A survey of classical methods and new trends in pansharpening of multispectral images

Israa Amro[1]*, **Javier Mateos**[1], **Miguel Vega**[2], **Rafael Molina**[1] **and Aggelos K. Katsaggelos**[3] **

[1] Departamento de Ciencias de la Computación e I.A., Universidad de Granada, Spain
[2] Departamento de Lenguajes y Sistemas Informáticos, Universidad de Granada, Spain
[3] Department of Electrical Engineering and Computer Science, Northwestern University, USA

Abstract   There exist a number of satellites on different earth observation platforms, which provide multispectral images together with a panchromatic image, that is, an image containing reflectance data representative of a wide range of bands and wavelengths. Pansharpening, is a pixel level fusion technique used to increase the spatial resolution of the multispectral image while simultaneously preserving its spectral information. In this paper we provide a review of the pansharpening methods proposed in the literature giving a clear classification of them and a description of their main characteristics. Finally 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.

1 Introduction

Nowadays, huge quantities of satellite images are available from many earth observation platforms, such as SPOT [1], Landsat 7 [2], IKONOS [3], Quick-Bird [4] and OrbView [5]. Moreover, due to the growing number of satellite sensors, the acquisition frequency of the same scene is continuously increasing. Remote sensing images are recorded in digital form and then processed by computers to produce image products useful for a wide range of applications.

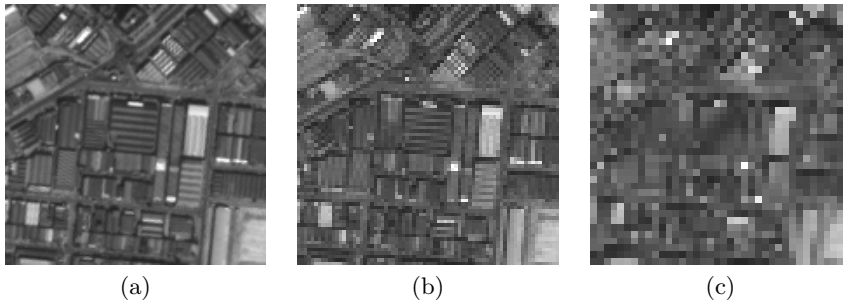* Israa Amro is also with Al-Quds Open University, West Bank (Palestine)

Fig. 1 Images of the same area with different spatial resolutions (a) Spatial resolution = 5m, (b) Spatial resolution = 10m, (c) Spatial resolution = 20m.

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 [6]. The instantaneous field of view (IFOV) is the ground area sensed at a given instant of time. 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 shows three images of the same ground area but with different spatial resolutions. The image at 5m depicted in Fig. 1a was captured by the SPOT 5 satellite while the other two images, at 10m and 20m, are simulated from the first image. As can be observed in these images, the detail information becomes clearer as the spatial resolution increases from 20m to 5m.

Spectral resolution is the electromagnetic bandwidth of the signals captured by the sensor producing a given image. The narrower the spectral bandwidth is, the higher the spectral resolution. If the platform captures images with a few spectral bands, typically 4 to 7, they are referred to as multispectral (MS) data, while if the number of spectral bands is measured in hundreds or thousands, they are referred to as hyperspectral (HS) data [7]. Together with the MS or HS image, satellites usually provide a panchromatic (PAN) image. This is an image that contains reflectance data representative of a wide range of wavelengths from the visible to the thermal infrared, that is, it integrates the chromatic information therefore the name is "pan" chromatic. A PAN image of the visible bands captures a combination of red, green and blue data into a single measure of reflectance.

Remote sensing systems are designed within often competing constraints, among the most important ones being the trade-off between IFOV and signal-to-noise ratio (SNR). Since MS, and to a greater extent HS, sensors have reduced spectral bandwidths compared to PAN sensors, they typically have for a given IFOV a reduced spatial resolution in order to collect

more photons and preserve the image SNR. Many sensors such as SPOT, ETM+, IKONOS, OrbView, and QuickBird have a set of MS bands and a co-registered higher spatial resolution PAN band. With appropriate algorithms it is possible to combine these data and produce MS imagery with higher spatial resolution. This concept is known as multispectral or multisensor merging, fusion, or pansharpening (of the lower-resolution image) [8].

Pansharpening, can consequently be defined as a pixel level fusion technique used to increase the spatial resolution of the MS image [9]. 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 [7]. 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 [10].

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 [9], 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 applications [10], like in the case of fusing different MS SPOT 5 images captured at different times with one PAN IKONOS image [11], 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, up-sampling 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 paper 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. To the best of our knowledge there is no recent paper providing a complete overview of the different pansharpening methods. However, some papers partially address the classification of pansharpening methods, see [12] for instance, or relate already proposed techniques of more global paradigms [13–15].

This paper is organized as follows. In section 2 data pre-processing techniques are described. In section 3 a classification of the pansharpening methods is presented, with a description of the methods related to each category and some examples. In this section we also point out open research problems in each category. In section 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 5 concludes the paper.

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 [8]. 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) [16]. 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 [8,16]), 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 [17] 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 data

resulting 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. For pansharpening, where the accuracy of the input data is crucial, at least radiometric and geometric corrections need to be performed on the satellite data. Radiometric correction rectifies defective columns and missing lines and reduces the non-uniformity of the sensor response among detectors. The geometrical correction deals with systematic effects such as panoramic effect, earth curvature and rotation. Note, however, that even with geometrically registered PAN and MS images, differences might appear between images as described in [10]. These differences include object disappearance or appearance and contrast inversion due to different spectral bands or different times of acquisition. Besides, both sensors do not aim exactly at the same direction and acquisition times are not identical which have an impact on the imaging of fast-moving objects.

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.

2.1 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 [8]. 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 [16]. 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 [18] 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 objects, include Fourier methods, cross-correlation and mutual information methods [19]. Since gray-level values of the images to be matched may be quite different, and taking into account that for any two different image modalities, neither the correlation or the mutual information are maximal when the images are spatially aligned, area-based techniques are not well adapted to the multisensor image registration problem[18]. 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 [19].

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. The up-sampling process may involve interpolation, usually performed via convolution of the image with an interpolation kernel [20]. In order to reduce the computational cost, preferably separable interpolants have been considered [19]. Many interpolants for various applications have been proposed in the literature. A brief discussion of interpolation methods used for image resampling is provided in [19]. Interpolation methods specific to remote sensing, as the one described in [21], have been proposed. In [22] the authors study the application of different interpolation methods to remote sensing imagery. These methods include *nearest neighbor interpolation* which only considers the closest pixel to the interpolated point, thus requiring 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 and *interpolation with smoothing filter* which produces a weighted average of the pixels contained in the area spanned by the filter mask. This process produces images with smooth transitions in gray level while *interpolation with sharpening filter* enhances details that have been blurred and highlights fine details. However sharpening filters produce aliasing in the output image, an undesirable effect which can be avoided applying *interpolation with unsharp masking* that subtracts a blurred version of an image from the image itself. The authors of [22] conclude that only bilinear interpolation, interpolation with smoothing filter and interpolation with unsharp masking have the potential to be used to interpolate remote sensing images. 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.

2.3 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 [16] and those pansharpening methods are prone to spectral distortions. Matching the histograms of the PAN image and MS bands will minimize brightness mismatching during the fusion process, which may help to reduce the spectral distortion in the pansharpened image. Although there are general purpose histogram matching techniques, as the ones described, for instance in [16] and [20], that could be used in remote sensing, specific techniques like the one presented in [23] are expected to provide more appropriate images for the application of pansharpening techniques. The technique in [23] 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

$$Stretched_{PAN}(i, j) = (PAN(i, j) - \mu_{PAN})\frac{\sigma_b}{\sigma_{PAN}} + \mu_b, \qquad (1)$$

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.

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 refer to image fusion. In 1980, Wong *et al.* [24] proposed a technique for the integration of Landsat Multispectral Scanner (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 [25], resolution enhancement [26] or coincident [27] 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.* [28] 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 intensity modulation (IM) methods. Welch *et al.* [29] 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.* [30] to merge Landsat MSS with Return Beam Vidicon (RBV) data and Landsat MSS with Heat Capacity Mapping Mission data. In 1988, Chavez *et al.* [31] used SPOT panchromatic data to "sharpen" Landsat Thematic Mapper (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) [32] and Principal Component Substitution (PCS) [33,34], among others, can be found in [9].

In 1987, Price [35] 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 [36] and Bayesian-based methods [37,38].

More recently, multi-resolution analysis employing the generalized Laplacian pyramid (GLP) [39,40], the discrete wavelet transform [41,42] and the contourlet transform [43–45] 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 [8] classified them into spectral domain, spatial domain, and scale space techniques. Ranchin and Wald [46] 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 [13], but Tu *et al.* [47] found that the PCS, BT, and AWT methods could be also considered as IHS-like image fusion methods. Meanwhile Bretschneider *et al.* [12] classified IHS and PCA methods as transformation based methods, in a classification that also included 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 transforms, had been classified by Sheftigara [48] as component substitution methods. Recently, two comprehensive frameworks that generalize previously proposed fusion methods such as IHS, BT, PCA, HPF or AWT and study the relationships between different methods have been proposed in [14,15].

Although it is not possible to find an 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 Gram-Schmidt (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 inject high frequencies details extracted 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 which includes generalized Laplacian pyramid, wavelet and contourlet methods and any combination of multiresolution analysis with methods from other categories.

Note that although the proposed classification defines five categories, as we have already mentioned, some methods can be classified in several categories and, so, the limits of each category are not sharp and there are many relations among them. The relations will be explained when the categories are described.

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 images 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 1. This algorithm has been generalized by Tu *et al.* [47], where the authors also prove that the forward and backward transforms are not needed and steps 2-5 of Alg. 1 can be summarized as finding a new component C_l and adding the difference between the PAN and this new component to each upsampled MS image band. This framework has been further extended by Wang *et al.* [14] and Aiazzi *et al.* [15] in the so

called General Image Fusion (GIF) and extended GIF (EGIF) protocol, respectively.

Algorithm 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 [48,49], each of them involving a different transformation of the MS image. CS techniques are attractive because they are fast and easy to implement and allow users expectations to be fulfilled most of the time, since they provide pansharpened images with good visual/geometrical quality in most cases [50]. However, the results obtained by these methods highly depend on the correlation between the bands, and since the same transform is applied to the whole image, it does not take into account local dissimilarities between PAN and MS images [10, 51].

A single type of transform does not always obtain the optimal component required for substitution and it would be difficult to choose the appropriate spectral transformation method for diverse 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* [52–54].

The ***Intensity-Hue-Saturation (IHS)*** pansharpening method [31, 55] 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 transform 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 [14]. The ability of IHS transform to separate effectively spatial information (band I) and spectral information (bands H and S) [20] makes it very applicable in pansharpening. There are different models of IHS transform, differing in the method used to compute the intensity value. Smith's hexacone and triangular models are two of the most widely used ones [7]. An example of pansharpened image using IHS method is shown in Figure. 2(b).

The major limitation of this technique is that only three bands are involved. Tu *et al.* [47] proposed a Generalized IHS transform which surpass the dimensional limitation. In any case, since the spectral response of I, as

synthesized from the MS bands, does not generally match the radiometry of the histogram-matched PAN [50], 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, Tu *et al.* [56] proposed a new adaptive IHS method in which the intensity band approximates the PAN image for IKONOS images as closely as possible. This adaptive IHS has been extended by Rahmani *et al.* [52] to deal with any kind of image by determining the coefficients α_i that best approximate

$$PAN = \sum_i \alpha_i MS_i, \qquad (2)$$

subject to the physical constraint of non-negativity of the coefficients α_i. Note that, although this method reduces spectral distortion, local dissimilarities between MS and PAN images might remain [10].

Another method in the CS family is ***principal component substitution (PCS)*** which relies on the principal component analysis (PCA) mathematical transformation. The PCA, also known as the Karhunen-Loéve transform or the Hotelling transform, is widely used in signal processing, statistics, and many other areas. This transformation generates a new set of rotated axes, in which the new image spectral 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 ones. 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 [20]. 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 [42,33]. 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 existing among the MS bands [33], resulting in the same problems that IHS. To overcome this drawback, Shah *et al.* [53] proposed a new Adaptive PCA-based pansharpening method that determines, using cross-correlation, the appropriate PC component to be substituted by the PAN image. By replacing this PC component by the high spatial resolution PAN component, adaptive PCA method will produce better results than traditional ones [53].

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 [57]. The GS transformation, as described in [58], 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 [57]. 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 [57] 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.* [54] 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. GS is more general than PCA which can be understood as a particular case of GS in which LR PAN is the first principal component [15].

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

$$HRMS_b^h(i,j) = \frac{MS_b^h(i,j)PAN^h(i,j)}{\sum_b MS_b^h(i,j)}, \tag{3}$$

Algorithm 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. (3).*

---

where $b = 1, 2, \cdots, B$ and B is the number of MS bands. The process flow diagram of RSC sharpening techniques is shown in Algorithm 2. This family does not tell what to do when MS_b^l lies outside the spectral range of PAN^h. In Eq. (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 [46]. These methods include the Brovey transform (BT) [32], the P+XS [59, 60] and the Intensity modulation (IM) method [61].

The Brovey Transform (BT), named after its author, is a simple method to merge data from different sensors based on the chromaticity transform [32], with the limitation that only three bands are involved [42, 14]. A pansharpened image using the BT method is shown in Fig. 2(c).

The Brovey transform provides excellent contrast in the image domain but greatly distorts the spectral characteristics [62]. The Brovey sharpened image is not suitable for pixel-based classification as the pixel values are changed drastically [7]. A variation of the BT method subtracts the intensity of the MS image from the PAN image before applying Eq. (3) [14]. 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.* [24] in 1980 for integrating Landsat MSS and Seasat SAR images. Later, this method was used by Cliche *et al.* [28] 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. (3), by replacing the sum of all MS bands, by the intensity component of the IHS transformation [6]. 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 [63]. In the literature, different versions based on the IM concept, have been used [28, 6, 63].

The relations between RSC and CS families have been deeply studied in [47, 14] where these families are considered as a particular case of the GIHS and GIF protocols, respectively. The authors also found that RSC methods are closely CS, with the difference, as already commented, that the contribution of the PAN varies locally.
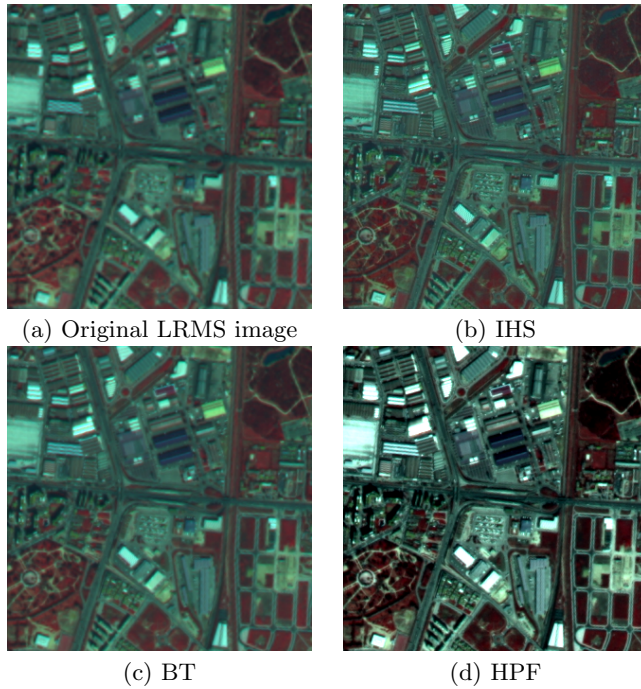
(a) Original LRMS image   (b) IHS

(c) BT   (d) HPF

Fig. 2 Results of some classical pansharpening methods using SPOT 5 images.

3.3 High-frequency injection family

The high-frequency injection family methods were first proposed by Schowengerdt [64], 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 [65] 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"[31]. 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,

$$image(i,j) = LP(i,j) + HP(i,j), \qquad (4)$$

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. (4)) [8].

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 [66]. 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.* [34] 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 [67] 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,

$$filtered_{PAN} = h_0 * PAN, \qquad (5)$$

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 [31,68]. There are many different filters that can be used, like Box filter, Gaussian, Laplacian, and so on. Recently, the use of the modulation transfer function (MTF) of the sensor as the low-pass filter has been proposed in [69]. The MTF is the amplitude spectrum of the system point spread function (PSF) [70]. In [69], 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 [14]. 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* [46].

In the **High Pass Modulation (HPM)**, also known as **High Frequency Modulation (HFM)** algorithm [8], the 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 [14]. 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,

$$k_b(i,j) = \frac{LRMS_b(i,j)}{filtered_{PAN}(i,j)},\tag{6}$$

where $filtered_{PAN}$ is a low-pass filtered version of PAN image (see Eq. (5)) [8]. According to [14] (where HFI has also been formulated into the GIF framework and relations with CS, RSC and some multiresolution family methods are explored) 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 3. Also a pansharpened image using the HPM method is shown in Fig. 2(d). Note that the HFI methods are closely related, as we will see later, to the multiresolution family. The main differences are the types of filter used, that a single level of decomposition is applied to the images, and the different origins of the approaches.

Algorithm 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. (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 $k_b(i,j)$ in Eq. (6) in the case of HPM).

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 [35] 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. Later, the method was improved by Price [71] by computing the local statistics of the images and by Park *et al.* [36] in the so called spatially adaptive algorithm.

Price's method [71] 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

$$z_b - \hat{y_b} = \hat{a}(x - \hat{x}), \tag{7}$$

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 low 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 [71] is shown in Fig. 4 (a).

The ***spatially adaptive algorithm*** [36] starts from Price's method [71], 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 [36].

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 [37], a number of pansharpening methods have been proposed using the Bayesian framework (see [72,73] 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,

$$p(z|y, x) = \frac{p(y, x|z)p(z)}{p(y, x)}, \tag{8}$$

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 [73], 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.* [73] assumed 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.* [74]. In [37] 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 [75] 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

$$p(z) = \prod_{b=1}^{B} p(z_b) \propto \prod_{b=1}^{B} exp\left\{ -\frac{1}{2}\alpha_b \|Cz_b\|^2 \right\} \tag{9}$$

where C denotes the Laplacian operator and $1/\alpha_b$ is the variance of the Gaussian distribution of z_b, $b = 1, \ldots, B$, 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 [38], Total Variation (TV) [76], Markov Random Fields (MRF) [77] based models and Stochastic Mixing Models (SMM) [78]. Note that the described models do not take into account the correlations between the MS bands. In [79] the authors propose a TV prior model to take into account spatial pixel relationships and a quadratic 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

$$y = g_s(z) + n_s, \tag{10}$$

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

$$x = g_p(z) + n_p, \tag{11}$$

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

$$p(y, x|z) = p(y|z)p(x|z) \tag{12}$$

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 $et\ al.$ [73] 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. (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 $y = z + n_s$ [73] 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 $y = Hz + n_s$ 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

$$p(y|z) \propto \exp\left\{-\frac{1}{2}\beta\|y - Hz\|^2\right\}. \tag{13}$$

This model has been extensively used (see [77, 78, 80]) and it is the base for the so called super-resolution based methods [81] as the ones described, for instance, in [38, 76]. The degradation model in [37] can be also written in this way. A pansharpened image using the super-resolution method proposed in [76] is shown in Fig. 4 (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 $et\ al.$ [82] (and also [38, 76, 77] 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 Fig. 3 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,

$$p(x|z) \propto \exp\left\{-\frac{1}{2}((x - g_p(z))^t C_p^{-1}(x - g_p(z)))\right\}. \tag{14}$$

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
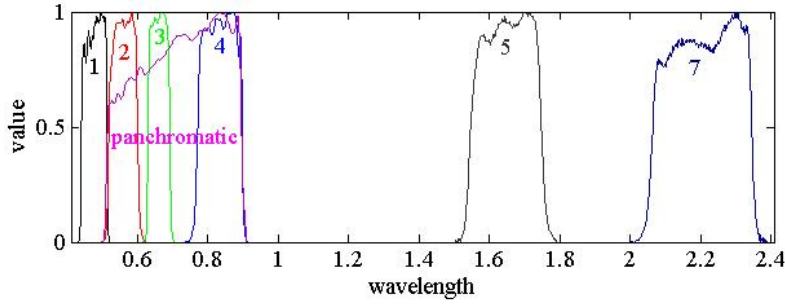
Fig. 3 *Landsat 7 ETM+ band spectral response*

is sharpened by a higher resolution MS or PAN image. In this context, Eismann and Hardie [80, 78] and other authors later (see for instance [83]) 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 [82].

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) [73], linear minimum mean square error (LMMSE) [83], Maximum a Posteriori (MAP) [74], the Variational approach [38, 76] or Simulated Annealing [77] are some of the techniques used. Bayesian methods usually end up with an estimation of the HRMS image z that is a convex combination of the LRMS image, upsampled to the size of the HRMS image by inverting the degradation model, the PAN image, and the prior knowledge about the HRMS image. The combination factors usually are pixel adaptive and related to the spectral characteristics of the images.

Although all approaches already mentioned use the hypothesis of Gaussian additive noise for mathematical convenience, in practice, remote sensing imagery noise shows non-Gaussian characteristics [84]. In some applications, such as astronomical image restoration, Poisson noise is usually used, or a shaping filter [85] may be used in order to transform non-Gaussian noise into Gaussian. Recently, Niu *et al.* [84] proposed the use of a mixture of Gaussian (MoG) noise for multisensor fusion problems.
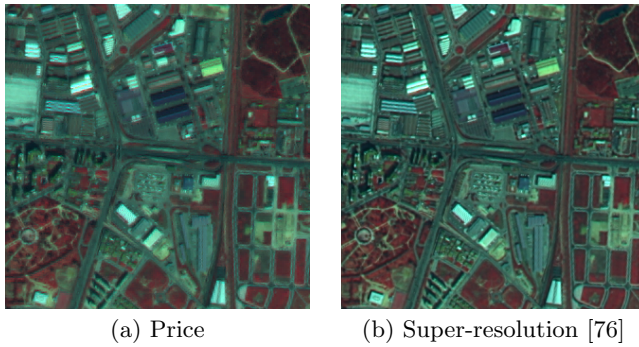
<div align="center">(a) Price   (b) Super-resolution [76]</div>

Fig. 4 Results of some statistical pansharpening methods using SPOT 5 images.

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 others 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 [8]. This increasingly important category of scale-space filters utilize multiscale decomposition techniques such as Laplacian pyramids [86], wavelet transform [41], contourlet transform [43] and curvelets transform. These techniques are used in pansharpening to decompose MS and PAN images in 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 [87]. This is the idea behind the methods based on the successful ARSIS (from French "Amèlioration de la Résolution by Structure Injection") concept [46].

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 [86], 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 decomposes the image into nearly disjoint bandpass channels in the spatial frequency domain, without losing the spatial connectivity of its edges [88]. Figure 5 shows the concept of GP and its relation to LP. The Generalized Laplacian Pyramid (GLP) is an extension of the LP where a scale factor different of two is used [89]. 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 restora-
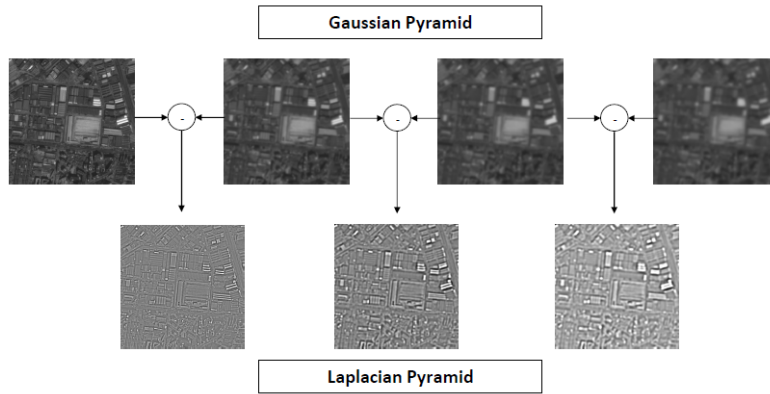
Fig. 5 Laplacian Pyramid created from Gaussian pyramid by subtraction

---

Algorithm 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.

---

tion 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 4, where different injection methods can be used with GLP [40, 90]. 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 an approximation of the PAN image at the resolution of the MS bands. Details are only injected 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 [40]. Pansharpened images using wavelet/contourlet-based methods are shown in Fig. 6.

The Ranchin-Wald-Mangolini (RWM) injection model [40], unlike the SDM and CBD models, is calculated on bandpass details instead of approximations. RWM models the MS details as an space and spectral-varying linear combination of the PAN image coefficients.

Another popular category of multiresolution pansharpening methods is the one based on **Wavelet and Contourlet**. Wavelet provide a framework for the decomposition of images into a hierarchy with a decreasing degree of resolution, separating detailed spatial information between successive levels [91]. 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 [42, 46, 92, 93] and the "a' trous" [94, 13, 95] 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 [91]. 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. It is important to note that undecimated, shift invariant decompositions, and specifically "a' trous" wavelets, where subband and original image pixels corresponds to the same locations, produce fewer artifacts and better preserve the linear continuity of features that do not have a horizontal or vertical orientation [96], and hence are better suited for image fusion.

Contourlets provide a new representation system for image analysis [43]. The contourlet transform is so called because of its ability to capture and link the discontinuity points into linear structures (contours). The two-stage process used to derive the contourlet coefficients involves a multiscale transform and a local directional transform. First, a multiscale LP which detect discontinuities is applied. Then, a 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 orientations. This flexibility of having different numbers of directions at each scale makes contourlets different from other available multiscale and directional image representations [53]. Similarly to wavelets, contourlets also have different implementations of the subsampled and non-subsampled [43] 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 follow the process in Algorithm 5. In the wavelet/contourlet-based approach, the MS and PAN images need to be decomposed multiple times in step 1 of Algorithm 5.
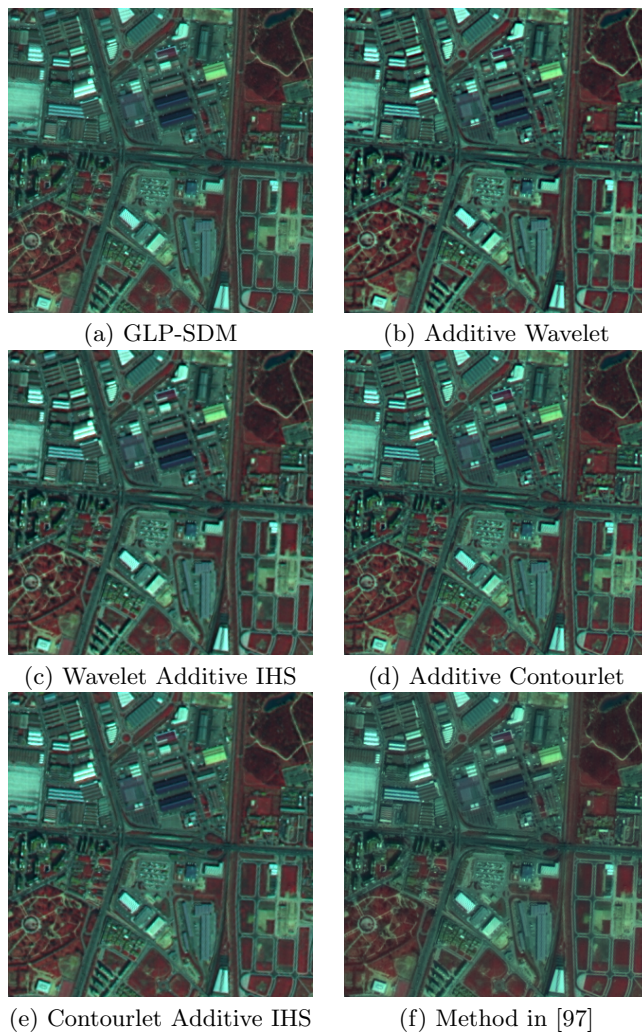
(a) GLP-SDM

(b) Additive Wavelet

(c) Wavelet Additive IHS

(d) Additive Contourlet

(e) Contourlet Additive IHS

(f) Method in [97]

Fig. 6 Results of some multiresolution pansharpening methods using SPOT 5 images.

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.* [98] attempt in their work to determine the optimal number of decomposition levels for the wavelet-based pansharpening, producing the optimal spatial and spectral quality.

Algorithm 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.

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 dependent on 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 [91] /contourlet [44] transform in steps 1 and 3 in Algorithm 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 [94] for wavelet and [99] 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** [91] and **contourlet** [100,101] **additive IHS** pansharpening. If the IHS transform is followed by the substitutive wavelet method, the **wavelet substitutive IHS** [102] pansharpening method is obtained.

Similarly to the IHS wavelet/contourlet methods, the **PCA wavelet [68,91] /contourlet [53]** 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 [53].

The **WiSpeR** [103] 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 ob-

served object. A new contourlet panshapening method named ***CiSper*** was proposed in [45] 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 [104, 97]. Pansharpened images using wavelet and contourlet-based methods are shown in Fig. 6 (b-f).

Some authors [41, 42] 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 that wavelet transform fusion methods easily control this trade-off. The trade-off idea, however, is just a convenient simplification, as discussed in [10], 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.

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.* [67] 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 [105]. The consistency property is the same as the first property and the synthesis property combines the second and third properties defined by [67]. 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.* [67]

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 [106]. 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.

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 [105]: (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.

4.2 Quantitative Analysis

A set of measures has been proposed to quantitatively assess 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 for the different bands 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 [107].

2. *Relative-shift mean (RM):* The RM [108] of each band of the fused image helps to visualize the change in the histogram of fused image and is defined in [108] 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 [62]. However, CC is insensitive to a constant gain and bias between two images and does not allow for subtle discrimination of possible pansharpening artifacts [14]. 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 [67]. RMSE is a very good indicator of the spectral quality when it considered along homogeneous regions in the image [108]. RMSE should be as close to 0 as possible.

5. *Structure Similarity Index (SSIM):* SSIM [109] 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 [110], 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 [111], is a global quality index sensitive to mean shifting and dynamic range change [112]. 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 [109,103] 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 [113] 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 [42] 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, a Laplacian filter is applied to the images under comparison. 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 [114,10] and the HR MS image has to be used, as done by Otazu *et al.* in [103]. 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 [98], 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 [98]. Thus, a Sobel edge operator is applied on the image in order to detect 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 [114].

Additionally, some spectral quality measures have been adapted to spatial quality assessment. Pradhan *et al.* [98] suggested the use of structural information in SSIM measure between panchromatic and pansharpened images as a spatial quality measure. Lillo-Saavedra *et al.* [115] 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.

Although an exhaustive comparison of all the aforementioned pansharpening methods is out of the scope of this paper, for the sake of reference we have included, in Table 1, the figures of merit for some of the pansharpened images from the observed multispectral image shown in Fig. 2a already presented in this paper. The two best values for each measure have been highlighted in Table 1. From the obtained results it is clear that BT, HPF, and Price methods, depicted in Figs. 2(c)–(d) and 4(a), respectively, suffer the highest spectral distortions, with the lowest SSIM and MSSIM values and the highest ERGAS value. In this case, the IHS method (Fig. 2(b)) and IHS-Wavelet (Fig. 6(c)) produce good numerical results but note that these results have been obtained considering only the first 3 bands, the ones involved in the IHS transform. Methods based in multiresolution approaches, GLP (Fig. 6(a)), Additive Wavelets (Fig. 6(b)), and the method described in [97] (Fig. 6(f)), provide the best results with lower spectral distortion and

Method	Band	SCC	SSIM	MSSIM	ERGAS
IHS[14]	B1	0.97	**0.91**	0.89	5.02
	B2	0.96	0.87		
	B3	0.97	0.90		
BT[32]	B1	0.77	0.79	0.76	6.40
	B2	0.78	0.72		
	B3	0.79	0.78		
HPF[31]	B1	**0.99**	0.82	0.85	10.04
	B2	**0.98**	0.90		
	B3	**0.99**	0.87		
	B4	**0.96**	0.82		
Price[71]	B1	0.60	0.73	0.71	6.59
	B2	0.64	0.66		
	B3	0.62	0.71		
	B4	0.58	0.74		
GLP[40]	B1	**0.98**	**0.92**	0.89	**3.88**
	B2	**0.99**	**0.94**		
	B3	**0.98**	**0.93**		
	B4	**0.98**	0.76		
Add-Wavelet [91]	B1	0.96	**0.91**	0.89	4.00
	B2	0.97	**0.93**		
	B3	0.97	**0.91**		
	B4	**0.96**	0.80		
IHS-Wavelet [91]	B1	0.96	0.90	**0.90**	4.21
	B2	0.95	0.88		
	B3	0.97	**0.91**		
Method in [97]	B1	**0.98**	0.90	**0.91**	**2.76**
	B2	0.96	**0.93**		
	B3	0.94	**0.93**		
	B4	0.94	0.90		

Table 1 Numerical results with the presented pansharpened methods utilizing the multispectral image in Fig. 2a.

higher SCC values. These results are consistent with the ones reported in [116] where the multiresolution methods perform better than other methods. Note that the lower values of SSIM for the GLP and AW methods are due to the ratio between band 4, which has a spatial resolution of 20 m per pixel, and the PAN image, that covers 5 m per pixel, while the three first bands have a spatial resolution of 10 m per pixel.

For a comparison between some of the reported methods the reader is referred, for instance, to the results of the 2006 GRS-S Data-Fusion Contest [116] where a set of 8 methods, mainly CS and MRA based, were tested on a common set of images, or [10] where the authors discuss from a theoretical point of view the strengths and weakness of CS, RSC and ARSIS concept implementations (which includes HFI and multi-resolution families), and perform a comparison based on the fusion rule and the effect of

this rule on the spectral and spatial distortion. Another comparative analysis was developed in [14], where a general image fusion (GIF) framework for pansharpening IKONOS images is proposed and the performance of several image fusion method is analyzed based on the way they compute, from the LR MS image, an LR approximation of the PAN image, and how the modulation coefficients for the detail information are defined. This work has been extended in [15] to consider also context adaptive methods in the so called extended GIF protocol. Recently, a comparison of pansharpening methods was carried out in [117] based on their performance on automatic classification and visual interpretation applications.

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 [118] comprises two indices, one pertaining to spectral distortion and the other to spatial distortion. As proposed in [118] and [119] the two distortions may be combined to yield a unique global quality measure. While the QNR measure proposed in [118] is based on the UIQI index, the one proposed in [119] is based on the measure of the Mutual Information (MI) between the different images. The spectral distortion index defined in [118] can be derived from the difference of inter-band UIQI values calculated among all the fused MS bands and the resampled LR MS bands. The spatial distortion index defined in [118] is based on differences between the UIQI of each band of the fused image and the PAN image and the UIQI of each band of the LR MS 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 [119] 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.* [69] 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 fused image. This filtered image, once it has been decimated, will give a degraded LR MS image. For comparing the degraded and original low-resolution MS images, the Q4 index [113] 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.

5 Conclusion

In this paper we have provided a complete overview of the different methods proposed in the literature to tackle the pansharpening problem and classified them into different categories according to the main technique they use. As previously described in sections 3.1 and 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 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 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 in-

corporating the knowledge available about the HR MS image. Those methods allow, as explained in section 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 3.5, we can see that the generalized Laplacian pyramid and redundant shift-invariant 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 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 4.3 set the trend for new measures specific for pansharpening that have to be considered.

References

1. *SPOT Web Page.* [Online]. Available: www.spotimage.fr
2. *Landsat 7 Web Page.* [Online]. Available: http://geo.arc.nasa.gov/sge/landsat/
3. *IKONOS Web Page.* [Online]. Available: http://www.geoeye.com/CorpSite/products/imagery-sources/
4. *QuickBird Web Page.* [Online]. Available: http://www.digitalglobe.com/index.php/85/QuickBird
5. *OrbView Web Page.* [Online]. Available: http://www.tec.army.mil/tio/ORBVIEW.htm
6. 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.
7. V. Vijayaraj, "A quantitative analysis of pansharpened images," Master's thesis, Mississippi State Univ., 2004.
8. R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*, 3rd ed. Orlando, FL: Academic, 1997.
9. 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.

10. C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, "Synthesis of multi-spectral 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.

11. 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.*, ser. Part B4, vol. XXXVII, 2008, pp. 499–505.

12. T. Bretschneider and O. Kao, "Image fusion in remote sensing," Technical University of Clausthal, Germany.

13. 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.

14. 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.

15. B. Aiazzi, S. Baronti, F. Lotti, and M. Selva, "A comparison between global and context-adaptive pansharpening of multispectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 2, pp. 302 –306, 2009.

16. J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, 4th ed.   Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2005.

17. 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.

18. Y. Yang and X. Gao, "Remote sensing image registration via active contour model," *Int. J. Electron. Commun.*, vol. 63, pp. 227–234, 2009.

19. B. Zitova and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.

20. R. C. Gonzalez and R. E. Woods, *Digital image processing*, 3rd ed.   Prentice Hall, 2008.

21. S. Takehana, M. Kashimura, and S. Ozawa, "Predictive interpolation of remote sensing images using vector quantization," in *Communications, Computers and Signal Processing, 1993., IEEE Pacific Rim Conference on*, vol. 1, 1993, pp. 51–54.

22. 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.

23. 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.

24. 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, pp. 161–164.

25. P. Rebillard and P. T. Nguyen, "An exploration of co-registered SIR-A, SEASAT and Landsat images," in *Proceeidngs International Symposium on RS of Environment, RS for Exploration Geology*, 1982, pp. 109–118.

26. R. Simard, "Improved spatial and altimetric information from SPOT composite imagery," in *Proceedings International Symposium of Photogrammetry and Remote Sensing)*, 1982, pp. 434 – 440.

27. E. P. Crist, "Comparison of coincident Landsat-4 MSS and TM data over an agricultural region," in *Proceedings of the 50th Annual Meeting ASP-ACSM Symposium, ASPRS*, 1984, pp. 508–517.

28. 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, no. 3, pp. 311–316, 1985.

29. R. Welch and M. Ehlers, "Merging multiresolution SPOT HRV and landsat TM data," *Photogramm. Eng. Remote Sens*, vol. 53, no. 3, pp. 301–303, 1987.

30. R. Haydn, G. W. Dalke, and J. Henkel, "Application of the IHS color transform to the processing of multisensor data and image enhancement," in *International Symposium on Remote Sensing of Arid and Semi-Arid Lands*, 1982, pp. 599–616.

31. 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.

32. 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.

33. 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.

34. P. Chavez, S. Sides, and J. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 295–303, 1991.

35. J. C. Price, "Combining panchromatic and multispectral imagery from dual resolution satellite instruments," *Remote Sensing Of Environment*, vol. 21, pp. 119–128, 1987.

36. J. Park and M. Kang, "Spatially adaptive multi-resolution multispectral image fusion," *Int. J. Remote Sensing*, vol. 25, no. 23, pp. 5491– 5508, 2004.

37. N. D. A. Mascarenhas, G. J. F. Banon, and A. L. B. Candeias, "Multispectral image data fusion under a Bayesian approach," *Int. J. Remote Sensing*, vol. 17, pp. 1457–1471, 1996.

38. M. Vega, J. Mateos, R. Molina, and A. Katsaggelos, "Super-resolution of multispectral images," *The Computer Journal*, vol. 1, pp. 1–15, 2008.

39. B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis," *IEEE Trans. Geosc. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, 2002.

40. A. Garzelli and F. Nencini, "Interband structure modeling for pansharpening of very high-resolution multispectral images," *Information Fusion*, vol. 6, pp. 213–224, 2005.

41. S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions On Pattern Analysis And Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

42. 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.

43. 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.

44. 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.

45. 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.

46. 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.

47. 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.

48. 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, no. 5, pp. 561–567, 1992.

49. 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.

50. B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS + Pan data," *IEEE Trans. Geosc. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, 2007.

51. 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.

52. 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.

53. 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. Geosc. Remote Sens.*, vol. 46, no. 5, pp. 1323–1335, 2008.

54. 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*, 2006, pp. 3806–3809.

55. 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.

56. T.-M. Tu, P. S. Huang, C.-L. Hung, and C.-P. Chang, "A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery," *IEEE Geoscience And Remote Sensing Letters*, vol. 1, no. 4, pp. 309–312, 2004.

57. C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pansharpening," U.S. Patent 6 011 875, 2000.

58. R. W. Farebrother, "Gram-schmidt regression," *Applied Statistics*, vol. 23, pp. 470–476, 1974.

59. *SPOT Users Handbook*, Centre National Etude Spatiale (CNES) and SPOT Image, Toulouse, France, 1988.

60. 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.

61. 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, no. 3, pp. 311–316, 1985.

62. 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.

63. 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*, 2004, pp. 637–643.

64. R. A. Schowengerdt, "Reconstruction of multispatial, multispectral image data using spatial frequency contents," *Photogrammetric Engineering & Remote Sensing*, vol. 46, no. 10, pp. 1325–1334, 1980.

65. P. S. Chavez, Jr., "Digital merging of Landsat TM and digitized NHAP data for 1:24,000 scale image mapping," *Photogramm. Eng. Remote Sens.*, vol. 52, no. 10, pp. 1637–1646, 1986.

66. V. J. D. Tsai, "Frequency-based fusion of multiresolution images," in *2003 IEEE International Geoscience and Remote Sensing Symposium IGARSS 2003*, vol. 6, 2003, pp. 3665–3667.

67. 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.

68. 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.

69. 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.

70. 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*, 2006, pp. 1–7.

71. J. C. Price, "Combining multispectral data of differing spatial resolution," *IEEE Trans. Geosc. Remote Sens.*, vol. 37, no. 3, pp. 1199–1203, May 1999.

72. O. Punska, "Bayesian approaches to multi-sensor data fusion," Master's thesis, University of Cambridge, 1999.

73. 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.

74. 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.

75. 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.

76. 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.*, 2008, pp. 408–415.

77. H. G. Kitaw, "Image pan-sharpening with Markov random field and simulated annealing," Ph.D. dissertation, International Institute for Geo-information Science and Earth Observation, NL, 2007.

78. 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, no. 3, pp. 455–465, 2005.

79. 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*, 2009, pp. 1–6.

80. 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, no. 9, pp. 1924–1933, Sep 2004.

81. A. Katsaggelos, R. Molina, and J. Mateos, *Super Resolution Of Images And Video*. Synthesis Lectures on Image, Video, and Multimedia Processing, Morgan & Claypool, 2007.

82. 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)*, 2005, pp. 280–283.

83. 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.

84. 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.

85. L. Brandenburg and H. Meadows, "Shaping filter representation of nonstationary colored noise," *IEEE Transactions on Information Theory*, vol. 17, pp. 26–31, 1971.

86. P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions On Communications*, vol. COM-3l, no. 4, pp. 532–540, 1983.

87. 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.

88. 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, 2001, pp. 852–854.

89. M. G. Kim, I. Dinstein, and L. Shaw, "A prototype filter design approach to pyramid generation," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 15, no. 12, pp. 1233–1240, 1993.

90. 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*, 2004, p. 20.1.

91. 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.

92. 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.

93. D. A. Yocky, "Image merging and data fusion by means of the discrete two-dimensional wavelet transform," *Optical Society of America*, vol. 12, no. 9, pp. 1834–1841, Septemper 1995.

94. 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.

95. 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.

96. 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.

97. I. Amro, J. Mateos, and M. Vega, "General contourlet pansharpening method using Bayesian inference," in *2010 European Signal Processing Conference (EUSIPCO-2010)*, 2010, pp. 294–298.

98. 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 Trans. Geosc. Remote Sens.*, vol. 44, no. 12, pp. 3674–3686, 2006.

99. 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*, 2009, pp. 77–80.

100. 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.

101. 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.

102. 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*, 2004, pp. 564–568.

103. 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.

104. Y. Zhang, S. D. Backer, and P. Scheunders, "Bayesian fusion of multispectral and hyperspectral image in wavelet domain," *IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, vol. 5, pp. 69–72, 2008.

105. V. Meenakshisundaram, "Quality assessment of IKONOS and Quickbird fused images for urban mapping," Master's thesis, University of Calgary, 2005.

106. L. Wald, *Data Fusion Definitions and Architectures: Fusion of Images of Different Spatial Resolutions.* Les Presses de l'Ecole des Mines, Paris, 2002.

107. 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.

108. V. Vijayaraj, N. H. Younan, and C. G. O'Hara, "Quantitative analysis of pansharpened images," *Optical Engineering*, vol. 45, no. 4, p. 046202, 2006.

109. 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.

110. Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, 2002.

111. L. Wald, "Quality of high resolution synthesized images: Is there a simple criterion?" *Proc. Int. Conf. Fusion of Earth Data*, vol. 1, pp. 99 –105, 2000.

112. Q. Du, N. H. Younan, R. King, and V. P. Shah, "On the performance evaluation of pan-sharpening techniques," *IEEE Trans. Geosc. Remote Sens.L*, vol. 4, no. 4, pp. 518–522, 2007.

113. 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.

114. 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"*, 2007, pp. 101–111.

115. 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.

116. L. Alparone, L. Wald, J. Chanussot, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosc. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, 2007.

117. Y. Jinghui, Z. Jixian, L. Haitao, S. Yushan, and P. Pengxian, "Pixel level fusion methods for remote sensing images: A current review," in *ISPRS TC VII Symposium 100 Years ISPRS*, W. Wagner and B. Szkely, Eds., 2010, pp. 680–686.

118. 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.

119. L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, and F. Nencini, "A new method for MS + Pan image fusion assessment without reference," in *Proc. IEEE Int. Conf. Geoscience and Remote Sensing Symp. IGARSS 2006*, 2006, pp. 3802–3805.