which are satisfied at the maximum. We have

$$\frac{a_{\alpha_b^r} + \frac{p-1}{2}}{(\alpha_b^r)^{k+1}} - c_{\alpha_b^r} = \left\|Q(\hat{y}_b^r)^k\right\|^2 + trace\left[COV^{-1}((\alpha_b^r)^k,(\beta_b^r)^k)Q^tQ\right], \tag{6.34}$$

and

$$\frac{a_{\beta_b^r} + \frac{p}{2}}{(\beta_b^r)^{k+1}} - c_{\beta_b^r} = \left\|s_b^r - (\hat{y}_b^r)^k\right\|^2 + trace\left[COV^{-1}((\alpha_b^r)^k,(\beta_b^r)^k)\right], \tag{6.35}$$

consequently we have the following estimation for $(\alpha_b^r)^{k+1}$ and $(\beta_b^r)^{k+1}$

$$(\alpha_b^r)^{k+1} = \frac{a_{\alpha_b^r} + \frac{p-1}{2}}{c_{\alpha_b^r} + \left\|Q(\hat{y}_b^r)_k\right\|^2 + trace(COV^{-1}((\alpha_b^r)^k,(\beta_b^r)^k)Q^tQ)}, \tag{6.36}$$

$$(\beta_b^r)^{k+1} = \frac{a_{\beta_b^r} + \frac{p}{2}}{c_{\beta_b^r} + \left\|s_b^r - (\hat{y}_b^r)_k\right\|^2 + trace(COV^{-1}((\alpha_b^r)^k,(\beta_b^r)^k))}. \tag{6.37}$$

The following algorithm is proposed for the simultaneous estimation of the parameters and the residual band of the HRMS image.

Algorithm 6.2 *Posterior parameter and residual image estimation*

```
for $k = 1, 2, \ldots$, until a stopping criterion is met
1. Choose $(\alpha_b^r)^1$ and $(\beta_b^r)^1$.
2. Compute $(\hat{y}_b^r)^1$ using Eq. (6.28) with $\hat{\alpha}_b^r = (\alpha_b^r)^1$ and $\hat{\beta}_b^r = (\beta_b^r)^1$.
3. For $k = 1, 2, \ldots$ until a stopping criterion is met,
    a. Calculate $(\alpha_b^r)^{k+1}$ and $(\beta_b^r)^{k+1}$ by substituting $(\alpha_b^r)^k$ and $(\beta_b^r)^k$ in the right hand size of
       Eqs. (6.36) and (6.37), respectively.
    b. Compute $(\hat{y}_b^r)^{k+1}$ using Eq. (6.28) with $\hat{\alpha}_b^r = (\alpha_b^r)^k$ and $\hat{\beta}_b^r = (\beta_b^r)^k$.
```

6.4.2 Bayesian Inference for the coefficients

Now we will move to do the inference for the coefficients. The set of all unknowns is given by $(\Omega_b^{ld}, y_b^{ld}) = (\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}, y_b^{ld})$. The Bayesian paradigm dictates that inference on $(\Omega_b^{ld}, y_b^{ld})$ should be based on

$$p(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}, y_b^{ld} | s_b^{ld}, x^{ld}) = \frac{p(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}, y_b^{ld}, s_b^{ld}, x^{ld})}{p(s_b^{ld}, x^{ld})} \tag{6.38}$$

Once $p(\Omega_b^{ld}, y_b^{ld} | s_b^{ld}, x^{ld})$ has been calculated, y_b^{ld} can be integrated out to obtain $p(\Omega_b^{ld} | s_b^{ld})$. This distribution is then used to simulate or select the hyperparameters. If a point estimate,

$$\hat{\Omega}_b^{ld} = (\hat{\alpha}_b^{ld}, \hat{\beta}_b^{ld}, \hat{\gamma}_b^{ld}), \tag{6.39}$$

is required, then the mode or the mean of this posterior distribution can be used. Finally, a point estimate of the original multispectral image $\hat{y}_b^{ld}$ can be obtained by maximizing $p(y_b^{ld} | s_b^{ld}, x^{ld}, \hat{\Omega}_b^{ld})$. Alternatively the mean value of this posterior distribution can be selected as the estimate of the multispectral image.

From the above discussion it is clear that in order to perform inference we need to either calculate or approximate the posterior distribution $p(\Omega_b^{ld}, y_b^{ld} | s_b^{ld}, x^{ld})$. Since $p(\Omega_b^{ld}, y_b^{ld} | s_b^{ld}, x^{ld})$ can not be found in closed form, because $p(s_b^{ld}, x^{ld})$ cannot be calculated analytically, we will apply the variational methodology to approximate the posterior distribution by another distribution, $q(\Omega_b^{ld}, y_b^{ld})$, that minimizes the Kullback-Leibler(KL) divergence [173], defined as

$$\begin{aligned} &C_{KL}(q(\Omega_b^{ld}, y_b^{ld}) || p(\Omega_b^{ld}, y_b^{ld} | s_b^{ld}, x^{ld})) \\ &= \int_{\Omega_b^{ld}, y_b^{ld}} q(\Omega_b^{ld}, y_b^{ld}) \log\left(\frac{q(\Omega_b^{ld}, y_b^{ld})}{p(\Omega_b^{ld}, y_b^{ld} | s_b^{ld}, x^{ld})}\right) d\Omega_b^{ld} \, dy_b^{ld}, \end{aligned} \tag{6.40}$$

which is always non negative and equal to zero only when $q(\Omega_b^{ld}, y_b^{ld}) = p(\Omega_b^{ld}, y_b^{ld} | s_b^{ld}, x^{ld})$.

We choose to approximate the posterior distribution $p(\Omega_b^{ld}, y_b^{ld} | s_b^{ld}, x^{ld})$ by the distribution $q(\Omega_b^{ld}, y_b^{ld}) = q(\Omega_b^{ld}) q(y_b^{ld})$, and so we can write

$$\begin{aligned} &C_{KL}(q(\Omega_b^{ld}, y_b^{ld}) || p(\Omega_b^{ld}, y_b^{ld} | s_b^{ld}, x^{ld})) \\ &= \int_{\Omega_b^{ld}} q(\Omega_b^{ld}) \left(\int_{y_b^{ld}} q(y_b^{ld}) \log\left(\frac{q(\Omega_b^{ld}) q(y_b^{ld})}{p(\Omega_b^{ld}) p(y_b^{ld}, s_b^{ld}, x^{ld})}\right) dy_b^{ld} \right) d\Omega_b^{ld} + const, \\ &= \int_{y_b^{ld}} q(y_b^{ld}) \left(\int_{\Omega_b^{ld}} q(\Omega_b^{ld}) \log\left(\frac{q(\Omega_b^{ld}) q(y_b^{ld})}{p(\Omega_b^{ld}) p(y_b^{ld}, s_b^{ld}, x^{ld})}\right) d\Omega_b^{ld} \right) dy_b^{ld} + const. \end{aligned} \tag{6.41}$$

Unfortunately, the integral in Eq. (6.41) cannot be directly evaluated due to the TV prior but, as we already did in chapter 5, we can approximate it by using the Majorization-Minimization approach [181] that converts a non-quadratic problem to a quadratic one by the introduction of a new parameter that also needs to be estimated. Thus, the TV prior in Eq. (6.18) is majorized by the functional ,

$$M(\alpha_b^{ld}, y_b^{ld}, u_b^{ld}) = c.(\alpha_b{}^{ld})^{p/2} \exp\left[-\alpha_b^{ld} \sum_{i=1}^{p} \frac{(\Delta_i^h(y_b^{ld}))^2 + (\Delta_i^v(y_b^{ld}))^2 + (u_b^{ld})(i)}{2\sqrt{(u_b^{ld})(i)}} \right], \tag{6.42}$$

where u_b^{ld} is a $p-$dimensional vector, $u_b^{ld} \in (R^+)^p$, with components $(u_b^{ld})(i)$, $i = 1,...,p$, are quantities that need to be computed and have an intuitive interpretation related to the unknown images y_b^{ld}.

Let us consider the inequality, also used in [96], which states that, for any $w \geq 0$ and $z > 0$

$$\sqrt{w} \leq \frac{w+z}{2\sqrt{z}}, \tag{6.43}$$

Now, using the inequality in Eq. (6.43) with $w = (\Delta_i^h(y_b^{ld}))^2 + (\Delta_i^v(y_b^{ld}))^2$ and $z = (u_b)(i)$ and comparing Eq. (6.42) with Eq. (6.18), we obtain $p(y_b^{ld}, \alpha_b^{ld}) \geq c.M(\alpha_b^{ld}, y_b^{ld}, u_b^{ld})$. This leads to the following lower bound for the joint probability distribution

$$\begin{aligned} p(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}, y_b^{ld}, s_b^{ld}, x^{ld}) \geq c. \\ p(\alpha_b^{ld})p(\beta_b^{ld})p(\gamma_b^{ld})M(\alpha_b^{ld}, y_b^{ld}, u_b^{ld})p(s_b^{ld}|\beta_b^{ld}, y_b^{ld})p(x^{ld}|\gamma_b^{ld}, y_b^{ld}) \\ = F(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}, y_b^{ld}, s_b^{ld}, x^{ld}, u_b^{ld}), \end{aligned} \tag{6.44}$$

By defining,

$$\tilde{\mathcal{M}}(q(\Omega_b^{ld}, y_b^{ld}), u_b^{ld}) = \int_{\Omega_b^{ld}, y_b^{ld}} q(\Omega_b^{ld}, y_b^{ld}) \log\left(\frac{q(\Omega_b^{ld}, y_b^{ld})}{F(\Omega_b^{ld}, y_b^{ld}, s_b^{ld}, x^{ld})}\right) d\Omega_b^{ld}\, dy_b^{ld}, \tag{6.45}$$

and using Eq. (6.44), we obtain

$$\mathcal{M}(q(\Omega_b^{ld}, y_b^{ld})) \leq \min_{u_b^{ld}} \tilde{\mathcal{M}}(q(\Omega_b^{ld}, y_b^{ld}), u_b^{ld}). \tag{6.46}$$

Therefore, by finding a sequence of distributions $q^k(\Omega_b^{ld}, y_b^{ld})$ that monotonically decreases $\tilde{\mathcal{M}}(q(\Omega_b^{ld}, y_b^{ld})$ for a fixed u_b^{ld} a sequence of an ever decreasing upper bound of

$$C_{KL}(q(\Omega_b^{ld}, y_b^{ld})||p(\Omega_b^{ld}, y_b^{ld}|s_b^{ld}, x^{ld}))$$

is also obtained due to (6.41). Even more, also minimizing $\tilde{\mathcal{M}}(q(\Omega_b^{ld}, y_b^{ld})$ with respect to u_b^{ld} generates a sequence of vectors $\{(u_b^{ld})^k\}$ that tightens the upper-bound for each distribution $q^k(\Omega_b^{ld}, y_b^{ld})$. Therefore, the two sequences $\{(u_b^{ld})^k\}$ and $q^k(\Omega_b^{ld}, y_b^{ld})$ are coupled. We develop an iterative algorithm (Algorithm 6.3) to find such sequences.

Algorithm 6.3 *Posterior parameter and image distributions estimation in TV reconstruction using* $q(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}, y_b^{ld}) = q(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}) q(y_b^{ld})$

Given $u^1 \in (R^+)^p$ and $q^1(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld})$, an initial estimate of the distribution $q(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld})$, *for* $k = 1, 2, \ldots$, until a stopping criterion is met.

1. *Find*

$$q^k(y_b^{ld}) = \arg\min_{q(y_b^{ld})} \int \int q^k(\Omega_b^{ld}) q(y_b^{ld}) \times \log\left(\frac{q^k(\Omega_b^{ld}) q(y_b^{ld})}{F(\Omega_b^{ld}, y_b^{ld}, s_b^{ld}, x^{ld}, (u_b^{ld})^k)}\right) d\Omega_b^{ld}\, dy_b^{ld}. \tag{6.47}$$

2. *Find*

$$(u_b^{ld})^{k+1} = \arg\min_{u_b^{ld}} \int \int q^k(\Omega_b^{ld}) q^k(y_b^{ld}) \times \log\left(\frac{q^k(\Omega_b^{ld}) q^k(y_b^{ld})}{F(\Omega_b^{ld}, y_b^{ld}, s_b^{ld}, x^{ld}, u_b^{ld})}\right) d\Omega_b^{ld}\, dy_b^{ld}. \tag{6.48}$$

3. *Find*

$$(\Omega_b^{ld})^{k+1} = \arg\min_{\Omega_b^{ld}} \int \int q(\Omega_b^{ld}) q^k(y_b^{ld}) \times \log\left(\frac{q(\Omega_b^{ld}) q^k(y_b^{ld})}{F(\Omega_b^{ld}, y_b^{ld}, s_b^{ld}, x^{ld}, (u_b^{ld})^{k+1})}\right) d\Omega_b^{ld}\, dy_b^{ld}. \tag{6.49}$$

Set $q(\Omega_b^{ld}) = \lim_{k\to\infty} q^k(\Omega_b^{ld}), q(y_b^{ld}) = \lim_{k\to\infty} q^k(y_b^{ld})$.

Inequality (6.43) provides a local quadratic approximation to the TV prior. Had a fixed $(u_b^{ld})^o$ with same elements been used, a global conditional auto-regression model approximating the TV prior would have been obtained. Clearly, the procedure which updates u_b^{ld} will provide a tighter upper bound for $\mathcal{M}(q(\Omega_b^{ld}, y_b^{ld}))$, since we are using $\min_{u_b^{ld}} \tilde{\mathcal{M}}(q(\Omega_b^{ld}, y_b^{ld}), u_b^{ld})$ instead of $\tilde{\mathcal{M}}(q(\Omega_b^{ld}, y_b^{ld}), (u_b^{ld})^0)$. Finally, we note that the process to find the best posterior distribution approximation of the image in combination with u_b^{ld} is a very natural extension of the majorization-minimization approach to function optimization and that local majorization has also been applied to variational logistic regression, as well as, to the inference of its parameters.

Let us now further develop each of the steps of the above algorithm. To calculate $q^k(y_b^{ld})$, we observe that differentiating the integral on the right-hand side of (6.47) with respect to $q(y_b^{ld})$ and setting it equal to zero, we obtain

$$q^k(y_b^{ld}) \propto \exp\left\{E_{q^k(\Omega_b^{ld})}\left[\ln F(\Omega_b^{ld}, y_b^{ld}, s_b^{ld}, x^{ld}, (u_b^{ld})^k)\right]\right\}, \tag{6.50}$$

which represents the Gaussian distribution

$$q^k(y_b^{ld}) = \mathcal{N}(y_b^{ld} | E_{q^k(y_b^{ld})}[y_b^{ld}], cov_{q^k(y_b^{ld})}[y_b^{ld}]), \tag{6.51}$$

with parameters

$$E_{q^k(y_b^{ld})}[y_b^{ld}] = cov_{q^k(y_b^{ld})}[y_b^{ld}]\left[\beta_b^{ld} s_b^{ld} + \gamma^{ld} x^{ld}\right], \tag{6.52}$$

$$cov_{q^k(y_b^{ld})}[y_b^{ld}] = \left[\alpha_b^{ld}\varsigma(u_b^{ld})^k + \beta_b^{ld} I_p + \gamma_b^{ld} I_p\right]^{-1}, \tag{6.53}$$

with

$$\varsigma(u_b^{ld})^k = (\Delta^h)^t W(u_b^{ld})^k(\Delta^h) + (\Delta^v)^t W(u_b^{ld})^k(\Delta^v), \tag{6.54}$$

where $W((u_b^{ld})^k)$ is an $p \times p$ diagonal matrix of the form

$$W((u_b^{ld})^k) = diag\left((u_b^{ld})^k(i)^{-\frac{1}{2}}\right), \tag{6.55}$$

and Δ^h and Δ^v represent the $p \times p$ convolution matrices associated to the first order horizontal and vertical differences, respectively.

To calculate $(u_b^{ld})^{k+1}$, we have from (6.48) that

$$(u_b^{ld})^{k+1} = \arg\min_{u_b^{ld}} \sum_{i=1}^{p} \frac{E_{q^k(y_b^{ld})}\left[(\Delta_i^h(y_b^{ld}))^2 + (\Delta_i^v(y_b^{ld}))^2\right] + (u_b^{ld})(i)}{\sqrt{(u_b^{ld})(i)}}, \tag{6.56}$$

and, consequently

$$(u_b^{ld})^{k+1}(i) = E_{q^k(y_b^{ld})}\left[(\Delta_i^h(y_b^{ld}))^2 + (\Delta_i^v(y_b^{ld}))^2\right], i = 1, \cdots, p \tag{6.57}$$

By differentiating the integral on the right hand side of (6.49) with respect to $q(\Omega_b^{ld})$ and setting it equal to zero, we obtain

$$q^{k+1}(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}) \propto \exp\left\{E_{q^k(y_b^{ld})}\left[\log F(\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}, y_b^{ld}, s_b^{ld}, x^{ld}, (u_b^{ld})^{k+1})\right]\right\} \tag{6.58}$$

and thus

$$q^{k+1}(\alpha_b^{ld},\beta_b^{ld},\gamma_b^{ld}) = q^{k+1}(\alpha_b^{ld})q^{k+1}(\beta_b^{ld})q^{k+1}(\gamma_b^{ld}) \tag{6.59}$$

where these gamma distributions given by

$$q^{k+1}(\alpha_b^{ld}) \propto (\alpha_b^{ld})^{(p/2+a_{\alpha_b^{ld}})-1} \times \exp\left[-\alpha_b^{ld}\left(E_{q^k(y_b^{ld})}\left[TV(y_b^{ld})\right]+c_{\alpha_b^{ld}}\right)\right], \tag{6.60}$$

$$q^{k+1}(\beta_b^{ld}) \propto (\beta_b^{ld})^{(p/2+a_{\beta_b^{ld}})-1} \times \exp\left[-\beta_b^{ld}\left(\frac{E_{q^k(y_b^{ld})}\left\|s_b^{ld}-y_b^{ld}\right\|^2}{2}+c_{\beta_b^{ld}}\right)\right], \tag{6.61}$$

$$q^{k+1}(\gamma_b^{ld}) \propto (\gamma_b^{ld})^{(p/2+a_{\gamma^{ld}})-1} \times \exp\left[-\gamma_b^{ld}\left(\frac{E_{q^k(y_b^{ld})}\left\|x^{ld}-y_b^{ld}\right\|^2}{2}+c_{\gamma^{ld}}\right)\right]. \tag{6.62}$$

From the definition of the gamma distribution in Eq. (6.15), previous distributions have the following means,

$$E[\alpha_b^{ld}]_{q^k(\Omega_b^{ld})} = \frac{a_{\alpha_b^{ld}}+\frac{p}{2}}{c_{\alpha_b^{ld}}+E_{q^k(y_b^{ld})}\left[TV(y_b^{ld})\right]}, \tag{6.63}$$

$$E[\beta_b^{ld}]_{q^k(\Omega_b^{ld})} = \frac{a_{\beta_b^{ld}}+\frac{p}{2}}{c_{\beta_b^{ld}}+\frac{1}{2}E_{q^k(y_b^{ld})}\left[\left\|s_b^{ld}-y_b^{ld}\right\|^2\right]}, \tag{6.64}$$

$$E[\gamma_b^{ld}]_{q^k(\Omega_b^{ld})} = \frac{a_{\gamma_b^{ld}}+\frac{p}{2}}{c_{\gamma_b^{ld}}+\frac{1}{2}E_{q^k(y_b^{ld})}\left[\left\|x^{ld}-y_b^{ld}\right\|^2\right]}, \tag{6.65}$$

Having into account that the mean of the prior distribution on the parameters are $\overline{\alpha_b^{ld}} = a_{\alpha_b^{ld}}/c_{\alpha_b^{ld}}, \overline{\beta_b^{ld}} = a_{\beta_b^{ld}}/c_{\beta_b^{ld}}, \overline{\gamma_b^{ld}} = a_{\gamma_b^{ld}}/c_{\gamma_b^{ld}}$ and

$$\lambda_{\alpha_b^{ld}} = \frac{a_{\alpha_b^{ld}}}{a_{\alpha_b^{ld}}+p/2}, \lambda_{\beta_b^{ld}} = \frac{a_{\beta_b^{ld}}}{a_{\beta_b^{ld}}+p/2}, \lambda_{\gamma_b^{ld}} = \frac{a_{\gamma_b^{ld}}}{a_{\gamma_b^{ld}}+p/2}, \tag{6.66}$$

after simple calculations, we can rewrite the above means as follows,

$$\begin{aligned}(E[\alpha_b^{ld}]_{q^k(\Omega_b^{ld})})^{-1} &= \frac{\lambda_{\alpha_b^{ld}}}{\overline{\alpha_b^{ld}}} + (1-\lambda_{\alpha_b^{ld}})\frac{\sum_{i=1}^{p}\sqrt{(u_b^{ld})_i^{k+1}}}{p/2} \\ &= \frac{\lambda_{\alpha_b^{ld}}}{\overline{\alpha_b^{ld}}} + (1-\lambda_{\alpha_b^{ld}}) \\ &\times 2\sum_{i=1}^{p}\frac{\sqrt{\left(\Delta_i^h\left(E_{q^k(y)}[y_b^{ld}]\right)\right)^2 + \left(\Delta_i^v\left(E_{q^k(y)}[y_b^{ld}]\right)\right)^2 + \frac{1}{p}trace\left[(COV_b^{ld})^{-1}\times((\Delta^h)^t(\Delta^h)+(\Delta^v)^t(\Delta^v))\right]}}{p},\end{aligned} \tag{6.67}$$

$$\begin{aligned}(E[\beta_b^{ld}]_{q^k(\Omega_b^{ld})})^{-1} &= \frac{\lambda_{\beta_b^{ld}}}{\overline{\beta_b^{ld}}} + (1-\lambda_{\beta_b^{ld}})\frac{E_{q^k(y)}\left[\|s_b^{ld}-y_b^{ld}\|^2\right]}{p} \\ &= \frac{\lambda_{\beta_b^{ld}}}{\overline{\beta_b^{ld}}} + (1-\lambda_{\beta_b^{ld}})\frac{\|s_b^{ld}-E_{q^k(y)}[y_b^{ld}]\|^2 + trace\left((COV_b^{ld})^{-1}\right)}{p},\end{aligned} \tag{6.68}$$

$$\begin{aligned}(E[\gamma_b^{ld}]_{q^k(\Omega_b^{ld})})^{-1} &= \frac{\lambda_{\gamma_b^{ld}}}{\overline{\gamma_b^{ld}}} + (1-\lambda_{\gamma_b^{ld}})\frac{E_{q^k(y)}\left[\|x^{ld}-y_b^{ld}\|^2\right]}{p} \\ &= \frac{\lambda_{\gamma_b^{ld}}}{\overline{\gamma_b^{ld}}} + (1-\lambda_{\gamma_b^{ld}})\frac{\|x^{ld}-E_{q^k(y)}[y_b^{ld}]\|^2 + trace\left((COV_b^{ld})^{-1}\right)}{p},\end{aligned} \tag{6.69}$$

where $(COV_b^{ld})^{-1}$ is the approximation of the covariance of $q(x)$, following the same procedure we followed in chapter 5, and is equal to

$$\begin{aligned}\left(COV_b^{ld}\right)^{-1} = {} & E_{q^k(\alpha)}[\alpha_b^{ld}]z((u_b^{ld})^k)(\Delta^h)^t(\Delta^h) + E_{q^k(\alpha)}[\alpha_b^{ld}]z((u_b^{ld})^k)(\Delta^v)^t(\Delta^v) \\ & + E_{q^k(\beta)}[\beta_b^{ld}]I_p + E_{q^k(\gamma)}[\gamma_b^{ld}]I_p,\end{aligned} \tag{6.70}$$

where $z(u_b^k)$, as we mentioned before, is calculated as the mean value of the diagonal values in $W((u_b^{ld})^k)$, that is

$$z((u_b^{ld})^k) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{\sqrt{(u_b^{ld})^k(i)}}. \tag{6.71}$$

These means are then used to recalculate the distributions of y_b^{ld} in Algorithm 6.3. As we discussed before in chapter 5, equation (6.66) indicates that $\lambda_{\alpha_b^{ld}}$, $\lambda_{\beta_b^{ld}}$ and $\lambda_{\gamma_b^{ld}}$ are taking values in the interval $[0,1)$, can be understood as normalized confidence parameters. As can be seen from Eqs. (6.67),(6.68) and (6.69) the inverse of the means of the hyper-priors are represented as convex combinations of their initial values and their maximum likelihood (ML) estimates.

These ML estimates have been derived before either empirically or by using regularization formulations [182,183]. As we already mentioned in chapter 5, when the values of λ_α, λ_β and λ_γ are equal to zero, according to Eqs. (6.67),(6.68) and (6.69), no confidence is placed on the given values of the hyper-parameters and ML estimates are used, while when they are asymptotically equal to one, the prior knowledge of the mean is fully enforced (i.e., no estimation of the hyper-parameters is performed).

A summary of the steps of the proposed pansharpening method is shown in Algorithm 6.4.

Algorithm 6.4 *Proposed pansharpening method steps*

1. Upsample each band of the MS image, Y_b, to the size of the PAN, x, and register them obtaining s_b, $b = 1,\dots,B$.
2. Apply NSCT decomposition on the PAN image x and registered MS image $\{s_b\}$,

$$x = x^r + \sum_{l=1}^{L}\sum_{d=1}^{D} x^{ld}, \tag{6.72}$$

$$s_b = s_b^r + \sum_{l=1}^{L}\sum_{d=1}^{D} s_b^{ld}, \quad b = 1,\dots,B, \tag{6.73}$$

where we are using the superscript r to denote the residual (low pass filtered version) NSCT coefficients band and the superscript ld to refer to the detail bands, with $l = 1,\dots,L$, representing the scale and $d = 1,\dots,D$, representing the direction for each coefficient band.

3. Select the $\hat{y}_b^r$ band as the output of the Algorithm 6.2
4. Select $\hat{y}_b^{ld}$ bands as the mean of the posterior distribution $q(y_b^{ld})$ calculated according to Algorithm 6.3.
5. Apply the inverse NSCT to merge the MS band coefficients $\{\hat{y}_b^r\}$, $\{\hat{y}_b^{ld}\}$, getting $\{\hat{y}_b\}$,

$$\hat{y}_b = \hat{y}_b^r + \sum_{l=1}^{L}\sum_{d=1}^{D} \hat{y}_b^{ld}, b = 1,\dots,B. \tag{6.74}$$

6.5 Experimental results

In this section, the proposed NSCT-based pansharpening method using the Bayesian inference is tested. Experiments on a synthetic color image and a real SPOT5 and QuickBird images are conducted to test the proposed method.

λ_α	λ_β	λ_γ	**ERGAS**	**COR**	**SSIM**
0.0	0.0	0.0	2.5380	0.9812	0.9606
0.0	0.0	0.4	4.0420	0.7068	0.8890
0.0	0.0	1.0	4.4107	0.6260	0.8654
0.0	0.4	0.0	2.6618	0.9925	0.9559
0.0	0.4	0.4	3.6324	0.8247	0.9099
0.0	0.4	1.0	4.3728	0.6529	0.8679
0.0	1.0	1.0	4.5835	0.7278	0.8556
0.4	0.0	0.0	2.5522	0.9798	0.9602
0.4	0.4	0.0	2.7097	0.9915	0.9560
0.4	0.4	0.4	3.2359	0.9158	0.9352
1.0	0.4	0.4	4.8240	0.6954	0.8260
1.0	1.0	0.4	5.1565	0.7216	0.8050
1.0	1.0	1.0	5.3380	0.6295	0.7908

Table 6.1: ERGAS, Mean COR and Mean SSIM for synthetic HRMS image with selected values of λ_α, λ_β and λ_γ .

The observations of the synthetic multispectral are obtained from the color image, displayed in Figure 6.13(a), by convolving it with mask $0.25 \times 1_{2\times2}$ to simulate sensor integration, and then downsampling it by a factor of two by discarding every other pixel in each direction and adding zero mean Gaussian noise with variance 16. For the PAN image we used the luminance of the original color image and zero mean Gaussian noise of variance 9 was added. These images are the inputs of Algorith 6.4. In this algorithm, bicubic interpolation is used to upsample the observed MS image to the size of the PAN image in step 1 of the algorithm and then three levels of NSCT decomposition was applied on each upsampled MS band and PAN image with 4 directions for the first level and 8 directions for the second and the third decomposition levels, in step 2. Algorithms 6.2 and 6.3 were run on the resulting coefficients bands until the criterion $\|(y_b^{ld})^k - (y_b^{ld})^{k-1}\|^2/\|(y_b^{ld})^{k-1}\|^2 < 10^{-4}$ was satisfied, where $(y_b^{ld})^k$ denotes the mean of $q^k(y_b^{ld})$ at iteration k or the residual band estimation at iteration k, $(\hat{y}_b^r)^k$. The algorithm typically converges within 4 iterations. The values of parameters were automatically estimated by the proposed method. For the residual band, the initial parameters in Algorithm 6.2 were estimated

Measure	**Band**	**NSCT [62]**	**SR [96]**	**Proposed** ($\lambda = 0$)
COR	b1	0.91	0.84	**0.99**
	b2	0.91	0.98	**0.99**
	b3	0.90	0.62	**0.98**
SSIM	b1	0.79	0.90	**0.97**
	b2	0.81	0.94	**0.97**
	b3	0.81	0.85	**0.96**
PSNR	b1	26.75	32.68	**38.17**
	b2	27.17	35.50	**39.51**
	b3	27.65	30.15	**36.93**
ERGAS	-	5.76	3.12	**1.61**

Table 6.2: Synthetic Image Quantative Results

based on the observed MS image as $(\alpha_b^r)^1 = (p-1)/(2\|Q(\hat{y}_b^r)_k\|^2)$ and $(\beta_b^r) = p/(2\|s_b^r - (\hat{y}_b^r)_k\|^2)$. The initial distribution on the parameters, $q^1(\Omega_b^{ld})$, in Algorithm 6.3 was estimated based on the observed MS and PAN images coefficients as $\alpha_b^{ld} = p/(2TV(y_b^{ld}))$, $\beta_b^{ld} = p/\|s_b^{ld} - y_b^{ld}\|^2$ and $\gamma_b^{ld} = p\|x^{ld} - y_b^{ld}\|^2$ for all $b = 1,\dots,B$, assuming that $q^0(y)$ is a degenerate distribution on the bicubic interpolation of the observed MS image, that is, we used the observation to initialize the hyperparameter distribution. For this first experiment we chose $\lambda_\alpha = \lambda_\beta = \lambda_\gamma = 0$ for all the bands and directions, i.e., no prior information on the value of the hyperparameters is provided. The initial value of u is calculated from this observed image. Note that the algorithms are initialized automatically without any manual input. We compared the proposed NSCT using Bayesian inference method with the SR method in [96] and the additive NSCT method [62]. The resulted images corresponding to each method are displayed in Figure 6.13(d)-(h).

To assess the spatial improvement of the pansharpened images we use the correlation of the high frequency components (COR) [25] which takes values between zero and one (the higher the value the better the quality of the pansharpened image). Spectral fidelity was assessed by means of the peak signal-to-noise ratio (PSNR), the Structural Similarity Index Measure (SSIM) [131], an index ranging from -1 to $+1$ with $+1$ corresponding to exactly equal images, and the *erreur relative globale adimensionnelle de synthése* (ERGAS) [89] index, a global criterion for what the lower the value, specially a value lower than the number of bands in the image, the higher the quality of the pansharpened image. Table 6.2 shows the corresponding quantitative results. The proposed method provides better results for each measure. The COR values reflect that all methods are able to incorporate the details of the PAN image into the pansharpened one, although the SR method in [96], see Figure 6.13(e), introduced less details in the band 3 (blue) since it contributes less to the PAN image and more into the band 2 (green) since it has the highest contribution, which is reflected as a greenish color near the edges of the image. The NSCT method in [62] incorporates details in all the bands but produces a noisy image, see Figure 6.13(d). The proposed method (Figure 6.13(f)) is able to incorporate details in all the bands while controls the noise and avoids the color bleeding effect. The spectral fidelity measures show that the proposed method performs better than the competing methods, which is also clear from the image in Figure 6.13(f). The PSNR for the proposed method is about $10dB$ higher than NSCT method in [62] and from $2dB$ to almost $6dB$ higher than for the SR method in [96], with a remarkable high SSIM and low ERGAS values which reflect the high quality of the resulting images.

We next examine the effect of the introduction of additional information about the unknown hyperparameters through the use of the confidence parameters $\lambda_{\alpha_b^{ld}}$, $\lambda_{\beta_b^{ld}}$ and $\lambda_{\gamma_b^{ld}}$ on the performance of the algorithm. As we have already explained before, in the case of $\lambda_{\alpha_b^{ld}} = \lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0$, no information about the hyperparameters is available, and the observed image is responsible for the estimation of the hyperparameters and the image. In our experiments, we run the algorithm varying the confidence parameters $\lambda_{\alpha_b^{ld}}$, $\lambda_{\beta_b^{ld}}$ and $\lambda_{\gamma_b^{ld}}$ from 0 to 1 in 0.1 steps, using the initial parameters values for $\overline{\alpha}$, $\overline{\beta}$ and $\overline{\gamma}$. Table 6.1 shows the ERGAS, the mean of COR and SSIM for all the bands, using the proposed algorithm for selected values of the confidence parameters. The confidence values are selected to demonstrate the behavior of the proposed algorithm in the following cases:

1. when full information about the HRMS image and noise variance is available,
2. when no information is provided, i.e., the observation is fully responsible for the reconstruction,
3. when some information about the image prior variance α is provided,
4. when some information about the noise variance is provided.

Moreover, the evolution of ERGAS, COR and SSIM for the full set of confidence parameters are depicted in the Figures 6.7-6.11. From these figures we can extract the ideal values of $\lambda_{\alpha_b^{ld}}$, $\lambda_{\beta_b^{ld}}$ and $\lambda_{\gamma_b^{ld}}$ that can help the algorithm to find the best possible reconstruction of the HRMS image. Regarding to these figures, it is clear that best results are obtained when $\lambda_{\alpha_b^{ld}} = \lambda_{\beta_b^{ld}} = \lambda_{\gamma} = 0$, and this may occurred due to the use of the information from the observed image as a prior knowledgement, which is not good for the reconstruction since it comes from the bicubic interpolation. We can notice also from these figures that high values of $\lambda_{\alpha_b^{ld}}$ decrease the spatial and spectral quality of the reconstructed images.

We already proved that although the levels and directions are independent they are related so we have studied the relationships between them. More concretely, we are going study the relationship of the different parameters. For that, we generated images following a Gaussian distribution with zero mean and different values for the variance, applied the contourlet decomposition to the images and calculated the variance of the noise at each level and direction. Studying those values we found that, for a given level, the calculated variances are very similar. Also, we found that there is a ratio between different levels. The relationship between the inverse of the variance of the coefficients for three level of decompositions, the first with 4 directions and the second and third with 8 directions, is shown in Table 6.4. Also we found experimentally that the sum of the noise variance for all the coefficients at different levels and directions is almost equal to the variance of the noise before contourlet decomposition. In the same way we tested the relationship between α_b^{ld} at different decomposition levels, by estimating the TV parameter for the coefficients, but we could not find a clear relationship.

Having studied the relationship between the variance of the coefficients of the different levels we used this relations as prior knowledgement to guide the estimation of the noise parameters at each level. More concretely, we set $\lambda_{\alpha_b^{ld}}, \lambda_{\beta_b^{ld}}$ and $\lambda_{\gamma_b^{ld}}$ to zero at the coarser level, relying only on the data to estimate the parameters, for the finer decomposition levels, we set the confidence parameters $\lambda_{\alpha_b^{ld}} = 0$ and vary $\lambda_{\beta_b^{ld}}$ and $\lambda_{\gamma_b^{ld}}$ from 0 to 1 in 0.1 intervals,with the values for $\overline{\beta_b^{ld}}$ and $\overline{\gamma_b^{ld}}$ calculated as the value of the parameter estimated at the coarser level multiplied by the corresponding factor according to Table 6.4. Figure 6.12 shows the evolution of ERGAS, mean COR and mean SSIM for the synthetic HRMS image with different values of $\lambda_{\beta_b^{ld}}$ and $\lambda_{\gamma_b^{ld}}$. In all the cases, we set $\lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0$ for the coarser decomposition level and $\lambda_{\alpha^{ld}} = 0$ for all decomposition levels. Table 6.3 shows the values of ERGAS, the mean COR and mean SSIM, using the proposed relationship between the coefficients in the algorithm for selected values of the confidence parameters. Comparing the results in Table 6.3 with the ones

λ_β	λ_γ	**ERGAS**	**COR**	**SSIM**
0.0	0.0	2.5380	0.9812	0.9606
0.0	0.5	2.4445	0.9944	0.9605
0.5	0.0	2.6893	0.9675	0.9551
0.0	1.0	2.4621	0.9951	0.9581
0.5	1.0	**2.4433**	**0.9953**	**0.9585**
0.5	0.0	2.7612	0.9601	0.9528
0.0	0.9	2.4622	0.9951	0.9580
0.5	0.9	2.4436	0.9954	0.9583
0.9	0.9	**2.4425**	**0.9952**	**0.9587**
1.0	1.0	2.4405	0.9950	0.9589

Table 6.3: ERGAS, Mean COR and Mean SSIM for synthetic HRMS image with selected values of $\lambda_{\beta_b^{ld}}$ and $\lambda_{\gamma_b^{ld}}$, while $\lambda_{\alpha_b^{ld}} = 0$.

in Table6.1, we can found that providing the algorithm with prior knowledge related to the relationship between the coefficients, in other words high values of $\lambda_{\beta^{ld}}$ and $\lambda_{\gamma^{ld}}$, improved the resulted images and reduced the calculation time. The highlighted values in Table 6.3 show the best ERGAS, mean COR and mean SSIM values of the resulted images. We can see that this correspond with $\lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0.9$. Note also that these values are not crucial and that setting then to, for instance $\lambda_{\beta_b^{ld}} = 0.6$ and $\lambda_{\gamma_b^{ld}} = 1$ will obtain also very good results. Although the numerical values are better than those using $\lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0.0$ the resulting images, depicted in Figure 6.13(f)-(h) are almost indistinguishably. However the computation time is greatly reduced by including this prior knowledgement.

In the last experiment with this synthetic image we tested the change of $\lambda_{\alpha_b^r}$ and $\lambda_{\beta_b^r}$ in the proposed algorithm, and we found that we get the best results when $\lambda_{\alpha_b^r} = 1$ and $\lambda_{\alpha_b^r} = 0$, in other words not modifying the residual image. However, in the case of very noisy images, small values of $\lambda_{\alpha_b^r}$ can improve the constructed HRMS image.

In the next experiment, the method was tested on real SPOT5 dataset, where the MS image covers a region of interest of 80 by 80 pixels with pixel resolution of 10 m, while the PAN image is 160 by 160 pixels with a pixel resolution of 5 m. The MS image consists of four bands from the visible and infrared region corresponding to green (0.50-0.59 μm), red (0.61-0.68

	Second Level	**Third Level**
First Level	1:2	1:8
Second Level	-	1:4

Table 6.4: The relationship between contourlet coefficients for three level of decompositions, the first with 4 directors, and second and third with 8 directions.

μm), Near IR (0.78-0.89 μm), Mid IR(1.58-1.75 μm), while the PAN image consists of a single band covering the visible and NIR (0.48-0.71 μm). Figure 6.14(a) shows a region of the RGB color image representing bands 1 to 3 of the MS image. Its corresponding PAN and bicubic interpolation image are depicted in Figure 6.14(b) and (c), respectively. The resulting image after running the proposed method when no prior information about the parameters value is introduced, $\lambda_{\alpha_b^{ld}} = \lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0$, is depicted in Figure 6.14(h), this figure also shows the resulting images for bicubic interpolation, CiSper in [29], NSCT method in [62] and SR method in [96]. Following the results, for the synthetic image, we also run the method with introducing prior knowledgement on the value of the parameters. We used $\lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0$ for the coarser decomposition level, $\lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0.9$ and $\lambda_{\beta_b^{ld}} = 0.6$, $\lambda_{\gamma_b^{ld}} = 1.0$, following the best results in Table 6.1, for the finer decomposition levels, and $\lambda_{\alpha^{ld}} = 0$ for all decomposition levels, with the values for $\overline{\beta}_b^{ld}$ and $\overline{\gamma}_b^{ld}$ calculated as the value of the parameter estimated at the coarser level multiplied by the corresponding factor according to Table 6.4. The resulting images is depicted in Figure 6.14(i)-(j). Table 6.5 shows the quantitative results corresponding to these images. The highlighted value in the table presents the highest value for each measure.

The resulted images, displayed in Figures 6.14(c)-(j), reveal similar conclusions to the obtained for the synthetic image, from the visual inspections. The NSCT method in [62] (Figure 6.14(d)) provides a detailed image but quite noisy. Figure 6.14(e) show the result of CiSper method in [29], presents the best spectral values but lowest COR values, that is, it preserves spectral properties but does not incorporate much of the details of the PAN image. This is clear in the image which is not as sharp as the one obtained with the proposed method. Also, from PSNR values, it is clear that for band 3 and 4 in SPOT 5, these bands are not changed by the method. The SR method in [96] provides good details for bands 1 and 2, see Figure 6.14(g), but not for bands 3 and 4 since the PAN image does not cover those bands. This is why the blue color in Figure 6.14(g), seems to be vanished.

Measure	**Band**	**NSCT** [62]	**CiSper** [29]	**Method** in Ch. 5	**SR** [96]	**Proposed** $\lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0$	**Proposed** $\lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0.9$
COR	b1	0.91	0.92	**0.99**	0.87	0.97	**0.99**
	b2	0.95 0.95	**0.99**	0.95	0.98	**0.99**	**0.99**
	b3	0.95	0.95	**0.99**	0.96	0.97	**0.99**
	b4	0.94	0.96	**0.98**	0.78	0.89	0.97
SSIM	b1	**0.79**	**0.84**	0.59	0.74	0.62	0.62
	b2	0.75	0.86	**0.87**	0.78	0.86	0.86
	b3	0.76	0.84	**0.85**	0.78	**0.85**	**0.85**
	b4	0.67	0.81	0.64	**0.89**	0.70	0.70
PSNR	b1	29.40	**29.77**	28.23	27.23	24.90	25.05
	b2	24.76	26.20	**29.07**	26.36	28.29	28.63
	b3	26.47	∞	27.98	27.10	**29.03**	**29.03**
	b4	24.77	∞	25.35	**27.00**	26.21	25.66
ERGAS	-	6.36	**3.03**	6.12	6.02	5.78	5.65

Table 6.5: SPOT 5 Image Quantative Results

The proposed method in Figures 6.14(h)-(j), provides the best result, preserving the spectral properties of MS image while incorporating the high frequencies from the panchromatic image and controlling the noise in the image. Table 6.5 shows the corresponding quantitative results. The highlighted value in the table presents the highest value for each measure.

Finally, the method was tested on real QuickBird dataset, where the MS image covers a region of interest of 60 by 60 pixels with pixel resolution of 2.44 m, while the PAN image is 240 by 240 pixels with a pixel resolution of 61 cm to 72 cm. Figure 6.15(a) shows a region of the RGB color image representing bands 1 to 3 of the MS image. Its corresponding PAN and bicubic interpolation image are depicted in Figure 6.15(b) and (c), respectively. The resulted images, displayed in Figures 6.15(c)-(h), reveal similar conclusions to the obtained for the synthetic and SPOT 5 images, from the visual inspections. Table 6.6 shows the corresponding quantitative results. The highlighted value in the table presents the highest value for each measure.

Measure	**Band**	**NSCT** [62]	**CiSper** [29]	**Method** in Ch. 5	**SR** [96]	**Proposed** $\lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0$
COR	b1	0.96	0.97	**0.99**	0.39	0.95
	b2	0.97	0.92	**0.99**	0.87	0.87
	b3	0.97	0.96	**0.99**	0.86	0.94
	b4	**0.96**	0.79	0.95	0.88	0.76
SSIM	b1	0.80	0.75	0.51	**0.96**	0.82
	b2	0.61	0.78	0.52	0.58	**0.79**
	b3	0.58	**0.79**	0.52	0.62	**0.79**
	b4	0.39	**0.81**	0.55	0.44	0.77
PSNR	b1	34.68	30.28	27.77	31.44	**35.15**
	b2	29.21	30.12	27.08	20.75	**33.40**
	b3	28.37	30.44	26.85	21.97	**33.04**
	b4	23.43	29.34	25.89	18.51	**31.94**
ERGAS	-	2.80	**2.65**	5.93	14.76	2.67

Table 6.6: QuickBird Image Quantative Results

6.6 Conclusions

In this chapter we propose a new pansharpening method that generalizes the fusion strategy of the panchromatic and multispectral images in contourlet based methods. The proposed fusion algorithm is based on the Bayesian modeling and incorporates a solid way to incorporate the details in the panchromatic into the multispectral image while controlling the noise. Particular cases of the proposed fusion method are substitution, additive and weighted contourlets methods.

The proposed pansharpening method has been compared with other methods both in synthetic and real images and its performance has been assessed both numerically and visually.

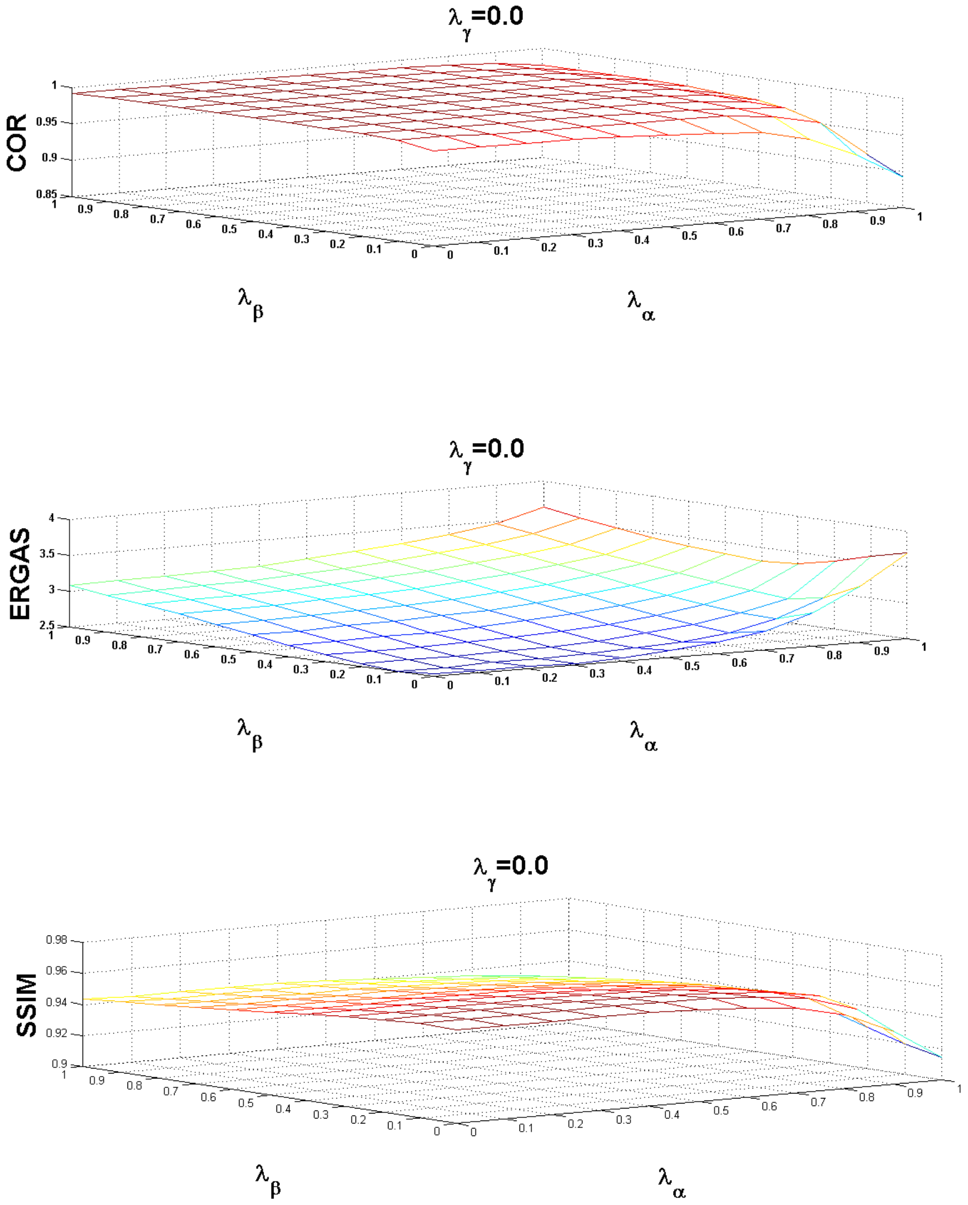


Figure 6.7: *Mean COR, ERGAS and mean SSIM evolution for the synthetic HRMS image with* $\lambda_{\gamma_b^{ld}} = 0.0$ *for all levels* l *and directions* d.

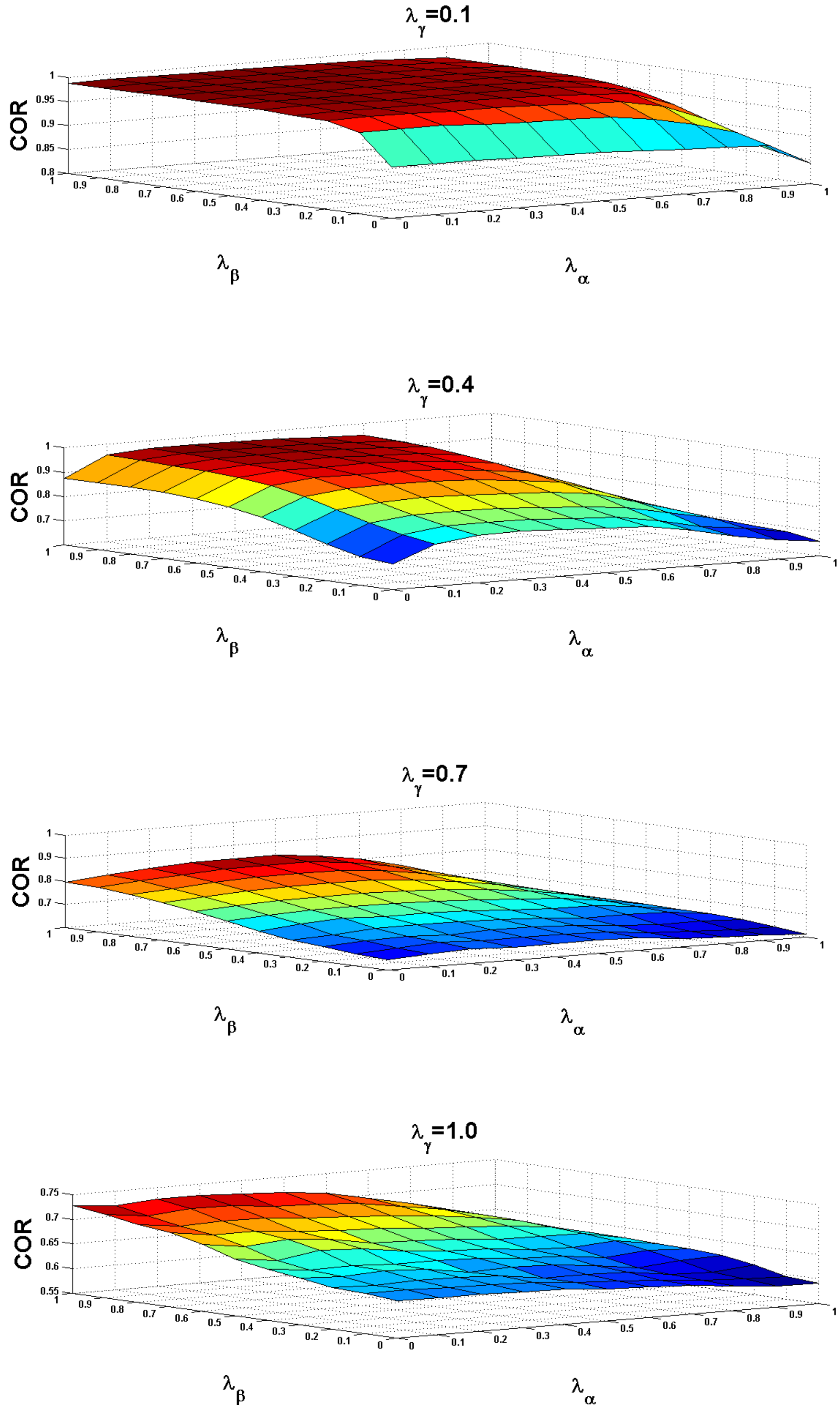
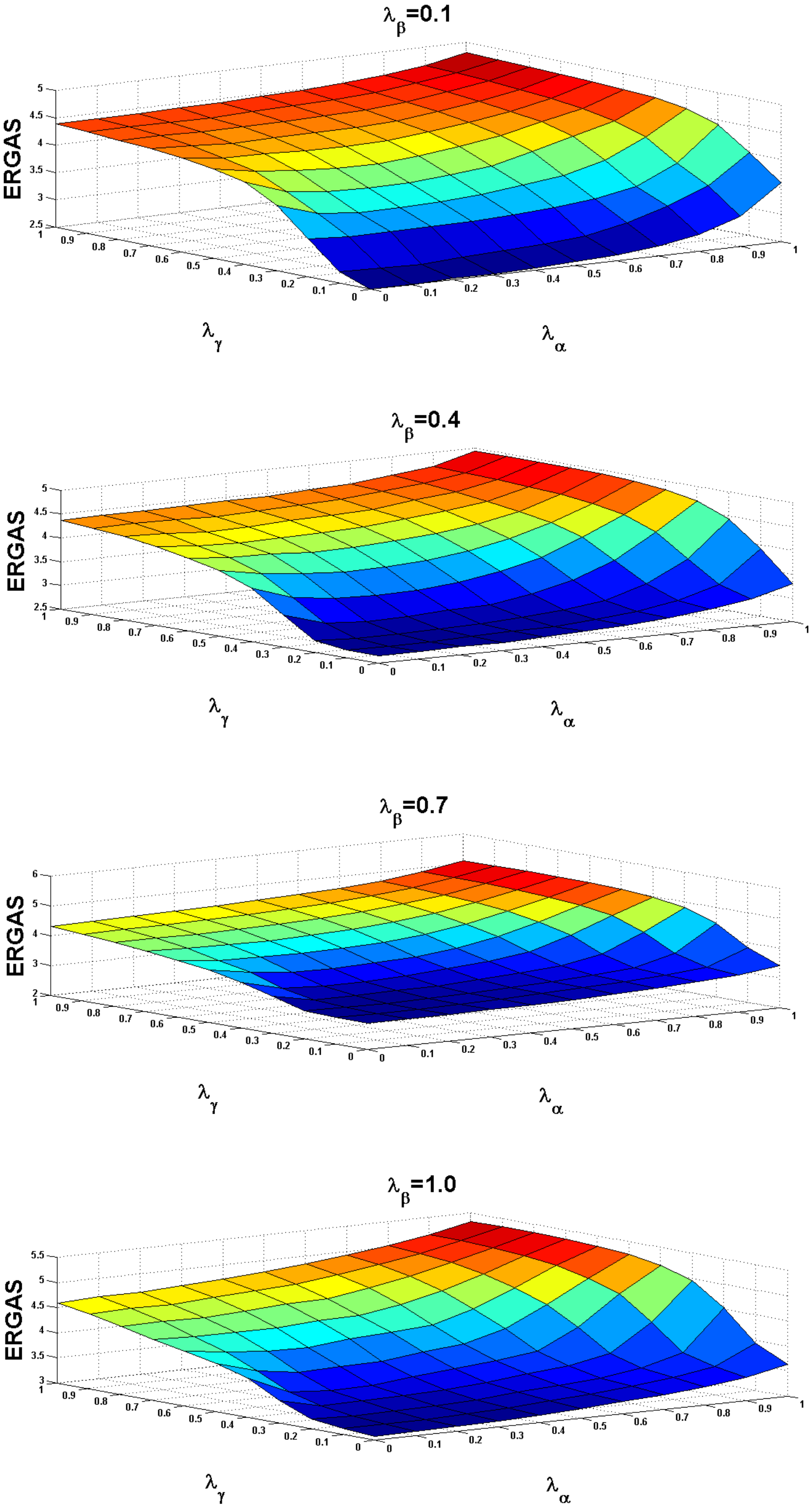


Figure 6.8: *Mean COR evolution for the synthetic HRMS image with different values of* $\lambda_{\gamma_b^{ld}}$ *for all levels l and directions d.*



λ_β=0.1

λ_β=0.4

λ_β=0.7

λ_β=1.0

Figure 6.9: *ERGAS evolution for the synthetic HRMS image with different values of* $\lambda_{\beta_b^{ld}}$ *for all*

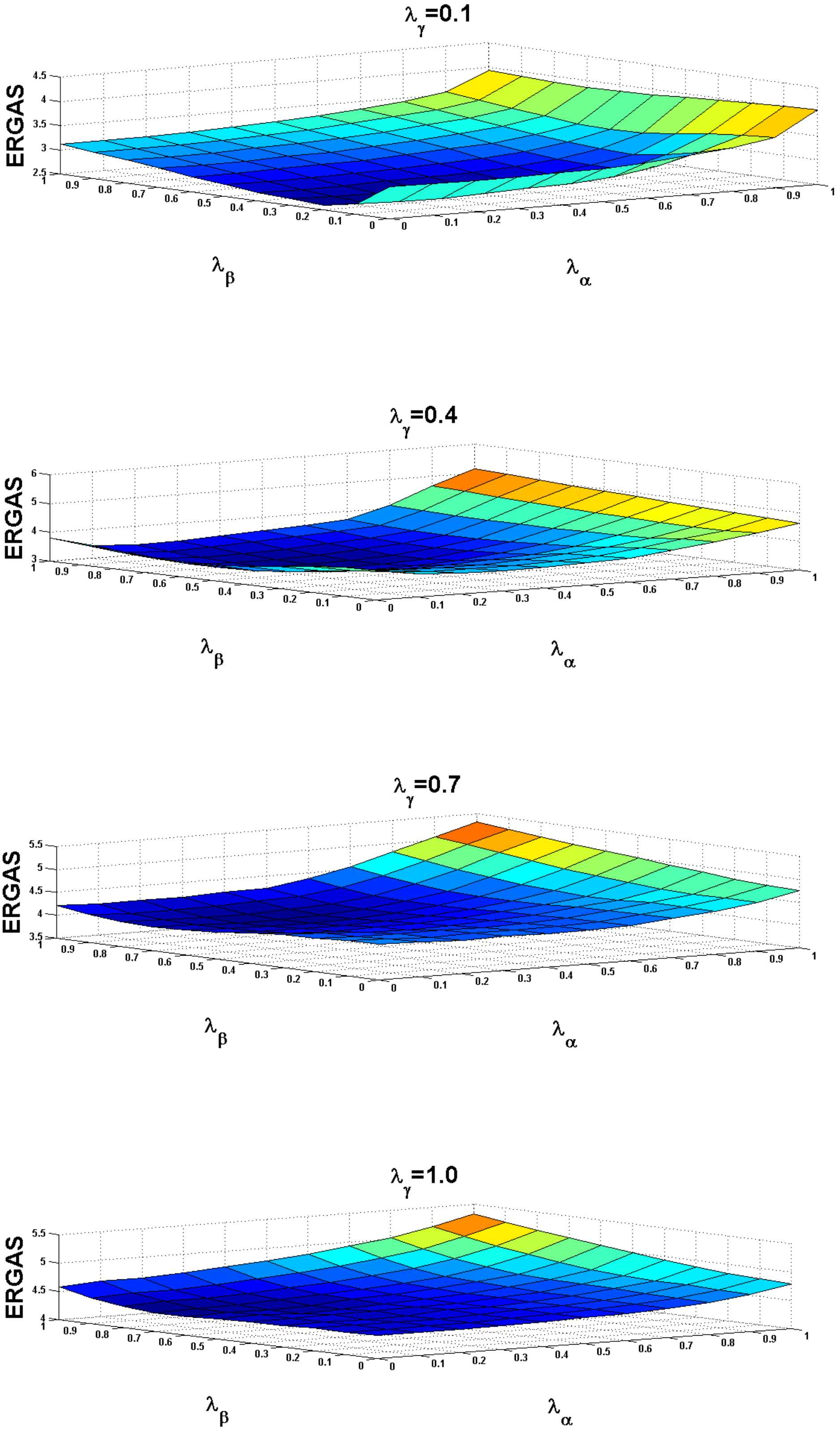


Figure 6.10: *ERGAS evolution for the synthetic HRMS image with different values of* $\lambda_{\gamma_b^{ld}}$ *for all levels* l *and directions* d*.*

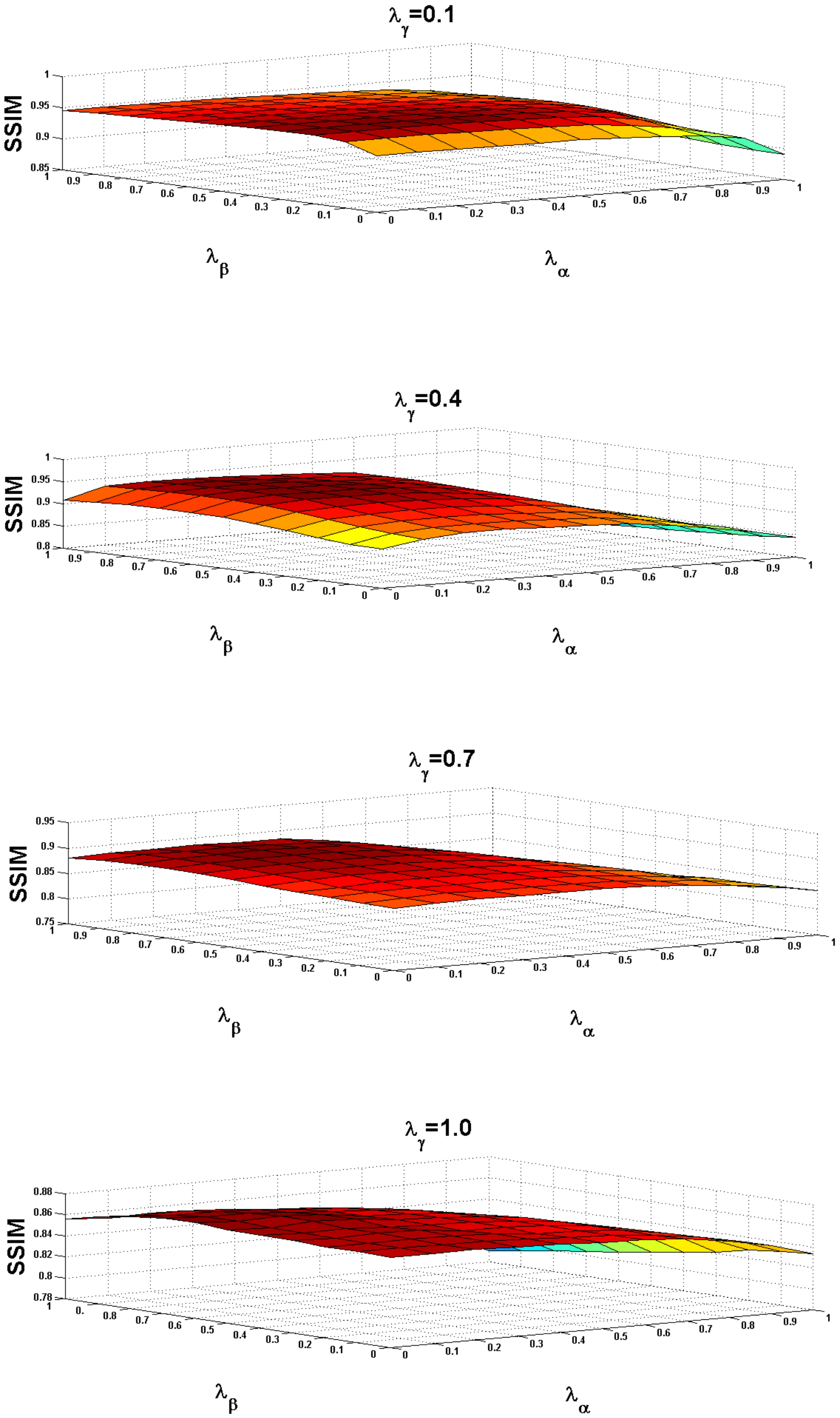


Figure 6.11: *Mean SSIM evolution for the synthetic HRMS image with different values of* $\lambda_{\gamma_b^{ld}}$ *for all levels* l *and directions* d.

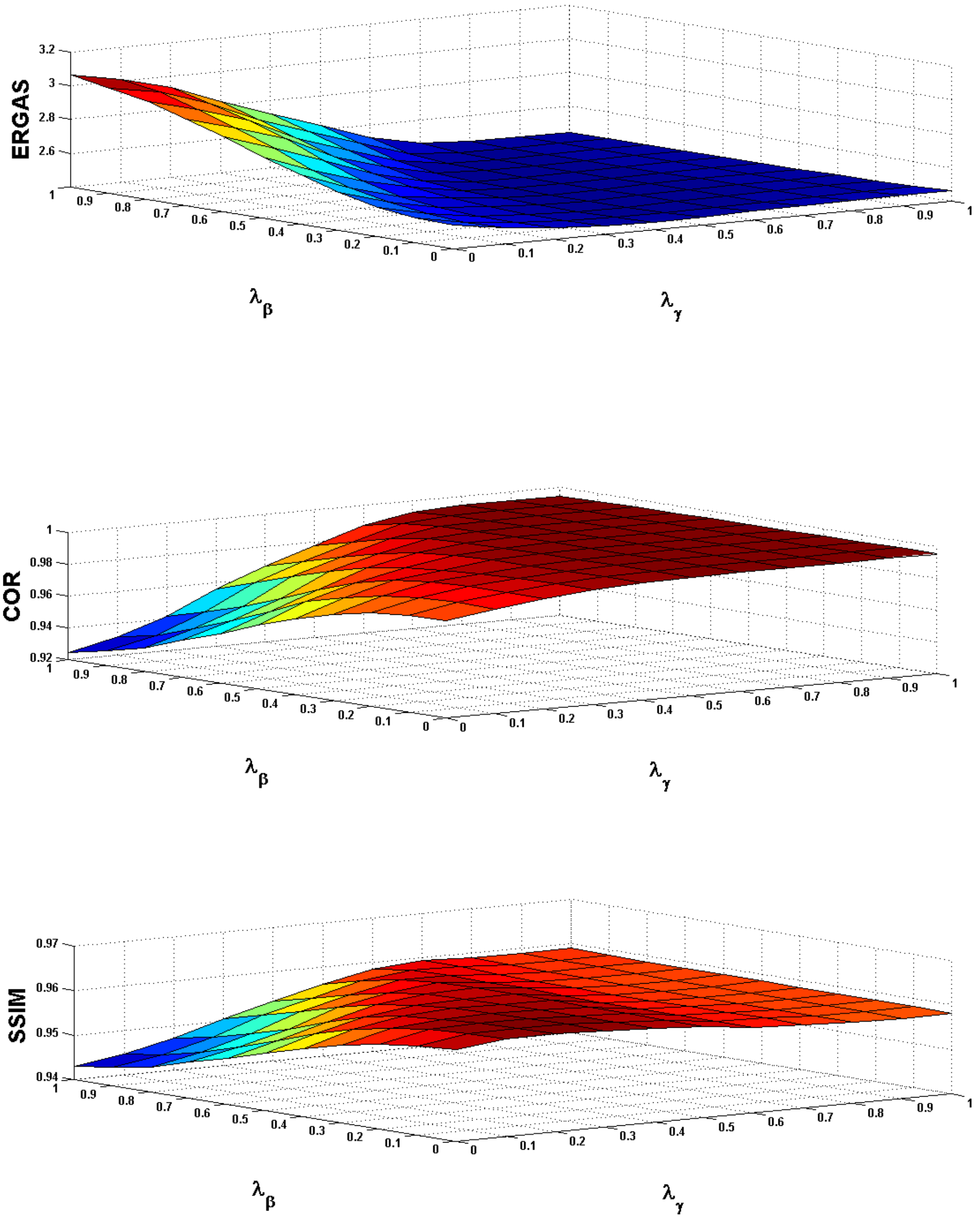
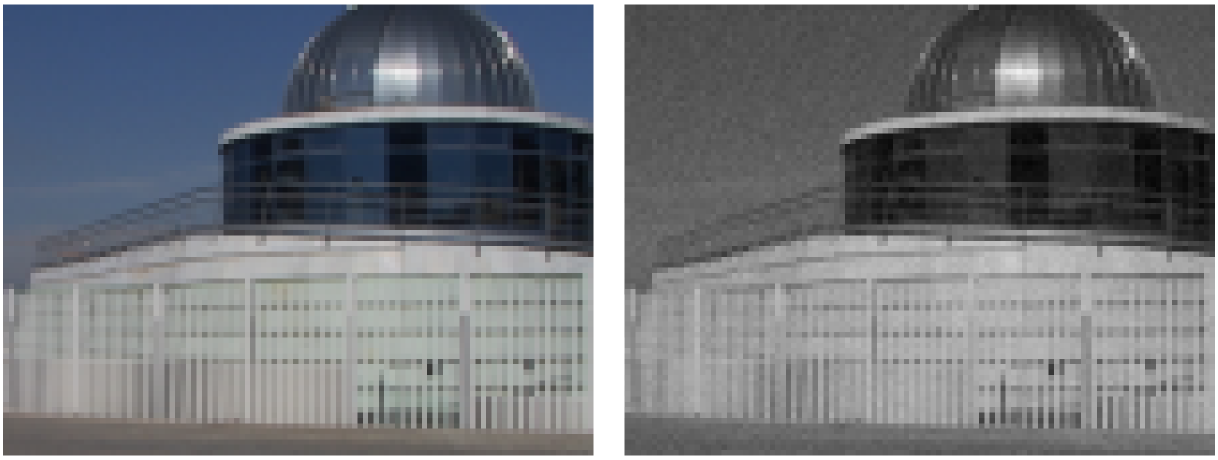
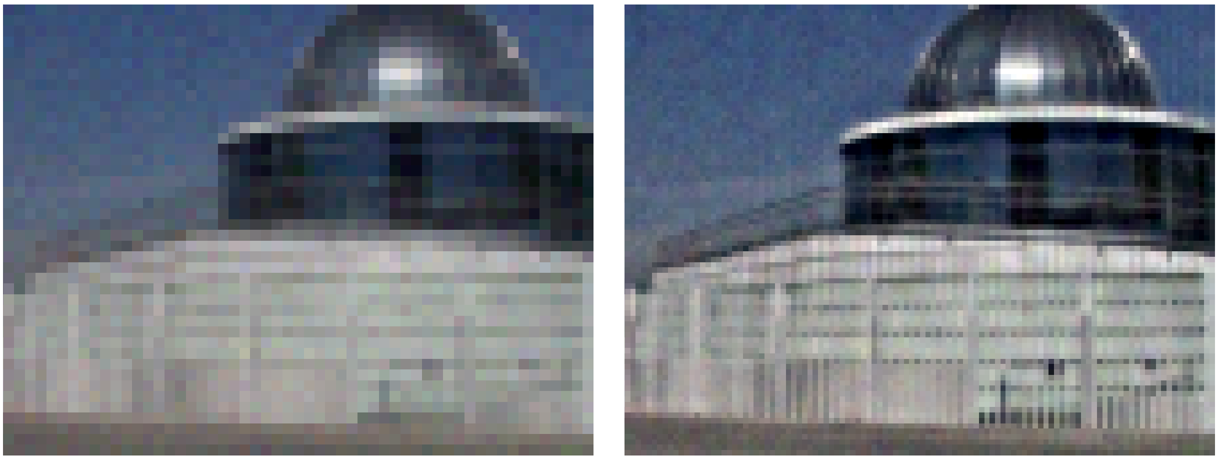


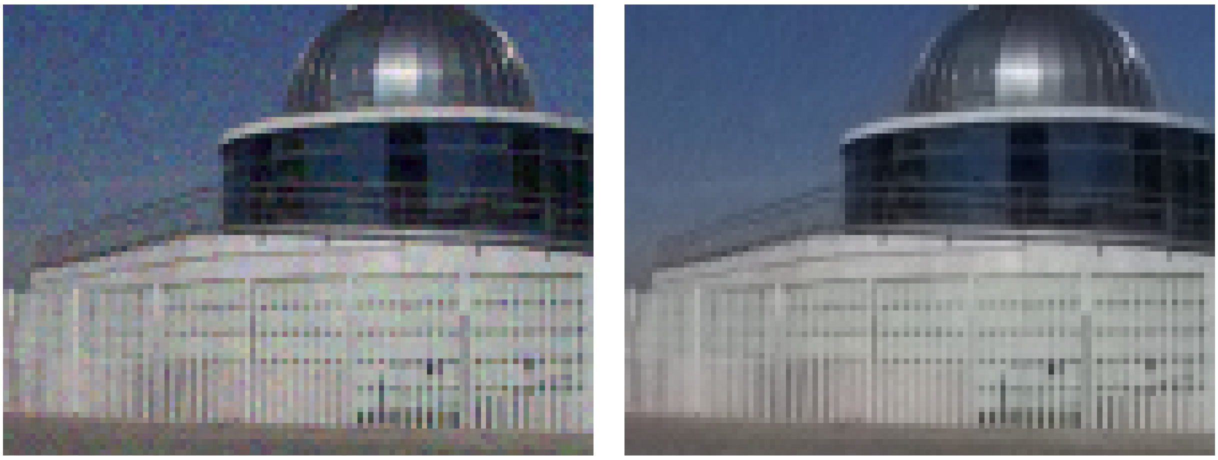
Figure 6.12: *ERGAS, mean COR and mean SSIM evolution for the synthetic HRMS image with* $\lambda_{\beta_b^{ld}} = \lambda_{\gamma_b^{ld}} = 0$ *for the coarser decomposition level and* $\lambda_{\alpha_b^{ld}} = 0$ *for all decomposition levels.*



(a) Original image (b) Observed PAN image

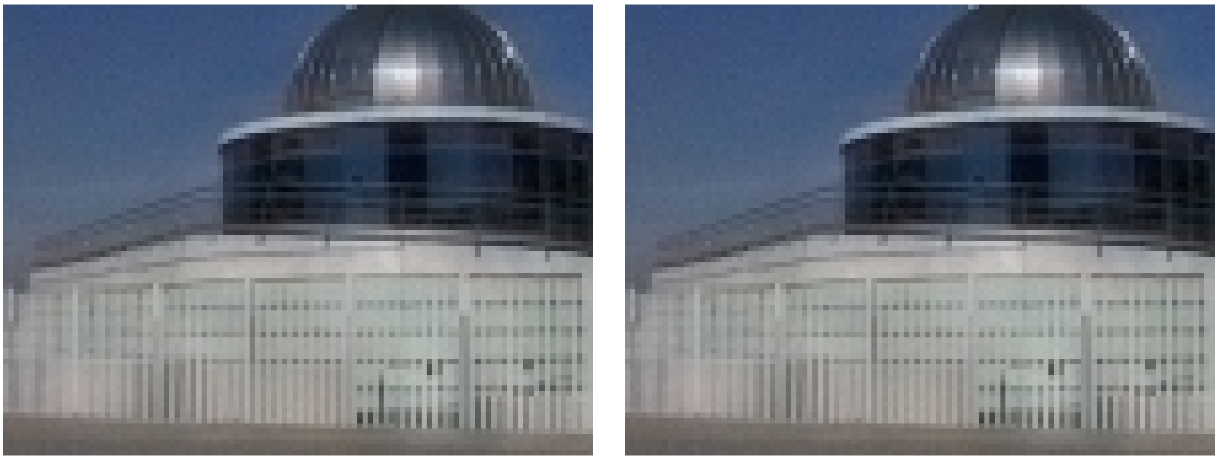


(c) Observed MS image (d) NSCT method in [62]



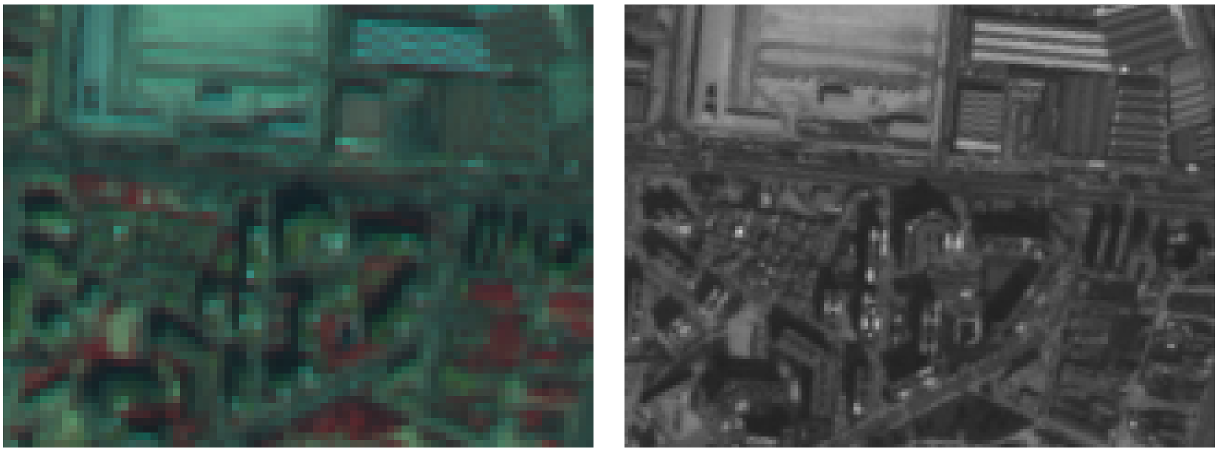
(e) SR method in [96] (f) proposed method($\lambda_\alpha = \lambda_\beta = \lambda_\gamma = 0$) .

Figure 6.13: *Results for the synthetic image*



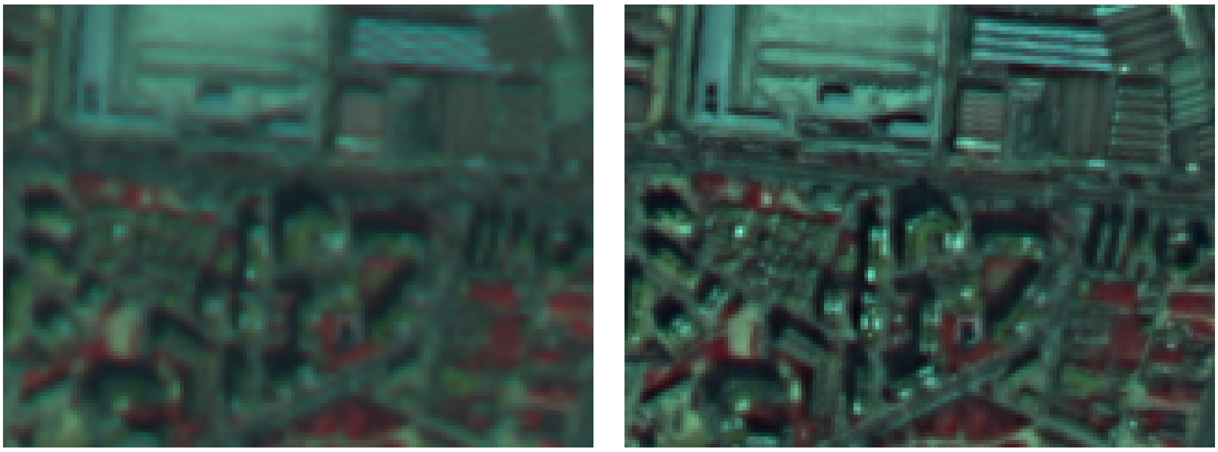
(g) proposed method ($\lambda_\alpha = 0, \lambda_\beta = \lambda_\gamma = 0.9$) (h) proposed method ($\lambda_\alpha = 0, \lambda_\beta = 0.6, \lambda_\gamma = 1.0$).

Figure 6.13: *Results for the synthetic image (contd.)*



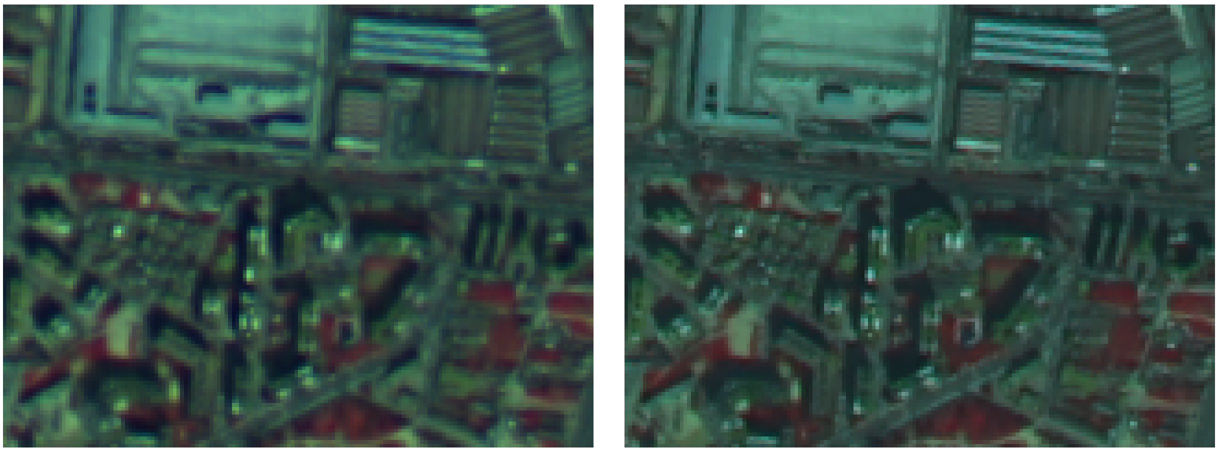
(a) Observed MS image

(b) Observed PAN image



(c) Bicubic Interpolation

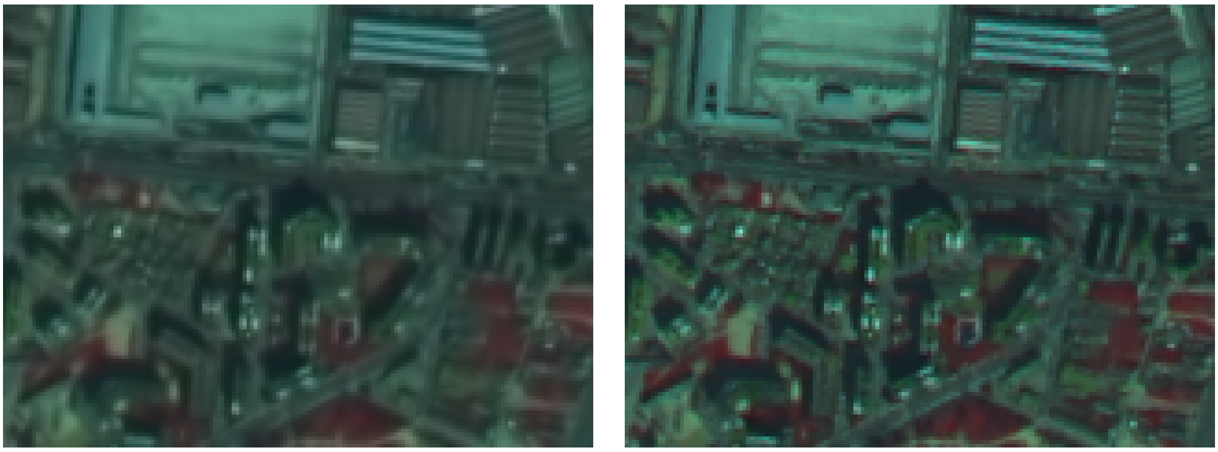
(d) NSCT method in [62]



(e) CiSper in [29]

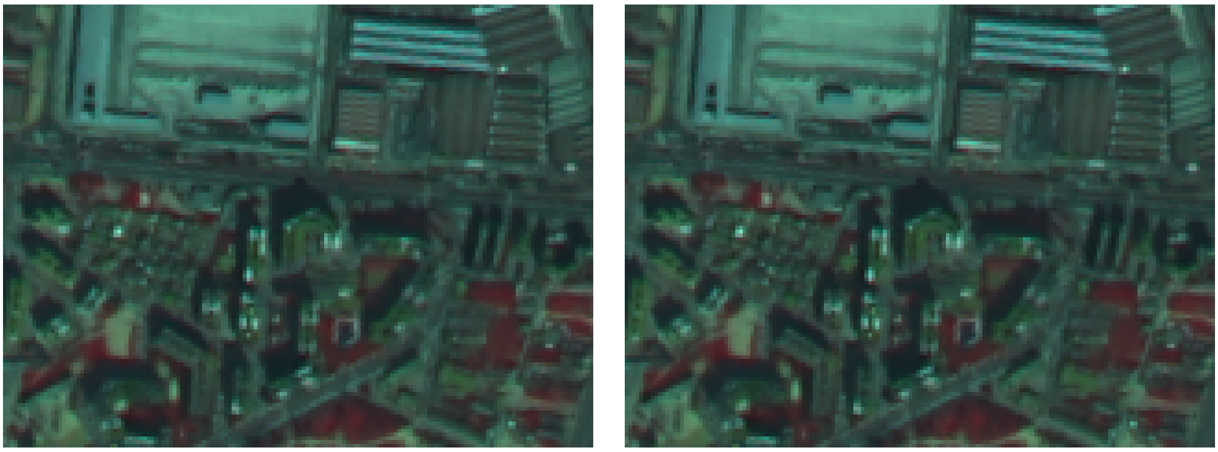
(f) the method proposed in Chapter 5

Figure 6.14: *Results for the SPOT5 image*



(g) SR method in [96]

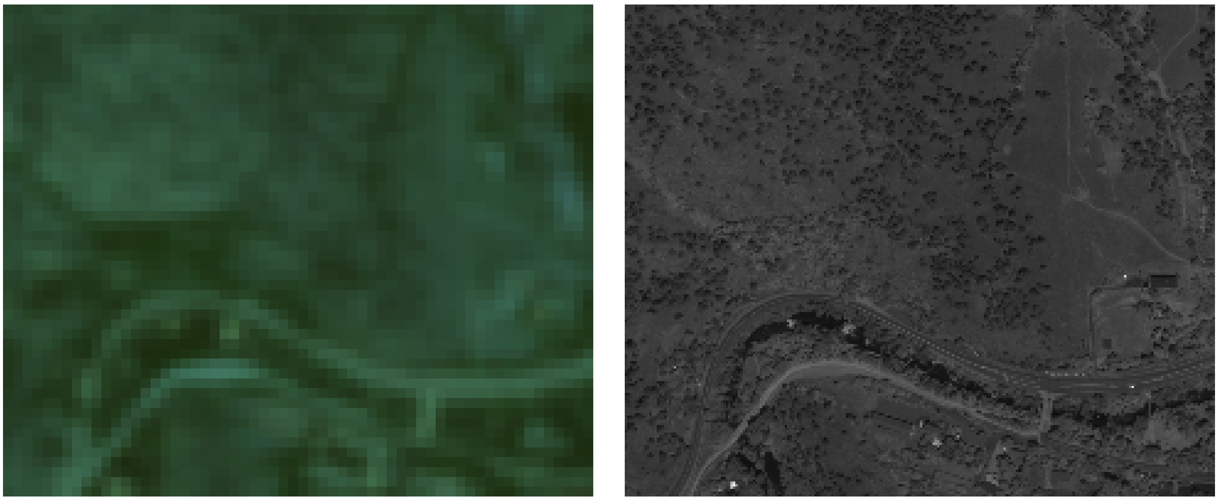
(h) proposed method ($\lambda_\alpha = \lambda_\beta = \lambda_\gamma = 0$).



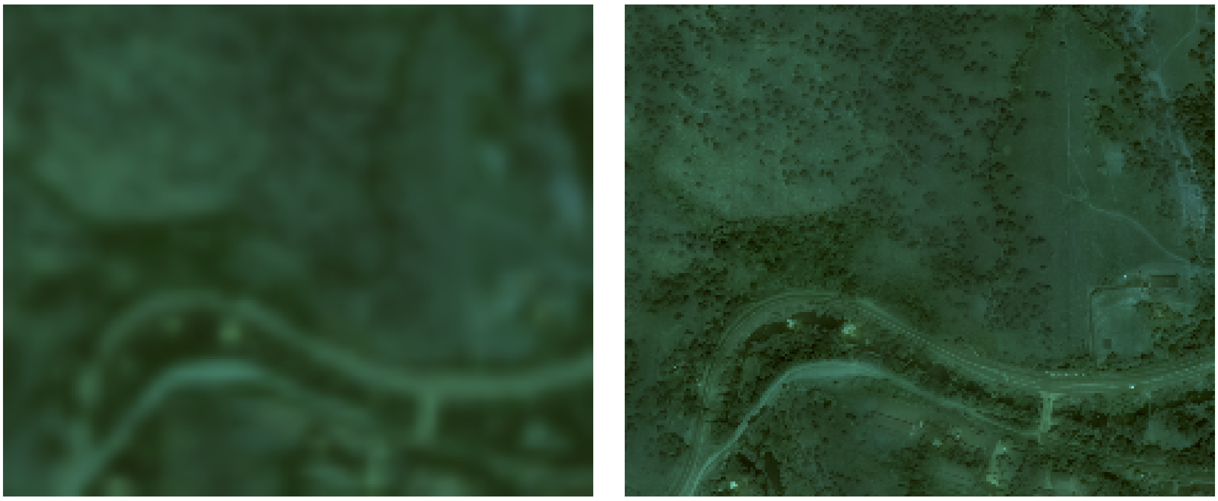
(i) proposed method ($\lambda_\alpha = 0, \lambda_\beta = \lambda_\gamma = 0.9$)

(j) proposed method ($\lambda_\alpha = 0, \lambda_\beta = 0.6, \lambda_\gamma = 1.0$).

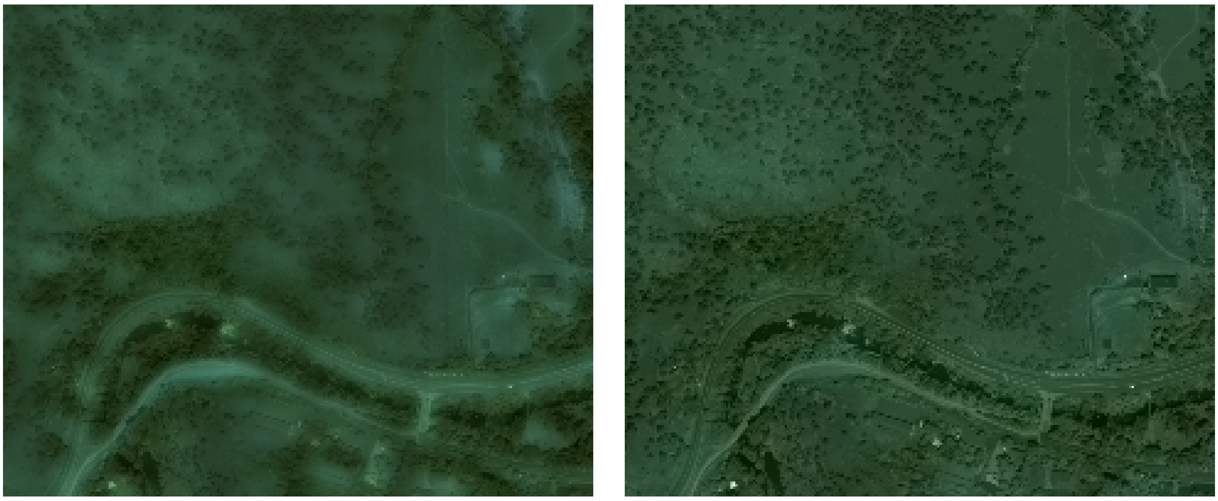
Figure 6.14: *Results for the SPOT5 image (contd.)*



(a) Observed MS image (b) Observed PAN image

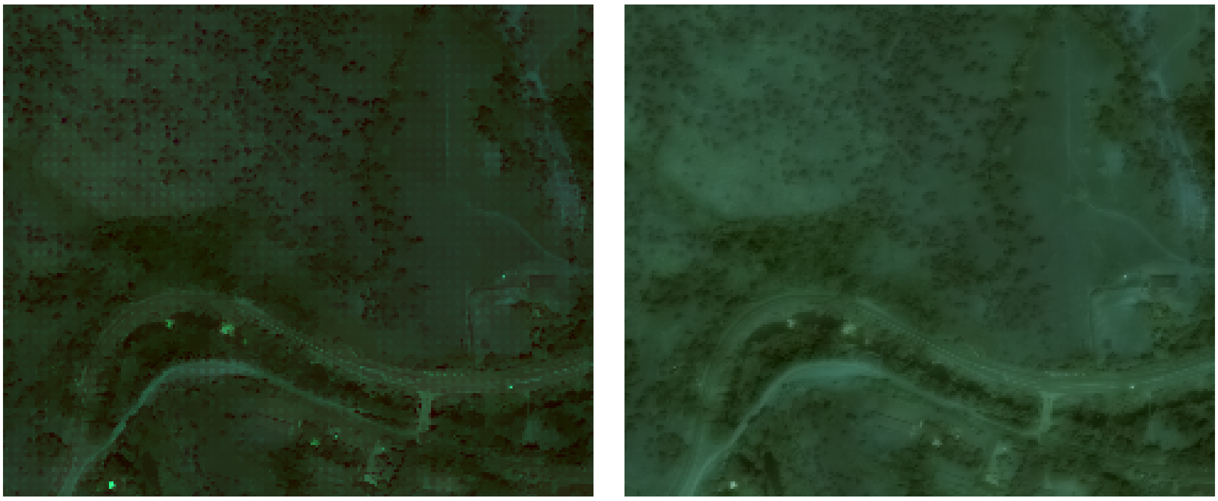


(c) Bicubic Interpolation (d) NSCT method in [62]



(e) CiSper in [29] (f) the method proposed in Chapter 5

Figure 6.15: *Results for the QuickBird image*



(g) SR method in [96]

(h) proposed method ($\lambda_\alpha = \lambda_\beta = \lambda_\gamma = 0$).

Figure 6.15: *Results for the QuickBird image (contd.)*

Chapter 7

Conclusions and Future Works

7.1 Summary

- We have provided a complete overview of the different methods proposed in the literature to tackle the pansharpening prolblem and classified them into different categories according to the main technique they use. An important category are the methods based on the statistics of the image that provides a flexible and powerful way to model the image capture system as well as incorporating the knowledge available about the HR MS image. Those methods allow, as explained in chapter 2, to accurately model the relationship between the HR MS image and the MS and PAN images incorporating the physics of the sensors (MTF, spectral response, or noise properties, for instance) and the conditions in which the images were taken.
- The multiresolution analysis, as already mentioned, is one of the most successful approaches for the pansharpening problem. Most of those techniques have been previously classified into techniques relevant to the ARSIS concept. Decomposing the images at different resolution levels allows to inject the details of the PAN image into the MS one depending on the context. From the methods described in chapter 2, we can see that the generalized Laplacian pyramid and redundant wavelet and contourlet transforms are the most popular multiresolution techniques applied to this fusion problem.
- Different image pansharpening methods based on the undecimated wavelet and contourlet transform (Additive, IHS and PCA) have been experimentally compared in chapter 4. Also the new proposed method CiSper was compared with these methods and with

WiSper. In all methods, contourlet-based pansharpened images present, visually and numerically, better results than those obtained by wavelet for both Landsat 7, SPOT 5 and QuickBird imagery, injecting spatial information from the PAN image missed in the MS image, without modifying its spectral information content.

- A new pansharpening method based on super-resolution reconstruction and non subsampled contourlet transform has been presented. The proposed method preserves the spectral properties of MS image while incorporating the high frequencies from the panchromatic image and controlling the noise in the image. The efficiency of pansharpening methods has been evaluated by means of visual and quantitative analysis, for synthetic and real data such as SPOT5, Landsat7 and QuickBird. Based on the presented experiments, the proposed method does significantly outperform NSCT-based and TV-based super-resolution methods.

- Another new pansharpening method that generalizes the fusion strategy of the panchromatic and multispectral images in contourlet based methods has been proposed. The proposed fusion algorithm is based on the Bayesian modeling and incorporates a solid way to incorporate the details in the panchromatic into the multispectral image while controlling the noise. Particular cases of the proposed fusion method are classical substitution, additive and weighted contourlets methods. The proposed pansharpening method has been compared with other methods both in synthetic and real images such as SPOT5, Landsat7 and QuickBird and its performance has been assessed both numerically and visually.

7.2 Future Works

The research presented in this thesis seems to have raised more questions that it has answered. There are several lines of research arising from this work which should be pursued. The development of new pansharpening methods for remote sensing images that can preserve both spatial and spectral details and be implemented by fast algorithms, is one of the most active research areas in the image processing community. The Contourlet transform in this thesis is very promising and leads to a variety of possible extensions. In the following, we give an overview of ongoing and future research directions:

In most cases it is preferable to convert satellite image data to physical quantities before using the data to interpret the landscape. Important physical quantities include spectral radiance and spectral reflectance. The term radiance is used to characterize the entire solar spectrum while spectral radiance is used to characterize the light at a single wavelength or band interval. Hence, we need to work on these pansharpening methods using spectral reflectance values of the images instead of the digital number (DN). This should allow better extraction for HRMS images. Still the models used are not very sophisticated thus presenting an open research area in this family.

A second line of research, which follows from chapter 4, is to study the optimum number of contourlet decomposition levels for each proposed methods, and the dependency between these levels and the spatial resolution ratio between MS and PAN images, since we found in our experiments on different sensor images with different ratios, that there seems to be a relation between the ratio and the number of decompositions for all of the used pansharpening methods.

Following from Chapters 5 and 6 there is a need to investigate the incorporation of more prior knowledgement like, the correlation between the bands, $L1$ models and mixture of Gaussian, since as we mentioned before many studies conclude that the contourlet coefficients are non-Gaussian but conditionally Gaussian, so the contourlet coefficients of natural images may be accurately modeled, for instance, by these prior distributions whose variances depend on their generalized neighborhood coefficients.

Also following from Chapters 5, is the possibility of reducing the number of parameters, by approximating the distribution parameters in each level of decomposition depending on the Contourlet transform matrix and the parameters of the image before the contourlet decompositions. While for Chapter 6 we found that there is a relation between different decomposition levels, but we need to study more these relationships with different levels and directions and justify it.

Propose new pansharpening methods that takes into account the physical characteristics of the Remote sensor and MTF. Also methods like adaptive GS which can be modified to work in the contourlet domain. Moreover, the study of pansharpening based on other multiscale transforms like Curvelets, wavelets and Gaussian Laplacian pyramid.

Finally, in Chapter 2, we discussed the Quality without Reference quantitative measures, as a concept, but we need to investigate more on these measures and compare it with used ones in this thesis.

Bibliography

[1] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, “A comparative analysis of image fusion methods,” *IEEE Trans. Geosci.Remote Sens.*, vol. 43, no. 6, pp. 1391–1402, 2005.

[2] L. Wald, “Some terms of reference in data fusion,” *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1190–1193, 1999.

[3] C. Pohl and J. L. V. Genderen, “Multi-sensor image fusion in remote sensing: Concepts, methods, and applications,” *Int. J. Remote Sens.*, vol. 19, no. 5, pp. 823–854, 1998.

[4] T. M. Tu, S. C. Su, H. C. Shyu, and P. S. Huang, “A new look at IHS-like image fusion methods,” *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, 2001.

[5] X. Otazu, M. González-Audícana, O. Fors, and J. Núñez, “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, 2005.

[6] I. Couloigner, T. Ranchin, V. P. Valtonen, and L. Wald, “Benefit of the future SPOT-5 and of data fusion to urban roads mapping,” *Int. J. Remote Sensing*, vol. 19, no. 8, pp. 1519–1532, 1998.

[7] L. Alparone, L. Wald, J. Chanussot, P. Gamba, and L. M. Bruce, “Comparison of pan-sharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest,” *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, 2007.

[8] Google Earth, http://earth.google.com/, *Google Earth Web Site*.

[9] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*. Orlando, FL: Academic, third ed., 1997.

[10] *Landsat Web Page*. http://www.landsat.org.

[11] *SPOT Web Page*. http://www.spotimage.fr.

[12] *Remote Sensing Tutorial*. http://www.gisdevelopment.net/tutorials/tuman008.htm.

[13] D. Landgrebe, "The evolution of Landsat data analysis," *Photogrammetric Engineering and Remote Sensing*, vol. 63, pp. 859–867, July 1997.

[14] *NASA Web Page*. http://geo.arc.nasa.gov/sge/landsat/l7.html.

[15] *EOS Web Page*. http://eospso.gsfc.nasa.gov/.

[16] *AVIRIS Web Page*. http://aviris.jpl.nasa.gov/.

[17] *MERIS Web Page*. http://envisat.esa.int/handbooks/meris/.

[18] *MODIS Web Page*. http://modis.gsfc.nasa.gov/.

[19] *IKONOS Web Page*. http://www.geoeye.com/CorpSite/products/imagery-sources/.

[20] *QuickBird Web Page*. http://www.digitalglobe.com/index.php/85/QuickBird.

[21] *OrbView Web Page*. http://www.tec.army.mil/tio/ORBVIEW.htm.

[22] H. J.Kramer, *Observation Of The Earth And Its Environment: Survey Of Missions And Sensors.* Berlin: Springer-Verlag, fourth ed., 2002.

[23] *Landsat 7 Web Page*. http://geo.arc.nasa.gov/sge/landsat/.

[24] J. G. 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.

[25] V. Vijayaraj, "A quantitative analysis of pansharpened images," Master's thesis, Mississippi State Univ., 2004.

[26] *Pan Sharpening Tutorial*. http://www.geospatial.overwatch.com/.

[27] 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. Remote Sens.*, vol. 46, pp. 1301–1312, 2008.

[28] I. Amro and J. Mateos, “Multispectral image pansharpening based on the contourlet transform,” in *Advances in Information Optics and Photonics* (B. Javidi and T. Fournel, eds.), Springer, 2010.

[29] I. Amro and J. Mateos, “Multispectral image pansharpening based on the contourlet transform,” in *8th International Workshop on Information Optics (WIO’ 09) - Journal of Physics: Conference Series*, vol. 206, pp. 012031, DOI:10.1088/1742–6596/206/1/012031, Paris (France), July 2009.

[30] I. Amro, J. Mateos, and M. Vega, “Bayesian super-resolution pansharpening using contourlets,” in *International Conference on Image Processing 2010 - Finalist of the Huawei Best Student Paper Award*, pp. 809–812, Honk Kong (China), September 2010.

[31] I. Amro, J. Mateos, and M. Vega, “General contourlet pansharpening method using Bayesian inference,” in *2010 European Signal Processing Conference (EUSIPCO-2010)*, pp. 294–298, 2010.

[32] I. Amro, J. Mateos, M. Vega, R. Molina, and A. Katsaggelos, “Classical methods and new trends in pansharpening of multispectral images.” Journal of Selected Topics in Signal Processing "Submitted", 2010.

[33] M. Ehlers, S. Klonus, and P. Astrand, “Quality assessment for multi-sensor multi-date image fusion,” in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.*, vol. XXXVII of *Part B4, Beijing*, 2008.

[34] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., fourth ed., 2005.

[35] C. L. Parkinson, A. Ward, and M. D. King, eds., *Earth Science Reference Handbook A Guide to NASA’s Earth Science Program and Earth Observing Satellite Missions*. National Aeronautics and Space Administration, 2006.

[36] Y. Yang and X. Gao, “Remote sensing image registration via active contour model,” *Int. J. Electron. Commun.*, vol. 63, pp. 227 –234, 2009.

[37] B. Zitova and J. Flusser, “Image registration methods: a survey,” *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.

[38] K. K. Teoh, H. Ibrahim, and S. K. Bejo, “Investigation on several basic interpolation methods for the use in remote sensing application,” in *Proceedings of the 2008 IEEE Conference on Innovative Technologies*, 2008.

[39] R. C. Gonzalez and R. E. Woods, *Digital image processing*. Prentice Hall, third ed., 2008.

[40] J. Parker, R. Kenyon, and D. Troxel, “Comparison of interpolating methods for image resampling,” *IEEE Transactions on Medical Imaging*, vol. 2, pp. 31–39, 1983.

[41] W. Dou and Y. Chen, “An improved IHS image fusion method with high spectral fidelity,” *The Int. Archiv. of the Photogramm., Rem. Sensing and Spat. Inform. Sciences*, vol. XXXVII, pp. 1253–1256, 2008. part.B7.

[42] F. H. Wong and R. Orth, “Registration of SEASAT/LANDSAT composite images to UTM coordinates,” in *Proceedings of the Sixth Canadian Syinposium on Remote Sensing*, 1980.

[43] P. Rebillard and P. T. Nguyen, “An exploration of co-registered SIR-A, SEASAT and landsat images,” in *RS of Environment, RS for Exploration Geology, Proceedings International Symposium, Second Thematic Conference, Forth Worth, U.S.A. (Forth Worth: ISRS)*, pp. 109 – 118, 1982.

[44] R. Simard, “Improved spatial and altimetric information from SPOT composite imagery,” in *Proceedings ISPRS Conference, Forth Worth, U.S.A. (Forth Worth: International Archive of Photogrammetry and Remote Sensing)*, pp. 434 – 440, 1982.

[45] E. P. Crist, “Comparison of coincident Landsat-4 MSS and TM data over an agricultural region,” in *Technical Papers, Proceedings 50th Annual Meeting ASP-ACSM Symposium, ASPRS, Washington DC, U.S.A. (Washington: ASPRS)*, pp. 508–517., 1984.

[46] G. Cliche, F. Bonn, and P. Teillet, “Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancement,” *Photogramm. Eng. Remote Sens*, vol. 51, pp. 311–316, March 1985.

[47] R. Welch and M. Ehlers, “Merging multiresolution SPOT HRV and landsat TM data,” *Photogramm. Eng. Remote Sens*, vol. 53, pp. 301–303, March 1987.

[48] R. Haydn, G. W. Dalke, and J. Henkel, “Application of the IHS color transform to the processing of multisensor data and image enhancemen,” in *lnternational Symposium on Remote Sensing of Arid and Semi-Arid Lands, Cairo, Egypt*, 1982.

[49] P. S. Chavez, Jr and J. A. Bowell, “Comparison of the spectral information content of Landsat Thematic Mapper and SPOT for three different sites in the Phoenix, Arizona region,” *Photogramm. Eng. Remote Sens.*, vol. 54, no. 12, pp. 1699–1708, 1988.

[50] A. R. Gillespie, A. B. Kahle, and R. E. Walker, “Color enhancement of highly correlated images. II. Channel Ratio and "Chromaticity" Transformation Techniques,” *Remote Sensing Of Environment*, vol. 22, pp. 343–365, 1987.

[51] P. S. Chavez and A. Y. 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, 1989.

[52] P. Chavez, S. Sides, and J. Anderson, “Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic,” *Photogrammetric Engineering and Remote Sensing*, vol. 57, pp. 295–303, March 1991.

[53] J. C. Price, “Combining panchromatic and multispeetral imagery from dual resolution satellite instruments,” *Remote Sensing Of Environment*, vol. 21, pp. 119–128, 1987.

[54] J. Park and M. Kang, “Spatially adaptive multi-resolution multispectral image fusion,” *Int. J. Remote Sensing*, vol. 25, pp. 5491– 5508, December 2004.

[55] N. D. A. Mascarenhas, G. J. F. Banon, and A. L. B. Candeias, “Multispectral image data fusion under a bayesian approach,” *International Journal of Remote Sensing*, vol. 17, pp. 1457–1471, 1996.

[56] M. Vega, J. Mateos, R. Molina, and A. Katsaggelos, “Super-resolution of multispectral images,” *The Computer Journal*, vol. 1, pp. 1–15, 2008.

[57] 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. J. GRS.*, vol. 40, pp. 2300–2312, Oct 2002.

[58] A. Garzelli and F. Nencini, “Interband structure modeling for pan-sharpening of very high-resolution multispectral images,” *Information Fusion*, vol. 6, pp. 213–224, 2005.

[59] S. G. Mallat, “A theory for multiresolution signal decomposition: The wavelet representation,” *IEEE Transactions On Pattern Analysis And Machine Intelligence*, vol. 11, pp. 674–693, July 1989.

[60] J. Zhou, D. L. Civco, and J. A. Silander, “A wavelet transform method to merge Landsat TM and SPOT panchromatic data,” *Int. J. Remote Sens.*, vol. 19, no. 4, pp. 743 – 757, 1998.

[61] A. L. da Cunha, J. Zhou, and M. N. Do, “The nonsubsampled contourlet transform: Theory, design, and applications,” *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, 2006.

[62] M. Lillo-Saavedra and C. Gonzalo, “Multispectral images fusion by a joint multidirectional and multiresolution representation,” *Int. J. Remote Sens.*, vol. 28, no. 18, pp. 4065 – 4079, 2007.

[63] I. Amro and J. Mateos, “Multispectral image pansharpening based on the contourlet transform,” *Journal of Physics. Conference Series*, vol. 206, no. 1, pp. 1–3, 2010.

[64] T. Ranchln and L. Wald, “Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation,” *Photogramm. Eng. Remote Sens.*, vol. 66, pp. 49–61, 2000.

[65] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, and L. Wald, “Image fusion: The ARSIS concept and some successful implementation schemes,” *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 58, pp. 4–18, 2003.

[66] T. Bretschneider and O. Kao, “Image fusion in remote sensing.” Technical University of Clausthal, Germany.

[67] V. K. Sheftigara, “A generalized component substitution technique for spatial enhancement of multispectral lmages using a higher resolution data set,” *Photogramm. Eng. Remote Sens*, vol. 58, pp. 561 – 567, May 1992.

[68] W. Dou, Y. Chen, X. Li, and D. Z. Sui, “A general framework for component substitution image fusion: An implementation using the fast image fusion method,” *Computers & Geosciences*, vol. 33, pp. 219–228, 2007.

[69] B. Aiazzi, S. Baronti, and M. Selva, “Improving component substitution pansharpening through multivariate regression of MS + Pan data,” *IEEE Transactions On Geoscience And Remote Sensing,*, vol. 45, pp. 3230–3239, October 2007.

[70] L. Alparone, B. Aiazzi, S. B. A. Garzelli, and F. Nencini, “A critical review of fusion methods for true colour display of very high resolution images of urban areas,” in *1st EARSeL Workshop of the SIG Urban Remote Sensing, Humboldt-Universität zu Berlin*, 2006.

[71] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, “An adaptive IHS pan-sharpening method,” *IEEE Geoscience And Remote Sensing Letters*, vol. 7, no. 3, pp. 746–750, 2010.

[72] V. P. Shah, N. H. Younan, and R. L. 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, 2008.

[73] B. Aiazzi, S. Baronti, M. Selva, and L. Alparone, “Enhanced Gram-Schmidt spectral sharpening based on multivariate regression of MS and pan data,” in *IEEE International Conference on Geoscience and Remote Sensing Symposium, IGARSS 2006*, pp. 3806–3809, 2006.

[74] W. J. Carper, T. M. Lillesand, and R. W. Kiefer, “The use of Intensity-Hue-Saturation transform for merging SPOT panchromatic and multispectral image data,” *Photogramm. Eng. Remote Sens*, vol. 56, no. 4, pp. 459–467, 1990.

[75] A. R. Smith, “Color gamut transform pairs,” *Computer Graphics*, vol. 13, no. 3, pp. 12–19, 1978.

[76] F. C. Silva and L. V. Dutra, “Urban remote sensing image enhancement using a generalized IHS fusion technique, PUC-Rio, Rio de Janeiro, Brazil,” in *Eleventh Ursi Commission For Open Symposium On Radio Wave Propagation And Remote Sensing*, 2007.

[77] K. Castleman, *Digital Image Processing*. Prentice Hall , Upper Saddle River, NewJersey, 1995.

[78] C. A. Laben and B. V. Brower, “Process for enhanceing the spatial resolution of multispectral imagery using pansharpening,” 2000.

[79] R. W. Farebrother, “Gram-schmidt regression,” *Applied Statistics*, vol. 23, pp. 470–476, 1974.

[80] *SPOT Users Handbook*, vol. 1-3. Centre National Etude Spatiale (CNES) and SPOT Image, Toulouse, France, 1988.

[81] C. Ballester, V. Caselles, L. Igual, and J. Verdera, “A variational model for P+XS image fusion,” *International Journal of Computer Vision*, vol. 69, pp. 43–58,, 2006.

[82] G. Cliche, F. Bonn, and P. Teillet, “Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancement,” *Photogrammetric Engineering & Remote Sensing*, vol. 51, pp. 311–316, March 1985.

[83] V. Vijayaraj, C. G. O’Hara, and N. H. Younan, “Quality analysis of pansharpened images,” in *Proc. IEEE Int. Geosc. Remote Sens. Symp. IGARSS ’04*, vol. 1, 20–24 2004.

[84] J. G. LIU, “Smoothing filter-based intensity modulation: a spectral preserve image fusion technique for improving spatial details,” *Int. J. Remote Sensing*, vol. 21, pp. 3461–3472, 2000.

[85] L. Alparone, L. Facheris, S. Baronti, A. Garzelli, and F. Nencini, “Fusion of multispectral and SAR images by intensity modulation,” in *Proceedings of the 7th International Conference on Information Fusion*, pp. 637–643, 2004.

[86] R. A. Schowengerdt, “Reconstruction of multispatial, multispectral image data using spatial frequency contents,” *Photogrammetric Engineering & Remote Sensing*, vol. 46, pp. 1325–1334, October 1980.

[87] Pat S. Chavez, Jr., “Digital merging of Landsat TM and digitized NHAP data for 1:24,000 scale image mapping,” *Photogrammetric Engineering And Remote Sensing*, vol. 52, pp. 1637–1646, October 1986.

[88] V. J. D. Tsai, “Frequency-based fusion of multiresolution images,” *Geoscience and Remote Sensing Symposium, 2003. IGARSS ’03. Proceedings. 2003 IEEE International*, vol. 6, pp. 3665– 3667, 2003.

[89] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogramm. Eng. Remote Sens.*, vol. 63, pp. 691 – 699, 1997.

[90] M. González-Audícana, J. Saleta, R. García Catalán, and R. García, "Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition," *IEEE Trans. Geosc. Remote Sens.*, vol. 42, no. 6, pp. 1291–1298, 2004.

[91] M. M. Khan, L. Alparone, and J. Chanussot, "Pansharpening quality assessment using the modulation transfer functions of instruments," *IEEE Trans. Geosc. Remote Sens.*, vol. 47, no. 11, pp. 3880–3891, 2009.

[92] C. Thomas and L. Wald, "A MTF-based distance for the assessment of geometrical quality of fused products," in *2006 9th International Conference on Information Fusion*, pp. 1–7, 2006.

[93] J. C. Price, "Comparison of the information content of data from the Landsat-4 Thematic Mapper and the Multispectral Scanner," *IEEE Trans. Geosci. Remote Sens.*, vol. 22, pp. 272–281, 1984.

[94] J. C. Price, "Statistical analysis of SPOT simulation data and comparison with Landsat-4 Thematic Mapper data," in *SPOT Simulation Applications Handbook, Am. Soc. of Photogrammetry, Falls Church, VA.*, 1984.

[95] J. C. Price, "Combining multispectral data of differing spatial resolution," *IEEE. J. GRS*, vol. 37, pp. 1199–1203, May 1999.

[96] M. Vega, J. Mateos, R. Molina, and A. Katsaggelos, "Super resolution of multispectral images using TV image models," in *2th Int. Conf. on Knowledge-Based and Intelligent Information & Eng. Sys.*, pp. 408–415, 2008.

[97] O. Punska, "Bayesian approaches to multi-sensor data fusion," Master's thesis, Signal Processing and Communications Laboratory, Department of Engineering, University of Cambridge, 1999.

[98] D. Fasbender, J. Radoux, and P. Bogaert, "Bayesian data fusion for adaptable image pansharpening," *IEEE Transactions On Geoscience And Remote Sensing*, vol. 46, pp. 1847–1857, 2008.

[99] 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, 2004.

[100] R. Molina, M. Vega, J. Mateos, and A. K. Katsaggelos, “Variational posterior distribution approximation in Bayesian super resolution reconstruction of multispectral images,” *Applied And Computational Harmonic Analysis*, vol. 12, pp. 1–27, 2007.

[101] H. G. Kitaw, *Image pan-sharpening with Markov random field and simulated annealing*. PhD thesis, International Institute for Geo-information Science and Earth Observation, NL, 2007.

[102] M. T. Eismann and R. C. Hardie, “Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions,” *IEEE Transactions On Geoscience And Remote Sensing*, vol. 43, pp. 455–465, Mar 2005.

[103] M. Vega, J. Mateos, R. Molina, and A. Katsaggelos, “Super resolution of multispectral images using l1 image models and interband correlations,” in *2009 IEEE International Workshop on Machine Learning for Signal Processing*, pp. 1–6, Grenoble (France), September 2009.

[104] M. T. Eismann and R. C. Hardie, “Application of the stochastic mixing model to hyperspectral resolution enhancement,” *IEEE Transactions On Geoscience And Remote Sensing*, vol. 42, pp. 1924–1933, Sep 2004.

[105] A. Katsaggelos, R. Molina, and J. Mateos, *Super Resolution Of Images And Video*. Synthesis Lectures on Image, Video, and Multimedia Processing, Morgan & Claypool, 2007.

[106] R. Molina, J. Mateos, A. K. Katsaggelos, and R. Z. Milla, “A new super resolution Bayesian method for pansharpening Landsat ETM+ imagery,” in *9th International Symposium on Physical Measurements and Signatures in Remote Sensing (ISPMSRS)* (S. Liang, X. L. J. Liu, R. Liu, and M. Schaepman, eds.), pp. 280–283, October 2005.

[107] G. Z. Rong, W. Bin, and Z. L. Ming, “Remote sensing image fusion based on Bayesian linear estimation,” *Science in China Series F: Information Sciences*, vol. 50, no. 2, pp. 227–240, 2007.

[108] W. Niu, J. Zhu, W. Gu, and J. Chu, “Four statistical approaches for multisensor data fusion under non-gaussian noise,” in *IITA International Conference on Control, Automation and Systems Engineering*, 2009.

[109] L. Brandenburg and H. Meadows, “Shaping filter representation of nonstationary colored noise,” *IEEE Transactions on Information Theory*, vol. 17, pp. 26–31, 1971.

[110] P. J. Burt and E. H. Adelson, “The laplacian pyramid as a compact image code,” *IEEE Transactions On Communications*, vol. COM-3l, pp. 532–540, April 1983.

[111] M. N. Do and M. Vetterli, “The contourlet transform: an efficient directional multiresolution image representation,” *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, 2005.

[112] J. Zhang, “Multi-source remote sensing data fusion: status and trends,” *International Journal of Image and Data Fusion*, vol. 1, no. 1, pp. 5–24, 2010.

[113] L. Alparone, S. Baronti, and A. Garzelli, “Assessment of image fusion algorithms based on noncritically-decimated pyramids and wavelets,” in *Proc. IEEE 2001 International Geoscience and Remote Sensing Symposium IGARSS ’01*, vol. 2, pp. 852–854, 9–13 July 2001.

[114] M. G. Kim, I. Dinstein, and L. Shaw, “A prototype filter design approach to pyramid generation,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, no. 12, pp. 1233–1240, 1993.

[115] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, “Spectral information extraction by means of MS+PAN fusion,” in *Proceedings of ESA-EUSC 2004 - Theory and Applications of Knowledge-Driven Image Information Mining with Focus on Earth Observation*, p. 20.1, Mar 2004.

[116] M. González-Audícana and X. Otazu, “Comparison between Mallat’s and the a’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, 2005.

[117] B. Garguet-Duport, J. Girel, J.-M. Chassery, and G. Pautou, “The use of multiresolution analysis and wavelets transform for merging SPOT panchromatic and multispectral image data,” *Photogramm.Eng.Remote Sens.*, vol. 62, no. 9, pp. 1057–1066, 1996.

[118] D. A. Yocky, “Image merging and data fusion by means of the discrete two-dimensional wavelet transform,” *Optical Society of America*, vol. 12, pp. 1834 –1841, Septemper 1995.

[119] J. Nuñez, 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, 1999.

[120] Y. Chibani and A. Houacine, “The joint use of IHS transform and redundant wavelet decomposition for fusing multispectral and panchromatic image,” *Int. J. Remote Sensing*, vol. 23, no. 18, pp. 3821–3833, 2002.

[121] P. S. Pradhan, R. L. King, N. H. Younan, and D. W. Holcomb, “Estimation of the number of decomposition levels for a wavelet-based multiresolution multisensor image fusion,” *IEEE Transactions On Geoscience And Remote Sensing*, vol. 44, no. 12, pp. 3674–3686, 2006.

[122] M. Song, X. Chen, and P. Guo, “A fusion method for multispectral and panchromatic images based on HSI and contourlet transformation,” in *Proc. 10th Workshop on Image Analysis for Multimedia Interactive Services WIAMIS ’09*, pp. 77–80, 6–8 May 2009.

[123] A. M. ALEjaily, I. A. E. Rube, and M. A. Mangoud, “Fusion of remote sensing images using contourlet transform,” *Springer Science*, pp. 213 – 218, 2008.

[124] S. Yang, M. Wang, Y. X. Lu, W. Qi, and L. Jiao, “Fusion of multiparametric SAR images based on SW-nonsubsampled contourlet and PCNN,” *Signal Processing*, vol. 89, pp. 2596–2608, 2009.

[125] J. Wu, H. Huang, J. Liu, and J. Tian, “Remote sensing image data fusion based on IHS and local deviation of wavelet transformation,” in *Proc. IEEE Int. Conf. on Robotics and Biomimetics ROBIO 2004*, pp. 564 – 568, 2004.

[126] Y. Zhang, S. D. Backer, and P. Scheunders, “Bayesian fusion of multispectral and hyperspectral image in wavelet domain,” *Geoscience and Remote Sensing Symposium, IGARSS IEEE International*, vol. 5, pp. V–69 –V–72, 2008.

[127] V. Meenakshisundaram, “Quality assessment of Ikonos and Quickbird Fused Images for Urban Mapping,” Master’s thesis, UNIVERSITY OF CALGARY, 2005.

[128] L. Wald, *Data Fusion Definitions and Architectures: Fusion of Images of Different Spatial Resolutions*. Les Presses de l-Ecole des Mines, Paris, 2002.

[129] F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, “Remote sensing image fusion using the curvelet transform,” *Information Fusion*, vol. 8, pp. 143–156, 2007.

[130] V. Vijayaraj, N. H. Younan, and C. G. O’Hara, “Quantitative analysis of pansharpened images,” *Optical Engineering*, vol. 45, no. 4, pp. 046202 (1–12), 2006.

[131] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, pp. 600 – 612, 2004.

[132] Z. Wang and A. C. Bovik, “A universal image quality index,” *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, 2002.

[133] L. Wald, “Quality of high resolution synthesized images: Is there a simple criterion?,” *Proc. Int. Conf. Fusion of Earth Data. Nice, France*, vol. 1, pp. 99 –105, 2000.

[134] Q. Du, N. H. Younan, R. King, and V. P. Shah, “On the performance evaluation of pan-sharpening techniques,” *IEEE. J. GRSL.*, vol. 4, pp. 518–522, Oct. 2007.

[135] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, “A global quality measurement of pan-sharpened multispectral imagery,” *IEEE Geoscience And Remote Sensing Letters*, vol. 1, no. 4, pp. 313–317, 2004.

[136] C. Thomas and L. Wald, “Comparing distances for quality assessment of fused images,” in *Proceedings of the 26th EARSeL Symposium “New Strategies for European Remote Sensing”*, pp. 101–111, 2007.

[137] M. Lillo-Saavedra, C. Gonzalo, A. Arquero, and E. Martinez, “Fusion of multispectral and panchromatic satellite sensor imagery based on tailored filtering in the fourier domain,” *Int. J. Remote Sens.*, vol. 26, pp. 1263–1268, 2005.

[138] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, “Multispectral and panchromatic data fusion assessment without reference,” *Photogrammetric Engineering & Remote Sensing*, vol. 74, pp. 193–200, 2008.

[139] L. Alparone, B. Aiazzi, S. B. A. Garzelli, and F. Nencini, “A critical review of fusion methods for true colour display of very high resolution images of urban areas,” in *1st EARSeL Workshop of the SIG Urban Remote Sensing, Humboldt-Universität zu Berlin*, 2006.

[140] A. Duijster, P. Scheunders, and S. De Backer, “Wavelet-based em algorithm for multispectral-image restoration,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3892–3898, 2009.

[141] G. Aubert and P. Kornprobst, *Mathematical problems in image processing: partial differential equations and the calculus of variations*. Spriger Verlag New York, 2000.

[142] R. Lagendijk and J.Biemond, *Iterative Identification and Restoration of Images*. Kluwer Academic Publishers, 1991.

[143] R. Molina, “On the hierarchical bayesian approach to image restoration: applications to astronomical images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 11, pp. 1122–1128, 1994.

[144] R. Molina, A. K. Katsaggelos, and J. Mateos, “Bayesian and regularization methods for hyperparameter estimation in image restoration,” *IEEE Transactions on Image Processing*, vol. 8, no. 2, pp. 231–246, 1999.

[145] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, “Mtf-tailored multiscale fusion of high-resolution MS and pan imagery,” *Photogrammetric Engineering & Remote Sensing*, vol. 72, pp. 591–596, 2006.

[146] N. K. Bose and K. J. Boo, “High-resolution image reconstruction with multisensors,” *John Wiley & Sons, Inc.*, vol. 9, pp. 294–304, 1998.

[147] M. K. Ng and A. M. Yip, “A fast map algorithm for high-resolution image reconstruction with multisensors,” *Multidimens. Syst. Signal Process.*, vol. 12, pp. 143–164, 2001.

[148] W. Buntine, *A Theory of Learning Classification Rules*. PhD thesis, University of Technology, Sydney, Australia, 1991.

[149] W. Buntine, “Theory refinement on bayesian networks,” in *Seventh Conference on Uncertainty in Artificial Intelligence*, pp. 52–60, 1991.

[150] G. Cooper and E. Herkovsits, “A bayesian method for the induction of probabilistic networks from data,” *Machine Learning*, vol. 9, pp. 309–347, 1992.

[151] D. J. Spiegelhalter and S. Lauritzen, “Sequential updating of conditional probabilities on directed graphical structures,” *Networks*, vol. 20, pp. 579–605, 1990.

[152] W. Buntine and A. Weigund, “Bayesian back-propagation,” *Complex Systems*, vol. 5, pp. 603–643, 1991.

[153] D. J. C. MacKay, “A practical bayesian framework for backprop networks,” *Neural Computation*, vol. 4, pp. 448–472, 1992.

[154] S. Gull, *Developments in Maximum Entropy Data Analysis*, ch. in Maximum Entropy and Bayesian Methods, pp. 53–71. Kluwer, 1989.

[155] D. J. C. MacKay, “Bayesian interpolation,” *Neural Computation*, vol. 4, pp. 415–447, 1992.

[156] A. Lopez, R. Molina, A. Katsaggelos, and J. Mateos, “SPECT image reconstruction using compound models,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 16, pp. 317–330, 2002.

[157] A. Lopez, R. Molina, A. K. Katsaggelos, A. Rodriguez, J. Lopez, and J. M. Llamas, “Parameter estimation in bayesian reconstruction of SPECT images: An aid in nuclear medicine diagnosis,” *Wiley Periodicals*, vol. 14, pp. 21–27, 2004.

[158] J. Mateos, A. K. Katsaggelos, and R. Molina, “A bayesian approach for the estimation and transmission of regularization parameters for reducing blocking artifacts,” *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1200–1215, 2000.

[159] N. P. Galatsanos, V. Z. Mesarovic, R. Molina, and A. K. Katsaggelos, “Hierarchical bayesian image restoration from partially known blurs,” *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1784–1797, 2000.

[160] J. Abad, J. Mateos, R. Molina, and F. Cortijo, “A study of methods for choosing the unknown parameters in the bayesian and regularization approaches to image restoration,” in *VI Simposium Nacional de Reconocimiento de Formas y Analisis de Imagenes*, pp. 358–365, Cordoba (Spain), 1995.

[161] R. Molina, A. del Olmo, J. Perea, and B. Ripley, “Bayesian deconvolution in optical astronomy,” *The Astronomical Journal*, vol. 103, pp. 666–675, 1992.

[162] R. Molina, A. Katsaggelos, J. Mateos, and J. Abad, “Compound gauss-markov random fields for astronomical image restoration,” *Vistas in Astronomy (Special issue on "Vision Modeling and Information Coding")*, vol. 40, pp. 539–546, 1997.

[163] R. Molina, J. Mateos, and J. Abad, “Prior models and the richardson-lucy restoration method,” in *The Restoration of HST Images and Spectra II*, pp. 118–122, The Space Telescope Science Institute, Baltimore, Maryland (USA), 1993.

[164] J. Miskin, *Ensemble learning for independent component analysis*. PhD thesis, Astrophysics Group, Univ. Cambridge, Cambridge, U.K., 2000.

[165] M. Beal, *Variational Algorithms for Approximate Bayesian Inference*. PhD thesis, Gatsby Computational Neuroscience Unit, University College London., 2003.

[166] V. Smidl and A. Quinn, *The Variational Bayes Method in Signal Processing*. New York: Springer Verlag, 2005.

[167] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.

[168] M. I. Jordan, Z. Ghahramani, T. S. Jaakola, and L. K. Saul, *An Introduction to variational methods for graphical models*, ch. in Learn. Graph. Models., pp. 105–162. Cambridge, MA: MIT Press, 1998.

[169] S. Kullback, *Information Theory and Statistics*. New York: Dover, 1959.

[170] C. M. Bishop and M. E. Tipping, “Variational relevance vector machine,” in *Proc. 16th Conf. Uncertainty in Articial Intelligence*, pp. 46–53, 2000.

[171] R. Molina, J. Mateos, and A. K. Katsaggelos, “Blind deconvolution using a variational approach to parameter, image, and blur estimation,” *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3715–3727, 2006.

[172] J. W. Miskin and D. J. C. MacKay, *Ensemble learning for blind image separation and deconvolution*, ch. Advances in Independent Component Analysis. New York: Springer-Verlag, 2000.

[173] S. Kullback and R. A. Leibler, "On information and sufficiency," *Anls. Math. Stat.*, vol. 22, pp. 79–86, 1951.

[174] V. P. Shah, N. H. Younan, and R. King, "Pan-sharpening via the contourlet transform," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. IGARSS 2007*, pp. 310–313, 23–28 2007.

[175] X.-H. Yang and L.-C. Jiao, "Fusion algorithm for remote sensing images based on nonsubsampled contourlet transform," *Acta Automatica Sinica*, vol. 34, no. 2, pp. 274–281, 2008.

[176] M. Vega, J. Mateos, R. Molina, and A. Katsaggelos, "Super resolution of multispectral images using l1 image models and interband correlations," *Journal of Signal Processing Systems. Published online at http://www.springerlink.com/content/nt412p75831n0661/*, 2010.

[177] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.

[178] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Parameter estimation in tv image restoration using variational distribution approximation," *IEEE Transactions on Image Processing*, vol. 17, no. 3, pp. 326–339, 2008.

[179] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag, 1985.

[180] J. Bioucas-Dias, M. Figueiredo, and J. Oliveira., "Totalvariation image deconvolution: A majorization-minimization approach," in *Proc. of the 2006 Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'2006)*, vol. 2, pp. II–861–II–864, 2006.

[181] N. Galatsanos, "A majorization-minimization approach to total variation reconstruction of super-resolved images," in *16th European Signal Processing Conf.*, 2008.

[182] A. K. Katsaggelos, K. T. Lay, and N. P. Galatsanos, "A general framework for frequency domain multi-channel signal processing," *IEEE Transactions on Image Processing*, vol. 2, no. 3, pp. 417–420, 1993.

[183] M. G. K. A. K. Katsaggelos, "Spatially adaptive iterative algorithm for the restoration of astronomical images," *Int. J. Imaging Syst. Technol.*, vol. 6, pp. 305–313, 1995.

[184] K. Amolins, Y. Zhang, and P. Dare, “Wavelet based image fusion techniques: An introduction, review and comparison,” *ISPRS. J. Photogramm.*, vol. 62, pp. 249–263, 2007.

[185] D. D.-Y. Po and M. N. Do, “Directional multiscale modeling of images using the contourlet transform,” *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1610–1620, 2006.

[186] S. D. Babacan, J. Wang, R. Molina, and A. K. Katsaggelos, “Bayesian blind deconvolution from differently exposed image pairs,” *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2874–2888, 2010.

[187] M. Belge and E. L. Miller, “Wavelet domain image restoration using edge preserving prior models,” in *Proc. Int. Conf. Image Processing ICIP 98*, vol. 2, pp. 103–107, 1998.