Robust fusion of multi-band images with different spatial and spectral resolutions for change detection

Vinicius Ferraris, Nicolas Dobigeon, Qi Wei, Marie Chabert

Robust Fusion of Multiband Images With Different Spatial and Spectral Resolutions for Change Detection

Vinicius Ferraris, *Student Member, IEEE*, Nicolas Dobigeon, *Senior Member, IEEE*, Qi Wei, *Member, IEEE*, and Marie Chabert, *Member, IEEE*

***Abstract*—Archetypal scenarios for change detection generally consider two images acquired through sensors of the same modality. However, in some specific cases such as emergency situations, the only images available may be those acquired through different kinds of sensors. More precisely, this paper addresses the problem of detecting changes between two multiband optical images characterized by different spatial and spectral resolutions. This sensor dissimilarity introduces additional issues in the context of operational change detection. To alleviate these issues, classical change detection methods are applied after independent preprocessing steps (e.g., resampling) used to get the same spatial and spectral resolutions for the pair of observed images. Nevertheless, these preprocessing steps tend to throw away relevant information. Conversely, in this paper, we propose a method that more effectively uses the available information by modeling the two observed images as spatial and spectral versions of two (unobserved) latent images characterized by the same high spatial and high spectral resolutions. As they cover the same scene, these latent images are expected to be globally similar except for possible changes in sparse spatial locations. Thus, the change detection task is envisioned through a robust multiband image fusion method, which enforces the differences between the estimated latent images to be spatially sparse. This robust fusion problem is formulated as an inverse problem, which is iteratively solved using an efficient block-coordinate descent algorithm. The proposed method is applied to real panchromatic, multispectral, and hyperspectral images with simulated realistic and real changes. A comparison with state-of-the-art change detection methods evidences the accuracy of the proposed strategy.**

***Index Terms*—Change detection, different resolutions, hyperspectral imagery, image fusion, multispectral imagery.**

V. Ferraris, N. Dobigeon, and M. Chabert are with IRIT/INP-ENSEEIHT, University of Toulouse, 31062 Toulouse, France (e-mail: vinicius.ferraris@enseeiht.fr; nicolas.dobigeon@enseeiht.fr; marie.chabert@enseeiht.fr).

Q. Wei is with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA (e-mail: qi.wei@duke.edu).

Digital Object Identifier 10.1109/TCI.2017.2692645

I. Introduction

REMOTE sensing is a reliable technique for Earth surface monitoring and observation [2]. One of the most important applications using remotely sensed data is the so-called change detection (CD) problem. CD has many definitions and it is generally considered as the ability of analyzing two or more multi-date (i.e., acquired at different time instants) and possibly multi-source (i.e., acquired by different sensors) images of the same scene to detect areas where potential changes have occurred [3], [4]. Because of the increasing number of satellites and of new policies for data distribution, more multi-temporal data becomes available. While this increases the amount of information on the present scene, it highlights some additional issues when designing operational change detection techniques.

Each remotely sensed observation image is intimately connected to the acquisition modality providing a particular excerpt of the observed scene according to the sensor specifications. For instance, optical images are generally well suited to map horizontal structures, e.g., land-cover type at large scales [5]. More particularly, remote sensing images acquired by multiband optical sensors can be classified according to their spectral and spatial resolutions. The spectral resolution is related to the capability of sensing the electromagnetic spectrum. This term can also refer to the number of spectral bands [2], [6], which generally leads to a commonly adopted classification of these images: *panchromatic* (PAN) images, characterized by a low spectral resolution, *multispectral* (MS) and *hyperspectral* (HS) images which sense part of the spectrum with higher precision. Alternatively, multi-band optical images can be classified with respect to (w.r.t.) their spatial resolution [2], [5]. The concept of spatial resolution should be understood as the capability of representing the smallest object that can be resolved up to a specific pixel size. Images having fine resolution and finer details are generally identified as of *high resolution* (HR) in contrast to *low resolution* (LR) images where only coarse features are observable. Because of the physical limitations of optical passive sensors, multi-band optical images suffer from a trade-off between spectral and spatial resolution [7], [8]. To ensure that any sensor has sufficient amount of energy to guarantee a proper acquisition (in terms of, e.g., signal-to-noise ratio), one of the resolutions must be decreased allowing the other to be increased. For this reason, PAN images are generally

characterized by higher spatial resolution and lower spectral resolution than MS or HS images.

Optical images have been the most studied remote sensing modality for CD since the widely admitted additive Gaussian modeling of optical optical sensor noises allows CD techniques to be implemented through a simple operation of image differencing [3], [4]. Originally designed for single-band images, CD differencing methods have been adapted to handle multi-band images by considering spectral change vectors [9], [10] and transform analysis [11], [12]. The possibility of detecting changes by exploiting both spatial and spectral information is one of the greatest advantages of these multi-band optical images. Nevertheless, images of same spatial and spectral resolutions, are not always available. In some specific situations, for instance consecutive to natural disasters, acquisitions must be performed in emergency, with the available devices. In this case, CD may require the comparison of images acquired by different kinds of sensors. Such disadvantageous situations yet require fast, flexible and accurate methods able to handle also the incompatibilities introduced by each sensor modality [13]–[16]. Most of the CD classical methods do not support differences in resolutions. Generally, each observed image is independently preprocessed in order to get the same resolution and then classical CD techniques are applied. However, resampling operations independently applied to each image do not take into account their joint characteristics and thus crucial information may be missed. To overcome this limitation, a general CD framework has been recently proposed in [17] to deal with multi-band images with different spatial and spectral resolutions based on a 3-step procedure (fusion, prediction, detection). Instead of independently preprocessing each observed image, this approach consists in recovering a latent (i.e., unobserved) high spatial and spectral resolution image containing changed and unchanged regions by fusing both observed images. Then, it predicts pseudo-observed images by artificially degrading this estimated latent image using the same forward models underlying the actually observed images. As the pairs of predicted and observed observations have the same spatial and spectral resolutions, any classical multi-band CD method can be finally applied to build a change map. Albeit significantly improving detection performance when compared to crude methods relying on independent preprocessing, the 3-step sequential formulation appears to be non-optimal for the following twofold reasons: i) any inaccuracies in the fusion step are propagated throughout the subsequent degradation and detection steps, ii) relevant information regarding the change may be lost during the prediction steps, since it consists in spatially or spectrally degrading the latent images to estimate the pseudo-observed images. Thus, significant improvements in terms of change detection performance may be expected provided one is able to overcome both limitations.

In this paper, capitalizing on the general framework developed in [17], we show that the CD task can be formulated as a particular instance of the multi-band image fusion problem. However, contrary to the 3-step procedure in [17], the proposed approach jointly estimates a couple of distinct latent images corresponding to the two acquisition times as well as the change image. Since the two latent images of high spatial and spectral resolutions are supposed to represent the same scene, they are expected to share a high level of similarity or, equivalently, to differ only in a few spatial locations. Thus, akin to numerous robust factorizing models such as robust principal component analysis [18] and robust nonnegative matrix factorization [19], the two observed images are jointly approximated by a standard linear decomposition model complemented with an outlier term corresponding to the change image. This so-called CD-driven *robust* fusion of multi-band images is formulated as an inverse problem where, in particular, the outlier term is characterized by a spatial sparsity-inducing regularization. The resulting objective function is solved through the use of a block coordinate descent (BCD) algorithm, which iteratively optimizes w.r.t. one latent image and the change image. Remarkably, optimizing w.r.t. the latent image boils down to a classical multi-band image fusion step and can be efficiently conducted following the algorithmic solutions proposed in [20]. The CD map can be finally generated from the recovered change image.

The paper is organized as follows. Section II-B formulates the change detection problem for multi-band optical image. Section III presents the solution for the formulated problem based on robust fusion. Various experimental scenarios are considered in Section IV, based on realistic physically-motivated experimental setups and a pair of real LANDSAT 8 images. Section V concludes the paper.

II. From Change Detection to Robust Fusion

A. Generic Forward Model

Let us consider the image formation process as a sequence of transformations, denoted $T[\cdot]$, of the original scene into an output image. The output image of a particular sensor is referred to as the observed image and denoted $\mathbf{Y} \in \mathbb{R}^{n_\lambda \times m}$ where m and n_λ are the numbers of pixels and spectral bands in the observed image, respectively. It provides a limited version of the original scene with characteristics imposed by the image signal processor (ISP) characterizing the sensor. The original scene can be conveniently represented by an (unknown) latent image of higher spatial and spectral resolutions, $\mathbf{X} \in \mathbb{R}^{m_\lambda \times n}$, where $n \geq m$ and $m_\lambda \geq n_\lambda$ are the numbers of pixels and spectral bands, respectively, related to the observed image following

$$\mathbf{Y} = T[\mathbf{X}]. \tag{1}$$

The intrinsic sequence of transformations of the sensor over the latent image $\mathbf{X}$ can be typically classified as spectral or spatial degradations. On one hand, spatial degradations are related to the spatial characteristics of the sensor such as sampling scheme and optical transfer function. On the other hand, spectral degradations refer to the wavelength sensitivity and the spectral sampling. There are many ways to represent the degradation process. In this paper, it is considered as a sequence of linear operations leading to the following generic forward model [21]–[23]

$$\mathbf{Y} = \mathbf{LXR} + \mathbf{N} \tag{2}$$

where

1) $\mathbf{L} \in \mathbb{R}^{n_\lambda \times m_\lambda}$ is the spectral degradation matrix,

2) $\mathbf{R} \in \mathbb{R}^{n \times m}$ is the spatial degradation matrix,
3) $\mathbf{N}$ is the additive term comprising sensor noise and modeling errors.

In (2), the left-multiplying matrix $\mathbf{L} \in \mathbb{R}^{n_\lambda \times m_\lambda}$ degrades the latent image by combination of some spectral bands for each pixel while the right-multiplying matrix $\mathbf{R} \in \mathbb{R}^{n \times m}$ degrades the latent image by linear combination of pixels within the same spectral band. The former degradation corresponds to a spectral resolution reduction with respect to the latent image $\mathbf{X}$ as in [20], [22], [23]. In practice, this degradation models an intrinsic characteristic of the sensor, namely the spectral response. It can be either learned by cross-calibration or known *a priori* [23], [24]. Conversely, the spatial degradation matrix $\mathbf{R}$ models the combination of different transformations which are specific of the sensor architecture taking into account external factors including warp, blurring, translation and decimation [20], [24]. In this work, since geometrical transformations such as warp and translations can be corrected using image co-registration techniques in pre-processing steps, only a spatially invariant blurring and a decimation (i.e., subsampling) will be considered. A space-invariant blur can be modeled by a symmetric convolution kernel associated with a sparse symmetric Toeplitz matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ which operates a cyclic convolution on each individual band [25]. The decimation operation, denoted by the $n \times m$ matrix $\mathbf{S}$, corresponds to a uniform downsampling operator[1] of factor $d = d_r \times d_c$ with $m = n/d$ ones on the block diagonal and zeros elsewhere, such that $\mathbf{S}^T\mathbf{S} = \mathbf{I}_m$ [20]. To summarize, the overall spatial degradation process corresponds to the matrix composition $\mathbf{R} = \mathbf{BS} \in \mathbb{R}^{n \times m}$.

The noise corrupting multi-band optical images is generally modeled as additive and Gaussian [4], [8], [20], [26]. Thus the noise matrix $\mathbf{N}$ in (2) is assumed to be distributed according to the following matrix normal distribution[2]

$$\mathbf{N} \sim \mathcal{MN}_{n_\lambda, m}(\mathbf{0}_{n_\lambda \times m}, \mathbf{\Lambda}, \mathbf{\Pi}). \tag{3}$$

The row covariance matrix $\mathbf{\Lambda}$ carries information regarding the between-band spectral correlation. Following [20], in what follows, this covariance matrix $\mathbf{\Lambda}$ will be assumed to be diagonal, which implies that the noise is independent from one band to the other and characterized by a specific variance in each band. Conversely, the column covariance matrix $\mathbf{\Pi}$ models the noise correlation w.r.t. to the pixel locations. Following a widely admitted hypothesis of the literature, this matrix is assumed to be identity, $\mathbf{\Pi} = \mathbf{I}_m$, to reflect the fact the noise is spatially independent. In real applications, both matrices $\mathbf{\Lambda}$ and $\mathbf{\Pi}$ can be estimated by previous calibrations [24].

[1]The corresponding operator $\mathbf{S}^T$ represents an upsampling transformation by zero-interpolation from m to n.

[2]The probability density function $p(\mathbf{X}|\mathbf{M}, \mathbf{\Sigma}_r, \mathbf{\Sigma}_r)$ of a matrix normal distribution $\mathcal{MN}_{r,c}(\mathbf{M}, \mathbf{\Sigma}_r, \mathbf{\Sigma}_c)$ is given by [27]

$$p(\mathbf{X}|\mathbf{M}, \mathbf{\Sigma}_r, \mathbf{\Sigma}_r) = \frac{\exp\left(-\frac{1}{2} tr\left[\mathbf{\Sigma}_c^{-1}(\mathbf{X}-\mathbf{M})^T \mathbf{\Sigma}_r^{-1}(\mathbf{X}-\mathbf{M})\right]\right)}{(2\pi)^{rc/2} |\mathbf{\Sigma}_c|^{r/2} |\mathbf{\Sigma}_r|^{c/2}}$$

where $\mathbf{M} \in \mathbb{R}^{r \times c}$ is the mean matrix, $\mathbf{\Sigma}_r \in \mathbb{R}^{r \times r}$ is the row covariance matrix and $\mathbf{\Sigma}_c \in \mathbb{R}^{c \times c}$ is the column covariance matrix.

B. Problem Statement

Let us denote t_j and t_i the acquisition times of two co-registered multi-band optical images. It is not assumed any specific information about time ordering, either $t_i < t_j$ or $t_i > t_j$ are possible cases. Hence, without loss of generality, the HR image acquired at time t_i is assumed to be a low spectral resolution (e.g., PAN or MS) image of high spatial resolution denoted $\mathbf{Y}_{\mathrm{HR}}^{t_i} \in \mathbb{R}^{n_\lambda \times n}$. The image acquired at time t_j is a LR image (i.e., MS or HS) denoted $\mathbf{Y}_{\mathrm{LR}}^{t_j} \in \mathbb{R}^{m_\lambda \times m}$. The problem addressed in this paper consists of detecting significant changes between these two images. This is a challenging task mainly due to the spatial and spectral resolution dissimilarity which prevents any use of simple yet efficient differencing operation [3], [4]. To alleviate this issue, this work proposes to generalize the CD framework introduced in [17]. More precisely, following the widely admitted forward model described in Section II-A and adopting consistent notations, the observed images $\mathbf{Y}_{\mathrm{HR}}^{t_i}$ and $\mathbf{Y}_{\mathrm{LR}}^{t_j}$ can be related to two HR latent images $\mathbf{X}^{t_i}$ and $\mathbf{X}^{t_j}$, respectively, as follows

$$\mathbf{Y}_{\mathrm{HR}}^{t_i} = \mathbf{L}\mathbf{X}^{t_i} + \mathbf{N}_{\mathrm{HR}} \tag{4a}$$

$$\mathbf{Y}_{\mathrm{LR}}^{t_j} = \mathbf{X}^{t_j}\mathbf{BS} + \mathbf{N}_{\mathrm{LR}}. \tag{4b}$$

Note that (4a) and (4b) are a specific double instance of (2). Indeed, the HR (resp., LR) image $\mathbf{Y}_{\mathrm{HR}}^{t_i}$ (resp., $\mathbf{Y}_{\mathrm{LR}}^{t_j}$) is assumed to be only a spectrally (resp., spatially) degraded version of the HR multi-band latent image $\mathbf{X}^{t_i}$ (resp., $\mathbf{X}^{t_i}$) such that both latent images $\mathbf{X}^{t_i} \in \mathbb{R}^{m_\lambda \times n}$ and $\mathbf{X}^{t_j} \in \mathbb{R}^{m_\lambda \times n}$ share the same spectral and spatial resolutions which correspond to the highest resolutions of both observed images. Thereby, provided these two latent images can be efficiently inferred, any classical differencing technique can be subsequently implemented on them to detect changes, notably at a high spatial resolution. More specifically, it would consist of evaluating an HR change image denoted $\Delta\mathbf{X} = [\Delta\mathbf{x}_1, \ldots, \Delta\mathbf{x}_n]$ that would gather information related to any change between the two observed images

$$\Delta\mathbf{X} = \mathbf{X}^{t_i} - \mathbf{X}^{t_j} \tag{5}$$

where $\Delta\mathbf{x}_p \in \mathbb{R}^{m_\lambda}$ denotes the spectral change vector in the pth pixel ($p = 1, \ldots, n$). This spectral change image can be exploited by conducting a pixel-wise change vector analysis (CVA) [28] which exhibits the polar coordinates (i.e., magnitude and direction) of the spectral change vectors. To spatially locate the changes, a natural approach consists of monitoring the information contained in the magnitude part of this representation [9], [10], [29], by considering the corresponding HR spectral change energy image

$$\mathbf{e} = [e_1, \ldots, e_n] \in \mathbb{R}^n \tag{6}$$

with

$$e_p = \|\Delta\mathbf{x}_p\|_2, \quad p = 1, \ldots, n. \tag{7}$$

When the CD problem in the pth pixel is formulated as the binary hypothesis testing

$$\begin{cases} \mathcal{H}_{0,p} & : \text{ no change occurs in the } p\text{th pixel} \\ \mathcal{H}_{1,p} & : \text{ a change occurs in the } p\text{th pixel} \end{cases} \tag{8}$$

the pixel-wise statistical test can be written for a given threshold τ as

$$e_p \underset{\mathcal{H}_{0,p}}{\overset{\mathcal{H}_{1,p}}{\gtrless}} \tau. \tag{9}$$

The final binary HR CD map denoted $\mathbf{d} = [d_1, \ldots, d_n] \in \{0,1\}^n$ can be derived as

$$d_p = \begin{cases} 1 & \text{if } e_p \geq \tau \quad (\mathcal{H}_{1,p}) \\ 0 & \text{otherwise} \quad (\mathcal{H}_{0,p}). \end{cases} \tag{10}$$

When complementary information needs to be extracted from the change image $\Delta\mathbf{X}$, e.g., to identify different types of changes, the whole polar representation (i.e., both magnitude and direction) can be fully exploited [9], [10]. As a consequence, to solve the multi-band image CD problem, the key issue lies in the joint estimation of the pair of HR latent images $\{\mathbf{X}^{t_i}, \mathbf{X}^{t_j}\}$ from the forward model (4) or, equivalently, the joint estimation of one of this latent image and the difference image, e.g., $\{\mathbf{X}^{t_j}, \Delta\mathbf{X}\}$. The next paragraph shows that this problem can be formulated as a particular instance of multi-band image fusion.

C. Robust Multiband Image Fusion

Linear forward models similar to (4) have been extensively investigated in the image processing literature for various applications. When a unique LR image (e.g., MS or HS) $\mathbf{Y}_{\mathrm{LR}}^{t_j}$ has been observed at time t_j, recovering the HR latent image $\mathbf{X}^{t_j}$ from the direct model (4b) can be cast as a superresolution problem [30], [31]. Besides, when a complementary HR image $\mathbf{Y}_{\mathrm{HR}}^{t_i}$ of lower spectral resolution (i.e., PAN or MS) has been simultaneously acquired at time $t_i = t_j$ under (4a), the two corresponding latent images are expected to represent exactly the same scene, i.e., $\Delta\mathbf{X} = \mathbf{0}$ or, equivalently, $\mathbf{X}^{t_i} = \mathbf{X}^{t_j} = \mathbf{X}$ where the time index can be omitted. In such scenario, estimating the common HR latent image $\mathbf{X}$ from the two observed images $\mathbf{Y}_{\mathrm{HR}}$ and $\mathbf{Y}_{\mathrm{LR}}$ is a multi-band image fusion problem addressed in [20]–[23], [25], [32]–[34], also referred to as MS or HS pansharpening when the available HR image $\mathbf{Y}_{\mathrm{HR}}^{t_i}$ is a PAN image [26]. Whether the problem consists in increasing the resolution of a single image or fusing multiple images of different spatial and spectral resolutions, the underlying objective consists in compensating the energy trade-off of optical sensors to get highly spatially and spectrally resolved images. Those problems are often formulated as an inverse problem, which is generally ill-posed or, at least, ill-conditioned. To overcome this issue, a classical approach consists of penalizing the data fitting terms derived from the linear models (4) and the noise statistics (3) with additional regularizing terms exploiting any prior information on the latent image. Various penalizations have been considered in the literature, including Tikhonov regularizations expressed in the image domain [21], [35] or in a transformed (e.g., gradient) domain [36], [37], dictionary- or patch-based regularizations [25], [30], total variation (TV) [23], [38] or regularizations based on sparse wavelet representations [39], [40].

In this work, we propose to follow a similar route by addressing, in a first step, the CD problem as a linear inverse problem derived from (4). However, the CD problem addressed here differs from the computational imaging problems discussed above by the fact that two distinct HR latent images $\mathbf{X}^{t_i}$ and $\mathbf{X}^{t_j}$ need to be inferred, which makes the inverse problem highly ill-posed. However, this particular applicative scenario of CD yields a natural reparametrization where relevant prior knowledge can be conveniently exploited. More precisely, since the two HR latent images are related to the same scene observed at two time instants, they are expected to share a high level of similarity, i.e., the change image $\Delta\mathbf{X}$ is expected to be spatially sparse. Thus, instead of jointly estimating the pair $\{\mathbf{X}^{t_i}, \mathbf{X}^{t_j}\}$ of HR latent images, we take benefit from this crucial information to rewrite the joint observation model (4) as a function of $\{\mathbf{X}^{t_j}, \Delta\mathbf{X}\}$, i.e.,

$$\mathbf{Y}_{\mathrm{HR}}^{t_i} = \mathbf{L}\left(\mathbf{X}^{t_j} + \Delta\mathbf{X}\right) + \mathbf{N}_{\mathrm{HR}} \tag{11a}$$

$$\mathbf{Y}_{\mathrm{LR}}^{t_j} = \mathbf{X}^{t_j}\mathbf{B}\mathbf{S} + \mathbf{N}_{\mathrm{LR}}. \tag{11b}$$

It is worthy to note that this dual observation model parametrized by the new pair $\{\mathbf{X}^{t_j}, \Delta\mathbf{X}\}$ of images to be inferred can be straightforwardly associated with a particular instance of the multi-band image fusion discussed earlier. Indeed, given the HR change image $\Delta\mathbf{X}$ and the HR image observed at time t_i, an HR *corrected* image denoted $\mathbf{Y}_{\mathrm{cHR}}^{t_j}$ that would be acquired by the HR sensor at time t_j can be defined as

$$\mathbf{Y}_{\mathrm{cHR}}^{t_j} = \mathbf{Y}_{\mathrm{HR}}^{t_i} - \mathbf{L}\Delta\mathbf{X}. \tag{12}$$

In such case, the HR forward model (11a) can be easily rewritten, leading to

$$\mathbf{Y}_{\mathrm{cHR}}^{t_j} = \mathbf{L}\mathbf{X}^{t_j} + \mathbf{N}_{\mathrm{HR}} \tag{13a}$$

$$\mathbf{Y}_{\mathrm{LR}}^{t_j} = \mathbf{X}^{t_j}\mathbf{B}\mathbf{S} + \mathbf{N}_{\mathrm{LR}}. \tag{13b}$$

This observation model (13) defines a standard multi-band image fusion problem for the LR observed image $\mathbf{Y}_{\mathrm{LR}}^{t_j}$ and the corrected HR image $\mathbf{Y}_{\mathrm{cHR}}^{t_j}$. Consequently, since the change image $\Delta\mathbf{X}$ can be considered as an outlier term, akin to those encountered in several robust factorizing models such as robust principal component analysis (RPCA) [18] and robust nonnegative factorization [19] which relies on a similar sparse outlier term, the joint observation model (11) naturally defines a so-called *robust fusion* scheme whose objective function is detailed in the next paragraph.

D. Robust Fusion Objective Function

Because of the additive nature and the statistical properties of the noise $\mathbf{N}_{\mathrm{HR}}$ and $\mathbf{N}_{\mathrm{LR}}$, both observed images $\mathbf{Y}_{\mathrm{HR}}^{t_i}$ and $\mathbf{Y}_{\mathrm{LR}}^{t_j}$ can be assumed matrix normally distributed

$$\begin{aligned} \mathbf{Y}_{\mathrm{HR}}^{t_i}|\mathbf{X}^{t_j}, \Delta\mathbf{X} &\sim \mathcal{MN}_{n_\lambda, n}\left(\mathbf{L}\left(\mathbf{X}^{t_j} + \Delta\mathbf{X}\right), \mathbf{\Lambda}_{\mathrm{HR}}, \mathbf{I}_n\right) \\ \mathbf{Y}_{\mathrm{LR}}^{t_j}|\mathbf{X}^{t_j} &\sim \mathcal{MN}_{m_\lambda, m}\left(\mathbf{X}^{t_j}\mathbf{B}\mathbf{S}, \mathbf{\Lambda}_{\mathrm{LR}}, \mathbf{I}_m\right). \end{aligned}$$

Besides, since both observations are acquired by different modality sensors, the noise, which is sensor-dependent, can be assumed statistically independent. Thus, $\mathbf{Y}_{\mathrm{HR}}^{t_i}|\mathbf{X}^{t_j}, \Delta\mathbf{X}$ and $\mathbf{Y}_{\mathrm{LR}}^{t_j}|\mathbf{X}^{t_j}$ are also statistically independent and the joint likelihood function $p(\mathbf{Y}_{\mathrm{HR}}^{t_i}, \mathbf{Y}_{\mathrm{LR}}^{t_j}|\mathbf{X}^{t_j}, \Delta\mathbf{X})$ can be written as a simple product of the conditional distributions $p(\mathbf{Y}_{\mathrm{HR}}^{t_i}|\mathbf{X}^{t_j}, \Delta\mathbf{X})$ and $p(\mathbf{Y}_{\mathrm{LR}}^{t_j}|\mathbf{X}^{t_j})$.

A Bayesian formulation of the robust multi-band image fusion problem allows prior information to be introduced to regularize the underlying estimation problem [41]. Bayesian estimators can be derived from the joint posterior distribution

$$p(\mathbf{X}^{t_j}, \Delta\mathbf{X}|\mathbf{Y}_{\mathrm{HR}}^{t_i}, \mathbf{Y}_{\mathrm{LR}}^{t_j}) \propto p(\mathbf{Y}_{\mathrm{HR}}^{t_i}, \mathbf{Y}_{\mathrm{LR}}^{t_j}|\mathbf{X}^{t_j}, \Delta\mathbf{X}) p(\mathbf{X}^{t_i}) p(\Delta\mathbf{X}) \quad (14)$$

where $p(\mathbf{X}^{t_i})$ and $p(\Delta\mathbf{X})$ correspond to the prior distributions associated with the latent and change images, respectively, assumed to be a priori independent. Under a maximum a posteriori (MAP) paradigm, the joint MAP estimator $\{\hat{\mathbf{X}}_{\mathrm{MAP}}^{t_j}, \boldsymbol{\Delta}\hat{\mathbf{X}}_{\mathrm{MAP}}\}$ can be derived by minimizing the negative log-posterior, leading to the following minimization problem

$$\left\{\hat{\mathbf{X}}_{\mathrm{MAP}}^{t_i}, \Delta\hat{\mathbf{X}}_{\mathrm{MAP}}\right\} \in \underset{\mathbf{X}^{t_j}, \Delta\mathbf{X}}{\mathrm{Argmin}}\, \mathcal{J}\left(\mathbf{X}^{t_j}, \Delta\mathbf{X}\right) \quad (15)$$

with

$$\begin{aligned}\mathcal{J}\left(\mathbf{X}^{t_j}, \Delta\mathbf{X}\right) = &\frac{1}{2}\left\|\boldsymbol{\Lambda}_{\mathrm{HR}}^{-\frac{1}{2}}\left(\mathbf{Y}_{\mathrm{HR}}^{t_i} - \mathbf{L}\left(\mathbf{X}^{t_j} + \Delta\mathbf{X}\right)\right)\right\|_F^2 \\ &+ \frac{1}{2}\left\|\boldsymbol{\Lambda}_{\mathrm{LR}}^{-\frac{1}{2}}\left(\mathbf{Y}_{\mathrm{LR}}^{t_j} - \mathbf{X}^{t_j}\mathbf{B}\mathbf{S}\right)\right\|_F^2 \\ &+ \lambda\phi_1\left(\mathbf{X}^{t_j}\right) + \gamma\phi_2\left(\Delta\mathbf{X}\right). \end{aligned} \quad (16)$$

The regularizing functions $\phi_1(\cdot)$ and $\phi_2(\cdot)$ can be related to the negative log-prior distributions of the latent and change images, respectively, and the parameters λ and γ tune the amount of corresponding penalizations in the overall objective function $\mathcal{J}(\mathbf{X}^{t_j}, \Delta\mathbf{X})$. These functions should be carefully designed to exploit any prior knowledge regarding the parameters of interest. As discussed in Section II-C, numerous regularizations can be advocated for the HR latent image $\mathbf{X}^{t_j}$. In this work, a Tikhonov regularization proposed in [21] has been adopted

$$\phi_1\left(\mathbf{X}^{t_j}\right) = \left\|\mathbf{X}^{t_j} - \bar{\mathbf{X}}^{t_j}\right\|_F^2 \quad (17)$$

where $\bar{\mathbf{X}}^{t_j}$ refers to a crude estimate of $\mathbf{X}^{t_j}$, e.g., resulting from a naive spatial interpolation of the observed LR-HS image $\mathbf{Y}_{\mathrm{LR}}^{t_j}$. This choice has been proven to maintain computational efficiency while providing accurate results [26]. Additionally, a subspace-based representation can also be adopted to enforce $\mathbf{X}^{t_j}$ to live in a previously identified subspace, as advocated in [23] and [42].

Conversely and more critically, a specific attention should be paid to the regularizing function $\phi_2(\cdot)$. This function should reflect the fact that most of the pixels are expected to remain unchanged in $\mathbf{X}^{t_i}$ and $\mathbf{X}^{t_j}$, i.e., most of the columns of the change image $\Delta\mathbf{X}$ are expected to be null vectors. This noticeable property can be easily translated by promoting the sparsity of the spectral change energy image $\mathbf{e}$ defined by (6). As a consequence, the regularizing function $\phi_2(\cdot)$ is chosen as the sparsity-inducing ℓ_1-norm of the change energy image $\mathbf{e}$ or, equivalently, as the $\ell_{2,1}$-norm of the change image

$$\phi_2\left(\Delta\mathbf{X}\right) = \left\|\Delta\mathbf{X}\right\|_{2,1} = \sum_{p=1}^{n}\left\|\Delta\mathbf{x}_p\right\|_2. \quad (18)$$

This regularization is a specific instance of the non-overlapping group-lasso penalization [43] which has been considered in various applications to promote structured sparsity [19], [44]–[49]. The next section describes an iterative algorithm which solves the minimization problem in (15).

III. Minimization Algorithm

Computing the joint MAP estimator of the HR latent image $\mathbf{X}^{t_j}$ at time t_j and of the change image $\Delta\mathbf{X}$ can be achieved by solving the minimization problem in (15). However, no closed-form solution can be derived for this problem. Thus this section presents a minimization algorithm which iteratively converges to this solution. It consists in sequentially solving the problem w.r.t. to each individual variables $\mathbf{X}^{t_j}$ and $\Delta\mathbf{X}$. This block coordinate descent algorithm is summarized in Algorithm 1 whose main steps (fusion and correction) are detailed in what follows.

Algorithm 1: BCD Algorithm for Robust Multi-Band Image Fusion.

Input: $\mathbf{Y}_{\mathrm{LR}}^{t_j}$, $\mathbf{Y}_{\mathrm{HR}}^{t_i}$, $\mathbf{L}$, $\mathbf{B}$, $\mathbf{S}$, $\boldsymbol{\Lambda}_{\mathrm{HR}}$, $\boldsymbol{\Lambda}_{\mathrm{LR}}$.

1: Set $\Delta\mathbf{X}_1$.
2: **for** $k = 1, \ldots, K$ **do**
3: $\quad\mathbf{X}_{k+1}^{t_j} = \arg\min_{\mathbf{X}^{t_j}} \mathcal{J}(\mathbf{X}^{t_j}, \Delta\mathbf{X}_k)$
4: $\quad\Delta\mathbf{X}_{k+1} = \arg\min_{\Delta\mathbf{X}} \mathcal{J}(\mathbf{X}_{k+1}^{t_j}, \Delta\mathbf{X})$
5: **end for**

Output: $\hat{\mathbf{X}}_{\mathrm{MAP}}^{t_j} \triangleq \mathbf{X}_{K+1}^{t_j}$ and $\Delta\hat{\mathbf{X}}_{\mathrm{MAP}} \triangleq \Delta\hat{\mathbf{X}}_{K+1}$

A. Fusion: Optimization w.r.t $\mathbf{X}^{t_j}$

At the kth iteration of the BCD algorithm, let assume that the current value of the change image is denoted $\Delta\mathbf{X}_k$. As suggested in Section II-C, a corrected HR image $\mathbf{Y}_{\mathrm{cHR},k}^{t_j}$ that would be observed at time t_j given the HR image $\mathbf{Y}_{\mathrm{HR}}^{t_i}$ observed at time t_i and the HR change image $\Delta\mathbf{X}_k$ can be introduced as

$$\mathbf{Y}_{\mathrm{cHR},k}^{t_j} = \mathbf{Y}_{\mathrm{HR}}^{t_i} - \mathbf{L}\Delta\mathbf{X}_k. \quad (19)$$

Updating the current value of the HR latent image consists in minimizing w.r.t. $\mathbf{X}^{t_j}$ the partial function

$$\begin{aligned}\mathcal{J}_1\left(\mathbf{X}^{t_j}\right) &\triangleq \mathcal{J}\left(\mathbf{X}^{t_j}, \Delta\mathbf{X}_k\right) \\ &= \left\|\boldsymbol{\Lambda}_{\mathrm{LR}}^{-\frac{1}{2}}\left(\mathbf{Y}_{\mathrm{LR}}^{t_j} - \mathbf{X}^{t_j}\mathbf{B}\mathbf{S}\right)\right\|_F^2 \\ &\quad + \left\|\boldsymbol{\Lambda}_{\mathrm{HR}}^{-\frac{1}{2}}\left(\mathbf{Y}_{\mathrm{cHR},k}^{t_j} - \mathbf{L}\mathbf{X}^{t_j}\right)\right\|_F^2 + \lambda\phi_1\left(\mathbf{X}^{t_j}\right). \end{aligned} \quad (20)$$

As noticed earlier, this sub-problem boils down to the multi-band image fusion which has received considerable attention in the recent image processing and remote sensing literature [20], [21], [23], [25], [26], [42]. The two difficulties arising from this formulation lies in the high dimension of the optimization problem and in the fact that the sub-sampling operator $\mathbf{S}$ prevents any fast resolution in the frequency domain by diagonalization of

the spatial degradation matrix $\mathbf{R} = \mathbf{BS}$. However, with the particular choice (17) of the regularization function $\phi_1(\cdot)$ adopted in this paper, a closed-form solution can still be derived and efficiently implemented. It consists in solving a matrix Sylvester equation [20] of the form

$$\mathbf{C}_1\mathbf{X}^{t_j} + \mathbf{X}^{t_j}\mathbf{C}_2 = \mathbf{C}_3 \tag{21}$$

where the matrices $\mathbf{C}_1$, $\mathbf{C}_2$ and $\mathbf{C}_3$ depend on the quantities involved in the problem, i.e., the virtual and observed images, the degradation operators, the noise covariance matrices and the spatially interpolated image defined in (17) (see [20] for more details). Note that when a more complex regularization function $\phi_1(\cdot)$ is considered (e.g., TV or sparse representation over a dictionary), iterative algorithmic strategies can be adopted to approximate the minimizer of $\mathcal{J}_1\left(\mathbf{X}^{t_j}\right)$.

B. Correction: Optimization w.r.t $\Delta\mathbf{X}$

Following the same strategy as in [17], let introduce the *predicted* HR image

$$\mathbf{Y}_{\mathrm{pHR},k}^{t_j} = \mathbf{L}\mathbf{X}_k^{t_j} \tag{22}$$

that would be observed at time index t_j by the HR sensor given its spectral response $\mathbf{L}$ and the current state of the HR latent image $\mathbf{X}_k^{t_j}$ at the kth iteration of the BCD algorithm. Similarly to (5), the predicted HR change image can thus be defined as

$$\Delta\mathbf{Y}_{\mathrm{pHR},k} = \mathbf{Y}_{\mathrm{HR}}^{t_i} - \mathbf{Y}_{\mathrm{pHR},k}^{t_j}. \tag{23}$$

The objective function (16) w.r.t $\Delta\mathbf{X}$ is then rewritten by combining (22) and (23) with (16), leading to

$$\begin{aligned}\mathcal{J}_2(\Delta\mathbf{X}) &\triangleq \mathcal{J}(\mathbf{X}_k^{t_j}, \Delta\mathbf{X}) \\ &= \left\|\mathbf{\Lambda}_{\mathrm{HR}}^{-\frac{1}{2}}\left(\Delta\mathbf{Y}_{\mathrm{pHR},k} - \mathbf{L}\Delta\mathbf{X}\right)\right\|_F^2 + \gamma\phi_2\left(\Delta\mathbf{X}\right).\end{aligned} \tag{24}$$

With the specific CD-driven choice of $\phi_2(\cdot)$ in (18), minimizing $\mathcal{J}_2(\Delta\mathbf{X})$ is an $\ell_{2,1}$-penalized least square problem. It is characterized by the sum of a convex and differentiable data fitting term with β-Lipschitz continuous gradient $\nabla f(\cdot)$

$$f\left(\Delta\mathbf{X}\right) \triangleq \left\|\mathbf{\Lambda}_{\mathrm{HR}}^{-\frac{1}{2}}\left(\Delta\mathbf{Y}_{\mathrm{pHR},k} - \mathbf{L}\Delta\mathbf{X}\right)\right\|_F^2 \tag{25}$$

and a convex but non-smooth penalization

$$g\left(\Delta\mathbf{X}\right) \triangleq \gamma\phi_2\left(\Delta\mathbf{X}\right) = \gamma\left\|\Delta\mathbf{X}\right\|_{2,1}. \tag{26}$$

Various algorithms have been proposed to solve such convex optimization problems including forward-backward splitting [50], [51], Douglas-Rachford splitting [51], [52] and alternating direction method of multipliers [53], [54]. Since the proximal operator related to $g(\cdot)$ can be efficiently computed, in this work, we propose to resort to an iterative forward-backward algorithm which has shown to provide the fastest yet reliable results. This algorithmic scheme is summarized in Algorithm 2. It relies on a forward step which consists in conducting a gradient descent using the data-fitting function $f(\cdot)$ in (25), and a backward step relying on the proximal mapping associated with the penalizing function $g(\cdot)$ in (26).

Algorithm 2: Correction Step: Forward-Backward Algorithm.

Input: $\Delta\mathbf{X}_k$, $\Delta\mathbf{Y}_{\mathrm{pHR},k}$, $\mathbf{\Lambda}_{\mathrm{HR}}$, $\mathbf{L}$, $\{\eta_j\}_{j=1}^J$

Set $\mathbf{V}_1 \triangleq \Delta\mathbf{X}_k$

2: **for** $j = 1, \ldots, J$ **do**

% forward step

4: $\mathbf{U}_{j+1} = \mathbf{V}_j - \eta_j\nabla f\left(\mathbf{V}_j\right)$

% backward step

6: $\mathbf{V}_{j+1} = \mathrm{prox}_{\eta_j g}\left(\mathbf{U}_{j+1}\right)$

end for

Output: $\Delta\mathbf{X}_{k+1} \triangleq \mathbf{V}_{J+1}$

Since the HR observed image has only a few spectral bands (e.g., $n_\lambda \sim 10$), the spectral degradation matrix $\mathbf{L} \in \mathbb{R}^{n_\lambda \times m_\lambda}$ is a fat (and generally full-row rank) matrix. Thus, the corresponding gradient operator $\nabla f(\cdot)$ defining the forward step (see line 2 of Algorithm 2) can be easily and efficiently computed. Conversely, the proximal operator associated with $g(\cdot)$ in (26) and required during the backward step (see line 2 of Algorithm 2) is defined as

$$\mathrm{prox}_{\eta g}(\mathbf{U}) = \arg\min_{\mathbf{Z}}\left(\gamma\left\|\mathbf{Z}\right\|_{2,1} + \frac{1}{2\eta}\left\|\mathbf{Z} - \mathbf{U}\right\|_F^2\right) \tag{27}$$

for some $\eta > 0$. The function $g(\mathbf{U})$ in (26) can be split as $\sum_{p=1}^n g_p(\mathbf{u}_p)$ with, for each column, $g_p(\cdot) = \gamma\left\|\cdot\right\|_2$. Based on the separability property of proximal operators [54], the operator (27) can be decomposed and computed for each pixel location p ($p = 1, \ldots, n$) as

$$\left[\mathrm{prox}_{\eta g}(\mathbf{U})\right]_p = \mathrm{prox}_{\eta g_p}(\mathbf{u}_p) \tag{28}$$

where the notations $[\cdot]_p$ stands for the pth column. Thus, only the proximal operator associated with the Euclidean distance induced by the ℓ_2-norm is necessary. The Moreau decomposition [54]

$$\mathbf{u}_p = \mathrm{prox}_{\eta g}\left(\mathbf{u}_p\right) + \eta\mathrm{prox}_{\eta^{-1}g_p^*}\left(\eta^{-1}\mathbf{u}_p\right) \tag{29}$$

establishes a relationship between the proximal operators of the function $g_p(\cdot)$ and its conjugate $g_p^*(\cdot)$. When the function $g(\cdot)$ is a general norm, its conjugate corresponds to the indicator function into the ball $\mathbb{B}$ defined by its dual norm [47], [54], leading to

$$\mathrm{prox}_{\eta g}(\mathbf{u}_p) = \mathbf{u}_p - \eta\mathcal{P}_{\mathbb{B}}\left(\frac{\mathbf{u}_p}{\eta}\right) \tag{30}$$

where $\mathcal{P}_{\mathbb{B}}(\cdot)$ denotes the projection. When $g(\cdot)$ is defined by (26), since the ℓ_2-norm is self-dual, this projection is

$$\mathcal{P}_{\mathbb{B}}\left(\mathbf{u}_p\right) = \begin{cases}\dfrac{\gamma\mathbf{u}_p}{\left\|\mathbf{u}_p\right\|_2} & \text{if } \left\|\mathbf{u}_p\right\|_2 > \gamma \\ \mathbf{u}_p & \text{otherwise.}\end{cases} \tag{31}$$

Consequently, replacing (31) in (30), the proximal operator associated with the function $g_p(\cdot)$ in (28) is

$$\mathrm{prox}_{\eta g_p}\left(\mathbf{u}_p\right) = \begin{cases}\left(1 - \dfrac{\eta\gamma}{\left\|\mathbf{u}_p\right\|_2}\right)\mathbf{u}_p & \text{if } \left\|\mathbf{u}_p\right\|_2 > \eta\gamma \\ 0 & \text{otherwise.}\end{cases} \tag{32}$$

To conclude, the correction procedure can be interpreted as a gradient descent step for spectral deblurring of the HR change image from the HR predicted change image (forward step), followed by a soft-thresholding of the resulting HR change image to promote spatial sparsity (backward step).

IV. Experiments

A. Simulation Framework

Real dataset for assessing performance of CD algorithms is rarely available. Indeed, this assessment requires couples of images acquired at two different dates, geometrically and radiometrically pre-corrected, presenting changes and, for the scenario considered in this paper, coming from two different optical sensors. In addition, these pairs should be accompanied by a ground truth (i.e., a binary CD mask locating the actual changes) to allow quantitative figures-of-merit to be computed. To alleviate this issue, inspired by the well-known Wald's evaluation protocol dedicated to pansharpening algorithms [55], a framework has been proposed in [17] to assess the performance of CD algorithms when dealing with optical images of different spatial and spectral resolutions. This framework only requires a single HR-HS reference image $\mathbf{X}^{\text{ref}}$ and generates a pair of latent HR-HS images $\mathbf{X}^{t_i}$ and $\mathbf{X}^{t_j}$ resulting from a unmixing-mixing process. This process allows synthetic yet realistic changes to be incorporated within one of these latent images, w.r.t. a predefined binary reference HR change mask $\mathbf{d}_{\text{HR}} \in \mathbb{R}^n$ locating the pixels affected by these changes and further used to assess the performance of the CD algorithms. This procedure allows various physically-inspired changes to be considered, e.g., by tuning the relative abundance of each endmember or replacing one of them by another. This protocol is briefly described below (see [17] for more details).

1) Reference image: The HR-HS reference image $\mathbf{X}^{\text{ref}}$ used in the experiments reported in this paper is a $610 \times 330 \times 115$ HS image of the Pavia University, Italy, acquired by the reflective optics system imaging spectrometer (ROSIS) sensor. This image has undergone a pre-precessing to smooth the atmospheric effects of vapor water absorption by removing some bands. Thus the final HR-HS reference image is of size $610 \times 330 \times 93$.

2) Generating the Changes: Using the same procedure proposed in [17], the HR-HS reference image $\mathbf{X}^{\text{ref}} \in \mathbb{R}^{m_\lambda \times n}$ has been linearly unmixed to define the reference matrix $\mathbf{M}^{\text{ref}} \in \mathbb{R}^{m_\lambda \times R}$ of R endmember spectral signatures and the corresponding reference abundance matrix $\mathbf{A}^{\text{ref}} \in \mathbb{R}^{R \times n}$ such that $\mathbf{X}^{\text{ref}} \approx \mathbf{M}^{\text{ref}}\mathbf{A}^{\text{ref}}$. The two latent HR-HS images $\mathbf{X}^{t_i}$ and $\mathbf{X}^{t_j}$ are then computed as linear mixture of the endmembers in $\mathbf{M}^{\text{ref}}$ with corresponding abundance matrices $\mathbf{A}^{t_i}$ and $\mathbf{A}^{t_j}$, respectively, derived from the reference abundances $\mathbf{A}^{\text{ref}}$ and the change mask $\mathbf{d}_{\text{HR}}$, i.e.,

$$\mathbf{X}^{t_i} = \mathbf{M}^{\text{ref}}\mathbf{A}^{t_i} \quad \text{and} \quad \mathbf{X}^{t_j} = \mathbf{M}^{\text{ref}}\mathbf{A}^{t_j}$$

$$\mathbf{A}^{t_i} = \vartheta^{t_i}\left(\mathbf{A}^{\text{ref}}, \mathbf{d}_{\text{HR}}\right) \quad \text{and} \quad \mathbf{A}^{t_j} = \vartheta^{t_j}\left(\mathbf{A}^{\text{ref}}, \mathbf{d}_{\text{HR}}\right)$$

where the two change-inducing functions $\vartheta^{t_\cdot} : \mathbb{R}^{R \times n} \times \mathbb{R}^n \to \mathbb{R}^{R \times n}$ are defined to simulate realistic changes in some pixels of the HR-HS latent images. Three sets of 75 predefined change masks have been designed according to three specific change rules introduced in [17]. For each simulated pair $\{\mathbf{X}^{t_i}, \mathbf{X}^{t_j}\}$, one of the two functions $\vartheta^{t_\cdot}(\cdot, \mathbf{d}_{\text{HR}})$ is defined as a "no-change" operator, i.e., $\vartheta^{t_\cdot}(\mathbf{A}^{\text{ref}}, \mathbf{d}_{\text{HR}}) = \mathbf{A}^{\text{ref}}$, which leads to an overall set of 450 simulated pairs $\{\mathbf{X}^{t_i}, \mathbf{X}^{t_j}\}$ of HR-HS latent images.

3) Generating the Observed Images: The observed images are generated under 3 distinct scenarios involving 3 pairs of images of different spatial and spectral resolutions, namely,

1) Scenario 1 considers HR-PAN and LR-HS images,
2) Scenario 2 considers HR-PAN and LR-MS images,
3) Scenario 3 considers HR-MS and LR-HS images.

The HR-PAN or HR-MS observed image $\mathbf{Y}_{\text{HR}}^{t_i}$ is obtained by spectrally degrading the corresponding HR-HS latent image $\mathbf{X}^{t_i}$. In Scenarios 1 and 2, the degradation process consists in averaging the first 43 bands of the HR-HS latent image to produce an HR-PAN image. Conversely, in Scenario 3, the degradation process consists in spectrally filtering the HR-HS latent image $\mathbf{X}^{t_i}$ with a 4-band LANDSAT-like spectral response.

Besides, to generate a spatially degraded LR-HS image $\mathbf{Y}_{\text{LR}}^{t_j}$ in Scenarios 1 and 3, the corresponding latent image $\mathbf{X}^{t_j}$ has been blurred by a 5×5 Gaussian kernel and subsequently equally down-sampled in the vertical and horizontal directions with a down-sampling ratio $d = 5$. In Scenario 2, the LR-MS image has been obtained by spectrally filtering the LR-HS image used in Scenarios 1 and 3 with the 4-band LANDSAT-like spectral response.

To illustrate, Fig. 1 shows one of the 450 simulation configurations used during the experiments to assess the performance of the proposed CD technique. Note that, in this particular example representative of Scenario 3, the induced changes are visible in the HR-MS image $\mathbf{Y}_{\text{HR}}^{t_i}$ since the change mask affects the HR-HS latent image $\mathbf{X}^{t_i}$. However, among the 450 simulation configurations, half of them corresponds to changes incorporated into the HR-HS latent image $\mathbf{X}^{t_i}$ (thus visible in the HR-MS image $\mathbf{Y}_{\text{HR}}^{t_i}$) whereas the remaining half corresponds to changes incorporated into the HR-HS latent image $\mathbf{X}^{t_j}$ (thus visible in the LR-HS image $\mathbf{Y}_{\text{LR}}^{t_j}$).

B. Compared Methods

The proposed robust fusion-based CD technique has been compared to four methods able to deal with optical images of different spatial and spectral resolutions. The first one has been proposed in [17] and also relies on a fusion-based approach. Up to the authors' knowledge, it was the first operational CD technique able to operate with multi-band optical images of different spatial and spectral images. Contrary to the model (4) proposed in this paper, it consists in recovering a common latent image by fusing the two observed images and then predicting an HR (PAN or MS) image $\hat{\mathbf{Y}}_{\text{HR}}^{\text{F},t_i}$ from the underlying forward model. An HR change image $\Delta\mathbf{Y}_{\text{HR}}^{\text{F},t_i}$ has been then computed as in (5) from the pair of HR observed and predicted images $\{\mathbf{Y}_{\text{HR}}^{t_i}, \hat{\mathbf{Y}}_{\text{HR}}^{\text{F},t_i}\}$. Finally, as recommended in [17], a spatially-regularized CVA (sCVA) similar to the decision rule detailed in Section II-B has been conducted on $\Delta\mathbf{Y}_{\text{HR}}^{\text{F},t_i}$ to produce an estimated HR CD mask denoted $\hat{\mathbf{d}}_{\text{F}}$.

The second method aims at producing an HR predicted image by successive spatial superresolution and spectral degradation.

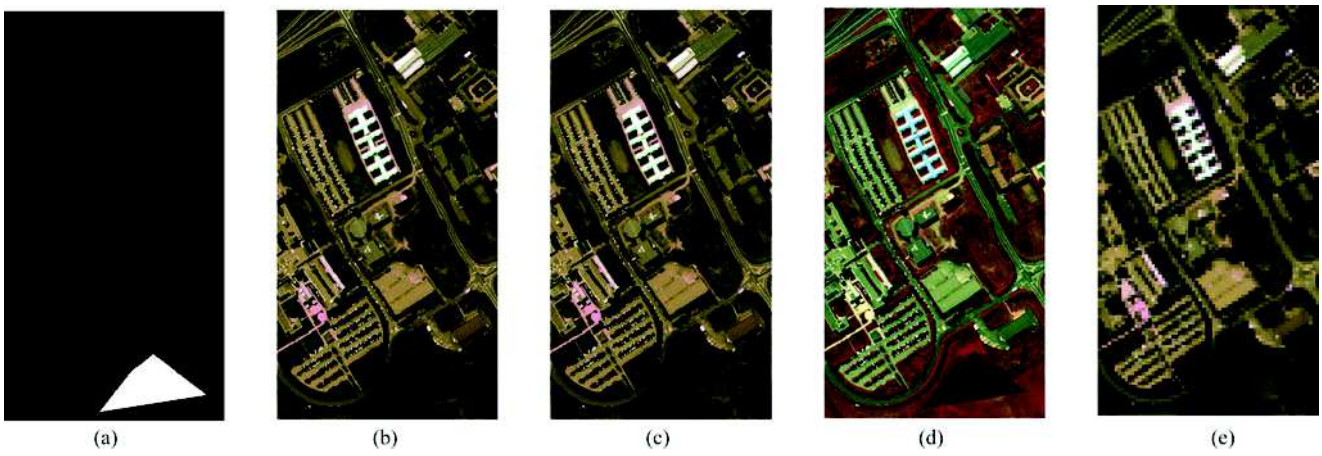


Fig. 1. One particular simulation configuration (Scenario 3). (a) The HR change mask $\mathbf{d}_{\mathrm{HR}}$, (b) and (c) the HR-HS latent images $\mathbf{X}^{t_i}$ and $\mathbf{X}^{t_j}$, and (d) and (e) the spectrally degraded version HR-MS observed image $\mathbf{Y}_{\mathrm{HR}}^{t_i}$ and the spatially degraded LR-HS observed image $\mathbf{Y}_{\mathrm{LR}}^{t_j}$. Note that, in this particular configuration, the changes are visible in the HR-MS observed image $\mathbf{Y}_{\mathrm{HR}}^{t_i}$. Indeed, the change-inducing function $\vartheta^{t_j}(\cdot, \mathbf{d}_{\mathrm{HR}})$ is the identity operator (i.e., $\mathbf{A}^{t_j} = \mathbf{A}^{\mathrm{ref}}$) since it does not apply any change into the corresponding HR-HS latent image $\mathbf{X}^{t_j}$ while the function $\vartheta^{t_i}(\cdot, \mathbf{d}_{\mathrm{HR}})$ includes a triangular region of pixels in $\mathbf{X}^{t_i}$ affected by changes. Moreover, the HR observed image is here an MS image.

More precisely, an HR latent image is first recovered by conducting a band-wise spatial superresolution of the observed LR $\mathbf{Y}_{\mathrm{LR}}^{t_j}$ following the fast method in [31]. Then this latent image is spectrally degraded according to produce an HR predicted image $\hat{\mathbf{Y}}_{\mathrm{HR}}^{\mathrm{SD},t_j}$. Similarly to the previous fusion-based method, sCVA has been finally conducted on the pair $\{\mathbf{Y}_{\mathrm{HR}}^{t_i}, \hat{\mathbf{Y}}_{\mathrm{HR}}^{\mathrm{SD},t_j}\}$ to produce an HR CD mask denoted $\hat{\mathbf{d}}_{\mathrm{SD}}$. The third CD method applies the same procedure with a reverse order of spatial superresolution and spectral degradation, and produces produces an HR change mask denoted $\hat{\mathbf{d}}_{\mathrm{DS}}$ from the pair of HR images $\{\mathbf{Y}_{\mathrm{HR}}^{t_i}, \hat{\mathbf{Y}}_{\mathrm{HR}}^{\mathrm{DS},t_j}\}$. The fourth CD method, referred to as the worst-case (WC) as in [17], build a LR change mask $\hat{\mathbf{d}}_{\mathrm{WC}}$ by crudely conducting a sCVA on a spatially degraded version of the HR image and a spectrally degraded version of the LR image.

C. Figures-of-Merit

The CD performances of these four methods, as well as the performance of the proposed robust fusion-based method whose HR change mask is denoted $\hat{\mathbf{d}}_{\mathrm{RF}}$, have been visually assessed from empirical receiver operating characteristics (ROC), representing the estimated pixel-wise probability of detection (PD) as a function of the probability of false alarm (PFA). Moreover, two quantitative criteria derived from these ROC curves have been computed, namely, i) the area under the curve (AUC), corresponding to the integral of the ROC curve and ii) the distance between the no detection point $(PFA = 1, \mathrm{PD} = 0)$ and the point at the interception of the ROC curve with the diagonal line defined by $\mathrm{PFA} = 1 - \mathrm{PD}$. For both metrics, greater the criterion, better the detection.

D. Results

1) Scenario 1 (HR-PAN and LR-HS): The ROC curves displayed in Fig. 2 with corresponding metrics in Table I (first two rows) correspond to the CD results obtained from a pair of

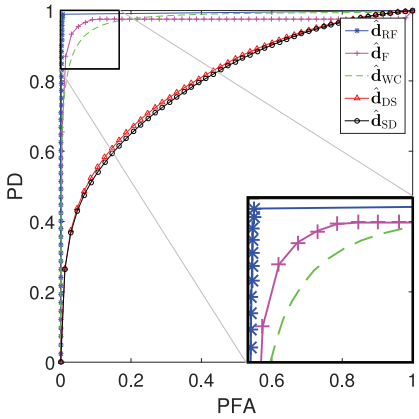


Fig. 2. Scenario 1 (HR-PAN and LR-HS): ROC curves.

TABLE I
SCENARIOS 1 , 2, AND 3: QUANTITATIVE DETECTION PERFORMANCE (AUC AND DISTANCE)

		$\hat{\mathbf{d}}_{\mathrm{RF}}$	$\hat{\mathbf{d}}_{\mathrm{F}}$	$\hat{\mathbf{d}}_{\mathrm{WC}}$	$\hat{\mathbf{d}}_{\mathrm{DS}}$	$\hat{\mathbf{d}}_{\mathrm{SD}}$
Scenario 1	AUC	0.9936	0.9886	0.9762	0.8088	0.8000
	Dist.	0.9896	0.9523	0.9297	0.7270	0.7187
Scenario 2	AUC	0.9948	0.9711	0.9783	0.8137	0.8070
	Dist.	0.9915	0.9534	0.9296	0.7323	0.7247
Scenario 3	AUC	0.9974	0.9895	0.9775	0.8477	0.8521
	Dist.	0.9944	0.9589	0.9268	0.7694	0.7741

HR-PAN and LR-HS observed images. Clearly, the proposed robust fusion-based CD technique outperforms the four other CD techniques. More importantly, it provides almost perfect detections even for very low PFA, i.e., for very low energy changes. Note that the CD mask $\mathbf{d}_{\mathrm{WC}}$ estimated by the worst-case method is defined at a LR.

2) Scenario 2 (HR-PAN and LR-MS): Applying the same procedure as for Scenario 1 but now considering an LR-MS observed image instead of the LR-HS observed image leads to very similar overall performance. The ROC plot is displayed in

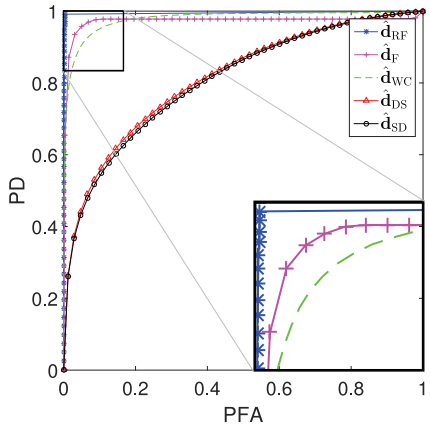


Fig. 3. Scenario 2 (HR-PAN and LR-MS): ROC curves.

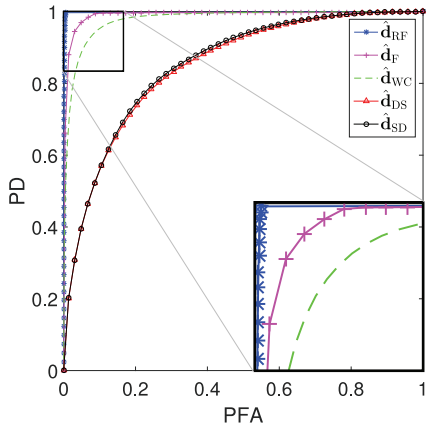


Fig. 4. Scenario 3 (HR-MS and LR-HS): ROC curves.

Fig. 3 with corresponding metrics in Table I (3rd and 4th rows). As in Scenario 1, comparing curves in Fig. 2 shows that the proposed method offers a higher precision even when analyzing a lower spectral resolution HR observed image.

3) Scenario 3 (HR-MS and LR-HS): Following the same strategy, as for Scenario 1 but now considering an HR-MS observed image instead of the HR-PAN observed image, the results lead to very similar overall performance. The ROC plot is displayed in Fig. 4 with corresponding metrics in Table I (last two rows). As in Scenarios 1 and 2, comparing curves in Fig. 4 shows that the proposed method still offers outstanding detection accuracy.

As an additional result, for Scenario 3, Fig. 5 compares the abilities of detecting changes of decreasing size of the proposed method against the fusion-based CD method [17] and the worst-case CD method. Fig. 5(a) and (b) shows a particular example of observed image pair $\mathbf{Y}_{\mathrm{HR}}^{t_i}$ and $\mathbf{Y}_{\mathrm{LR}}^{t_j}$ containing multiple changes with size varying from 1×1-pixel to 61×61-pixels, with the corresponding change mask $\mathbf{d}_{\mathrm{HR}}$ presented in Fig. 5(c). Fig. 5(d) and (e) present the change masks $\hat{\mathbf{d}}_{\mathrm{F}}$ and $\hat{\mathbf{d}}_{\mathrm{WC}}$ recovered by the two competing methods, respectively, while the CD mask $\hat{\mathbf{d}}_{\mathrm{RF}}$ recovered by the proposed robust fusion-based method is reported in Fig. 5(f) shows the proposed CD. For each technique, the decision threshold τ required in the CVA in (10) has been

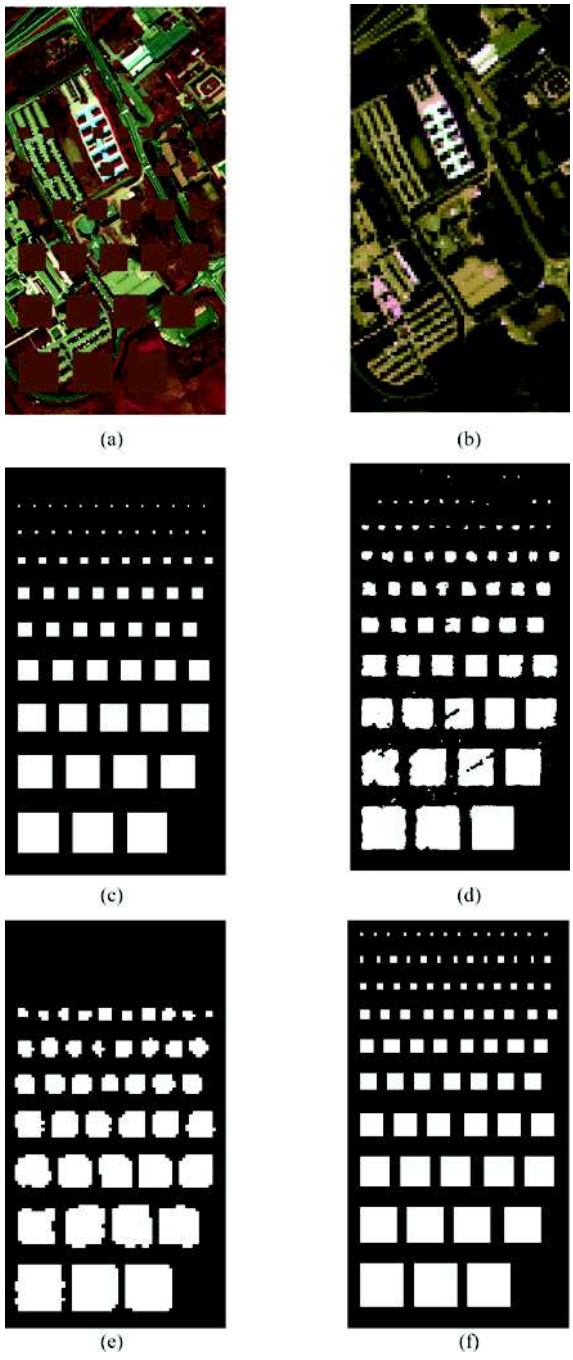


Fig. 5. CD precision for Scenario 2 (HR-MS and LR-HS). (a) HR-MS observed image $\mathbf{Y}_{\mathrm{HR}}^{t_i}$, (b) LR-HS observed image $\mathbf{Y}_{\mathrm{LR}}^{t_j}$, (c) actual change mask $\mathbf{d}_{\mathrm{HR}}$, (d) change mask $\hat{\mathbf{d}}_{\mathrm{F}}$ estimated by the fusion-based approach [17], (e) change mask $\hat{\mathbf{d}}_{\mathrm{WC}}$ estimated by the worst-case approach, and (f) change mask $\hat{\mathbf{d}}_{\mathrm{RF}}$ estimated by the proposed robust fusion-based approach.

tuned to reach the higher distance value in the corresponding ROC curves. The first advantage of the proposed method is a significant decrease of the number of false alarm which are due to propagated errors when implementing the two other methods. Moreover, these results prove once again that the proposed method achieves a better detection rate with a higher resolution, even when considering extremely localized change regions. Remaining false alarms only occur near edges between change and

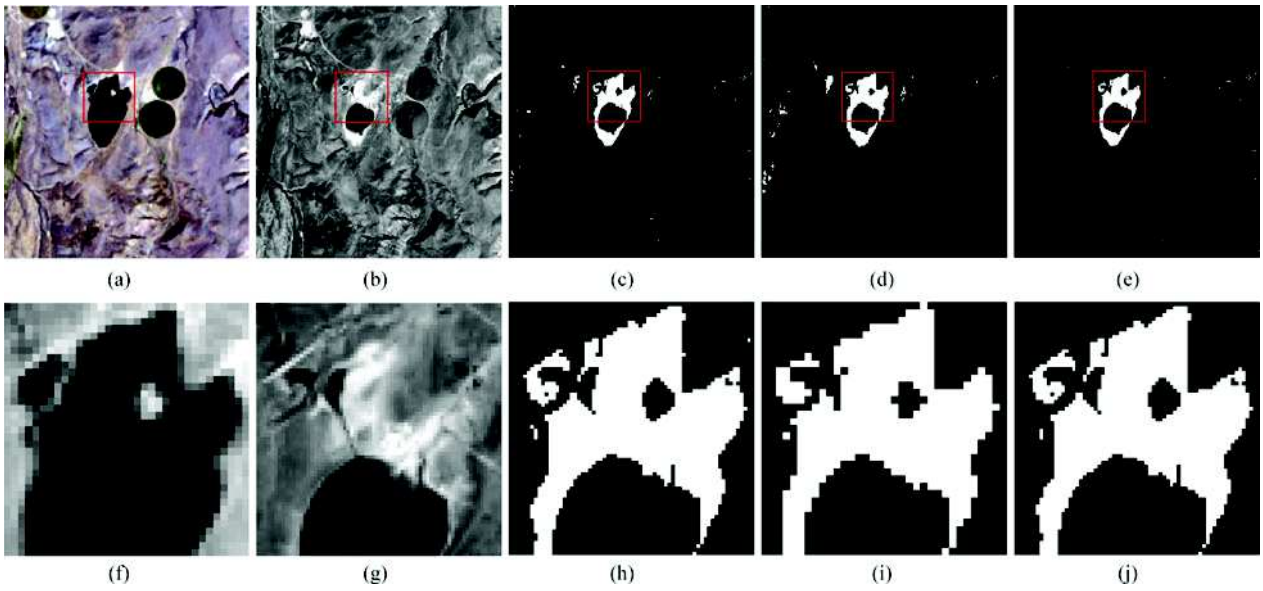


Fig. 6. Real scenario (LR-MS and HR-PAN). (a) LR-MS observed image $\mathbf{Y}_{\mathrm{LR}}^{t_1}$, (b) HR-PAN observed image $\mathbf{Y}_{\mathrm{HR}}^{t_2}$, (c) change mask $\hat{\mathbf{d}}_{\mathrm{F}}$ estimated by the fusion-based approach [17], (d) change mask $\hat{\mathbf{d}}_{\mathrm{WC}}$ estimated by the worst-case approach, and (e) change mask $\hat{\mathbf{d}}_{\mathrm{RF}}$ estimated by the proposed robust fusion-based approach. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

no-change regions of small size due to the difference of spatial resolutions and the width of the blur kernel. Note also that the CD mask estimated by the worst-case method is of coarse scale since based on the comparison of two LR-MS images.

E. Application to Two Real Multidate LANDSAT Images

Finally, to emphasize the reliability of the proposed CD method, a pair of real LR-MS and HR-PAN images, acquired at different dates has been considered. These images $\mathbf{Y}_{\mathrm{LR}}^{t_1}$ and $\mathbf{Y}_{\mathrm{HR}}^{t_2}$ have been acquired by LANDSAT 8 over the Lake Tahoe region (CA, USA), respectively, on April 15th and September 22th, 2015 [56]. The LR-MS image $\mathbf{Y}_{\mathrm{LR}}^{t_1}$ is of size 175×180 characterized by a spatial resolution of 30m. According to the spectral response of the LANDSAT 8 sensor, the HR-PAN image $\mathbf{Y}_{\mathrm{HR}}^{t_2}$ is of size 350×360 with a spatial resolution of 15m and has a spectral range from 0.5 μm to 0.68 μm covering 3 bands of the LR-MS image. Fig. 6(a) and (b) shows the two multidate LR-MS and HR-PAN images that have been manually geographically aligned. The resulting CD binary masks recovered by the fusion-based approach in [17], the worst-case approach and the proposed robust fusion-based CD method are depicted in Fig. 6(c)–(e). For this pair of images, the ground truth information (i.e., in term of a binary map of actual changes) is not available. However, a visual inspection reveals that all methods succeeds in recovering the most significant changes between the pair of images, namely, the pixels corresponding to the lake drought. Nevertheless, the proposed and the fusion approaches have the advantage of providing CD binary masks at HR, which helps to detect finer details than the worst-case method, as illustrated by the zoomed regions in Fig. 6(f)–(j). Moreover, the proposed method seems to lead to a significantly lower of false alarms when compared to the two other methods, which, again, confirms its overall higher accuracy. Note that this low number of false alarms in spite of a manual geographical alignment demonstrates the robustness against misalignment.

V. Conclusion

This paper proposed a robust fusion-based change detection technique to handle two multi-band optical observed images of different spatial and spectral resolutions. The technique was based on the definition of two high resolution hyperspectral latent images related to the observed images via a double physically-inspired forward model. The difference between these two latent images was assumed to be spatially sparse, implicitly locating the changes at a high resolution scale. Inferring these two latent images was formulated as an inverse problem which was solved within a 2-step iterative scheme. This algorithmic strategy amounted to solve a standard fusion problem and an $\ell_{2,1}$-penalized spectral deblurring step. Contrary to the methods already proposed in the literature, modeling errors were not anymore propagate in-between steps. A simulation protocol allowed the performance of the proposed technique in terms of detection and precision to be assessed and compared with the performance of various algorithms. Future works include the assessment of the robustness of the proposed technique w.r.t. nonlinear effects (e.g., due to atmospheric effects, geometric and radiometric distortions) and the problem of detecting changes between optical and non-optical data.

References

[1] V. Ferraris, N. Dobigeon, Q. Wei, and M. Chabert, "Change detection between multi-band images using a robust fusion-based approach," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, New Orleans, LA, USA, 2017.

[2] J. B. Campbell and R. H. Wynne, *Introduction to Remote Sensing*, 5th ed. New York, NY, USA: Guilford Press, 2011.

[3] A. Singh, "Review article digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, Jun. 1989.

[4] F. Bovolo and L. Bruzzone, "The time variable in data fusion: A change detection perspective," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 8–26, Sep. 2015.

[5] M. D. Mura, S. Prasad, F. Pacifici, P. Gamba, J. Chanussot, and J. A. Benediktsson, "Challenges and opportunities of multimodality and data fusion in remote sensing," *Proc. IEEE*, vol. 103, no. 9, pp. 1585–1601, Sep. 2015.

[6] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 17–28, Jan. 2002.

[7] J. C. Price, "Spectral band selection for visible-near infrared remote sensing: Spectral-spatial resolution tradeoffs," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 5, pp. 1277–1285, Sep. 1997.

[8] C. Elachi and J. Van Zyl, *Introduction to the Physics and Techniques of Remote Sensing (Wiley Series in Remote Sensing)*, 2nd ed. Hoboken, NJ, USA: Wiley-Interscience, 2006.

[9] F. Bovolo and L. Bruzzone, "A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 218–236, Jan. 2007.

[10] F. Bovolo, S. Marchesi, and L. Bruzzone, "A framework for automatic and unsupervised detection of multiple changes in multitemporal images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2196–2212, Jun. 2012.

[11] A. A. Nielsen, K. Conradsen, and J. J. Simpson, "Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies," *Remote Sens. Environ.*, vol. 64, no. 1, pp. 1–19, 1998.

[12] A. A. Nielsen, "The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 463–478, Feb. 2007.

[13] J. Inglada, "Similarity measures for multisensor remote sensing images," in *Proc. IEEE Int. Conf. Geosci. Remote Sens.*, 2002, vol. 1, pp. 104–106.

[14] V. Alberga, M. Idrissa, V. Lacroix, and J. Inglada, "Performance estimation of similarity measures of multi-sensor images for change detection applications," in *Proc. IEEE Int. Workshop Anal. Multitemporal Remote Sens. Images*, 2007, pp. 1–5.

[15] G. Mercier, G. Moser, and S. Serpico, "Conditional copula for change detection on heterogeneous SAR data," in *Proc. IEEE Int. Conf. Geosci. Remote Sens.*, 2007, pp. 2394–2397.

[16] J. Prendes, M. Chabert, F. Pascal, A. Giros, and J.-Y. Tourneret, "A new multivariate statistical model for change detection in images acquired by homogeneous and heterogeneous sensors," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 799–812, Mar. 2015.

[17] V. Ferraris, N. Dobigeon, Q. Wei, and M. Chabert, "Detecting changes between optical images of different spatial and spectral resolutions: A fusion-based approach," 2016, submitted for publication.

[18] E. J. Candés, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, 2011, Art. no. 11.

[19] C. Févotte and N. Dobigeon, "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4810–4819, Dec. 2015.

[20] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Fast fusion of multi-band images based on solving a sylvester equation," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4109–4121, Nov. 2015.

[21] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Bayesian fusion of multi-band images," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1117–1127, Sep. 2015.

[22] 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.

[23] M. Simões, J. Bioucas Dias, L. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 6, no. 53, pp. 3373–3388, Jun. 2015.

[24] N. Yokoya, N. Mayumi, and A. Iwasaki, "Cross-calibration for data fusion of EO-1/hyperion and terra/ASTER," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 419–426, Apr. 2013.

[25] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3658–3668, Jul. 2015.

[26] L. Loncan *et al.*, "Hyperspectral pansharpening: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.

[27] A. K. Gupta and D. K. Nagar, *Matrix Variate Distribution (ser. Monographs and Surveys in Pure and Applied Mathematics)*. London, U.K.: Chapman & Hall, 1999, no. 104.

[28] R. D. Johnson and E. S. Kasischke, "Change vector analysis: A technique for the multispectral monitoring of land cover and condition," *Int. J. Remote Sens.*, vol. 19, no. 3, pp. 411–426, Jan. 1998.

[29] A. Singh, "Digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, 1989.

[30] J. Yang, J. Wright, T. S. Huang, and Yi Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[31] N. Zhao, Q. Wei, A. Basarab, N. Dobigeon, D. Kouame, and J.-Y. Tourneret, "Fast single image super-resolution using a new analytical solution for $\ell_2 - \ell_2$ problems," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3683–3697, Aug. 2016.

[32] R. C. Hardie, M. T. Eismann, and G. L. Wilson, "MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1174–1184, Sep. 2004.

[33] M. T. Eismann and R. C. Hardie, "Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions," *IEEE Trans. Image Process.*, vol. 43, no. 3, pp. 455–465, Mar. 2005.

[34] Y. Zhang, S. De Backer, and P. Scheunders, "Noise-resistant wavelet-based Bayesian fusion of multispectral and hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3834–3843, Nov. 2009.

[35] M. Ebrahimi and E. R. Vrscay, "Regularization schemes involving self-similarity in imaging inverse problems," *J. Phys.: Conf. Ser.*, vol. 124, 2008, Art. no. 012021.

[36] Y.-W. Tai, S. Liu, M. S. Brown, and S. Lin, "Super resolution using edge prior and single image detail synthesis," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2010, pp. 2400–2407.

[37] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum, "Gradient profile prior and its applications in image super-resolution and enhancement," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1529–1542, Jun. 2011.

[38] H. A. Aly and E. Dubois, "Image up-sampling using total-variation regularization with a new observation model," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1647–1659, Oct. 2005.

[39] C. V. Jiji, M. V. Joshi, and S. Chaudhuri, "Single-frame image super-resolution using learned wavelet coefficients," *Int. J. Imag. Syst. Technol.*, vol. 14, pp. 105–112, 2004.

[40] J. M. Bioucas-Dias, "Bayesian wavelet-based image deconvolution: A GEM algorithm exploiting a class of heavy-tailed priors," *IEEE Trans. Image Process.*, vol. 15, no. 4, pp. 937–951, Apr. 2006.

[41] J. Idier, *Bayesian Approach to Inverse Problems*, J. Idier, Ed. New York, NY, USA: Wiley-ISTE, 2008.

[42] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Bayesian fusion of multispectral and hyperspectral images using a block coordinate descent method," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Process.: Evol. Remote Sens.*, 2015.

[43] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, *Convex Optimization with Sparsity-Inducing Norms (ser. Optimization for Machine Learning)*. Cambridge, MA, USA: MIT Press, 2011.

[44] S. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.

[45] C. Ding, D. Zhou, X. He, and H. Zha, "R1-PCA: Rotational invariant L1-norm principal component analysis for robust subspace factorization," in *Proc. Int. Conf. Mach. Learn.*, 2006, pp. 281–288.

[46] J. Liu, S. Ji, and J. Ye, "Multi-task feature learning via efficient L2,1-norm minimization," in *Proc. 25th Conf. Uncertainty Artif. Intell.*, 2009, pp. 339–348.

[47] S. Wright, R. Nowak, and M. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, Jul. 2009.

[48] F. Nie, H. Huang, X. Cai, and C. H. Ding, "Efficient and robust feature selection via joint l2, 1-norms minimization," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, 2010, pp. 1813–1821.

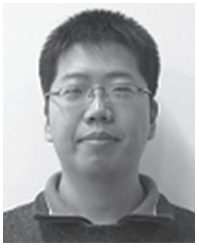
[49] H. Lu, X. Long, and J. Lv, "A fast algorithm for recovery of jointly sparse vectors based on the alternating direction methods," in *Proc. Int. Conf. Artif. Intell. Statist.*, 2011, pp. 461–469.

[50] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1168–1200, 2005.

[51] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," *Fixed-Point Algorithm for Inverse Problems in Science and Engineering (Springer Optimization and Its Applications)*. Berlin, Germany: Springer, 2011.
[52] P. L. Combettes and J.-C. Pesquet, "A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 564–574, Dec. 2007.
[53] S. Boyd, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.
[54] N. Parikh and S. Boyd, "Proximal Algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 123–231, 2013.
[55] 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, 1997.
[56] USGS. Landsat 8, 2017. [Online]. Available: https://landsat.usgs.gov/landsat-8

Vinicius Ferraris (S'16) was born in Rio de Janeiro, Brazil, in 1991. He received the Engineering degree in electronics engineering from ENSEEIHT, Toulouse, France, and the M.Sc. degree in signal processing from National Polytechnic Institute of Toulouse, Toulouse, both in September 2013. He also received the Electrical Engineer (Hons.) degree from the Federal University of Rio Grande do Norte, Natal, Brazil, in February 2014. He is currently working toward the Ph.D. degree with the Signal and Communications Group, IRIT Laborator, National Polytechnic Institute of Toulouse. His recent research activities have been focused on statistical signal and image processing, with a particular interest in Bayesian inverse problems with applications to remote sensing imaging.

Nicolas Dobigeon (S'05–M'08–SM'13) was born in Angoulême, France, in 1981. He received the Engineering degree in electrical engineering from ENSEEIHT, Toulouse, France, and the M.Sc. degree in signal processing from the National Polytechnic Institute of Toulouse (INP Toulouse), Toulouse, both in 2004, and the Ph.D. degree and Habilitation à Diriger des Recherches in signal processing from INP Toulouse in 2007 and 2012, respectively. From 2007 to 2008, he was a Postdoctoral Research Associate with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor. Since 2008, he has been at INP Toulouse, University of Toulouse, Toulouse, where he is currently a Professor and conducts his research within the Signal and Communications Group, IRIT Laboratory, and he is also an Affiliated Faculty Member of the TeSA Laboratory. His recent research activities have been focused on statistical signal and image processing, with a particular interest in Bayesian inverse problems with applications to remote sensing and biomedical imaging.



Qi Wei (S'13–M'15) was born in Shanxi, China, in 1989. He received the bachelor's degree in electrical engineering from Beihang University (BUAA), Beijing, China in July 2010, and the Ph.D. degree in signal and image processing from National Polytechnic Institute of Toulouse (INP-ENSEEIHT), University of Toulouse, Toulouse, France, in September 2015. His doctoral thesis "Bayesian fusion of multi-band images: A powerful tool for super-resolution" was rated as one of the best theses (awarded Prix Léopold Escande) at the University of Toulouse, in 2015. From August 2015 to September of 2016, he worked on multiband image processing as a Research Associate with the Signal Processing Laboratory, Department of Engineering, University of Cambridge, Cambridge, U.K. Since October 2016, he has been working with Prof. L. Carin as a Research Associate at Duke University, Durham, NC, USA. His research interests include machine/deep learning, computer vision/image processing, and Bayesian statistical inference.

Marie Chabert received the Eng. degree in electronics and signal processing from ENSEEIHT, Toulouse, France, in 1994, and the M.Sc. degree in signal processing, the Ph.D. degree in signal processing, and the Habilitation a Diriger les Recherches from the National Polytechnic Institute of Toulouse, Toulouse, in 1994, 1997, and 2007, respectively. She is currently a Professor of signal and image processing at the University of Toulouse, Toulouse. She is with the Engineering School INPT-ENSEEIHT, part of the University of Toulouse. She is conducting her research with the Signal and Communication Team, Institut de Recherche en Informatique de Toulouse (UMR 5505 of the CNRS). Her research interests include nonuniform sampling, time–frequency diagnosis and condition monitoring, and statistical modeling of heterogeneous data in remote sensing.