Pansharpening Based on Semiblind Deconvolution

Gemine Vivone, Miguel Simões, Mauro Dalla Mura, *Member, IEEE*, Rocco Restaino, *Member, IEEE*, José M. Bioucas-Dias, *Member, IEEE*, Giorgio A. Licciardi, *Member, IEEE*, and Jocelyn Chanussot, *Fellow, IEEE*

***Abstract*—Many powerful pansharpening approaches exploit the functional relation between the fusion of PANchromatic (PAN) and MultiSpectral (MS) images. To this purpose, the modulation transfer function of the MS sensor is typically used, being easily approximated as a Gaussian filter whose analytic expression is fully specified by the sensor gain at the Nyquist frequency. However, this characterization is often inadequate in practice. In this paper, we develop an algorithm for estimating the relation between PAN and MS images directly from the available data through an efficient optimization procedure. The effectiveness of the approach is validated both on a reduced scale data set generated by degrading images acquired by the IKONOS sensor and on full-scale data consisting of images collected by the QuickBird sensor. In the first case, the proposed method achieves performances very similar to that of the algorithm that relies upon the full knowledge of the degrading filter. In the second, it is shown to outperform several very credited state-of-the-art approaches for the extraction of the details used in the current literature.**

***Index Terms*—Data fusion, deconvolution, modulation transfer function (MTF), pansharpening, remote sensing.**

I. Introduction

In RECENT years, pansharpening has become a task of great importance in the field of data fusion, as demonstrated by the increasing number of scientific contributions to this topic and by the success of a contest issued by the IEEE Geoscience and Remote Sensing Society [1]. Pansharpening addresses the fusion of two optical remote sensing images characterized by different spectral and spatial features. Specifically, a MultiSpectral (MS) image with high spectral but low spatial resolution is considered along with a PANchromatic (PAN) image, which is obtained by sensing a single wide electromagnetic spectrum covering the visible and near infrared (VNIR) frequencies and has complementary characteristics with respect to MS: lower spectral but greater spatial resolution.

The objective of pansharpening algorithms is the generation of a fused product characterized by the spectral content of the MS image and the spatial details of the PAN image. The product generated by pansharpening finds its use in many applications. In particular, the demand for these products is becoming more frequent both for user-oriented commercial products, such as Google Earth and Microsoft Bing Map, and as a preliminary step for many automated signal processing methodologies (e.g., change detection [2]). Since a single imaging device is not able to achieve both the required spectral and spatial performances in resolution due to physical constraints and practical reasons (e.g., limited onboard storage capabilities), a raw data fusion step is mandatory.

Nowadays, many pansharpening algorithms exist in the literature. It is possible to group them into two main families [3]: 1) the methods based on the projection of the MS image into a new space and the substitution of a component with a histogram matched version of the PAN image (the so-called *Component Substitution* class) and 2) the approaches based on the extraction of spatial details from the PAN image and their injection into the MS one (this class is called *MultiResolution Analysis* because details are usually obtained through a multiscale decomposition of the original image). Popular examples of the former are intensity–hue–saturation [4], principal component analysis [5], [6] and Gram–Schmidt (GS) [7] techniques, while examples of the latter are simple spatial low-pass filters such as Box filters [4] or more complex decompositions based on Laplacian [8] or wavelet transformations [9]. A more systematic overview of pansharpening methodologies belonging to these two categories can be found in [3] and [10]. Moreover, in the last years, approaches that do not fit in this classification have started to appear in the literature. Among those, we can list Bayesian methods based on parameter estimation [11], total variation penalization terms [12], and sparse signal representation (or compressive sensing theory) [13], [14].

This paper is focused on a particular issue that is crucial in many pansharpening approaches: the characterization of the relationship between the MS and the PAN image through the *modulation transfer function* (*MTF*) of the MS sensor. We recall that the MTF is defined as the absolute value of the

Manuscript received January 16, 2014; revised June 27, 2014; accepted August 12, 2014. This work was supported in part by the Portuguese Science and Technology Foundation under Projects PEst-OE/EEI/LA0008/2013 and PTDC/EEIPRO/1470/2012.

G. Vivone is with the North Atlantic Treaty Organization Science and Technology Organization Centre for Maritime Research and Experimentation, 19126 La Spezia, Italy (e-mail: vivone@cmre.nato.int).

M. Simões is with the Instituto de Telecomunicações, Instituto Superior Técnico, 1049-001 Lisbon, Portugal, and also with Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique (GIPSA-Lab), Grenoble Institute of Technology, 38402 Saint Martin d'Heres Cedex, France (e-mail: miguel.simoes@gipsa-lab.grenoble-inp.fr).

M. Dalla Mura and G. A. Licciardi are with Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique (GIPSA-Lab), Grenoble Institute of Technology, 38402 Saint Martin d'Heres Cedex, France (e-mail: mauro.dalla-muragipsa-lab.grenoble-inp.fr; giorgio-antonino.licciardi@gipsa-lab.grenoble-inp.fr).

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

J. M. Bioucas-Dias is with the Instituto de Telecomunicações, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal (e-mail: bioucas@lx.it.pt).

J. Chanussot is with Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique (GIPSA-Lab), Grenoble Institute of Technology, 38402 Saint Martin d'Heres Cedex, 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).

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.2014.2351754

Fourier transform of the sensor's *point spread function* (*PSF*).[1] In this regard, the lower spatial resolution of the MS with respect to the one of the PAN can be modeled as the effect of a convolution of the former with a filter having the MTF as amplitude response. This assumption relies upon the hypothesis of ideality of the frequency response of the PAN sensor (i.e., it does not introduce any blur) and of the interpolation applied to the MS for upsampling. The well-established pansharpening technique described in [15] is based on this model. In more details, this technique employs a classical scheme consisting of two parts: 1) detail extraction, in which the spatial content missing in the MS is extracted from the PAN image, and 2) detail injection, in which such information is introduced into the MS image. Specifically, in [15], the suitability of using a filter matched to the MS sensor's MTF for detail extraction is justified by the purpose of restoring all the spatial details not resolved by the MS but visible in the PAN. Knowing the relation between the MS and the PAN is also necessary to pansharpening algorithms that require a version of the PAN image at the resolution of the original MS image, such as the band-dependent spatial detail [16], or the GS mode 2 [7]. In addition, we cite the approach recently proposed in [17] and based on compressive sensing theory [18], which exploits the PAN–MS relationship for ensuring the consistency of the high resolution dictionary with the low resolution available data. Moreover, even some quality assessment procedures rely on the knowledge of this relation [19].

The design of a filter matching the real (and unknown) MTF is not straightforward. Gaussian-shaped filters are usually employed to reproduce the MTF of the MS [20], [21], whose form closely resembles a Normal bell in many optical acquisition systems [20]. Thus, in this approximation, the definition of the filter response is given by the setting of a single parameter (i.e., the standard deviation). Usually, the value of the standard deviation of the filter is determined from the gain at the Nyquist frequency since this latter is often provided by the sensor's manufacturer (derived by construction specifications or, more properly, by onboard measurements) [15], [21]. Alternatively, "à trous" wavelet transforms have been also used for modeling the MTF [22]. Indeed, it was seen that, when these filters are designed for matching the sensor specifications, their shape becomes very close to a Gaussian [15].

A. Motivations

Unfortunately, in many cases, resorting to the gain at the Nyquist frequency does not lead to a specification of the sensor MTF with the desirable accuracy. A first issue concerns the modeling of the MTF by a Gaussian-shaped filter, which is sometimes unconfirmed in practice. Even when it constitutes a reasonable approximation, its specification through the MTF gain at the Nyquist frequency is not always possible since it is unavailable for many sensors. Even when available, the gain at the Nyquist frequency provided by the sensor manufacturer has typically been measured just after the launch of the sensor. However, due to the aging of the optical and electronics payload, this may change. A final concern regards the symmetry hypothesis of the MTF; in fact, for several sensors, the MTF shape varies between the along-track and cross-track directions [23], and thus, it cannot be fully characterized through the single value of the standard deviation [15].

In this paper, we propose to estimate the MTF by modeling the relation between the PAN and the MS image directly from the data in order to overcome the aforementioned limitations. To the best of our knowledge, the pansharpening literature lacks in methods that propose to estimate the filter used for the extraction of the PAN details only by considering the images themselves. The comparison between the proposed approach and those based on the determination of the MTF from the sensor knowledge is thus a very intriguing topic to investigate.

In order to perform the estimation of the filter, we assume that each MS band is a spatially degraded version of the PAN image, all blurred by the same filter, and we model the estimation as a problem of blind image deblurring. We recall that blind image deblurring, a classical challenge in image processing, is the inverse problem of recovering sharp images from degraded ones by both estimating the blur affecting the image and the target high resolution (deblurred) image. In general, this problem is hard to solve since it is ill-posed (i.e., there are multiple pairs of estimated image and blurring operator that can produce the degraded image) and the blurring operators are usually very ill-conditioned in practice. For these reasons, in order to obtain a well-defined solution, additional hypotheses (or *a priori* information) about the original image and/or the blurring operator are needed. The problem can thus be recasted into the Bayesian framework or as a regularization problem, both leading to a similar optimization problem form (please refer to [24] and [25] and to the more recent paper [26] for a review of the topic). In this paper, only the estimation of the blurring operator is required (i.e., the filter approximating the MTF of the MS sensor), thus making the solution of the problem significantly simpler.

B. Contribution and Outline

Starting from the motivations presented earlier, the objectives of this paper are as follows: 1) provide an estimation procedure for the degradation filter based only on the acquired images, with the objective of obtaining an adapted system for detail extraction from the PAN image, and 2) compare this approach with techniques for detail extraction used in state-of-the-art pansharpening methods and, in particular, with those designed by some specific knowledge of the sensor MTF. To this end, two different validation procedures are performed. The first one aims at evidencing the capabilities of the estimation procedure to properly approximate the actual response of the imaging systems; the second one points out the advantages of extracting details through the filters derived according to the proposed procedure, instead of those employed in classical approaches.

The remaining sections of this paper are organized as follows. In Section II, the formulation of the estimation problem and its solution are provided. In Section III, we will describe the employed algorithm for spatially enhancing the MS image, which is based on a well-known injection model, while in

[1]The PSF is the inverse Fourier transform of the optical transfer function, which is the product between the phase transfer function (PTF) and the MTF.

Section IV, the two validation procedures are fully described and exploited to assess the performances of the proposed method with respect to state-of-the-art methods. In the end, in Section V, the conclusions and future developments are drawn.

C. Notation

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

II. Filter Estimation

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

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

$$y = x * h + n \tag{1}$$

in which y, h, and n are the observed degraded image, the blurring filter, and the observation noise, respectively, and $*$ denotes the convolution operator. This equation can also be expressed in matrix-vector notation by

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{n} \tag{2}$$

where $\mathbf{h}$, $\mathbf{y}$, and $\mathbf{n}$ are the lexicographic ordering of h, y, and n all belonging to $\mathbb{R}^q$ and $\mathbf{X} \in \mathbb{R}^{q \times q}$ is a matrix operator, which is constructed by properly arranging the elements of $\mathbf{x}$ [27]. In the case of the 2-D cyclic convolution, which implicitly assumes that x and h are periodic signals, $\mathbf{X}$ is a block circulant with circulant blocks (BCCB) matrix. The latter can be spectrally decomposed as $\mathbf{X} = \mathbf{F}^T\mathbf{\Lambda}\mathbf{F}$, where $\mathbf{F}$ is the 2-D unitary ($\mathbf{F}^T = \mathbf{F}^{-1}$) discrete Fourier transform matrix, and $\mathbf{\Lambda} = \sqrt{q}\,\mathrm{diag}\,(\mathbf{F}\mathbf{x}_1)$, where $\mathrm{diag}(\mathbf{a})$ denotes a diagonal matrix with diagonal $\mathbf{a}$, $\mathbf{x}_1$ is the first column of $\mathbf{X}$, and $q > 0$ is a constant that depends on the size of the images. This decomposition has two main advantages. First, by using fast Fourier transforms (FFTs), computing matrix-vector multiplications with $\mathbf{F}^T$ and $\mathbf{F}$ can be done without explicitly constructing them; this implies significant computational savings since the cost of the FFT algorithm is $O(N \log_2 N)$ for power-of-two length vectors, where N is the length of the vector. Second, the diagonal matrix $\mathbf{\Lambda}$ is easily invertible (assuming it is nonsingular), which is very convenient, as it will be seen later. Furthermore, we restrict the estimation of the blur vector $\mathbf{h}$ to a given nonempty convex set $\mathcal{H}$ of $\mathbb{R}^q$ that is used to impose a finite support to $\mathbf{h}$ (namely, to limit the number of nonzero values of $\mathbf{h}$).

The simplest and most intuitive approach to apply the deconvolution problem to pansharpening consists in searching for a relation between each MS band and the PAN image. In other words, the estimation problem is separated into L independent parts, where L represents the number of the MS spectral bands. The intrinsic suboptimal nature of this method, which will be referred to as *Filter Estimation with Multi Spectral optimization (FE MS)*, entails the lack of correlation among the MS bands leading to worse performances (as shown in Section IV). In particular, this incoherence can cause the wrong estimation of some filters, as it is the case, for example, of the blur filter relative to the blue channel of the IKONOS sensor (see Fig. 4). For this reason, other more sophisticated approaches have to be devised to solve the problem. In the following, we discuss in more detail our proposed method in light of the previous considerations.

A different method, aimed at preserving the coherence among details of the MS and PAN images, is the object of our current proposal and hence will be simply denoted as *Filter Estimation (FE)*. Our approach assumes that the same spatial degradation affects all bands. Accordingly, we estimate the blur by applying (1) to the original PAN image and a low resolution version of it. The low resolution version of the PAN is obtained through a linear combination of the MS bands, as it is common for the class of relative-spectral-contribution methods [10]. The equivalent panchromatic image $\mathbf{p}_e$ is thus written as a function of the MS image upsampled to the PAN resolution, $\widetilde{\mathbf{M}}$, and of a vector $\boldsymbol{\alpha}$ containing the weighting coefficients. In order to include a constant term accounting for the radiation collected by the PAN sensor and not by the MS ones, we write $\mathbf{p}_e$

$$\mathbf{p}_e = \widetilde{\mathbf{M}}_{aug}^T\boldsymbol{\alpha} \tag{3}$$

where $\widetilde{\mathbf{M}}_{aug} \equiv [\widetilde{\mathbf{M}}^T, \mathbf{1}^T]^T$ is obtained by stacking $\widetilde{\mathbf{M}}$ and a row vector $\mathbf{1}^T$ composed by all ones. The choice of a linear relationship between the MS bands and the PAN image, which is implied by (3), allows for an efficient implementation of the proposed algorithm, as detailed in the following. It constitutes a widely employed model for pansharpening applications, which led to state-of-the-art pansharpening algorithms [28], [29], and is the basis of the most promising approaches based on the Bayesian theory [11], [12], [17], [30].

More specifically, the filter estimation problem is formulated as follows:

$$\begin{aligned} \underset{\mathbf{h},\alpha}{\text{minimize}} \big\{ &\|\mathbf{p}_e - \mathbf{P}_C\mathbf{h}\|^2 + \lambda\|\mathbf{h}\|^2 + \\ &+\mu\left(\|\mathbf{D}_v\mathbf{h}\|^2 + \|\mathbf{D}_h\mathbf{h}\|^2\right)\big\} \\ \text{subject to } &\mathbf{h}^T\mathbf{1} = 1, \mathbf{h} \in \mathcal{H}. \end{aligned} \tag{4}$$

The first term is known as the *data-fitting term*, which imposes that the blurred version of the panchromatic image must be close to the equivalent panchromatic. Note that $\mathbf{P}_C$ is defined such that $\mathbf{P}_C\mathbf{h}$ represents the linear convolution in matrix form between the panchromatic image $\mathbf{p}$ and the blur $\mathbf{h}$, as in (2). The second and third addends act as regularization terms (in the sense of Tikhonov) aimed at dealing with the

$$\mathbf{h} = \mathcal{F}^{-1}\left\{\circ\frac{\mathcal{F}\{\mathbf{p}\}^* \circ \mathcal{F}\{\mathbf{p_e}\}}{\mathcal{F}\{\mathbf{p}\}^* \circ \mathcal{F}\{\mathbf{p}\} + \lambda + \mu\left(\mathcal{F}\{\mathbf{d_h}\}^* \circ \mathcal{F}\{\mathbf{d_h}\} + \mathcal{F}\{\mathbf{d_v}\}^* \circ \mathcal{F}\{\mathbf{d_v}\}\right)}\circ\right\} \tag{7}$$

ill-posedness of the inverse problem [31]. These terms can be also seen as *a priori* information under a Bayesian framework [25]. Here, $\mathbf{D}_h$, $\mathbf{D}_v \in \mathbb{R}^{q\times q}$, which are BCCB matrices, stand for the first-order finite difference operator in the horizontal and vertical directions, respectively, and can also be interpreted as the convolution between $\mathbf{h}$ and the derivative filters $\mathbf{D}_h \in \mathbb{R}^q$ and $\mathbf{D}_v \in \mathbb{R}^q$. The constraints induce the normalization of the blur (i.e., $\mathbf{h}^T\mathbf{1} = 1$) and its finite support (i.e., $\mathbf{h} \in \mathcal{H}$), respectively. The selection of the squared ℓ_2 norm $\|\cdot\|^2$ allows for a closed-form solution, which can be computed efficiently in the frequency domain. The same choice proves favorable also in the regularization terms, matching some desirable physical conditions. Specifically, its use in the second addend forces the obtained solution to have limited energy. This makes sense since the blur degrading effect is usually confined to a small region. Analogously, when acting on the filter finite differences, it assures smooth transitions among the filter values. This is desirable since blurs experienced in the practice tend to be smooth and have Gaussian-like shapes. The two regularization terms are properly weighted by coefficients λ and μ, which can be used as input parameters to the algorithm.

The solution of the optimization problem follows an alternated minimization with a projection scheme, where each variable is minimized separately. First, the coefficients $\boldsymbol{\alpha}$ are estimated, and subsequently, the blur is optimized; this sequence is repeated for every iteration. The optimization (4) is strictly convex. Therefore, its solution is unique and does not depend on the initialization. For faster convergence, we initialize the blur with a reasonable guess. We start from the Starck and Murtagh (S&M) low-pass filter [32], which is commonly used in pansharpening within the "à trous" algorithm. This algorithm implements a wavelet decomposition based on a B_3 spline scaling function that was observed to closely match the typical MTF shape in the VNIR spectrum with a cutoff value of 0.185 [15]. Since, in our test cases, the resolution ratio between MS and PAN is 4, the initial filter $\mathbf{h}_0$ is obtained as the cascade of the first two steps of the dyadic wavelet decomposition, namely, it is defined as the convolution of the filter with coefficients [1 4 6 4 1] and its expanded version interlaced by zeros (i.e., [1 0 4 0 6 0 4 0 1]) [22]. After optimizing with respect to $\boldsymbol{\alpha}$ and with respect to $\mathbf{h}$, the estimated filter is normalized after each iteration, and values outside a given window (typically being very close to zero) are set to zero.

We remark that, because the optimization problem is convex, the objective function is quadratic, and the constraints are easy to deal with, we could use an algorithm to compute the exact solution. In an array of experiments, we concluded, however, that the proposed suboptimal scheme yields, from the practical point of view, solutions comparable with the exact ones with far lower computational complexity than that of the exact solvers.

This allows us to find an approximate solution to the optimization problem in a simple way. For each iteration, the first step consists in computing $\boldsymbol{\alpha}$, given the current value of $\mathbf{h}$. This corresponds to finding the solution of a simple least squares problem

$$\widetilde{\mathbf{M}}_{aug}\widetilde{\mathbf{M}}_{aug}^T\boldsymbol{\alpha} = \widetilde{\mathbf{M}}_{aug}\mathbf{P}_C\mathbf{h}. \tag{5}$$

In the second step, the computation of $\mathbf{h}$ given $\boldsymbol{\alpha}$ involves the solution of the optimization problem described by (4) with respect to $\mathbf{h}$. Being the cost function quadratic, it has a global minimum achieved when

$$\left[\mathbf{P}_C^T\mathbf{P}_C + \lambda\mathbf{I} + \mu\mathbf{D}_v^T\mathbf{D}_v + \mu\mathbf{D}_h^T\mathbf{D}_h)\right]\mathbf{h} = \mathbf{P}_C^H\mathbf{p}_e. \tag{6}$$

As seen before, the BCCB matrices $\mathbf{P}_C$, $\mathbf{D}_h$, and $\mathbf{D}_v$ are diagonalized by the 2-D discrete fourier transform (DFT) matrix, $\mathcal{F}\{\cdot\}$. This accelerates the computation of this solution as can be seen in (7) shown at the top of the page, where $\mathcal{F}^{-1}\{\cdot\}$ denotes the inverse Fourier transform, $\mathcal{F}\{\cdot\}^*$ is the complex conjugate of the Fourier transform, $\circ$ is the componentwise multiplication, and $\circ - \circ$ is the componentwise division. This equation involves a diagonal inversion, which has cost $O(k)$, but is dominated by the FFTs—with cost $O(k\log k)$; note that $\mathcal{F}\{\mathbf{p}\}$, $\mathcal{F}\{\mathbf{D}_h\}$, and $\mathcal{F}\{\mathbf{D}_v\}$ can be computed in advance. In order to take advantage of the properties of the FFT, it is necessary to work under periodic boundary conditions, as discussed earlier. However, when dealing with real-world images, this assumption is usually too strong since it is highly improbable that an image's external borders (unobserved) are repeated periodically. Processing these images as they are usually leads to the development of undesirable artifacts, and in order to reduce them, a preprocessing step is usually taken, namely, by blurring the borders of the images. This allows the discontinuities to be smoothed out [33].

The followed procedure can be summarized as in Algorithm 1.

Algorithm 1: Filter Estimation

Data: The observed PAN and MS data, $\mathbf{p}$, $\mathbf{M}$, respectively; the regularizer weights λ and μ; an initial filter estimate $\mathbf{h}^0$, dimension of the filter's support, threshold value for the stopping criterion.
Result: Degradation filter $\mathbf{h}$.
begin
  Upsample the MS image $\mathbf{M}$
  $j = 0$
  while *some stopping criterion is satisfied* **do**
    $j = j + 1$;
    Set $\boldsymbol{\alpha}^j$ as indicated in (5)
    Set $\mathbf{h}^j$ as indicated in (7)
    $\mathbf{h}^j$: Normalize, threshold and use a mask to ignore values outside the filter's support
  Set $\mathbf{h} = \mathbf{h}^j$

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

III. Pansharpening Method Based on Filter Estimation

One of the main issues for pansharpening algorithms is how to properly inject into the MS the spatial details extracted from the PAN in order to obtain the required spatial enhancement. Many approaches have been proposed in the literature, and a widespread classification distinguishes between local and global approaches [34]. An example of the former is the context-based-decision algorithm [20], which relies upon the partitioning of the fusing images in blocks. The details extracted by the PAN image are injected, if needed, after an equalization phase based on the relationship between the standard deviations calculated within the corresponding blocks of the two images. Furthermore, detail injection is performed only if the correlation between the low-pass version of the PAN image and each of the expanded MS bands is greater than a given threshold [20]. The injection schemes belonging to the second class are based on global rules, as it is the case of the simple additive injection model [22]. In this method, a histogram matching procedure between the PAN image and each MS band is advisable before fusing data from different sensors. Again, we want to stress that the objective of this work is the study of the extraction detail phase and the comparison of several approaches under the same conditions. To reach this scope, we selected a well-known injection model based on the concept of modulation [often called *high-pass modulation* (HPM)]. In [35], it is shown that the HPM scheme is related to the local contrast of an image and, in general, outperforms the high pass filter (HPF) approach. Moreover, some powerful fusing procedures are also based on the HPM, as in [21].

More in detail, let us define $\mathbf{P}^{hm} \in \mathbb{R}^{L\times r}$ as the PAN image $\mathbf{p}$ after histogram matching with the L MS spectral bands $\mathbf{M}_l$, with $l \in [1,\ldots,L]$. For the lth band, $\mathbf{P}_l^{hm}$ is given by

$$\mathbf{P}_l^{hm} = [\mathbf{p} - m_{\mathbf{p}}] \cdot \frac{\sigma_{\mathbf{M}_l}}{\sigma_{\mathbf{p}}} + m_{\mathbf{M}_l} \tag{8}$$

with $m_{\mathbf{X}}$ and $\sigma_{\mathbf{X}}$ denoting the mean and standard deviation of image $\mathbf{X}$, respectively. The HPM injection model defines the fused product $\widehat{\mathbf{M}} \in \mathbb{R}^{L\times r}$ (in the lexicographic ordering) according to the formula

$$\widehat{\mathbf{M}} = \widetilde{\mathbf{M}} \circ \left(\circ \frac{\mathbf{P}^{hm}}{\mathbf{P}_{LP}^{hm}} \circ \right) \tag{9}$$

with $\circ$ being the componentwise multiplication and $\circ - \circ$ being the componentwise division. The $\mathbf{P}_{LP}^{hm} \in \mathbb{R}^{L\times r}$ is the low resolution version of $\mathbf{P}^{hm}$ at the same scale of the MS image. A widely employed procedure for properly generating $\mathbf{P}_{LP}^{hm}$ resorts to a multiresolution analysis (MRA) [36] implemented through a pyramidal decomposition scheme. It consists in repeated applications of an analysis operator, which has the purpose of successively reducing the contained information. This process is typically completed through the application of a smoothing low-pass linear filter, which has to be matched to the sensor MTF in order to obtain the best results [15], [21]. This approach constitutes a perfect scenario for testing the capabilities of the proposed filter estimation scheme since the latter can be applied to estimate the required analysis filter. The consequent proposed pansharpening procedure is summarized as in Algorithm 2. As in [15] and [21], at each decomposition step, the output of the smoothing filter is decimated by a factor equal to the ratio R between the spatial resolution of the MS and the PAN to get the size of $\mathbf{M}$ and finally upsampled by the same factor R to reconvert the data to the PAN size, in order to reduce the aliasing effects.

Algorithm 2: Pansharpening Algorithm

Data: The observed PAN and MS data, $\mathbf{p}$, $\mathbf{M}$, respectively; the degradation filter $\mathbf{h}$.
Result: Pansharpened image $\widehat{\mathbf{M}}$.
begin
- Upsample the MS image to get $\widetilde{\mathbf{M}}$ with the same size of $\mathbf{p}$
- Equalize $\mathbf{p}$ with each band of $\mathbf{M}$ using Eq. (8) to get $\mathbf{P}^{hm}$
- Convolve $\mathbf{P}^{hm}$ with the degradation filter $\mathbf{h}$ to get $\mathbf{P}_{LP}^{hm}$
- Use Eq. (9) to obtain the final product $\widehat{\mathbf{M}}$

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

IV. Experimental Results

In order to validate the proposed method and point out its advantages with respect to filters conventionally used for detail extraction in state-of-the-art pansharpening techniques, an accurate experimental analysis has been performed by considering two data sets. The first one is an image acquired by IKONOS, which senses in the visible and near infrared spectrum range. The sensor acquires a multispectral image with four bands [Blue, Green, Red, and near-infraRed (NIR)] and a panchromatic channel. The resolution cell is 4 m $\times$ 4 m for the multispectral bands and 1 m $\times$ 1 m for the panchromatic channel. The employed data set[2] represents a mountainous and vegetated area of the China-Sichuan region and hence is here named *China data set*. The second data set is acquired by QuickBird. Analogously to IKONOS, QuickBird acquires almost simultaneously [37] a four-band multispectral image (Blue, Green, Red, and NIR) and a panchromatic channel. The resolution cell is 2.4 m $\times$ 2.4 m for the multispectral bands and 0.6 m $\times$ 0.6 m for the panchromatic channel. The data set represents an urban area of Indianapolis and is named, for this reason, *Indianapolis data set*. The ratio between the spatial resolution of the PAN and MS for both the data sets is 4.

All the algorithms considered for the comparison are based on the HPM injection model presented in Section III. Thus, the pansharpening algorithms differ only by the detail extraction

[2]Available at http://glcf.umiacs.umd.edu

phase, which is defined by the filter used to generate the equivalent panchromatic with the same resolution of the MS image [i.e., $\mathbf{P}_{LP}$ in (9)]. More in detail, we analyze the following options for the analysis filter: a 5×5 *Box* filter, which leads to the *smoothing filter-based intensity modulation* (*SFIM*) method [38]; the *S&M* filter described in Section III, which characterizes the *"à trous" wavelet transform* (*ATWT*) method [22]; and a *Gaussian*-shaped filter [20], [21], designed to match the sensor MTF by exploiting the knowledge of the gain at the Nyquist frequency [15]. Each method is referred to through an acronym, which recalls the employed filter, namely, we identify the three approaches as *Box*, *S&M*, and *Gauß*, respectively. For the sake of comparison, the *EXP* method, which does not involve details injection, is also included in the evaluation. In this case, only the interpolation phase is performed to reach the same image size of the PAN. The interpolation algorithm exploits the polynomial kernel with 23 coefficients, as in [20]. The same method is also used to generate the $\widetilde{\mathbf{MS}}$ in (9) for all the algorithms.

Two validation strategies are considered.

1) The first one is performed by following Wald's protocol. A reduced scale MS image is generated by applying a Gaussian MTF matched filter and a subsampling step to the original MS image that is also used as a reference for validating the fused products. The main objective of this test is to show that the proposed method is able to properly approximate the LPF used for degrading the images in the simulation phase.
2) The other experiment is performed at full scale with no degradation. In this case, no reference image is available, and thus, the spectral quality of the fused image has to be evaluated against the original MS image, while the spatial quality is evaluated by comparing the details of the final product with those of the original PAN image. Moreover, in this case, a visual analysis is a mandatory step to appreciate the quality of the pansharpened images.

A. Wald Protocol

In the literature, there are two well-known techniques to validate pansharpening products. One of the most used (see the pansharpening contest organized by the IEEE Data Fusion Committee in 2006 [1]) is derived from Wald's protocol and aims at verifying the *synthesis* property [39]. In this case, reduced scale MS and PAN images are simulated; the two images are then fused, and the pansharpened image is compared with the original image, used then as a reference. A problem implied by this validation procedure regards the quantification of the similarity between the two MS images or, in another words, the choice of the quality indexes to use in order to verify the spectral and the spatial consistency of the fused product with the given reference image. The literature is plenty of quality indexes that, for the pansharpening applications, can be divided into scalar and vector ones. The former can only evaluate the radiometric distortions, while the latter are able to measure the spectral distortions, too. The most used in this category are the following: the *spectral angle mapper* (*SAM*) [40] (in degrees) that evaluates the spectral distortion and the *erreur relative globale adimensionnelle de synthese* (*ERGAS*) [41] that is a generalization of the root-mean-square error (RMSE) and is able to measure both the spectral and radiometric distortions. This feature is shared by the *Q4* index [42], which represents a vectorial extension of the Q index proposed in [43], relevant to four band data sets.

TABLE I
MTF GAINS AT NYQUIST CUTOFF FREQUENCY

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

The last question that arises in this protocol is related to the procedure for simulating the low resolution MS and PAN images. Proper low-pass spatial filters have to be applied to the two images. Usually, in the case of the MS image, the filter is matched with the MTF shape of the MS sensor, designed by exploiting the hypothesis of Gaussian shape and the knowledge of the gains at the Nyquist frequency of the MS sensor (see Table I for values of sensors involved in this analysis) [15]. On the other hand, an ideal filter is applied in this phase to the PAN image [15].

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

In this validation scenario, the experiments are initially pointed at evaluating the role of the parameters characterizing the proposed filter estimation method. The first analysis is carried out by varying the λ and μ coefficients, fixing the support dimensions to a value equal to the double of the supposed real support size (i.e., 12 pixels, which can be a reasonable size for the filter support, as detailed afterward). A large range of values is analyzed in the absence of noise, i.e., by employing the model described by (1) for $n = 0$. The performances are evaluated by using the error in angle, which is measured by the SAM (in degrees), between the filter used in the simulation phase and the one estimated by Algorithm 1, with both converted in vectors by lexicographic ordering. The results are shown in Fig. 1(a) and (b). From experiments, it can be stated that the proposed method shows no particular sensitivity with respect to the regularization coefficients for a wide range of values.

Subsequently, an analysis for evaluating the robustness of the technique to different filter support sizes is performed. A reasonable choice for the support dimension is considered to be at least $3 \cdot R$, assuming a blur with Gaussian shape (i.e., for $R = 4$, the support size is at least equal to $3 \cdot R = 12$). The performances obtained for different sizes of the filter support are computed by fixing $\mu = \lambda = 0$ and $\mu = \lambda = 10^5$. The results, reported in Fig. 1(c), show that this rule of thumb for setting the size of the support is reasonable. By investigating more in detail the obtained results, we can state that errors in angle are comparable, but when the support is large, the robustness increases for larger values of the parameters. For this reason and for the

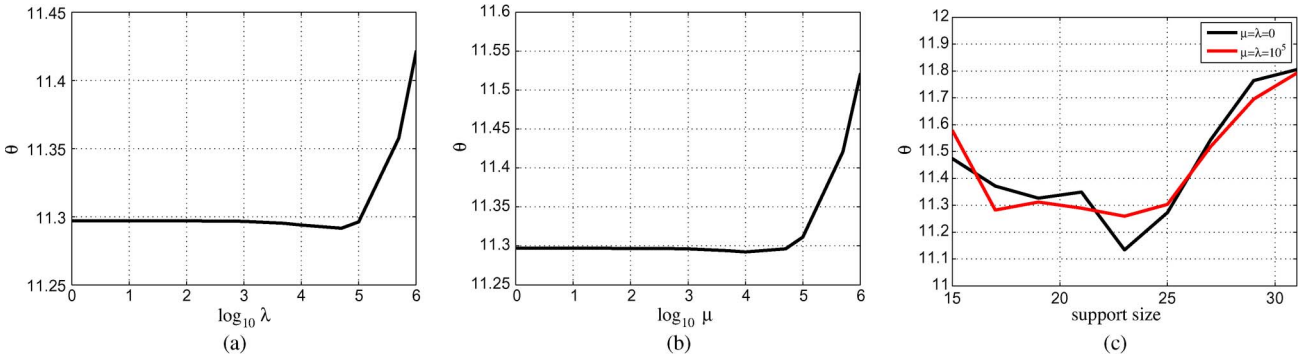

Fig. 1. Error in angle over (a) λ and (b) μ variations (with support size equal to 25) and (c) support size, using the model in (1) without additional noise.

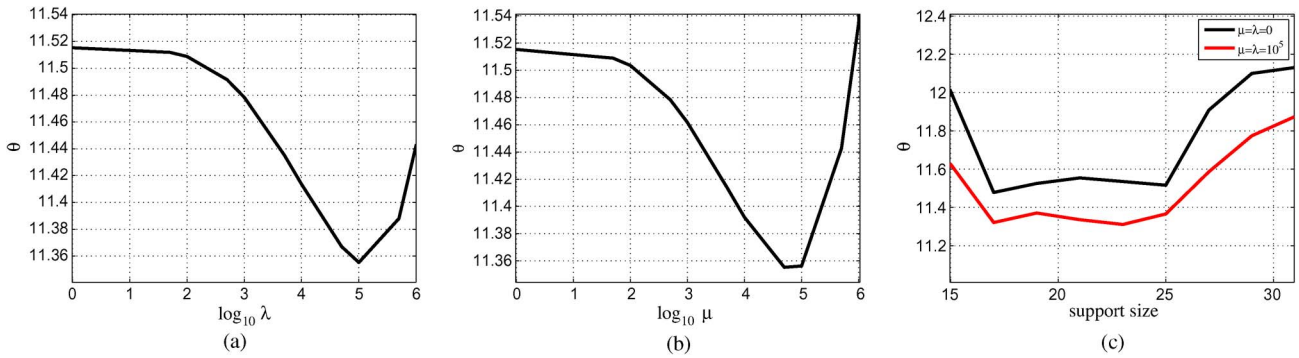

Fig. 2. Error in angle over (a) λ and (b) μ variations (with support size equal to 25) and (c) support size, using the model in (1) with Gaussian noise. The SNR is equal to 45 dB.

greater robustness to noise of this configuration, we choose the latter setting. This characteristic is particularly useful in real scenarios, in which the signal-to-noise ratio (SNR) is typically lower than that in a simulated case. As a further support to this choice, we show in Fig. 2 the curves of the angle error versus λ and μ coefficients and support dimensions obtained by using independent and identically distributed (i.i.d.) white Gaussian noise in the degradation model described by (1). We set the SNR value to 45 dB that is well suited for simulating real scene acquisitions by MS sensors. The effectiveness of employing higher values for the regularization coefficients to face the additional noise contribution is evident from the results.

Regarding the stopping criterion, a fixed number of iterations is chosen. The estimation in (5) is not very sensitive to the variation of the degradation filter applied to the PAN image, leading to a fast convergence of the proposed iterative approach for the filter estimation. A couple of iterations are usually enough to guarantee the convergence. The maximum number of iterations is set to 10 in order to ensure the stop of the algorithm.

Another interesting analysis is carried out by varying the LPF needed by the interpolation procedure that yields the upsampled version $\widetilde{\mathbf{M}}$ of the MS image. In fact, the estimation problem in (5) takes into account both the degradation introduced by the MS sensor and the one created by the nonideality of the LPF used in the upsampling step. In Fig. 3, we can see the different blurs estimated by the algorithm with different interpolation methods. It is worth noticing that, when the bicubic interpolator (i.e., the farthest one from ideal) is used, the estimated degradation is more severe than the one achieved when the half-band polynomial with 23 coefficients (i.e., closer to the ideal filter)

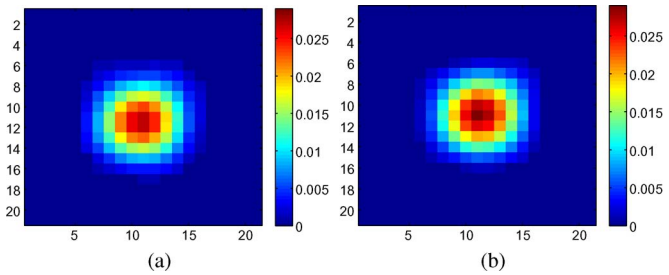

Fig. 3. Blur functions estimated through the proposed method when $\widetilde{\mathbf{M}}$ is obtained with (a) the bicubic interpolator or (b) the polynomial with 23-coefficient interpolation.

TABLE II
PERFORMANCE EVALUATION FOR THE *Gauß*, *FE MS*, AND *FE* METHODS WHEN $\widetilde{\mathbf{M}}$ IS OBTAINED BY THE BICUBIC INTERPOLATOR (ON THE LEFT) AND BY THE POLYNOMIAL CONVOLUTION KERNEL WITH 23 COEFFICIENTS (ON THE RIGHT)

Bicubic			
Algorithm	$Q4$	SAM(°)	ERGAS
Gauß	0.8744	3.1595	2.5908
FE MS	0.8674	3.1990	2.6888
FE	**0.8758**	**3.0990**	**2.5745**

23-taps			
Algorithm	$Q4$	SAM(°)	ERGAS
Gauß	**0.8826**	2.9925	**2.4573**
FE MS	0.8774	3.0679	2.5828
FE	0.8825	**2.9807**	2.4623

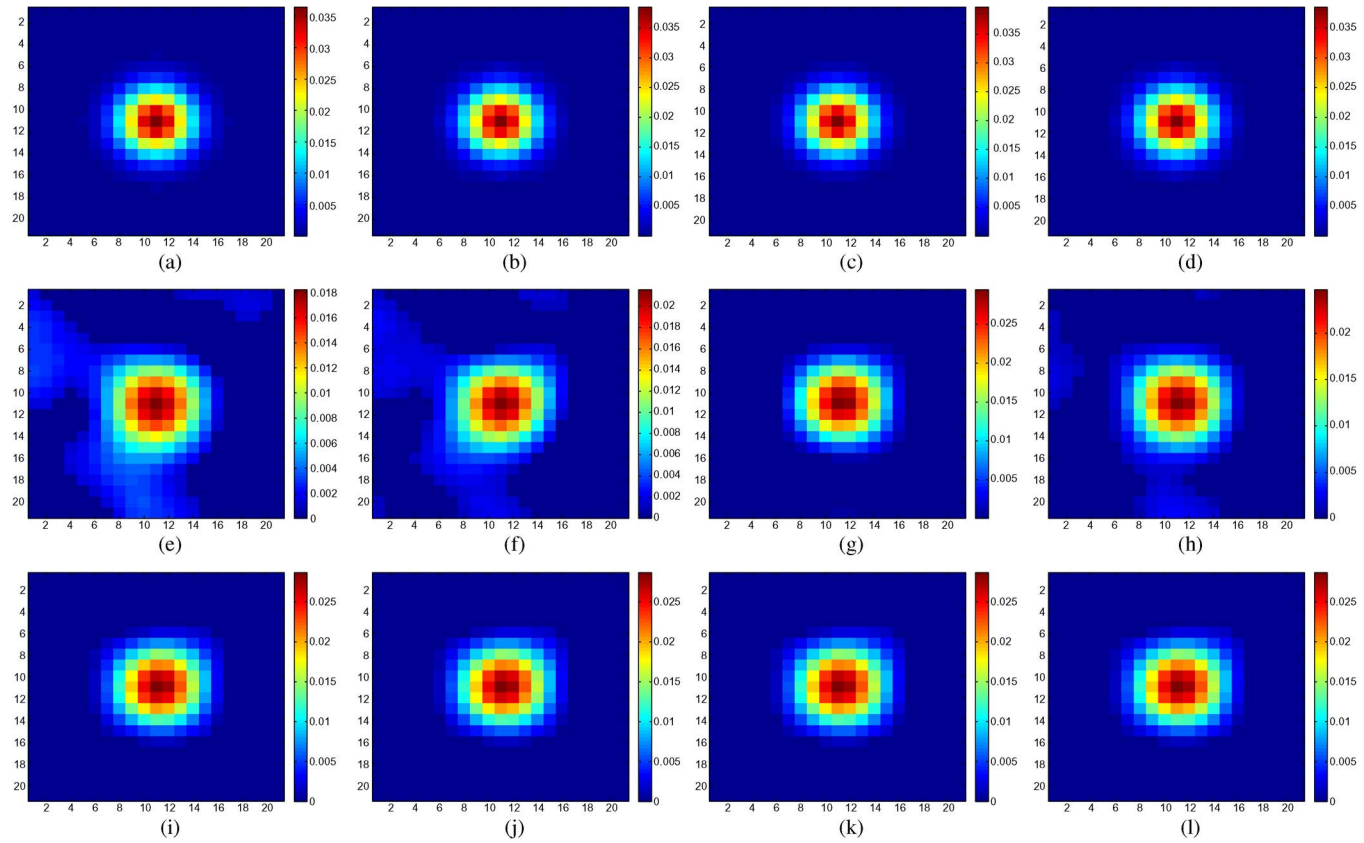

Fig. 4. From the top to the bottom: (a–d) Expected blurs. (e–h) Estimated blurs by the *FE MS* method. (i–l) Estimated blurs by the *FE* approach. The columns are ordered by wavelengths (i.e., Blue, Green, Red, and NIR).

is exploited. This confirms that the filter estimation algorithm also compensates the blurring due to the nonideality of the upsampling filter. The quality indexes reported in Table II show that the proposed approach (*FE*) overcomes the method based on the same filter used for simulating the degraded images (which supposedly should obtain the best performances) when the interpolation is far from ideal. On the contrary, by using an almost ideal LPF in the upsampling phase, the filter estimation procedure tends to approximate the system applied during the simulation, as confirmed by the very similar values achieved by MTF and FE methods. Furthermore, the results evidence that the band by band estimation procedure (*FE MS*) is not able to produce a good approximation of the degradation filters and, consequently, high-quality fused images.

A visual comparison between the actual MTF-based blurs and the estimated ones is shown in Fig. 4. The first row reports the ideal shapes of the degradation filters, and the second contains the results obtained by applying the estimation procedure band by band (*FE MS*). Finally, the third row reports the filter achieved by examining the relation between the PAN image and its equivalent low-pass version, which constitutes the proposed method of this paper (*FE*). The *FE MS* method yields a different filter for each band. This is the simplest and more intuitive choice, but it obtains poor results due to the incoherence between a MS band and the PAN image (in particular, the Blue and Green bands) [10]. In fact, the coherence between a MS channel and the PAN image increases with the contribution that the band gives to the PAN image (see the relative spectral responses for the IKONOS sensor depicted in Fig. 5); a better

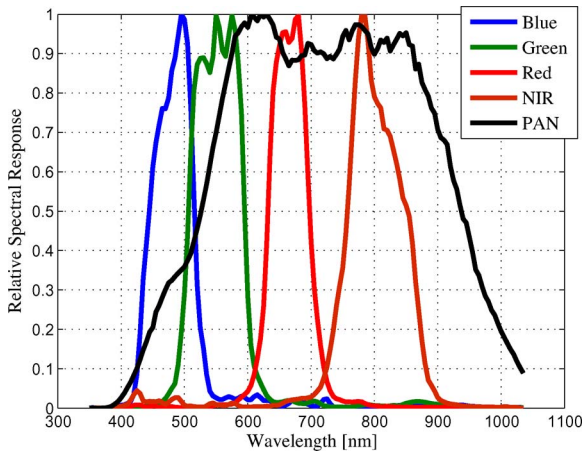

Fig. 5. IKONOS sensor: Relative spectral responses.

estimation of the degradation filter can be achieved for the NIR and Red channel or by resorting to the *FE* approach. The results obtained by the *FE MS*, the *FE*, and the *Gauß* methods are shown on the left in Table II and refer to the final products reported in Fig. 6. The advantages of the *FE* method with respect to the *FE MS* approach are evident in terms of all the performance indexes. Furthermore, it can be underlined that the proposed method constitutes a good approximation of the optimal *Gauß* approach. A further remark concerns the blur function estimated through the *FE* method, which is larger than the actual one. This effect is mainly related to the nonideality of the interpolation filter, which is implicitly addressed by

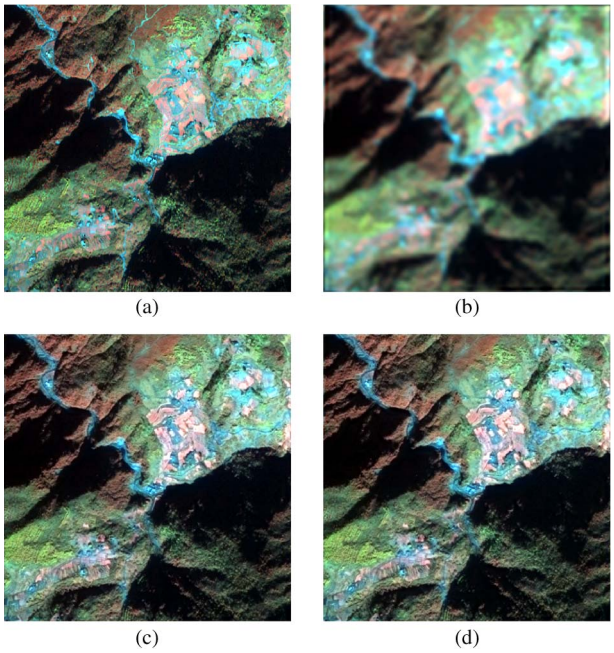

Fig. 6. (a) MS full resolution (4 m) image (used as reference image). (b) *EXP*. (c) *Gauß*. (d) *FE* methods.

TABLE III
PERFORMANCE EVALUATION FOR THE *EXP*, *Box*, *S&M*, AND *Gauß* METHODS, WHEN $\widetilde{\mathbf{M}}$ IS OBTAINED WITH THE 23-COEFFICIENT INTERPOLATOR

Algorithm	$Q4$	SAM(°)	ERGAS
EXP	0.7423	4.4100	3.8337
Box	0.8744	3.1871	2.5659
S&M	**0.8829**	3.0308	2.4659
Gauß	0.8826	**2.9925**	**2.4573**

TABLE IV
EXECUTION TIMES FOR THE COMPARED METHODS ON THE *China Data Set*

Algorithms	*EXP*	*Box*	*S&M*	*Gauß*	*FE*
Execution Times [s]	**0.30**	0.53	2.05	1.04	1.74

TABLE V
DATA SIMULATED THROUGH ANISOTROPIC DEGRADATION FILTER: PERFORMANCE EVALUATION FOR THE *Gauß*, *FE MS*, AND *FE* METHODS, EMPLOYING THE 23-COEFFICIENT INTERPOLATOR FOR ACHIEVING $\widetilde{\mathbf{M}}$

Algorithm	$Q4$	SAM(°)	ERGAS
Gauß	**0.8816**	3.0062	**2.4709**
FE MS	0.8754	3.1038	2.6114
FE	0.8813	**2.9985**	2.4769

the proposed estimation procedure, as it has been previously shown. Indeed, also the frequency response of the 23-tap cubic polynomial filter is significantly different from that of the ideal LPF [20]. Moreover, we have also checked that the estimated blur is even wider if rougher interpolation filters are employed.

The comparison with other very popular detail extraction filters is reported in Table III. Naturally, the best results are obtained by the MTF approach (since exactly the same filters

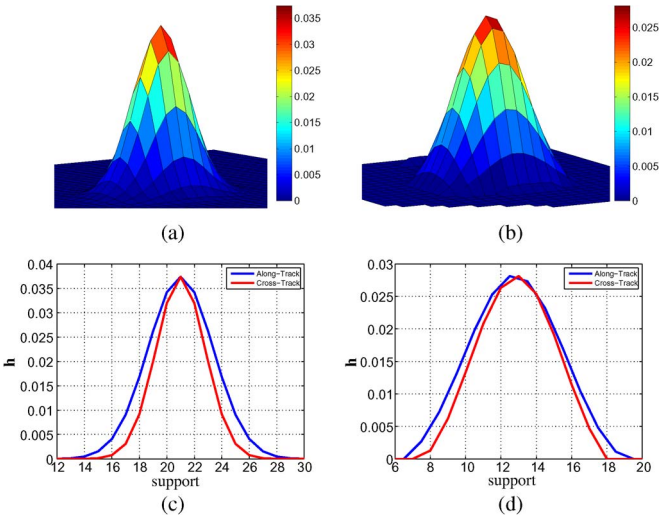

Fig. 7. (a) Gaussian-shaped impulse response of the degradation filter used for the simulation of the Red channels and (b) the corresponding estimated blur functions; the related sections in the cross- and along-track directions are reported in (c) and (d), respectively.

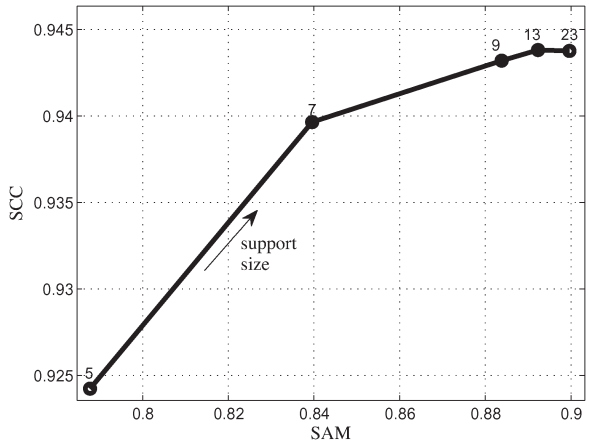

Fig. 8. Results in terms of *SAM* and *SCC* indexes of the proposed method obtained by varying the support sizes.

TABLE VI
FULL-SCALE PERFORMANCE EVALUATION USING SAM AND SCC INDEXES FOR *EXP*, *Box*, *Gauß*, AND *FE* METHODS

Algorithm	SAM(°)	SCC
EXP	**0**	0.4864
Box	0.6808	0.9159
S&M	0.7904	0.9351
Gauß	0.9772	0.9366
FE	0.8923	**0.9438**

used for the artificial degradation of the data set are used for extracting the details). Indeed, the performances of the Gaussian approaches are slightly better than wavelet-based ones, and in general, they are preferred in the detail extraction phase, due to a greater robustness to the aliasing effects [44]. Poorer performances are instead shown by the *Box* method, which is based on the *Box* filter. They are caused by the shape of its frequency response that does not perfectly match the MTF of the MS sensor and is characterized by considerable

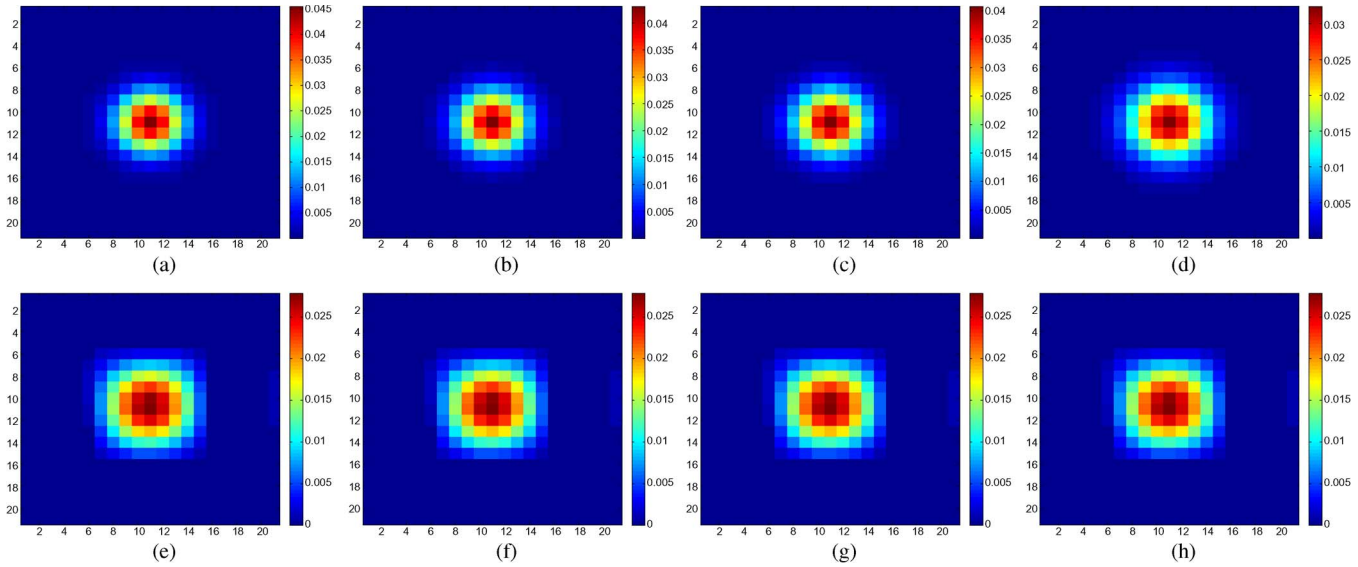

Fig. 9. From the top to the bottom: (a–d) Expected blurs and (e–h) estimated blurs by the *FE* approach. The columns are ordered by wavelengths (i.e., Blue, Green, Red, and NIR).

ripples in the obscure band; In Table IV, the execution times are shown with reference to the *China data set*, using an image with dimensions equal to 300 $\times$ 300 and an AMD Athlon 1.6-GHz processor. The interpolation phase with the half-band polynomial with 23-coefficient interpolator requires 0.3 s. The fastest approach is *Box* followed by *Gauß*, *FE* that takes more time due to the filter estimation step, and *S&M*. However, all the approaches obtain good results in terms of this index, also owing to the chosen simple but efficient injection rule.

Finally, in order to further highlight the capabilities of the proposed technique, we perform a test based on simulated data obtained through an anisotropic degradation filter. Indeed, real images hardly show isotropic blurs since the motion of the platform induces a more severe degradation in the along-track direction (see Section I-A). Thus, the simulation is conducted by using an MTF-matched filter to degrade the MS image, but in this case, a stronger attenuation in one direction is imposed. Again, remarkable performances are shown by the proposed approach, as evidenced by the overall results presented in Table V and by the comparison of the actual and estimated blurs illustrated by Fig. 7 with reference to the Red channel. Even in this case, the estimated blur has a wider support because of the previously described reasons. Nevertheless, the shapes of the two blurs are comparable, and remarkably, an appreciable match between the actual and the estimated anisotropy is achieved, as demonstrated in Fig. 7(c) and (d).

B. Full-Scale Validation

The second test case is carried out at full scale by employing the *Indianapolis data set*. The most straightforward consequence of the reference image unavailability concerns the numerical assessment of the fused product quality. Thus, visual inspection constitutes the key step to understand the effectiveness of a pansharpening algorithm. Nevertheless, as further information, we quantify the performances of the considered approaches through the spectral angle mapper (SAM) [40] between the MS image and the fused product and the spatial correlation coefficient (SCC) [45] between the details of the PAN image and the ones of the fused product. The former is a measure of the spectral distortion with respect to the original MS product, but its value has to be considered *cum grano salis*. Indeed, since the details are characterized by particular spectral features, a small SAM could eventually indicate a limited quantity of injected information [10]. Accordingly, the SAM has to be analyzed in conjunction with the SCC index that, on the other side, quantifies the amount of spatial details. It is worth to underline that other approaches are often employed in the literature to evaluate the fused products at full scale (e.g., the Quality with No Reference index [46]), but the applicability to MRA approaches could be compromised by the presence of severe aliasing effects [47], as it is the case of QuickBird data.

Fig. 8 shows the performances of the proposed method, achieved by varying the support dimension and fixing $\mu = \lambda = 10^5$. A good balance between spectral consistency and spatial enhancement is provided by a support dimension equal to 13. In Table VI, the results of the *Gauß* and *FE* algorithms are shown together with a comparison with state-of-the-art pansharpening algorithms based on different detail extraction procedures. The proposed approach obtains the best *SCC*, which indicates a remarkable detail extraction capability, and a reasonable *SAM* value that evidences a good spectral consistency.

Fig. 9 shows for every band the comparison between the blurs hypothesized by the *Gauß* method and those estimated by the proposed technique. Furthermore, Fig. 10 shows a more detailed comparison between the blur used by the *Gauß* method and that derived by the proposed estimation procedure, with reference to the Red channel. The along- and cross-track sections are reported, thus allowing to verify the expected anisotropic shape, which is particularly evident by analyzing Fig. 10(c) and (d).

Fig. 11 depicts the Red Green Blue (RGB) representation of the fused results achieved by the *Gauß* and *FE*, while Fig. 12 reports the differences (in absolute value) between the

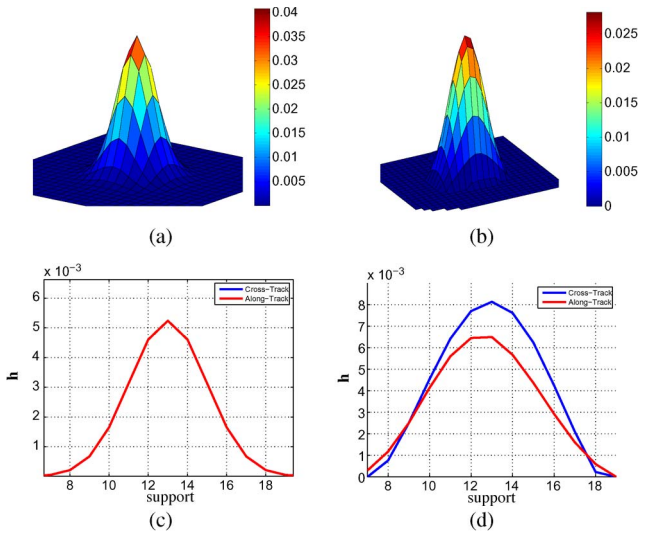

Fig. 10. (a) *Gauβ*-based and (b) estimated blur functions for the Red channel and the relative sections in the cross- and along-track directions in (c) and (d), respectively.

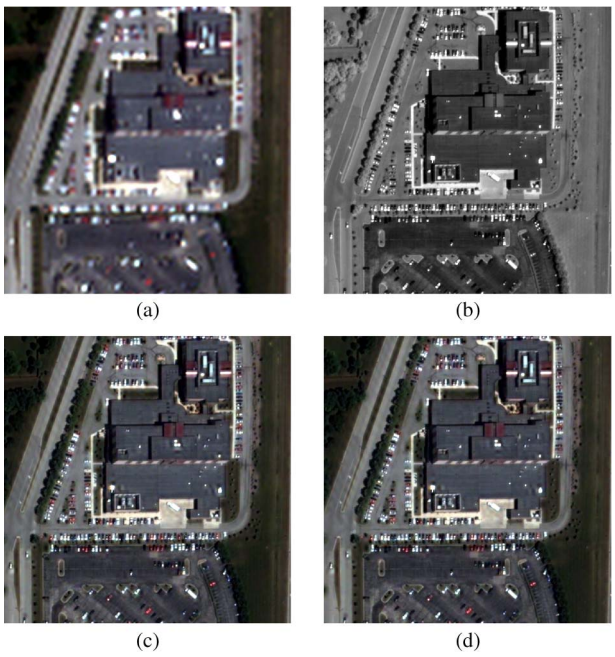

Fig. 11. Full-scale fused results: (a) *EXP*, (b) *PAN*, (c) *Gauβ*, and (d) *FE* images.

details extracted by the two methods over the bands. They are particularly evident on the edges of small structures (such as paths and cars) and on the larger ones (such as buildings). The main differences can be highlighted by focusing on the blue band, while the results are closer for the NIR channel. This similarity is mainly due the fact that the NIR band is the one that most significantly contributes to the generation of the equivalent PAN image used for estimation.

To help a visual inspection, a small area of the *Indianapolis data set* is shown in Fig. 13. Since the same injection model is adopted, only the details for both the algorithms are reported. An unique image is presented for the proposed method, while

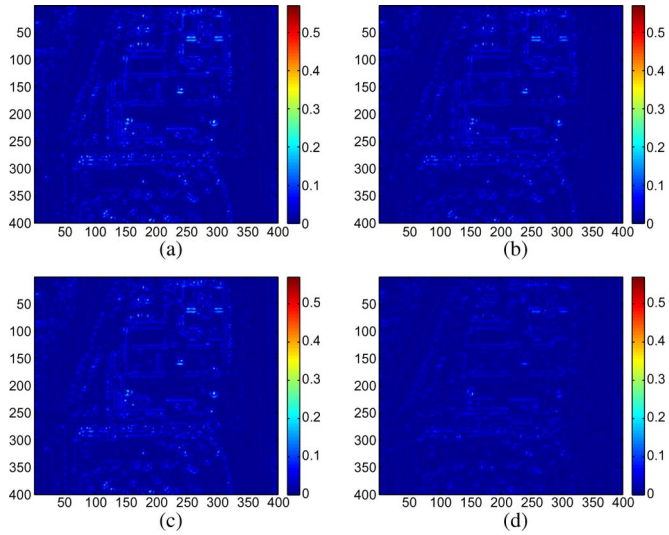

Fig. 12. Differences in absolute value between the details achieved by the *FE* and *Gauβ* methods over the bands, ordered by wavelengths (i.e., B, G, R, and NIR).

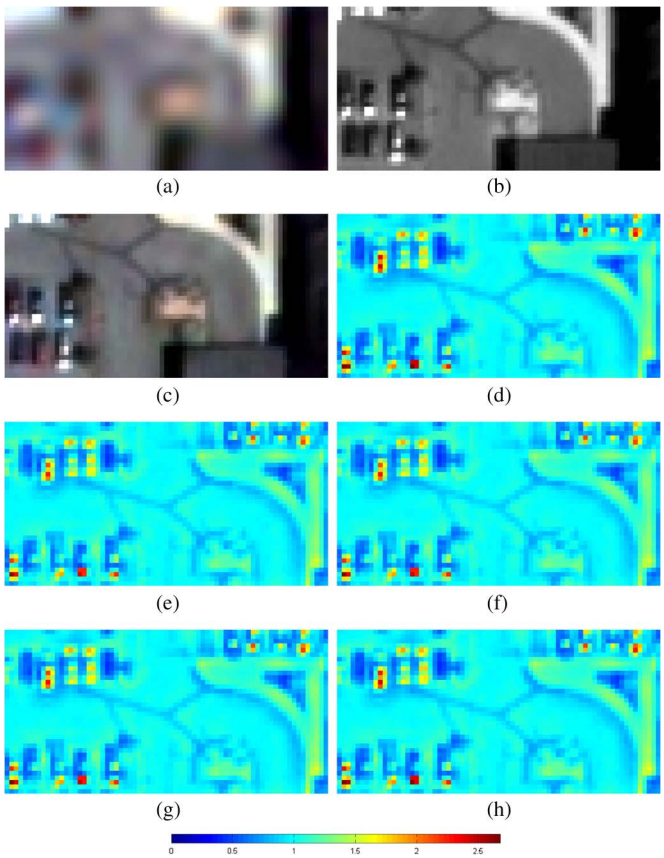

Fig. 13. Small area in the *Indianapolis data set*. (a) MS image. (b) PAN image. (c) RGB fusion result using the *FE* approach. Details extracted as $(\mathbf{p}/\mathbf{p}_{LP})$ using (d) *FE* and (e–h) *Gauβ* (different details for every spectral band) approaches.

four images, one for each band, are shown for the *Gauβ* approach. The greater evidence in the details of the proposed method compared to the ones of the *Gauβ* approach is clear, in particular, for some zones such as the path and cars present in the area under study.

V. Conclusion and Future Developments

In the last years, the relevance of pansharpening in data fusion is constantly increasing. A pansharpening algorithm is usually divided into two steps: The extraction of the details from the PAN image and their subsequent injection into the MS image. The first phase is often carried out by filters matching the MTF of the MS sensor in order to maximize the amount of useful details. Gaussian filters matched with the sensor MTF represent the state of the art for this step. In this case, a prior knowledge on the sensor characteristics is exploited to define the filter. The unavailability (or the inaccuracy) of this information results in a strong limitation of this approach. In addition, with filters matching the MTF and an approximation of the real (and unknown) MTF, the details extracted can be incorrect due to this residual mismatch. To overcome this problem, in this paper, we have proposed a procedure for *estimating* the filter that models the blur between the MS and the PAN image by only using the available images and no additional information (e.g., the MTF as specified by the sensor manufacturer). The proposed technique was compared to the state-of-the-art pansharpening technique extracting the details through filters. The experimental results were carried out by exploiting two different validation procedures: at reduced scale and at full image scale. The former validation underlined the capability of the proposed method of correctly approximating the unknown blur filter and its robustness with respect to the tuning of its free parameters. The second protocol pointed out the advantages of the proposed method with respect to the state-of-the-art pansharpening techniques, both from a numerical and a visual analysis.

The proposed approach estimates a unique filter for all the MS bands. This could represent a limitation for older sensors as QuickBird or, in future applications, for hyperspectral sensors since they could be characterized by strongly band-dependent MTFs because of a wider spectral range covered. In both cases, a different detail extraction filter for each band is advisable. Accordingly, forthcoming developments include the modification of the image blur estimation algorithm in order to take into account these diversities among spectral bands.

References

[1] L. Alparone *et al.*, "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.

[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] 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*, C.-H. Chen, Ed., 2nd ed. Boca Raton, FL, USA: CRC Press, 2012, pp. 533–548.

[4] 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.

[5] 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.

[6] 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.

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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] S. Li, H. Yin, and L. Fang, "Remote sensing image fusion via sparse representations over learned dictionaries," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4779–4789, Sep. 2013.

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

[19] M. M. Khan, L. Alparone, and J. Chanussot, "Pansharpening quality assessment using the modulation transfer functions of instruments," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3880–3891, Nov. 2009.

[20] 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.

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

[22] J. Núñez *et al.*, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1204–1211, May 1999.

[23] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*., 3rd ed. Amsterdam, The Netherlands: Elsevier, 2007.

[24] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.*, vol. 13, no. 3, pp. 43–64, May 1996.

[25] P. Campisi and K. Egiazarian, *Blind Image Deconvolution: Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2007.

[26] A. Levin, Y. Weiss, F. Durand, and W. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 1964–1971.

[27] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.

[28] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rouge, "A variational model for P + XS image fusion," *Int. J. Comput. Vis.*, vol. 69, no. 1, pp. 43–58, Aug. 2006.

[29] 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.

[30] C. Jiang, H. Zhang, H. Shen, and L. Zhang, "A practical compressed sensing-based pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 629–633, Jul. 2012.

[31] A. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*. Hoboken, NJ, USA: Wiley, 1977.

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

[33] S. Reeves, "Fast image restoration without boundary artifacts," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1448–1453, Oct. 2005.

[34] 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.
[35] 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.
[36] S. Mallat, "Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$," *Trans. Amer. Math. Soc.*, vol. 315, no. 1, pp. 69–87, Sep. 1989.
[37] W. Liu and F. Yamazaki, "Speed detection of moving vehicles from one scene of QuickBird images," in *Proc. Joint Urban Remote Sens. Event*, 2009, pp. 1–6.
[38] J. Liu, "Smoothing filter based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, Dec. 2000.
[39] 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.
[40] 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.
[41] 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.
[42] 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.
[43] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
[44] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "Advantages of Laplacian pyramids over "à trous" wavelet transforms," in *Proc. SPIE Image Signal Process. Remote Sens. XVIII*, L. Bruzzone, Ed., 2012, vol. 8537, pp. 853 704-1–853 704-10.
[45] 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.
[46] L. Alparone *et al.*, "Multispectral and panchromatic data fusion assessment without reference," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200, Feb. 2008.
[47] 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.

Gemine Vivone received the B.Sc. (*cum laude*), the M.Sc. (*cum laude)*, and the Ph.D. degrees in information engineering from the University of Salerno, Salerno, Italy, in 2008, 2011, and 2014, respectively.

He is currently a Research Fellow at the North Atlantic Treaty Organization (NATO) Science and Technology Organization Centre for Maritime Research and Experimentation, La Spezia, Italy. In 2013, he was as a Visiting Scholar with the Grenoble Institute of Technology (INPG), Grenoble, France, where he was conducting his research at the Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique (GIPSA-Lab). In 2012, he was as a Visiting Researcher at the NATO Undersea Research Centre, La Spezia. His main research interests focus on statistical signal processing, detection of remotely sensed images, data fusion, and tracking algorithms.

Dr. Vivone serves as a referee for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Miguel Simões was born in Lisbon, Portugal, in 1987. He received the M.Sc. degree in electrical and computer engineering from the Instituto Superior Técnico, University of Lisbon, Lisbon, in 2010. He is currently working toward the joint Ph.D. degree in electrical and computer engineering and signal and image processing at the Instituto de Telecomunicaçõs, Instituto Superior Técnico, University of Lisbon, Lisbon, and at the Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique (GIPSA-Lab), University of Grenoble, Grenoble, France, respectively.

Previously, he has worked as an Information Technology (IT) consultant in the field of telecommunications. His main areas of research interest are image processing, optimization, and remote sensing.

Mauro Dalla Mura (S'08–M'11) received the laurea (B.E.) and laurea specialistica (M.E.) degrees in telecommunication engineering from the University of Trento, Trento, Italy, in 2005 and 2007, respectively. He obtained a joint Ph.D. degree in information and communication technologies (telecommunications area) and electrical and computer engineering from the University of Trento and the University of Iceland, Reykjavík, Iceland, in 2011.

In 2011, he was a Research Fellow at Fondazione Bruno Kessler, Trento, Italy, conducting research on computer vision. He is currently an Assistant Professor at the Grenoble Institute of Technology (Grenoble INP), Grenoble, France. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). His main research activities are in the fields of remote sensing, image processing, and pattern recognition. In particular, his interests include mathematical morphology, classification, and multivariate data analysis.

Dr. Dalla Mura was the recipient of the IEEE Geoscience and Remote Sensing Society Second Prize in the Student Paper Competition of the 2011 IEEE International Geoscience and Remote Sensing Symposium 2011 (Vancouver, CA, USA, July 2011). He is a Reviewer of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND REMOTE SENSING, IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, *Pattern Recognition Letters*, *International Society for Photogrammetry and Remote Sensing (ISPRS) Journal of Photogrammetry and Remote Sensing*, and *Photogrammetric Engineering and Remote Sensing*. He is a member of the IEEE Geoscience and Remote Sensing Society (GRSS) and IEEE GRSS Data Fusion Technical Committee and the Secretary of the IEEE GRSS French Chapter (2013–2016). He was a Lecturer at the Remote Sensing Summer School 2012 (organized by the IEEE GRSS), Munich, Germany.

Rocco Restaino (M'14) received the Laurea degree in electronic engineering from the University of Naples, Naples, Italy, in 1998 and the Ph.D. degree in information engineering from the University of Salerno, Fisciano, Italy, in 2002.

He is currently an Assistant Professor at the University of Salerno. His research interests include probability theory, stochastic geometry, and signal processing for remote sensing and networking.

José M. Bioucas-Dias (S'87–M'95) received the E.E., M.Sc., Ph.D., and "Agregado" degrees in electrical and computer engineering from the Instituto Superior Técnico (IST), the engineering school of the Technical University of Lisbon, Lisbon, Portugal, in 1985, 1991, 1995, and 2007, respectively.

Since 1995, he has been with the Department of Electrical and Computer Engineering, IST, where he was an Assistant Professor from 1995 to 2007 and an Associate Professor since 2007. Since 1993, he is also a Senior Researcher with the Pattern and Image Analysis group of the Instituto de Telecomunicações, which is a private nonprofit research institution. His research interests include inverse problems, signal and image processing, pattern recognition, optimization, and remote sensing. He has authored or coauthored more than 230 scientific publications, including more than 60 journal papers (48 of them published in IEEE journals) and 170 peer-reviewed international conference papers and book chapters.

Dr. Bioucas-Dias was an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS (1997–2000), and he is an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING and IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He was a Guest Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING for the Special Issue on Spectral Unmixing of Remotely Sensed Data, of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING for the Special Issue on Hyperspectral Image and Signal Processing, and of the IEEE SIGNAL PROCESSING MAGAZINE for the Special Issue on Signal and Image Processing in Hyperspectral Remote Sensing. He was the General Cochair of the 3rd IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS'2011) and has been a member of program/technical committees of several international conferences.

Giorgio A. Licciardi (M'11) received the M.S. degree in telecommunication engineering and the Ph.D. degree in "geoinformation" from the Tor Vergata University, Rome, Italy, in 2005 and 2010, respectively.

He is currently a Postdoctoral Fellow with the Grenoble Institute of Technology (INPG), Grenoble, France, where he is conducting his research at the Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique (GIPSA-Lab). His research includes information extraction from remote sensing data and multispectral and hyperspectral image analysis. He is also a European Space Agency Category-1 Principal Investigator for Earth observation data.

Dr. Licciardi serves as a Referee for several scientific journals such as the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

Jocelyn Chanussot (M'04–SM'04–F'12) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995 and the Ph.D. degree from Savoie University, Annecy, France, in 1998.

In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA—French National Defense Department). Since 1999, he has been with Grenoble INP, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and is currently a Professor of signal and image processing. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). Since 2013, he has been an Adjunct Professor of the University of Iceland, Reykjavík, Iceland. His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing.

Dr. Chanussot is the founding President of the IEEE Geoscience and Remote Sensing French chapter (2007–2010) which received the 2010 IEEE GRS-S Chapter Excellence Award. He was the corecipient of the 7th Nordic Signal Processing Symposium (NORSIG) 2006 Best Student Paper Award, the IEEE GRSS 2011 Symposium Best Paper Award, the IEEE GRSS 2012 Transactions Prize Paper Award, and the IEEE GRSS 2013 Highest Impact Paper Award. He was a member of the IEEE Geoscience and Remote Sensing Society Administrative Committee (2009–2010), in charge of membership development. He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote Sensing. He was the Chair (2009–2011) and Cochair of the GRS Data Fusion Technical Committee (2005–2008). He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006–2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing (2009). He was an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (2005–2007) and for *Pattern Recognition* (2006–2008). Since 2007, he has been an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. Since 2011, he has been the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. He was a Guest Editor for the PROCEEDINGS OF THE IEEE in 2013 and a Guest Editor for the IEEE SIGNAL PROCESSING MAGAZINE in 2014. He is a member of the Institut Universitaire de France (2012–2017).