Hyperspectral super-resolution of locally low rank images from complementary multisource data

Miguel Veganzones, Miguel Simoes, Giorgio Licciardi, Naoto Yokoya, Jose Bioucas-Dias, Jocelyn Chanussot

HAL Id: hal-01117253

https://hal.archives-ouvertes.fr/hal-01117253v3

Submitted on 12 Nov 2015

Hyperspectral super-resolution of locally low rank images from complementary multisource data

M.A. Veganzones, *Member, IEEE*, M. Simões, G. Licciardi, *Member, IEEE*, N. Yokoya, *Member, IEEE*, J.M. Bioucas-Dias, *Senior Member, IEEE*, and J. Chanussot, *Fellow, IEEE*

Abstract—Remote sensing hyperspectral images (HSI) are quite often low rank, in the sense that the data belong to a low dimensional subspace/manifold. This has been recently exploited for the fusion of low spatial resolution HSI with high spatial resolution multispectral images (MSI) in order to obtain super-resolution HSI. Most approaches adopt an unmixing or a matrix factorization perspective. The derived methods have led to state-of-the-art results when the spectral information lies in a low dimensional subspace/manifold. However, if the subspace/manifold dimensionality spanned by the complete data set is large, *i.e.*, larger than the number of multispectral bands, the performance of these methods decrease mainly because the underlying sparse regression problem is severely ill-posed. In this paper, we propose a local approach to cope with this difficulty. Fundamentally, we exploit the fact that real world HSI are locally low rank, that is, pixels acquired from a given spatial neighborhood span a very low dimensional subspace/manifold, *i.e.*, lower or equal than the number of multispectral bands. Thus, we propose to partition the image into patches and solve the data fusion problem independently for each patch. This way, in each patch the subspace/manifold dimensionality is low enough such that the problem is not ill-posed anymore. We propose two alternative approaches to define the hyperspectral super-resolution via local dictionary learning using endmember induction algorithms (HSR-LDL-EIA). We also explore two alternatives to define the local regions, using sliding windows and binary partition trees. The effectiveness of the proposed approaches is illustrated with synthetic and semi real data.

Index Terms—Hyperspectral imagery, multispectral imagery, super-resolution, data fusion, dictionary learning, spectral unmixing, binary partition tree.

I. INTRODUCTION

IN recent years, there has been a huge improvement on the spectral and spatial resolutions in the design of remote sensing sensors. However, it is not possible to acquire images with relatively high spectral and high spatial resolution

M.A. Veganzones and G. Licciardi are with the Department Image and Signal (DIS), GIPSA-lab, Grenoble-INP, F-38402 Saint Martin d'Heres Cedex, France. (e-mail: miguel-angel.veganzones@gipsa-lab.fr; giorgio.licciardi@gipsa-lab.fr). M. Simões is with the Department Image and Signal (DIS) GIPSA-lab, Grenoble-INP, F-38402 Saint Martin d'Heres Cedex, France; and, with the Instituto de Telecomunicações and the Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal. (e-mail: miguel.simoes@gipsa-lab.fr). N. Yokoya is with the Department of Advanced Interdisciplinary Studies (AIS), University of Tokyo, Tokyo, Japan. (e-mail: yokoya@sal.rcast.u-tokyo.ac.jp). J.M. Bioucas-Dias is with the Instituto de Telecomunicações and the Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal. (e-mail: bioucas@lx.it.pt). J. Chanussot is with the Department Image and Signal (DIS), GIPSA-lab, Grenoble-INP, F-38402 Saint Martin d'Heres Cedex, France; and, with the Faculty of Electrical and Computer Engineering, University of Iceland, Reykjavik, Iceland. (e-mail: jocelyn.chanussot@gipsa-lab.fr).

simultaneously. This is due, on one hand, to the system trade-off related to data volume and signal-to-noise ratio (SNR) limitations and, on the other hand, to the specific requirements of different applications [1]. Consequently, there is a need for super-resolution techniques that fuse high spectral resolution images, such as hyperspectral images (HSI), with high spatial resolution images, such as multispectral images (MSI) or panchromatic images, in order to obtain high spectral and spatial (super) resolution images.

Recently, some techniques dedicated to the fusion of HSIs and MSIs have been proposed. A general trend is to associate this problem with either linear spectral unmixing [2], which assumes that the underlying data can be described by a mixture of a relatively small number of "pure" spectral signatures corresponding to the materials present in the scene, inferred by any of the multiple endmembers induction algorithms (EIA) available on the literature [3], or with the learning of a spectral dictionary that codifies the information present on the images. Since both HSIs and MSIs capture the same scene, the underlying materials (the so-called *endmembers*) or the dictionaries should be the same. Therefore, the spectral information extracted from one of the images should also be able to explain the other one. Due to the high spectral resolution of the HSIs, the endmembers or the dictionary are extracted from these data, and are then used to reconstruct the MSI. Since MSIs have high spatial resolution, the code of the dictionary or the spatial fractional abundances for an endmembers approach provide the high spatial resolution information to reconstruct the super-resolution HSI.

A. Related work

Zurita *et al.* [4], introduced one of the first unmixing-based approaches to the fusion of remote sensing multiband images. A related approach is proposed in [5], where a very high-resolution hyperspectral image is estimated from a lower-resolution hyperspectral image and a high-resolution RGB image. The method starts by identifying an unmixing matrix used to represent the hyperspectral spectra and then uses this matrix in conjunction with the RGB input to compute, via sparse regression, representation coefficients for the high-resolution hyperspectral image. This methodology can be viewed as a factorization of the input into a mixing matrix and a set of maximally sparse coefficients. An approach with similar flavour is proposed in [6]. The main difference is that the mixing matrix is replaced by a dictionary learnt using a non-negative matrix factorization with sparsity regularization on the code. In [7], the hyperspectral data is unmixed via the

K-SVD algorithm, and the multispsectral data is reconstructed using orthogonal matching pursuit to induce sparsity. The authors in [1] proposed a method where two dictionaries were learnt from the two different datasets and, then, used a dictionary-pair learning method to establish the correspondence between them. A similar and older technique introduced in [8] alternately unmixes both sources of data to find the signatures and the abundances of the endmembers.

The main limitation of these dictionary approaches comes from the difficulties of estimating the code from the MSI. Since the number of multispectral bands is usually lower than the number of entries in the dictionary or the number of induced endmembers, solving the undetermined system of equations is an ill-posed problem. This issue opens the door to all sort of sparse regression techniques, many of them recently introduced in compressive sensing applications [9]. Nevertheless, there is still a difficulty: in hyperspectral applications, the columns of the dictionary or the spectral signatures of the endmembers tend to be highly correlated, implying that their *mutual coherence* is close to 1. This makes the undetermined system of equations ill-posed even when the codes/fractional abundances are sparse [3], [10].

Moreover, the spectral responses of both the hyperspectral and the multispectral sensors are often assumed to be known *a priori* (see [8], [11]–[13], for example). However, estimating this response directly from the data can be advantageous due to a number of reasons. The information made available by the manufacturers can be incomplete or hard to precisely adapt to the problem under analysis. Additionally, atmospheric conditions, post-processing artefacts, and even the variability within the observed scene can cause a mismatch between the real spectral response and the data supplied by the manufacturers [14]. Some works have addressed this question directly. For example, Yokoya *et al.* [15] have estimated the relative spectral response between the Hyperion (HSIs) and ASTER (MSIs) sensors, which are aboard two different satellites. Recently, the authors of [16] introduced a method to estimate the spectral responses by formulating a convex problem.

B. Contribution

In real-world HSI, it is very likely that in a small spatial neighbourhood the number of different materials is small, *i.e.*, these images are locally low rank. We take advantage of this property to propose two HSI super-resolution methodologies by local dictionary learning using endmember induction algorithms (HSR-LDL-EIA). The first one, termed HSR-LDL-EIA(I), is a straightforward adaptation of the state-of-the-art HSR by global dictionary learning using an endmember induction algorithm approach (HSR-GDL-EIA) [5], where the dictionaries are learnt by estimating a set of endmembers from patches of the observed hyperspectral image. The second proposed local method, termed HSR-LDL-EIA(II), extracts the dictionaries from patches of a spatially upscaled and spectrally downscaled version of the observed hyperspectral image.

In order to define the set of patches of the image such that the set of pixels with indices in each patch span a subspace of lower dimensionality than that of the whole image, we propose two approaches: one using a square sliding window

of fixed size and a second one using a binary partition tree (BPT) representation. Sliding windows have been broadly employed in the local processing of images, while the BPT representation [17], [18] has recently proved itself to be meaningful for the partitioning of hyperspectral images using spectral unmixing information [19].

We show, through experiments using synthetic and real datasets, that the proposed local approaches outperform the baseline global approach when the HSI and MSI are locally low rank. We also explore the impact of using an estimation of the spectral response instead of using the sensor specifications, to validate the proposed approaches when the spectral response is unknown or unreliable.

C. Outline

The remainder of the paper is organized as follows: in Sec. II, the super-resolution problem is formulated and we overview the state-of-the-art HSR-GDL-EIA approach. In Sec. III, the proposed HSR-LDL-EIA methodologies are introduced. In Sec. IV, the patches definition by sliding windows and a BPT representation are explained. The use of spectral response estimators is given in Sec. V. Finally, the experimental methodology and the results are provided in Sec. VI and VII, respectively. Sec. VIII gives some conclusion remarks.

II. PROBLEM FORMULATION

Let $\mathbf{X} \in \mathbb{R}^{n_h \times n}$ denote a HSI with n_h spectral bands (rows of $\mathbf{X}$) and $n = n_x \times n_y$ pixels (columns of $\mathbf{X}$). We may interpret $\mathbf{X}$ either as a collection of n_h 2D images (or bands) of size $n_x \times n_y$, each one associated to a given wavelength interval, or as a collection of n spectral vectors of size n_h, each one associated with a given pixel. In this work, we are concerned with the estimation of $\mathbf{X}$, which we term the original HSI, from two degraded observations of $\mathbf{X}$: a) a low spatial resolution HSI, $\mathbf{Y}_h \in \mathbb{R}^{n_h \times (n/d^2)}$, where $d > 1$ denotes a spatial downsampling factor, and b) a MSI, $\mathbf{Y}_m \in \mathbb{R}^{n_m \times n}$, where $n_m \ll n_h$. We assume that $\mathbf{Y}_h$ is generated as

$$\mathbf{Y}_h = \mathbf{XBM} + \mathbf{N}_h, \qquad (1)$$

where $\mathbf{B} \in \mathbb{R}^{n \times n}$ is a matrix modelling band independent sensor blur, $\mathbf{M} \in \mathbb{R}^{n \times (n/d^2)}$ is a masking matrix accounting for spatial downsampling of size d on both spatial dimensions, and $\mathbf{N}_h$ is an additive perturbation. Concerning the MSI, $\mathbf{Y}_m$, we assume the generation model

$$\mathbf{Y}_m = \mathbf{RX} + \mathbf{N}_m, \qquad (2)$$

where the matrix $\mathbf{R} \in \mathbb{R}^{n_m \times n_h}$ holds in its columns the n_h spectral responses of the multispectral sensor, and $\mathbf{N}_m$ is an additive perturbation.

Let us suppose that it is possible to learn a dictionary $\mathbf{D} \in \mathbb{R}^{n_h \times n_d}$ from the observed hyperspectral image $\mathbf{Y}_h$, and that the columns $\mathbf{X}$, denoted by $\mathbf{x}_i$, for $i \in \mathcal{S} = \{1, \ldots, n\}$, may be sparsely represented as linear combinations of the columns of $\mathbf{D}$. That is, given $\mathbf{x}_i$, $i \in \mathcal{S}$, there is a sparse vector $\boldsymbol{\alpha}_i \in \mathbb{R}^{n_d}$ (*i.e.*, only a few components of $\boldsymbol{\alpha}_i$ are non-zero) such that:

$$\mathbf{x}_i = \mathbf{D}\boldsymbol{\alpha}_i. \qquad (3)$$

By introducing (3) into (2), we obtain:

$$\mathbf{Y}_m = \mathbf{RDA} + \mathbf{N}_m, \tag{4}$$

where $\mathbf{A} \equiv [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n]$, is often termed as *code* in dictionary learning and sparse regression applications. If equation (4) can be solved with respect to $\mathbf{A}$, then we may plug its solution into (3) and thereby obtain an estimate of $\mathbf{X}$. This methodology does not need to estimate the spatial blur, as far as it is constant across bands. This reasoning led to the conventional dictionary-based methodology for hyperspectral super-resolution [5]. Given the observed hyperspectral image, $\mathbf{Y}_h$, the first step is to learn the dictionary, $\mathbf{D}$, from it. Then, the dictionary is projected onto the multispectral domain by a spectral response, $\mathbf{R}$. The projected dictionary, $\mathbf{RD}$, is used to estimate the code $\hat{\mathbf{A}}$ from the observed multispectral image, $\mathbf{Y}_m$. Finally, the dictionary and the code are combined to reconstruct the hyperspectral super-resolution image, $\hat{\mathbf{X}} = \mathbf{D}\hat{\mathbf{A}}$.

The success of a dictionary-based approach depends fundamentally on the ability to solve (4) with respect to $\mathbf{A}$. The difficulty in solving this system comes from the fact that the system matrix $\mathbf{RD} \in \mathbb{R}^{n_m \times n_d}$ is often fat, *i.e.*, $n_m < n_d$, yielding an undetermined system of equations. A typical multispectral sensor has less than 10 bands, quite often 4 in the wavelength interval $[0.4, 2.5]$ microns where most hyperspectral sensors operate. The number of entries in a typical dictionary, n_d, is often of the order of a few tens. Eq. (2) is ill-posed when $n_m < n_d$, but it may be solved by exploiting the sparsity of the codes $\boldsymbol{\alpha}_i$, for $i \in \mathcal{S}$. Nevertheless, in hyperspectral applications, the columns of $\mathbf{D}$ tend to be highly correlated, implying that the *mutual coherence* between the columns of $\mathbf{RD}$ is close to 1. This makes (2) ill-posed even when codes $\boldsymbol{\alpha}_i$, for $i \in \mathcal{S}$, are sparse [3], [10].

A. HSI Super Resolution via Global Dictionary Learning using Endmember Induction Algorithms (HSR-GDL-EIA)

A particular implementation of the general dictionary-based methodology is the one in which the dictionaries are defined as the set of endmembers induced from the image by some EIA, hereafter referred to as HSI super resolution via global dictionary learning using EIAs (HSR-GDL-EIA). Fig. 1 shows the pseudo-code of the HSR-GDL-EIA algorithm. This algorithm takes as inputs an observed hyperspectral image, $\mathbf{Y}_h$, an observed multispectral image, $\mathbf{Y}_m$, the spectral response, $\mathbf{R}$, and the number of entries in the dictionary, n_d. The dictionary, $\mathbf{D}$, is defined as a set of n_d endmembers induced from $\mathbf{Y}_h$ by means of an EIA. Then, the code, $\mathbf{A}$, is estimated by the spectral unmixing of the multispectral image, $\mathbf{Y}_m$:

$$\hat{\mathbf{A}} := \arg \min_{\mathbf{A} \geq 0} \|\mathbf{Y}_m - \mathbf{RDA}\|_F^2, \tag{5}$$

where the inequality $\mathbf{A} \geq 0$ is to be understood in the component-wise sense, and $\|\cdot\|_F$ denotes the Frobenius norm. The constraint $\mathbf{A} \geq 0$ in (5) is used because, in the linear mixing model, the codes $\mathbf{A}$ represent abundances of materials which are necessarily non-negative [20]. It is possible to add the abundances sum-to-one constraint, $\mathbf{A}^T \mathbf{1}_{n_d}^T = \mathbf{1}_n$, but this constraint is usually dropped due to possible scale model mismatches [3]. We remark that, since our observations

---

Algorithm *HSR-GDL-EIA*
Input:
1. $\quad \mathbf{Y}_h, \mathbf{Y}_m, \mathbf{R}, n_d$
2. $\quad \mathbf{D} := \text{EIA}(\mathbf{Y}_h, n_d)$
3. $\quad \hat{\mathbf{A}} := \arg \min_{\mathbf{A} \geq 0} \|\mathbf{Y}_m - \mathbf{RDA}\|_F^2$
4. $\quad \hat{\mathbf{X}} := \mathbf{D}\hat{\mathbf{A}}$

Fig. 1. Pseudo-code of the HSI Super Resolution algorithm via Global Dictionary Learning using Endmember Induction Algorithms (HSR-GDL-EIA).

are spectral vectors and, thus, non-negative (apart from the noise contribution), the non-negativity constraint in (5) is equivalent to a form of constrained ℓ_1 regularization and, therefore, to some kind of sparsity enforcement (see [21] for details). Finally, the super-resolution HSI is estimated by the linear combination of the induced set of endmembers and the estimated fractional abundances.

III. HSI SUPER-RESOLUTION VIA LOCAL DICTIONARY LEARNING USING ENDMEMBER INDUCTION ALGORITHMS (HSR-LDL-EIA)

Here, we introduce two HSI super-resolution via local dictionary learning using EIAs (HSR-LDL-EIA) techniques conceived to cope with the ill-posedness with origin in the matrix system $\mathbf{RD}$. The main idea, in the vein of the local approaches to image restoration, is to decompose the HSI and MSI into patches and build patch-dependent dictionaries such that $n_m \geq n_d$ in each patch, implying that the image is locally low rank. Thereby, solving the inverse problem (4) in each patch is well-posed. The way to actually define the patches will be presented in Section IV.

A. HSR-LDL-EIA (I)

The first proposed local technique, termed as HSR-LDL-EIA (I), is a straightforward adaptation of the global approach introduced in Section II-A to work with patches defined from the HSI. Let us define a set of patches, $\mathcal{P}_j$, $j \in \{1, \ldots, P\}$, obtained from the observed HSI, $\mathbf{Y}_h$. The set of pixels with indices in each patch, $\mathbf{Y}_{h,\mathcal{P}_j}$, hereby denoted as $\mathbf{Y}_{h,j}$ with a little abuse of notation, spans a subspace of lower dimensionality than that of $\mathbf{Y}_h$. For each patch, $\mathcal{P}_j$, we identify a low rank dictionary, $\mathbf{D}_j$, from the set of hyperspectral pixels, $\mathbf{Y}_{h,j}$, indicated by $\mathcal{P}_j$. Then, the low rank dictionary identified for the patch is used to estimate the patch code, $\hat{\mathbf{A}}_j$, from the set of multispectral pixels, $\mathbf{Y}_{m,j}$, indicated by $\mathcal{P}_j$. The dictionary and the code are linearly combined to obtain the HSI super-resolution patch, $\hat{\mathbf{X}}_j$. Hence, the conventional HSR-GDL-EIA approach is independently applied to each patch. Finally, the HSI super-resolution patches are combined to build the HSI super-resolution image, $\hat{\mathbf{X}}$.

Fig. 2 presents the pseudo-code of the proposed HSR-LDL-EIA (I) approach. This is analogous to the HSI-GDL-EIA approach presented in Section II-A. Here, however, the dictionaries obtained by an EIA are extracted from the hyperspectral patches, $\mathbf{Y}_{h,j}$, instead of from the whole image, $\mathbf{Y}_h$. Having identified the patches dictionaries, $\mathbf{D}_j$, the code $\mathbf{A}$ is

Algorithm *HSR-LDL-EIA (I)*
Input:
1. $\mathbf{Y}_h$, $\mathbf{Y}_m$, $\mathbf{R}$, $\mathcal{P}_j$, $n_d \leq n_m$
2. **for** $j = 1$ **to** P
3. $\qquad \mathbf{Y}_{h,j} \leftarrow I_h(\mathcal{P}_j)$
4. $\qquad \mathbf{Y}_{m,j} \leftarrow I_m(\mathcal{P}_j)$
5. $\qquad \mathbf{D}_j := \text{EIA}(\mathbf{Y}_{h,j}, n_d)$
6. $\qquad \widehat{\mathbf{A}}_j := \arg\min_{\mathbf{A}_j \geq 0} \|\mathbf{Y}_{m,j} - \mathbf{R}\mathbf{D}_j\mathbf{A}_j\|_F^2$
7. $\qquad \widehat{\mathbf{X}}_j := \mathbf{D}_j \widehat{\mathbf{A}}_j$
8. $\widehat{\mathbf{X}} := \left\{ \widehat{\mathbf{X}}_j \right\}, \; j = 1, \ldots, P$

Fig. 2. Pseudo-code of the first proposed HSI Super Resolution algorithm via Local Dictionary Learning using Endmember Induction Algorithms (HSR-LDL-EIA (I)).

estimated by solving the following constrained least squares (CLS) optimization problem:

$$\min_{\mathbf{A}_j \geq 0} \|\mathbf{Y}_{m,j} - \mathbf{R}\mathbf{D}_j\mathbf{A}_j\|_F^2 \quad j = 1, \ldots, P, \qquad (6)$$

where $\mathbf{A}_j \in \mathbb{R}^{n_d \times P}$ and $\mathbf{Y}_{m,j} \in \mathbb{R}^{n_m \times P}$ gather the columns of $\mathbf{A}$ and $\mathbf{Y}_m$ corresponding to the multispectral pixels indicated by $\mathcal{P}_j$, respectively. As in the global approach, the code is defined as fractional abundances so the positivity constraint, $\mathbf{A} \geq 0$, is enforced. The number of estimated endmembers in the dictionary should be lower or equal than the dimensionality of the multispectral data, $n_d \leq n_m$, so that the optimization problem defined in (6) is well posed. This could be achieved in practice, for instance, by fixing the estimated number of endmembers on each patch to be equal to the number of available multispectral bands, $n_d = n_m$.

In order to apply the proposed HSR-LDL-EIA (I) approach the indicator functions, $\mathbf{Y}_{h,j} \leftarrow I_h(\mathcal{P}_j)$ and $\mathbf{Y}_{m,j} \leftarrow I_m(\mathcal{P}_j)$ that map the patches to the observed HSI and MSI, respectively, should be defined. Let $\mathcal{I}_h = [1, (n_x/d)] \times [1, (n_y/d)] \subset \mathbb{N}^2$ denote the spatial support of $\mathbf{Y}_h$. Hereafter, the row components of a patch index will be denoted by $i_x \in [1, (n_x/d)]$. The column components will be denoted by $i_y \in [1, (n_y/d)]$. Thus, the index can be represented as the pair, $i = (i_x, i_y) \in \mathcal{I}_h$. Given $\mathbf{Y}_h$ with size $n_h \times (n/d^2)$, with $n = n_x n_y$, a patch will be defined as a set of indexes lying inside the spatial support of $\mathbf{Y}_h$:

$$\mathcal{P}_j = \{i_l \subseteq \mathcal{I}_h\}_{l=1}^L, \qquad (7)$$

where $L \in \mathbb{N}$ denotes the number of indexes in the j-th patch. The definition of the function, $I_h(\mathcal{P}_j)$, that selects a set of pixels from the observed hyperspectral image, $\mathbf{Y}_h$, using the indices in the patch, $\mathcal{P}_j$, is straightforward:

$$\mathbf{Y}_{h,j} \leftarrow I_h(\mathcal{P}_j) = \{\mathbf{y}_{h,i} \in \mathbf{Y}_h \mid i \in \mathcal{P}_j\}. \qquad (8)$$

However, in order to define the multispectral mapping function, $I_m(\mathcal{P}_j)$, we need to take into account the spatial downsampling factor, d. The spatial support of the observed multispectral image, $\mathbf{Y}_m$, is given by $\mathcal{I}_m = [1, n_x] \times [1, n_y] \subset \mathbb{N}^2$. Then, for each hyperspectral index, $i \in \mathcal{I}_h$, there is a set, $\mathcal{N}(i) \subset \mathcal{I}_m$, associated to it. These sets are obtained by:

$$\mathcal{N}(i) = [(i_x - 1)d + 1, i_x d] \times [(i_y - 1)d + 1, i_y d], \qquad (9)$$

and hold the non-overlapping property:

$$\mathcal{N}(i_\alpha) \bigcap \mathcal{N}(i_\beta) = \emptyset, \; \forall i_\alpha, i_\beta; \; \alpha \neq \beta. \qquad (10)$$

Therefore, the multispectral mapping function is defined as:

$$\mathbf{Y}_{m,j} \leftarrow I_m(\mathcal{P}_j) = \{\mathbf{y}_{m,k} \in \mathbf{Y}_m \mid k \in \mathcal{N}(i), \; i \in \mathcal{P}_j\}. \qquad (11)$$

B. HSR-LDL-EIA (II)

Besides the image rank, the quality of the super-resolution image can be strongly influenced by the choice of the dictionary. In particular, the elements of the dictionary $\mathbf{D}$ should be consistent with the multispectral image $\mathbf{Y}_m$. This means that, since the elements of $\mathbf{D}$ are derived from $\mathbf{Y}_h$, to match the spectral resolution of $\mathbf{Y}_m$ they should be spectrally downsampled by the sensor's spectral response, $\mathbf{R}$. Thus, the spectral downsampling of the dictionary is a critical point in terms of quality of the super-resolution image. From a practical point of view, the elements of the spectrally downsampled dictionary, $\mathbf{R}\mathbf{D}$, should be as similar as possible to the endmembers that can be extracted from the multispectral image. However, a non-perfect model of the spectral responses of the sensors may lead to elements of $\mathbf{D}$ that could not match the endmembers extracted from $\mathbf{Y}_m$. Moreover, due to the spectral differences between $\mathbf{Y}_h$ and $\mathbf{Y}_m$, the endmembers extracted from the hyperspectral image $\mathbf{Y}_h$ may not correspond to those extracted from the multispectral image with the same approach.

In order to overcome these problems, we propose a novel methodology where the dictionary entries are obtained from a spatially upscaled and spectrally downscaled version of the observed HSI. The pseudo-code of the proposed HSR-LDL-EIA (II) approach is presented in Fig. 3. First, we obtain a spatially upscale version of the observed HSI:

$$\mathbf{Y}_{f(h)} = f(\mathbf{Y}_h), \qquad (12)$$

where $f(\cdot)$ is an upscaling function, for example a bi-cubic interpolation. Then, we spectrally downscale the resulting image to obtain a spatially upscaled and spectrally downscaled version of the observed HSI:

$$\mathbf{Y}_{hm} = \mathbf{R}\mathbf{Y}_{f(h)}. \qquad (13)$$

In order to be compared with the multispectral image, the histogram of each band of $\mathbf{Y}_{hm}$ is matched with the corresponding band of $\mathbf{Y}_m$. Then, $\mathbf{Y}_{hm}$ is partitioned into patches and different sets of endmembers are extracted independently from each patch, $\mathcal{P}_j$, $j \in \{1, \ldots, P\}$. For each patch $\mathcal{P}_j$, we identify a mixing matrix $\mathbf{D}_{m,j}$ by means of an EIA algorithm. After determining the local dictionary $\mathbf{D}_{m,j}$, the code $\mathbf{A}$ is estimated by solving the following constrained least squares (CLS) optimization problem:

$$\min_{\mathbf{A}_j > 0} \|\mathbf{Y}_{m,j} - \mathbf{D}_{m,j}\mathbf{A}_j\|_F^2, \qquad (14)$$

where the dictionary does not need to be spectrally downsampled by the spectral response, $\mathbf{R}$, since it has been induced from $\mathbf{Y}_{hm}$.

We are taking advantage of that the set of endmembers induced by certain EIA algorithms, i.e. the Vertex Component

Algorithm *HSR-LDL-EIA (II)*
Input:
1. $\mathbf{Y}_{f(h)}$, $\mathbf{Y}_{hm}$, $\mathbf{Y}_m$, $\mathbf{R}$, $\mathcal{P}_j$, $n_d \le n_m$
2. **for** $j = 1$ **to** P
3. $\qquad \mathbf{Y}_{hm,j} \leftarrow I_{hm}(\mathcal{P}_j)$
4. $\qquad \mathbf{Y}_{m,j} \leftarrow I_m(\mathcal{P}_j)$
5. $\qquad \mathbf{k}_j, \mathbf{D}_{m,j} := \text{EIA}(\mathbf{Y}_{hm,j}, n_d)$
6. $\qquad \widehat{\mathbf{A}}_j := \arg\min_{\mathbf{A}_j \ge 0} \|\mathbf{Y}_{m,j} - \mathbf{D}_{m,j}\mathbf{A}_j\|_F^2$
7. $\qquad \mathbf{D}_j \leftarrow \mathbf{Y}_{f(h)}(\mathbf{k}_j)$
8. $\qquad \widehat{\mathbf{X}}_j := \mathbf{D}_j \widehat{\mathbf{A}}_j$
9. $\widehat{\mathbf{X}} := \left\{ \widehat{\mathbf{X}}_j \right\}$, $j = 1, \ldots, P$

Fig. 3. Pseudo-code of the second proposed HSI Super Resolution algorithm via Local Dictionary Learning using Endmember Induction Algorithms (HSR-LDL-EIA (II)).

Algorithm *Patches definition using sliding windows*
Input:
1. $\mathbf{Y}$, s, t
2. $i_x = 0$, $i_y = 0$, $j = 0$
3. **while** $i_x < n_x/d$
4. $\qquad$ **while** $i_y < n_y/d$
5. $\qquad\qquad \mathcal{P}_j = [i_x, i_x + s] \times [i_y, i_y + s] \bigcap \mathcal{I}$
6. $\qquad\qquad i_y = i_y + s - t$, $j = j + 1$
7. $\qquad i_x = i_x + s - t$

Fig. 4. Pseudo-code of the patches definition algorithm using sliding windows.

Analysis (VCA) [22], are obtained by selecting a subset of the original pixels. The indexes of these endmembers, $\mathbf{k}_j$, are used to define the dictionary, $\mathbf{D}_j$, from $\mathbf{Y}_{f(h)}$. Finally, the super-resolution hyperspectral patch can be reconstructed by combining the dictionary $\mathbf{D}_j$ with the estimated code $\widehat{\mathbf{A}}_j$:

$$\widehat{\mathbf{X}}_j := \mathbf{D}_j \widehat{\mathbf{A}}_j. \tag{15}$$

In this case, the indicator functions, $\mathbf{Y}_{hm,j} \leftarrow I_{hm}(\mathcal{P}_j)$ and $\mathbf{Y}_{m,j} \leftarrow I_m(\mathcal{P}_j)$, are easily defined since the patches are obtained from $\mathbf{Y}_{hm,j}$, and both $\mathbf{Y}_{hm,j}$ and $\mathbf{Y}_{m,j}$ share the same spatial support $\mathcal{I} = [1, n_x] \times [1, n_y] \subset \mathbb{N}^2$. Then, a patch will be defined as a set of indexes lying inside the spatial support $\mathcal{I}$:

$$\mathcal{P} = \{i_l \subseteq \mathcal{I}\}_{l=1}^L, \tag{16}$$

where $L \in \mathbb{N}$ denotes the number of indexes in the patch. The definitions of the functions, $I_{hm}(\mathcal{P}_j)$ and $I_m(\mathcal{P}_j)$, are similar:

$$\mathbf{Y}_{hm,j} \leftarrow I_{hm}(\mathcal{P}_j) = \{\mathbf{y}_{hm,i} \in \mathbf{Y}_{hm} \mid i \in \mathcal{P}_j\}, \tag{17}$$

$$\mathbf{Y}_{m,j} \leftarrow I_m(\mathcal{P}_j) = \{\mathbf{y}_{m,i} \in \mathbf{Y}_m \mid i \in \mathcal{P}_j\}. \tag{18}$$

IV. Local spatial patches definition

In this section, we introduce two approaches to obtain the patches definition from a given image $\mathbf{Y}$ with spatial support $\mathcal{I} \subset \mathbb{N}^2$. The first approach uses a conventional sliding window to define the patches, while the second approach relies on a binary partition tree (BPT) representation [17].

A. Patches definition using a sliding window of fixed size

The sliding window methodology has been broadly used in image processing. It consists in defining a square window of fixed size, $s \in \mathbb{N}$, that identifies a patch of the image, $\mathcal{P}$. The window slides over the whole image in a standard zig-zag way, be it with some overlapping or not, eventually covering the whole image and defining the image patches. Fig. 4 shows the pseudo-code of the patches definition algorithm using sliding windows. The algorithm takes as inputs an image, $\mathbf{Y}$, the sliding window size $s > 0$, and the overlapping, t, $0 \le t < s$. The pseudocode of the patches definition by means of sliding

windows is presented in Fig. 4. The sliding windows cover the image in a zig-zag way controlled by the two loops. Occasionally, a sliding window could partially lie outside the borders of the image. This issue is addressed by the definition of the patch as an intersection to the image support:

$$\mathcal{P}_j = [i_x, i_x + s\} \times \,]i_y, i_y + s] \bigcap \mathcal{I}. \tag{19}$$

Thus, the number of pixels, L, contained inside a patch will be upper bounded by the size of the sliding windows, $L \le s^2$.

When the patches are defined by overlapping sliding windows, the local approaches presented in Section III must deal with the fact that some super-resolution pixels could be estimated from different patches. A simple way of addressing this issue is to average the estimated super-resolution spectra. Formally, being $\hat{\mathbf{x}}_{i,j}$ the super-resolution pixel i estimated from patch j, then the super-resolution pixel $\hat{\mathbf{x}}_i$ is given by:

$$\hat{\mathbf{x}}_i = \frac{1}{\sum_{j=1}^P I_j(\hat{\mathbf{x}}_i)} \sum_{j=1}^P I_j(\hat{\mathbf{x}}_i) \hat{\mathbf{x}}_{i,j}, \tag{20}$$

where $I_j(\hat{\mathbf{x}}_i)$ is an indicator function taking the value 1 if the pixels $\hat{\mathbf{x}}_i$ is being estimated from patch $\mathcal{P}_j$, and the value 0 otherwise. Other approaches could be used, for instance, selecting the median pixel.

B. Patches definition using a BPT representation

The BPT is a hierarchical region-based representation of an image in a tree structure [17]. In the BPT representation, the leaf nodes correspond to an initial partition of the image, which can be the individual pixels, or a coarser segmentation map. From this initial partition, an iterative bottom-up region merging algorithm is applied until only one region remains. This last region represents the whole image and corresponds to the root node. All the nodes between the leaves and the root result of the merging of two adjacent children regions.

Two notions are of prime importance when defining a BPT, the *region model*, $\mathcal{M}_\mathcal{R}$, which specifies how a region $\mathcal{R}$ is modelled, and the *merging criterion*, $\mathcal{O}(\mathcal{M}_{\mathcal{R}_\alpha}, \mathcal{M}_{\mathcal{R}_\beta})$, which is a similarity measure between the region models of any two regions $\mathcal{R}_\alpha$ and $\mathcal{R}_\beta$. Each merging iteration involves the search of the two neighbouring regions which achieve the lowest pair-wise similarity among all the pairs of neighbouring regions in the current segmentation map. Those two regions are consequently merged. To build the BPT representation

Algorithm *Patches definition using BPT*
Input:
1. $\mathbf{Y}_h, n_d, n_r$
2. $\mathcal{T} \leftarrow \text{BPT}(\mathbf{Y}_h)$
3.   **for all** $\mathcal{R}_j \in \mathcal{T}$
4. $\mathbf{D}_j := \text{EIA}(\mathbf{Y}_{h,j}, n_d)$
5. $\widehat{\mathbf{A}}_j := \arg\min_{\mathbf{A}_j \geq 0} \|\mathbf{Y}_{h,j} - \mathbf{D}_j \mathbf{A}_j\|_F^2$
6. $\mathcal{P} \leftarrow \text{pruning}(\mathcal{T}, n_r)$

Fig. 5. Pseudo-code of the patches definition algorithm using a BPT representation.

from a hyperspectral image [18], [19], [23], $\mathbf{Y}_h$, we use the first-order parametric model $\mathcal{M}_{\mathcal{R}}$:

$$\mathcal{M}_{\mathcal{R}} \overset{d}{=} \bar{\mathbf{x}} = \frac{1}{N_{\mathcal{R}}} \sum_{i=1}^{N_{\mathcal{R}}} \mathbf{x}_i, \tag{21}$$

where $N_{\mathcal{R}}$ is the number of pixels on the region; and, in order to merge regions, the spectral angle distance:

$$\mathcal{O}\left(\mathcal{M}_{\mathcal{R}_\alpha}, \mathcal{M}_{\mathcal{R}_\beta}\right) \overset{d}{=} d_{\text{SAM}}\left(\bar{\mathbf{x}}_\alpha, \bar{\mathbf{x}}_\beta\right) = \arccos\left(\frac{\bar{\mathbf{x}}_\alpha \bar{\mathbf{x}}_\beta}{\|\bar{\mathbf{x}}_\alpha\| \|\bar{\mathbf{x}}_\beta\|}\right). \tag{22}$$

Once the BPT representation is built, the BPT is pruned to achieve a partition of the image such that the regions of the partition define the image patches, $\mathcal{P}_j$. We propose to use the unmixing-based pruning strategy introduced in [19]. Let $\mathcal{P}$ be a partition of the image (a pruning of the BPT) and Ω be the set of all possible partitions. Then, the spectral unmixing of each region, $\mathcal{R}$, is computed and for each pixel, $\mathbf{x} \in \mathcal{R}$, its reconstruction error, $\epsilon_{\mathcal{R}}(\mathbf{x}, \hat{\mathbf{x}})$, is calculated by means of the average RMSE, where $\hat{\mathbf{x}}$ is the reconstructed pixel using the unmixing information. The unmixing-based pruning criterion defines the optimal partition as the one minimizing the overall average RMSE regularized by the number of regions in the partition:

$$\mathcal{P}^\star = \arg\min_{\mathcal{P} \in \Omega} \frac{1}{N} \sum_{\mathcal{R} \in \mathcal{P}} \sum_{\mathbf{x} \in \mathcal{R}} \epsilon_{\mathcal{R}}(\mathbf{x}, \hat{\mathbf{x}}) + \lambda |\mathcal{P}|, \tag{23}$$

where N denotes the number of pixels in the image, and $|\mathcal{P}|$ denotes the number of regions in the partition. Fig. 5 shows the pseudo-code of the patches definition by means of a BPT representation. Given the image $\mathbf{Y}_h$ and the number of endmembers n_d, a BPT representation of $\mathbf{Y}_h$ is built using (21) and (22). Then, the dictionary and the abundances are obtained and stored for each region in the BPT representation in order to later use this information in the pruning of the BPT (23). The regions in the optimal partition $\mathcal{P}$ obtained by the pruning process define the patches. A user can give as input an approximate number of regions, n_r, that the optimal partition is expected to contain. Then, the λ value in (23) is automatically obtained so the number of regions in the optimal partition is as close to the input parameter n_r as possible.

V. ESTIMATION OF THE SPECTRAL RESPONSE

The authors in [16] estimate the spectral response $\mathbf{R}$ from the data by means of an optimization problem:

$$\underset{\mathbf{R}}{\text{minimize}} \quad \|\mathbf{R}\mathbf{Y}_h - \mathbf{Y}_m \mathbf{B}\mathbf{M}\|^2 + \lambda_R \phi_R(\mathbf{R}), \tag{24}$$

where $\phi_R(\cdot)$ is a quadratic regularizer and $\lambda_R \geq 0$ is the respective regularization parameter. Recall that HSIs generally have a large correlation between bands. Consequently, the spectral vectors, of size n_h, usually "live" in a subspace of dimension much lower than n_h [3]. This implies that, when using the observed data to estimate the matrix $\mathbf{R}$, it is not possible to fully estimate it, and only the component of $\mathbf{R}$ parallel to the mentioned subspace can be found. Conveniently, the orthogonal component of $\mathbf{R}$ has essentially no influence on the result of the image fusion. Due to this, the regularizer $\phi_R(\cdot)$ is used to deal with the indetermination of this orthogonal component, and to reduce estimation noise. Furthermore, if information on the spectral coverage of the sensors is available, this estimate can be improved by taking the correspondence between bands of the HSIs and MSIs into account. In this case, the elements of $\mathbf{R}$ that correspond to non-overlapping bands are constrained to be zero.

We estimate the spectral response independently for each of the MSI bands. Let $\mathbf{r}_i^T$ denote a row vector containing the ith row of $\mathbf{R}$ without the elements that are known to correspond to hyperspectral bands that do not overlap the ith multispectral band, $\mathbf{Y}_{h,i}$ denote the matrix $\mathbf{Y}_h$ without the rows corresponding to those same bands, and $\mathbf{Y}_{m,i:}$ denote the ith row of $\mathbf{Y}_m$. The regularizer mentioned previously is given by $\|\mathbf{H}\mathbf{r}_i\|^2$, where the product by $\mathbf{H}$ computes the differences between the elements in $\mathbf{r}_i$ corresponding to contiguous hyperspectral bands. Note that this choice of regularizer is connected to the knowledge that the spectral response of the sensors should be somewhat smooth between these contiguous hyperspectral bands. Taking this into account, the solution of (24) is given by

$$\mathbf{r}_i^* = \left[\mathbf{Y}_{h,i}\mathbf{Y}_{h,i}^T + \lambda_R \mathbf{H}^T \mathbf{H}\right]^{-1} \mathbf{Y}_{h,i} \left[\mathbf{Y}_{m,i:} \mathbf{B}\mathbf{M}\right]^T. \tag{25}$$

Note that the estimation of $\mathbf{R}$, as presented so far, requires explicit knowledge of matrix $\mathbf{B}$. Since we do not know $\mathbf{B}$ we adapt the technique described in [16] to deal with this. If both the observed HSIs and MSIs are blurred with a strong spatial blur, the effect of $\mathbf{B}$ becomes negligible. Following this, the estimate of the spectral response $\mathbf{R}$ is made using (25) on the spatially blurred versions of the observed data, setting the kernel of the spatial blur associated with $\mathbf{B}$ to a delta impulse.

VI. MATERIALS AND EXPERIMENTAL METHODOLOGY

In what follows, we describe the datasets, the experimental methodology and the quantitative quality measures employed to compare the proposed HSR-LDL-EIA approaches to the baseline HSR-GDL-EIA approach.

A. Datasets

We run the experiments over three different datasets. In the first one we synthesize a super-resolution HSI and then
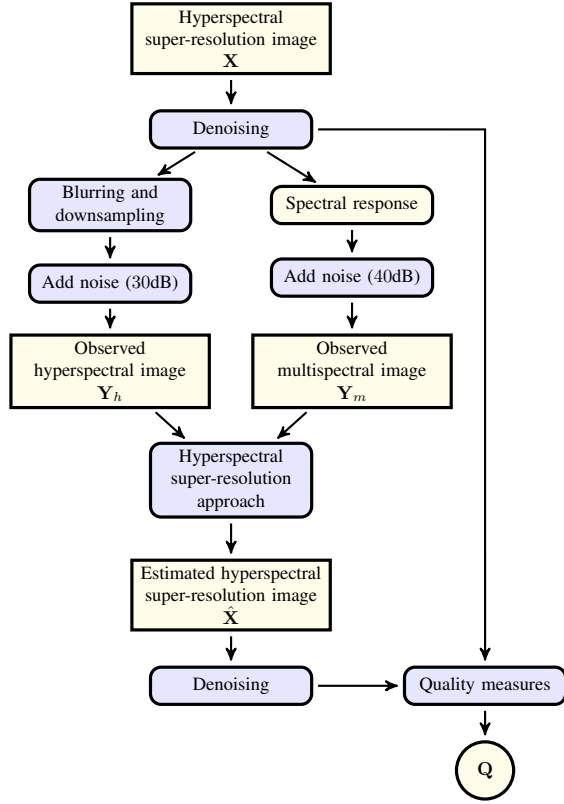
Fig. 6. Flow diagram of the experimental methodology, derived from Wald's protocol (simulated observations), for synthetic and semi-real datasets.
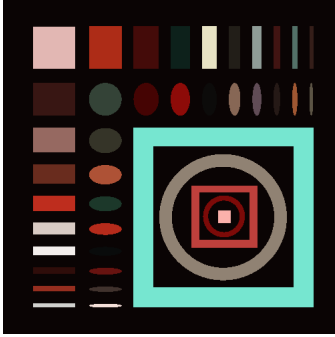

Fig. 7. False color representations of the synthetic dataset.

TABLE I
VALUES OF THE p_{GLOBAL} AND p_{LOCAL} VARIABLES FOR EACH DATASET.

Dataset	p_{global}	p_{local}	d
Synthetic	45	3	4
Paris	{10,30,50}	3	3
San Francisco	10	3	3

image, $\mathbf{X}$, has been denoised, the observed low-spatial resolution hyperspectral image, $\mathbf{Y}_h$, is simulated by applying a Gaussian blurring and then, by downsampling the blurred image selecting one of each d pixels, where d denotes the downsampling factor. A Gaussian noise of 30dB is added to the image to simulate the hyperspectral sensor acquisitions SNR. 2) The observed high-spatial resolution multispectral image, $\mathbf{Y}_m$, is also simulated by applying a spectral response matrix to the denoised super-resolution image. A noise of 40dB is added to the image to simulate the multispectral sensors higher SNR. 3) A super-resolution methodology is applied to the simulated observations in order to obtain the estimated hyperspectral super-resolution image, $\hat{\mathbf{X}}$. 4) Finally, the denoised HSR reference image and the estimated HSR can be compared to obtain quantitative quality measures of the applied super-resolution approach.

1) Synthetic dataset: A synthetic image composed of multiple geometric shapes (ellipses and rectangles) of different sizes and orientations, where each geometrical element and the background are formed using a different linear mixture of 5 endmembers randomly extracted from the U.S. Geological Survey Digital Spectral Library splib06[1], for a total of 45 different endmembers in the image. Fig. 7 shows a false color representation of the synthetic dataset. The purpose of this toy example is to simulate a scenario with perfect conditions for the proposed super-resolution approaches, that is, a high global spectral variability and a local low rank. Being the assumptions valid, the proposed approaches should outperform the state-of-the-art global approach. In addition, it allows us to study the impact that the different parameters of the different strategies to define the patches and the estimation of the spectral response have in the results.

2) Paris dataset: This dataset consists of images taken above Paris (see Fig. 13(a)) by two instruments on board of the Earth Observing-1 Mission (EO-1) satellite, the Hyperion instrument and the Advanced Land Imager (ALI). Hyperion is a grating imaging spectrometer providing 242 hyperspectral bands (from 0.4 to 2.5 μm) with a 30 meter spatial resolution. The ALI instrument provides 9 spectral bands (from 0.43 to 2.35 μm) with 30-meter resolution. We made use only of the ALI spectral bands 4, 7 and 9 to make the problem more challenging [25][2]. Since the two sensors are carried by the same satellite, and the images are acquired simultaneously, then the super-resolution image will not be affected by differences in terms of angle of view, atmospheric path, illumination as well as miss-registration.

we simulate the observed HSI and MSI using a version of the Wald protocol [24] detailed below. The second and third datasets are scenes captured by two sensors (one hyperspectral and another multispectral one) operating at the same spatial resolution. Thus, we consider the captured hyperspectral images as the super-resolution references and simulate only the observed hyperspectral images (which we call half-Wald protocol).

Wald's protocol [24] is a methodological pipeline developed for the evaluation of super-resolution approaches. It works by simulating the observed images from a reference hyperspectral super-resolution image. We further describe the version of the Wald protocol, depicted in Fig. 6, that we employed in the experiments: 1) Once the super-resolution reference

[1]Available at http://speclab.cr.usgs.gov/spectral-lib.html.
[2]More information is available at http://eo1.gsfc.nasa.gov/, http://eo1.usgs.gov/sensors/ali and http://eo1.usgs.gov/sensors/hyperioncoverage.
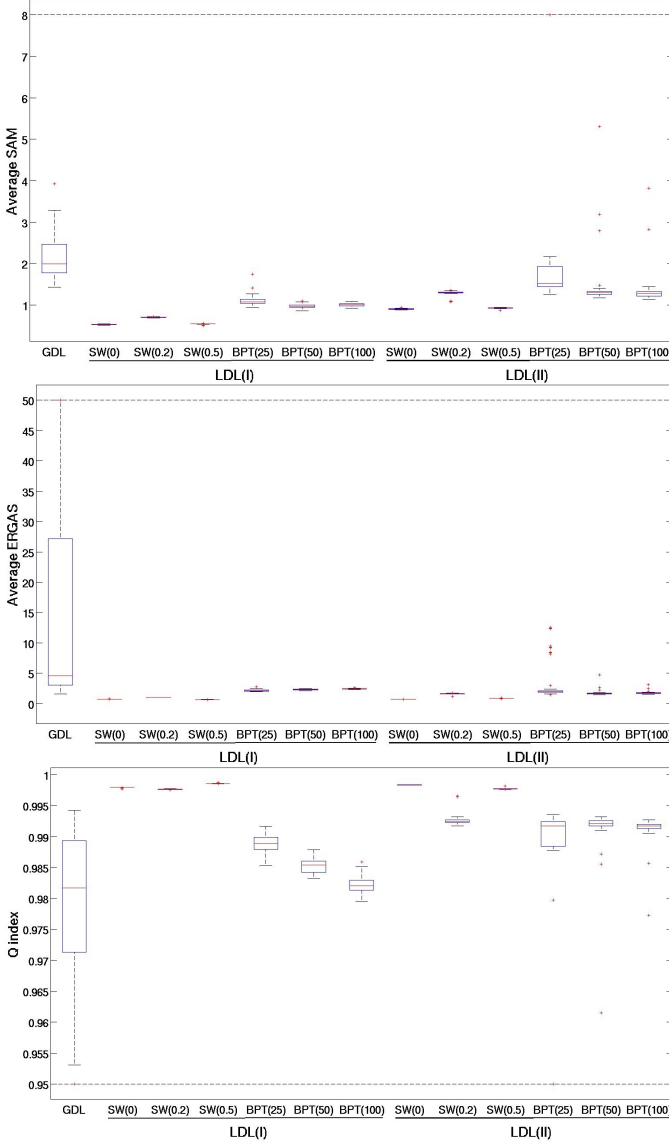
Fig. 8. Reconstruction errors for the synthetic dataset using the actual spectral response matrix. The global approach (GDL) is compared to the proposed local approaches LDL(I) and LDL(II): (top) Average SAM, (middle) Average ERGAS and (bottom) UIQI index.
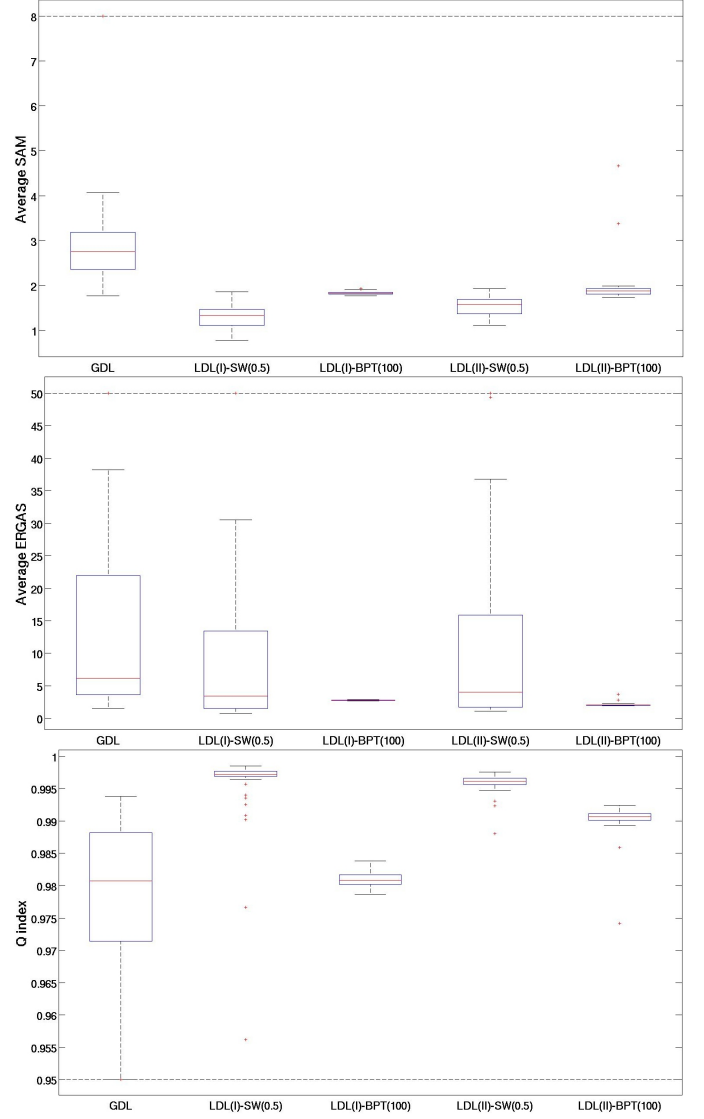

Fig. 9. Reconstruction errors for the synthetic dataset using the spectral response matrix estimation [16]. The global approach (GDL) is compared to the proposed local approaches LDL(I) and LDL(II) for SW with $s = 60$ and $t = 0.5s$, and BPT with approximately 100 regions: (top) Average SAM, (middle) Average ERGAS and (bottom) UIQI index.

3) San Francisco dataset: This dataset (see Fig. 15(a)) was acquired by Hyperion and ASTER sensors over San Francisco on 31th July, 2002. On that date, two sensors were in the same orbit with a 30 min time difference (Hyperion: 10:35 am, ASTER: 11:05 am). We assume that the observation conditions for these two images were the same. The Hyperion/VNIR data with 30 m ground sampling distance and 50 bands (spectral channels 8 to 57) and ASTER/VNIR data with 15 m ground sampling distance and three bands were used for the fusion of hyperspectral and multispectral data.

B. Experimental methodology

We compared the two proposed local super-resolution approaches, HSR-LDL-EIA (I) and HSR-LDL-EIA (II), to the baseline global approach, HSR-GDL-EIA. For the proposed local approaches, we also considered the patches definition by either, sliding windows (SW) or a BPT representation. We also studied the impact of estimating the spectral response with respect to the use of the actual spectral response. Next, we detail the implementation aspects required for running the experiments and the employed quantitative quality measures.

For each experiment, 50 Monte Carlo runs are obtained in order to estimate the robustness of the competing approaches with respect to the noise and the stochastic components, *i.e.*, the spectral unmixing. For each image, $\mathbf{I}$, we performed denoising by factorizing it via Singular Value Decomposition (SVD). This denoising step ensures that the actual super-resolution image is (almost) free of noise, so the evaluation of the competing methods is done according only to their capacity to estimate the super-resolution signal. To apply SVD, the p_{global} eigenvectors, $\mathbf{U}_p$, with higher eigenvalues were retained. Then, the image is projected onto the subspace spanned by
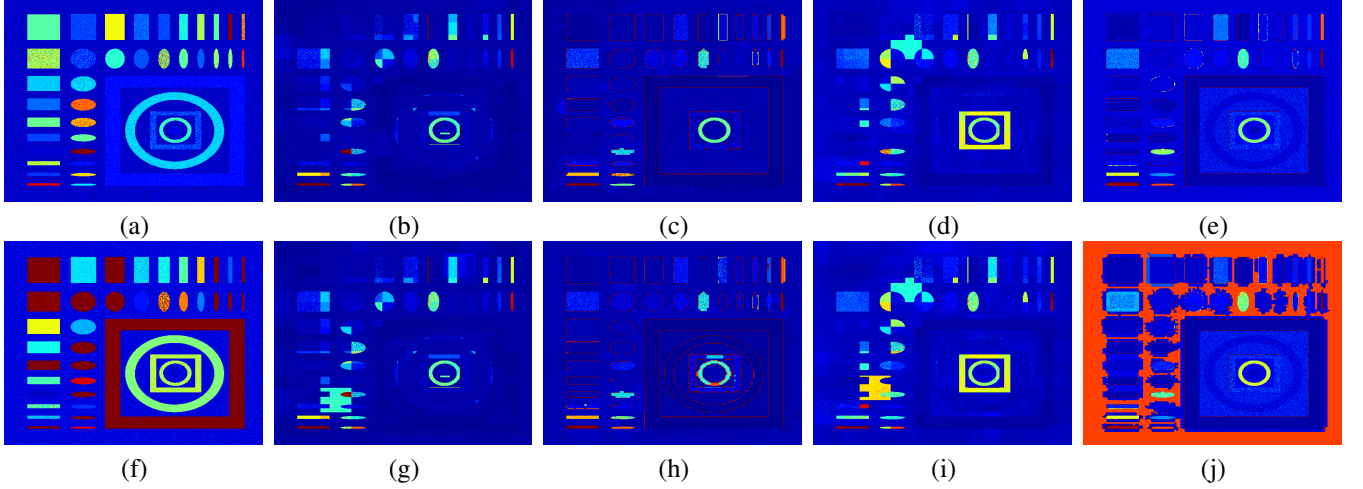
Fig. 10. False color images of the angular errors between the actual super-resolution image and the best (top row) and worst (bottom row) super-resolution hyperspectral images estimated from the syntehtic dataset using: (a,f) HSR-GDL-EIA, (b,g) HSR-LDL-EIA(I)-SW, (c,h) HSR-LDL-EIA(I)-BPT, (d,i) HSR-LDL-EIA(II)-SW, (e,j) HSR-LDL-EIA(II)-BPT. The color scale goes from 0 degrees error (dark blue) to 10 degrees error (dark red).

$\mathbf{U}_p$: $\mathbf{I} \leftarrow \mathbf{U}_p^T \mathbf{U}_p \mathbf{I}$. The value of p_{global} varies with each image (see Table I). For the synthetic dataset, p_{global} is set to the number of different spectral signatures used to build the data. For the remainder datasets, p_{global} has been set according to the estimated intrinsic dimensionality (ID) using the HySime algorithm [26]. The estimated ID for the Paris dataset is 31, and it is 11 for the San Francisco dataset. To explore the effect of the p_{global} parameter, we run experiments using $p_{\text{global}} = \{10, 30, 50\}$ for the Paris dataset.

Dictionary learning is done by the Vertex Component Analysis (VCA) algorithm [22]. Due to its stochastic nature, every time a set of endmembers is estimated, we run VCA ten times and retain the set of endmembers with maximum volume among the ten runs [27]. The number of estimated endmembers depends on each image as well. For the global HSR-GDL-EIA approach the values given by p_{global} are used. For the proposed local approaches, the values given by p_{local} are used instead. The value of p_{local} is set in each dataset to the number of available multispectral bands (see Table I). The setting of this parameter is hence not critical. In order to solve the optimization problems in (5), (6) and (14), we used the SUnSAL algorithm [28], which is an instance of the C-SALSA methodology introduced in [29] to effectively solve a large number of constrained least-squares problems sharing the same matrix system. On the datasets where the observed multispectral image, $\mathbf{Y}_m$ is simulated we used the spectral response of the IKONOS sensor. This sensor captures both a panchromatic $(0.45 - 0.90\mu m)$ and four multispectral bands $(0.45 - 0.52, 0.52 - 0.60, 0.63 - 0.69$ and $0.76 - 0.90\mu m)$ [30]. In the experiments we employed the first three multispectral bands only.

C. Quantitative quality measures

The competing HSI super-resolution approaches are evaluated by quantitative quality measures obtained comparing the original image, $\mathbf{X}$, to the estimated super-resolution image, $\hat{\mathbf{X}}$. We made use of three different image comparison quality measures [1], [31]: the average spectral angle distance (SAD), the Universal Image Quality Index (UIQI) [32] and the *Erreur Relative Globale Adimensionnelle de Synthèse* (ERGAS) quality measure [31].

The SAD measures the average angular spectral reconstruction error:

$$\text{avgSAD}\left(\mathbf{X}, \hat{\mathbf{X}}\right) = \frac{1}{n} \sum_{i=1}^{n} \epsilon_{\text{SAD}}\left(\mathbf{x}_i, \hat{\mathbf{x}}_i\right), \qquad (26)$$

where $\epsilon_{\text{SAD}}\left(\mathbf{x}_i, \hat{\mathbf{x}}_i\right) = \arccos\left(\frac{\mathbf{x}^T \hat{\mathbf{x}}}{\|\mathbf{x}\| \|\hat{\mathbf{x}}\|}\right)$ is the spectral angle distance.

The UIQI index measures the average correlation between the original and the estimated images. Lets $\mathbf{x}^{(l)}, \hat{\mathbf{x}}^{(l)} \in \mathbb{R}^n$ denote the l-th band of the original and estimated images respectively, the band correlation Q index is defined as:

$$Q\left(\mathbf{x}^{(l)}, \hat{\mathbf{x}}^{(l)}\right) = \frac{\sigma_{\mathbf{x}^{(l)}\hat{\mathbf{x}}^{(l)}}}{\sigma_{\mathbf{x}^{(l)}}\sigma_{\hat{\mathbf{x}}^{(l)}}} \frac{2\mu_{\mathbf{x}^{(l)}}\mu_{\hat{\mathbf{x}}^{(l)}}}{\mu_{\mathbf{x}^{(l)}}^2 + \mu_{\hat{\mathbf{x}}^{(l)}}^2} \frac{2\sigma_{\mathbf{x}^{(l)}}\sigma_{\hat{\mathbf{x}}^{(l)}}}{\sigma_{\mathbf{x}^{(l)}}^2 + \sigma_{\hat{\mathbf{x}}^{(l)}}^2}, \quad (27)$$

where $\mu_{\mathbf{x}^{(l)}}$ and $\mu_{\hat{\mathbf{x}}^{(l)}}$ denote the mean vectors of the original and estimated images respectively; $\sigma_{\mathbf{x}^{(l)}}$ and $\sigma_{\hat{\mathbf{x}}^{(l)}}$ denote the variances, and $\sigma_{\mathbf{x}^{(l)}\hat{\mathbf{x}}^{(l)}}$ the covariance. The UIQI is the average Q index over all the bands:

$$\text{UIQI}\left(\mathbf{X}, \hat{\mathbf{X}}\right) = \frac{1}{n_h} \sum_{l=1}^{n_h} Q\left(\mathbf{x}^{(l)}, \hat{\mathbf{x}}^{(l)}\right). \qquad (28)$$

The ERGAS evaluates both spectral and spatial divergences:

$$\text{ERGAS}\left(\mathbf{X}, \hat{\mathbf{X}}\right) = \frac{100}{S} \sqrt{\frac{1}{n_h} \sum_{l=1}^{n_h} \left(\frac{\epsilon_{\text{RMSE}}\left(\mathbf{x}^{(l)}, \hat{\mathbf{x}}^{(l)}\right)}{\mu_{\mathbf{x}^{(l)}}}\right)^2}, \tag{29}$$

where S denotes the spatial ratio between the observed hyperspectral and the observed multispectral images, $\mathbf{x}^{(l)}, \hat{\mathbf{x}}^{(l)} \in \mathbb{R}^n$ denote the l-th band of the original and estimated images, respectively, and $\mu_{\mathbf{x}^{(l)}} \in \mathbb{R}$ denotes the mean value of the original l-th spectral band, $\mathbf{x}^{(l)}$.
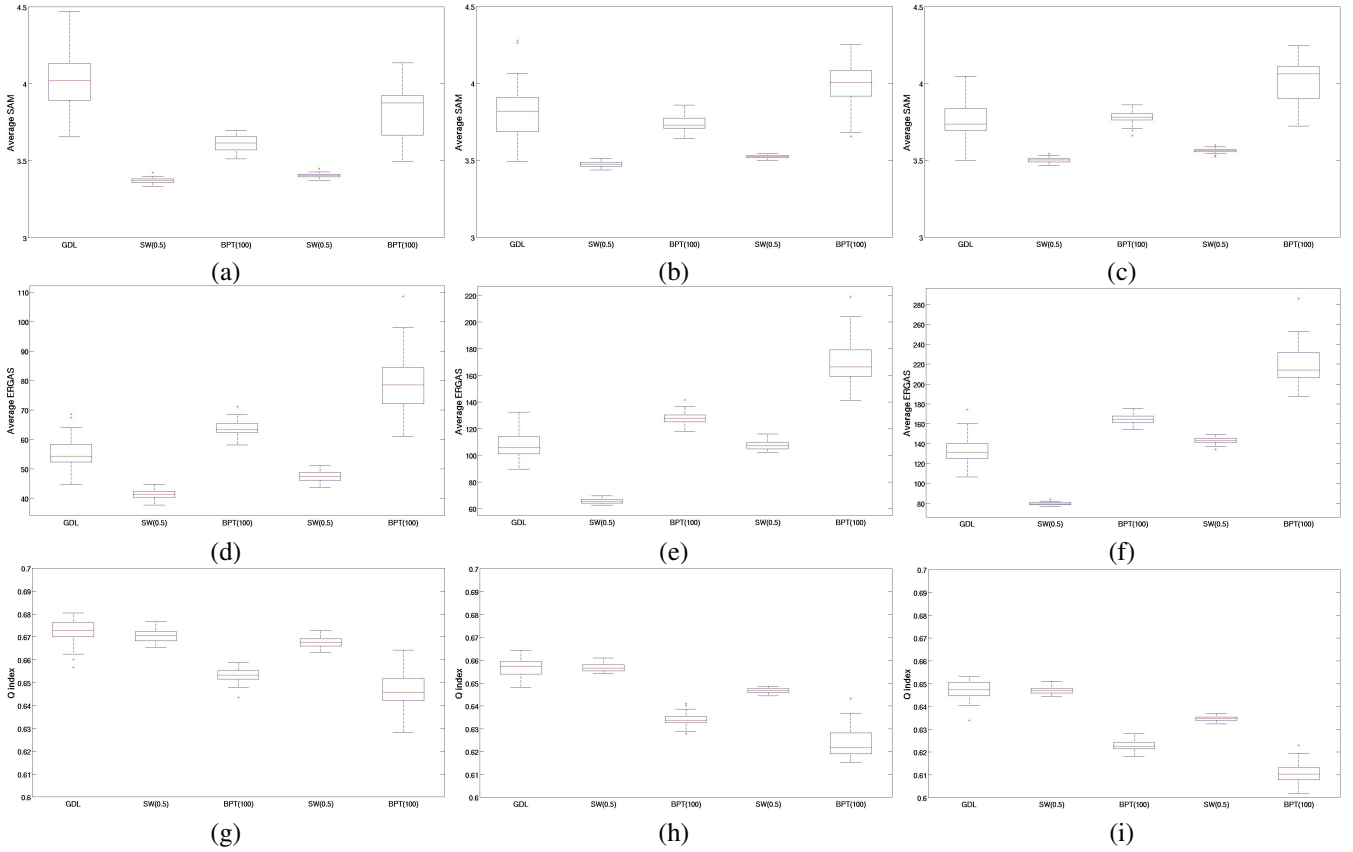
Fig. 11. Reconstruction errors for the Paris dataset using the estimated spectral response matrix [16]. The global approach (GDL) is compared to the proposed local approaches LDL(I) and LDL(II): (top) Average SAM, (middle) Average ERGAS and (bottom) UIQI index; (a,d,g) $p_{\text{global}} = 10$, (b,e,h) $p_{\text{global}} = 30$, and (c,f,i) $p_{\text{global}} = 50$.

VII. EXPERIMENTAL RESULTS

A. Results for the synthetic dataset

Fig. 8 shows the boxplots of the quality measures calculated from the 50 Monte Carlo runs of the competing global and local approaches using either the SW or the BPT patches definition. For the sliding windows, we experimentally selected a window size of $s = 60$, that is a 60×60 pixels window, with three different overlapping factors: $t = [0, 0.2s, 0.5s]$. For the BPT approach, the approximately expected number of regions in the optimal partitions have been set to: $n_r = [25, 50, 100]$. The actual spectral response matrix, $\mathbf{R}$, is assumed to be known. Results show that, as expected, the local approaches clearly outperform the global baseline approach, taking advantage of the local low rank.

Fig. 9 shows the results obtained for the competing approaches over the synthetic dataset when the spectral response is unknown and has to be estimated from the data. We set the SW parameters to $s = 60$ and $t = 0.5s$, and the expected number of regions in the BPT partition to be 100. Results show that the local approaches also outperform the global approach in a global high rank but locally low rank scenario, when the spectral response should be estimated from the data. In this case, it could be appreciated that the BPT-based patches definition achieves more stable results than the SW one.

The SAD image errors of the best and worst Monte Carlo runs using the actual spectral response matrix, depicted in Fig. 10, help to understand the differences between the SW and the BPT patches definition approaches. The SW produces error images with grid patterns, specially on the transitions between the geometrical shapes. However, the patches provided by the BPT representation form a partition of the image which is adapted to the homogeneities of the geometrical shapes. Thus, the error is concentrated on the edges of the geometrical shapes or on the slim geometrical areas which are difficult to represent as a single patch. This is more evident in the error images of the worst runs using the HSR-LDL-EIA(II) approach. Since the patches are obtained from the spatially upscaled and spectrally downscaled, $\mathbf{Y}_{hm}$ image, the BPT representation is more complex than in the global and the HSR-LDL-EIA(I) approaches, and occasionally fails to achieve low rank patches (i.e., the background). The SW overlapping factor does not significantly affect the results, neither the number of approximate regions in the partition obtained from the BPT representation once the number of regions is high enough to achieve low rank regions.

B. Results for Paris and San Francisco datasets

Since both semi-real datasets present smaller objects than the synthetic data, specially for the urban areas, we experimentally selected a window size of 60×60 pixels ($s = 60$), and fixed the overlapping factors to $t = 0.5s$. For the BPT
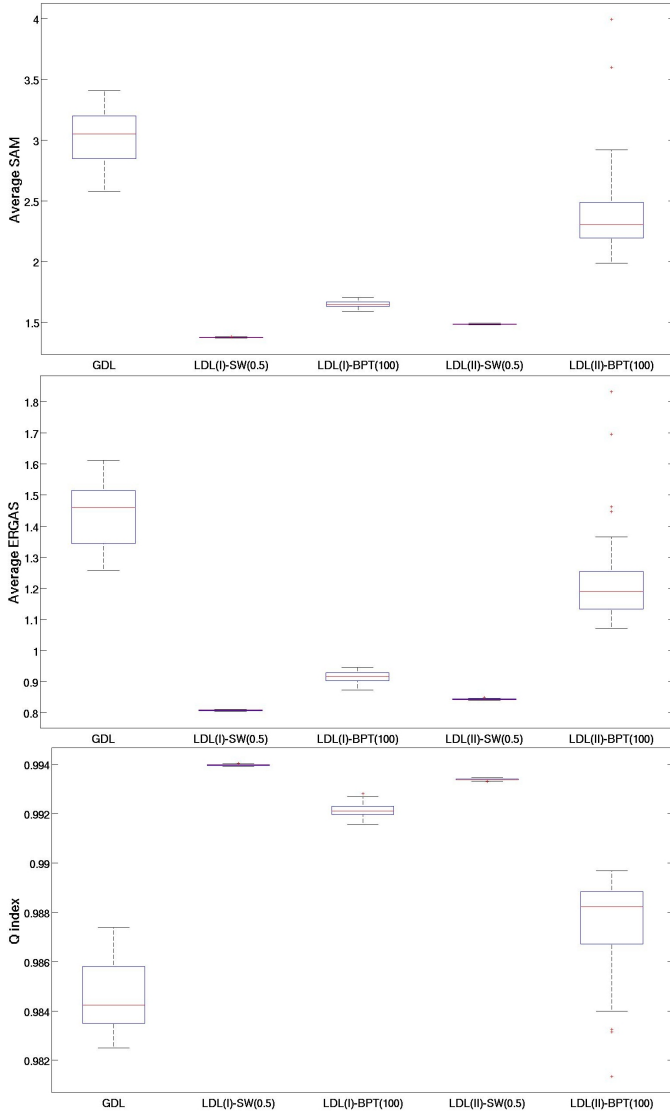
Fig. 12. Reconstruction errors for the San Francisco dataset using the estimated spectral response matrix [16]. The global approach (GDL) is compared to the proposed local approaches LDL(I) and LDL(II): (top) Average SAM, (middle) Average ERGAS and (bottom) UIQI index.

approach, the approximately expected number of regions in the optimal partitions have been set to $n_r = 100$. In both datasets, the spectral response is unknown and has been estimated from the data. The quality measures obtained for the 50 Monte Carlo runs of the competing approaches using the semi-real Paris and San Francisco datasets, are depicted in Figs. 11 and 12, respectively.

For the latter, the proposed local approaches clearly outperform the global approach in all the three quality measures. For the former, the proposed local approaches outperform the global baseline approach in terms of angular error. However, there is a significant difference in terms of ERGAS and Q index errors, depending on the patches definition approach. The SW achieves better results than the BPT approach, probably because the Paris dataset is an urban dataset with many small details and edges, while the San Francisco dataset has more natural landscapes, which are easier to capture by the

BPT representation. Again, the HSR-LDL-EIA(II) approach in combination with the BPT-based patches definition, achieves worse results than the other local approaches for the semi-real datasets. This evidences that the BPT representation fails to obtain a good representation of the low rank patches from the spatially upscaled and spectrally downscaled image. False colour images of the best and worst HSR obtained by the global and the local approaches of the Paris and San Francisco datasets, are shown in Figs. 13 and 15, respectively. The p_{global} parameter seems to not affect the trend on the results comparison. This suggests that as far as we do not introduce noise in the reference super-resolution image by setting p_{global} to a too high value, or we reduce the global information by setting p_{global} to a too small value (close to p_{local}), the competing approaches give consistent results.

Figs. 14 and 16, show the angular error images for the best and worst results obtained using the competing approaches from the Paris and San Francisco datasets, respectively. Although the visual assessment is more difficult than in the synthetic dataset, the grid pattern of the SW approach is still recognizable. The BPT approach presents more homogeneous error areas than the SW approach, although sometimes failing to achieve a partition with low rank regions, specially for the HSR-LDL-EIA(II) approach. It can also be noticed that the water areas present the higher angular errors, possibly due to the low spectral response of the water.

VIII. Conclusions

We have proposed two novel methods for hyperspectral super-resolution via local dictionary learning using endmember induction algorithms, termed HSR-LDL-EIA(I) and HSR-LDL-EIA(II). We have also provided two alternative methodologies to define the local patches, either by using sliding windows (SW) or a BPT-representation. The experimental results show that the proposed approaches are useful for the estimation of super-resolution hyperspectral images from locally low rank complementary hyperspectral and multispectral observations, even if the actual spectral response is unknown and should be estimated from the data. The SW patches definition approach performs well in all the datasets that we have tested, but it requires to set an appropriate window size and overlapping factor, although we did not evidence that the latter was of big relevance in our experiments. The BPT-based patches definition seems to be less stable than the SW. However, it naturally adapts to the geometry of the objects on the scene, once a conservative estimation of the number of regions in the image is provided.

Further research will focus on improving the results of the BPT-based approach. A plausible research avenue is to make use of a set of hierarchical partitions of the image, provided by the BPT representation, in order to increase the robustness of the proposed local HSR approaches. We will also foresee the application of local approaches to other data fusion problems, *i.e.*, pansharpening.
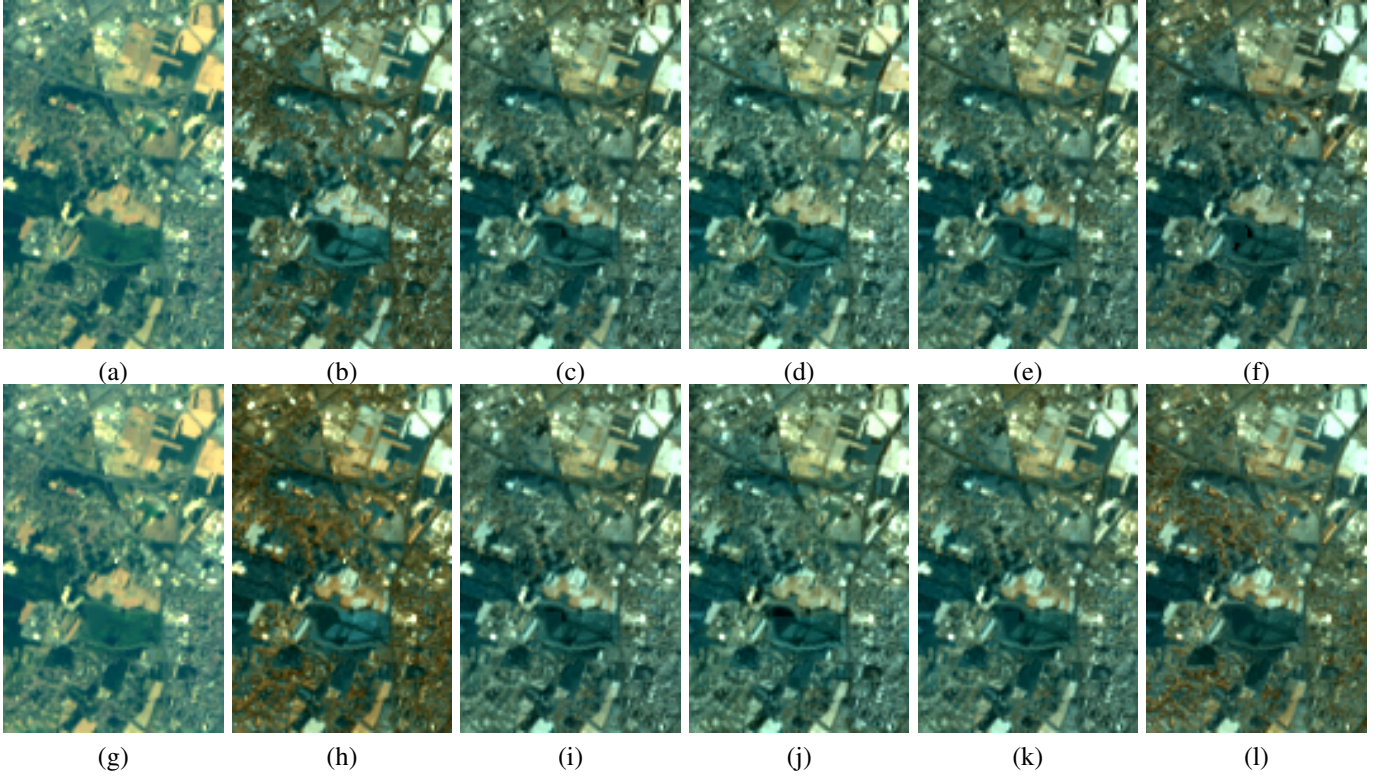
Acknowledgments

Fig. 13. False color images of the best (top row) and worst (bottom row) super-resolution hyperspectral images estimated from the Paris dataset using: (b,h) HSR-GDL-EIA, (c,i) HSR-LDL-EIA(I)-SW, (d,j) HSR-LDL-EIA(I)-BPT, (e,k) HSR-LDL-EIA(II)-SW, (f,l) HSR-LDL-EIA(II)-BPT. (a,g) Super-resolution reference image.
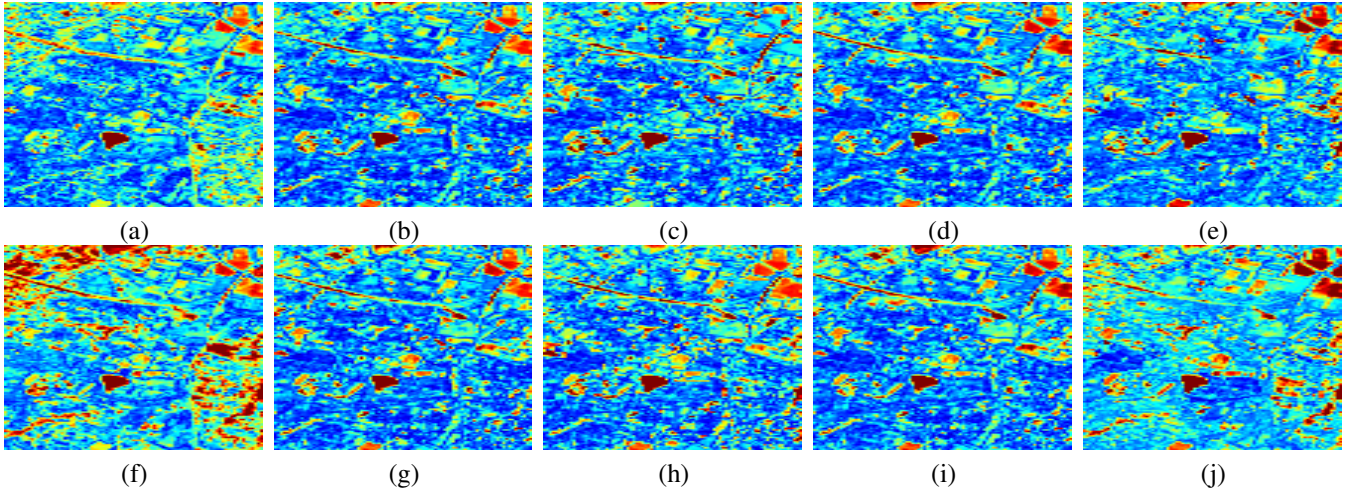

Fig. 14. False color images of the angular errors between the actual super-resolution image and the best (top row) and worst (bottom row) super-resolution hyperspectral images estimated from the Paris dataset using: (a,f) HSR-GDL-EIA, (b,g) HSR-LDL-EIA(I)-SW, (c,h) HSR-LDL-EIA(I)-BPT, (d,i) HSR-LDL-EIA(II)-SW, (e,j) HSR-LDL-EIA(II)-BPT. The color scale goes from 0 degrees error (dark blue) to 10 degrees error (dark red).

REFERENCES

[1] H. Song, B. Huang, K. Zhang, and H. Zhang, "Spatio-spectral fusion of satellite images based on dictionary-pair learning," *Information Fusion*, vol. 18, no. 0, pp. 148 – 160, 2014.
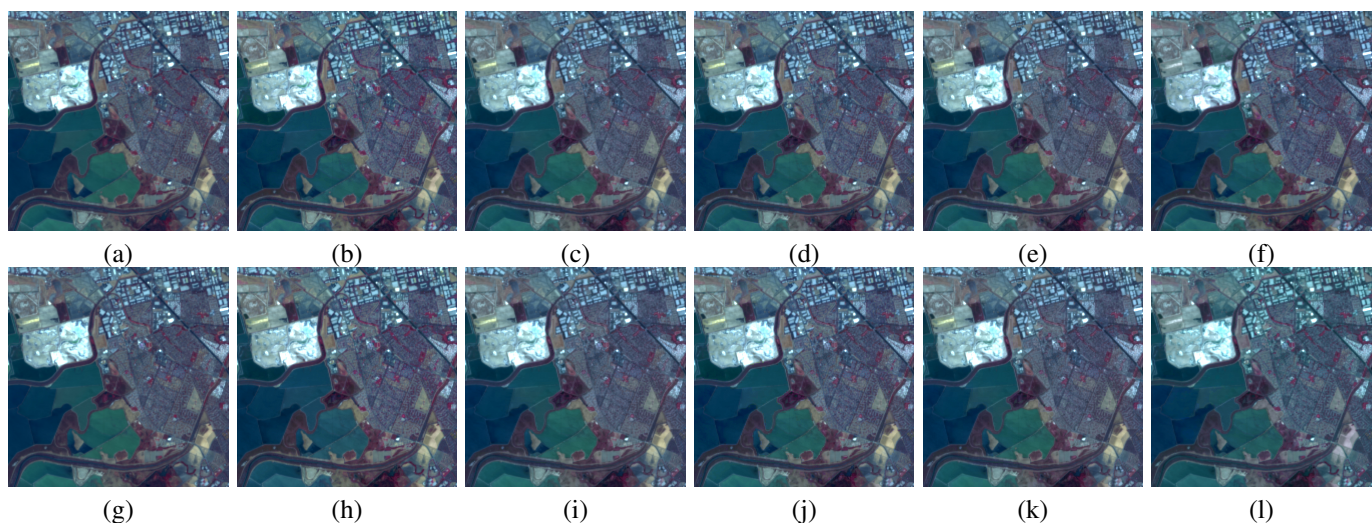
Fig. 15. False color images of the best (top row) and worst (bottom row) super-resolution hyperspectral images estimated from the San Francisco dataset using: (b,h) HSR-GDL-EIA, (c,i) HSR-LDL-EIA(I)-SW, (d,j) HSR-LDL-EIA(I)-BPT, (e,k) HSR-LDL-EIA(II)-SW, (f,l) HSR-LDL-EIA(II)-BPT (a,g) Super-resolution reference image.
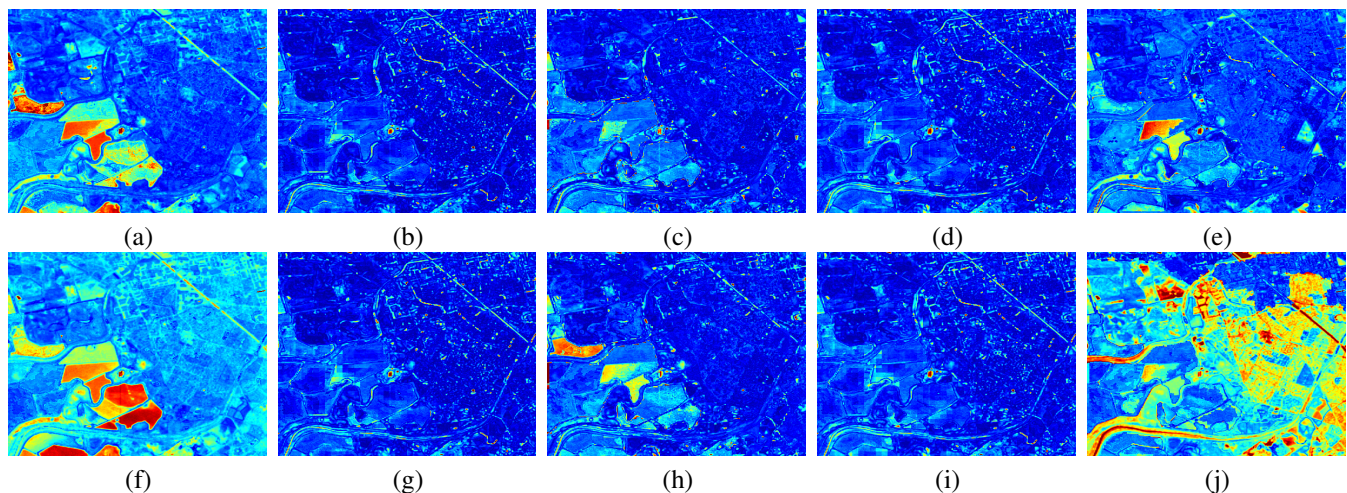

Fig. 16. False color images of the angular errors between the actual super-resolution image and the best (top row) and worst (bottom row) super-resolution hyperspectral images estimated from the San Francisco dataset using: (a,f) HSR-GDL-EIA, (b,g) HSR-LDL-EIA(I)-SW, (c,h) HSR-LDL-EIA(I)-BPT, (d,i) HSR-LDL-EIA(II)-SW, (e,j) HSR-LDL-EIA(II)-BPT. The color scale goes from 0 degrees error (dark blue) to 10 degrees error (dark red).

[2] J. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 2, pp. 6–36, June 2013.

[3] J. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, April 2012.

[4] R. Zurita-Milla, J. Clevers, and M. Schaepman, "Unmixing-based landsat tm and meris fr data fusion," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 3, pp. 453–457, July 2008.

[5] R. Kawakami, J. Wright, Y. Tai, Y. Matsushita, M. Ben-Ezra, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2011, pp. 2329–2336.

[6] A. Charles, B. Olshausen, and C. Rozell, "Learning sparse codes for hyperspectral imagery," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 963–978, 2011.

[7] B. Huang, H. Song, H. Cui, J. Peng, and Z. Xu, "Spatial and spectral image fusion using sparse matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 3, pp. 1693–1704, March 2014.

[8] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 528–537, Feb 2012.

[9] R. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[10] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2014–2039, 2011.

[11] Y. Zhang, S. De Backer, and P. Scheunders, "Noise-Resistant Wavelet-Based Bayesian Fusion of Multispectral and Hyperspectral Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3834–3843, Nov. 2009.

[12] W. Zhang and W.-K. Cham, "Single-image refocusing and defocusing." *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 873–82, Feb. 2012.

[13] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Bayesian fusion of hyperspectral and multispectral images," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp.
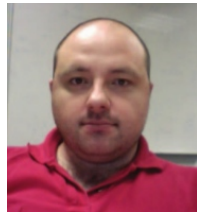
3176–3180.

[14] C. Wang, L.-F. Sun, Z.-Y. Chen, J.-W. Zhang, and S.-Q. Yang, "Multi-scale blind motion deblurring using local minimum," *Inverse Problems*, vol. 26, no. 1, p. 015003, Jan. 2010.

[15] N. Yokoya, N. Mayumi, and A. Iwasaki, "Cross-calibration for data fusion of eo-1/hyperion and terra/aster," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 419–426, April 2013.

[16] M. Simoes, J. Bioucas-Dias, L. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 53, no. 6, pp. 3373–3388, June 2015.

[17] P. Salembier and L. Garrido, "Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval," *IEEE Transactions on Image Processing*, vol. 9, no. 4, pp. 561–576, Apr 2000.

[18] S. Valero, P. Salembier, and J. Chanussot, "Hyperspectral image representation and processing with binary partition trees," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1430–1443, April 2013.

[19] M. Veganzones, G. Tochon, M. Dalla-Mura, A. Plaza, and J. Chanussot, "Hyperspectral image segmentation using a new spectral unmixing-based binary partition tree representation," *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3574–3589, Aug 2014.

[20] W.-K. Ma, J. Bioucas-Dias, T.-H. Chan, N. Gillis, P. Gader, A. Plaza, A. Ambikapathi, and C.-Y. Chi, "A signal processing perspective on hyperspectral unmixing: Insights from remote sensing," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 67–81, Jan 2014.

[21] A. Bruckstein, M. Elad, and M. Zibulevsky, "On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4813–4820, 2008.

[22] J. Nascimento and J. Bioucas-Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, April 2005.

[23] S. Valero, P. Salembier, and J. Chanussot, "Comparison of merging orders and pruning strategies for binary partition tree in hyperspectral data," in *17th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2010, pp. 2565–2568.

[24] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of Satellite Images of Different Spatial Resolutions : Assessing the Quality of Resulting Images," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 6, pp. 691–699, 1997.

[25] E. Middleton, S. Ungar, D. Mandl, L. Ong, S. Frye, P. Campbell, D. Landis, J. Young, and N. Pollack, "The earth observing one (eo-1) satellite mission: Over a decade in space," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 243–256, April 2013.

[26] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 8, pp. 2435–2445, Aug 2008.

[27] M. Winter, "N-findr: an algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Proceedings of SPIE*, vol. 3753, 1999, pp. 266–275.

[28] J. M. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. IEEE, 2010, pp. 1–4.

[29] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 681–695, 2011.

[30] H. Kramer, *Observation of the Earth and Its Environment: Survey of Missions and Sensors*, 4th ed. Springer Science & Business Media, 2002, 00210.

[31] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?" in *Third conference on Fusion of Earth data: merging point measurements, raster maps and remotely sensed images, SEE/URISCA*, Sophia Antipolis, France, 2000, pp. 99–103.

[32] Z. Wang and A. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84, 2002.

Miguel A. Veganzones (M'12) received the M.Sc. and Ph.D. degrees in computer science from the Basque Country University (UPV/EHU), Donostia, Spain, in 2005 and 2012 respectively. Since 2012, he is a posdoctoral researcher at the Images-Signal department in the GIPSA-lab, Grenoble, France. His current main research activities focus in the analysis of hyperspectral images by means of computational intelligence, statistical techniques and tensor data analysis. He is a Reviewer of IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND REMOTE SENSING and the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING.

Miguel Simões was born in Lisbon, Portugal, in 1987. He received the M.Sc. degree in electrical and computer engineering from the Instituto Superior Técnico, University of Lisbon, Lisbon, in 2010. He is currently working toward the joint Ph.D. degree in electrical and computer engineering, and signal and image processing at the Instituto de Telecomunicações, Instituto Superior Técnico, University of Lisbon, Lisbon, and at the Grenoble Images Parole Signal Automatique (GIPSA-lab), University of Grenoble, Grenoble, France, respectively. Previously, he has worked as an Information Technology (IT) consultant in the field of telecommunications. His main areas of research interest are image processing, optimization, and remote sensing.

Giorgio Licciardi received the M.S. degree in telecommunication engineering and the Ph.D. degree in geoinformation from the Tor Vergata University, Rome, Italy, in 2005 and 2010, respectively. In 2010 he joined the Laboratoire Grenoblois de lImage, de la Parole, du Signal et de lAutomatique (GIPSA-Lab) as a Postdoctoral Fellow. His main research is focused on hyperspectral image processing, including feature extraction techniques, spectral unmixing, super-resolution and pansharpening. From 2014 he is Ingnieur de recherch at the Grenoble Institute of Technology. He is also a European Space Agency Category-1 Principal Investigator for Earth observation data. Dr. Licciardi serves as a Referee for several scientific journals such as the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS and the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

Naoto Yokoya (S'10-M'13) received the M.Sc. and Ph.D. degrees in aerospace engineering from the University of Tokyo, Tokyo, Japan, in 2010 and 2013, respectively. He was a Research Fellow with the Japan Society for the Promotion of Science from 2012 to 2013. He is currently an Assistant Professor with the University of Tokyo. His current research interests include image analysis and data fusion in remote sensing.

José Bioucas-Dias (SM'15) received the EE, MSc, PhD, and "Agregado" degrees from Instituto Superior Técnico (IST), Technical University of Lisbon (TULisbon, now University of Lisbon), Portugal, in 1985, 1991, 1995, and 2007, respectively, all in electrical and computer engineering.

Since 1995, he has been with the Department of Electrical and Computer Engineering, IST, where he was an Assistant Professor from 1995 to 2007 and an Associate Professor since 2007. Since 1993, he is also a Senior Researcher with the Pattern and Image Analysis group of the Instituto de Telecomunicações, which is a private non-profit research institution. His research interests include inverse problems, signal and image processing, pattern recognition, optimization, and remote sensing. Dr. Bioucas-Dias has authored or co-authored more than 250 scientific publications including more than 77 journal papers (64 of which published in IEEE journals) and 180 peer-reviewed international conference papers and book chapters.

Dr. Bioucas-Dias was an Associate Editor for the IEEE Transactions on Circuits and Systems (1997-2000) and IEEE Transactions on Image Processing and he is an an Associate Editor for the IEEE Transactions on Geoscience and Remote Sensing. He was a Guest Editor of IEEE Transactions on Geoscience and Remote Sensing for the *Special Issue on Spectral Unmixing of Remotely Sensed Data*, of IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing for the *Special Issue on Hyperspectral Image and Signal Processing*, of IEEE Signal Processing Magazine for the *Special Issue on Signal and Image Processing in Hyperspectral Remote Sensing*, of IEEE Journal of Selected Topics in Signal processing for the *Advances in Hyperspectral Data Processing and Analysis*, and of IEEE Geoscience and Remote Sensing Magazine for the *Special Issue on Advances in Machine Learning for Remote Sensing and Geosciences* He was the General Co-Chair of the 3rd IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS'2011) and has been a member of program/technical committees of several international conferences.

Jocelyn Chanussot (M'04-SM'04-F'12) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995, and the Ph.D. degree from Savoie University, Annecy, France, in 1998. In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA - French National Defense Department). Since 1999, he has been with Grenoble INP, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and is currently a Professor of signal and image processing. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing. He is a member of the Institut Universitaire de France (2012-2017). Since 2013, he is an Adjunct Professor of the University of Iceland. Dr. Chanussot is the founding President of IEEE Geoscience and Remote Sensing French chapter (2007-2010) which received the 2010 IEEE GRS-S Chapter Excellence Award. He was the co-recipient of the NORSIG 2006 Best Student Paper Award, the IEEE GRSS 2011 Symposium Best Paper Award, the IEEE GRSS 2012 Transactions Prize Paper Award and the IEEE GRSS 2013 Highest Impact Paper Award. He was a member of the IEEE Geoscience and Remote Sensing Society AdCom (2009-2010), in charge of membership development. He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Chair (2009-2011) and Cochair of the GRS Data Fusion Technical Committee (2005-2008). He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006-2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing, (2009). He was an Associate Editor for the IEEE Geoscience and Remote Sensing Letters (2005-2007) and for Pattern Recognition (2006-2008). Since 2007, he is an Associate Editor for the IEEE Transactions on Geoscience and Remote Sensing. Since 2011, he is the Editor-in-Chief of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. In 2013, he was a Guest Editor for the Proceedings of the IEEE and in 2014 a Guest Editor for the IEEE Signal Processing Magazine. He is a Fellow of the IEEE and a member of the Institut Universitaire de France (2012-2017).