A Variational Pansharpening Approach Based on Reproducible Kernel Hilbert Space and Heaviside Function

Liang-Jian Deng, Gemine Vivone, *Member, IEEE*, Weihong Guo, Mauro Dalla Mura, *Senior Member, IEEE*, and Jocelyn Chanussot, *Fellow, IEEE*

Abstract—**Pansharpening is an important application in remote sensing image processing. It can increase the spatial-resolution of a multispectral image by fusing it with a high spatial-resolution panchromatic image in the same scene, which brings great favor for subsequent processing such as recognition, detection, *etc*. In this paper, we propose a continuous modeling and sparse optimization based method for the fusion of a panchromatic image and a multispectral image. The proposed model is mainly based on reproducing kernel Hilbert space (RKHS) and approximated Heaviside function (AHF). In addition, we also propose a Toeplitz sparse term for representing the correlation of adjacent bands. The model is convex and solved by the alternating direction method of multipliers which guarantees the convergence of the proposed method. Extensive experiments on many real datasets collected by different sensors demonstrate the effectiveness of the proposed technique as compared with several state-of-the-art pansharpening approaches.**

Index Terms—**Pansharpening, remote sensing image, sparse model, RKHS, Heaviside, Toeplitz sparsity, alternating direction method of multipliers.**

I. INTRODUCTION

PANSHARPENING refers to the fusion of a *high* spatial-resolution panchromatic (PAN) image and a *low* spatial-resolution multispectral (MS) image to recover a high spatial-resolution MS image. PAN and MS images are acquired almost simultaneously by sensors mounted on many optical satellites, such as IKONOS and QuickBird. Due to physical constraints, the fusion of the PAN and MS products represents the only viable solution to attain a high resolution in both the spatial and spectral domains. The interest of the scientific community in pansharpening can be attested from the contest launched by the Data Fusion Committee of the IEEE Geoscience and Remote Sensing Society in 2006 [1], [2] and the review papers in [3] and [4]. Commercial products, see for instance Google Earth, are exploiting pansharpened images making pansharpening an important preliminary step for several image analysis tasks, see *e.g.*, change detection [5].

Most of the pansharpening papers in the literature are based on paradigms such as component substitution (CS) and Multiresolution analysis (MRA). The former relies upon the substitution of a component of the image, after a spectral transformation of the MS data, with the PAN image, see *e.g.*, the intensity-hue-saturation [6], the principal component analysis [7], and the Gram-Schmidt (GS) spectral sharpening [8]. These methods first project the upsampled MS images into a new space, then substitute the image components by the image details of PAN image, and finally apply an inverse projection to get the high spatial resolution pansharpened image. CS methods have in general a low computational burden, but the results can be affected by spectral distortions. The MRA approach is based on the injection into the MS image of spatial details (*i.e.*, high spatial frequencies not resolved in the MS image) that are obtained from the PAN image, *e.g.*, additive wavelet luminance proportional [9], Laplacian pyramid [10], generalized Laplacian pyramid [11], "à-trous" wavelet transform [12], and so on. In contrast to CS techniques, MRA approaches mainly suffer from spatial distortions while preserving spectral information well.

Besides CS and MRA based methods, variational regularization based ones are also popular [13]–[27]. Ballester *et al.* [13] proposed a variational regularization pansharpening method based on two assumptions: one is that PAN image is the linear combination of high spatial resolution multispectral bands and the other one is that the PAN image could provide high spatial resolution information for the multispectral image. However, this method sometimes suffers from spectral distortion due to the first unrealistic assumption [3]. To overcome this

limitation, Möller *et al.* [14] combine wavelet based fusion and the pansharpening. Jiang *et al.* [24] presented a sparsity-promoting regularization model for pansharpening tasks, in which a Hyper-Laplacian penalty[1] using $\ell_{1/2}$ norm is employed for spectral preservation. This non-convex model can be efficiently solved by the alternating direction method of multipliers (ADMM) approach. Pansharpening based on unmixing turns out to be effective as well. Work in this direction includes that of [28] and [29]. For instance, Yokoya *et al.* [28] proposed an image fusion approach from the perspective of coupled nonnegative matrix factorization (CNMF) unmixing. This method can produce high-quality fusion results, both spectrally and spatially.

Recently, convolutional neural networks (CNNs) based methods have been proposed and show very powerful ability for the pansharpening application, see *e.g.* [30]–[35]. These CNNs models are generally based on the assumption which the relationship between HR/LR multispectral image patches is the same as that between the corresponding HR/LR panochromatic image patches. By this assumption, one can learn a mapping through a neural network to finally determine the pansharpened image. To the best of our knowledge, the first work of using CNNs for pansharpening is proposed by Huang *et al.* [30]. Afterwards, Masi *et al.* [31] adopt a simple and effective three-layer architecture, which is recently proposed for super-resolution [36], to pansharpening application. This method can obtain the state-of-the-art pansharpening performance. Yang *et al.* [34] present a deep network for the pansharpening, which is called PanNet. This method designs the PanNet architecture by incorporating domain-specific knowledge, as well as mainly focuses on two important issues, *i.e.*, spectral and spatial preservation. In summary, the CNNs based approaches have become a new trend in the pansharpening problem, and shown the excellent fusion ability.

In this paper, we extend our previous work on single image super-resolution [37] to pansharpening in a variational format. In [37], the super-resolution problem was solved from an intensity estimation perspective. It was assumed that the high-resolution image is a discretization of an underlying image intensity function defined on a continuous domain. Since an image usually contains smooth components and discontinuous components such as edges. The intensity function is assumed to be the sum of two functions, one describing the continuous component and one describing the edges. The smooth function is chosen from RKHS and the edge function is represented using AHF. When extending the idea to pansharpening, we model each band of the high resolution multispectral image using a redundant basis with the help of RKHS and AHF. Differences between coefficients of adjacent bands are put in the minimization energy functional to enforce correlation between adjacent bands, which results in a Toeplitz sparse term. The proposed sparse model is mainly based on three terms which are derived from RKHS, AHF and the Toeplitz sparse term. We propose an ADMM-based algorithm that is guaranteed to converge to solve the proposed model. Moreover, to compensate for the errors coming from using a not perfect basis and parameter selection, we also conduct an outer iterative strategy to pick up more image details. Furthermore, the proposed approach is positively compared with several state-of-the-art pansharpening approaches using real data acquired by different sensors such as Pléiades, IKONOS, WorldView-2.

A preliminary version of this paper is just accepted by IEEE ICIP2017 [38]. It is worth mentioning that this paper is a significantly expanded version. First, we add more method-ological details to illustrate the motivation, modeling and algorithm of the proposed method. Second, we add one Toeplitz sparse term to enforce the correlation between adjacent bands, which can obtain more promising results, e.g., to the data by IKONOS sensor. Third, more data by different sensors, including four 4-bands and one 8-bands data, are employed for experiments to verify the effectiveness of proposed method. Finally, more discussions and comparisons from different perspectives are given in the results section, which can also demonstrate the effectiveness of our method.

The remainder of the paper is organized as follows. The related work is introduced in Sect. II. The proposed approach is detailed in Sect. III. Sect. IV is devoted to the description of the experimental results. Finally, conclusions are drawn in Sect. V.

II. REVIEW ON THE RELATED WORK

Deng *et al.* [37] proposed to use a *continuous modeling* approach to enhance image spatial resolution from *one* low spatial resolution grayscale image. Let f represent the intensity function of the underlying image defined on a continuous domain. Without loss of generality, we may assume $f : [0, 1] \times [0, 1] \longrightarrow \mathbb{R}$. The low and high resolution digital images are just discretization of f on some coarse and fine grids of $[0, 1] \times [0, 1]$ respectively. We assume f can be well represented as the sum of the smooth components and the edge components. They model the intensity of the image of interest as a function defined on a continuous domain, and assume the *intensity function* f of an image can be well represented as the sum of two functions: $f = f_s + f_e$ where f_s is the *smooth component* lying in a space called Reproducible Kernel Hilbert Space (RKHS); f_e describes image *edges* at various locations with different orientations represented by a group of redundant approximated Heaviside functions.

A. Smooth Component

We adopt 2D thin-plate spline based RKHS for the smooth component f_s of a 2D image [39]. We first review its application in the noise removal problem.

Let

$$\mathbf{g} = \mathbf{f} + \mathbf{\eta}, \tag{1}$$

where $\mathbf{g}$ is the observed noisy image, $\mathbf{f} = (f(\mathbf{z}_1), f(\mathbf{z}_2), \cdots, f(\mathbf{z}_n))^T$ is the underlying clean digital image on the discretization grids $\mathbf{z}_i = (x_i, y_i) \in [0, 1] \times [0, 1], i = 1, 2, \cdots, n$, $\mathbf{\eta}$ represents an additive noise.

In [39], an optimal estimate of f for spline smoothing problems can be obtained by minimizing the following model

$$\min_f \frac{1}{n} \int_\Omega (g - f)^2 + \lambda J_m(f), \tag{2}$$

---
[1]Which corresponds to ℓ_p-norm, $0 < p < 1$.

where λ is the regularization parameter, g is the noisy function, m is a parameter to control the total degree of polynomial, and the penalty term is defined as follows

$$J_m(f) = \sum_{v=0}^{m} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} C_m^v \left(\frac{\partial^m f}{\partial x^v \partial y^{m-v}} \right)^2 dx dy, \quad (3)$$

From [39, Ch. 2], the null space of the penalty function $J_m(f)$ is a $M = C_{d+m-1}^d$ dimensional space spanned by the polynomials of degree no more than $m-1$. In the experiments, we let $d = 2$ (for 2D image), $m = 3$, then $M = C_{d+m-1}^d = 6$, so the null space of the penalty function $J_m(f)$ can be spanned by the following terms: $\phi_1(x, y) = 1, \phi_2(x, y) = x, \phi_3(x, y) = y, \phi_4(x, y) = xy, \phi_5(x, y) = x^2, \phi_6(x, y) = y^2$. Since these polynomials are smooth functions, we mainly use them to describe the smooth components of an image. Duchon (see [40]) has proved that if there exists $\{z_i\}_{i=1}^n$ so that the least squares regression on $\{\phi_v\}_{v=1}^M$ is unique, then the optimization model (2) has a unique solution as follows

$$f_\lambda(t) = \sum_{v=1}^{M} d_v \phi_v(\mathbf{z}) + \sum_{i=1}^{n} c_i E_m(\mathbf{z}, \mathbf{z}_i), \quad (4)$$

where $\mathbf{z} = (x, y)$, $E_m(\mathbf{z}, \mathbf{z}_i)$ is a Green's function for the m-iterated Laplacian defined as:

$$E_m(s, t) = E_m(|s - t|) = \theta_{m,d} |s - t|^{2m-d} \ln|s - t|,$$

where $\theta_{m,d} = \frac{(-1)^{d/2+m+1}}{2^{2m-1} \pi^{d/2} (m-1)! (m-d/2)!}$ (refer to the chapter 2 of [39] for more details).

The result $f_\lambda(t)$ from the denoising problem tends to be smooth and we therefore define the smooth component as $f_s := f_\lambda$. Notice that the polynomial functions in f_s are obviously smooth and one may understand the $E_m(s, t)$ part as any smooth component that can not be described by those polynomials.

B. Edge Component

A group of redundant AHFs is employed to represent the edge function f_e. The Heaviside function is defined as

$$\phi(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases} \quad (5)$$

It is singular at $x = 0$ and describes a jump at $x = 0$. We usually use its smooth approximation for practical problems. Deng et al. [37] use the following smoothed version

$$\psi(x) = \frac{1}{2} + \frac{1}{\pi} \arctan(\frac{x}{\xi}), \quad (6)$$

which approximates to $\phi(x)$ when $\xi \to 0$ and $\xi \in \mathbb{R}$ actually controls the smoothness. The *edges* in 2D images can be expressed by the approximated Heaviside function $\psi((\cos\theta_j, \sin\theta_j) \cdot \mathbf{z} + c_\rho)$ which represents an edge with θ_j elevation from the horizontal axis at location c_ρ. Refer to [37] for details. In this work, we set $\xi = 5 \times 10^{-3}$, and 45 angles, i.e.,

$$\theta_j \in \{2\pi/45, 4\pi/45, 6\pi/45, \cdots, 88\pi/45, 2\pi\},$$

while $c_\rho \in \{0, \frac{1}{n-1}, \frac{2}{n-1}, \cdots, 1\}$, n is the number of pixels of low-resolution image, $m = k \cdot n$ where k is the number of orientations $\{\theta_j\}$, i.e. $k = 45$.

Putting the smooth component and the edge component together, we have the intensity function $f(\mathbf{z}) = \sum_{v=1}^{M} d_v \phi_v(\mathbf{z}) + \sum_{i=1}^{n} c_i E_m(\mathbf{z}, \mathbf{z}_i) + \sum_{j=1}^{k} \sum_{\rho=1}^{n} \beta_j \psi((\cos\theta_j, \sin\theta_j) \cdot \mathbf{z} + c_\rho)$ with k orientations and n pixels. Therefore, the set $\{\phi_v(\mathbf{z}), E_m(\mathbf{z}, \mathbf{z}_i), \psi((\cos\theta_j, \sin\theta_j) \cdot \mathbf{z} + c_\rho)\}$ forms a redundant "basis" to represent the intensity function.

Let $\mathbf{r} \in \mathbb{R}^n$ be the given vector-form low-resolution image defined on a coarse grid, the corresponding discretization for $\phi_v(\cdot), E_m(\cdot), \psi(\cdot)$ leads to matrices $\mathbf{T}^l \in \mathbb{R}^{n \times M}$, $\mathbf{K}^l \in \mathbb{R}^{n \times n}$ and $\mathbf{\Psi}^l \in \mathbb{R}^{n \times m}$ ($m = k \cdot n$) which are obtained respectively from the 2D thin-plate spline based RKHS and the AHF on the coarse grid. Similarly, we can also get the matrices on a finer grid, i.e., $\mathbf{T}^h \in \mathbb{R}^{\tilde{n} \times M}$, $\mathbf{K}^h \in \mathbb{R}^{\tilde{n} \times n}$ and $\mathbf{\Psi}^h \in \mathbb{R}^{\tilde{n} \times m}$ where $\tilde{n}$ is the pixels number of high-resolution image and $\tilde{n} > n$.

In [37], the intensity information of the given low-resolution image defined on a coarser grid is used to estimate coefficients $\mathbf{d}, \mathbf{c}, \boldsymbol{\beta}$ of the redundant "basis":

$$\min_{\mathbf{d}, \mathbf{c}, \boldsymbol{\beta}} \frac{1}{2\tilde{n}} \|\mathbf{r} - \mathbf{DB}(\mathbf{T}^h \mathbf{d} + \mathbf{K}^h \mathbf{c} + \mathbf{\Psi}^h \boldsymbol{\beta})\|_2^2 + \lambda \mathbf{c}^T \mathbf{K}^l \mathbf{c} + \alpha \|\boldsymbol{\beta}\|_1,$$

$$(7)$$

where λ and α are two positive regularization parameters, the first term is the fidelity term, the second term is to regularize the coefficient $\mathbf{c}$, and the coefficient $\boldsymbol{\beta}$ of AHF redundant basis is considered as being sparse (using ℓ_1 to depict) since edges are pretty sparse in an image; $\mathbf{D}$ is a downsampling operator and $\mathbf{B}$ is a blurring operator. To focus on the main contribution of the proposed idea and to avoid errors from the blurring operator, the authors in [37] assume the deblurring is taken care of already and let $\mathbf{B} = \mathbf{I}$ where $\mathbf{I}$ is an identity matrix. The continuous modeling nature also makes it convenient to remove the downsampling operator D as downsampling of the high resolution image could be considered as a discretization of the intensity function on a coarser grid. Thus $\mathbf{DB}(\mathbf{T}^h \mathbf{d} + \mathbf{K}^h \mathbf{c} + \mathbf{\Psi}^h \boldsymbol{\beta})$ can be considered as $\mathbf{T}^l \mathbf{d} + \mathbf{K}^l \mathbf{c} + \mathbf{\Psi}^l \boldsymbol{\beta}$,

$$\min_{\mathbf{d}, \mathbf{c}, \boldsymbol{\beta}} \frac{1}{2n} \|\mathbf{r} - (\mathbf{T}^l \mathbf{d} + \mathbf{K}^l \mathbf{c} + \mathbf{\Psi}^l \boldsymbol{\beta})\|_2^2 + \lambda \mathbf{c}^T \mathbf{K}^l \mathbf{c} + \alpha \|\boldsymbol{\beta}\|_1. \quad (8)$$

After gaining the coefficients $\mathbf{d} \in \mathbb{R}^M$, $\mathbf{c} \in \mathbb{R}^n$ and $\boldsymbol{\beta} \in \mathbb{R}^m$, it is easy to get the high-resolution image by $\mathbf{h} = \mathbf{T}^h \mathbf{d} + \mathbf{K}^h \mathbf{c} + \mathbf{\Psi}^h \boldsymbol{\beta}$ (vector form).[2]

Recently, remote sensing image processing has attracted the more attentions in the community of image and signal processing. Many related applications have been applied, see for instance detection [41], [42], unmixing [43], destriping [44], pansharpening [21], [23], [29], etc. The work is also focusing on the remote sensing image application, i.e., pansharpening. In what follows, we will show how to extend the single image super-resolution idea to pansharpening problem of multispectral images.

[2]Note that the model (8) is applied to image patches which can be viewed as square images with significantly smaller sizes than the whole image, thus for every patch, it has one set of $\mathbf{d}, \mathbf{c}, \boldsymbol{\beta}$.

III. THE PROPOSED MODEL AND ITS SOLUTION

A. Related Notations

Before introducing the proposed method, it is necessary to state related notations along this paper.

- Low-resolution multispectral image (LRMS): $\mathbf{MS} \in \mathbb{R}^{m_1 \times n_1 \times N}$ with N spectral images $\mathbf{MS}_i \in \mathbb{R}^{m_1 \times n_1}$, $i = 1, 2, \cdots, N$.
- High-resolution multispectral image (HRMS): $\widehat{\mathbf{MS}} \in \mathbb{R}^{m_2 \times n_2 \times N}$ with N spectral images $\widehat{\mathbf{MS}}_i \in \mathbb{R}^{m_2 \times n_2}$, $i = 1, 2, \cdots, N$, where $m_2 = s \cdot m_1$, $n_2 = s \cdot n_1$ with the scale factor s. In addition, $\widetilde{\mathbf{MS}} \in \mathbb{R}^{m_2 \times n_2 \times N}$ represents the initial upscaled multispectral image, which is generally formulated by the interpolation of $\mathbf{MS}$.
- Panchromatic (PAN) image: $\mathbf{P} \in \mathbb{R}^{m_2 \times n_2}$.

In our work, the proposed model mainly involves vector forms, thus in the following context we denote the above notations as vector ones, *i.e.*, $\mathbf{s}_i \in \mathbb{R}^{m_1 \, n_1 \times 1}$ (matrix form $\mathbf{MS}_i$), $\widehat{\mathbf{s}}_i \in \mathbb{R}^{m_2 \, n_2 \times 1}$ (matrix form $\widehat{\mathbf{MS}}_i$), $\widetilde{\mathbf{s}}_i \in \mathbb{R}^{m_2 \, n_2 \times 1}$ (matrix form $\widetilde{\mathbf{MS}}_i$) and $\mathbf{p} \in \mathbb{R}^{m_2 \, n_2 \times 1}$ (matrix form $\mathbf{P}$).

The *main contributions* of the proposed method are summarized as follows:

- Extend the continuous modeling based single image super-resolution method to pansharpening. To the best of authors' knowledge, it is probably the first continuous modeling approach on pansharpening from the perspective of image super-resolution. It considers the similarity between adjacent bands and thus helps improving the spectral correlation.
- Most of the existing variational regularization pansharpening approaches involve downsampling and blurring operators, but the blur kernel is generally not accurately available. Computations involving the two operators are usually more complicated. The proposed method bypass the two operators by continuous modeling as well as a cheap preprocessing that somehow mimics the deblurring.
- We also propose an outer iterative scheme to pick up more image details lost to imperfect basis and parameter selection.

B. The Proposed Model

We extend the single image SR model (7) to the pansharpening of multispectral images. The main differences between single image SR and multispectral pansharpening are two fold: 1) single image SR only has one low resolution image as the input while pansharpening problems have two inputs: the low resolution multi-spectral images and a high resolution panchromatic image; 2) the goal of the single image SR is to achieve high spatial resolution only as it doesn't involve spectral information while the pansharpening problem needs to increase spatial resolution while reducing the spectral distortion.

Next, we present the proposed model for the pansharpening of multispectral image.

1) RKHS and Heaviside Based Sparse Term: A natural extension of method (7) is to model each band of the MS image by the same redundant basis explained in the previous section but with different coefficients. Let $\mathbf{d}_i$, $\mathbf{c}_i$, $\boldsymbol{\beta}_i$ represent the coefficients for the i-th band. The first energy term (denoted as $Eng^{(1)}$) of our optimization model is presented as follows

$$Eng^{(1)} = \frac{1}{2N_{um}} \sum_{i=1}^{N} \|(\mathbf{T}^h \mathbf{d}_i + \mathbf{K}^h \mathbf{c}_i + \Psi^h \boldsymbol{\beta}_i) - \widetilde{\mathbf{s}}_i\|_2^2$$
$$+ \frac{\mu}{2} \sum_{i=1}^{N} \mathbf{c}_i^T \mathbf{K}^l \mathbf{c}_i + \frac{\lambda_1}{2} \sum_{i=1}^{N} \|\boldsymbol{\beta}_i\|_1, \quad (9)$$

where N_{um} represents the total number of pixels in the high-resolution multispectral (HRMS) image (*i.e.*, $m_2 \cdot n_2 \cdot N$), μ and λ_1 are two positive regularization parameters, with $\mathbf{T}^h \in \mathbb{R}^{m_2 n_2 \times M}$, $\mathbf{K}^h \in \mathbb{R}^{m_2 n_2 \times m_1 n_1}$, $\Psi^h \in \mathbb{R}^{m_2 n_2 \times m}$, $\mathbf{K}^l \in \mathbb{R}^{m_1 n_1 \times m_1 n_1}$, $\mathbf{d}_i \in \mathbb{R}^{M \times 1}$, $\mathbf{c}_i \in \mathbb{R}^{m_1 n_1 \times 1}$, $\boldsymbol{\beta}_i \in \mathbb{R}^{m \times 1}$, $i = 1, 2, \cdots, N$. In our work, we do not put blurring or downsampling operator into the modeling but use a fast component substitution or MRA approach to reverse both the downsampling and blurring operations to some extent and feed our algorithm with this preliminary pansharpening result to further improve the result. In this paper, we adopt a classical Gram-Schmidt (GS) [8] method to obtain the initial multispectral image $\widetilde{\mathbf{s}}_i \in \mathbb{R}^{m_2 \, n_2 \times 1}$, aiming to roughly depict the blurring and downsampling processes. It basically combines the low resolution multispectral image $\widetilde{\mathbf{s}}$ and the panchromatic $\mathbf{p}$ for a higher resolution image. Any other CS or MRA approach could be used as well (see the discussion in the results section). Note that the term (9) is different with the term in model (7) mainly on two aspects: one is the summation over all bands, and the other aspect is the difference about the downsampling and blurring operators.

2) The Fidelity Term Between Panchromatic Image and Multispectral Image: The panchromatic image $\mathbf{p}$ contains spatial details. To use it to help increase the spatial resolution, we adopt it using the fact that $\mathbf{p}$ can be viewed as the weighted sum of the latent HRMS image, where the weights ω_i, $i = 1, 2, \cdots, N$ can be automatically estimated by the linear regression between the known LRMS image and the spatially degraded panchromatic image, e.g., [45]. Therefore, the fidelity term between panchromatic image and multispectral image can be written as follows,

$$Eng^{(2)} = \| \sum_{i=1}^{N} \omega_i (\mathbf{T}^h \mathbf{d}_i + \mathbf{K}^h \mathbf{c}_i + \Psi^h \boldsymbol{\beta}_i) - \mathbf{p}\|_2^2, \quad (10)$$

where $\mathbf{T}^h \mathbf{d}_i + \mathbf{K}^h \mathbf{c}_i + \Psi^h \boldsymbol{\beta}_i$, $i = 1, 2, \cdots, N$ represents one band of latent HRMS image.

a) The Toeplitz Sparsity: It is observed that there is similarity between adjacent bands in the MS image. For instance, even though the color looks different, adjacent bands represent same objects which should be characterized by similar smoothness and edges across bands. Since we use a redundant basis to represent each individual band, the similarity between adjacent bands can be described by imposing similarity on the coefficients. Therefore, we define the third term in our energy function as

$$Eng^{(3)} = \|\mathbf{e}\|_1, \quad (11)$$

where

$$\mathbf{e} = \begin{pmatrix} \mathbf{d}_1 - \mathbf{d}_2 \\ \mathbf{c}_1 - \mathbf{c}_2 \\ \boldsymbol{\beta}_1 - \boldsymbol{\beta}_2 \\ \vdots \\ \mathbf{d}_{N-1} - \mathbf{d}_N \\ \mathbf{c}_{N-1} - \mathbf{c}_N \\ \boldsymbol{\beta}_{N-1} - \boldsymbol{\beta}_N \end{pmatrix} = \mathbf{T}_{oe} \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{c}_1 \\ \boldsymbol{\beta}_1 \\ \vdots \\ \mathbf{d}_N \\ \mathbf{c}_N \\ \boldsymbol{\beta}_N \end{pmatrix}, \quad (12)$$

and $\mathbf{T}_{oe}$ is a non-symmetric Toeplitz matrix $\mathbf{T}_{oe}(\mathbf{s}_1, \mathbf{s}_2) \in \mathbb{R}^{(M+m_1 n_1 + m)(N-1) \times (M+m_1 n_1 + m)N}$ with the vectors

$$\mathbf{s}_1 = (0, 0, \cdots, 0, 0, 0, \cdots, 0)^T \in \mathbb{R}^{(M+m_1 n_1 + m)(N-1) \times 1},$$

and

$$\mathbf{s}_2 = (1, 0, \cdots, 0, -1, 0, \cdots, 0) \in \mathbb{R}^{1 \times (M+m_1 n_1 + m)N},$$

where the value "-1" locates at the $(M + m_1 n_1 + m + 1)$-th location of $\mathbf{s}_2$.

We combine (9), (10) and (11) to formulate the final pansharpening model,

$$\min_{\mathbf{c}_i, \mathbf{d}_i, \boldsymbol{\beta}_i} \frac{1}{2N_{um}} \sum_{i=1}^N \|(\mathbf{T}^h \mathbf{d}_i + \mathbf{K}^h \mathbf{c}_i + \Psi^h \boldsymbol{\beta}_i) - \widetilde{\mathbf{s}}_i\|_2^2$$
$$+ \frac{\mu}{2} \sum_{i=1}^N \mathbf{c}_i^T \mathbf{K}^l \mathbf{c}_i + \frac{\lambda_1}{2} \sum_{i=1}^N \|\boldsymbol{\beta}_i\|_1 + \frac{\kappa}{2} \|\mathbf{e}\|_1$$
$$+ \frac{\lambda_2}{2} \|\sum_{i=1}^N \omega_i (\mathbf{T}^h \mathbf{d}_i + \mathbf{K}^h \mathbf{c}_i + \Psi^h \boldsymbol{\beta}_i) - \mathbf{p}\|_2^2, \quad (13)$$

where μ, λ_1, λ_2 and κ are regularization parameters.

After computing the coefficients $\mathbf{d}_i, \mathbf{c}_i, \boldsymbol{\beta}_i, i = 1, 2, \cdots, N$, the final HRMS image can be estimated by $\widehat{\mathbf{s}}_i = \mathbf{T}^h \mathbf{d}_i + \mathbf{K}^h \mathbf{c}_i + \Psi^h \boldsymbol{\beta}_i$.

C. Rewrite the Model (13)

For computation convenience, we write (13) into a simpler matrix-vector form as below:

$$\min_{\mathbf{x}} \frac{1}{2N_{um}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \frac{\mu}{2} \|\mathbf{B}\mathbf{x}\|_2^2 + \frac{\lambda_1}{2} \|\mathbf{C}\mathbf{x}\|_1$$
$$+ \frac{\lambda_2}{2} \|\mathbf{D}\mathbf{x} - \mathbf{p}\|_2^2 + \frac{\kappa}{2} \|\mathbf{T}_{oe} \mathbf{x}\|_1, \quad (14)$$

where

$$\mathbf{y} = (\widetilde{\mathbf{s}}_1^T, \widetilde{\mathbf{s}}_2^T, \cdots, \widetilde{\mathbf{s}}_N^T)^T \in \mathbb{R}^{m_2 n_2 N \times 1},$$

$$\mathbf{x} = \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{c}_1 \\ \boldsymbol{\beta}_1 \\ \vdots \\ \mathbf{d}_N \\ \mathbf{c}_N \\ \boldsymbol{\beta}_N \end{pmatrix} \in \mathbb{R}^{(M+m_1 n_1 + m)N \times 1},$$

the diagonal block matrix

$$\mathbf{A} = \text{diag}(\mathbf{E}, \mathbf{E}, \cdots, \mathbf{E}) \in \mathbb{R}^{m_2 n_2 N \times (M+m_1 n_1 + m)N}, \quad (15)$$

where $\mathbf{E} = (\mathbf{T}^h, \mathbf{K}^h, \Psi^h)$, and

$$\mathbf{B} = \text{diag}(\mathbf{F}, \mathbf{F}, \cdots, \mathbf{F}), \quad (16)$$

where $\mathbf{F} = (\mathbf{O}_1, (\mathbf{K}^l)^{1/2}, \mathbf{O}_2)$ with zero matrices $\mathbf{O}_1 \in \mathbb{R}^{m_1 n_1 \times M}$ and $\mathbf{O}_2 \in \mathbb{R}^{m_1 n_1 \times m}$.

$$\mathbf{C} = \text{diag}(\mathbf{G}, \mathbf{G}, \cdots, \mathbf{G}). \quad (17)$$

where $\mathbf{G} = \text{diag}(\mathbf{O}_3, \mathbf{O}_4, \mathbf{I})$ with zero matrices $\mathbf{O}_3 \in \mathbb{R}^{M \times M}$, $\mathbf{O}_4 \in \mathbb{R}^{m_1 n_1 \times m_1 n_1}$, and identity matrix $\mathbf{I} \in \mathbb{R}^{m \times m}$.

$$\mathbf{D} = (\omega_1 \mathbf{T}^h, \omega_1 \mathbf{K}^h, \omega_1 \Psi^h, \cdots, \omega_N \mathbf{T}^h, \omega_N \mathbf{K}^h, \omega_N \Psi^h), \quad (18)$$

where the matrix block $(\omega_i \mathbf{T}^h, \omega_i \mathbf{K}^h, \omega_i \Psi^h)$ will appear N times, where $i = 1, \cdots, N$.

After computing the coefficient $\mathbf{x}$, it is easy to estimate the final pansharpening image with vector form by

$$\widehat{\mathbf{s}} = \mathbf{A}\mathbf{x}. \quad (19)$$

D. The Solution of the Proposed Model

The proposed model (14) is a convex and nonsmooth ℓ_1 optimization problem that can be solved using various methods such as alternating direction method of multipliers (ADMM) [46]–[48] and a primal-dual approach [49]. We adopt ADMM here. By introducing two auxiliary variables $\mathbf{u} = \mathbf{C}\mathbf{x}$ and $\mathbf{w} = \mathbf{T}_{oe}\mathbf{x}$, we get the following augmented Lagrangian problem

$$\mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{w}, \mathbf{b}_1, \mathbf{b}_2) = \frac{1}{2N_{um}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \frac{\mu}{2} \|\mathbf{B}\mathbf{x}\|_2^2$$
$$+ \frac{\lambda_2}{2} \|\mathbf{D}\mathbf{x} - \mathbf{p}\|_2^2 + \frac{\lambda_1}{2} \|\mathbf{u}\|_1$$
$$+ \frac{\eta_1}{2} \|\mathbf{u} - \mathbf{C}\mathbf{x} + \mathbf{b}_1\|_2^2 + \frac{\kappa}{2} \|\mathbf{w}\|_1$$
$$+ \frac{\eta_2}{2} \|\mathbf{w} - \mathbf{T}_{oe}\mathbf{x} + \mathbf{b}_2\|_2^2, \quad (20)$$

where $\mathbf{b}_1$ and $\mathbf{b}_2$ are Lagrangian multipliers with proper size, η_1 and η_2 are two positive parameters.

The problem of minimizing $\mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{w}, \mathbf{b}_1, \mathbf{b}_2)$ can be solved by iteratively and alternatively solving the following three simpler subproblems:

a) The $\mathbf{x}$-subproblem is given as follows

$$\min_{\mathbf{x}} \frac{1}{2N_{um}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \frac{\mu}{2} \|\mathbf{B}\mathbf{x}\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{D}\mathbf{x} - \mathbf{p}\|_2^2$$
$$+ \frac{\eta_1}{2} \|\mathbf{u} - \mathbf{C}\mathbf{x} + \mathbf{b}_1\|_2^2 + \frac{\eta_2}{2} \|\mathbf{w} - \mathbf{T}_{oe}\mathbf{x} + \mathbf{b}_2\|_2^2, \quad (21)$$

which is a least squares problem and has the following closed-form solution:

$$\mathbf{x}^{k+1} = \mathbf{H}^{-1} \mathbf{z}^k, \quad (22)$$

where

$$\mathbf{H} = \mathbf{A}^T \mathbf{A} + \mu N_{um} \mathbf{B}^T \mathbf{B} + \lambda_2 N_{um} \mathbf{D}^T \mathbf{D}$$
$$+ \eta_1 N_{um} \mathbf{C}^T \mathbf{C} + \eta_2 N_{um} \mathbf{T}_{oe}^T \mathbf{T}_{oe}, \quad (23)$$

and

$$\mathbf{z}^k = \mathbf{A}^T \mathbf{y} + \lambda_2 N_{um} \mathbf{D}^T \mathbf{p} + \eta_1 N_{um} \mathbf{C}^T (\mathbf{u}^k + \mathbf{b}_1^k)$$
$$+ \eta_2 N_{um} \mathbf{T}_{oe}^T (\mathbf{w}^k + \mathbf{b}_2^k). \quad (24)$$

Algorithm 1 The Solution of Model (14) by ADMM

Input: Initial image $\widetilde{\mathbf{s}}$ (*i.e.*, $\mathbf{y}$), panchromatic image
    $\mathbf{p}$, and weights $\omega_i, i = 1, 2, \cdots, N$
Output: The coefficient $\mathbf{x}$
Initialize:
    1) $j \leftarrow 0$, $\mathbf{x}^0 \leftarrow \mathbf{0}$, $\mathbf{u}^0 \leftarrow \mathbf{0}$, $\mathbf{w}^0 \leftarrow \mathbf{0}$
    2) Construct $\mathbf{T}^h, \mathbf{K}^h, \mathbf{\Psi}^h$ as in Sect. II
    3) Recall Eqs. (15)-(18) to construct $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ by $\mathbf{T}^h$,
        $\mathbf{K}^h$, $\mathbf{\Psi}^h$, $\mathbf{K}^l$ and $\mathbf{T}_{oe}$
While not converged do
    4) $j \leftarrow j + 1$
    5) Solve $\mathbf{x}^{(j)}$ by Eq. (22)
    6) Solve $\mathbf{u}^{(j)}$ by Eq. (26)
    7) Solve $\mathbf{w}^{(j)}$ by Eq. (28)
    8) Update Lagrangian multipliers $\mathbf{b_1}$, $\mathbf{b_2}$ by Eq. (29)
Endwhile

Please note that the computation of (22) is done on patches whose sizes are significantly smaller than the whole image, thus computing $\mathbf{x}^{k+1}$ by using inverse is not a challenging problem.

b) The $\mathbf{u}$-subproblem is solved accurately by minimizing the following problem:

$$\min_{\mathbf{u}} \frac{\lambda_1}{2}\|\mathbf{u}\|_1 + \frac{\eta_1}{2}\|\mathbf{u} - \mathbf{Cx} + \mathbf{b}_1\|_2^2, \tag{25}$$

which has a closed-form solution in terms of soft-thresholding [50]:

$$\mathbf{u}^{k+1} = \mathbf{shrink}\left(\mathbf{Cx}^{k+1} - \mathbf{b}_1^k, \frac{\lambda_1}{2\eta_1}\right), \tag{26}$$

where $\mathbf{shrink}(a, b) = \mathbf{sign}(a)\mathbf{max}(|a| - b, 0)$ and

$$\mathbf{sign}(a) = \begin{cases} 1, & a > 0, \\ 0, & a = 0, \\ -1, & a < 0. \end{cases}$$

c) Similarly, the $\mathbf{w}$-subproblem is shown as follows

$$\min_{\mathbf{w}} \frac{\kappa}{2}\|\mathbf{w}\|_1 + \frac{\eta_2}{2}\|\mathbf{w} - \mathbf{T}_{oe}\mathbf{x} + \mathbf{b}_2\|_2^2, \tag{27}$$

and has the following closed form solution:

$$\mathbf{w}^{k+1} = \mathbf{shrink}\left(\mathbf{T}_{oe}\mathbf{x}^{k+1} - \mathbf{b}_2^k, \frac{\kappa}{2\eta_2}\right). \tag{28}$$

d) Update the Lagrangian multipliers $\mathbf{b_1}$ and $\mathbf{b_2}$ by:

$$\begin{aligned} \mathbf{b}_1^{k+1} &= \mathbf{b}_1^k + \left(\mathbf{u}^{k+1} - \mathbf{Cx}^{k+1}\right), \\ \mathbf{b}_2^{k+1} &= \mathbf{b}_2^k + \left(\mathbf{w}^{k+1} - \mathbf{T}_{oe}\mathbf{x}^{k+1}\right). \end{aligned} \tag{29}$$

In Algorithm 1, we list all steps of the ADMM algorithm used to iteratively and alternatively solve (14).

Note that the convergence of Algorithm 1 for the separable problem (20) is guaranteed by [51].

Although the proposed model (14) obtains competitive results, due to errors in basis selection and computation, there

Algorithm 2 The Summarized Algorithm for the Proposed Pansharpening Method

---

Input: LRMS image $\mathbf{s}$, panchromatic image $\mathbf{p}$, ω_i,
        $i = 1, 2, \cdots, N$, the number of outer iterations τ
Output: HRMS image $\widehat{\mathbf{s}}^{final}$ (vector form)
Initialize:
    1) Set $\mathbf{s}^{(1)} \leftarrow \mathbf{s}$, $\mathbf{p}^{(1)} \leftarrow \mathbf{p}$
For k = 1: τ
    2) Upscale $\mathbf{s}^{(k)}$ to get $\widetilde{\mathbf{s}}^{(k)}$ via an initial method
    3) Obtain the coefficient $\mathbf{x}$ by **Algorithm 1**
    4) Compute the pansharpened image by Eq. (19)
        $$\widehat{\mathbf{s}}^{(k)} = \mathbf{Ax}$$
    5) Downgrade $\widehat{\mathbf{s}}^{(k)}$ to coarse grid to get $\mathbf{d}_{MS}$
    6) Compute the residual image on the coarse grid by
        $$\mathbf{s}^{(k+1)} = \mathbf{s}^{(k)} - \mathbf{d}_{MS}$$
    7) Compute the residual of panchromatic image by
        $$\mathbf{p}^{(k+1)} = \mathbf{p}^{(k)} - \sum_{i=1}^{N} \omega_i\widehat{\mathbf{s}}_i^{(k)}$$
Endfor
    8) Compute the final high-resolution image:
        $$\widehat{\mathbf{s}}^{final} = \sum_{k=1}^{\tau} \widehat{\mathbf{s}}^{(k)}$$

---

might still be errors in the result. The difference between the given low spatial-resolution MS image (LRMS) and the downgraded version of the computed high spatial-resolution MS image still contains many important details, such as edge information. Thus, inspired by the iterative regularization methods [37], [52], we try to take this difference, as well as with corresponding panchromatic image residual, into the model (14) again to iteratively recover more spectral and spatial details. In particular, the following Algorithm 2 is summarized for the pansharpening problem.

In Algorithm 2, we propose an iterative algorithm based on model (14) and Algorithm 1 for better performance. When $\tau = 1$, Algorithm 2 is the same as Algorithm 1. When $\tau > 1$, Algorithm 2 iteratively picks up more details which is added to the final high-resolution image. Steps 2)-4) obtain a preliminary high-resolution image and step 5) downgrades the preliminary high-resolution image to a low-resolution. If the preliminary image is already good enough, the downgraded result should be close to the given low-resolution image. Otherwise, differences, mainly details of images, will be observed. Step 6) and 7) compute the difference and feed it as input to Algorithm 1 to pick up more details in the next iteration. The detailed picked up will be added to form the final result.

In particular, if we start from the computed high-resolution MS image and manually downgrade it, errors in the computed high-resolution MS image will be reflected in a large difference between this simulated low-resolution MS image and given low-resolution MS image (LRMS). We propose to pick up more image details from the difference (step 6 in Algorithm 2). Accordingly, the residual of panchromatic image can be estimated by the difference between the original panchromatic image and the weighted sum of estimated HRMS image (see step 7 in Algorithm 2). This iterative strategy has been used by some image applications, *e.g.*, natural

image super-resolution [53], remote sensing image pansharpening [54]. The final HRMS image $\widehat{\mathbf{s}}^{final}$ (vector form) is generated by the sum of estimated HRMS images $\widehat{\mathbf{s}}^{(k)}$ on each iterations.

Note that the "Downgrade" in step 5 of Algorithm 2 is done through the following procedure: $\widehat{\mathbf{s}}^{(k)}$ is first filtered by a Gaussian filter matched with the modulation transfer function (MTF) of the MS sensor used to acquire the image and is then downscaled to the size of $\mathbf{s}$ by downsampling strategy. Furthermore, we take GS method [8] as the initial guess method in the step 2 of Algorithm 2. For simplicity purpose, we set the outer layer iteration number τ to be 5. More iterations might lead to slightly better results but cause a greater computation burden. One can also make it automatic by stopping the outer layer loop when the relative error between two adjacent iterations is small.

Algorithm 2 can be applied on the whole image or image patches. In our experiments, we only apply the algorithm to image patches to reduce computation time and storage. In particular, the inverse of $\mathbf{H}$ in (22) can be efficiently computed for patches. In our work, we set patch size to be 6×6 with 3 pixels overlap between patches.

In what follows, we will compare the proposed approach with some competitive methods.

IV. RESULTS AND DISCUSSIONS

In this section, we compare the proposed method with extensive pansharpening methods on different datasets, *i.e.*, Pléiades, Toulouse, China and Rio, acquired by three different sensors: Pléiades, IKONOS and WorldView-2. Note that, due to the space limit, we have to move the results of 8 bands WorldView-2 data (Rio) to supplementary materials, please find it. We utilize various panchromatic images and multispectral images with different sizes in our experiments. The scale factors are all set as 4 in all experiments. The number of bands is 4 for Pléiades, Toulouse and China datasets (*i.e.*, $N = 4$) and 8 for Rio dataset (*i.e.*, $N = 8$). The experiments are implemented in MATLAB(R2013b) on a computer of 16Gb RAM and Intel(R) Core(TM) i5-4590 CPU: @3.30 GHz 3.30 GHz.

For the parameters in the proposed method, we empirically set $\lambda_1 = 2.4 \times 10^{-6}$, $\lambda_2 = 1.2 \times 10^{-5}$, $\mu = 1.2 \times 10^{-7}$, $\eta_1 = 2.4 \times 10^{-7}$, $\eta_2 = 2.4 \times 10^{-5}$, $\kappa = 2.4 \times 10^{-5}$, the outer layer iteration number $\tau = 5$ for all experiments. Note that, fine tuning of parameters for different datasets may lead to better results, but we unify the parameter selection to illustrate the stability of the proposed method. More discussion on the selection of parameters can be found at the part *D* of this section. For the weights ω_i in model (13), we take an automatic approach which has been presented in [45] to estimate them. This approach estimates the weights by a linear regression between the multispectral image and the spatially degraded panchromatic image. Moreover, we employ some standard indexes to estimate the performance of different methods, *i.e.*, Q [55], Q4 [56], Q8 [57], SAM [58], ERGAS [59] and SCC (*i.e.*, spatial correlation coefficient). In particular, the ideal values for Q, Q4, Q8 and SCC are 1 whereas the ideal values for ERGAS and SAM are 0.

A. Dataset

In our experiments, we utilize different datasets to evaluate the performance of different methods. These datasets are all common and publicly available, including: **a)** *Pléiades1 dataset (4 bands)*[3] which is kindly made available by CNES for the 2006 contest [1]. For this dataset, the *high-resolution panchromatic image* is simulated only by the average of green and red bands, *i.e.*, thus holding the weights $\boldsymbol{\omega} = [0, 0.5, 0.5, 0]$ on Blue, Green, Red and Near-Infrared (NIR) Channel, respectively. The *low-resolution multispectral bands* with spatial resolution of four times lower than that of the panchromatic (*i.e.*, LRMS image) are simulated according to the Walds protocol, namely by MTF filtering and decimation. The radiometric resolution is 11-bits. **b)** *Pléiades2 dataset (4 bands)*. Different with (a), the *high-resolution panchromatic image* for the second Pléiades dataset (*i.e.*, Pléiades2) is obtained as follows: 1) averaging the green and red channels; 2) applying the nominal MTF of the panchromatic camera; 3) resampling the outcome to 80 cm; 4) adding the instrument noise; and 5) recovering the ideal image by means of inverse filtering and wavelet denoising. The weights are unknown in this case, thus we here estimate the weights by the linear regression of the multispectral image and the spatially degraded panchromatic image [45]. **c)** *Toulouse dataset (4 bands)* which is acquired on the urban area of Toulouse by the IKONOS sensor, in France, on May 15, 2000. For this dataset, the weights are also acquired automatically by the same way as (b). **d)** *China dataset (4 bands)* [4] which is acquired by the IKONOS sensor represents a mountainous and vegetated area of the Sichuan region in China. Besides, the weights are also obtained by the same way as (b). Readers can find more details about these datasets from the literature [4].

B. Algorithms

- EXP: MS image interpolation, using a polynomial kernel with 23 coefficients [11].
- PCA: Principal Component Analysis [7].
- IHS: Generalized version of the Intensity-Hue-Saturation (GIHS) image fusion [60].
- Brovey: Brovey transform [61].
- BDSD: Band-Dependent Spatial-Detail with local parameter estimation [62].
- GS: Gram Schmidt (Mode 1) [8].
- PRACS: Partial Replacement Adaptive Component Substitution [63].
- HPF: High-Pass Filtering with 5×5 box filter for $1 : 4$ fusion [6].
- SFIM: Smoothing Filter-based Intensity Modulation [64], [65].
- Indusion: Decimated Wavelet Transform using an additive injection model [66].
- ATWT: Additive a Trous Wavelet Transform with unitary injection model [67].

[3]http://openremotesensing.net/knowledgebase/a-critical-comparison-among-pansharpening-algorithms/
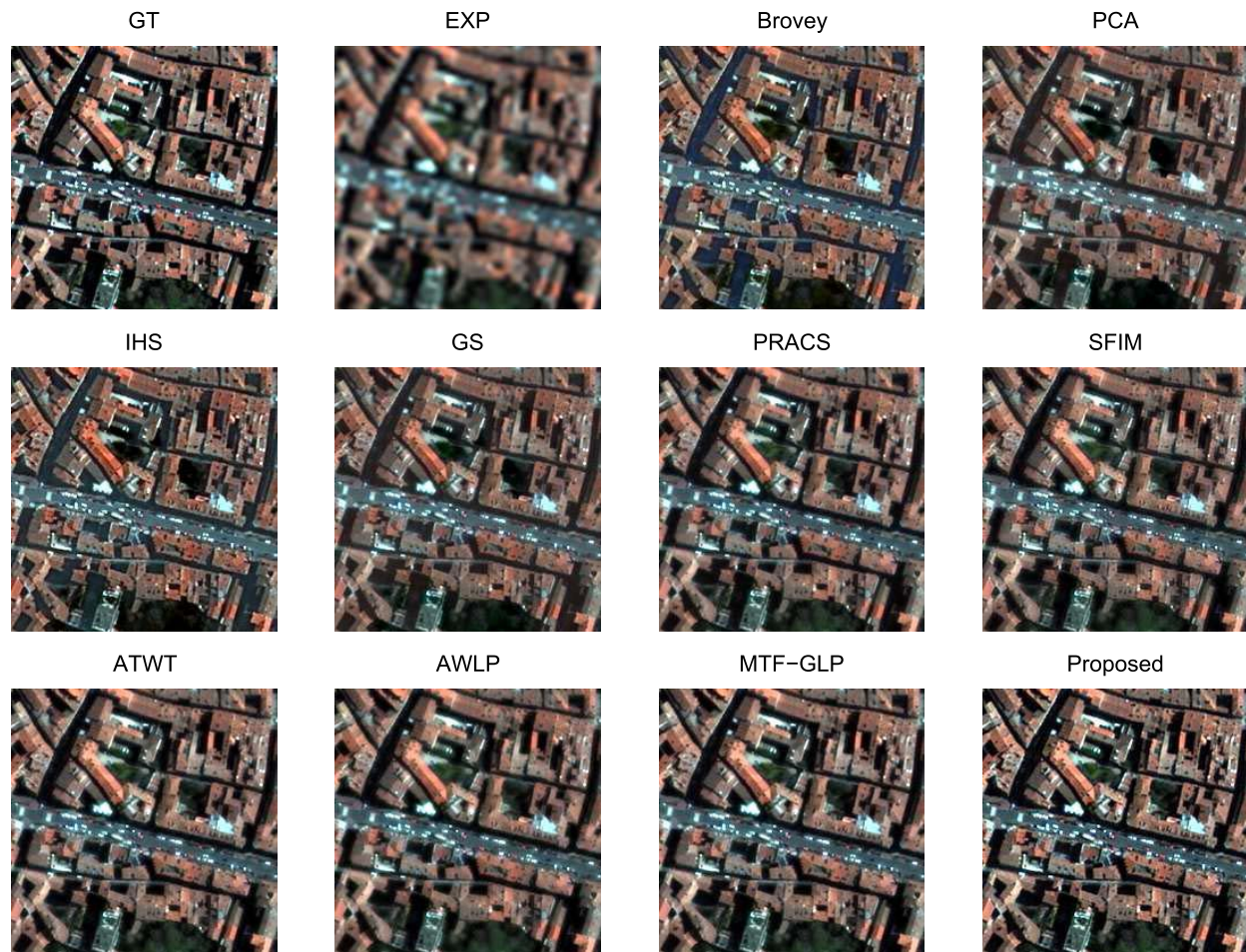
[4]http://glcf.umiacs.umd.edu

Fig. 1. Visual results of **Pléiades2** dataset. Readers are recommended to zoom in all figures for better visibility.

TABLE I

QUANTITATIVE RESULTS FOR 4 BANDS PLÉIADES1 DATASET. (BOLD: BEST; UNDERLINE: SECOND BEST, ITALIC: THIRD BEST)

	Q4	Q	SAM	ERGAS	SCC
EXP	0.7878	0.7629	4.6685	5.7915	0.6500
PCA	0.8811	0.8940	5.0165	4.3882	0.9218
IHS	0.8748	0.8936	5.2947	4.1695	0.9155
Brovey	0.8834	0.8966	4.6685	4.0771	0.9292
BDSD	*0.9765*	0.9734	3.4606	*2.3104*	0.9413
GS	0.8975	0.9040	4.7511	3.9996	0.9325
PRACS	0.9511	0.9449	3.8740	3.0323	0.9283
HPF	0.9462	0.9397	3.6215	3.1337	0.9234
SFIM	0.9473	0.9427	3.5802	3.0730	0.9338
Indusion	0.8924	0.8894	4.5325	4.3842	0.8658
ATWT	0.9631	0.9571	3.2633	2.6712	0.9366
AWLP	0.9617	0.9533	*3.2526*	2.7201	0.9294
ATWT-M2	0.8585	0.8449	4.8020	4.7684	0.8796
ATWT-M3	0.8868	0.8740	5.1776	4.5425	0.8868
MTF-GLP	0.9684	0.9627	3.1959	2.5103	0.9406
MTF-GLP-HPM-PP	0.9552	0.9504	3.6976	2.9398	0.9319
MTF-GLP-HPM	0.9707	0.9677	3.2566	2.3875	0.9487
MTF-GLP-CBD	0.9773	*0.9730*	3.1081	2.2222	0.9422
CNMF	0.9563	0.9548	4.4880	2.8997	*0.9452*
RKHSPan	**0.9789**	**0.9770**	**3.0070**	**1.9979**	**0.9517**

TABLE II

QUANTITATIVE RESULTS FOR 4 BANDS PLÉIADES2 DATASET

	Q4	Q	SAM	ERGAS	SCC
EXP	0.7878	0.7629	4.6685	5.7915	0.6500
PCA	0.8772	0.8896	5.0484	4.4843	0.9150
IHS	0.8750	0.8931	5.2664	4.2207	0.9109
Brovey	0.8829	0.8955	4.6685	4.1415	0.9239
BDSD	*0.9733*	*0.9703*	3.6455	*2.4611*	0.9358
GS	0.8943	0.9004	4.7790	4.0834	0.9264
PRACS	0.9419	0.9354	3.9436	3.2840	0.9210
HPF	0.9413	0.9348	3.7075	3.2726	0.9185
SFIM	0.9421	0.9373	3.6474	3.2253	0.9287
Indusion	0.8897	0.8866	4.5589	4.4465	0.8612
ATWT	0.9585	0.9525	3.3764	2.8409	0.9305
AWLP	0.9571	0.9487	3.3590	2.8760	0.9236
ATWT-M2	0.8550	0.8411	4.8226	4.8283	0.8738
ATWT-M3	0.8795	0.8664	5.2011	4.6612	0.8780
MTF-GLP	0.9638	0.9582	*3.3154*	2.6900	0.9344
MTF-GLP-HPM-PP	0.9505	0.9456	3.7659	3.1030	0.9255
MTF-GLP-HPM	0.9661	0.9630	3.3328	2.5790	0.9428
MTF-GLP-CBD	0.9735	*0.9694*	3.3009	2.4051	0.9367
CNMF	0.9513	0.9497	4.5083	3.0692	*0.9399*
RKHSPan	**0.9754**	**0.9738**	**3.1935**	**2.2201**	**0.9456**

- AWLP: Additive Wavelet Luminance Proportional [9], a generalization of AWL [68] to more than three bands.
- ATWT-M2: A Trous Wavelet Transform using the Model 2 proposed in [69].
- ATWT-M3: A Trous Wavelet Transform using the Model 3 proposed in [69].
- MTF-GLP: Generalized Laplacian Pyramid (GLP) [11] with MTF-matched filter [70] with unitary injection model [67].
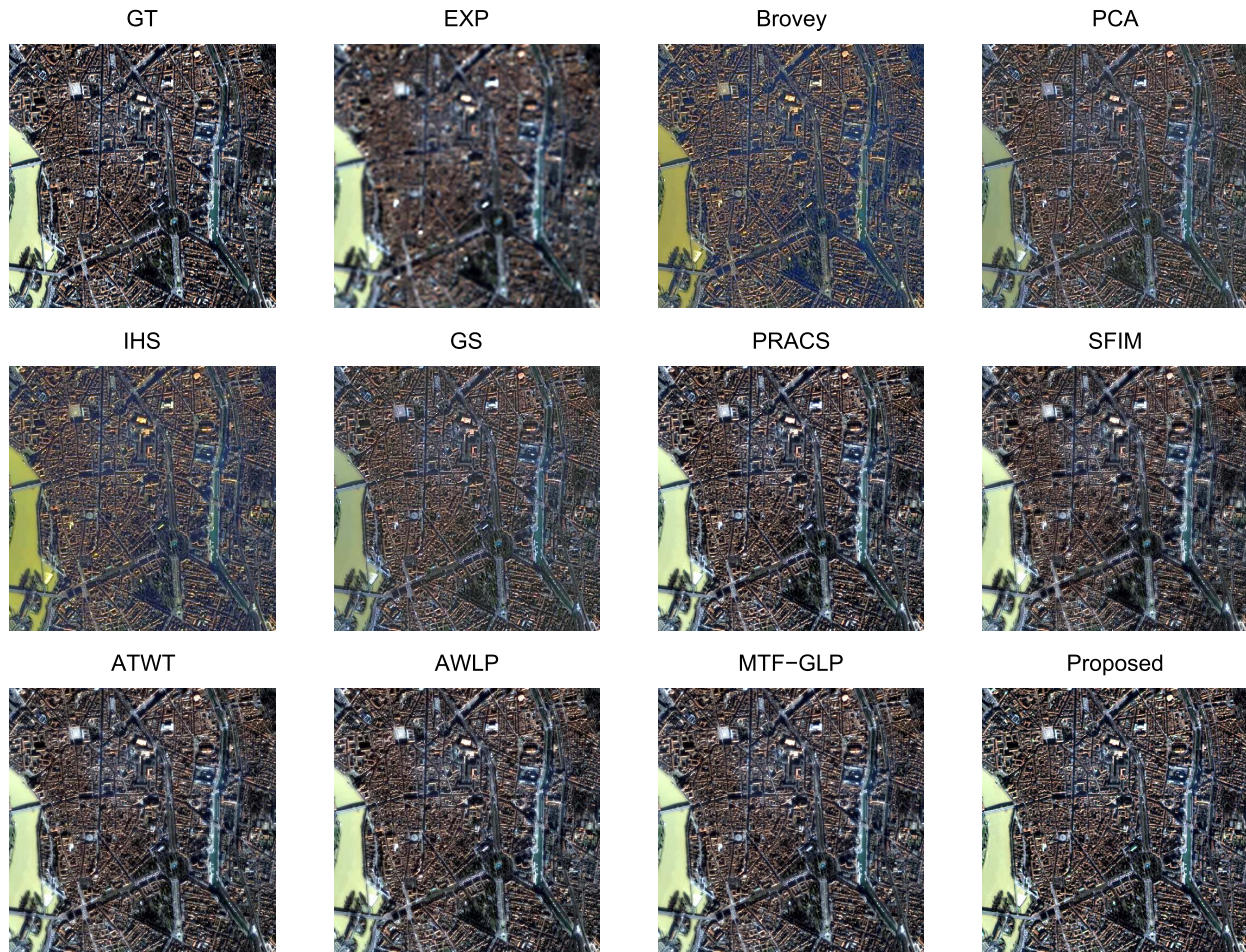
Fig. 2. Visual results of **Toulouse** dataset.

- MTF-GLP-HPM: GLP with MTF-matched filter [70] and multiplicative injection model [67].
- MTF-GLP-HPM-PP: GLP with MTF-matched filter [70], multiplicative injection model and Post-Processing [71].
- MTF-GLP-CBD: GLP with MTF-matched filter [70] and regression based injection model [1], [11].
- CNMF: Coupled nonnegative matrix factorization method [28].
- RKHSPan: The proposed RKHS and Heaviside based pansharpening method.

Note that the source codes of all comparing methods can be found from the website.[5]

C. The Results on Different Datasets

1) Pléiades Dataset: In this subsection, we compare our method with several methods on the two Pléiades datasets, *i.e.*, Pléiades1 and Pléiades2. From Tab. I, we see that the proposed method obtains the best quantitative performance on the Pléiades1 dataset than other comparing methods which include extensive CS-based methods, MRA-based methods and CNMF method. For the Pléiades2 dataset, it is clear that the proposed method also outperforms other all comparing

methods on all evaluation indexes, *i.e.*, Q4, Q, SAM, ERGAS and SCC, see Tab. II. Furthermore, the proposed method also obtains the best visual performance and holds more spatial details than other comparing methods, see Fig. 1.[6] In particular, the weights are automatically estimated as [0.1015, 0.4225, 0.4586, 0.0069] via the strategy mentioned before. CS-based methods, *e.g.*, PCA and GS, drop spectral information significantly while keeping spatial information well. Moreover, MRA-based methods, *e.g.*, AWLP, do not perform excellently on preserving spatial details but keeping spectral information well. Our method not only recovers more spatial details significantly due to the sharp AHF and the iterative algorithm, but also resists the spectral distortion effectively.

2) Toulouse Dataset: For Toulouse dataset, Fig. 2 shows that our method holds more spatial image details than others, as well as keep better spectral performance. In Tab. III, the proposed method ranks the second best on Q4, Q and ERGAS indexes among 20 comparing methods, while the method BDSD has the best quantitative performance.

---

[5]http://openremotesensing.net/kb/codes/pansharpening/

[6]Note that, to save space, we only show the visual results of some representative and competitive methods in our experiments.

GT EXP Brovey PCA
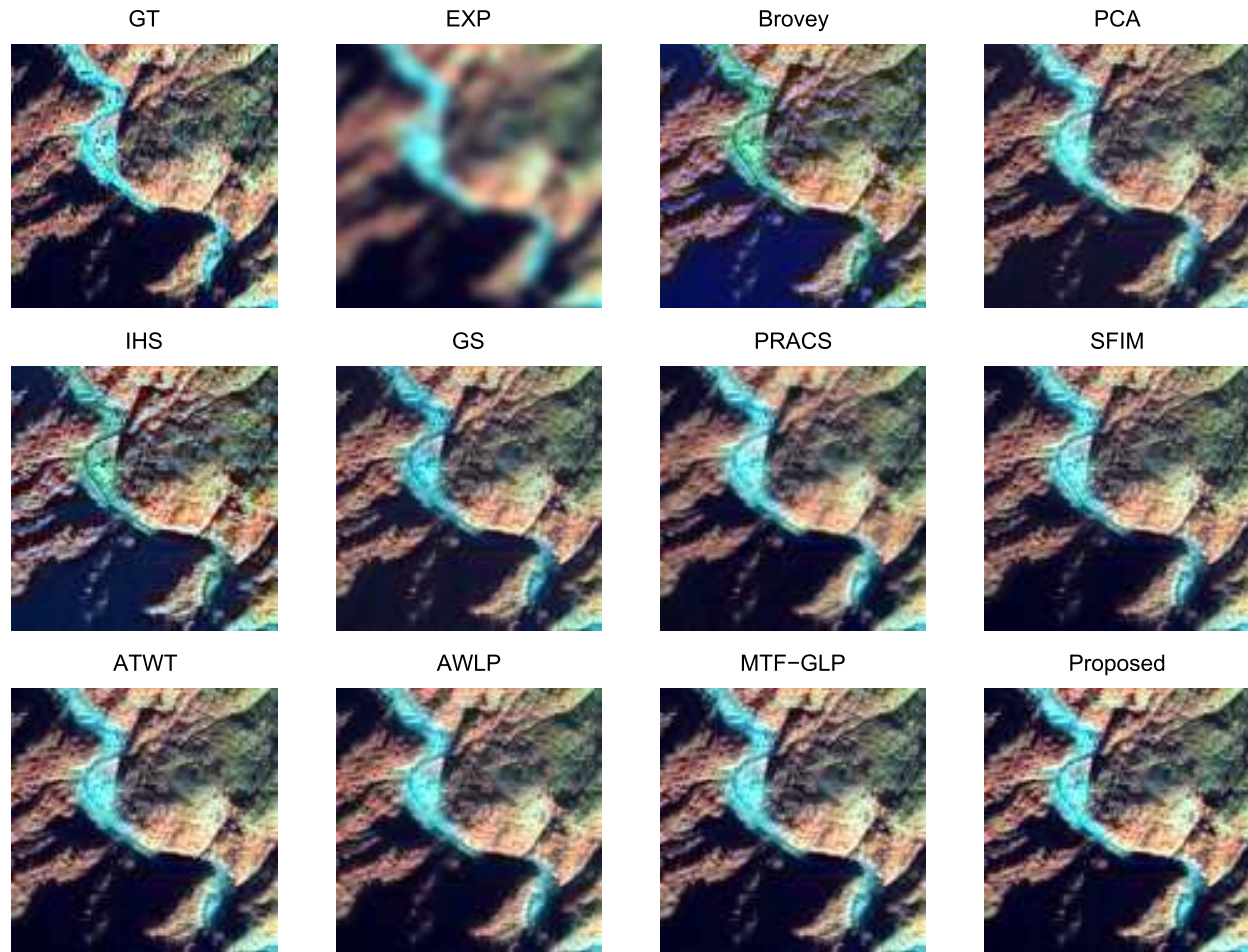
IHS GS PRACS SFIM

ATWT AWLP MTF−GLP Proposed

Fig. 3. Visual results of **China** dataset.

TABLE III
QUANTITATIVE RESULTS FOR 4 BANDS TOULOUSE DATASET

	Q4	Q	SAM	ERGAS	SCC
EXP	0.4827	0.4787	5.2005	6.3637	0.4751
PCA	0.7732	0.7798	4.7874	4.7327	0.8915
IHS	0.7797	0.7602	4.9259	4.7575	0.8774
Brovey	0.7726	0.7428	5.2005	4.8975	0.8481
BDSD	**0.9107**	**0.9113**	**3.0077**	**2.9360**	**0.9347**
GS	0.7733	0.7805	4.7220	4.6864	0.8963
PRACS	0.8849	0.8755	3.8562	3.5389	0.9174
HPF	0.8519	0.8462	3.9420	3.8554	0.9014
SFIM	0.8536	0.8485	3.9152	3.8003	0.9062
Indusion	0.7427	0.7450	4.4990	4.8642	0.8261
ATWT	0.8731	0.8670	3.7897	3.6419	0.9179
AWLP	0.8757	0.8588	4.2781	3.8341	0.9035
ATWT-M2	0.6347	0.6328	4.7549	5.3756	0.8336
ATWT-M3	0.7297	0.7184	4.7695	4.9937	0.8732
MTF-GLP	0.8828	0.8766	3.7160	3.5216	0.9231
MTF-GLP-HPM-PP	0.8593	0.8533	4.0018	3.7688	0.9155
MTF-GLP-HPM	0.8878	0.8824	3.7015	3.4034	*0.9288*
MTF-GLP-CBD	*0.9047*	*0.9016*	<u>3.2428</u>	*3.0744*	0.9223
CNMF	0.8868	0.8863	3.6128	3.251	<u>0.9302</u>
RKHSPan	<u>0.9098</u>	<u>0.9047</u>	*3.2528*	<u>2.9653</u>	0.9255

TABLE IV
QUANTITATIVE RESULTS FOR 4 BANDS CHINA DATASET

	Q4	Q	SAM	ERGAS	SCC
EXP	0.8006	0.7916	5.2953	4.4520	0.6793
PCA	0.8998	0.8992	4.2719	3.0363	0.9181
IHS	0.7639	0.7944	5.9671	3.9498	0.8119
Brovey	0.7522	0.8061	5.2953	3.5781	0.8613
BDSD	0.9237	<u>0.9149</u>	3.3210	2.8000	<u>0.9215</u>
GS	0.8934	0.8890	4.3686	3.0929	0.9186
PRACS	0.9196	0.9085	3.6495	2.8542	0.9104
HPF	0.9141	0.9018	3.7200	2.9162	0.9066
SFIM	0.9153	0.9038	3.6534	2.8818	0.9087
Indusion	0.8374	0.8473	4.6643	3.8194	0.8705
ATWT	0.9244	0.9114	3.4225	2.7483	0.9182
AWLP	0.9247	0.9107	3.3312	2.7629	0.9170
ATWT-M2	0.8675	0.8545	4.7461	3.6397	0.8630
ATWT-M3	0.8825	0.8617	5.3248	3.7712	0.8445
MTF-GLP	0.9248	0.9121	3.3569	2.7210	0.9205
MTF-GLP-HPM-PP	0.9150	0.9016	3.6357	2.9628	0.9114
MTF-GLP-HPM	<u>0.9267</u>	*0.9148*	*3.3196*	**2.6731**	**0.9243**
MTF-GLP-CBD	*0.9251*	0.9129	**3.1978**	*2.7015*	*0.9214*
CNMF	0.8792	0.8769	3.9344	3.0044	0.9126
RKHSPan	**0.9276**	**0.9176**	<u>3.2520</u>	<u>2.6971</u>	0.9207

In particular, the estimated weights are [0.0930, 0.1188, 0.3514, 0.4224] fo

3) China Dataset: For China dataset, the proposed method still performs competitively and its weights are estimated as [0.1595, 0.3443, 0.4430, 0.0437]. From Tab. IV, we can see

that our method obtains the best performance on the overall quality index Q4, which takes into account of both radiometric and spectral distortions, and also on the Q index, the second best SAM and ERGAS indexes. The visual performance shown in Fig. 3 also demonstrates the superior of the given method.
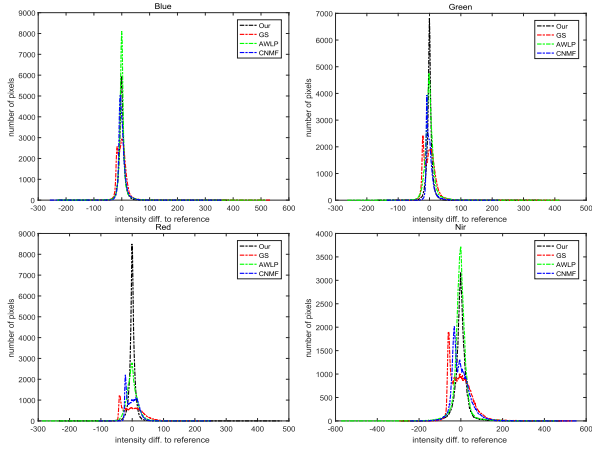
Fig. 4. Intensity difference to ground-truth *v.s.* the number of pixels on each bands (Blue, Green, Red, NIR) on Pléiades1 dataset for GS, AWLP, CNMF and our method. The range of intensity difference to ground-truth is shown in Tab. VI. Please see pdf file for better color visibility.

It is worth mentioning that each of the five metrics has a specific meaning. For instance, SCC only measures the spatial enhancement while SAM measures spectral distortion. Q4 is the unique one that measures both spatial and spectral distortions. According to this metric, the results of the proposed method are the best almost everywhere (3 out of 4). The only exception is the Toulouse dataset, where our approach is ranked in the second place only after the state-of-the-art BDSD approach, which usually works really fine when we have urban scenarios acquired by 4-bands sensors like IKONOS. The difference between the two Q4âŁ™s is 0.0009 which indicates the two approaches get practically the same performance. Thus, our approach can be considered as the best because it is able to adapt to all data without tuning parameters.

D. More Discussions

- **The influence of using different initial pre-processing methods**: Tab. V shows the results by different initial methods (*i.e.*, step 2 of Algorithm 2), it is clear that taking GS method as the initial approach obtains the best results. Thus, we choose GS method as the initial method in the paper. A more sophisticated initialization probably would lead to a slightly better final result but we have preferred to keep the proposed method simple, since we would like to focus on the methodological aspects of this contribution rather that with the fine tuning driven by specific data or application. Generally speaking, a rough initial guess does not always produce a low quality outcome of the proposed approach. The proposed Algorithm 2 actually has a self-correction scheme that can absorb some errors introduced in the initial. When there is error, it will be reflected in the difference between the degraded preliminary high-resolution image and the low-resolution input. The iterative scheme in Algorithm 2 will correct accordingly.
- **Intensity difference to ground-truth**: From Fig. 4, the proposed method is more consistent with the true green and red bands (sharper shape), while AWLP method
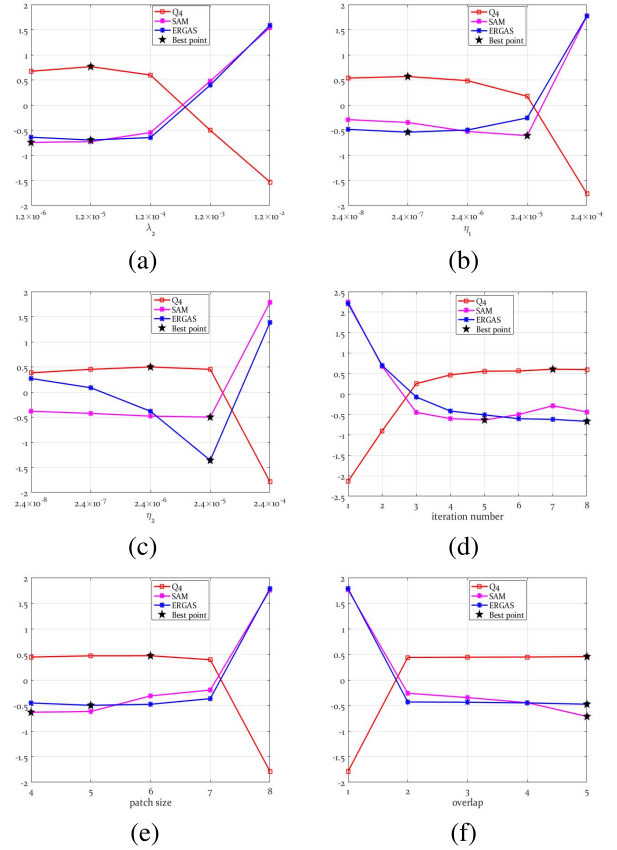

Fig. 5. The influence of different parameters to the proposed method. Although the selected parameters in the work are not always the best ones for all indexes (Q4: red; SAM: pink; ERGAS: blue), they can get the competitive results by fully considering the balance of computation and performance. Note that for better comparisons, here we deal with the obtained indexes by $(\text{index} - \text{mean}(\text{index}))/\text{std}(\text{index})$ where mean and std represent the mean value and standard deviation, respectively.
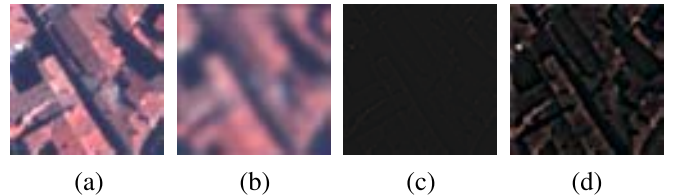

Fig. 6. The estimated components by the proposed method on Pléiades2 dataset. The values are normalized to [0, 1]. For the convenience of computing, we only conduct our method on the top-left part with size $60 \times 60 \times 4$. (a) The estimated image by our method; (b) $\mathbf{T}^h \mathbf{d}_i$; (c) $\mathbf{K}^h \mathbf{c}_i$; (d) $\Psi^h \beta_i$, $i = 1, 2, 3, 4$. For better visibility, we add 0.1 intensity to (c) and (d).

obtains better performance on blue and NIR bands. However, the range of intensity difference to ground-truth for AWLP method is larger than our approach on blue and NIR bands (see Tab. VI), thus this may result in the worse quantitative performance than our method.
- **The influence of related parameters**: In Fig. 5, we take the Pléiades2 dataset to exhibit the performance with varying parameters. For the convenience of computing, we only take the top-left part with size $128 \times 128 \times 4$ for the test. In addition, for better comparisons, here we scale the obtained indexes by $(\text{index} - \text{mean}(\text{index}))/\text{std}(\text{index})$
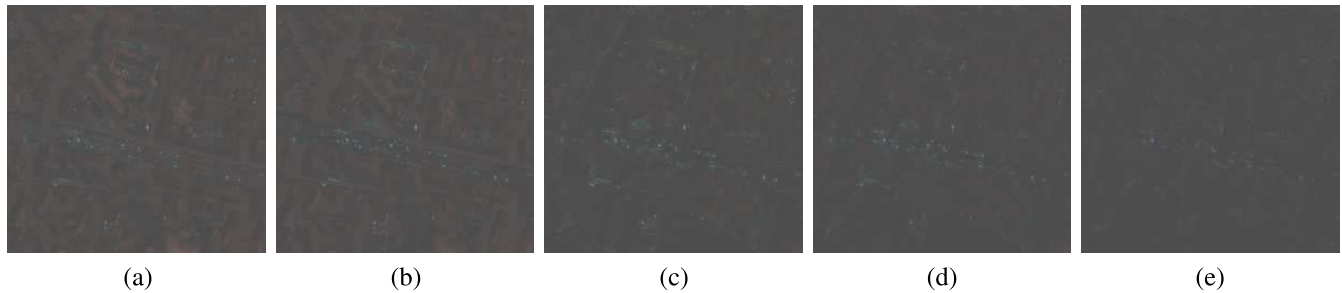
Fig. 7.   Difference maps of some representative methods in Fig. 1 (RGB channels are chosen to be 3 (red), 2 (green), and 1 (blue), respectively): (a) PCA; (b) GS; (c) AWLP; (d) MTF-GLP; (e) Proposed. For better visibility, we add 0.3 to the difference maps which are normalized to [0, 1].

TABLE V
TAKING DIFFERENT METHODS AS INITIAL APPROACHES
ON THE PLÉIADES1 DATASET. (BOLD: BEST)

Initial	Method	Q4	Q	SAM	ERGAS	SCC
EXP	**EXP**	0.7878	0.7629	4.6685	5.7915	0.6500
	RKHSPan	0.9415	0.9187	5.2634	3.5750	0.8232
SFIM	SFIM	0.9421	0.9373	3.6474	3.2253	0.9288
	RKHSPan	0.9745	0.9737	3.0179	2.1255	0.9405
PCA	PCA	0.8772	0.8896	5.0484	4.4843	0.9150
	RKHSPan	0.9729	0.9702	3.0262	2.2544	0.9406
GS	GS	0.8975	0.9040	4.7511	3.9996	0.9325
	RKHSPan	**0.9789**	**0.9770**	**3.0070**	**1.9979**	**0.9517**

TABLE VI
THE RANGE OF INTENSITY DIFFERENCE TO
GROUND-TRUTH, E.G., [MIN, MAX].

Band	GS	AWLP	CNMF	Ours
Blue	[-116, 533]	[-234, 516]	[-259, 361]	[-229, 385]
Green	[-116, 415]	[-263, 418]	[-137, 219]	[-167, 352]
Red	[-106, 240]	[-288, 247]	[-107, 251]	[-235, 477]
NIR	[-296, 534]	[-574, 467]	[-244, 557]	[-287, 484]

where mean and std represent the mean value and standard deviation, respectively. In particular, we fix all parameters suggested in this section except the parameter we want to test, aiming to independently evaluate the parameter sensitivity. From Fig. 5 (a) - (c), we know that the quantitative results (*i.e.*, Q4, SAM and ERGAS) only have slight changes around our suggested parameters λ_2, η_1 and η_2, which demonstrates the stability of our method to parameters. In addition, when λ_1, μ and κ changing, the quantitative results are almost not affected, here we do not present the results of them for the goal of saving space. For Fig. 5 (d), the results with increasing iterations are exhibited. When the iteration number is 1 (*i.e.*, the model (13)), it performs worse, whereas when the iteration number increases, we can get better results. Although we can get the best quantitative result by the maximum iteration, it is quite expensive. To balance the computation and performance, we finally set 5 iterations in this study. Fig. 5 (e) presents the results with different patch size, which indicates the suggested patch size performs competitively for all indexes. Fig. 5 (f) presents the results when setting different patch overlaps. Similar as Fig. 5 (d), although it can get the best quantitative results by the maximum overlap, it is quite time consuming. We therefore finally choose the overlap 3.

- **The estimated components by the proposed method**: In Fig. 6, we show the estimated components by the proposed method. From the figure, it is clear that $\mathbf{T}^h \mathbf{d}_i$, $i = 1, 2, 3, 4$, which is derived from the polynomials of RKHS, depicts the primary image information but without sharp image edge, see Fig. 6 (b). $\Psi^h \boldsymbol{\beta}_i$ which is from AHFs mainly represents the sharp image edges, see Fig. 6 (d). Moreover, $\mathbf{K}^h \mathbf{c}_i$ that is derived from high-order term of RKHS describes some high-frequency image information, see Fig. 6 (c).
- **The comparisons of difference maps**: Fig. 7 reports the difference maps on Pléiades2 dataset for some representative methods. These difference maps are obtained by the corresponding errors on red, blue and green channels. The darker image shows the better details preserving, thus our method significantly performs better than other methods from this figure.

V. CONCLUSIONS

In this paper we presented a variational pansharpening technique based on the image super-resolution method of Deng et al. [37]. The novelty of the proposed pansharpening approach is in the representation of the pansharpened image as a continuous function in which image edges are modeled as approximated Heaviside functions and are considered to appear sparsely in the spatial domain. Considering the correlation of adjacent bands, we designed a Toeplitz sparse term to attain more spectral information. Experiments were conducted in order to prove the effectiveness of the proposed technique against state-of-the-art methods belonging to both the Component Substitution and MultiResolution Analysis approaches [4]. Many acknowledged datasets acquired by different sensors, including Pléiades, IKONOS, WorldView-2, were considered for the tests. The quantitative and visual results showed that the proposed technique outperformed reference pansharpening techniques in terms of spatial and spectral fidelity with the reference image. As future developments we plan to make full use of the toeplitz sparsity by making spectral analysis and extend the proposed technique to hyperspectral image pansharpening.

REFERENCES

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

[2] M. Dalla Mura, S. Prasad, F. Pacifici, P. Gamba, and J. Chanussot, "Challenges and opportunities of multimodality and data fusion in remote sensing," in *Proc. 22nd Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2014, pp. 106–110.

[3] 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.

[4] G. Vivone *et al.*, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.

[5] 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, 2003.

[6] 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, 1991.

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

[8] 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, Jan. 4, 2000.

[9] 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.

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

[11] 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.

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

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

[14] M. Moller, T. Wittman, and A. L. Bertozzi, "A variational approach to hyperspectral image fusion," *Proc. SPIE*, vol. 7334, p. 73341E, Apr. 2009.

[15] 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.

[16] 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.

[17] F. Fang, F. Li, C. Shen, and G. Zhang, "A variational approach for pansharpening," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2822–2834, Jul. 2013.

[18] J. Duran, A. Buades, B. Coll, and C. Sbert, "A nonlocal variational model for pansharpening image fusion," *SIAM J. Imag. Sci.*, vol. 7, no. 2, pp. 761–796, 2014.

[19] 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.

[20] H. A. Aly and G. Sharma, "A regularized model-based optimization framework for pan-sharpening," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2596–2608, Jun. 2014.

[21] X. He, L. Condat, J. M. 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.

[22] C. Chen, Y. Li, W. Liu, and J. Huang, "Image fusion with local spectral consistency and dynamic gradient sparsity," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2014, pp. 2760–2765.

[23] C. Chen, Y. Li, W. Liu, and J. Huang, "SIRF: Simultaneous satellite image registration and fusion in a unified framework," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4213–4224, Nov. 2015.

[24] Y. Jiang, X. Ding, D. Zeng, Y. Huang, and J. Paisley, "Pan-sharpening with a hyper-Laplacian penalty," in *Proc. Int. Conf. Comput. Vis. (ICCV)*, 2015, pp. 540–548.

[25] Q. Wei, N. Dobigeon, J.-Y. Tourneret, J. M. Bioucas-Dias, and S. Godsill, "R-FUSE: Robust fast fusion of multiband images based on solving a Sylvester equation," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1632–1636, Nov. 2016.

[26] M. Ghahremani and H. Ghassemian, "Nonlinear IHS: A promising method for pan-sharpening," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 11, pp. 1606–1610, Nov. 2016.

[27] X. Yang, J. Wang, and R. Zhu, "Random walks for synthetic aperture radar image fusion in framelet domain," *IEEE Trans. Image Process.*, vol. 27, no. 2, pp. 851–865, Feb. 2018.

[28] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.

[29] R. Restaino, G. Vivone, M. Dalla Mura, and J. Chanussot, "Fusion of multispectral and panchromatic images based on morphological operators," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2882–2895, Jun. 2016.

[30] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, "A new pan-sharpening method with deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 1037–1041, May 2015.

[31] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, p. 594, Jul. 2016.

[32] Y. Rao, L. He, and J. Zhu, "A residual convolutional neural network for pan-shaprening," in *Proc. Int. Workshop Remote Sens. Intell. Process. (RSIP)*, 2017, pp. 1–4.

[33] Y. Wei, Q. Yuan, H. Shen, and L. Zhang, "Boosting the accuracy of multispectral image pansharpening by learning a deep residual network," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1795–1799, Oct. 2017.

[34] J. Yang, X. Fu, Y. Hu, Y. Huang, X. Ding, and J. Paisley, "PanNet: A deep network architecture for pan-sharpening," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1753–1761.

[35] N. Li, N. Huang, and L. Xiao, "PAN-sharpening via residual deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 5133–5136.

[36] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.

[37] L.-J. Deng, W. Guo, and T.-Z. Huang, "Single-image super-resolution via an iterative reproducing kernel Hilbert space method," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 11, pp. 2001–2014, Nov. 2016.

[38] L.-J. Deng, G. Vivone, W. Guo, M. Dalla Mura, and J. Chanussot, "A variational pansharpening approach based on reproducible kernel Hilbert space and Heaviside function," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 535–539. [Online]. Available: http://www.escience.cn/people/dengliangjian/paper.dhome

[39] G. Wahba, *Spline Methods for Observational Data*. Philadelphia, PA, USA: SIAM, 1990.

[40] J. Duchon, "Splines minimizing rotation-invariant semi-norms in Sobolev spaces," in *Constructive Theory of Functions of Several Variables*. Springer, 1977, pp. 85–100.

[41] Z. G. Liu, G. Li, G. Mercier, Y. He, and Q. Pan, "Change detection in heterogenous remote sensing images via homogeneous pixel transformation," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 1822–1834, Apr. 2018, doi: 10.1109/TIP.2017.2784560.

[42] Z. Zou and Z. Shi, "Random access memories: A new paradigm for target detection in high resolution aerial remote sensing images," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1100–1111, Mar. 2018.

[43] L. Drumetz, M.-A. Veganzones, S. Henrot, R. Phlypo, J. Chanussot, and C. Jutten, "Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3890–3905, Aug. 2016.

[44] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1852–1866, Jun. 2015.

[45] 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.

[46] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 323–343, 2009.

[47] B. He, M. Tao, and X. Yuan, "Alternating direction method with Gaussian back substitution for separable convex programming," *SIAM J. Optim.*, vol. 22, no. 2, pp. 313–340, 2012.

[48] X.-L. Zhao, F. Wang, T.-Z. Huang, M. K. Ng, and R. J. Plemmons, "Deblurring and sparse unmixing for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4045–4058, Jul. 2013.

[49] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, no. 1, pp. 120–145, 2011.

[50] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[51] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*. Philadelphia, PA, USA: SIAM, 1989.

[52] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 460–489, 2005.

[53] L.-J. Deng, W. Guo, and T.-Z. Huang, "Single image super-resolution by approximated Heaviside functions," *Inf. Sci.*, vol. 348, pp. 107–123, Jun. 2016.

[54] M. R. Vicinanza, R. Restaino, G. Vivone, M. Dalla Mura, G. Licciardi, and J. Chanussot, "A method for improving the consistency property of pansharpening algorithms," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2014, pp. 2534–2537.

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

[56] 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.

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

[58] 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.

[59] L. Wald, *Data Fusion. Definitions and Architectures—Fusion of Images of Different Spatial Resolutions*. Paris, France: Presses de l'Ecole, Ecole des Mines de Paris, 2002.

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

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

[62] 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.

[63] 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.

[64] 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.

[65] L. Wald and T. Ranchin, "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, 2002.

[66] M. M. Khan, J. Chanussot, L. Condat, and A. Montanvert, "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.

[67] 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.

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

[69] 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.

[70] 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.

[71] J. Lee and C. Lee, "Fast and efficient panchromatic sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 155–163, Jan. 2010.
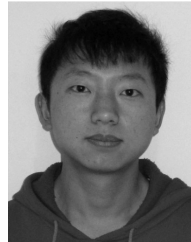
Liang-Jian Deng received the B.S. and Ph.D. degree from the School of Mathematical Sciences, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2010 and 2016, respectively. He is currently a Lecturer with the School of Mathematical Sciences, UESTC. His research interests include image processing and computer vision.

Gemine Vivone (M'17) received the B.Sc. *(summa cum laude)*, M.Sc. *(summa cum laude)*, and Ph.D. degrees in information engineering from the University of Salerno, Salerno, Italy, in 2008, 2011, and 2014, respectively. He is currently with the Department of Information Engineering, Electrical Engineering and Applied Mathematics, University of Salerno. In 2014, he joined the North Atlantic Treaty Organization (NATO) Science and Technology Organization Centre for Maritime Research and Experimentation, La Spezia, Italy, as a Scientist. In 2013, he was a Visiting Scholar with the Grenoble Institute of Technology, Grenoble, France, 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 a Visiting Researcher with the NATO Undersea Research Centre. He serves as a referee for several remote sensing and image processing journals. He received the Symposium Best Paper Award at the IEEE International Geoscience and Remote Sensing Symposium in 2015 and the Best Reviewer Award of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2017. He was a Lead Guest Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS Special Stream Advances in image processing for very high resolution optical remote sensing images.

Weihong Guo received the B.S. degree in computational mathematics from the Minzu University of China in 1999, and the M.S. degree in statistics and the Ph.D. degree in applied mathematics from the University of Florida in 2007. She was an Assistant Professor of mathematics at the University of Alabama from 2007 to 2009. She is currently an Associate Professor of applied mathematics at Case Western Reserve University, Cleveland, OH, USA. Her research interests include image processing, image analysis, image fusion, computer vision and their applications in geoscience, remote sensing, biology and medicine. She served as a Guest Editor for the *International Journal of Biomedical Imaging* in 2011. She has been on the Editorial Board of the *International Journal Inverse Problems and Imaging* since 2012.

Mauro Dalla Mura (S'08–M'11–SM'18) received the B. E. (laurea) and M.E. (laurea specialistica) degrees in telecommunication engineering from the University of Trento, Italy, in 2005 and 2007, respectively, and the joint Ph.D. degree in information and communication technologies (telecommunications area) from the University of Trento, Italy, in 2011 and in electrical and computer engineering from the University of Iceland, Iceland. In 2011, he was a Research Fellow at Fondazione Bruno Kessler, Trento, Italy, where he was conducting research on computer vision.

He is currently an Assistant Professor at the Grenoble Institute of Technology, France. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory. His main research activities include remote sensing, image processing, and pattern recognition. In particular, his research interests include mathematical morphology, classification, and multivariate data analysis.

Dr. Dalla Mura was a recipient of the IEEE GRSS Second Prize in the Student Paper Competition of the 2011 IEEE IGARSS 2011 and a co-recipient of the Best Paper Award of the *International Journal of Image and Data Fusion* from 2012 to 2013 and the Symposium Paper Award for the IEEE IGARSS 2014.

Dr. Dalla Mura has been the President of the IEEE GRSS French Chapter since 2016 (he previously served as Secretary from 2013 to 2016). In 2017, the IEEE GRSS French Chapter was a recipient of the IEEE GRSS Chapter Award and the Chapter of the year 2017 from the IEEE French Section. He is on the Editorial Board of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (J-STARS) since 2016.

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 the Université de Savoie, Annecy, France, in 1998. In 1999, he was with the Geography Imagery Perception Laboratory, Delegation Generale de l'Armement (French National Defense Department). Since 1999, he has been with Grenoble INP, where he is currently a Professor of signal and image processing. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory. He has been a Visiting Scholar at Stanford University, USA, KTH, Sweden, and NUS, Singapore. From 2015 to 2017, he was a Visiting Professor with the University of California at Los Angeles, Los Angeles, CA, USA. Since 2013, he has been an Adjunct Professor with the University of 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 from 2007 to 2010, which received the 2010 IEEE GRS-S Chapter Excellence Award. He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society from 2006 to 2008. He was a member of the IEEE Geoscience and Remote Sensing Society AdCom from 2009 to 2010, in charge of membership development. He is a member of the Institut Universitaire de France from 2012 to 2017. He was a co-recipient of the NORSIG 2006 Best Student Paper Award, the IEEE GRSS 2011 and 2015 Symposium Best Paper Award, the IEEE GRSS 2012 Transactions Prize Paper Award, and the IEEE GRSS 2013 Highest Impact Paper Award. 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 from 2009 to 2011 and the Co-Chair of the GRS Data Fusion Technical Committee from 2005 to 2008. He was the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing in 2009. He was an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS from 2005 to 2007 and *Pattern Recognition* from 2006 to 2008. Since 2007, he has been an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He was the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING from 2011 to 2015. In 2013, he was a Guest Editor of the PROCEEDINGS OF THE IEEE and in 2014, he was a Guest Editor of the IEEE SIGNAL PROCESSING MAGAZINE.