Open access • Journal Article • DOI:10.1109/JSTARS.2016.2528339

Fusion of Hyperspectral and Multispectral Images Using Spectral Unmixing and Sparse Coding — Source link

Zahra Hashemi Nezhad, Azam Karami, Rob Heylen, Paul Scheunders

Institutions: Shahid Bahonar University of Kerman, University of Antwerp

Published on: 03 Mar 2016 - IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (IEEE)

Topics: Hyperspectral imaging, Multispectral image and Image resolution

Fusion of hyperspectral and multispectral images using spectral unmixing and sparse coding

FUSION OF HYPERSPECTRAL AND MULTISPECTRAL IMAGES USING SPECTRAL UNMIXING AND SPARSE CODING

Zahra Hashemi Nezhad[a], *Azam Karami*[b,c], *Rob Heylen*[c], *Paul Scheunders*[c]

[a]Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran
[b] Faculty of Physics, Shahid Bahonar University of Kerman, Kerman, Iran
[c] Visionlab, University of Antwerp, Belgium

ABSTRACT

Unlike multispectral (MSI) and panchromatic (PAN) images, generally the spatial resolution of hyperspectral images (HSI) is limited, due to sensor limitations. In many applications, HSI with a high spectral as well as spatial resolution are required. In this paper, a new method for spatial resolution enhancement of a HSI using spectral unmixing and sparse coding (SUSC) is introduced. The proposed method fuses high spectral resolution features from the HSI with high spatial resolution features from a MSI of the same scene. Endmembers are extracted from the HSI by spectral unmixing and the exact location of the endmembers is obtained from the MSI. This fusion process by using spectral unmixing is formulated as an ill-posed inverse problem which requires a regularization term in order to convert it into a well-posed inverse problem. As a regularizer, we employ sparse coding, for which a dictionary is constructed using high spatial resolution MSI or PAN images from unrelated scenes. The proposed algorithm is applied to real Hyperion and ROSIS datasets. Compared with other state-of-the-art algorithms based on pansharpening, spectral unmixing and sparse coding methods, the proposed method is shown to significantly increase the spatial resolution while perserving the spectral content of the HSI.

Index Terms— Fusion, Sparse coding, Spectral unmixing, Hyperspectral images, Multispectral images

1. INTRODUCTION

Remote sensing images have been widely used in different practical applications such as earth surface monitoring, agriculture, forest monitoring, environmental studies and military applications [1]. The main types of remote sensing images are panchromatic, multispectral and hyperspectral images. PAN images have a high spatial resolution and spatial structures are well defined, but they are limited to one grayscale image band. MSI have lower spatial resolution than PAN images and contain a limited number of spectral bands. HSI usually have lower spatial resolution than MSI and PAN images but have a high spectral resolution [2, 3]. HSI have been used in many different practical applications. In various applications, HSI with high spectral and spatial resolutions are required [2].

In recent years, many techniques have been proposed to enhance the spatial resolution of HSI or MSI. The first type of methods are pansharpening methods and are based on the fusion of MSI (HSI) and PAN images. Two popular groups of pansharpening methods are component substitution (CS) and multiresolution analysis (MRA) [4]. The CS family includes many popular pansharpening approaches such as principal component analysis [5] and Gramm-Schmidt orthogonalization [6, 7]. The MRA approach is based on the injection of spatial details of the PAN image into the MSI (HSI), which are obtained through a multiscale decomposition (such as decimated wavelet transform [8], Laplacian pyramid [9], curvelets [10, 11] etc.). In general, pansharpening methods are not suited for the fusion of HSI and PAN images or HSI and MSI, because the spectral range of a HSI is not covered by the PAN image or the MSI, which causes spectral distortion. In addition, these methods have high computational load because of the high dimensionality of the HSI [12].

Recently, new techniques have been proposed specifically for spatial resolution enhancement of HSI based on the fusion of HSI and MSI. Some of these methods are based on sparse coding [2, 13, 14]. In these methods, the initial dataset is spatially patched into sub images and sparse coding is applied to each sub image, by constructing a dictionary. However, in these methods the dictionary doesn't consider the spectral correlation of the HSI and spectral distortion is introduced in the reconstructed images.

In order to overcome this limitation, spectral unmixing methods can be used [15–18]. In the linear spectral unmixing model (LMM), mixed pixel spectra are decomposed into endmembers and abundance fractions [19]. Using this model, one can assume that spectral and spatial features are represented by endmembers and abundance fractions respectively. However, unmixing does not determine the exact location of the endmembers. For this, an image with a higher spatial resolution than the HSI can be used. If a MSI of the same scene is available, it should contain the same endmembers. Several HSI resolution enhancement methods are based on

this. In [18], high resolution HSI (HRHSI) are obtained using spectral features of the HSI represented by endmembers, and spatial features of the MSI represented by abundance fractions. However, in these methods, the number of endmembers is limited to be smaller than the number of MSI bands. Therefore in [20], local spectral unmixing (LSU) is used, where the HSI and MSI are partitioned into patches and the endmembers are independently extracted from each patch. Another method for the spatial resolution enhancement of HSI is Coupled Nonnegative Matrix Factorization (CNMF) [21], in which both HSI and MSI are alternately unmixed by NMF [22] and, taking into account the sensor observation models, the HRHSI image is produced by the endmember spectra of the HSI and the abundance fractions of the MSI. Generally, the spectral distortion in methods based on LSU is low but a lower spatial resolution is obtained in comparison to sparse coding methods.

In this paper, a new method for spatial resolution enhancement of HSI is proposed, based on the fusion of HSI and MSI using the LMM. In the proposed method, the HRHSI reconstruction problem is formulated as a linear inverse problem (LIP). A LIP is generally ill-posed and does not have a unique solution. A regularization term needs to be included to convert it into a well-posed inverse problem. In the proposed method, the regularization term is constructed based on sparse coding, for which dictionary is constructed by several high spatial resolution MSI or PAN images which are unrelated to the HSI. In this way, a method is obtained that simultaneously makes use of the LMM to avoid spectral distortion and sparse coding to optimize the spatial resolution improvement.

The proposed spectral unmixing and sparse coding algorithm (SUSC) is applied to real datasets and compared with state-of-the-art algorithms using pansharpening [6, 7, 23], sparse coding [13], LSU [20] and CNMF [21]. The results obtained by SUSC are superior to these methods from the state of the art. The rest of the paper is organized as follows. In Section 2, the proposed method is described. Section 3 presents the experimental results and Section 4 concludes the paper.

2. PROPOSED METHOD

We assume that a HSI and a MSI or Pan image of the same scene are available. As a first step, before the introduction of the proposed method, pre-processing is performed to reduce the noise in the HSI. For this, we apply a recently proposed method, presented in [24].

Then, LMM is applied. In this model, the endmembers are extracted using the spectral properties of the low resolution HSI (LRHSI, which is the observed HSI). After that the initial abundance fractions are calculated using the spatial properties from the MSI. From this, an initial estimate of the HRHSI is obtained. The proposed fusion method then iteratively updates the abundance fractions. At the end, the spectrum of each HRHSI pixel is reconstructed based on the LRHSI endmembers.

The fusion process of the HSI and MSI is formulated as an ill-posed inverse problem. A regularization term is used to convert it into a well-posed inverse problem. The regularization term is constructed based on sparse coding. We construct a proper dictionary with several high spatial MSI or PAN images from unrelated scenes. Based on this dictionary and the initial HRHSI obtained from LMM, the sparse coding is estimated. By using the sparse coding as the regularization term, the abundance fractions are calculated by solving the well-posed inverse problem. The final HRHSI is obtained from the obtained the abundance fractions and the endmembers. In the following, we will explain the different steps of the procedure in more detail.

2.1. OBSERVATION MODELS

Let us first define the observation models. In general, HSI and MSI have three dimensions, two spatial dimensions and one spectral. For notational convenience, these images are converted to two dimensions [12]. If we define Z as the HRHSI (desired image), Y_h as the low spatial resolution HSI and Y_m as the MSI from the same scene, their relationship can be expressed as [21]:

$$Y_h = ZBM + N_h; Z \in \mathbb{R}^{L_h \times n_m}, B \in \mathbb{R}^{n_m \times n_m}, \quad (1)$$
$$M \in \mathbb{R}^{n_m \times n_h}, Y_h \in \mathbb{R}^{L_h \times n_h}, N_H \in \mathbb{R}^{L_h \times n_h}$$
$$Y_m = RZ + N_m; Y_m \in \mathbb{R}^{L_m \times n_m}, R \in \mathbb{R}^{L_m \times L_h}, \quad (2)$$
$$N_m \in \mathbb{R}^{L_m \times n_m}, L_m \ll L_h, n_h \ll n_m$$

Eq.2 models the relationship between the HRHSI Z (with L_h bands and n_m pixels) and the LRHSI Y_h (size $L_h \times n_h$). B is a spatial blurring matrix representing the hyperspectral sensors point spread function. Matrix M accounts for a uniform subsampling of the HRHSI. Eq.3 models the relationship between Z and the high spatial resolution MSI Y_m (size $L_m \times n_m$). R holds in its rows the spectral responses of the multispectral instrument. In practice, the information that is available about the spatial and spectral responses is often scarce or somewhat inaccurate. For example, there may be discrepancies between the real spatial and spectral responses and the data supplied by the manufacturers [12]. Therefore, in the proposed method the matrices B and R are estimated from the observed images using the method presented in [12]. N_h, N_m are Gaussian noises with zero mean and variance δ_h^2, δ_m^2 respectively.

2.2. SPECTRAL UNMIXING MODEL AND DIMENSION REDUCTION

The first step is the use of the LMM:

$$Z = EX + N \quad (3)$$

where $E \in R^{L_h \times P}$ (P represents the number of endmembers) is the endmember signature matrix that can be e.g. extracted by Vertex Component Analysis (VCA) [25] or N-FINDR [26], $X \in \mathbb{R}^{P \times n_m}$ is the abundance fraction matrix and $N \in \mathbb{R}^{L_h \times n_m}$ represents the noise matrix. By substituting Eq.3 into the observation models defined by Eq.2 and Eq.3, we have approximately:

$$Y_h \approx EX_h, X_h = XBM \quad (4)$$

$$Y_m \approx E_m X, E_m = RE \quad (5)$$

According to Eq.3, the construction of Z requires the endmembers and abundance fractions. From Eq.4, we deduce that the HRHSI has the same endmembers as the HSI, and the abundance fractions of the HSI are obtained from the spatially degraded abundance fractions of the HRHSI. Therefore, the required endmembers can be extracted from the original HSI. A superresolution method then only requires the determination of the abundance fractions. An initial estimate of these fractions is obtained from the MSI by using Eq.5. In this work, we use the variable splitting and augmented (SUNSAL) algorithm [27] to do so. That the endmembers can be directly obtained from the HSI, not only means a reduction in the computational load, but the final HRHSI is also expected to have low spectral distortion.

2.3. OPTIMIZATION PROBLEM

The fusion of the the HSI and the MSI is formulated as a LIP:

$$\arg\min_X \frac{1}{2} \|Y_h - EXBM\|_F^2 + \frac{\lambda_m}{2} \|Y_m - REX\|_F^2 + \frac{\lambda}{2}\varphi(x) \quad (6)$$

where $\varphi(x)$ is a regularization term. The inverse problem of Eq.6 can be ill-posed or well-posed, depending on the reduced dimension of the images and the number of spectral bands [28]. If the product of the matrices R and E has a full column rank, the LIP problem is well-posed. However, if there are fewer bands in the MSI than the number of endmembers, the matrix (RE) cannot have full column rank, which turns the LIP problem in an ill-posed one. In this paper Eq.6 is assumed to be an ill-posed LIP. Therefore, a regularization term is required. The first two terms are the fidelity terms, describing that the estimated image is able to explain the observed data according to the models defined in Eqs.4 & 5. The last term is the regularizer. The parameters λ_m and λ control the relative importance of the various terms [12].

2.4. SPARSE CODING

Although image content can vary greatly from image to image, the micro-structures of images can be represented by a small number of structural primitives (e.g., edges, line segments and other elementary features). These micro-structures are the same for all images [2]. Sparse coding relies on this observation by constructing a dictionary of such primitives from a number of images and uses this dictionary to reconstruct a specific image from the smallest number of dictionary atoms.

In the proposed method, we use the concept of sparse coding for the regularization. A dictionary (ϕ) is constructed with micro-structures of unrelated MSI or PAN images of high spatial resolution. Then, a sparse code (S) for the HRHSI is calculated from the smallest number of dictionary atoms, where each pixel is represented by a linear combination of a few atoms. The optimization problem becomes:

$$\arg\min_{X,S} \frac{1}{2} \|Y_h - EXBM\|_F^2 + \frac{\lambda_m}{2} \|Y_m - REX\|_F^2 + \frac{\lambda}{2} \|\phi S - EX\|_F^2 \quad (7)$$

Usually, to generate the dictionary, the images are divided into several patches. Also, in order to construct the image with this dictionary, it is divided into several patches which are separately reconstructed by the dictionary atoms. In recent literature, overcomplete dictionaries with dimensions larger than the image dimensions have been considered. Different methods such as Online Dictionary Learning (ODL) [29] and Kmeans- Singular Value Decomposition (K-SVD) [30] have been developed for constructing an overcomplete dictionary. There is no unique rule to select the dictionary size and the number of atoms. Generally, the smaller the patches, the more atoms can be determined. However, too small patches are not efficient to properly capture the textures, edges, etc. With larger patch sizes, a larger number of atoms is required to guarantee the overcompleteness (which requires a larger computational cost). In general, the size of the patches is empirically selected [28]. In the proposed method, the dictionary was produced by the K-SVD algorithm, trained on a dataset of 10000 patches with size 8×8. Those patches are taken from an arbitrary set of natural images (unrelated to the test images). In fact, the resolution of these images should be sufficiently high to capture the image details. As an example, MSI from the Advanced Land Imager (ALI) sensor are shown in Fig.1. The constructed dictionary by using the K-SVD algorithm is shown in Fig.2.

After constructing the proper dictionary, the HRHSI can be obtained from a linear combination of a small number of atoms from the dictionary. The matrix with the obtained coefficients is called the sparse code [31]. It is obtained by:

$$\hat{S} \triangleq \arg\min_S \frac{1}{2} \|Z_{initial} - \phi S\| \quad \text{Subject to } \|S\|_0 \leq K \quad (8)$$

where K is the number of atoms needed for reconstructing $Z_{initial}$ patches. $Z_{initial}$ can be estimated by interpolation [16], but in the proposed method, it is obtained from Eq.3. For introducing sufficient sparsity, K is chosen much smaller than

Fig. 1: MSI used to construct the dictionary.

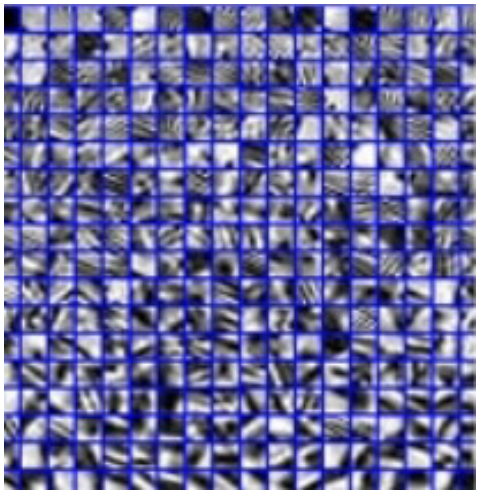


Fig. 2: Dictionary trained from the MSI of Fig. 1.

the number of dictionary atoms. In this paper, the sparse code is estimated by Orthogonal Matching Pursuit (OMP) [32].

With the proposed regularization term, the LIP can be rewritten as:

$$\arg\min_{X,\hat{S}} \frac{1}{2}\|Y_h - EXBM\|_F^2 + \frac{\lambda_m}{2}\|Y_m - REX\|_F^2 + \frac{\lambda}{2}\left\|\phi\hat{S} - EX\right\|_F^2 \quad (9)$$

2.5. SOLVING THE LINEAR INVERSE PROBLEM

The optimization problem Eq.9 can be solved by alternating optimizing w.r.t. the abundance matrix X and the sparse code ($\hat{S}$). The optimization w.r.t. X is achieved by the the Split Augmented Lagrangian Shrinkage Algorithm (SALSA) [33]. In SALSA algorithm an auxiliary variable besides the original optimization variable X is considered. By this algorithm, Eq.9 becomes:

$$\arg\min_{X,\hat{S}} \frac{1}{2}\|Y_h - EV_1M\|_F^2 + \frac{\lambda_m}{2}\|Y_m - REV_2\|_F^2 + \frac{\lambda}{2}\left\|\phi\hat{S} - EV_3\right\|_F^2 \quad (10)$$
$$\text{Subject to } V_1 = XB, V_2 = X, V_3 = X$$

The augmented Lagrangian associated with the optimization of X can be written as [34]:

$$\begin{aligned} L(X, V_1, V_2, V_3, G_1, G_2, G_3) = {} & \frac{1}{2}\|Y_h - EV_1M\|_F^2 + \\ & \frac{\mu}{2}\|XB - V_1 - G_1\|_F^2 + \frac{\lambda_m}{2}\|Y_m - REV_2\|_F^2 \\ & + \frac{\mu}{2}\|X - V_2 - G_2\|_F^2 + \frac{\lambda}{2}\left\|\phi\hat{S} - EV_3\right\|_F^2 + \\ & \frac{\mu}{2}\|X - V_3 - G_3\|_F^2 \end{aligned} \quad (11)$$

where G is the so-called scaled dual variable and μ is a positive constant penalty parameter. The optimal X is evaluated by Eq.11, and the sparse code is estimated by Eq.8. This process is repeated until a stopping criterion, in this case a maximal number of iterations, (max_iteration) is obtained. The pseudo-code of the proposed algorithm is summarized in the following:

In the proposed method, we make two assumptions. The first assumption is that the real datasets contain pure pixels (or endmembers). By this assumption, VCA is used for endmember extraction. If HSI do not contain pure pixels, other methods such as minimum volume constrained non-negative matrix factorization [35], the minimum volume simplex algorithm [36], the convex analysis-based minimum volume enclosing simplex algorithm [37] or simplex identification via split augmented Lagrangian [38]) can be used for virtual endmember extraction. The second assumption is that the number of endmembers equals the number of multispectral bands. If the number of endmembers is larger than the number of MSI bands, the HSI and MSI are patched until this assumption is met (the number of endmembers can be obtained using

Algorithm 1 Pseudo-code of SUSC

Input: Y_h,Y_m, R, B, M, λ_m, λ, μ,max_iteration.
Y_h: LRHSI, Y_m: MSI.
R = spectral response matrix, B = spatial blur matrix, M = down-sampling matrix.
Output: Z (HRHSI).
1- Apply VCA to LRHSI in order to extract E and calculate the initial abundance fractions using SUNSAL.
2- Construct dictionary from high resolution MSI using K-SVD.
3- Extract sparse code matrix using OMP.
For $T = 1, 2, ..., max\_iteration$.
  4- optimize X using SALSA $(*)$
  5- optimize S using Eq.8
END
6- $Z = EX$
END
$(*)$ Optimize X using SALSA (Eq.11)
  A. Initialization: $V_1^0, V_2^0, V_3^0, G_1^0, G_2^0, G_3^0$
  For $K = 0, 1, ...,$ stopping rule do $(\left\|X^{k+1} - X^k\right\| < \varepsilon)$
    B. Optimize w.r.t. X:
      $X^{k+1} = [(V_1^{(K)} + G_1^{(K)})B^T + (V_2^{(K)} + G_2^{(K)}) + (V_3^{(K)} + G_3^{(K)})][BB^T + 2I]^{-1}$
    C. Optimize w.r.t. V_1, V_2, V_3:
      $V_1^{(K+1)}M = [E^TE + \mu I]^{-1}[E^TY_h + \mu(X^{(k+1)}B - G_1^k)]M$
      $V_2^{(K+1)} = [\lambda_m E^T R^T RE + \mu I]^{-1}[\lambda_m E^T R^T Y_m + \mu(X^{(k+1)}B - G_2^k)]$
      $V_3^{(K+1)} = [\lambda_m E^T E + \mu I]^{-1}[\lambda E^T \phi\hat{S} + \mu(X^{(k+1)}B - G_3^k)]$
    D. Optimize w.r.t. G_1, G_2, G_3:
      $G_1^{(K+1)} = G_1^K - (X^{(K+1)}B - V_1^{(K+1)})$
      $G_2^{(K+1)} = G_2^K - (X^{(K+1)} - V_2^{(K+1)})$
      $G_3^{(K+1)} = G_3^K - (X^{(K+1)} - V_3^{(K+1)})$
  END
  E. $Z_{initial} = EX$

HySIME [39]) and the proposed method is applied to each patch separately.

2.6. COMPLEXITY ANALYSIS

The SALSA algorithm has a complexity of $O_1(Pn_{it}n_m log(Pn_m))$, where n_{it} is the number of SALSA iterations, P is the number of endmembers and n_m is the total number of pixels [28, 34]. The computational complexity of the sparse coding is $O_2(Kn_pL_h)$, where K is the number of dictionary atoms and n_p is the patch size. Therefore, the complexity of the proposed algorithm is $N_{th}(O_1(Pn_{it}n_m log(Pn_m)) + O_2(Kn_pL_h))$, where N_{th} is max_iteration.

3. EXPERIMENTS AND RESULTS

3.1. Quality Indices

In order to validate the quality of the obtained HRHSI, four image quality measures have been applied, based on the comparison with a high resolution ground truth hyperspectral image Z_G. The first index is the Peak Signal-to-Noise Ratio (PSNR):

$$PSNR = \frac{\sum_{i=1}^{L_h} 10\log_{10}\left(\frac{Max_i^2}{MSE_i}\right)}{L_h},$$
$$MSE_i = \frac{1}{n_m}\sum_{j=1}^{n_m}\left(Z_{G_{i,j}} - Z_{i,j}\right)^2 \quad (12)$$

where Max_i is the maximum pixel value in the i_{th} band. n_m is the number of pixels and L_h is the number of bands.

The second index is the Spectral Angle Mapper (SAM) that measures the spectral distortion between the reconstructed image and the actual ground truth image:

$$SAM = \frac{1}{n_m}\sum_{j=1}^{n_m}\arccos\left(\frac{Z_{:,j}, Z_{G_{:,j}}^T}{\|Z_{:,j}\|_2\left\|Z_{G_{:,j}}\right\|_2}\right) \quad (13)$$

where $Z_{:,j}$ and $Z_{G_{:,j}}$ are the spectra of pixel j of the estimated and ground truth image respectively.

The third index is the Error Relative Global Dimensional Synthesis index (ERGAS) which is also a measure for the

amount of spectral distortion in the images [40]:

$$ERGAS = 100\frac{1}{d}\sqrt{\frac{1}{L_h}\sum_{i=1}^{L_h}\frac{MSE\left(Z_{i,:}, Z_{G_{i,:}}\right)}{\mu^2_{Z_{G_{i,:}}}}} \quad (14)$$

where d is the ratio between the spatial resolutions of the HSI and MSI. $Z_{i,:}$ and $Z_{G_{i,:}}$ are the i_{th} bands of the reconstructed and ground truth images, respectively. $\mu_{Z_{G_{i,:}}}$ is the mean of $Z_{G_{i,:}}$. For a perfectly reconstructed image, $ERGAS = 0$.

The fourth index is the Cross Correlation (CC), which is a spatial measure for the geometric distortion:

$$CC = \frac{1}{L_h}\sum_{i=1}^{L_h} CCS\left(Z_{G_{i,:}}, Z_{i,:}\right) \quad (15)$$

where CCS is the cross correlation of 2 single-band images, defined as:

$$CCS(Z_{G_{i,:}}, Z_{i,:}) = \frac{\sum_{j=1}^{n_m}(Z_{G_{i,j}} - \mu_{Z_{G_{i,:}}})(Z_{i,j} - \mu_{Z_{i,:}})}{\sqrt{\sum_{j=1}^{n_m}(Z_{G_{i,j}} - \mu_{Z_{G_{i,:}}})^2 (Z_{i,j} - \mu_{Z_{i,:}})^2}} \quad (16)$$

The ideal value of CC is one.

3.2. DATA SET

The proposed method has been applied to two real datasets. The first dataset was acquired by the reflective optics system imaging spectrometer (ROSIS) optical sensor over the urban area of the University of Pavia, Italy[1]. The image size is $610\times 610 \times 103$ with a spatial resolution of $1.3m$. For these HSI, a MSI of the same scene does not exist. Therefore, we generate a MSI of four band by filtering the HSI with the IKONOS-like reflectance spectral responses. With these, R and B is estimated as in [12]. In fact, the IKONOS satellite captures both a panchromatic (0.45-0.90 μm) and four multispectral bands (0.45-0.52, 0.52-0.60, 0.63-0.69 and 0.76-0.90 μm).

On the HSI, the water vapor absorption bands are removed. The obtained HSI is applied as the ground truth image with high spatial and spectral resolutions. For constructing a low spatial resolution HSI (LRHSI), Gaussian blurring (B) (with dimension7 $\times$ 7 and $\sigma = 1.5$) is applied to the ground truth images and the blurred images are down-sampled by a factor of 4 (M). For the simulations, ground truth subimages are selected with size $120 \times 120 \times 93$ leading to LRHSI with size $30 \times 30 \times 93$ and MSI with size $120 \times 120 \times 4$. For the dictionary, twenty PAN QuickBird images[2] with a spatial resolution 0.7m, which are unrelated and do not overlap with the test images have been used. These images are down-sampled by a mean filter with a factor of 2 in order to have the same spatial resolution as the HRHSI. Fig.3 shows for one of the subimages band 10 of the ground truth image and the LRHSI, band 2 of the MSI and the Panchromatic image. SUSC is applied to these images and the reconstructed HRHSI is compared with the ground truth images.

The second dataset contains images taken above Shiraz city in Iran, and was obtained by two instruments, the Hyperion instrument and the Advanced Land Imager (ALI) [3] . Hyperion is a hyperspectral imager with a spatial resolution of $30m$, the entity ID of the Hyperion image is $EO1H1630392004316110PV\_1R1$. It has size $3858 \times 256 \times 242$. The ALI instrument provides MSI and PAN of the same scene at resolutions of 30 and 10 meters, respectively. The MSI are used in our experiment. The entity ID of the ALI data is $EO1A1630392004316110PV\_1GST$. It has size $4241 \times 256 \times 10$.

First the HSI and MSI are geometrically co-registered. Then the water absorption bands (1-7, 58-76, 121-128, 165-180, 221-242) are removed from the HSI and it is denoised using [24]. A LRHSI is constructed from it in the same way as described for the first dataset. Also, since the original HSI and MSI of this dataset have the same resolution, we set $B = I$ and estimate R. In the simulations, ground truth images are selected with size $120 \times 120 \times 170$ and LRHSI with size $30 \times 30 \times 170$ and MSI with size $120 \times 120 \times 9$ are obtained. Fig.4 shows for one of the subimages band 170 of the ground truth image and the LRHSI and band 2 of the MSI. The proposed algorithm is applied to these images and the reconstructed HRHSI are compared with the ground truth images.

In the proposed method, VCA is used for endmember extraction. Since VCA is not robust, we performed ten runs of the algorithm, and report the average of the corresponding results. The number of endmembers is equal to the number of MSI bands. For constructing the dictionary the PAN images are converted to 10000 patches with size of (8×8). Then 40 atoms of the dictionary are used for constructing the HRHSI. For solving the optimization problem, the regularization parameters are selected as: $\lambda_m = 1$ and $\mu = 5 \times 10^{-2}$ for both datasets [12, 28]. In order to select an appropriate value of λ, the performance of the proposed algorithm has been evaluated as a function of λ. Fig.5 and Fig.7 display the results for the first and second dataset respectively. The optimal value of λ is found to be one.

3.3. Comparison with other fusion methods

The proposed method is compared with state-of-the-art algorithms:

1) Pansharpening methods. For the spatial resolution enhancement of HSI, the Gramm-Schmidt (GS) procedure of [6], the adaptive GS method (GSA) of [7], the smooting filter-based intensity modulation method (SFIM) of [23] and

[1] http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes

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

[3] Available at http://earthexplorer.usgs.gov/

a Modulation Transfer Function Generalized Laplacian Pyramid with High-Pass Modulation (MTF_GLP_HPM) from [41] are applied to the first dataset along with the proposed SUSC method. All the methods are implemented in Matlab on a computer with an Intel(R) Core(TM) i5-3210 processor (3.1 GHz), 4GB of memory and a 64-bit Operating System.

Fig. 3 shows the obtained results for one specific subimage. Table 1 displays the quality measures. In Fig.6 the PSNR, ERGAS and CC values are shown in function of the wavelength (0.4 to 0.9 μm). The simulation results show that in component substitution-based fusion techniques (GS, GSA), the spatial resolution of the reconstructed image is low and the spectral distortion is high, because of the spectral mismatch between the PAN image and the HSI spectral range. In the multiresolution analysis approaches (SFIM, MTF_GLP_HPM) the spatial resolution is low but the spectral distortion is lower than with the CS methods. The reconstructed images by the proposed method SUSC are visually very close to the ground truth images. As is shown in Table 1, the spectral distortion is the lowest in the proposed method. The required computing time of the proposed method however is much higher than the Pansharpening methods. The construction of the dictionary and the estimation of the sparse code take a considerable amount of time. Moreover, the method iteratively updates the abundance fractions and the sparse code, which makes the proposed method time consuming.

2) Spectral unmixing and sparse coding methods. The Sparse Coding method (SC) from [13], CNMF from [21] and Local Spectral Unmixing (LSU) from [20] are applied to the second dataset. In CNMF, the maximum number of iterations in the inner and outer loops is selected as 10 and 300, respectively. In LSU, the image is divided into several patches, the proper size for the patches is related to the number of endmembers in each patch. In the experiment, different patch sizes were applied. Using SUSC, the same parameters as in the first dataset are applied.

Quality measures and computing time for the proposed algorithm and the other algorithms are reported in Table 2. Fig. 4 shows results for band 170 of a specific subimage. The simulation results show that sparse coding produces high spatial resolution HRHSI, because of the use of the high spatial resolution dictionary, but spectral distortion occurs. Using, LSU and CNMF, the spectral distortion is lower than for sparse coding, but the spatial resolution is limited. The reconstructed image by the proposed method is visually very close to the ground truth image.

The PSNR, ERGAS and CC in function of the wavelength (0.4 to 2.5 μm) are shown in Figs.8-10 respectively. Fig.11 shows the spectra of pixel (1,100) in the ground truth and reconstructed image; the spectral distortion value is lowest in the proposed method. In the proposed method, spectral unmixing is used to preserve the spectral content of the HSI and sparse coding is used to enhance the spatial resolution of the HSI. Therefore, the reconstructed HRHSI have, simultaneously, a higher spatial and spectral resolutions compared to the other methods.

4. CONCLUSION

In this paper a new method for enhancing the spatial resolution of HSI based on fusion with MSI is proposed. The method combines the spectral mixing model to reduce spectral distortions with sparse coding to inject high spatial information from a dictionary of unrelated high spatial resolution images. The problem is expressed as a linear inverse problem with the sparse coding as regularizer. The inverse problem is solved by iteratively updating the abundance fractions using SALSA and the sparse code using Orthogonal Matching Pursuit. Based on the visual and quantitative results, in the proposed method, the spatial resolution is significantly enhanced and the spectral distortion of the reconstructed image is low compared to state of the art reconstruction techniques based on local unmixing and sparse coding. In the future our aim is to reduce the computational complexity of the proposed algorithm.

ACKNOWLEDGMENT The authors would like to thank Prof. Jos Bioucas-Dias from Instituto de Telecomunicaes and Instituto Superior Tcnico, Universidade de Lisboa for sharing the IKONOS-like reflectance spectral responses used in the experiments. Rob Heylen is a postdoctoral researcher sponsored by the Flemish fund for scientific research (FWO - Vlaanderen).

5. REFERENCES

[1] Y. Gu, Y. Zhang, and J. Zhang, "Integration of spatial-spectral information for resolution enhancement in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 76, no. 5, pp. 1–10, May. 2008.

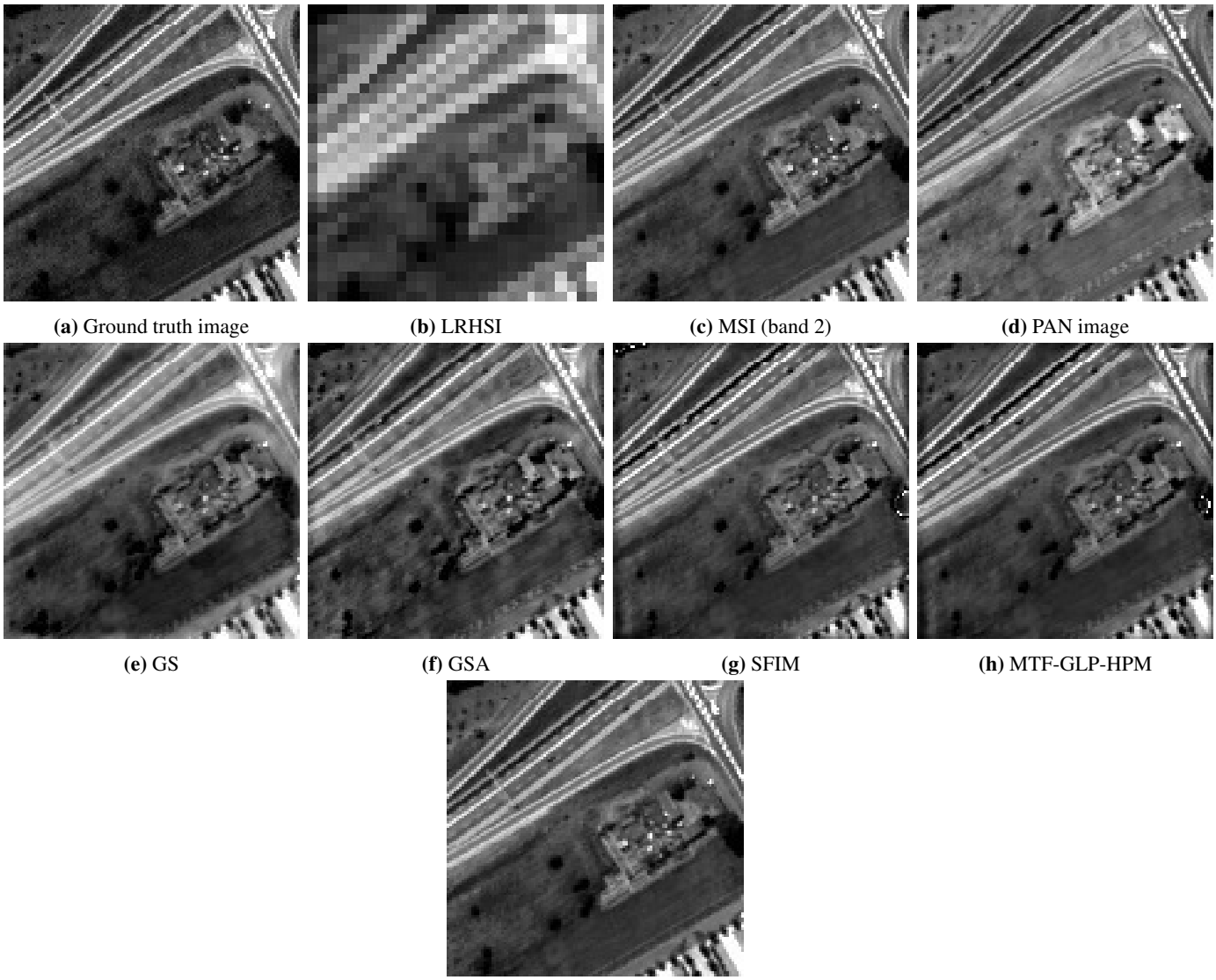
[2] Y. Zhao, J. Yang, Q. Zhang, L. Song, Y. Cheng, and Q. Pan, "Hyperspectral imagery super-resolution by sparse representation and spectral regularization," *EURASIP J. Adv. Signal Process.*, vol. 2011, no. 1, pp. 1–10, Oct. 2011.

[3] T. Akgun, Y. Altunbasak, and R. M. Mersereau, "Super-resolution reconstruction of hyperspectral images," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1860–1875, Nov. 2005.

[4] L. Loncan, L. B. Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, and M. Simes, "Hyperspectral pansharpening: a review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 27–46, 2015.

Table 1: simulation results of the different methods for the ROSIS dataset.

Method	$\mathbf{PSNR}_{dB}$	$\mathbf{SAM}^0$	ERGAS	CC	time
GS [6]	31.7388	4.3550	3.9310	0.9406	1.09
GSA [7]	30.0771	5.7625	4.5955	0.9135	1.22
SFIM [23]	30.0419	4.6117	3.7014	0.9396	0.80
MTF_GLP_HPM [41]	30.4563	4.6388	3.7988	0.9396	0.95
Proposed method (SUSC)	32.3052	3.2560	2.3815	0.9833	551.36

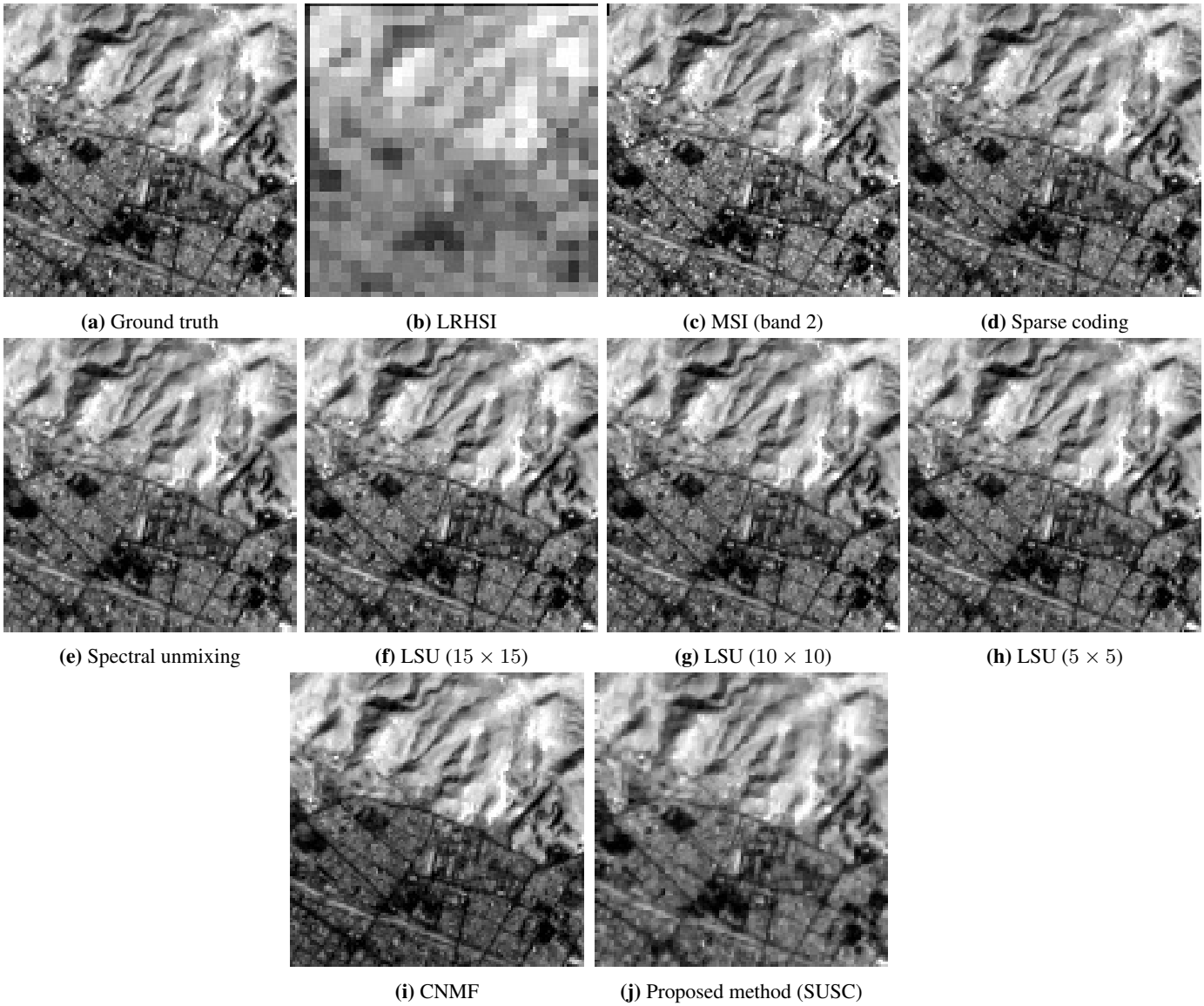


(a) Ground truth image **(b)** LRHSI **(c)** MSI (band 2) **(d)** PAN image

(e) GS **(f)** GSA **(g)** SFIM **(h)** MTF-GLP-HPM

(i) Proposed method (SUSC)

Fig. 3: (a): Band 10 of the ROSIS ground truth image; (b): LRHSI; (c): band 2 of the MSI; (d): PAN image; (e)-(i): Spatial resolution enhancement results of band 10 of the HSI.



(a) Ground truth **(b)** LRHSI **(c)** MSI (band 2) **(d)** Sparse coding

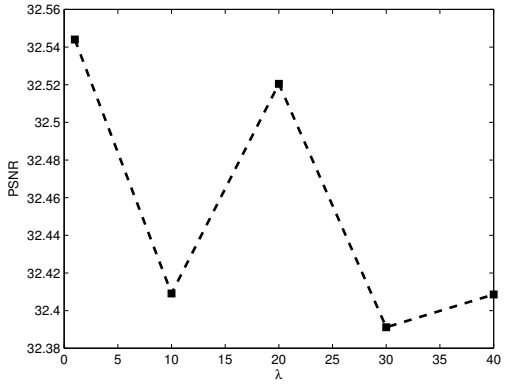
(e) Spectral unmixing **(f)** LSU (15×15) **(g)** LSU (10×10) **(h)** LSU (5×5)

(i) CNMF **(j)** Proposed method (SUSC)

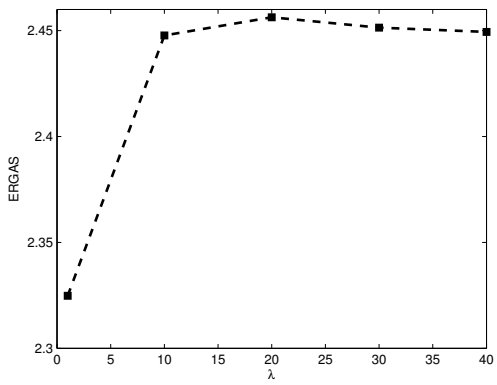
Fig. 4: (a): Band 170 of the Hyperion ground truth image; (b): LRHSI; (c): band 2 of the MSI; (d)-(j): Spatial resolution enhancement results of band 170 of the HSI.

Table 2: simulation results of different methods for the Hyperion dataset.

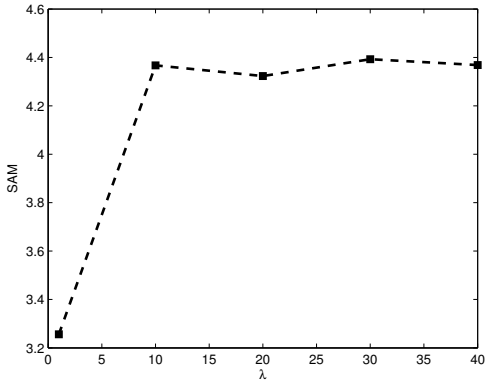
Method	**PSNR**$_{dB}$	**SAM**^{0}	**ERGAS**	**CC**	**Times (in second)**
Sparse Coding [13]	31.4750	7.2975	5.1038	0.9518	143.55
Spectral Unmixing [20]	30.8580	3.4904	1.7093	0.9872	58.82
CNMF [21]	31.8062	2.4170	1.5556	0.9919	61.69
Local Spectral Unmixing (15×15) [20]	32.1964	2.9629	1.4837	0.9876	57.73
Local Spectral Unmixing (10×10) [20]	32.0751	2.9576	1.4996	0.9874	93.35
Local Spectral Unmixing (5×5) [20]	31.9046	2.9539	1.5210	0.9866	65.14
Proposed method (SUSC)	34.5155	2.2376	1.1298	0.9916	617.38



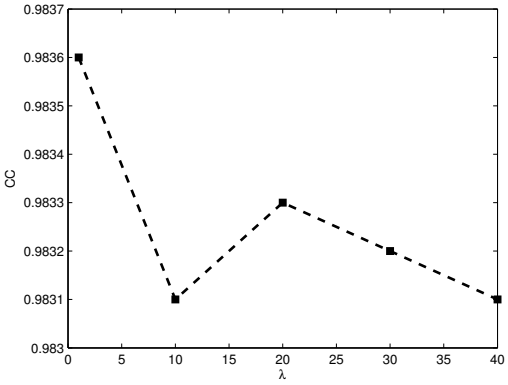
(a) PSNR



(b) ERGAS

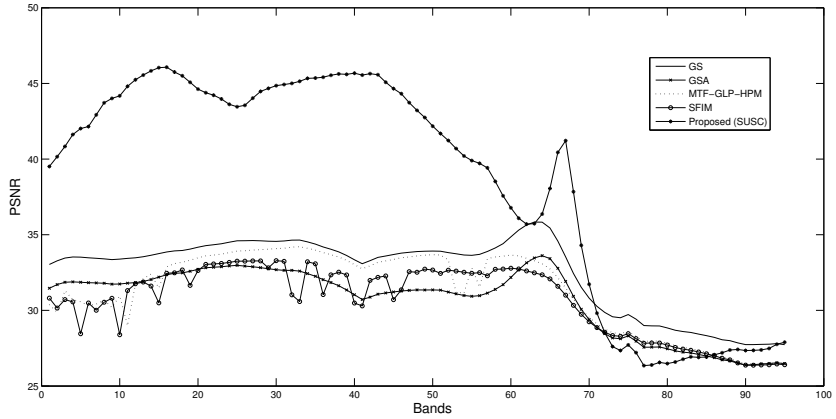


(c) SAM

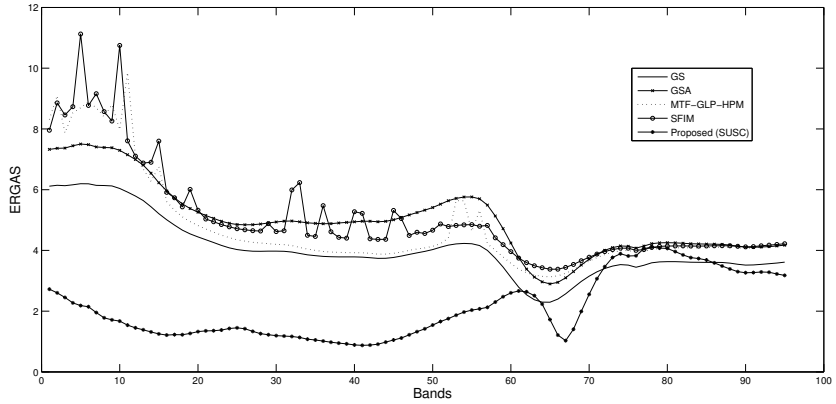


(d) CC

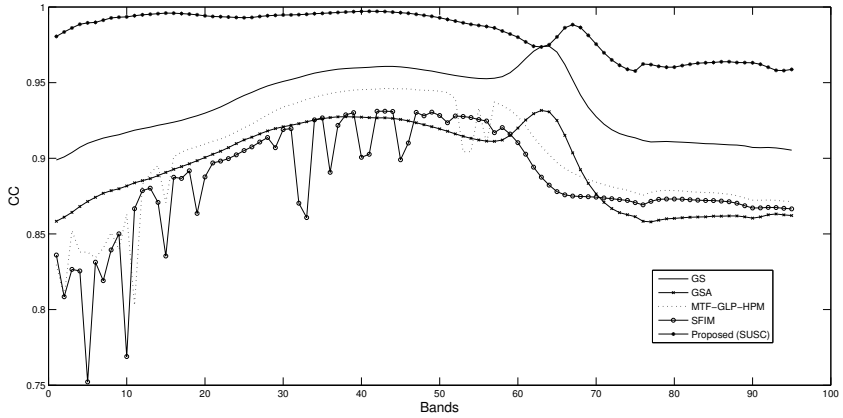
Fig. 5: Performance of the SUSC algorithm versus λ for the ROSIS dataset.



(a) PSNR



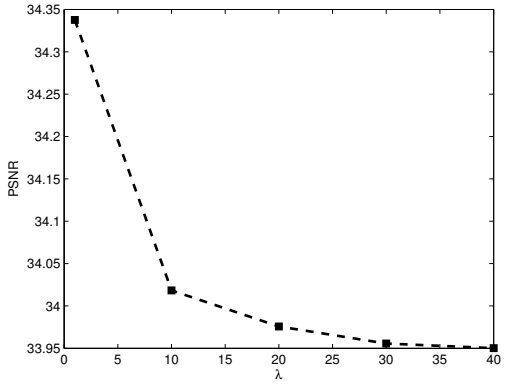
(b) ERGAS



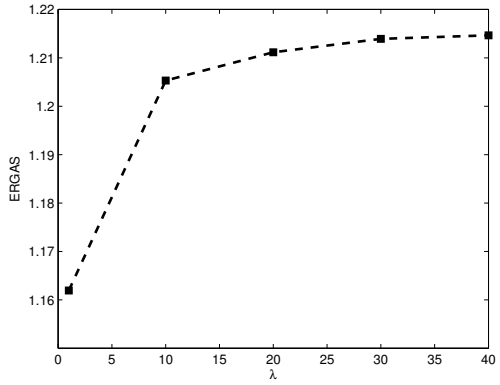
(c) CC

Fig. 6: PSNR, ERGAS and CC in function of wavelength for the ROSIS dataset.

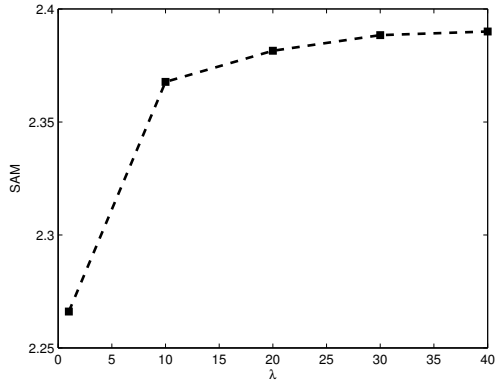
[5] P. Chavez and J. A. Sides, Stuart C andAnderson, "Comparison of three different methods to merge multiresolution and multispectral data- landsat tm and spot panchromatic," *Photogrammetric Engineering and remote sensing*, vol. 57, no. 3, pp. 295–303, Mar. 1991.



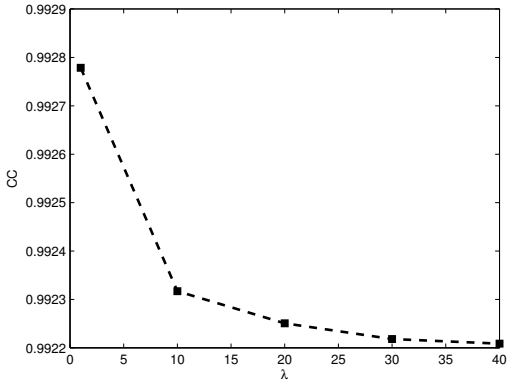
(a) PSNR



(b) ERGAS

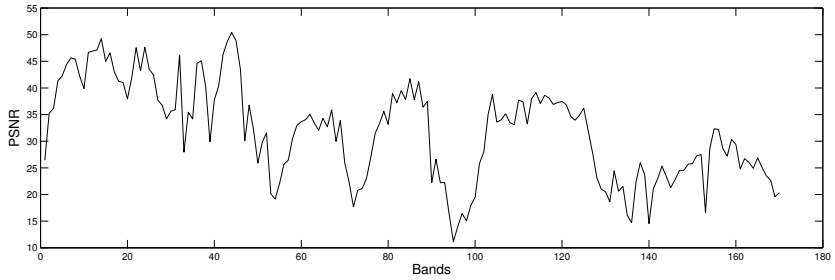


(c) SAM

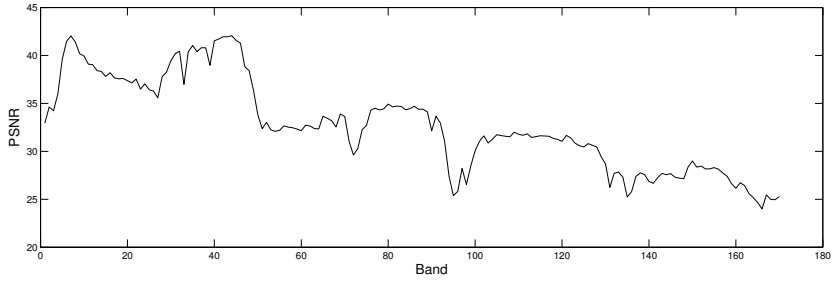


(d) CC

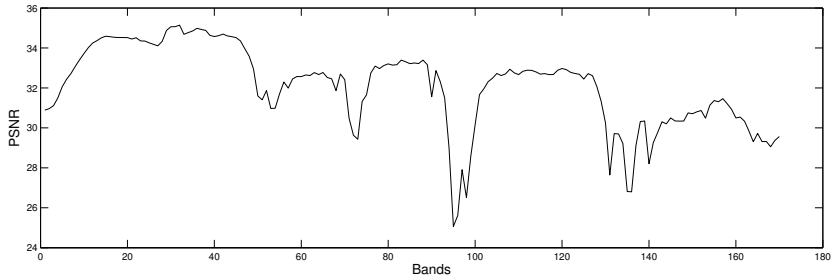
Fig. 7: Performance of the SUSC algorithm versus λ for the Hyperion dataset.



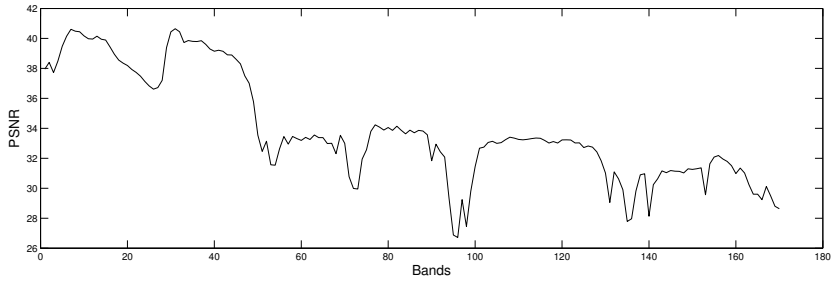
(a) sparse coding



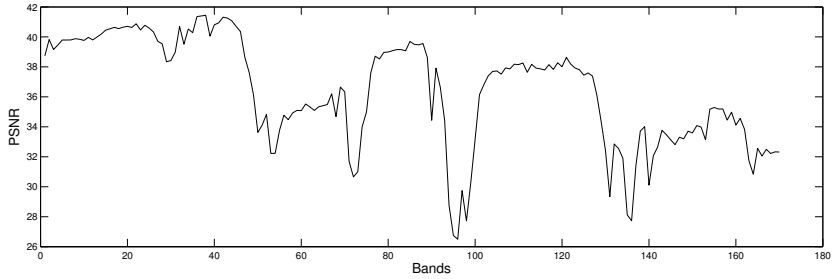
(b) spectral unmixing



(c) CNMF

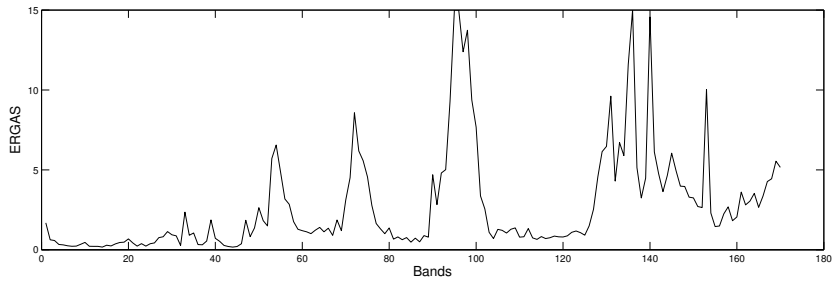


(d) LSU(15×15)

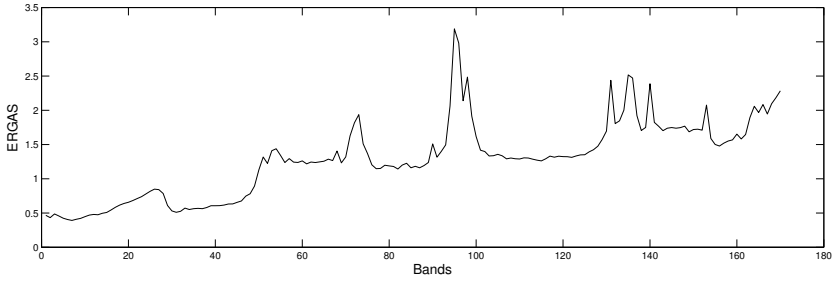


(e) Proposed (SUSC)

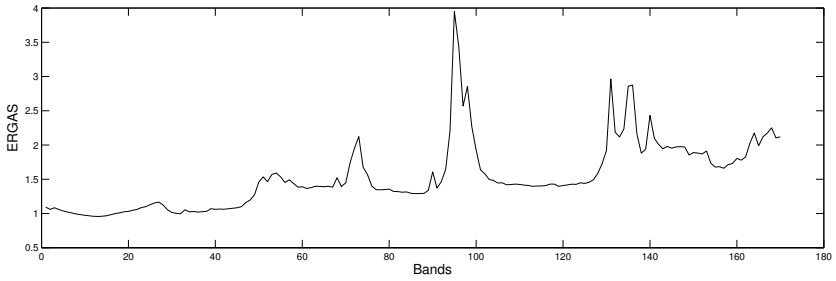
Fig. 8: PSNR in function of wavelength for the Hyperion dataset.



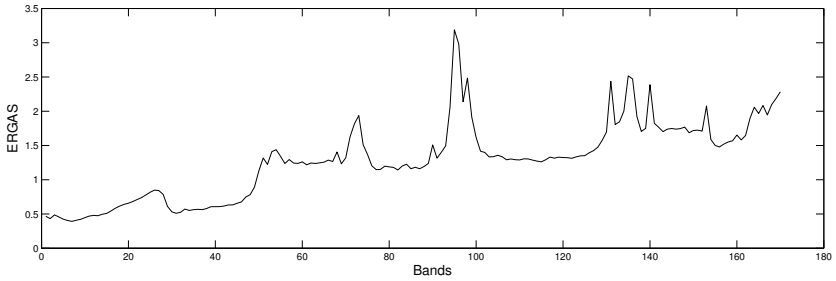
(a) sparse coding



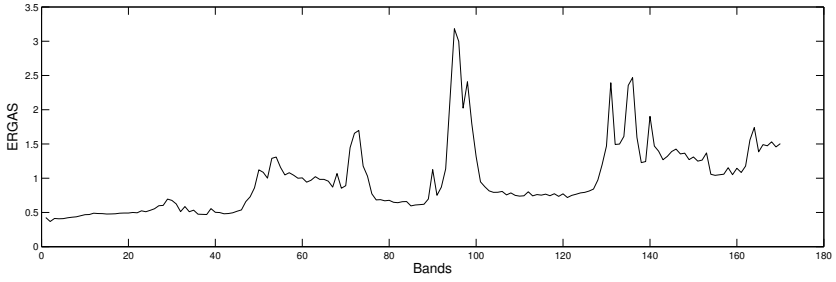
(b) spectral unmixing



(c) CNMF

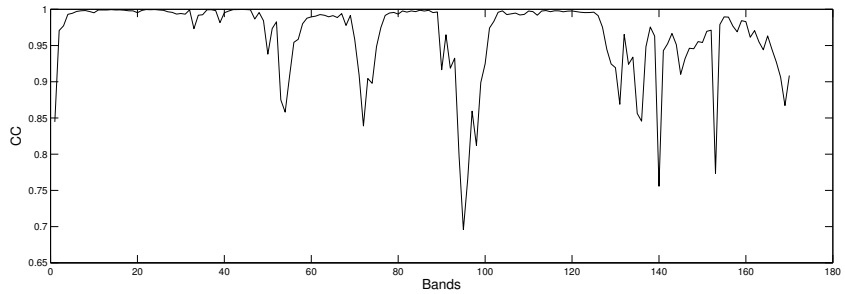


(d) LSU(15×15)

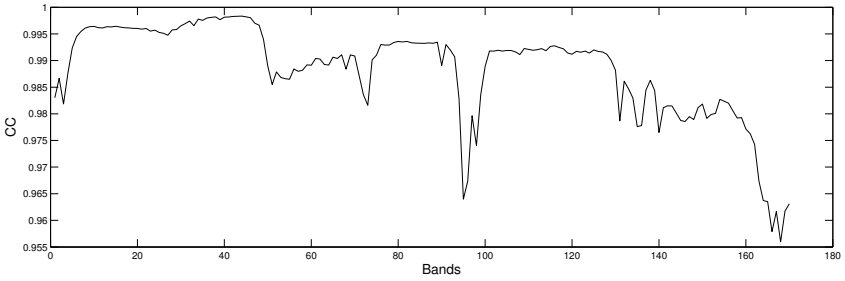


(e) Proposed (SUSC)

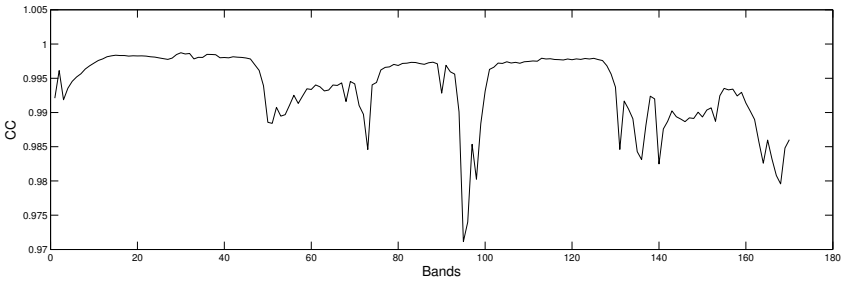
Fig. 9: ERGAS in function of wavelength for the Hyperion dataset.



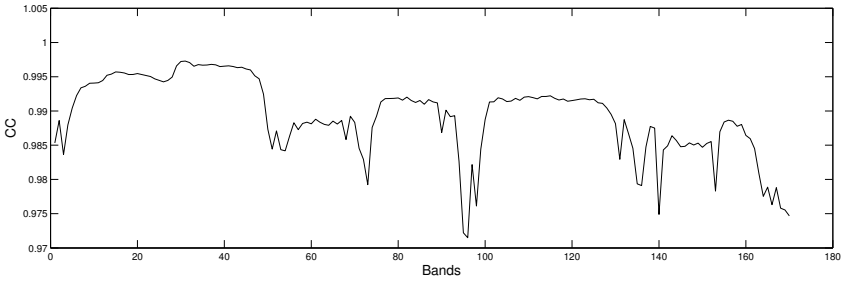
(a) sparse coding



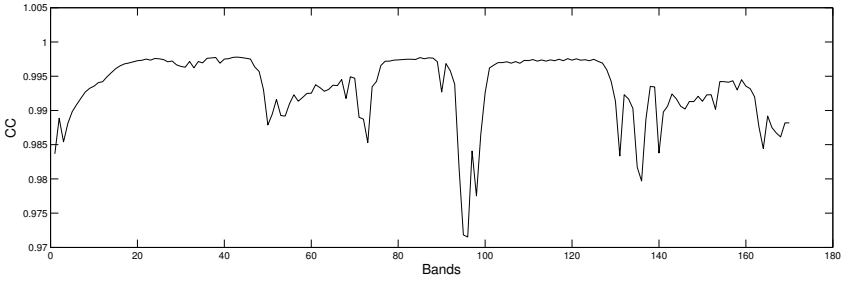
(b) spectral unmixing



(c) CNMF



(d) LSU(15×15)



(e) Proposed (SUSC)

Fig. 10: CC in function of wavelength for the Hyperion dataset.

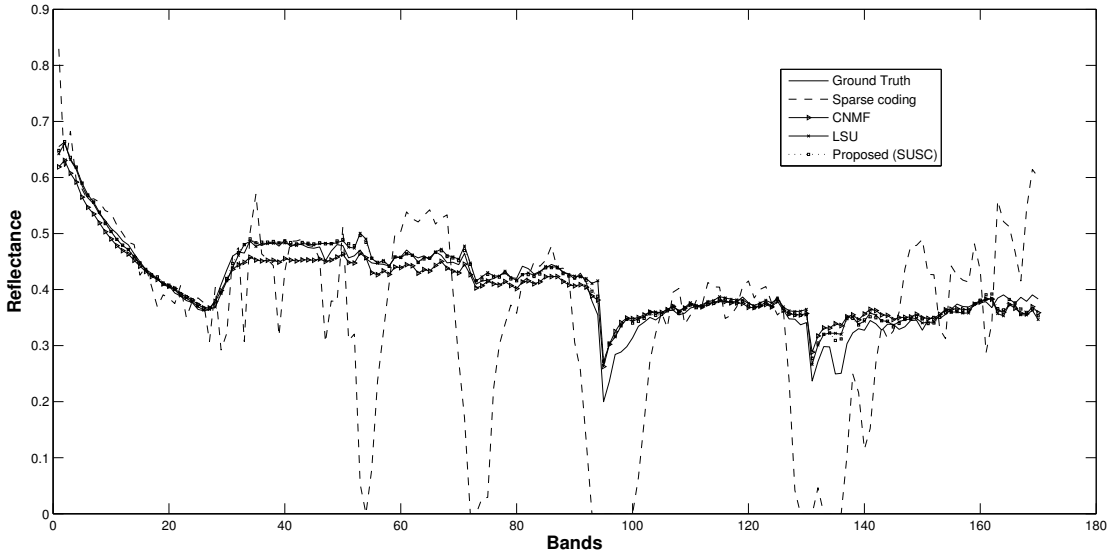


Fig. 11: Obtained pixel spectra of different methods for Hyperion dataset.

[6] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," *U.S. Patent US6 011 875.*, 2000.

[7] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of ms+ pan data," *IEEE Trans. Geosci. Remote Sens.,*, vol. 45, no. 10, pp. 3230–3039, Oct. 2007.

[8] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[9] P. J. Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," *Communications, IEEE Transactions on*, vol. 31, no. 4, pp. 5320–540, Apr. 1983.

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

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

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

[13] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," *Computer VisionECCV 2014*, pp. 63–78, 2014.

[14] C. Grohnfeldt, X. X. Zhu, and R. Bamler, "Jointly sparse fusion of hyperspectral and multispectral imagery," *in Proc. IEEE IGARSS*, pp. 4090–4093, 2013.

[15] C. Grohnfeldt and X. X. Zhu, "Towards a combined sparse representation and unmixing based hybrid hyperspectral resolution enhancement method," *in Proc. IEEE IGARSS*, pp. 2872–2875, 2015.

[16] M. Simes, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "Hyperspectral imagery super-resolution by spatialspectral joint nonlocal similarity," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2671–2679, Jun. 2014.

[17] Y. Zhao, J. Yang, and J. C.-W. Chan, "Hyperspectral image resolution enhancement based on spectral unmixing and information fusion," *in Proc. 2011.ISPRS Hannover, Germany*, pp. 33–36.

[18] M. A. Bendoumi, M. He, and S. Mei, "Hyperspectral image resolution enhancement using high-resolution multispectral image based on spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6574–6583, Oct. 2014.

[19] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE .Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[20] G. Licciardi, M. A. Veganzones, M. Simoes, J. M. Bioucas-Dias, and J. Chanussot, "Super-resolution of hyperspectral images using local spectral unmixing," *in Proc. WHISPERS*, pp. 1–4, 2014.

[21] 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.

[22] X. Liu, W. Xia, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 757–772, Feb. 2011.

[23] J. Liu, "Smoothing filter-based intensity modulation: a spectral preserve image fusion technique for improving spatial details," *International Journal of Remote Sensing*, vol. 21, no. 18, pp. 3461–3472, 2000.

[24] A. Karami, R. Heylen, and P. Scheunders, "Hyperspectral image noise reduction and its effect on spectral unmixing," *in Proc. WHISPERS*, pp. 1–4, 2014.

[25] J. M. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[26] M. E. Winter, "N-findr: an algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proc. SPIE*, vol. 3753, pp. 266–275, Jul. 1999.

[27] J. M. Bioucas-Dias and M. A. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," *in Proc. WHISPERS*, 2010.

[28] 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.

[29] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," *n Proc. 26th Annu. ICML, Montreal, QC, Canada*, pp. 689–696, 2009.

[30] M. Aharon, M. Elad, and A. Bruckstein, "The k-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006.

[31] A. S. Charles, B. Olshausen, and C. Rozell, "Learning sparse codes for hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 963–978, Sept. 2011.

[32] R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the k-svd algorithm using batch orthogonal matching pursuit," *CS Technion*, vol. 40, no. 8, pp. 1–15, 2008.

[33] M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *Image Processing, IEEE Transactions on*, vol. 19, no. 9, pp. 2345–2356, Sept. 2010.

[34] ——, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar . 2011.

[35] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar . 2007.

[36] J. Li and J. M. Bioucas-Dias, "Minimum volume simplex analysis: a fast algorithm to unmix hyperspectral data," *in Proc. IEEE Int. Conf. Geosci. Remote Sens. (IGARSS)*, vol. 3, pp. 250–253, 2008.

[37] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, Nov. 2009.

[38] J. M. Bioucas-Dias, "A variable splitting augmented lagrangian approach to linear spectral unmixing," *in Proc. IEEE GRSS Workshop Hyperspectral Image SIgnal Process.: Evolution in Remote Sens. (WHISPERS)*, pp. 1–4, 2009.

[39] J. M. Bioucas-Dias and J. M. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, 2008.

[40] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?" *in Proc. Int. Conf. Fusion Earth Data*, pp. 99–105, Jan. 2000.

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