© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

A Critical Comparison Among Pansharpening Algorithms

Gemine Vivone, Luciano Alparone, Jocelyn Chanussot, *Fellow, IEEE,* Mauro Dalla Mura, *Member, IEEE,* Andrea Garzelli, *Senior Member, IEEE,* Giorgio A. Licciardi, *Member, IEEE,* Rocco Restaino, *Member, IEEE,* and Lucien Wald

***Abstract*—Pansharpening aims at fusing a multispectral and a panchromatic image featuring the result of the processing with the spectral resolution of the former and the spatial resolution of the latter. In the last decades many algorithms addressing this task have been presented in the literature. However, the lack of universally recognized evaluation criteria, available image data sets for benchmarking, and standardized implementations of the algorithms makes a thorough evaluation and comparison of the different pansharpening techniques difficult to achieve. In this work, the authors attempt to fill this gap by providing a critical description and extensive comparisons of some of the main state of the art pansharpening methods. In greater details, several pansharpening algorithms belonging to the component substitution or multiresolution analysis families are considered. Such techniques are evaluated through the two main protocols for the assessment of pansharpening results, i.e., based on the full and reduced resolution validations. Five data sets acquired by different satellites allow for a detailed comparison of the algorithms, characterization of their performances with respect to the different instruments, and consistency of the two validation procedures. In addition, the implementation of all the pansharpening techniques considered in this paper and the framework used for running the simulations, comprising the two validation procedures and the main assessment indexes, are collected in a MatLab toolbox that is made available to the community.**

***Index Terms*—Benchmarking, component substitution, multiresolution analysis, multispectral pansharpening, quality assessment, very high resolution optical images.**

Manuscript received March 28, 2014; revised August 21, 2014; accepted September 18, 2014. A preliminary version of this paper was reported in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Quebec City, Canada, 20–25 July 2014.

G. Vivone and R. Restaino are with the Department of Information Engineering, Electrical Engineering and Applied Mathematics, University of Salerno, 84084 Fisciano, Italy (e-mail: gvivone@unisa.it; restaino@unisa.it).

L. Alparone is with the Department of Information Engineering, University of Florence, 50139 Florence, Italy (e-mail: luciano.alparone@unifi.it).

J. Chanussot is with GIPSA-Lab, Grenoble Institute of Technology, 38402 Saint Martin d'Hères, France, and also with the Faculty of Electrical and Computer Engineering, University of Iceland, 107 Reykjavík, Iceland (e-mail: jocelyn.chanussot@gipsa-lab.grenoble-inp.fr).

M. Dalla Mura and G. A. Licciardi are with GIPSA-Lab, Grenoble Institute of Technology, 38402 Saint Martin d'Hères, France (e-mail: {mauro.dalla-mura; giorgio-antonino.licciardi}@gipsa-lab.grenoble-inp.fr).

A. Garzelli is with the Department of Information Engineering and Mathematical Sciences, University of Siena, 53100 Siena, Italy (e-mail: andrea.garzelli@unisi.it).

L. Wald is with the Center Observation, Impacts, Energy, MINES ParisTech, 06904 Sophia Antipolis, France (e-mail: lucien.wald@minesparistech.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2014.2361734

I. Introduction

PANSHARPENING refers to the fusion of a *panchromatic* (PAN) and a *multispectral* (MS) image simultaneously acquired over the same area. This can be seen as a particular problem of data fusion since one would aim at combining the spatial details resolved by the PAN (but not present in the MS) and the several spectral bands of the MS image (against the single band of the PAN) in a unique product. With respect to the general problem of multisensor fusion, pansharpening may not require the challenging phase of spatial coregistration, since typically images are simultaneously captured, being the sensors acquiring the PAN and the MS both mounted on the same platform [1]. Nowadays, PAN and MS images can be obtained in bundle by several commercial optical satellites such as IKONOS, Geo-Eye, OrbView, Landsat, SPOT, QuickBird, WorldView, and Pléiades. The spatial resolution is even below half a meter for the PAN (for the commercial satellite product with the highest spatial resolution) and the spectral resolution can be up to eight bands captured in the visible and near infrared wavelengths for the MS product. The fusion of the PAN and MS images constitutes the sole possibility for achieving images with the highest resolutions in both the spatial and spectral domains. In fact, physical constraints preclude this goal from being achieved by using a single sensor. The demand for pansharpened data is continuously growing, thanks to the increasing availability of commercial products using high resolution images, as for example Google Earth and Bing Maps. Furthermore, pansharpening constitutes an important preliminary step for enhancing images for many remote sensing tasks, such as change detection [2], object recognition [3], visual image analysis and scene interpretation [4].

The interest of the community in pansharpening is evident by reviewing the recent technical literature. Detailed surveys of pansharpening algorithms can be found in [1] and in [5], [6], [7]. However, the authors believe that the comparison of the existing pansharpening methods has not been sufficiently addressed. The contest launched by the Data Fusion Committee of the IEEE Geoscience and Remote Sensing Society in 2006 [8] has made a first step tackling this issue, since it performed an explicit evaluation of several methods applied to the same data sets, assessed with the same validation procedure and using the same performance metrics.

This work moves in the same direction. It aspires at contributing to this goal in two ways: by i) performing a

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

critical comparison and analysis of many state of the art pansharpening techniques, and ii) providing a framework that gathers the implementation of the algorithms considered in this paper for easing their benchmarking, evaluation and aiming at reproducible researches.

The focus is on some widely used algorithms following the two main approaches, spectral and spatial, that are traditionally considered for pansharpening: The *component substitution* (CS) and the *multiresolution analysis* (MRA). The former relies on the substitution of a component (e.g., obtained by means of a spectral transformation of the MS data) with the PAN image. This class comprises such algorithms as *intensity-hue-saturation* (IHS) [9], [10], *principal component analysis* (PCA) [11], [12], [13] and *Gram-Schmidt* (GS) spectral sharpening [14]. Also recently developed methods belonging to this class can be found in the literature [15], [16]. The MRA approach is based on the injection of spatial details that are obtained through a multiresolution decomposition of the PAN image into the resampled MS bands. The spatial details can be extracted according to several modalities of MRA: *decimated wavelet transform* (DWT) [17], *undecimated wavelet transform* (UDWT) [18], *"à-trous" wavelet transform* (ATWT) [19], *Laplacian pyramid* (LP) [20], and nonseparable transforms, either based on wavelets (e.g., contourlet [21]) or not (e.g., curvelet [22]).

These two classes of methods will be detailed in Section II. Approaches different from the two aforementioned classes have also been proposed in the literature. In some cases, these alternative approaches are not fully investigated directions or are composed by isolated methods. For this reason, they will be briefly introduced in the following but not included in the comparison since here the analysis will focus on techniques belonging to the two more established and investigated CS and MRA classes. Algorithms have been proposed based on the Bayesian paradigm for carrying out the data fusion task [23]. The difficulty in finding a suitable statistical model to jointly characterize the pansharpening result and the available MS and PAN images [24] has strongly limited its use for pansharpening. However, many contributions based on Bayesian estimation theory have been presented in the recent literature. They are based on regularized solutions of the ill-posed problem consisting in the reconstruction of the (unknown) high resolution image from its coarse measurements. Some works rely on *Total Variation* penalization terms [25], [26] and others on recent developments in sparse signal representations, or *Compressive Sensing* theory [27], [28]. Among the latter, it is possible to refer to the seminal works [29], [30] that introduced this approach. More recent improvements have been achieved through the application of super-resolution techniques, which are already largely diffused in many image processing and computer vision applications [31]. A method belonging to this family has been presented in [32].

Pansharpening has been also proposed for fusion of panchromatic and hyperspectral (HS) data [33], [34]. Clearly, this task cannot be addressed by employing conventional methods, due to the particular issues that have to be faced, e.g., possible non-simultaneous acquisition, coregistration of the data, different spatial coverages and resolutions [1], [35].

A further contribution of this work is the comparison of the two main validation procedures used for performance assessment. Due to the lack of the reference image (i.e., the image estimated through pansharpening), universal measures for evaluating the quality of the enrichment introduced by pansharpening cannot be explicitly formulated. For this reason, a common practice is the verification of ideal criteria among which the most credited date back to Wald *et al.* [36]. These criteria define the characteristics required in the fused product and are formalized by the *consistency* and *synthesis* properties. The first, which is more easily achievable in practice, involves the reversibility of the pansharpening process; it states that the original MS image should be obtainable by simply degrading the pansharpened image. The synthesis property addresses the characteristics of the pansharpening result, by requiring that the final product has to reproduce the characteristics of the original MS image at a higher resolution. This condition entails that both the features of each single channel and the mutual relations among bands have to be preserved, justifying the original articulation in two distinguished statements.

The definition of a technique that fulfills the constraints defined in the protocol is still an open problem [6], [37] and relates closely to the general discussion regarding image quality assessment [38] and image fusion [39], [40]. Moreover, there are additional issues linked to the lack of a universally accepted evaluation index matching the human capability in assessing the difference of two images. For example the *mean square error* (MSE) has been proven to be inadequate for this task [41], which has given rise to the definition of many other indexes for assessing the image quality (some are presented in Section III). In addition, the unavailability of a reference high resolution MS image precludes the evaluation of the results regardless the chosen quality index. In order to face these aspects and perform a quantitative evaluation of the results, two main solutions have been proposed. The first relies on the reduction of the spatial resolution of both the original MS and PAN images and then the original MS image is used as reference for the evaluation of the results [8]. In this strategy the invariance among scales of the fusion procedures is assumed. This hypothesis is quite often verified in practice but not always [36], [42]. The second employs indexes that do not require the availability of the reference image [39], [40]. Clearly in this case, the evaluation is done at the native scale of the problem but the results are strongly dependent on the definition of such indexes.

The remainder of the paper is organized as follows. Section II presents a review of the CS and MRA approaches providing a detailed description of some algorithms belonging to these classes. Section III is devoted to describe the two assessment procedures operating at reduced and full resolutions, introducing the related quality indexes. Extensive simulations are reported in Section IV together with a detailed discussion of results. Finally, conclusions are drawn in Section V.

II. Pansharpening Techniques

The first goal of this work is the presentation and analysis of some widely used methods presented in the technical

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

TABLE I: LIST OF THE MAIN ACRONYMS.

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

TABLE II: LIST OF THE MAIN SYMBOLS.

Symbol	Description
$\mathbf{MS}$	MS image
$\widetilde{\mathbf{MS}}$	MS image interpolated at the scale of PAN
$\mathbf{P}$	PAN image
$\widehat{\mathbf{MS}}$	Pansharpened image
R	Spatial resolution ratio between MS and PAN
N	Number of MS bands

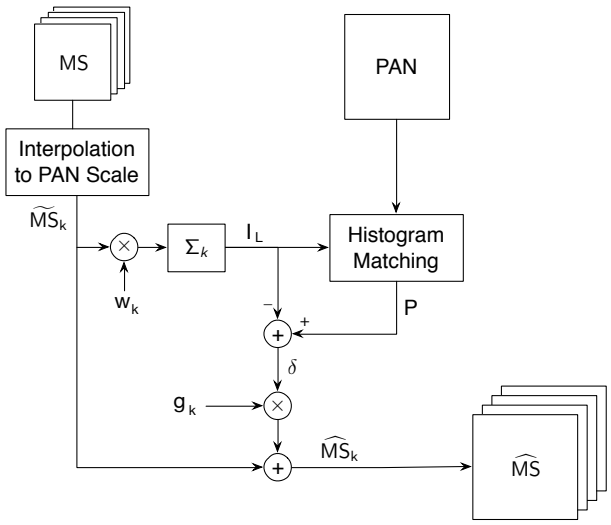

Fig. 1: Flowchart presenting the blocks of a generic pansharpening procedure based on the CS approach.

literature. Due to their widespread use they can be considered as reference state-of-the art algorithms for pansharpening. They can be grouped into two main classes, i.e., methods based on CS and methods based on MRA. Before the two approaches are described, the notation and conventions used in this paper are introduced.

A. Notation

Tables I and II report the lists of the main acronyms and symbols together with a brief description. The other acronyms and symbols will be defined within the paper upon need.

The notation and conventions used in the next sections are detailed in the following. Vectors are indicated in bold lowercase (e.g., $\mathbf{x}$) with the ith element indicated as x_i. Two- and three-dimensional arrays are expressed in bold uppercase (e.g., $\mathbf{X}$). An MS image $\mathbf{X} = \{\mathbf{X}_k\}_{k=1,\ldots,N}$ is a three dimensional array composed by N bands indexed by the subscript $k = 1, \ldots, N$; accordingly, $\mathbf{X}_k$ indicates the kth band of $\mathbf{X}$. A generic PAN image is a 2-D matrix and will be indicated as $\mathbf{Y}$.

B. CS

This family is based on the projection of the MS image into another space assuming that this transformation separates the spatial structure from the spectral information in different components [1]. Subsequently, the transformed MS image can be enhanced by replacing the component containing the spatial structure with the PAN image. Accordingly, the greater the correlation between the PAN image and the replaced component, the lower the distortion introduced by this fusion approach [1]. To this purpose, histogram matching of the panchromatic image to the selected component is performed before the substitution takes place. Thus, the histogram-matched PAN will exhibit same mean and variance as the component to replace. Finally, the pansharpening process is completed by bringing the data back to the original space through the inverse transformation.

This approach is global (i.e., it operates in the same way on the whole image) leading to advantages and limitations. In greater details, techniques belonging to this class are usually characterized by a high fidelity in rendering the spatial details in the final product [43] and they are in general fast and easy to implement. On the other side, they are not able to account for local dissimilarities between the PAN and MS images originated by the spectral mismatch between the PAN and MS channels of the instruments, which may produce significant spectral distortions [1], [6]. An alternative denomination of CS is *projection substitution* [1], to emphasize the two main steps of these techniques.

A new formalization of the CS approach was proposed by Tu *et al.* [44] and then analyzed in subsequent works [43], [45], [46], [47]. It was shown that under the hypothesis of a linear transformation and the substitution of a single component, the fusion process can be obtained without the explicit calculation of the forward and backward transformations, but through a proper injection scheme. This observation leads to a faster implementation of these methods. A general formulation of CS fusion is given by

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + g_k \left(\mathbf{P} - \mathbf{I}_L\right) \quad k = 1, \ldots, N \tag{1}$$

in which the subscript k indicates the kth spectral band, $\mathbf{g} = [g_1, \ldots, g_k, \ldots, g_N]$ is the vector of the *injection gains*, while, $\mathbf{I}_L$ is defined as

$$\mathbf{I}_L = \sum_{i=1}^{N} w_i \widetilde{\mathbf{MS}}_i \tag{2}$$

in which the weight vector $\mathbf{w} = [w_1, \ldots, w_i, \ldots, w_N]$ is the first row of the forward transformation matrix and may be chosen, whenever possible, so as to measure the degrees of spectral overlap among the MS and PAN channels [1], [45].

Fig. 1 shows a flowchart describing the fusion process of the CS approach. Specifically, it is possible to notice the presence of blocks aimed at: i) Interpolating the MS image for matching the scale of PAN; ii) Calculating the intensity component by (2); iii) Matching the histograms of the PAN

© 20XX IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

TABLE III: SPECTRAL WEIGHTS IN (2) AND INJECTION GAINS IN (1) FOR SEVERAL CS-BASED METHODS. IN $w_{k,i}$, SUBSCRIPTS k AND i REFER TO OUTPUT AND INPUT BANDS, RESPECTIVELY.

Method	$w_{k,i}$	g_k
IHS [9]	$1/N$ $(N=3)$	1
GIHS [44], [45], [46]	any $w_i \geq 0$	$(\sum_{i=1}^{N} w_i)^{-1}$
BT [48]	1/N	$\frac{\widetilde{\mathbf{MS}}_k}{\mathbf{I}_L}$
PCA [11]	$\mathbf{X}_{1,i}$	$\mathbf{X}_{1,k}$
GS [14]	1/N	$\frac{\mathrm{cov}(\mathbf{I}_L, \widetilde{\mathbf{MS}}_k)}{\mathrm{var}(\mathbf{I}_L)}$
GSA [43]	$\widehat{w}_i$ (Eq. 7)	$\frac{\mathrm{cov}(\mathbf{I}_L, \widetilde{\mathbf{MS}}_k)}{\mathrm{var}(\mathbf{I}_L)}$
BDSD [49]	$\widehat{w}_{k,i}$(Eqs. (9)-(11))	$\widehat{g}_k$ (Eqs. (9)-(11))
PRACS [47]	$\widehat{w}_i$ (Eq. 7)	Eqs. (13)-(14)

image and the intensity component; iv) Injecting the extracted details according to (1).

For CS methods interpolation must guarantee overlap of MS and PAN at the finer scale. Depending on the acquisition geometry of the imaging instruments, interpolation with conventional zero-phase linear finite-impulse response filters may require a manual realignment, e.g., after a bicubic interpolation. Alternatively, linear nonzero phase filters, having even numbers of coefficients, may be profitably used [50].

The CS family includes many popular pansharpening approaches, such as the IHS [44], [45], PCA [10], [11], [12] and GS [14], [43] methods, which differ by the projections of the MS image used in the process. Due to the lack of a unique transform for extracting the component most suited for substitution, methods based on its adaptive estimation have been proposed; they are known as *adaptive* CS [43], [49].

In the following, a more detailed description of the main CS methods is presented. The values of the spectral weights and injection gains indicated in (2) and (1) are summarized in Table III.

1) (G)IHS: The IHS pansharpening method [9], [10] exploits the transformation into the IHS color space that mimics the human visual system in processing the intensity (I), hue (H) and saturation (S) information. The IHS transform can be only applied to RGB true color images leading to a major limitation for processing MS images. In [44] the authors have generalized the concept of IHS to images with more than three bands (GIHS). Subsequent studies [46] have proven that GIHS can be can be formulated for any arbitrary set of nonnegative spectral weights as:

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \left(\sum_{i=1}^{N} w_i\right)^{-1} (\mathbf{P} - \mathbf{I}_L) \quad k = 1, \ldots, N \quad (3)$$

in which $\mathbf{I}_L$ follows from (2). In general, the coefficients $\{w_i\}_{k=1,\ldots,N}$ are all equal to $1/N$ [44]. Alternatively, they can be optimized to the responses of the spectral channels, MS and PAN [45]. The spectral weights must be nonnegative and may not sum to one. In that case the injection gains provide the proper rescaling [46]. IHS ($N = 3$) may be implemented as a *fast* IHS, which avoids the sequential computation of the direct transformation, substitution and the final backward step. GIHS is implicitly fast, because the transformation does not exist for $N > 3$.

In (1), setting the injection gains, $\mathbf{g}$, such that

$$g_k = \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{I}_L}, \quad k = 1, \ldots, N \quad (4)$$

yields

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{I}_L} \cdot (\mathbf{P} - \mathbf{I}_L) = \widetilde{\mathbf{MS}}_k \cdot \frac{\mathbf{P}}{\mathbf{I}_L} \quad (5)$$

which is the widely known *Brovey transform* (BT) pansharpening method [48]. Thus, BT may be achieved by the general model (1) with the choice of injection gains (4) and represents a *multiplicative* sharpening, that is, a spatial modulation of spectral pixels.

The group of methods defined by (5), varying with the choice of spectral weights in (2), is sometimes referred to as *relative spectral contribution* (RSC) [1], [6]. In this paper, however, following [7], all CS methods are considered as a unique class. According to (1) RSC can be seen as a particular case of the CS class, since such methods can be formalized and exhibit the same features as all CS methods [51].

2) PCA: PCA, a.k.a. *Karhunen-Loéve transform* or *Hotelling transform* [52], is another spectral transformation that has been widely employed for pansharpening [10], [11], [12]. PCA is achieved through a multidimensional rotation of the original coordinate system of the N-dimensional vector space, i.e., a linear transformation of the data, such that the projection of the original spectral vectors on the new axes, which are the eigenvectors of the covariance matrix along the spectral direction, produces a set of scalar images, called *principal components* (PCs), that are statistically uncorrelated to each other. PCs are generally sorted for decreasing variance, which quantifies their information content.

Specifically, the hypothesis underlying its use for pansharpening is that the spatial information (shared by all the channels) is concentrated in the first component (PC1), while the spectral information (specific to each single band) is accounted by the other $N - 1$ components. However, the *equivalent* spectral response of PC1, i.e., the combination of the spectral responses of the MS instrument according to the weights of the PCA transformation, may not match the spectral response of the PAN instrument. Actually, the spatial information is mapped to the first component, to an extent proportional to the correlation among the MS channels [6]. Again, the fusion process can be described by the general formulation stated by (1), where the $\mathbf{w}$ and $\mathbf{g}$ coefficient vectors are derived by the PCA procedure on the MS image. In general [44], [46], $\mathbf{w}$ is the first row of the forward transformation matrix; $\mathbf{g}$ is the first column of the backward transformation matrix, which is equal to its transpose for PCA, whose matrix $\mathbf{X}$ is data dependent.

3) GS: The GS orthogonalization procedure is the basis for defining a powerful pansharpening method, patented by its inventors [14] for Kodak and implemented in ENVI©, release 4.3 onward, as *Gram-Schmidt spectral sharpening*. The GS transformation is a common technique used in linear algebra and multivariate statistics to orthogonalize a set of vectors.

Before the fusion process starts, the MS bands are interpolated at the scale of PAN and all images are lexicographically ordered as vectors, whose dimension is the number of image

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

pixels at the scale of PAN. Also the mean (average) of each band (now a vector) is subtracted from all the components of the same vector. The orthogonalization procedure exploits a synthetic low resolution approximation of the PAN image, namely $\mathbf{I}_L$, as the first vector of the new orthogonal basis. GS orthogonalization proceeds one MS vector at the time, by finding its projection on the (hyper)plane defined by the previously found orthogonal vectors and its orthogonal component, such that the sum of the orthogonal and projection components is equal to the zero-mean version of the original vectorized band. Pansharpening is accomplished by replacing $\mathbf{I}_L$ with the histogram-matched $\mathbf{P}$ before the inverse transformation is performed. Therefore, GS constitutes an orthogonal decomposition more general than PCA, which can be obtained by using the first PC (PC1), instead of $\mathbf{I}_L$, as starting vector [53]. The process of finding projections and orthogonal components is formally identical (orthogonal zero-mean vectors is formally equivalent to uncorrelated components).

Since GS is a generalization of PCA, in which PC1 may be arbitrarily chosen and the remaining components are calculated so as to be orthogonal/uncorrelated to one another and to PC1, the procedure may be still be described by the flowchart in Fig. 1. Again, the fusion process is described by (1), with the injection gains given by [43]:

$$g_k = \frac{\mathrm{cov}(\widetilde{\mathbf{MS}}_k, \mathbf{I}_L)}{\mathrm{var}(\mathbf{I}_L)} \quad k = 1, \ldots, N \tag{6}$$

in which $\mathrm{cov}\,(\mathbf{X}, \mathbf{Y})$ indicates the covariance between two images $\mathbf{X}$ and $\mathbf{Y}$, and $\mathrm{var}\,(\mathbf{X})$ is the variance of $\mathbf{X}$.

Several versions of GS are achieved by changing the method for generating $\mathbf{I}_L$. The simplest way to obtain the low-resolution approximation of PAN consists of simply averaging the MS components (i.e., setting $w_i = 1/N$, for all $i = 1, \ldots, N$); this modality is simply called GS or *GS mode 1*[14]. In [43] the authors proposed an enhanced version, called *adaptive* GS (*GSA*), in which $\mathbf{I}_L$ is generated by a weighted average of the MS bands, with MSE-minimizing weights with respect to a low-pass filtered version of PAN, namely

$$\mathbf{P}_L = \sum_{k=1}^{N} w_k \widetilde{\mathbf{MS}}_k \tag{7}$$

The set of optimal weights $\{\widehat{w}_k\}_{k=1,\ldots,N}$ is calculated as the *minimum MSE* (MMSE) solution of (7).

A further option for generating the low resolution intensity image $\mathbf{I}_L$ entails the application of a low-pass filter to the original PAN image and is referred to as *GS mode 2* [14], leading to a hybrid approach, which is no longer based on CS, but actually belongs to the MRA-based class.

Other data-dependent self-adaptive approaches fall within the class of CS methods and thus are briefly described in the following. The *band dependent spatial detail* (BDSD) algorithm [49] starts from an extended version of the generic formulation (1):

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + g_k \left(\mathbf{P} - \sum_{i=1}^{N} w_{k,i} \widetilde{\mathbf{MS}}_i \right) \qquad k = 1, \ldots, N. \tag{8}$$

By defining the coefficients

$$\gamma_{k,i} = \begin{cases} g_k & \text{if } i = N+1 \\ -g_k \cdot w_{k,i} & \text{otherwise.} \end{cases} \tag{9}$$

Eq. (1) can be rewritten in compact matrix form as

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \mathbf{H}\boldsymbol{\gamma}_k \tag{10}$$

in which $\mathbf{H} = [\widetilde{\mathbf{MS}}_1, \ldots, \widetilde{\mathbf{MS}}_N, \mathbf{P}]$, $\boldsymbol{\gamma}_k = [\gamma_{k,1}, \ldots, \gamma_{k,N+1}]^T$ (all the images are organized by columns). The optimal MMSE joint estimation of the weights and gains vector $\boldsymbol{\gamma}$ would encompass the use of the unknown target image $\widetilde{\mathbf{MS}}_k$ and is thus performed at a reduced resolution. Consequently, the solution is found as

$$\widehat{\boldsymbol{\gamma}}_k = \left(\mathbf{H}_d^T \mathbf{H}_d\right)^{-1} \mathbf{H}_d^T \left(\widetilde{\mathbf{MS}}_k - \widetilde{\mathbf{MS}}_k^{\mathrm{LP}} \right) \tag{11}$$

in which $\mathbf{H}_d$ is the reduced resolution version of $\mathbf{H}$ and $\widetilde{\mathbf{MS}}_k^{\mathrm{LP}}$ is a version of $\widetilde{\mathbf{MS}}_k$ obtained through a low-pass filter, whose spatial frequency response matches the average *modulation transfer function* (MTF) of the MS sensor.

In [47] the concept of *partial replacement* of the intensity component is introduced. The PAN image is not directly used for CS; instead the algorithm utilizes $\mathbf{P}^{(k)}$, a weighted sum of PAN and of the kth MS band to calculate the kth sharpened band in (1). For this reason, this method is referred to as *partial replacement adaptive CS* (PRACS). For $k = 1, \ldots, N$, the band-dependent high-resolution sharpening image is calculated as

$$\mathbf{P}^{(k)} = \mathrm{CC}(\mathbf{I}_L, \widetilde{\mathbf{MS}}_k) \cdot \mathbf{P} + (1 - \mathrm{CC}(\mathbf{I}_L, \widetilde{\mathbf{MS}}_k)) \cdot \widetilde{\mathbf{MS}}_k' \tag{12}$$

in which $\widetilde{\mathbf{MS}}_k'$ is the kth MS band histogram-matched to PAN, $\mathrm{CC}\,(\mathbf{X}, \mathbf{Y})$ is the correlation coefficient between $\mathbf{X}$ and $\mathbf{Y}$, and $\mathbf{I}_L$ is given by (2) where the weights w_k are obtained through the linear regression of $\widetilde{\mathbf{MS}}_k'$, $k = 1, \ldots, N$, on $\mathbf{P}_L$, the PAN image spatially degraded to the MS resolution. The injection gains $\{g_k\}$ are obtained according to

$$g_k = \beta \cdot \mathrm{CC}\left(\mathbf{P}_L^{(k)}, \widetilde{\mathbf{MS}}_k\right) \frac{\mathrm{std}\left(\widetilde{\mathbf{MS}}_k\right)}{\frac{1}{N}\sum_{i=1}^{N} \mathrm{std}\left(\widetilde{\mathbf{MS}}_i\right)} L_k. \tag{13}$$

According to [47], (13) is given by the product of: 1) an empirically tuned parameter, β, that normalizes the high frequencies, so that they lie in the correct dynamic range; 2) a correlation coefficient that adjusts the relative magnitude of the high-frequency content, with the purpose of minimizing the global dissimilarity between the low-resolution image $\mathbf{P}_L^{(k)}$, achieved by low-pass filtering $\mathbf{P}^{(k)}$, and each MS band); 3) a scaling coefficient that takes into account the spectral distortion due to differences in standard deviations among the MS bands; 4) an adaptive factor L_k defined as

$$L_k = 1 - \left| 1 - \mathrm{CC}\left(\mathbf{I}_L, \widetilde{\mathbf{MS}}_k\right) \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{P}_L^{(k)}} \right| \tag{14}$$

and aimed at removing the local spectral instability error between the synthetic component image and the MS band.

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

C. MRA

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

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + g_k \left(\mathbf{P} - \mathbf{P}_L\right), \; k = 1, \ldots, N. \tag{15}$$

In [54] this paradigm has been denoted as *Amélioration de la Résolution Spatiale par Injection de Structures* (ARSIS) to highlight that the purposes of these methods are the preservation of the whole content of the LRMS image and the addition of further information obtained from the PAN image, through spatial filtering.

According to Tu *et al.* [44], who recast earlier wavelet-based pansharpening methods into (15), the different approaches belonging to this class are uniquely characterized by the algorithm employed for obtaining the image $\mathbf{P}_L$ and by the injection gains $\{g_k\}_{k=1,\ldots,N}$.

In a very general setting, $\mathbf{P}_L$ is achieved through an iterative decomposition scheme (called MRA) that aims at constructing a sequence of 2-D signals with progressively reduced information, through the repeated application of some analysis operators. Information is regarded as spatial frequency content, and thus possibly increased sampling step sizes may be used throughout the sequence; hence the alternative name of *pyramid*. The type of decomposition constitutes a first distinguishing feature of the various approaches; it can range from very simple methods based on a single level decomposition achieved through a simple low-pass filter to more complex techniques relying upon a formal MRA.

The mathematical expression of $\mathbf{g}$ assumes different forms in the literature. Noteworthy is the *high-pass modulation* (HPM) method [55], relying on a *multiplicative* combination of MS and PAN, which is defined, for $k = 1, \ldots, N$, by

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + \frac{\widetilde{\mathbf{MS}}_k}{\mathbf{P}_L}\left(\mathbf{P} - \mathbf{P}_L\right) = \widetilde{\mathbf{MS}}_k \cdot \frac{\mathbf{P}}{\mathbf{P}_L} \tag{16}$$

in which the details are weighted by the ratio of the MS and the PAN low-pass version $\mathbf{P}_L$ before they are injected, with the aim of reproducing the local intensity contrast of PAN in the fused image [56]. Furthermore, if a unique low-pass image $\mathbf{P}_L$ is used for all the MS bands, that is, if a unique filter or filter bank is used to process PAN in order to extract the details that shall be injected in all MS bands, this algorithm clamps the spectral distortion of $\widetilde{\mathbf{MS}}$, which can be quantified by the *spectral angle* between fused and interpolated MS; thus, it is an element of the *spectral distortion minimization* (SDM) family [57].

The general scheme of MRA fusion methods is reported in Fig. 2. Accordingly, the required blocks are devoted to: i) Interpolate the MS image to reach the panchromatic scale; ii) Calculate the low-pass version $\mathbf{P}_L$ of the PAN image by means of the equivalent filter for a scale ratio equal to R; iii) Compute the band dependent injection gains $\{g_k\}_{k=1,\ldots,N}$; iv) Inject the extracted details according to (15). Note that, apart from the filter, there is difference if $\mathbf{P}_L$ is decimated-interpolated or not. In the former case, corresponding to either GLP or DWT, it is

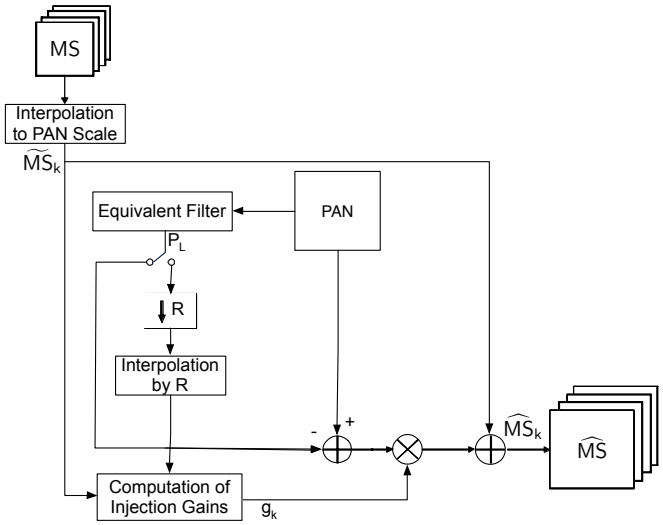

Fig. 2: Flowchart of a generic pansharpening algorithm belonging to the MRA class. The switch on the low-pass filtered PAN, P_L, enables a decimated (GLP, DWT) or an undecimated (ATWT, UDWT) type of analysis.

possible to compensate the aliasing of the MS image through the fusion process [58].

Interpolation is less crucial than for CS methods, because if the original data sets are intrinsically misaligned by constant values along rows and columns, e.g., by 1.5 pixels for 4:1 scale ratio between MS and PAN (see [50]), the subsequent step may realign the details extracted from PAN, provided that linear nonzero phase filters are used for MRA. Table IV reports a brief summary of the MRA-based methods that have been implemented, giving evidence to the paradigm used for calculating the injection gains required by (15) and the filtering scheme.

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

$$\widehat{\mathbf{MS}}_k = \widetilde{\mathbf{MS}}_k + g_k \left(\mathbf{P} - \mathbf{P} * h_{LP}\right), \; k = 1, \ldots, N \tag{17}$$

in which $*$ denotes the convolution operator. This approach can be performed by employing several different low-pass filters h_{LP}, among which the most diffused uses box, Gaussian and Laplacian masks [6].

Among the possible couples of filters and coefficients, we chose the simplest scheme achievable by using the box mask (i.e., a mask with uniform weights, implementing an average) and additive injection, which leads to the pansharpening algorithm known as *high-pass filtering* (HPF) [10], [55].

The corresponding methodology employing the HPM injection scheme has been proposed in [59], [60]; it is named *smoothing filter-based intensity modulation* (SFIM).

2) Pyramidal Decompositions: The resolution reduction needed to obtain the low-pass signal $\mathbf{P}_L$ at the original MS scale can be performed in one or more steps, namely by employing a single low-pass filter with cut frequency equal to $1/R$ and decimating by R, or by multiple fractional steps. This

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

TABLE IV: MRA-BASED PANSHARPENING METHODS AND RELATED MRA SCHEMES WITH FILTERS AND INJECTION GAINS.

Method	Type of MRA and Filter	gain g_k
HPF [10]	ATWT w/ Box Filter	1
HPM / SFIM [55] / [59], [60]	ATWT w/ Box Filter	$\frac{\widetilde{\mathbf{MS}}_k}{\mathbf{P}_L}$
Indusion [61]	DWT w/ CDF Bior. Filt.	1
MTF-GLP [62]	GLP w/ MTF Filter	1
MTF-GLP-CBD [62]	GLP w/ MTF Filter	$\frac{\mathrm{cov}(\mathbf{P}_L, \widetilde{\mathbf{MS}}_k)}{\mathrm{var}(\mathbf{P}_L)}$
ATWT [56]	ATWT w/ S&M Filter	1
ATWT-M2 [54]	ATWT w/ S&M Filter	see [54]
ATWT-M3 [54]	ATWT w/ S&M Filter	see [54]
MTF-GLP-HPM [57]	GLP w/ MTF Filter	$\frac{\widetilde{\mathbf{MS}}_k}{\mathbf{P}_L}$
MTF-GLP-HPM-PP [47]	GLP w/ MTF Filter	$\frac{\widetilde{\mathbf{MS}}_k}{\mathbf{P}_L}$
AWLP [63]	ATWT w/ S&M Filter	$\frac{\widetilde{\mathbf{MS}}_k}{\frac{1}{N}\sum_{i=1}^{N}\widetilde{\mathbf{MS}}_i}$

second method, that includes the first as a particular case, is commonly referred to as pyramidal decomposition and dates back to the seminal work of Burt and Adelson [20], which utilized Gaussian LPFs to carry out the analysis steps. The corresponding differential representation, achieved by calculating the differences between the Gaussian pyramid levels, is named Laplacian pyramid and has later been proven to be very valuable for pansharpening purposes [64]. Indeed the Gaussian filters can be tuned to closely match the sensor MTF. This allows to extract from the PAN image those details, which are not seen by the MS sensor, due to the coarser spatial resolution [62]. Since the Gaussian mask is defined by a single parameter (i.e., its standard deviation), its frequency response is fully specified by fixing it. To this aim, the value of the amplitude response at the Nyquist frequency is used, being it commonly provided by the manufacturer as a sensor specification, or obtained through on-orbit measurements. However it is useful to keep in mind that, components aging can induce a significant incertitude of this parameter, also in the more favorable case of on-orbit estimation.

Both the *general* (15) and the *signal dependent* injection scheme (16) [65] have been used in the literature and implemented in this work. They will be referred to as Modulation Transfer Function - Generalized Laplacian Pyramid (MTF-GLP) [62] and MTF-GLP with High-Pass Modulation (MTF-GLP-HPM) [65], respectively. Following the indications of the authors of [65], we also tested the MTF-GLP-HPM method followed by a PostProcessing (PP) phase aimed at correcting the noise generated in the presence of discontinuities. For the latter approach, whose details can be found in the cited reference, we use here the acronym MTF-GLP-HPM-PP.

As a further example of Gaussian pyramids exploiting an analysis filter matched with the MS sensor MTF we consider an algorithm that relies upon the optimization of the injection coefficients by least square fitting. In particular, it is based on (15), in which the coefficients are calculated as:

$$g_k = \frac{\mathrm{cov}(\widetilde{\mathbf{MS}}_k, \mathbf{P}_L)}{\mathrm{var}(\mathbf{P}_L)}, \tag{18}$$

where, in general, $\mathbf{P}_L$ depends on the kth band. This injection rule, also used by GS approaches [43], leads to a powerful algorithm. It is commonly known as MTF-GLP with Context Based Decision MTF-GLP-CBD [8], since the injection coefficient can be locally optimized by patching the image in non-overlapping zones. Actually, the consequent performance improvement is reflected in terms of a significant increase in the execution time.

In the same field, very appreciable results have been achieved by employing wavelet and contourlet decomposition pyramids [66], [13]. By focusing on the widely diffused wavelet decompositions, the decimated Mallat's method has been initially used for this purpose [67], [54]. As an example of this class, we selected for the comparison the more recent *Indusion* method [61], that profits from multiple equalization steps to achieve remarkable results.

However, the undecimated "à trous" method [68] has soon emerged as a very effective method [69]. Indeed, even if non-orthogonality (which implies that a wavelet plane could retain information for a neighboring plane) could compromise the spectral quality of the fused product [69], its beneficial characteristics such as the shift-invariance property [70] and the capability of being easily matched to the sensor MTF [62] produce accurate pansharpened image. We adopt here a widely used implementation of the "à trous" filter based on the sequential application (thanks to the separability property), in the vertical and horizontal direction, of 1-D filters [22]

$$h = [1 \quad 4 \quad 6 \quad 4 \quad 1], \tag{19}$$

that derives from the choice of a B_3 cubic spline as scaling function [71] (also called Starck and Murtagh (S&M) filter).

In this case, here denoted as ATWT, choosing the signal-dependent injection formula (16) for achieving the final product leads to the formulation as in [68]. Furthermore, alternative implementations using the *Model 2* and *Model 3* [54] have been added to the comparison as two further assessment methods and named ATWT-M2 and ATWT-M3, respectively. They consist in equalizing the first two moments of the details extracted by the PAN images, before injection. In both cases the mean and the standard deviation of the PAN and MS wavelet coefficient at the original MS scale are used. Whilst Model 2 employs a deterministic relation among the corresponding quantities, Model 3 is optimized by least square fitting.

All the previous wavelet methods are based on the choice of unitary injection coefficients $\{g_k\}_{k=1,\ldots,N}$. However, some further improvements can be achieved by injecting the details using the HPM paradigm (16) [56]. As an example of wavelet-based method employing a different choice we implemented the *additive wavelet luminance proportional* (AWLP) [63] that uses the more general fusion formula reported in (15), with the injection coefficients defined as

$$g_k = \frac{\widetilde{\mathbf{MS}}_k}{\frac{1}{N}\sum_{i=1}^{N}\widetilde{\mathbf{MS}}_i}, \quad k = 1, \ldots, N. \tag{20}$$

III. QUALITY ASSESSMENT OF FUSION PRODUCTS

As in most data fusion problems, the absence of a reference image is the main limitation for the evaluation of the results. In the context of pansharpening this lack prevents the direct application of the Wald protocol. Thus, two assessment

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

procedures have been proposed in order to circumvent this problem. The first one considers the images at a spatial resolution lower than the original (induced artificially) and uses the original MS image as a reference. Although this procedure allows for a precise evaluation of the results with established indexes, qualitatively there might be mismatches between the performances obtained at reduced resolution and the quality (interpreted by visual inspection) for the fusion product at the original scale [4]. Indeed, the performances are intrinsically related to way the resolution degradation is performed, especially in the case of pansharpening methods exploiting spatial filters [7].

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

Due to the sub-optimality of both quantitative evaluation procedures, a qualitative evaluation of the results through visual inspection is still a necessary procedure for appreciating local spectral distortions and precision in rendering the spatial details in the fused images [8].

A. Reduced Resolution Assessment

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

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

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

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

Clearly, the choice of the filter is crucial in this validation protocol. In general, the filter is defined for ensuring the *consistency* (defined by the first Wald's statement) of the pansharpening process. Since the pansharpened image (that here should match as close as possible the original image $\mathbf{MS}$), once degraded to its original resolution, should be identical to the original multispectral image (whose part is acted by $\mathbf{MS}^*$), it comes natural that the resolution reduction has to be performed by employing a filter simulating the transfer function of the remote sensor. In other terms the degradation filter has to match the MTF of the sensor [62]. In addition, the filter used for obtaining the panchromatic image $\mathbf{P}^*$ has to be designed in order to preserve the details that would have been seen if the image were acquired at the reduced resolution. Accordingly, it is commonly degraded through an ideal filter [62].

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

Vector measures are useful to quantify the spectral distortion. A simple index that has assumed a key role in the technical literature is the *Spectral Angle Mapper* (*SAM)* [72], which consists of calculating the angle between the corresponding pixels of the fused and reference images in the space defined by considering each spectral band as a coordinate axis. Let $\mathbf{I}_{\{n\}} = \left[I_{1,\{n\}}, \dots, I_{N,\{n\}}\right]$ be a pixel vector of the MS image $\mathbf{I}$ with N bands, the *SAM* between $\mathbf{I}_{\{i\}}$ and $\mathbf{J}_{\{i\}}$ is defined as:

$$SAM(\mathbf{I}_{\{i\}}, \mathbf{J}_{\{i\}}) = \arccos\left(\frac{\langle \mathbf{I}_{\{i\}}, \mathbf{J}_{\{i\}} \rangle}{\|\mathbf{I}_{\{i\}}\| \, \|\mathbf{J}_{\{i\}}\|}\right) \tag{21}$$

in which $\langle \cdot, \cdot \rangle$ denotes the scalar product (or inner product) and $\|\cdot\|$ the vector ℓ_2 norm. The global value of SAM for the whole image is obtained by averaging the single measures over all the pixels. The optimal value of the SAM index is 0.

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

$$RMSE\left(\mathbf{I}, \mathbf{J}\right) = \sqrt{E\left[\left(\mathbf{I} - \mathbf{J}\right)^2\right]}. \tag{22}$$

The ideal value of *RMSE* is zero and is achieved if and only if $\mathbf{I} = \mathbf{J}$. The main drawback of RMSE is that errors in each band are not related to the mean value of the band itself. Relative error measurements better match both visual analysis criteria and model based processing of MS pixel values, e.g., calculation of normalized differential vegetation index (NDVI).

A more credited global index is the *Erreur Relative Globale Adimensionnelle de Synthèse* (*ERGAS*) that was properly proposed for pansharpening in [73] and is defined as:

$$ERGAS = \frac{100}{R}\sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\frac{RMSE(\mathbf{I}_k, \mathbf{J}_k)}{\mu\left(\mathbf{I}_k\right)}\right)^2}, \tag{23}$$

where the *RMSE* is defined as in (22) and μ represents the mean (average on pixels) of the image. Since the ERGAS is composed by a sum of *RMSE* values, its optimal value is 0.

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

Another (scalar) index which was developed to overcome some limitations of the *RMSE* is the *Universal Image Quality Index* (*UIQI*) or *Q-index*, proposed by Wang and Bovik [38]. Its physical interpretation becomes straightforward by writing its expression in the form:

$$Q\left(\mathbf{I},\mathbf{J}\right)=\frac{\sigma_{\mathbf{IJ}}}{\sigma_{\mathbf{I}}\sigma_{\mathbf{J}}}\frac{2\overline{\mathbf{I}}\overline{\mathbf{J}}}{\left(\overline{\mathbf{I}}\right)^{2}+\left(\overline{\mathbf{J}}\right)^{2}}\frac{2\sigma_{\mathbf{I}}\sigma_{\mathbf{J}}}{(\sigma_{\mathbf{I}}^{2}+\sigma_{\mathbf{J}}^{2})}, \tag{24}$$

where $\sigma_{\mathbf{IJ}}$ is the sample covariance of $\mathbf{I}$ and $\mathbf{J}$ and $\overline{\mathbf{I}}$ is the sample mean of $\overline{\mathbf{I}}$. Accordingly, it comprises, in the order, an estimate of correlation coefficient (CC), the differences in the mean luminance and in the contrast. The *Q-index* varies in the range $[-1,1]$, with 1 denoting the best fidelity to reference.

A vector extension of the *Q-index* to vector data up to four bands, which accounts also for spectral distortion, has been proposed in [74]. In practice, the Q4 vector index can be used with data sets composed by four spectral bands (e.g., Blue, Green, Red and Near InfraRed) and is based on modeling each pixel $\mathbf{I}_{\{i\}}$ as a *quaternion*

$$\mathbf{I}_{\{i\}}=\mathbf{I}_{\{i\},1}+\mathbf{i}\mathbf{I}_{\{i\},2}+\mathbf{j}\mathbf{I}_{\{i\},3}+\mathbf{k}\mathbf{I}_{\{i\},4}. \tag{25}$$

A generalization of the $Q4$ index, suitable for the assessment of images with a number of spectral bands greater than four is presented in [75] and is called Q2^n-index.

B. Full Resolution Validation

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

$$QNR=\left(1-D_{\lambda}\right)^{\alpha}\left(1-D_{S}\right)^{\beta}. \tag{26}$$

It is thus composed by the product, weighted by the coefficients α and β, of two separate values D_λ and D_S, which quantify the spectral and the spatial distortion, respectively. The higher the QNR index, the better the quality of the fused product. The maximum theoretical value of this index is 1 when both D_λ and D_S are equal to 0.

The spectral distortion is estimated by

$$D_{\lambda}=\sqrt[p]{\frac{1}{N\left(N-1\right)}\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N}\left|d_{i,j}(\mathbf{MS},\widehat{\mathbf{MS}})\right|^{p}}, \tag{27}$$

where $d_{i,j}(\mathbf{MS},\widehat{\mathbf{MS}})=Q\left(\mathbf{MS}_i,\mathbf{MS}_j\right)-Q\left(\widehat{\mathbf{MS}}_i,\widehat{\mathbf{MS}}_j\right)$.

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

The spatial distortion is calculated by

$$D_{S}=\sqrt[q]{\frac{1}{N}\sum_{i=1}^{N}\left|Q\left(\widehat{\mathbf{MS}}_i,\mathbf{P}\right)-Q\left(\mathbf{MS}_i,\mathbf{P}_{LP}\right)\right|^{q}}, \tag{28}$$

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

IV. Experimental Results

This section is devoted to describing a set of experiments that have been carried out on very high resolution optical data. The results have been validated at both full and reduced resolution as described in the previous section. Firstly, the data sets used in the experiments are briefly presented along with a description of the experimental setup. Successively the experimental results are reported and discussed.

A. Data sets

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

a) China data set: The data set[1] represents a mountainous and vegetated area of the China-Sichuan region. The data set is of size 300×300 pixels. It is acquired by the IKONOS sensor which works in the visible and near infrared spectrum range; the multispectral sensor is characterized by four bands (blue, green, red and near infrared) and also a panchromatic channel is available. The resolution cell is 4 m $\times$ 4 m for the multispectral bands and 1 m $\times$ 1 m for the panchromatic channel. The resolution ratio R is therefore equal to 4. The radiometric resolution is 11 bits.

b) Toulouse data set: This data set represents an urban area of the city of Toulouse (France). The size is equal to 512×512 pixels. The sensor used to acquire the images is IKONOS, and, therefore, it has the same spectral and spatial characteristics as in the previous case.

c) Rome data set: This represents an urban area of the city of Rome (Italy). The dimension of the image is 300×300 pixels. The data set has been acquired by the WorldView-2 sensor, which provides a high resolution panchromatic channel and eight MS bands. Four standard colors (red, green, blue, and near-infrared 1) and four new bands (coastal, yellow, red edge, and near-infrared 2) are acquired. Though the native spatial resolution would be greater, the images are distributed with a pixel size of 0.5 m and 2 m for PAN and MS, respectively.

The resolution ratio R is equal to 4, again. The radiometric resolution is 11 bits.

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

e) Pléiades data set: The Pléiades data set, collected by an aerial platform, was used for the 2006 contest [8] and was provided by CNES, the French Space Agency. This represents an urban area of Toulouse (France) and it has a size of 1024×1024 pixels. The resolution of the four MS bands is 60 cm, but the data set lacks the availability of the panchromatic image, since the corresponding sensor was under

[1]The data set is available at http://glcf.umiacs.umd.edu

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

development. Accordingly, the high-resolution panchromatic data was simulated by the following procedure. The green and red channels were averaged and the result was filtered with a system characterized by the nominal MTF of the panchromatic sensor. After the resampling to 80 cm, thermal noise was added and the final simulated image was achieved by inverse filtering and wavelet denoising. As a consequence, also the low resolution MS image were simulated according to the Wald protocol, namely by MTF-filtering and decimation. A resolution ratio R of 4 was chosen. The radiometric resolution is 11-bits.

B. Algorithms

The algorithms used for the comparison are listed below. It is worth highlighting that MRA requires injection gains different from unity, because MS and PAN data sets are generally not equalized, i.e., they are not expressed in the same units. MRA methods in the literature employ their own injection gain to equalize the contribution of PAN to inject into the MS [59], [6].

In this work we adopted a preprocessing step of PAN, common to all methods that are compared, which consists of creating as many PAN images as MS bands. Each new PAN image is equalized to the corresponding MS band by matching the two first moments of the two images before applying the fusion algorithm.

Furthermore, when no authors' specifications are provided, the interpolation of the original MS image $\mathbf{MS}$ for generating the interpolated image $\widetilde{\mathbf{MS}}$ is carried out by using a polynomial kernel with 23 coefficients [64], followed by manual realignment by 1.5 pixels for data sets with even MS-PAN symmetry, in practice all commercial products. Only in the simulated Pléiades data set, the PAN image is simulated from the spectral components of the 80 cm airborne instrument, before they are spatially degraded to yield the simulated 2.4 m MS bands. Hence, MS and PAN exhibit odd symmetry and conventional zero-phase interpolators can be used without any loss of alignment. Interested readers are addressed to [50] for more information about the interpolation step.

- EXP: MS image interpolation, using a polynomial kernel with 23 coefficients [64]
- IHS: *Fast Intensity-Hue-Saturation* (GIHS) image fusion [44]
- BT: *Brovey transform* [48]
- PCA: *Principal Component Analysis* [10]
- BDSD: *Band-Dependent Spatial-Detail* with local parameter estimation [49]
- GS: *Gram Schmidt* (Mode 1) [14]
- GSA: *Gram Schmidt Adaptive* [43]
- PRACS: *Partial Replacement Adaptive Component Substitution* [47]
- HPF: *High-Pass Filtering* with 5 $\times$ 5 box filter for 1:4 fusion [10]
- SFIM: *High-Pass Modulation* with 5 $\times$ 5 box filter [55], a.k.a. *Smoothing Filter-based Intensity Modulation* (*SFIM*) [59], [60]
- Indusion: Decimated Wavelet Transform using an *additive* injection model [61]
- MTF-GLP: *Generalized Laplacian Pyramid* (GLP) [64] with MTF-matched filter [62] with unitary injection model
- MTF-GLP-HPM: GLP with MTF-matched filter [62] and *multiplicative* injection model [57]
- MTF-GLP-HPM-PP: GLP with MTF-matched filter [62], *multiplicative* injection model and Post-Processing [65]
- MTF-GLP-CBD: GLP [64] with MTF-matched filter [62] and regression based injection model [8]
- ATWT: Additive A Trous Wavelet Transform with unitary injection model [56]
- ATWT-M2: A Trous Wavelet Transform using the Model 2 proposed in [54]
- ATWT-M3: A Trous Wavelet Transform using the Model 3 proposed in [54]
- AWLP: *Additive Wavelet Luminance Proportional* [63], a generalization of AWL [68] to more than three bands

A more detailed description of the methods can be found in the relative references or in Section II.

C. Wald's Protocol

Three data sets are considered for evaluating the results at reduced resolution: The China, Rome and Pléiades data sets. The results of the fusion are shown in Figs. 3 - 5. The quality indexes used are the SAM (in degrees) to measure the spectral distortion, the $ERGAS$ and the $Q2^n$ (i.e., $Q4$ for four band data sets and $Q8$ for the WorldView-2 data set). The quantitative results obtained are reported in Tables V, VI and VII.

A first analysis of the results can be done per family of pansharpened algorithms. Among the CS approaches, the adaptive methods, namely the BDSD and PRACS, achieve the best performances with a good visual appearance of the final products and a reduced spectral distortion with respect to the other members of this family. They are followed by the GS, PCA and IHS approaches, ranked in this order for most of the data sets and the quality indexes. However, only for the case of vegetated area, i.e., the China data set, the PCA gives slightly better results than the GS algorithm. More in general, PCA and IHS-based approaches produce results with larger spectral distortion (see the greater values of the SAM) as it can also be confirmed by the visual inspection.

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

Considering the techniques belonging to the MRA family, the best accuracy is obtained by decimated approaches (e.g., Laplacian pyramids), in particular for urban areas, thanks to a greater robustness to aliasing with respect to undecimated methodologies, such as, the "à trous" wavelet transform. However, very good performances for both the above-mentioned approaches can be attributed to the similarity of the frequency response of the filters with that of the MS sensor MTF [62]. While, the application of the box filter leads to poorer performances. This is mainly caused by the spatial artifacts

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

due to the presence of ripples in the obscure band in the representation of this filter in the frequency domain and its greater dissimilarity with respect to the transfer function of the acquisition device. Even in the case of *Indusion*, severe artifacts introduced by the decimation can be noticed (see Figs. 3, 4 and 5).

Regarding the injection methodologies, the *multiplicative* injection model turns out to be better suited than the *additive* injection model. This can be explained by the greater flexibility of the former in setting the local weights that govern the detail injection and its close relationship with local contrast of the image [56].

ATWT-M2 and ATWT-M3 [54] do not lead to high performances both from a numerical and visual (leading to more blurred result) points of view when compared to the other approaches that use the same detail extraction filters but different injection methodologies. Finally, it is worth to underline that the results of the MTF-GLP-HPM-PP method are, in general, poorer than those obtained by MTF-GLP-HPM with no gain due to the PP phase and drawbacks provided by the iterative application of MTF-based filters and non-ideality of the interpolation. Furthermore, MTF-GLP-CBD shows comparable performances with respect to *additive* injection models and it represents the best choice when we deal with data with lower signal to noise ratio (i.e., the Pléiades data set).

In general, the application of a pre-processing such as the matching of the two first moments of the panchromatic image with those of each MS band, improves the performances of some MRA algorithms (e.g., approaches based on *additive* and *multiplicative* injection models), which do not consider a proper injection model to enhance the MS product after the extraction of the spatial details of the panchromatic image (i.e., they inject the extracted details in the same way for all the bands).

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

TABLE V: CHINA DATA SET: QUANTITATIVE RESULTS.

	Q4	SAM(°)	ERGAS	Time[s]
Reference	1	0	0	0
EXP	0.7398	4.4263	3.8471	0
PCA	0.8578	3.5433	2.6715	0.1189
IHS	0.7308	4.9892	3.5766	0.0079
BT	0.7314	4.4263	3.1722	**0.0053**
BDSD	**0.8869**	**2.9123**	2.4124	0.1146
GS	0.8500	3.5304	2.7982	0.0347
GSA	0.8756	2.9889	2.5521	0.0872
PRACS	0.8793	3.1514	2.5745	0.1355
HPF	0.8704	3.2533	2.6156	0.0228
SFIM	0.8730	3.2031	2.5778	0.0213
Indusion	0.8043	3.9059	3.2846	0.0457
ATWT	0.8791	3.0786	2.5178	0.0978
AWLP	0.8830	2.9424	**2.4073**	0.1031
ATWT-M2	0.8021	4.0493	3.2930	0.2790
ATWT-M3	0.8198	4.3388	3.3357	0.2888
MTF-GLP	0.8787	3.0387	2.5106	0.0683
MTF-GLP-HPM-PP	0.8643	3.3030	2.7540	0.1444
MTF-GLP-HPM	0.8819	3.0041	2.4624	0.0626
MTF-GLP-CBD	0.8780	2.9673	2.5067	0.0702

TABLE VI: ROME DATA SET: QUANTITATIVE RESULTS.

	Q8	SAM(°)	ERGAS	Time[s]
Reference	1	0	0	0
EXP	0.7248	4.9263	5.4171	0
PCA	0.8169	5.2153	4.4128	0.0939
IHS	0.7439	5.1455	4.1691	0.0076
BT	0.7487	4.9263	4.1407	**0.0051**
BDSD	0.8762	4.8717	3.8619	0.0997
GS	0.8335	4.8592	4.0144	0.0487
GSA	0.8907	4.1415	3.4062	0.0990
PRACS	0.8878	4.6678	3.6768	0.3215
HPF	0.8889	4.2813	3.5459	0.0341
SFIM	0.8950	4.0874	3.3979	0.0306
Indusion	0.8030	5.1415	4.8864	0.0700
ATWT	0.9013	4.1117	3.3237	0.1906
AWLP	0.9011	4.5146	3.3572	0.2013
ATWT-M2	0.7969	5.0277	4.6487	0.5510
ATWT-M3	0.8379	5.1042	4.3684	0.5734
MTF-GLP	0.9016	4.0957	3.2982	0.1142
MTF-GLP-HPM-PP	0.8900	4.3736	3.4777	0.2228
MTF-GLP-HPM	**0.9092**	**3.8871**	**3.1005**	0.1170
MTF-GLP-CBD	0.8940	4.1125	3.3479	0.1260

TABLE VII: PLÉIADES DATA SET: QUANTITATIVE RESULTS.

	Q4	SAM(°)	ERGAS	Time[s]
Reference	1	0	0	0
EXP	0.7782	4.6742	6.0826	0
PCA	0.8122	6.1435	6.0028	0.3840
IHS	0.8381	5.1788	5.0549	0.0444
BT	0.8411	4.6742	5.1506	**0.0428**
BDSD	**0.9650**	**4.0749**	**2.7811**	0.3688
GS	0.8448	5.3304	5.0468	0.3197
GSA	0.9572	4.4449	3.0013	0.7095
PRACS	0.9325	4.5157	3.6523	1.6298
HPF	0.9300	4.3125	3.7140	0.2469
SFIM	0.9284	4.2727	3.7128	0.2488
Indusion	0.8251	5.8322	5.6366	0.4018
ATWT	0.9479	4.1401	3.3157	1.5061
AWLP	0.9426	4.3356	3.5219	1.5550
ATWT-M2	0.8300	5.1505	5.4053	4.5053
ATWT-M3	0.8612	5.2915	5.0569	4.6316
MTF-GLP	0.9485	4.1529	3.3193	0.6176
MTF-GLP-HPM-PP	0.8786	5.2343	5.1103	1.0945
MTF-GLP-HPM	0.9488	4.1541	3.2741	0.6082
MTF-GLP-CBD	0.9549	4.4459	3.1521	0.6356

D. Full Resolution Validation

The evaluation of the algorithms at the original resolution of the images was on the Toulouse and the Rio data sets. In Figs. 6 - 7 the fusion results of a small area are shown by using an RGB representation. Tables VIII and IX report the values of the QNR, D_λ and D_S. Many results obtained in the analysis performed at reduced resolution are in line with those obtained

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

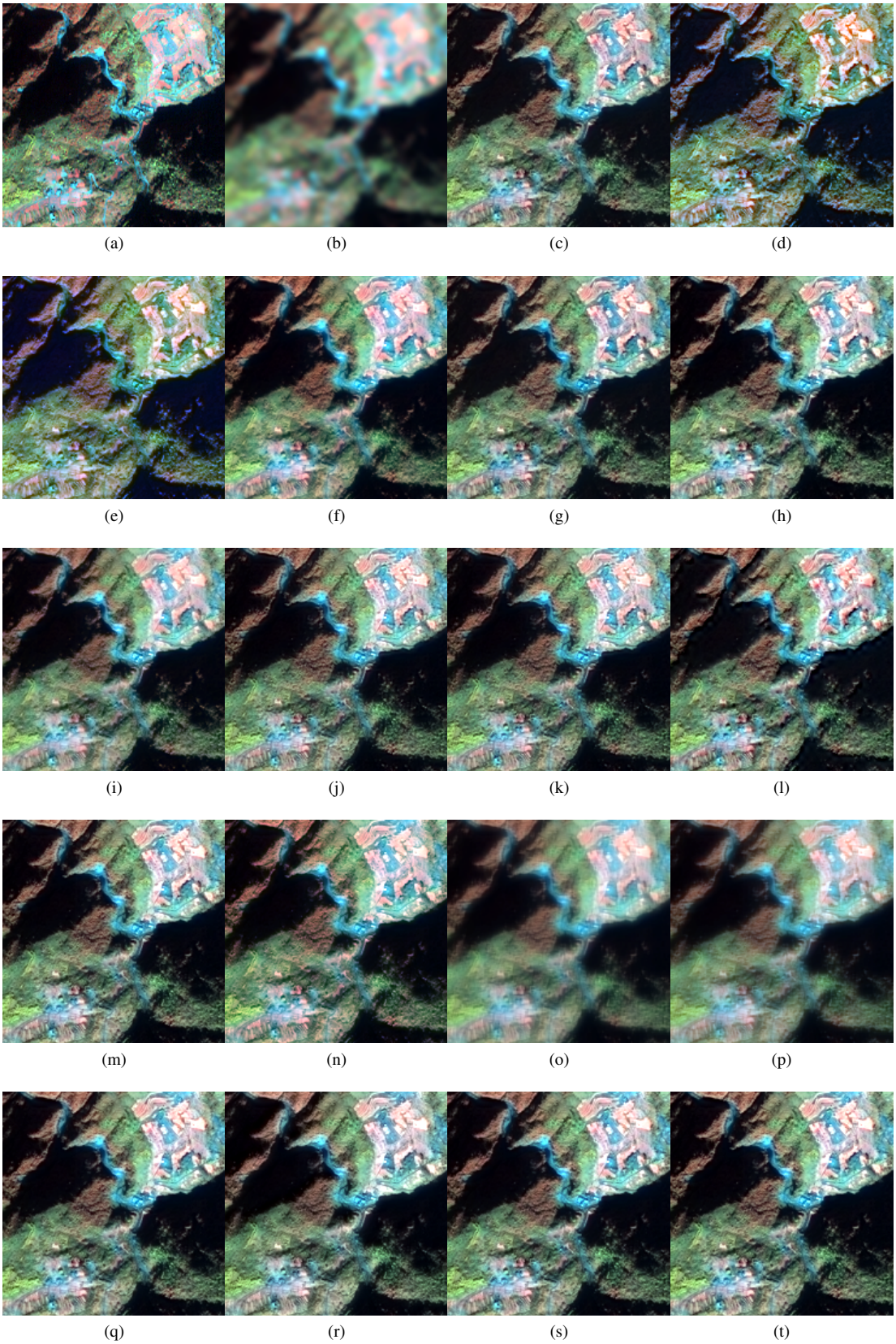

Fig. 3: China data set: (a) Reference Image; (b) EXP; (c) PCA; (d) IHS; (e) BT; (f) BDSD; (g) GS; (h) GSA; (i) PRACS; (j) HPF; (k) SFIM; (l) Indusion; (m) ATWT; (n) AWLP; (o) ATWT-M2; (p) ATWT-M3; (q) MTF-GLP; (r) MTF-GLP-HPM-PP; (s) MTF-GLP-HPM; (t) MTF-GLP-CBD.

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

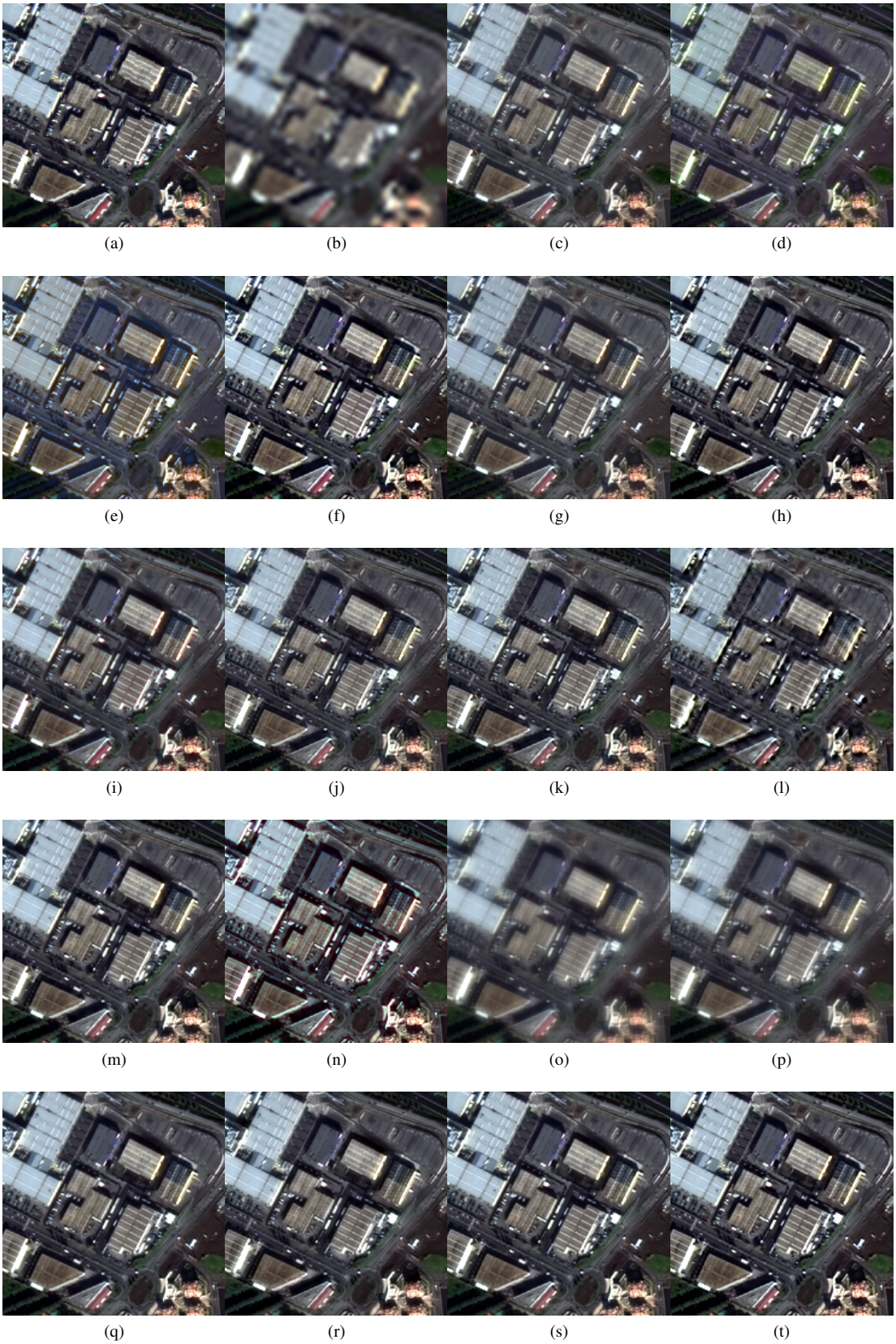

Fig. 4: Rome data set: (a) Reference Image; (b) EXP; (c) PCA; (d) IHS; (e) BT; (f) BDSD; (g) GS; (h) GSA; (i) PRACS; (j) HPF; (k) SFIM; (l) Indusion; (m) ATWT; (n) AWLP; (o) ATWT-M2; (p) ATWT-M3; (q) MTF-GLP; (r) MTF-GLP-HPM-PP; (s) MTF-GLP-HPM; (t) MTF-GLP-CBD.

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

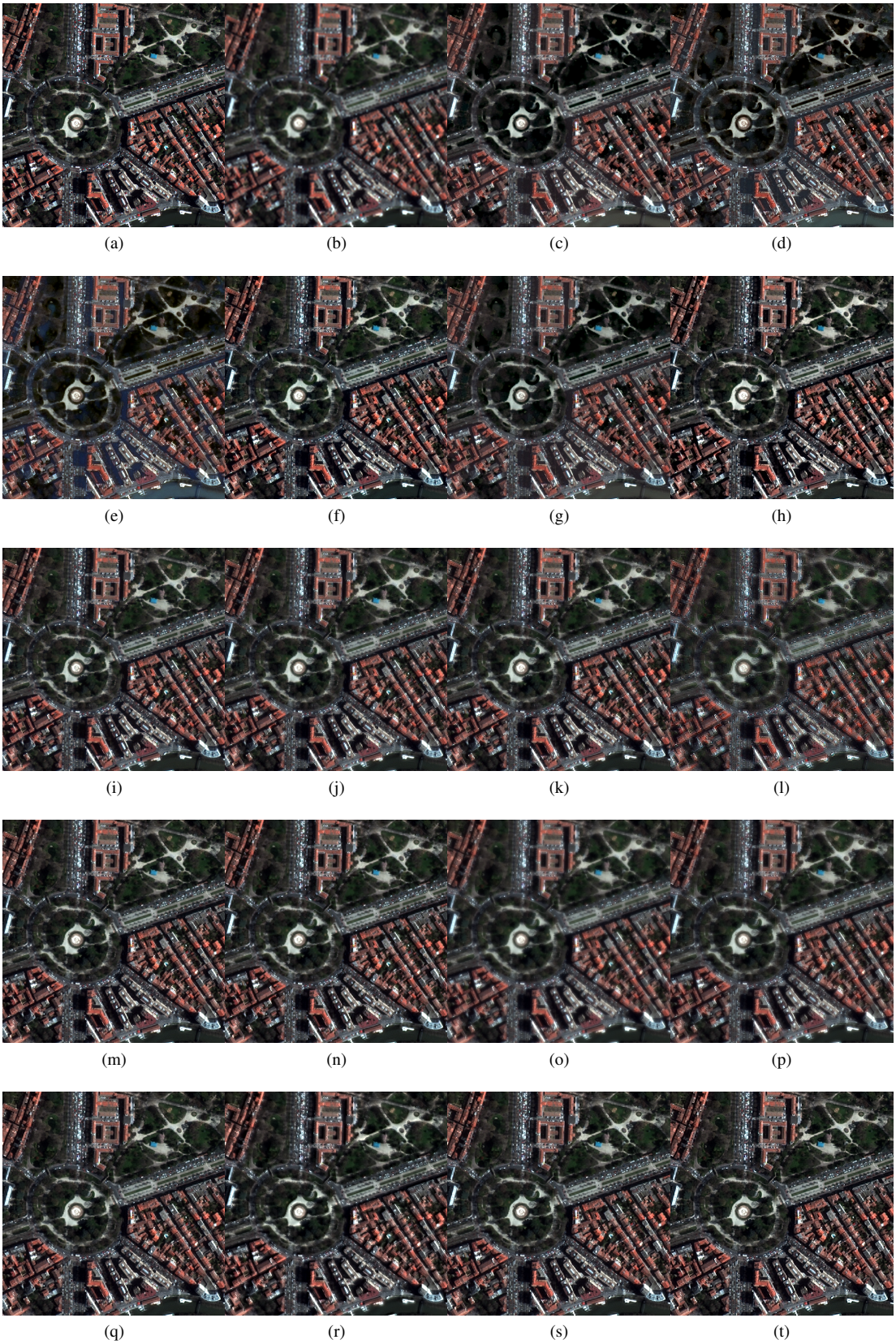

Fig. 5: Pléiades data set: (a) Reference Image; (b) EXP; (c) PCA; (d) IHS; (e) BT; (f) BDSD; (g) GS; (h) GSA; (i) PRACS; (j) HPF; (k) SFIM; (l) Indusion; (m) ATWT; (n) AWLP; (o) ATWT-M2; (p) ATWT-M3; (q) MTF-GLP; (r) MTF-GLP-HPM-PP; (s) MTF-GLP-HPM; (t) MTF-GLP-CBD.

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

at full resolution. In particular, the algorithms within the MRA category performing the best are those based on Gaussian and *"à trous"*. The superiority of the *multiplicative* injection scheme evidenced at reduced resolution is still confirmed here. Within the CS family, the best techniques in terms of obtained accuracy are the *adaptive* algorithms followed by the other CS approaches. As for the reduced resolution, the difference in accuracy between the two main families is more evident in the case of the eight band Rio data set. Again, the algorithms based on IHS show a higher spectral distortion confirmed by the values of D_λ.

It is worth to point out that the PCA and IHS-based approaches perform very well from a quantitative point of view, but this conclusion is not fully confirmed by the visual analysis since the pansharpened images are more spectral distorted if compared to the results of other approaches. Furthermore, a slightly lower accuracy of the full resolution validation procedure can be noted for the Rio data set. Indeed, for instance, the GS works better than the GSA and the AWLP led to lower performances with respect to the results at reduced resolution and the visual interpretation. This is mainly due to the acquisition modality of the WorldView-2 sensor, which captures scenes in eight MS bands arranged in two arrays of four bands each [76]. In fact, it causes misalignment among MS data that affects the effectiveness of the QNR index.

The obtained results give rise to some general considerations. Indeed, the validation at full resolution allows to avoid some issues and assumptions necessary for the reduced resolution protocol, but the values of the indexes are less reliable (since discrepancies between the quantitative values and the visual appearance of the results can be noticed). In more detail, the comparison of algorithms in terms of the QNR index is quite accurate within each category. However, it is more difficult to quantify the existing differences among algorithms belonging to different families.

TABLE VIII: TOULOUSE DATA SET: QUANTITATIVE RESULTS.

	D_λ	D_S	QNR	Time[s]
Reference	0	0	1	0
EXP	0	0.1975	0.8025	0
PCA	0.0114	0.0991	0.8906	0.1320
IHS	0.0367	0.0591	0.9063	0.0157
BT	0.0164	0.0579	0.9266	**0.0133**
BDSD	**0.0079**	0.0404	**0.9520**	0.0900
GS	0.0111	0.0912	0.8987	0.0795
GSA	0.0364	0.0555	0.9101	0.1859
PRACS	0.0137	0.0612	0.9259	0.3877
HPF	0.0301	0.0570	0.9146	0.0631
SFIM	0.0276	0.0561	0.9179	0.0622
Indusion	0.0274	0.0795	0.8952	0.1033
ATWT	0.0348	0.0428	0.9239	0.3491
AWLP	0.0412	0.0392	0.9212	0.3587
ATWT-M2	0.0303	0.0975	0.8752	0.9989
ATWT-M3	0.0544	0.0725	0.8771	1.0206
MTF-GLP	0.0384	0.0406	0.9225	0.1710
MTF-GLP-HPM-PP	0.0440	0.0407	0.9171	0.3075
MTF-GLP-HPM	0.0338	**0.0383**	0.9292	0.1611
MTF-GLP-CBD	0.0401	0.0484	0.9134	0.1670

E. Summary and Discussion

The obtained results give evidence that the good visual appearance and the spectral content preservation represent the main salient features of the CS and MRA methods,

TABLE IX: RIO DATA SET: QUANTITATIVE RESULTS.

	D_λ	D_S	QNR	Time[s]
Reference	0	0	1	0
EXP	0	0.0940	0.9060	0
PCA	0.0265	0.0582	0.9168	0.0857
IHS	0.0454	0.0560	0.9012	0.0080
BT	0.0326	0.0615	0.9080	**0.0054**
BDSD	0.0272	**0.0347**	**0.9390**	0.0672
GS	0.0261	0.0579	0.9176	0.0359
GSA	0.0369	0.0552	0.9100	0.0818
PRACS	**0.0158**	0.0647	0.9205	0.2506
HPF	0.0275	0.0527	0.9212	0.0297
SFIM	0.0261	0.0539	0.9215	0.0260
Indusion	0.0258	0.0600	0.9157	0.0653
ATWT	0.0361	0.0474	0.9182	0.1397
AWLP	0.0399	0.0568	0.9056	0.1484
ATWT-M2	0.0246	0.0906	0.8870	0.4270
ATWT-M3	0.0405	0.0713	0.8911	0.4396
MTF-GLP	0.0359	0.0421	0.9235	0.1060
MTF-GLP-HPM-PP	0.0525	0.0533	0.8970	0.1834
MTF-GLP-HPM	0.0348	0.0423	0.9244	0.0962
MTF-GLP-CBD	0.0375	0.0416	0.9225	0.0961

TABLE X: REDUCED RESOLUTION TOULOUSE DATA SET: QUANTITATIVE RESULTS.

	Q4	SAM(°)	ERGAS
Reference	1	0	0
EXP	0.4471	4.8218	6.4997
PCA	0.7466	3.9984	4.7532
IHS	0.7539	4.5023	4.9580
BT	0.7453	4.8218	5.0976
BDSD	**0.9440**	**2.0005**	**2.3237**
GS	0.7490	4.0063	4.7584
GSA	0.9353	2.0822	2.3892
PRACS	0.9370	2.3895	2.4377
HPF	0.8258	3.5976	4.1117
SFIM	0.8331	3.5878	3.9856
Indusion	0.7496	3.9087	4.7472
ATWT	0.8510	3.4438	3.8901
AWLP	0.8542	3.9489	4.1197
ATWT-M2	0.6140	4.2949	5.4643
ATWT-M3	0.7483	4.2257	4.8739
MTF-GLP	0.8649	3.3678	3.7513
MTF-GLP-HPM-PP	0.8330	3.7058	4.0392
MTF-GLP-HPM	0.8758	3.3658	3.5523
MTF-GLP-CBD	0.9350	2.0851	2.3795

TABLE XI: REDUCED RESOLUTION RIO DATA SET: QUANTITATIVE RESULTS.

	Q8	SAM(°)	ERGAS
Reference	1	0	0
EXP	0.6777	3.2843	5.0147
PCA	0.9017	2.5582	2.8980
IHS	0.8798	3.3457	3.1985
BT	0.8677	3.2843	3.0479
BDSD	0.8902	2.9798	2.9889
GS	0.9002	2.5672	2.9144
GSA	0.9286	2.1496	2.1178
PRACS	0.9363	2.5438	2.1188
HPF	0.9270	2.3713	2.3908
SFIM	0.9272	2.4329	2.3350
Indusion	0.8800	2.5508	3.7205
ATWT	0.9418	2.2471	1.9419
AWLP	0.9308	3.4671	2.4004
ATWT-M2	0.8426	3.0265	3.7720
ATWT-M3	0.8803	3.0677	3.3642
MTF-GLP	**0.9427**	2.2151	1.8878
MTF-GLP-HPM-PP	0.9270	2.5411	2.3516
MTF-GLP-HPM	0.9426	2.3141	**1.8834**
MTF-GLP-CBD	0.9302	**2.0906**	2.0705

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

Fig. 6: Toulouse data set: (a) PAN; (b) EXP; (c) PCA; (d) IHS; (e) BT; (f) BDSD; (g) GS; (h) GSA; (i) PRACS; (j) HPF; (k) SFIM; (l) Indusion; (m) ATWT; (n) AWLP; (o) ATWT-M2; (p) ATWT-M3; (q) MTF-GLP; (r) MTF-GLP-HPM-PP; (s) MTF-GLP-HPM; (t) MTF-GLP-CBD.

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

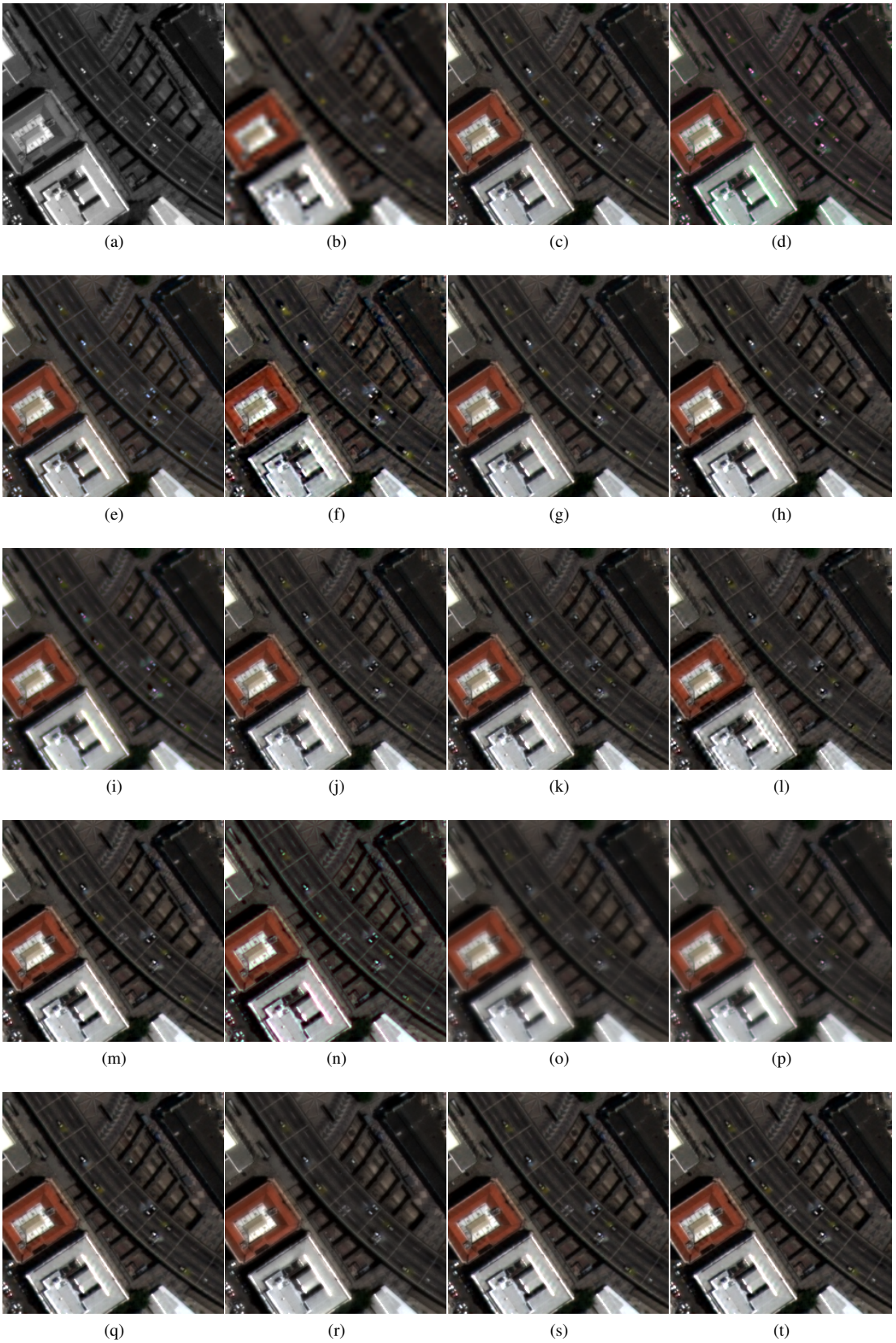

Fig. 7: Rio data set: (a) PAN; (b) EXP; (c) PCA; (d) IHS; (e) BT; (f) BDSD; (g) GS; (h) GSA; (i) PRACS; (j) HPF; (k) SFIM; (l) Indusion; (m) ATWT; (n) AWLP; (o) ATWT-M2; (p) ATWT-M3; (q) MTF-GLP; (r) MTF-GLP-HPM-PP; (s) MTF-GLP-HPM; (t) MTF-GLP-CBD.

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

TABLE XII: K-MEANS ($K = 3$) CLUSTERING OF PANSHARPENING ALGORITHMS APPLIED TO IKONOS DATA INTO HIGH, MEDIUM, AND LOW PERFORMANCE. INPUT DATA FOR CLUSTERING ARE THE REDUCED-RESOLUTION SCORE INDEXES IN TABLE X (LEFT) AND THE FULL-RESOLUTION SCORE INDEXES IN TABLE VIII (RIGHT).

High	Medium	Low
BDSD	**MTF-GLP-HPM**	**IHS**
PRACS	**MTF-GLP**	**Indusion**
GSA	**AWLP**	**GS**
MTF-GLP-CBD	**ATWT**	**ATWT-M3**
	SFIM	**PCA**
	MTF-GLP-HPM-PP	**BT**
	HPF	**ATWT-M2**

High	Medium	Low
BDSD	**MTF-GLP-HPM**	**GS**
PRACS	**BT**	**Indusion**
	ATWT	**PCA**
	MTF-GLP	**ATWT-M3**
	AWLP	**ATWT-M2**
	SFIM	
	MTF-GLP-HPM-PP	
	HPF	
	MTF-GLP-CBD	
	GSA	
	IHS	

TABLE XIII: K-MEANS ($K = 3$) CLUSTERING OF PANSHARPENING ALGORITHMS APPLIED TO WORLDVIEW-2 DATA INTO HIGH, MEDIUM, AND LOW PERFORMANCE. INPUT DATA FOR CLUSTERING ARE THE REDUCED-RESOLUTION SCORE INDEXES IN TABLE XI (LEFT) AND THE FULL-RESOLUTION SCORE INDEXES IN TABLE IX (RIGHT).

High	Medium	Low
MTF-GLP	**SFIM**	**ATWT-M3**
MTF-GLP-HPM	**HPF**	**Indusion**
ATWT	**PRACS**	**IHS**
AWLP	**PCA**	**BT**
MTF-GLP-CBD	**GS**	**ATWT-M2**
MTF-GLP-HPM-PP	**BDSD**	
GSA		

High	Medium	Low
BDSD	**GS**	**BT**
MTF-GLP-HPM	**PCA**	**AWLP**
MTF-GLP	**Indusion**	**IHS**
MTF-GLP-CBD	**GSA**	**MTF-GLP-HPM-PP**
SFIM		**ATWT-M3**
HPF		**ATWT-M2**
PRACS		
ATWT		

respectively. Accordingly, approaches of the first class aimed at improving the spectral quality and those of the second class properly designed for enhancing the spatial properties obtain the best results. Indeed, very interesting performances are attained by adaptive CS approaches (in particular on four band data sets), with the reduction of the spectral distortion, and by some MRA algorithms, which benefit from proper details extraction. In particular the match of the low-pass filter with the sensor MTF allows to significantly reduce the classical blur of the MRA final products.

Another interesting consideration arises from the performed tests. Specifically, for the case of eight band data sets, degradation of the performances for certain CS approaches is clear from both a numerical and visual point of view. This is mainly due to the larger spectral distortion affecting the results. This can be justified by the larger spectral misalignments between the MS and the PAN spectra and by the acquisition modality of the sensor (it captures scenes in eight MS bands arranged in two arrays of four bands each [76]). In fact, this acquisition modality can lead to small temporal misalignments among MS data, which, even though partially compensated, affect the effectiveness of the linear regression step usually exploited by the most powerful CS-based methods. For these reasons, the use of MRA approaches is, in general, as much advisable as the number of the bands gets greater.

Furthermore, it is worth underlining that CS approaches are often preferable with respect to the MRA ones thanks to their robustness to aliasing and misregistration errors. In fact, since they employ an aliasing-free version of the PAN image, the fusion rule adopted by the CS algorithms is able to compensate the aliasing present in the MS image. On the contrary, the aliasing patterns are visible when the MRA fusion rule is adopted [51]. However, such problem can be strongly reduced by employing MRA decimated approaches whose analysis filter is matched with the MTF of the MS sensor, as obtainable by properly designing Laplacian pyramids [58]. Temporal misalignments suggest the use of MRA methodologies with respect to CS ones [77]. This kind of robustness turns out to be very helpful when the images are acquired with a time delay, as, for example, when they are provided from sensors mounted on different remote sensing platforms.

From a computational point of view, the CS approaches are surely preferable to the MRA ones, since the filtering phase significantly slows down the algorithms. According to our tests, which were performed on a Pentium i7 Processor, 3.20 GHz, the former ones are about ten times faster than the latter ones, as evidenced by the computational times reported in Tables V-IX.

Final considerations arise from the comparison of the algorithms at reduced and full resolution. To this aim we assessed the performances of the analyzed methods at reduced resolution also with the reference used in the full resolution analysis for the Toulouse and Rio data sets. The obtained quantitative results at reduced and full resolution are reported in Tables X, XI and Tables VIII, IX, respectively. For these two data sets we grouped the pansharpening algorithms into three classes according to their performances (namely high, medium, and low performance methods). The results of this partition are shown in Tables XII and XIII referring to the IKONOS and WorldView-2 data sets, respectively. The obtained clusters correspond to a rather natural classification, as testified for example by the analysis of the $Q4$ histogram shown in Fig. 8 for the IKONOS case, which is reasonably described by a three-mode distribution. Analogous considera-

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

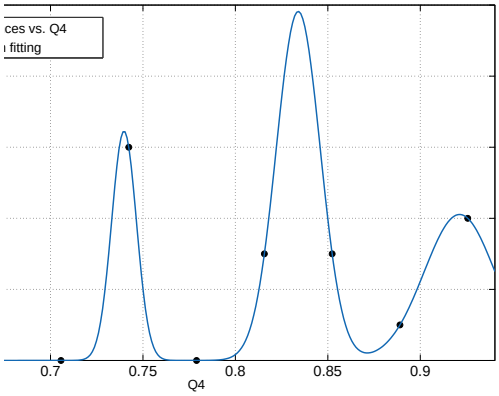

Fig. 8: Gaussian fitting of histogram of $Q4$ in Table X.

tions can be done both for the QNR index distribution and for the WorldView-2 case. Analogously, the 18 considered algorithms are automatically grouped into high, medium and low performance methods through the K-means procedure starting from the score indexes computed at reduced resolution in Tables X and XI, and at full resolution in Tables VIII and IX, which have been conveniently normalized with respect to their corresponding maximum values. The methods inside each group are subsequently ordered according to the values of the $Q2^n$ at reduced resolution, while, at full resolution the QNR is considered. This simple and automatic procedure allows one to easily identify the pansharpening methods that perform best for each sensor. The good match between the performances evaluated for both validation protocols can be appreciated. Indeed, the majority of the algorithms belong to the same performance class at the reduced and full resolution validation. Furthermore, it is worth pointing out that the worst and the best approaches are quite sensor independent. For instance, earlier methods as PCA, GS, ATWT-M2 and ATWT-M3 are poorly performing when considering images acquired by both the IKONOS and WorldView-2 sensors. Whereas more advanced ones, as BDSD and MTF-GLP-CBD evince their superiority, regardless the sensors. Moreover, from this classification, it is again possible to assess the superiority of the MRA approaches with respect to the CS ones for the WorldView-2 data set.

The analysis carried out in this work allowed us to confirm some features of the validation procedures: The reduced resolution protocol leads to a very accurate evaluation of the quality indexes, but the scale invariance hypothesis is not always verified in practice. Furthermore, the operation for generating a synthetic version of the MS image at reduced resolution introduces a strong bias in the analyzed algorithms, privileging those employing a similar procedure for extracting the spatial details. On the contrary, the full resolution validation gets quite accurate results with respect to algorithms of the same family, whereas being less reliable when the comparison includes algorithms belonging to different classes. This validation approach, which exploits the QNR index, has the advantage of validating the products at the original scale, thus avoiding any hypothesis on the behavior at different scales. However, due to the lower reliability of the quality indexes (since a reference is missing), the results of this analysis can be affected by some mismatches between the quantitative results and the visual appearance of the fused image (e.g., as for the IHS method).

V. Conclusions

In this paper, a comparison among several pansharpening algorithms presented in the literature has been performed. Specifically, a *hard* classification of methods into two main families, i.e., based on component substitution (CS) or multiresolution analysis (MRA), has been adopted and some widely used algorithms belonging to these families have been described and evaluated through extensive simulations.

Due to the differences in spectral and spatial characteristics of the products to fuse, the validation of the final product represents a difficult task. For this reason, the paper has been focused on the validation of pansharpening products. The different behaviors of the same algorithms under different validation procedures, on data sets acquired by different satellites, have been described. The two approaches currently considered in the literature for the assessment, i.e., analysis at reduced and at full resolutions, have been considered. The former follows Wald's protocol checking the *synthesis* property. To this aim the original image is used as a reference, whereas the pansharpening algorithms are applied to reduced resolution versions. Thanks to the availability of a target product, quality indexes quantifying the similarity of the fused product to its reference can be employed. Typically, indexes taking into account both radiometric and spectral distortions, e.g., $Q4$ and its extension $Q2^n$, are preferable in this phase. On the contrary, the full resolution validation protocol avoids reducing the resolution of the input images. Thus, since no reference image can be considered, appropriate indexes requiring no reference have to be exploited to verify the accuracy of the final product, e.g., the Quality w/No Reference protocol.

The particular characteristics of the two classes of pansharpening algorithms have been evidenced during the assessment phase. Specifically, the visually appealing features of CS methods have been highlighted by the absence of aliasing impairments. Such a favorable characteristic, together with the robustness of these methods to the errors induced by possible misregistration between the available multispectral and panchromatic data sets and a relatively low computational burden, has supported their widespread use. On the other side, very good overall performances are generally attained by MRA methods, which are characterized by a superior reproduction of the multispectral features and can be easily designed to match the behavior of the MS sensor for the extraction of spatial details. Furthermore, this class of algorithms is currently drawing an increasing attention; thanks to their temporal coherence, they can be employed for multiplatform data fusion.

The simulations presented in this paper have been performed by a MatLab implementation of the algorithms, quality indexes and validation procedures. The authors have chosen to make the MatLab toolbox they have developed available to the community. This allows for fair and easy comparisons with

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

some of the most widely used state of the art algorithms. They also hope that this toolbox could foster the development and validation, through the benchmarking with established algorithms on standardized data sets, of novel and ever more performing pansharpening methods.

REFERENCES

[1] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, "Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1301–1312, May 2008.

[2] C. Souza, Jr., L. Firestone, L. M. Silva, and D. Roberts, "Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models," *Remote Sens. Environ.*, vol. 87, no. 4, pp. 494–506, Nov. 2003.

[3] A. Mohammadzadeh, A. Tavakoli, and M. J. Valadan Zoej, "Road extraction based on fuzzy logic and mathematical morphology from pansharpened IKONOS images," *Photogramm. Rec.*, vol. 21, no. 113, pp. 44–60, Mar. 2006.

[4] F. Laporterie-Déjean, H. de Boissezon, G. Flouzat, and M.-J. Lefèvre-Fonollosa, "Thematic and statistical evaluations of five panchromatic/multispectral fusion methods on simulated PLEIADES-HR images," *Inform. Fusion*, vol. 6, no. 3, pp. 193–212, Sep. 2005.

[5] 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, Jun. 2005.

[6] I. Amro, J. Mateos, M. Vega, R. Molina, and A. K. Katsaggelos, "A survey of classical methods and new trends in pansharpening of multispectral images," *EURASIP J. Adv. Signal Process.*, vol. 2011, no. 1, pp. 79:1–79:22, Sep. 2011.

[7] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "Twenty-five years of pansharpening: A critical review and new developments," in *Signal and Image Processing for Remote Sensing*, 2nd ed., C.-H. Chen, Ed. Boca Raton, FL, USA: CRC Press, 2012, pp. 533–548.

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

[9] W. Carper, T. Lillesand, and R. Kiefer, "The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data," *Photogramm. Eng. Remote Sens.*, vol. 56, no. 4, pp. 459–467, Apr. 1990.

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

[11] P. S. Chavez, Jr. and A. W. Kwarteng, "Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis," *Photogramm. Eng. Remote Sens.*, vol. 55, no. 3, pp. 339–348, Mar. 1989.

[12] V. K. Shettigara, "A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set," *Photogramm. Eng. Remote Sens.*, vol. 58, no. 5, pp. 561–567, May 1992.

[13] V. P. Shah, N. H. Younan, and R. L. King, "An efficient pan-sharpening method via a combined adaptive-PCA approach and contourlets," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1323–1335, May 2008.

[14] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," 2000, U.S. Patent # 6,011,875.

[15] Q. Xu, B. Li, Y. Zhang, and L. Ding, "High-fidelity component substitution pansharpening by the fitting of substitution data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 7380–7392, Nov. 2014.

[16] X. Kang, S. Li, and J. A. Benediktsson, "Pansharpening with matting model," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5088–5099, Aug. 2014.

[17] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[18] G. P. Nason and B. W. Silverman, "The stationary wavelet transform and some statistical applications," in *Wavelets and Statistics*, A. Antoniadis and G. Oppenheim, Eds. New York, NY, USA: Springer-Verlag, 1995, vol. 103, pp. 281–299.

[19] M. J. Shensa, "The discrete wavelet transform: Wedding the à trous and Mallat algorithm," *IEEE Trans. Signal Process.*, vol. 40, no. 10, pp. 2464–2482, Oct. 1992.

[20] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM–31, no. 4, pp. 532–540, Apr. 1983.

[21] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.

[22] J.-L. Starck, J. Fadili, and F. Murtagh, "The undecimated wavelet decomposition and its reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 297–309, Feb. 2007.

[23] J. C. Price, "Combining panchromatic and multispectral imagery from dual resolution satellite instruments," *Remote Sens. Environ.*, vol. 21, no. 2, pp. 119–128, Mar. 1987.

[24] D. Fasbender, J. Radoux, and P. Bogaert, "Bayesian data fusion for adaptable image pansharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1847–1857, Jun. 2008.

[25] F. Palsson, J. R. Sveinsson, and M. O. Ulfarsson, "A new pansharpening algorithm based on total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 318–322, Jan. 2014.

[26] X. He, L. Condat, J. Bioucas-Dias, J. Chanussot, and J. Xia, "A new pansharpening method based on spatial and spectral sparsity priors," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 4160–4174, Sep. 2014.

[27] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[28] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[29] S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 738–746, Feb. 2011.

[30] X. X. Zhu and R. Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2827–2836, May 2013.

[31] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.

[32] Z. Pan, J. Yu, H. Huang, S. Hu, A. Zhang, H. Ma, and W. Sun, "Super-resolution based on compressive sensing and structural self-similarity for remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4864–4876, Sep. 2013.

[33] A. Garzelli, L. Capobianco, L. Alparone, B. Aiazzi, and S. Baronti, "Hyperspectral pansharpening based on modulation of pixel spectra," in *Proc. 2nd WHISPERS*, 2010, pp. 1–4.

[34] G. A. Licciardi, M. M. Khan, J. Chanussot, A. Montanvert, L. Condat, and C. Jutten, "Fusion of hyperspectral and panchromatic images using multiresolution analysis and nonlinear PCA band reduction," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, pp. 207:1–207:17, Sep. 2012.

[35] M. Ehlers, S. Klonus, P. J. Astrand, and P. Rosso, "Multi-sensor image fusion for pansharpening in remote sensing," *Int. J. Image Data Fusion*, vol. 1, no. 1, pp. 25–45, Feb. 2010.

[36] 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, no. 6, pp. 691–699, Jun. 1997.

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

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

[39] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, "Multispectral and panchromatic data fusion assessment without reference," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200, Feb. 2008.

[40] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. ICIP*, vol. 2, 2003, pp. III – 173–176.

[41] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

[42] C. Thomas and L. Wald, "Analysis of changes in quality assessment with scale," in *Proc. 9th Int. Conf. Inf. Fusion*, 2006, pp. 1–5.

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

[44] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," *Inform. Fusion*, vol. 2, no. 3, pp. 177–186, Sep. 2001.

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

[45] 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 Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 309–312, Oct. 2004.

[46] W. Dou, Y. Chen, X. Li, and D. Sui, "A general framework for component substitution image fusion: An implementation using fast image fusion method," *Comput. Geosci.*, vol. 33, no. 2, pp. 219–228, Feb. 2007.

[47] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 295–309, Jan. 2011.

[48] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly correlated images-II. Channel ratio and "Chromaticity" Transform techniques," *Remote Sens. Environ.*, vol. 22, no. 3, pp. 343–365, Aug. 1987.

[49] A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE pan sharpening of very high resolution multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 228–236, Jan. 2008.

[50] B. Aiazzi, S. Baronti, M. Selva, and L. Alparone, "Bi-cubic interpolation for shift-free pan-sharpening," *ISPRS J. Photogramm. Remote Sens.*, vol. 86, no. 6, pp. 65–76, Dec. 2013.

[51] S. Baronti, B. Aiazzi, M. Selva, A. Garzelli, and L. Alparone, "A theoretical analysis of the effects of aliasing and misregistration on pansharpened imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 446–453, Jun. 2011.

[52] J. Jolliffe, *Principal Component Analysis*. Hoboken, NJ, USA: Wiley, 2005, Online Library.

[53] B. Aiazzi, S. Baronti, F. Lotti, and M. Selva, "A comparison between global and context-adaptive pansharpening of multispectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 302–306, Apr. 2009.

[54] T. Ranchin and L. Wald, "Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 1, pp. 49–61, Jan. 2000.

[55] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*, 2nd ed. Orlando, FL, USA: Academic Press, 1997.

[56] G. Vivone, R. Restaino, M. Dalla Mura, G. Licciardi, and J. Chanussot, "Contrast and error-based fusion schemes for multispectral image pansharpening," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp. 930–934, May 2014.

[57] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "An MTF-based spectral distortion minimizing model for pan-sharpening of very high resolution multispectral images of urban areas," in *Proc. 2nd GRSS/ISPRS Joint Workshop Remote Sens. Data Fusion URBAN Areas*, 2003, pp. 90–94.

[58] ——, "Advantages of Laplacian pyramids over "à trous" wavelet transforms," in *Proc. SPIE Image Signal Process. Remote Sens. XVIII*, L. Bruzzone, Ed., vol. 8537, 2012, pp. 853 704–1–853 704–10.

[59] J. G. Liu, "Smoothing filter based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, Dec. 2000.

[60] L. Wald and T. Ranchin, "Comment: Liu 'Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details'," *Int. J. Remote Sens.*, vol. 23, no. 3, pp. 593–597, Jan. 2002.

[61] M. M. Khan, J. Chanussot, L. Condat, and A. Montavert, "Indusion: Fusion of multispectral and panchromatic images using the induction scaling technique," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 98–102, Jan. 2008.

[62] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and Pan imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.

[63] 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, Oct. 2005.

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

[65] J. Lee and C. Lee, "Fast and efficient panchromatic sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 155–163, Jan. 2010.

[66] L. Amolins, Y. Zhang, and P. Dare, "Wavelet based image fusion techniques - An introduction, review and comparison," *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 4, pp. 249–263, Sep. 2007.

[67] B. Garguet-Duport, J. Girel, J.-M. Chassery, and G. Pautou, "The use of multiresolution analysis and wavelet transform for merging SPOT Panchromatic and multispectral image data," *Photogramm. Eng. Remote Sens.*, vol. 62, no. 9, pp. 1057–1066, Sep. 1996.

[68] J. Núñez, X. Otazu, O. Fors, A. Prades, V. Palà, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1204–1211, May 1999.

[69] M. González-Audícana, X. Otazu, O. Fors, and A. Seco, "Comparison between Mallat's and the 'à trous' discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images," *Int. J. Remote Sens.*, vol. 26, no. 3, pp. 595–614, Feb. 2005.

[70] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ, USA: Prentice Hall, 1995.

[71] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, 2nd ed. Wellesley, MA: Wellesley Cambridge Press, 1996.

[72] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the Spectral Angle Mapper (SAM) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*, 1992, pp. 147–149.

[73] L. Wald, *Data Fusion: Definitions and Architectures — Fusion of images of different spatial resolutions*. Paris, France: Les Presses de l'École des Mines, 2002.

[74] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.

[75] A. Garzelli and F. Nencini, "Hypercomplex quality assessment of multi-/hyper-spectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 662–665, Oct. 2009.

[76] T. Updike and C. Comp, "Radiometric use of WorldView-2 imagery," DigitalGlobe, Longmont, CO, USA, Tech. Rep., Nov. 2010.

[77] B. Aiazzi, L. Alparone, S. Baronti, R. Carlà, A. Garzelli, L. Santurri, and M. Selva, "Effects of multitemporal scene changes on pansharpening fusion," in *Proc. 6th IEEE Int. Workshop Anal. MultiTemp. Remote Sens.*, 2011, pp. 73–76.