Spectral Super-Resolution of Hyperspectral Images via Coupled Dictionary Learning

Konstantina Fotiadou, Grigorios Tsagkatakis, and Panagiotis Tsakalides

Abstract—High spectral resolution imaging systems play a critical role in the identification and characterization of objects in a scene of interest. Unfortunately, multiple factors impair spectral resolution, as in the case of modern snapshot spectral imagers that associate each hyper-pixel with a specific spectral band. In this work, we introduce a novel post-acquisition computational technique aiming to enhance the spectral dimensionality of imaging systems by exploiting the mathematical frameworks of Sparse Representations and Dictionary Learning. We propose a coupled dictionary learning model which considers joint feature spaces, composed of low and high spectral resolution hypercubes, in order to achieve spectral super-resolution performance. We formulate our spectral coupled dictionary learning optimization problem within the context of the Alternating Direction Method of Multipliers, and we manage to update the involved quantities via closed-form expressions. Additionally, we consider a realistic spectral sub-sampling scenario, taking into account the spectral response functions of different satellites. Moreover, we apply our spectral super-resolution algorithm on real satellite data acquired by Landsat-8 and Sentinel-2 sensors. Finally, we have investigated the problem of hyperspectral image unmixing using the recovered high-spectral resolution data-cube, and we are able to demonstrate that the proposed scheme provides significant value in hyperspectral image understanding techniques. Experimental results demonstrate the ability of the proposed approach to synthesize high-spectral resolution three-dimensional hypercubes, achieving better performance compared to state-of-the-art resolution enhancement methods.

1 Introduction

HIGH-resolution remote sensing architectures including Hyperspectral Imagers (HSI) [1] offer valuable insights regarding the composition of a scene and singicifantly facilitate tasks such as object and material recognition [2], spectral unmixing [3]–[5], and region clustering [6]–[10], among others. To accomplish this goal, hyperspectral imaging systems must capture massive amounts of measurements, encoding the dynamics of the spatial and spectral variations in a scene. Currently deployed Earth Observation (EO) satellite platforms provide high frequency global coverage at a much finer spatial resolution compared to the past. Nevertheless, to achieve this goal, typically a small number of spectral observations is considered (ranging from 3 to 12 spectral bands). The work presented in this paper seeks to address this limitation by demonstrating the ability to increase spectral resolution, and thus the ability to perform spectral characterization of remote sensing imagery.

Achieving high spatial, spectral, and temporal resolution is extremely challenging, due to several architectural constraints and conflicting objectives. A characteristic example of this limitation pertains to the remote sensing community, where multispectral instruments such as QuickBird [11] or IKONOS [12] provide low-spectral (*i.e.* RGB, panchromatic, and near-IR) and high-spatial resolution imagery, while other instruments, such as the Hyperion [13] sensor, provide 3D datacubes with high-spectral and low spatial resolution. With regards to the temporal resolution of various satellites, the MODIS [14] multispectral sensor provides global coverage of the entire Earth with a revisit frequency of every one or two days. In contrast, EO-1's Hyperion sensor, has a low-temporal resolution, since it is designed to collect hyperspectral data according to tasking requests. Hence, the combination of different sources of information of the various instruments is vital for Earth observation applications. However, instead of using extra hardware components on the satellite instruments to enhance the spatio-spectral or temporal resolution, remote sensing imaging systems require the development of post-acquisition enhancement techniques, exploiting the already existing information captured by various satellites. Consider as an example the scenario where already acquired imagery from different resolution satellites, could be enhanced using training images of the same region, acquired by higher resolution spectrometers aboard newer platforms.

Additionally, another severe limitation concerns the commissioning phase of the various instruments. For instance, sensors such as QuickBird or IKONOS, have been decommissioned. As a result, these sensors provide only historical data. At this point, another limitation arises from instruments that were malfunctioning, such as the Soil Moisture Active and Passive (SMAP) [15] satellite, which was launched in order to map global soil moisture. SMAP satellite carried two instruments, the radar (active) and the radiometer (passive). Unfortunately, after a short time period the radar instrument has halted its transmissions. Consequently, the design of a generic algorithmic scheme that could exploit the historical information from the period that both instruments were commissioned or normally operating is crucial to the remote sensing community.

Finally, the demand of capturing simultaneously high spectral and spatial information, has led to the design of Spectrally Resolvable Detector Array (SRDA) architectures [16], a new generation of *snapshot* spectral imagers, which seek to acquire

K. Fotiadou is with the Department of Computer Science, University of Crete, Greece and the Institute of Computer Science - FORTH, Greece. E-mail: kfot@ics.forth.gr

G. Tsagkatakis is with the Institute of Computer Science - FORTH, Greece. E-mail: greg@ics.forth.gr

P. Tsakalides is with the Department of Computer Science, University of Crete, Greece and the Institute of Computer Science - FORTH, Greece. E-mail: tsakalid@ics.forth.gr

the entire three-dimensional hypercube over a single integration period. By employing advanced detector fabrication processes, SRDA architectures associate each pixel with a single spectral band according to a pattern that is repeated over the spatial dimensions of the detector. Despite the dramatic improvement these architectures offer with respect to acquisition time, they also lead to a reduction of the spatio-spectral resolution since only a single spectral band is captured by each spatial detector element [17], [18]. As a result, it is of great importance the design of a post-acquisition technique, able to overcome the trade-off between the spatial, spectral and temporal resolutions.

In this work, we propose a novel computational imaging technique that addresses the concept of *spectral super-resolution*, where low and high spectral resolution training examples are used within a machine learning framework to increase the spectral resolution of existing imaging systems. The main motivation in the spectral super-resolution problem is the inability of existing remote sensing imaging sensors to provide simultaneously high spatial and spectral imaging 3D data-cubes. Consequently, our algorithm can be relevant in a wide range of remote sensing applications for Earth Observation.

For instance, acquired imagery from low spectral resolution satellites, e.g. MODIS [14] or Sentinel-2 [19], could be enhanced using images acquired over the same region from higher resolution spectrometers, such as the EO-1 Hyperion sensor [13], as illustrated in Figure 1. Moreover, our algorithm can be valuable in enhancing the spectral dimension of existing historical remote sensing data acquired in the past from satellites that have been decommissioned. Additionally, our scheme could be considered for relaxing data communication requirements by training with high-resolution data during the commission phase and by reducing the required bandwidth during normal operation.

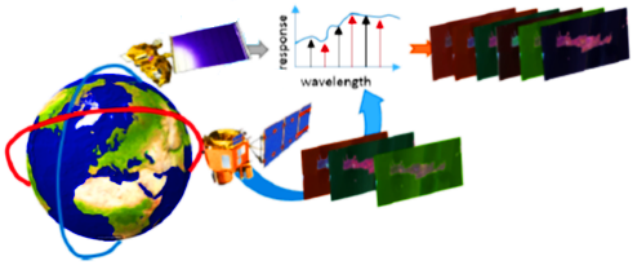


Fig. 1: *A case of spectral super resolution in Earth Observation. A model built on high-low resolution pairs from two instruments is introduced for increasing spectral resolution.*

The proposed *Spectral Coupled Dictionary Learning* (SCDL) algorithm capitalizes on the *Sparse Representations* framework [20] and extents it by introducing a *Coupled Dictionary Learning* process. Furthermore, we solve the SCDL problem within the highly efficient Alternating Direction Method of Multipliers optimization framework [21], [22]. The key contributions of this work can be summarized as follows:

- the formulation of a novel, post-acquisition approach for the enhancement of low-spectral resolution multi- and hyperspectral imagery;
- the design of an efficient coupled dictionary learning architecture, relying on the alternating direction method of multipliers;
- the investigation of a realistic spectral down-sampling scenario, using the spectral signatures of different satellites;
- the systematic evaluation of the proposed spectral resolution enhancement approach on real remote sensing multi- and hyperspectral datasets.

A key benefit of the proposed method is its flexibility, since it can be considered for the enhancement of various pairs of low and high resolution imagery.

The remainder of the paper is organized as follows: Section II provides an overview of the related state-of-the-art. Section III presents the proposed spectral super-resolution scheme of multispectral and hyperspectral imagery, whereas Section IV develops the coupled spectral dictionary learning methodology. Section V provides the data sources and the spectral down-sampling processes, along with the experimental results on both synthetic and real earth observation data. Additionally, in Section V we demonstrate the impact of the proposed SCDL technique on the hyperspectral image unmixing problem. Extensions of this work are discussed in Section VI.

2 Related Work

In this Section, we overview several representative approaches that address the problem of spatial and spectral resolution enhancement of hyperspectral imagery, as well as techniques for learning coupled feature spaces. Although enhancing the spatial, spectral, and temporal resolution of HSI imagery is a subject of significant research, most of the efforts have focused on improving the spatial resolution [23]. Unfortunately, only a few techniques have been proposed in literature that solve the problem of spectral super-resolution. In the following paragraphs we provide in great detail the most representative multi- and hyperspectral image enhancement techniques.

State-of-the-art spatial resolution enhancement approaches may be classified into two representative categories, namely, pan-sharpening [24]–[31] and spatio-spectral fusion [32]–[38]. On the one hand, pan-sharpening combines low-spatial resolution multi- and hyperspectral scenes, along with corresponding high spatial resolution panchromatic images, to synthesize spatially super-resolved 3D data cubes. This is achieved either by replacing the component containing the spatial structure from the HSI image with the panchromatic image [39], or by decomposing the panchromatic image and by re-sampling it to multispectral bands [40]. In both cases, pan-sharpening methods rely on a particular architecture where a high spatial resolution panchormatic camera shares the same field-of-view with a limited resolution spectral imaging system. Additionally, Guo *et al.* [41] tackled the image pan-sharpening problem by utilizing an online coupled dictionary learning technique, where a low-spatial resolution multispectral image is fused with a high-spatial resolution panchromatic image to obtain a high-spatial resolution multispectral image. More recent spatial super-resolution approaches based on learning and neural networks are reported in [42]–[44]. Contrary to the aforementioned techniques, in this work we propose a novel scheme that efficiently learns *coupled* feature spaces, overcoming the limitations arising from independent dictionary learning.

On the other hand, spatio-spectral fusion approaches improve spatial resolution by exploiting the relation between the spatial and the spectral variations of HSI scenes. In [34] the authors present a comparative review of several multi- and hyperspectral fusion techniques. In [45], Bieniarz *et al.* describe how to enhance the

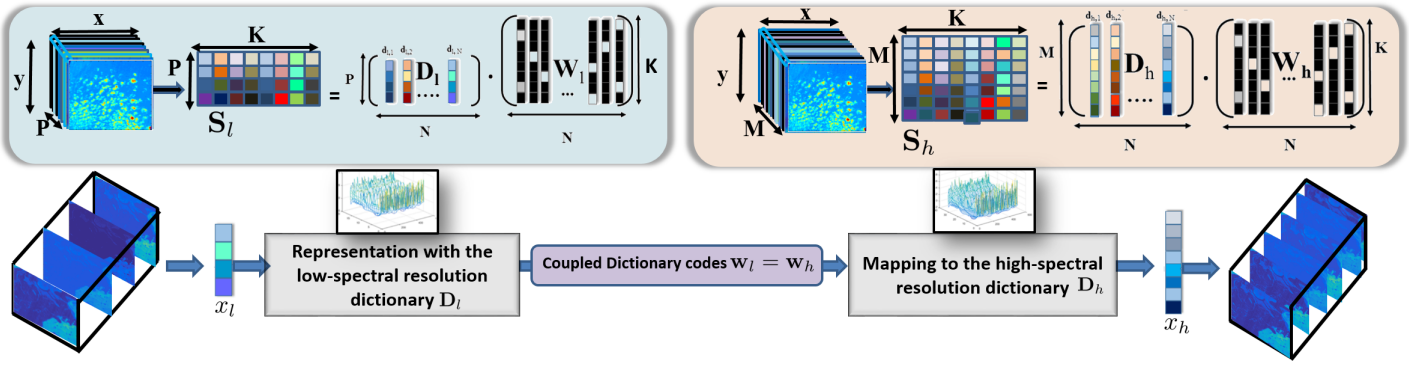


Fig. 2: *SCDL System Block Diagram: The system takes as input a hypercube acquired with a limited number of spectral bands and produces an estimate of an extended spatio-spectral hypercube. During training, multiple high and low-spectral resolution hyper-pixels are extracted from training hypercubes. Given these hyper-pixel pairs, a coupled dictionary learning scheme is employed for learning two sparsifying dictionaries, corresponding to the two resolution cases. During runtime, low-resolution hyper-pixels are mapped to the low-resolution dictionary and the identified sparse coding coefficients are combined with the high-resolution dictionary for producing the final estimates.*

spatial dimension of HSI by employing a sparse spectral unmixing technique and by fusing the results with the multispectral imagery. Similarly, a joint super-resolution and unmixing approach is proposed in [46], based on a sparse representation in the spatial domain and a spectral unmixing in the spectral domain. Erturk *et al.* develop in [47] a spatial super-resolution technique utilizing a fully constrained least squares spectral unmixing scheme with a spatial regularization based on modified binary particle swarm optimization. In [48], Dong *et al.* propose a non-negative sparsity-based hyperspectral super-resolution technique, combining a low-resolution hyperspectral image with a high-resolution RGB image and employing a single dictionary learning scheme to model the relations between the low-spectral resolution HSI and the corresponding high-resolution RGB images. The authors in [49] propose a Bayesian sparse coding method, utilizing Bayesian non-parametric dictionary learning to enhance the spatial variation of multi- and hyperspectral imagery. Finally, in [50] Yin *et al.* combine interpolated low-resolution images with fused images in order to learn their internal sparse representation and to reconstruct the high resolution version of the scene. Before learning the sparse representation, the authors extract the low and high frequency components of the interpolated low-resolution scenes.

An important class of image enhancement techniques considers transferring information between different feature spaces. For instance, Yang *et al.* [51] address the traditional RGB image super-resolution problem by constructing joint dictionaries for the low and the high-resolution spaces under the assumption that the two representations share the same sparse coding. As an extension, in [52] a coupled dictionary learning scheme based on bilevel optimization is proposed and applied to the problems of single image super-resolution and compressed sensing recovery. Although the specific bilevel dictionary learning approach achieves low reconstruction error, the same, possibly sub-optimal, sparse coding is still utilized among the different feature spaces. Consequently, accurate recovery is not assured by the jointly learned dictionaries. Additionally, He *et al.* [53] proposes a beta process based coupled dictionary learning approach by obtaining sparse representations with the same sparsity measure, but with different values in the coupled feature spaces. In contrast, the authors in [54] tackle the problem of spectroscopic data denoising, by exploiting a novel coupled dictionary learning scheme based on the Alternating Direction Method of Multipliers (ADMM) [21].

As opposed to spatial super-resolution, enhancing the *spectral* dimension of HSI scenes has drawn little attention. Charles and Rozell introduced a sparsity-based spectral super-resolution approach for hyperspectral images by learning a dictionary of spectral signatures that decomposes the spectral response of each hyper-pixel [55]. Specifically, they enhance the spectral dimension of multispectral to hyperspectral level by learning an approximation to the data manifold. As an extension of this work, the same authors introduced in [56] a re-weighted ℓ_1 spatial filtering technique that improves with greater accuracy the spectral super-resolution of remote sensing imagery. In this paper, we enhance directly the spectral dimension of remote sensing imagery without impairing the spatial super-resolution. As a result, no further spatial filtering is needed for improving the spatial resolution.

In [57], the authors consider geographically co-located multispectral and hyperspectral oceanic water-color images and they enhance the limited multispectral measurements utilizing a sparse-based approach. First, they use a spectral mixing formulation and they define the measured spectrum for each pixel as the sum of the weighted material spectra. The desired high-resolution spectra are expressed as the linear combination between a blurring matrix and the measured spectra. As a result, the authors take into consideration a blurring operator that represents how the measured (input) spectra is related to the desired (target) hyperspectral spectra. This blurring operator can be considered as an operator that either merges neighboring spectral bands together (i.e. a blurring operator) or omits completely the spectral bands. Consequently, the authors interpret the spectral super-resolution problem as a traditional image denoising problem, which is efficiently solved via a sparse decomposition technique. Our proposed coupled dictionary learning scheme directly learns the coupled feature spaces without requiring the knowledge of a blurring kernel. As a result, our approach can be applied in any type of remote sensing or terrestrial scenario that requires the enhancement of the spectral dimension.

Recently, we proposed several techniques exploiting the low-rank Matrix Completion framework for super-resolving low spatial resolution HSI scenes. In [17], we construct a high spatial and spectral resolution hypercube from undersampled snapshot mosaic imagery. In [58], we employ an alternating minimization coupled dictionary learning technique that is applicable to arbitrary low-high resolution pairs. The work in this paper is an extension of

an earlier sparsity-based approach [59] employing independent dictionaries that model the low and high spectral resolution feature spaces. To the best of our knowledge, this is the first work that applies a *coupled* sparse dictionary learning architecture to the problem of spectral resolution enhancement of HSI data.

3 Spectral Resolution Enhancement

The proposed approach synthesizes a high-spectral resolution hypercube from its low-spectral resolution acquired version by capitalizing on the *Sparse Representations* learning framework (SR) [20]. According to the SR framework, examples extracted from hyperspectral images, can be represented as sparse linear combinations of elements from learned over-complete dictionaries. An initial approach to this problem considers a set of low and high-spectral resolution hyperspectral image pairs and assumes that these images are generated by the same statistical process under a different spectral resolution, and as such, they share the same sparse coding with respect to their corresponding low $\mathbf{D}_l \in \mathbb{R}^{P\times N}$ and high $\mathbf{D}_h \in \mathbb{R}^{M\times N}$ spectral resolution dictionaries [59]. Each low-spectral resolution hyper-pixel $\mathbf{s}_l \in \mathbb{R}^P$ can thus be expressed as a sparse linear combination of elements from a dictionary matrix, $\mathbf{D}_l \in \mathbb{R}^{P\times N}$, composed of hyper-pixel atoms from low-spectral resolution training datacubes, according to $\mathbf{s}_l = \mathbf{D}_l\mathbf{w}$, where $\mathbf{w} \in \mathbb{R}^N$. Recovery of the sparse coding vector $\mathbf{w}$ is accomplished by solving the l_0-minimization problem

$$\min_{\mathbf{w}} ||\mathbf{w}||_0 \ \text{ subject to } \ ||\mathbf{s}_l - \mathbf{D}_l\mathbf{w}||_2^2 < \epsilon, \tag{1}$$

where ϵ denotes the approximation error modelling the system noise, and $||\mathbf{w}||_0 = \#(i|\mathbf{w}_i \neq 0)$ stands for the ℓ_0 pseudo-norm counting the number of non-zero elements in a vector. Although the ℓ_0-norm is theoretically the best regularizer for promoting sparsity, it leads to an intractable optimization. This problem is addressed by replacing the ℓ_0-norm by its convex surrogate ℓ_1-norm, where $\ell_1 = \sum_i |\mathbf{w}_i|$, leading to robust solutions and efficient optimization, as follows

$$\min_{\mathbf{w}} ||\mathbf{w}||_1 \ \text{ subject to } \ ||\mathbf{s}_l - \mathbf{D}_l\mathbf{w}||_2^2 < \epsilon, \tag{2}$$

The equivalent Lagrangian form of the aforementioned optimization problem is formulated as

$$\min_{\mathbf{w}} ||\mathbf{s}_l - \mathbf{D}_l\mathbf{w}||_2^2 + \rho||\mathbf{w}||_1, \tag{3}$$

where the parameter ρ controls the impact of the sparsity on the solution. To obtain the high-resolution signal, the optimal sparse code $\mathbf{w}^\star$ from (3), is directly mapped onto the high-spectral resolution dictionary $\mathbf{D}_h \in \mathbb{R}^{M\times N}$, to synthesize the high-spectral resolution hyper-pixel, according to $\mathbf{s}_h = \mathbf{D}_h\mathbf{w}^\star$. The concatenation of all the recovered high-spectral resolution hyper-pixels synthesizes the high-spectral resolution three-dimensional hypercube, as shown in Figure 2.

The two main challenges pertaining to the estimation of the high spectral resolution hypercubes are related to (i) the sufficient sparsity measure for the sparse coding vector $\mathbf{w}$ and (ii) the proper construction of the low and high spectral resolution dictionary matrices, $\mathbf{D}_l$ and $\mathbf{D}_h$, to efficiently sparsify the input signals.

4 Coupled Dictionary Learning

Generally, coupled dictionary learning refers to the problem of identifying two dictionary matrices standing for two different signal representations, for instance low- and high resolution RGB images [51], blurry and clean images [60], or low-light and well-illuminated scenes [61]. A straightforward strategy to create low and high spectral resolution dictionaries is to randomly sample multiple hyper-pixels extracted from corresponding low and high spectral resolution training scenes and to use this random selection as the sparsifying dictionary. However, this strategy is extremely inefficient since no information regarding the generative power of these examples is known. Alternatively, a joint feature space can be constructed and a single dictionary learning scheme, such as the K-SVD [62], can be considered [59].

The proposed Spectral Coupled Dictionary Learning (SCDL) algorithm relies on generating coupled dictionaries which jointly encode two coupled feature spaces, namely, the observation low-spectral resolution $\mathbf{S}_l \in \mathbb{R}^{P\times K}$, and the latent high-spectral resolution $\mathbf{S}_h \in \mathbb{R}^{M\times K}$. The main task is to find a coupled dictionary pair $\mathbf{D}_l$ and $\mathbf{D}_h$ for the spaces $\mathbf{S}_l$ and $\mathbf{S}_h$, respectively. Formally, the ideal pair of coupled dictionaries $\mathbf{D}_l$ and $\mathbf{D}_h$ can be estimated by solving the following set of sparse decompositions

$$\begin{aligned}&\underset{\mathbf{D}_h,\mathbf{D}_\ell,\mathbf{W}_h,\mathbf{W}_\ell}{\operatorname{argmin}} ||\mathbf{S}_h - \mathbf{D}_h\mathbf{W}_h||_F^2 + ||\mathbf{S}_l - \mathbf{D}_\ell\mathbf{W}_\ell||_F^2 + \\ &\lambda_h||\mathbf{W}_h||_1 + \lambda_\ell\,||\mathbf{W}_\ell||_1, \ \text{ subject to } \mathbf{W}_h = \mathbf{W}_\ell, \\ &||\mathbf{D}_h(:,i)||_2 \leq 1, \ ||\mathbf{D}_\ell(:,i)||_2 \leq 1\end{aligned} \tag{4}$$

where $\mathbf{W}_l$ is the sparse coefficient matrix corresponding to the low-spectral resolution feature space, $\mathbf{W}_h$ stands for the sparse coefficient matrix corresponding to the high-spectral resolution feature space, while λ_h and λ_l denote the parameters that control the sparsity penalty for each individual sub-problem.

Coupled dictionary learning considers the joint identification of two dictionary matrices $\mathbf{D}_h$, $\mathbf{D}_l$, representing the coupled feature spaces $\mathbf{S}_h$ and $\mathbf{S}_l$, such that both hyper-pixels $\mathbf{s}_h(i) \in \mathbf{S}_h$ and $\mathbf{s}_l(i) \in \mathbf{S}_l$ share exactly the same sparse coding vector in terms of $\mathbf{D}_h$ and $\mathbf{D}_l$, respectively. A straightforward approach is to concatenate the coupled feature spaces and utilize a common sparse representation $\mathbf{W}$, able to reconstruct both $\mathbf{S}_h$ and $\mathbf{S}_l$, by solving the optimization problem

$$\begin{aligned}&\underset{\mathbf{D},\mathbf{W}}{\operatorname{argmin}} ||\bar{\mathbf{S}} - \bar{\mathbf{D}}\mathbf{W}||_F^2 + \lambda||\mathbf{W}||_1 \\ &\text{subject to } ||\bar{\mathbf{D}}(:,i)||_2^2 \leq 1, \ \ i = \{1,...,K\},\end{aligned} \tag{5}$$

where $\bar{\mathbf{S}} = \begin{bmatrix}\mathbf{S}_h \\ \mathbf{S}_l\end{bmatrix}$, $\bar{\mathbf{D}} = \begin{bmatrix}\mathbf{D}_h \\ \mathbf{D}_\ell\end{bmatrix}$, and λ is the sparsity regularization term corresponding to the coupled feature space. In addition to sparsity, the elements of the learnt dictionary are also normalized to unit ℓ_2 norm. As a result, the problem posed in (5) is converted into a standard, single sparse decomposition problem, that can be efficiently solved via existing dictionary learning algorithms, such as the K-SVD [62]. However, such a strategy is optimal only in the concatenated feature space, and not in the individual feature spaces of $\mathbf{S}_h$ and $\mathbf{S}_\ell$. Thus, when presented only with examples from $\mathbf{S}_l$, the generated low spectral resolution dictionary $\mathbf{D}_l^*$ may adhere to different optimal space coding compared to $\bar{\mathbf{D}}$.

A major limitation of strategies relying either on a random collection of signal-pairs or on a single dictionary learning, is their inability to guarantee that the same sparse coding can be independently utilized by the different signal resolutions. In other

words, during the application of a spectral super-resolution process, only low-resolution signals are available. Thus, although one could consider only the low-resolution part of a learned dictionary, no constraints on the optimality of the identified sparse codes exist when high-resolution signals are considered. To overcome this limitation, we propose learning a compact dictionary from low and high-spectral resolution hyper-pixels.

We introduce a computationally efficient *coupled dictionary learning* technique, based on the Alternating Direction Method of Multipliers (ADMM) [21], [22], [63], [64] formulation, that converts the constrained dictionary learning problem posed in (5), into an unconstrained version which can be efficiently solved via alternating minimizations. Formally, we consider the observation signals, $\mathbf{S}_\ell = \{\mathbf{s}_\ell\}_{i=1}^N$, and $\mathbf{S}_h = \{\mathbf{s}_h\}_{i=1}^P$. The main task of coupled dictionary learning is to recover both the dictionaries $\mathbf{D}_h$ and $\mathbf{D}_\ell$ and their corresponding sparse codes $\mathbf{W}_h$ and $\mathbf{W}_\ell$, under the constraint, $\mathbf{W}_h = \mathbf{W}_\ell$, by solving the following individual sparse matrix decomposition problems

$$\begin{aligned}&\underset{\mathbf{D}_h,\mathbf{W}_h}{\operatorname{argmin}}||\mathbf{D}_h\mathbf{W}_h - \mathbf{S}_h||_F^2 + \lambda_h||\mathbf{W}_h||_1, ||\mathbf{D}_h(:,i)||_2^2 \leq 1 \\ &\underset{\mathbf{D}_\ell,\mathbf{W}_\ell}{\operatorname{argmin}}||\mathbf{D}_\ell\mathbf{W}_\ell - \mathbf{S}_\ell||_F^2 + \lambda_\ell||\mathbf{W}_\ell||_1, ||\mathbf{D}_\ell(:,i)||_2^2 \leq 1\end{aligned} \tag{6}$$

To apply the ADMM scheme in our spectral dictionary learning procedure, we reformulate the ℓ_1-minimization problem in (6) as

$$\begin{aligned}\min_{\mathbf{D}_h,\mathbf{W}_h,\mathbf{D}_\ell,\mathbf{W}_\ell} &||\mathbf{S}_h - \mathbf{D}_h\mathbf{W}_h||_F^2 + ||\mathbf{S}_\ell - \mathbf{D}_\ell\mathbf{W}_\ell||_F^2 \\ &+\lambda_\ell||\mathbf{Q}||_1 + \lambda_h||\mathbf{P}||_1 \\ \text{subject to } &\mathbf{P} - \mathbf{W}_h = 0, \mathbf{Q} - \mathbf{W}_\ell = 0, \mathbf{W}_h - \mathbf{W}_\ell = 0, \\ &||\mathbf{D}_h(:,i)||_2^2 \leq 1, ||\mathbf{D}_\ell(:,i)||_2^2 \leq 1\end{aligned} \tag{7}$$

Consequently, in comparison with the traditional coupled dictionary learning strategies, we impose the constraint that the two sparse representations $\mathbf{W}_\ell$ and $\mathbf{W}_h$, of the coupled feature spaces, $\mathbf{S}_\ell$ and $\mathbf{S}_h$, should be the same, directly into the optimization. The ADMM scheme takes into account the separate structure of each variable in (7), relying on the minimization of its unconstrained augmented Lagrangian function

$$\begin{aligned}&\mathcal{L}(\mathbf{D}_h, \mathbf{D}_\ell, \mathbf{W}_h, \mathbf{W}_\ell, \mathbf{P}, \mathbf{Q}, Y_1, Y_2, Y_3) = \frac{1}{2}||\mathbf{D}_h\mathbf{W}_h - \mathbf{S}_h||_F^2 + \\ &\frac{1}{2}||\mathbf{D}_\ell\mathbf{W}_\ell - \mathbf{S}_\ell||_F^2 + \lambda_h||\mathbf{P}||_1 + \lambda_\ell||\mathbf{Q}||_1 + < Y_1, \mathbf{P} - \mathbf{W}_h > \\ &+ < Y_2, \mathbf{Q} - \mathbf{W}_\ell > + < Y_3, \mathbf{W}_h - \mathbf{W}_\ell > + \frac{c_1}{2}||\mathbf{P} - \mathbf{W}_h||_F^2 + \\ &\frac{c_2}{2}||\mathbf{Q} - \mathbf{W}_\ell||_F^2 + \frac{c_3}{2}||\mathbf{W}_h - \mathbf{W}_\ell||_F^2\end{aligned} \tag{8}$$

where $\mathbf{Y}_1$, $\mathbf{Y}_2$, and $\mathbf{Y}_3$ stand for the Lagrange multiplier matrices, while $c_1 > 0$, $c_2 > 0$, and $c_3 > 0$ denote the step size parameters. Following the general algorithmic strategy of the ADMM scheme, we seek for the stationary point solving iteratively for each one of the variables while keeping the others fixed. As a result, we create the following sequence of update rules.

- *Sparse Coding Sub-problems*: For minimizing the augmented Lagrangian function with respect to the sparse coding matrices $\mathbf{W}_l$ and $\mathbf{W}_h$, we solve the individual sparse coding problems

$$\begin{aligned}\mathbf{W}_h^\star &= \underset{\mathbf{W}_h}{\operatorname{argmin}}\mathcal{L} \\ \mathbf{W}_\ell^\star &= \underset{\mathbf{W}_\ell}{\operatorname{argmin}}\mathcal{L}\end{aligned} \tag{9}$$

Setting, $\nabla_{\mathbf{W}_h}\mathcal{L} = \nabla_{\mathbf{W}_\ell}\mathcal{L} = 0$, the sub-problems admit closed-form solutions

$$\begin{aligned}&\mathbf{W}_h = (\mathbf{D}_h^T \cdot \mathbf{D}_h + c_1 \cdot I + c_3 \cdot I)^{-1} \cdot (\mathbf{D}_h^T \cdot \mathbf{S}_h + \\ &Y_1 - Y_3 + c_1 \cdot \mathbf{P} + c_3 \cdot \mathbf{W}_\ell) \\ &\mathbf{W}_\ell = (\mathbf{D}_\ell^T \cdot \mathbf{D}_\ell + c_2 \cdot I + c_3 \cdot I)^{-1} \cdot (\mathbf{D}_l^T \cdot \mathbf{S}_\ell + \\ &Y_2 + Y_3 + c_2 \cdot \mathbf{Q} + c_3 \cdot \mathbf{W}_h)\end{aligned} \tag{10}$$

- *Sub-problems* $\mathbf{P}$ and $\mathbf{Q}$

$$\begin{aligned}&\nabla_{\mathbf{P}}\Big(\lambda_h||\mathbf{P}||_1 + < Y_1, \mathbf{P} - \mathbf{W}_h > + \frac{c_1}{2}||\mathbf{P} - \mathbf{W}_h||_F^2\Big) \\ &\nabla_{\mathbf{Q}}\Big(\lambda_\ell||\mathbf{Q}||_1 + < Y_2, \mathbf{Q} - \mathbf{W}_\ell > + \frac{c_2}{2}||\mathbf{Q} - \mathbf{W}_\ell||_F^2\Big)\end{aligned} \tag{11}$$

Setting, $\nabla_{\mathbf{P}}\mathcal{L} = \nabla_{\mathbf{Q}}\mathcal{L} = 0$, the sub-problems can be re-formulated as

$$\begin{aligned}\mathbf{P} &= S_{\lambda_h}\Big(\Big|\mathbf{W}_h - \frac{Y_1}{c_1}\Big|\Big) \\ \mathbf{Q} &= S_{\lambda_\ell}\Big(\Big|\mathbf{W}_\ell - \frac{Y_2}{c_2}\Big|\Big),\end{aligned} \tag{12}$$

where S_{λ_h} and S_{λ_ℓ} denote the soft-thresholding operators, defined as

$$S_\lambda(x) = sign(x) \cdot max(|x| - \lambda, 0), \tag{13}$$

where $\lambda > 0$ stands for the threshold value.

- *Sub-problems* D_h and D_ℓ

For a fixed set of $\mathbf{W}_h$, $\mathbf{W}_\ell$, $\mathbf{P}$ and $\mathbf{Q}$, the dictionaries $\mathbf{D}_h$ and $\mathbf{D}_\ell$ can be updated as

$$\begin{aligned}\mathbf{D}_h^\star &= \underset{\mathbf{D}_h}{\operatorname{argmin}}\mathcal{L} \\ \mathbf{D}_\ell^\star &= \underset{\mathbf{D}_\ell}{\operatorname{argmin}}\mathcal{L} \Leftrightarrow\end{aligned} \tag{14}$$

$$\begin{aligned}\nabla_{\mathbf{D}_h}\Big(\frac{1}{2}||\mathbf{S}_h - \mathbf{D}_h\mathbf{W}_h||_F^2\Big) &= -(\mathbf{S}_h - \mathbf{D}_h\mathbf{W}_h)\mathbf{W}_h^T \\ \nabla_{\mathbf{D}_\ell}\Big(\frac{1}{2}||\mathbf{S}_\ell - \mathbf{D}_\ell\mathbf{W}_\ell||_F^2\Big) &= -(\mathbf{S}_\ell - \mathbf{D}_\ell\mathbf{W}_\ell)\mathbf{W}_\ell^T\end{aligned} \tag{15}$$

Setting $\nabla_{\mathbf{D}_h} = \nabla_{\mathbf{D}_\ell} = 0$, the high and the low-spectral resolution dictionaries are updated column by column adhering to the following iterative scheme

$$\begin{aligned}\phi_h &= \mathbf{W}_h(j,:) \cdot \mathbf{W}_h(j,:)^T \\ \phi_\ell &= \mathbf{W}_\ell(j,:) \cdot \mathbf{W}_\ell(j,:)^T,\end{aligned} \tag{16}$$

and

$$\begin{aligned}\mathbf{D}_h^{(k+1)}(:,j) &= \mathbf{D}_h(:,j)^{(k)}(:,j) + \frac{\mathbf{S}_h \cdot \mathbf{W}_h(j,:)}{\phi_h + \delta} \\ \mathbf{D}_\ell^{(k+1)}(:,j) &= \mathbf{D}_\ell(:,j)^{(k)}(:,j) + \frac{\mathbf{S}_\ell \cdot \mathbf{W}_\ell(j,:)}{\phi_\ell + \delta}\end{aligned} \tag{17}$$

where k denotes the number of iterations, δ stands for a small regularization factor, while $\mathbf{D}_h(:,j)$ and $\mathbf{D}_\ell(:,j)$ represent the j-th column of $\mathbf{D}_h$ and $\mathbf{D}_\ell$, respectively. Finally, the Lagrangian multiplier matrices are updated as

$$
\begin{aligned}
Y_1^{(k+1)} &= Y_1^{(k)} + c_1(\mathbf{P} - \mathbf{W}_h) \\
Y_2^{(k+1)} &= Y_2^{(k)} + c_2(\mathbf{Q} - \mathbf{W}_l) \\
Y_3^{(k+1)} &= Y_3^{(k)} + c_3(\mathbf{W}_h - \mathbf{W}_l)
\end{aligned} \tag{18}
$$

In our setup, we set $c_1 = c_3 = 0.8$ and $c_2 = 0.6$. The derivations of the individual sub-problems for the proposed SCDL-ADMM based dictionary learning scheme are shown in the Appendix. The overall algorithm for learning the coupled dictionaries, which correspond to the high and the low-spectral resolution feature spaces, is summarized in **Algorithm 1**.

Algorithm 1 Spectral Coupled Dictionary Learning

Input: training examples $\mathbf{S}_h$ and $\mathbf{S}_l$, number of iterations K and step size parameters c_1, c_2, c_3.
Initialize: $\mathbf{D}_h \in \mathbb{R}^{M\times N}$ and $\mathbf{D}_l \in \mathbb{R}^{P\times N}$ are initialized by a random selection of the columns of $\mathbf{S}_h$ and $\mathbf{S}_l$ with normalization; Initialize Lagrange multiplier matrices $\mathbf{Y}_1 = \mathbf{Y}_2 = \mathbf{Y}_3 = \mathbf{0}$.
for $k = 1, \cdots, K$ do

1) Update $\mathbf{W}_h$ and $\mathbf{W}_\ell$ via (10)
2) Update $\mathbf{P}$ and $\mathbf{Q}$ via (12)
3) **for** $j = 1, \cdots, N$ do

- Update ϕ_h and ϕ_l via (16)
- Update the two dictionaries $\mathbf{D}_h$ and $\mathbf{D}_l$ column by column via (17)

end

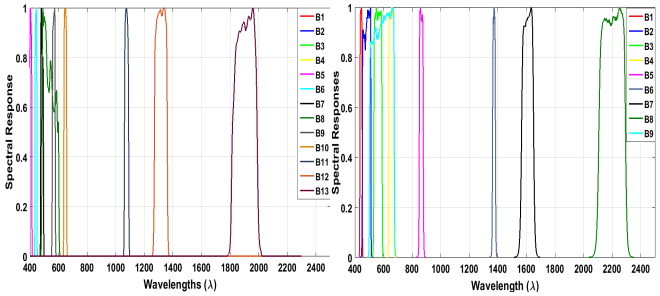
- Normalize $\mathbf{D}_h$ and $\mathbf{D}_l$ between $[0, 1]$
- Update Lagrange multiplier matrices Y_1, Y_2 and Y_3 via (18)

end

5 Experimental Evaluation

In this Section, we evaluate the performance of the proposed SCDL scheme when applied to the spectral super-resolution of hyperspectral imagery in terms of the quality of the estimated high spectral resolution hypercubes. The performance is quantified on both synthetic and real remote sensing data. Specifically, we have experimented with multi- and hyperspectral data acquired by: (i) NASA's EO-1 Hyperion [13] sensor, (ii) NASA's MODIS sensor [14], (iii) NASA's Landsat-8 OLI instrument [65], and (iv) ESA's Sentinel-2 satellite [19].

The Hyperion sensor resolves 224 spectral bands ranging from 0.4 to 2.5 μm, with a 30-meter spatial resolution. Due to its high spectral coverage, Hyperion scenes have been widely utilized in the remote sensing community for classification and spectral unmixing purposes [3], [6]. Specifically, we considered several hyperspectral scenes extracted from different Earth locations, acquired on June 16, 2016. We restrict ourselves to the 96 calibrated and high-resolution spectral bands in the visible and near infrared (VNIR) spectrum range, namely: (B9:B16, B18:B25, B28:B33, B42:B45, B49:B57, B77:B105, B106:B115, B141:B160, and B191:B192). Additionally, the MODIS [14] sensor acquires 36 spectral bands ranging between 0.4 and 14.4 μm, at varying spatial resolutions (2 bands at 250m, 5 bands at 500 m, and 29 bands at 1 km). Finally, we considered multispectral data scanned by the same region and extracted by the OLI sensor of the Landsat-8 satellite and the Sentinel-2 sensor, on September 17, 2017. The OLI sensor collects data at a 30-meter spatial resolution and resolves 8 spectral bands in the visible and near-infrared and in the shortwave infrared (SWIR) spectral regions of the electromagnetic spectrum, plus an additional panchromatic band at 15 meter spatial resolution, resulting into 9 spectral bands. On the other hand, the Sentinel-2 satellite provides high-spatial, spectral and temporal-resolution multispectral scenes, while it ensures the continuity of Landsat's observations. Sentinel-2 covers the visible and near-infrared and the shortwave infrared spectral regions, resolving 13 spectral observations. Figure 3 depicts the spectral response functions of the Sentinel-2 and the Landsat-8 multispectral sensors, where one can observe the overlapping and the non-overlapping spectral bands between the two different instruments.



(a) *Spectral Responses of Sentinel-2* (b) *Spectral Responses of Landsat-8*

Fig. 3: *Spectral signatures of the Sentinel-2 and the Landsat-8 multispectral sensors.*

5.1 Implementation and Evaluation Metrics

In order to validate the performance of the hyperspectral image enhancmenet algorithms, we employ the *Peak Signal to Noise Ratio* (PSNR) metric [66] given by

$$PSNR = 10\log_{10}[L_{max}^2/MSE(x, y, b)],$$

where L is the maximum pixel value of the scene, b denotes the spectral dimension, and MSE stands for the mean square error, defined as

$$\text{MSE}(x, y, b) = \frac{\sum_{x,y,b}\left[\mathbf{S}_{h_{(x,y,b)}} - \mathbf{S}_{l_{(x,y,b)}}\right]^2}{n_x n_y b}, \tag{19}$$

where x and y denote the spatial dimensions of the input and the synthesized images $\mathbf{S}_l$ and $\mathbf{S}_h$. Consequently, we evaluate the PSNR error metric across all the recovered spectral bands.

Additionally, we evaluate the performance of the recovered hyperspectral images in terms of the *Spectral Angular Mapper* (SAM) [67]–[69]. This metric determines the *spectral similarity* between two spectra by calculating the angle between the corresponding vectors of the testing and the reference hypercubes, formulated as

$$\theta = cos^{-1}\Big(\frac{\sum_{i=1}^{N} t_i r_i}{\sqrt{\sum_{i=1}^{N} t_i^2}\sqrt{\sum_{i=1}^{N} r_i^2}}\Big), \tag{20}$$

where t stands for the testing spectrum, r refers to the reference spectrum, while N denotes the number of the available spectral bands. Small values of SAM indicate a high similarity between the compared spectral vectors.

Regarding the dictionary training phase, pairs of low and high spectral resolution dictionaries were prepared, one for each sensor data set. In all cases, we utilized 10 training hypercubes from which 10^5 training hyper-pixels were randomly extracted. In the experimental results section, we further investigate and show graphically the impact of the dictionary size on the reconstruction performance.

6 Experimental Results

6.1 Synthetic Data Scenario

Concerning the synthetic data case, we use hyperspectral data acquired by NASA's EO-1 Hyperion satellite [13]. Instead of using predefined sub-sampling factors for synthesizing the low spectral resolution data-cubes [59], we consider a more realistic scenario of spectral down-sampling based on the spectral profiles of low-spectral resolution sensors. Specifically, we first construct the spectral calibration data, *i.e.* the spectral profiles (down-sampling) matrix that represents the relation between the two different representations, the high and the low spectral resolutions.

In order to generate the corresponding low-spectral resolution hypercubes, we consider the following scenario: for each spectral observation of Hyperion we find the corresponding wavelength value of another (target) instrument, that acquires hyperspectral data with a limited spectral resolution. Formally, let $\mathbf{s}_h \in \mathbb{R}^P$, be a high-spectral resolution hyper-pixel, acquired with P spectral bands, while $\mathbf{B} \in \mathbb{R}^{P \times M}$ stands for the spectral profiles (down-sampling) matrix, representing the spectral calibration data, and M denotes the number of limited spectral observations. The corresponding low-spectral resolution hyper-pixel, $\mathbf{s}_l \in \mathbb{R}^M$, is constructed as $\mathbf{s}_l = \mathbf{B}^T \mathbf{s}_h$. This procedure is performed for every spectral observation of the low-spectral resolution instrument.

In this work, we have experimented with ESA's Sentinel-2 and NASA's MODIS spectral response functions, and thus we have created two different spectral down-sampling matrices, one for each acquisition scenario. Concerning the Sentinel-2 satellite case, we considered the corresponding spectral observations that overlap with the calibrated and high-resolution Hyperion spectral bands. Specifically, the overlapping spectral bands among the two instruments are (B1:B4, B7, B8, and B12). The spectral calibration data for the Sentinel acquisition scenario form a matrix of size 96×7. Consequently, we reconstruct the 96 high-resolution bands of Hyperion from only 7 spectral observations of the Sentinel-2 sensor. Likewise, we used the spectral profiles of the multispectral MODIS sensor to synthesize the low-spectral resolution hypercubes. For each Hyperion spectral observation we find the corresponding wavelength values of the MODIS sensor. Similarly, the overlapping spectral observations between the Hyperion and the MODIS sensors are B1:B6, B9:B14, and B16:B19, resulting in 14 spectral observations for the low-resolution case. As a result, the spectral down-sampling matrix is of size 96×14. Figure 4 illustrates this spectral sub-sampling process.

In order to validate the merits of the proposed spectral super-resolution scheme, we first compare the synthesized 3D hypercubes against the ground truth cubes and against three state-of-art techniques, namely (i) the baseline approach of cubic interpolation among the available spectral bands; (ii) the sparse-based method of spectral super-resolution using the naive scenario of K-SVD coupled dictionary learning [59]; and (iii) the reweighted ℓ_1 spatial filtering algorithm (RWL1-SF) [55], [56] that learns a single dictionary of spectral signatures in order to decompose the spectral response of each hyper-pixel. To achieve a fair comparison with the K-SVD coupled dictionary learning technique, we utilize the same number of atoms for dictionary learning and the same sparsity constraints. For the reweighted ℓ_1 spatial filtering scheme, we chose the parameters that achieve the best possible reconstruction. In the following paragraphs, the performance of the competing spectral super-resolution techniques is assessed using the quantitative evaluation metrics PSNR and SAM and by qualitative visual inspection.

Figures 5 and 6 show representative bands from the reconstructed hypercubes obtained using the proposed SCDL method when applied on Cube 1 of the EO1H1860262006099110KF hyperspectral scene. In both figures, we employ the spectral down-sampling for the Sentinel-2 instrument, and thus we reconstruct the full spectrum composed of 96 spectral bands from only 7 available spectral observations. Figure 5 demonstrates four characteristic spectral bands from a full-spatial resolution hypercube $(3000 \times 900 \times 96)$, along with a specific region of interest. For this purpose, we depict the reconstructed 10^{th}, 25^{th}, 30^{th}, and 50^{th} spectral bands, along with their ground truth observations. We chose the specific spectral bands since they are able to discriminate the subtle differences of the reconstructed hyper-cube. The red-squared area corresponds to the zoomed region that is depicted on the upper left corner of the images, and it highlights the subtle differences among the various spectral observations. In terms of the PSNR metric, the proposed technique achieves a high performance with respect to the corresponding ground truth measurements, while the mean PSNR accross all the reconstructed spectral observations of Cube 1 hyperspectral scene is **46.90** dB. In terms of visual inspection, we note that the reconstructed spectral bands faithfully preserve important image features.

On the other hand, Figure 6 depicts a cropped-area of the Cube 1 hyperspectral scene. The spatial dimensions of the cropped area are 271×184 pixels. The results highlight the fact that important spatial features of the images, such as the field areas, are correctly synthesized. In this experiment, one can easily notice how different image regions, corresponding to different materials, are reliably estimated. The average PSNR values obtained by the proposed SCDL scheme for the recovery of the full and the cropped three-dimensional hyper-cubes are **46.90** dB and **30.7** dB, respectively.

In Figure 7, we compare the performance of SCDL with the three state-of-the-art methods when applied on the cropped (271×184) area of Cube 1 (EO1H1860262006099110KF). In this experiment, we use Sentinel's spectral profiles in order to down-sample the high-spectral resolution hyperspectral data, and thus we recover the full spectrum, composed of 96 spectral bands, from only 7 input spectral observations. We observe that the RWL1-SF and the K-SVD techniques are not able to reconstruct significant features such as the road areas at the top of the scenes. One may notice the subtle differences among the various techniques in the marked red-squared region of interest. The reweighted spatial filtering technique introduces severe artifacts on the reconstruction. In contrast, the K-SVD spectral super-resolution approach provides a smoother recovery compared to RWL1-SF, but not as good as the proposed SCDL scheme, which achieves the highest accuracy with the ground truth three-dimensional hypercube, both visually and quantitatively, in terms of the attained evaluation metrics. The PSNR values obtained by RWL1-SF, K-SVD, and SCDL are 25.72 dB, 26.85 dB, and **30.7** dB, respectively. In terms of the SAM error metric, the RWL1-SF algorithm achieves 0.188, while

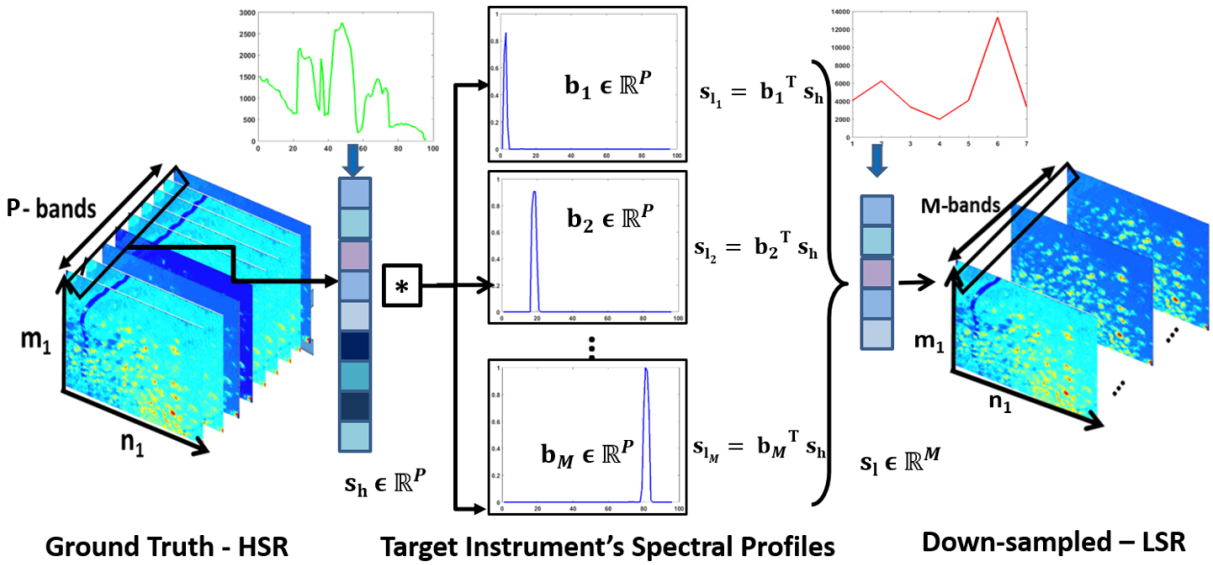
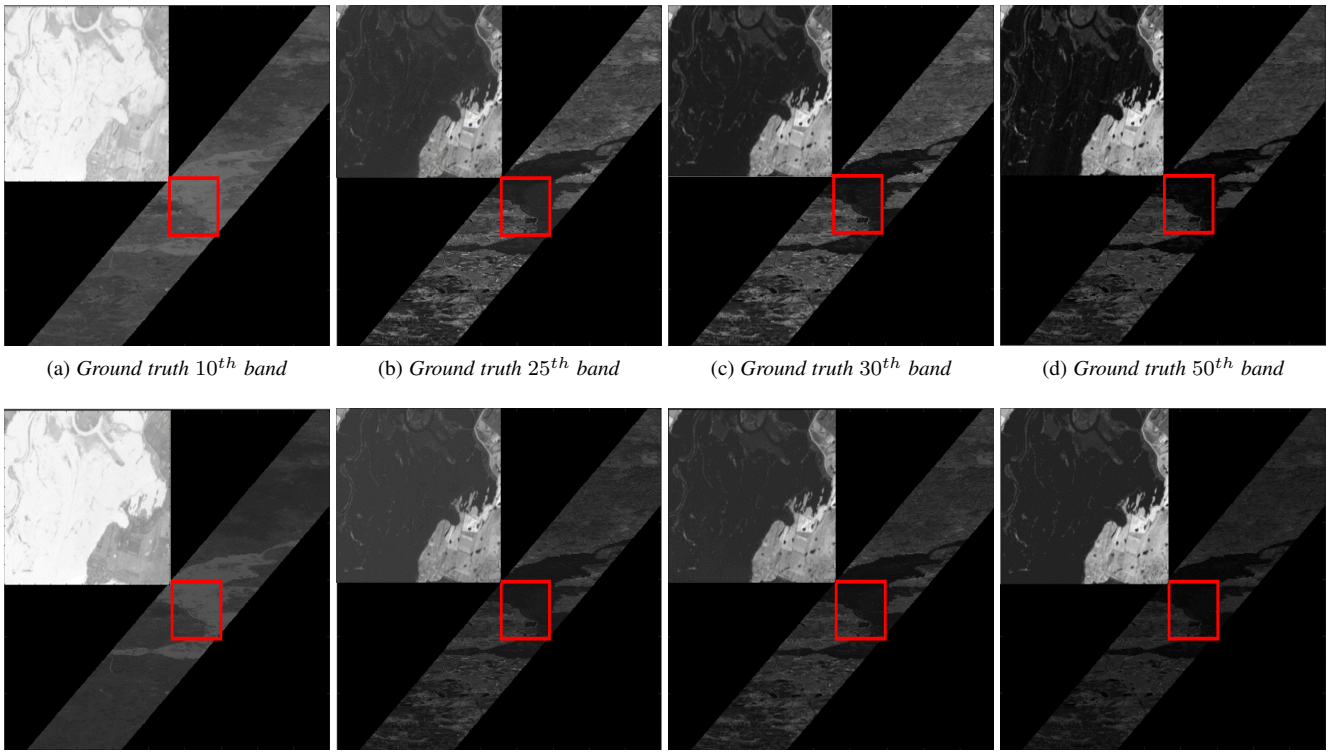


Fig. 4: *Hyperspectral data down-sampling process: High-resolution hyper-pixels are modulated by the spectral calibration data in order to synthesize the corresponding realistic low-spectral resolution hyper-pixels. The spectral profiles (down-sampling) matrix provides the corresponding wavelength values between the two spectral-resolution instruments. As illustrated in the figure, each column vector of this matrix is a spectral response function describing the relation between the two satellite sensors.*



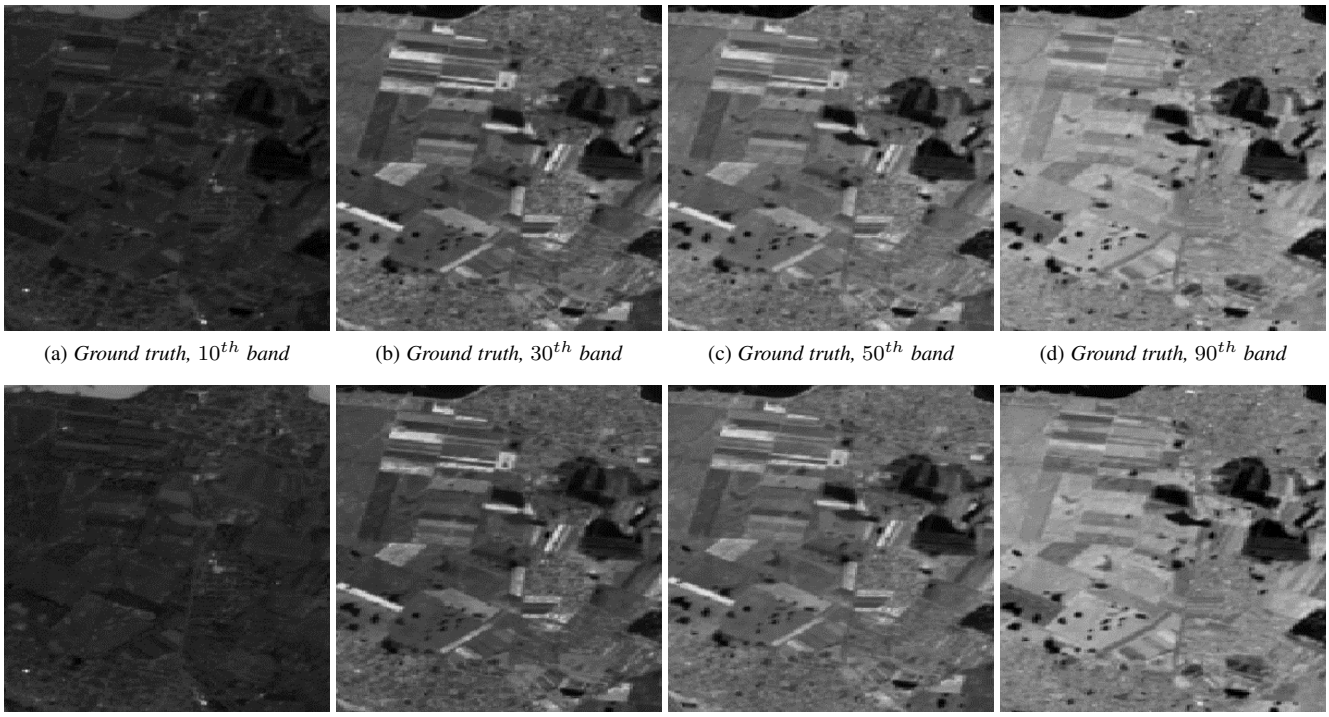
(a) *Ground truth* 10^{th} *band* (b) *Ground truth* 25^{th} *band* (c) *Ground truth* 30^{th} *band* (d) *Ground truth* 50^{th} *band*

(e) 10^{th} *band, PSNR:47.87 dB* (f) 25^{th} *band, PSNR:46.54 dB* (g) 30^{th} *band, PSNR:44.32 dB* (h) 50^{th} *band, PSNR:46.81 dB*

Fig. 5: *Sentinel to Hyperion spectral band reconstruction using SCDL: The full spectrum is composed of 96 bands in the VIS-NIR region, and the full resolution hypercube is estimated from 7 input spectral bands. In this simulation we illustrate four characteristic spectral bands of the proposed algorithm's reconstruction. Specifically, we demonstrate the* 10^{th}, 25^{th}, 30^{th}, *and* 50^{th} *spectral observations. Each spectral band corresponds to a different wavelength range. Consequently, we may notice the subtle differences among the spectral bands. Additionally, we observe that under real life conditions, the proposed SCDL scheme produces a significant quality improvement operating in satellite hyperspectral imagery.*

the K-SVD spectral super-resolution approach 0.035. In contrast, the proposed SCDL scheme achieves an even smaller SAM value, **0.027**, indicating a higher similarity among the reconstructed and ground truth hypercubes.

Another indicative set of reconstruction results is depicted in Figure 8, where the performance of the various methods is



(a) *Ground truth,* 10^{th} *band* (b) *Ground truth,* 30^{th} *band* (c) *Ground truth,* 50^{th} *band* (d) *Ground truth,* 90^{th} *band*

(e) 10^{th} *band, PSNR:* 38.97 *dB* (f) 30^{th} *band, PSNR:* 45.26 *dB* (g) 50^{th} *band, PSNR:* 41.23 *dB* (h) 90^{th} *band, PSNR:* 42.84 *dB*

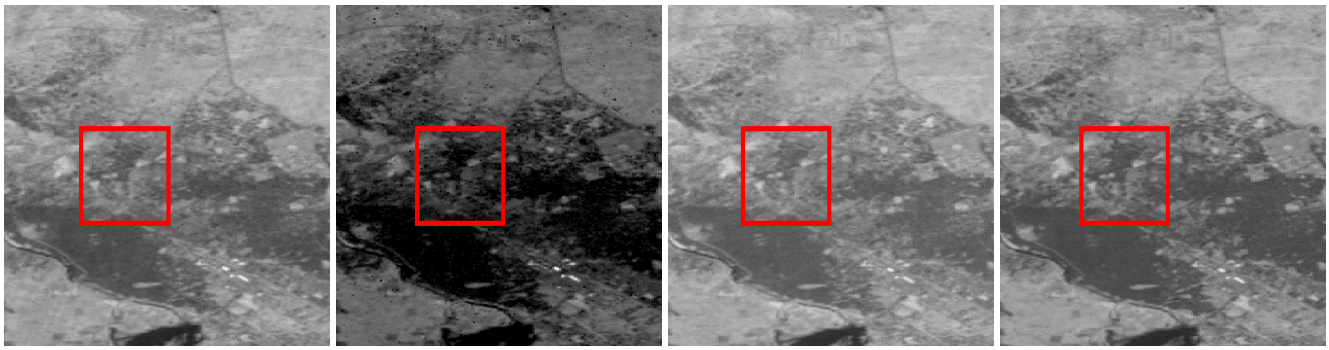
Fig. 6: *Sentinel to Hyperion cropped area reconstruction: In this experiment, we investigate the performance of our SCDL scheme when applied on a cropped area of Cube 1 (EO1H1860262006099110KF), with a spatial resolution of* (271×184) *pixels. Top row: Original spectral bands. Bottom row: SCDL reconstructed spectral bands. The full spectrum is composed of* 96 *bands, while the available spectral observations are only* 7*. In this experiment we demonstrate four characteristic spectral bands, each one corresponding to a different value of the spectrum.*

(a) *Ground Truth, 30th Band* (b) *RWL1-SF, PSNR:*25.72 *dB, SAM:0.188* (c) *SSR-K-SVD, PSNR:*26.85 *dB, SAM:0.035* (d) *SCDL, PSNR:**30.7** dB, SAM:**0.027***

Fig. 7: *Cube 1 (EO1H1860262006099110KF): Comparison with the state-of-the-art: In this experiment we consider the 30th band of Hyperion's hyperspectral scene. We recover the full hypercube from only* 7 *input spectral observations, using Sentinel's spectral profiles matrix. The SCDL scheme outperforms state-of-the art approaches both visually and in terms of the PSNR error metric.*

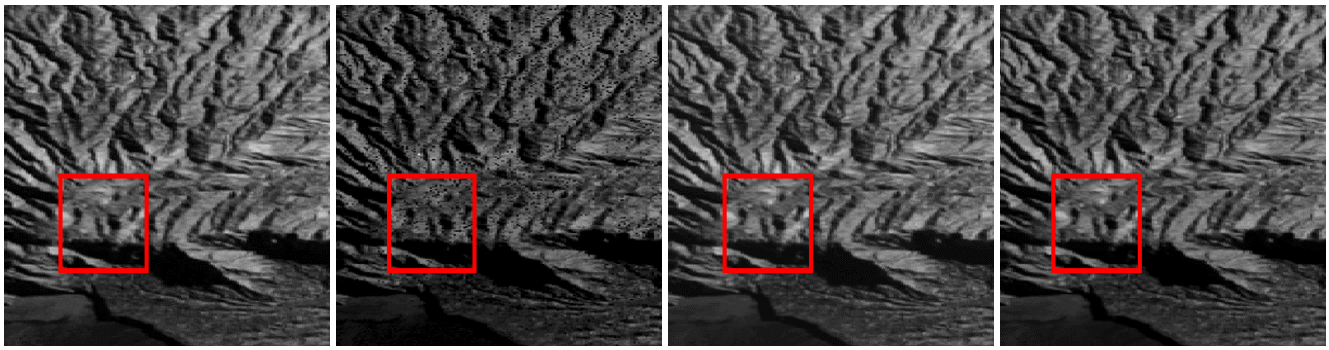
evaluated on a cropped region (400×200) of hyperspectral Cube 3 (EO1H1900512016017110KF). In this experiment, we used MODIS spectral profiles to synthesize the low-spectral resolution part. Hence, the full 96 spectral bands spectrum is recovered from 14 spectral observations. Figure 8 illustrates the 53th spectral band. Visual observation reveals that the RWL1-SF technique introduces severe artifacts at the high spatial frequency regions such as the marked red-squared area. On the other hand, K-SVD produces a smooth spectral representation of the scene without introducing spectral and noisy effects. Again, the proposed SCDL technique provides a more accurate and smooth approximation of the ground truth, revealing significant details over all regions of this challenging scene. In terms of the PSNR metric, the proposed scheme outperforms the other methods by achieving **31.83** dB, as compared to 22.23 dB and 30.22 dB, obtained by RWL1-SF and K-SVD, respectively. Similar observations can be made for the SAM index, where the proposed algorithm attains the smallest value of **0.021**, better than the RWL1-SF (0.103) and the K-SVD (0.035).

Figure 9 illustrates the performance of the various techniques when applied on hyperspectral Cube 4 (EO1H1180652016162110KF). Specifically, we depict a cropped (400×200) region of the 70^{th} acquired spectral band. In this scenario, we use the MODIS instrument spectral profiles and we recover the 96-band full spectrum from 14 input spectral observations. Outperforming the state-of-the-art approaches that



(a) *Ground Truth* 53^{th} *Band* (b) *RWL1-SF, 22.23 dB, SAM:0.103* (c) *K-SVD, PSNR:30.22 dB, SAM:0.035* (d) *SCDL, PSNR:**31.83** dB, SAM:0.021*

Fig. 8: *Cube 3 (EO1H1900512016017110KF) Comparison with state-of-the-art: In this experiment, we use MODIS spectral profiles in order to synthesize the low-spectral resolution part. The full spectrum is composed of 96 bands in the VIS-NIR region, while the full resolution hypercube is estimated from* 14 *input spectral bands. We observe that under real life conditions, the proposed SCDL method has a better performance than both RWL1-SF and K-SVD.*



(a) *Ground Truth* 70^{th} *Band* (b) *RWL1-SF, 28.93 dB, SAM:0.120* (c) *K-SVD, PSNR:31.64 dB, SAM:0.054* (d) *SCDL, PSNR:**35.81** dB, SAM:0.034*

Fig. 9: *Cube 4 (EO1H1180652016162110KF): Comparison with the state-of-the-art: In this experiment, we reconstruct the full spectrum, composed of 96 bands in the VIS-NIR region, from* 14 *input spectral bands using MODIS spectral profiles matrix. The proposed scheme provides an accurate high-quality reconstruction of the challenging scene, both quantitatively and visually.*

produce noisy reconstructions, SCDL provides a high resolution hypercube that has a better similarity with the original spectral data, both visually and quantitatively. Additionally, in the highlighted red-squared region of interest, one may observe that the proposed scheme faithfully preserves all the important image features. The PSNR values obtained by RWL1-SF, K-SVD and SCDL are 28.93 dB, 31.64 dB and **35.81** dB, respectively. In terms of the SAM metric, the proposed SCDL scheme achieves the best (smallest) value of **0.034**, followed by K-SVD (0.054) and RWL1-SF (0.120).

A complete set of PSNR and SAM values for all the evaluated techniques applied on several EO-1 Hyperion hyperspectral scenes, of full spatial resolution (3000×900), is provided in Table 1. In this table we show the mean reconstruction error and the SAM metric across all the recovered spectral observations. The results demonstrate that the proposed spectral resolution enhancement method outperforms consistently all other competing state-of-the-art techniques on a wide range of challenging hyperspectral scenes.

6.2 Real Data Scenario

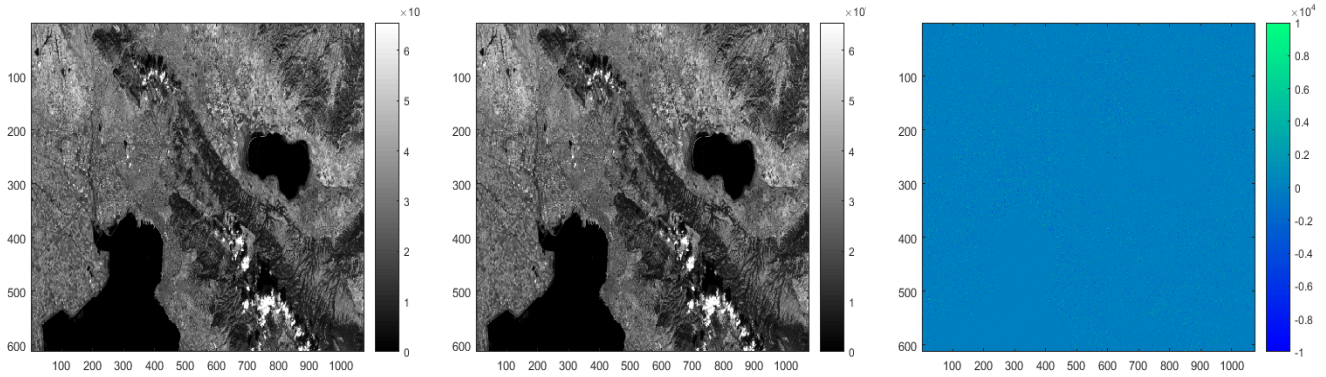
In this paragraph, we evaluate the performance of the proposed SCDL algorithm applied on a real data acquisition case. Specifically, we considered multispectral data scanned in the same region and extracted by the Landsat-8 OLI sensor and the Sentinel-2 sensor. The challenge is to recover the 13 spectral bands of the Sentinel-2 sensor using as input 9 spectral observations of the Landsat-8 sensor.

Figure 10 demonstrates the reconstruction result for the 12^{th} spectral band of the Sentinel-2 sensor using the proposed SCDL scheme. Specifically, we illustrate the actual 12^{th} spectral band and its reconstructed version by SCDL. As we may observe, the achieved reconstruction has a high similarity to its corresponding ground truth spectral band, both visually and quantitatively in terms of the PSNR (39.75 dB) and SAM (0.034) error metrics. Figure 10(c) illustrates the difference between the ground truth and the reconstructed 12^{th} spectral band and clearly demonstrates that the proposed scheme recovers accurately the ground truth spectral observation.

In Figure 11, we provide the comparison with the state-of-the-art algorithms when applied to the real data scenario. As the testing data-cube, we utilize a multispectral scene captured by Landsat-8 on September, 17th of 2017. Visual observation in both the full-spatial resolution image and the cropped red square region reveals that the RWL1-SF algorithm produces noisy reconstructions with severe artifacts. Although the K-SVD produces more faithful approximations of the ground truth spectral information, the proposed SCDL algorithm synthesizes spectral observations of

TABLE 1: *Quantitative performance evaluation of the proposed SCDL method against state-of-the-art techniques in terms of PSNR (dB) and SAM error metrics, using MODIS and Sentinel's spectral profiles.*

Image	Scale	Interpolation	RWL1-SF	K-SVD	SCDL
Cube 1 (EO1H1860262006099110KF)	MODIS to Hyperion	33.18, (0.093)	38.60, (0.261)	44.78, (0.034)	**46.40**, (**0.033**)
	Sentinel to Hyperion	37.19, (0.165)	40.57, (0.122)	42.85, (0.031)	**46.90**, (**0.024**)
Cube 2 (EO1H1830332004264110PX)	MODIS to Hyperion	24.66, (0.211)	31.77, (0.224)	32.84, (0.035)	**41.01**, (**0.017**)
	Sentinel to Hyperion	29.14, (0.220)	36.80, (0.203)	38.10, (0.030)	**40.18**, (**0.023**)
Cube 3 (EO1H1900512016017110KF)	MODIS to Hyperion	23.60, (0.105)	28.61, (0.097)	29.63, (0.037)	**38.15**, (**0.024**)
	Sentinel to Hyperion	28.61, (0.215)	35.18, (0.271)	36.26, (0.021)	**40.81**, (**0.019**)
Cube 4 (EO1H1180652016162110KF)	MODIS to Hyperion	33.49, (0.150)	38.50, (0.120)	38.74, (0.063)	**43.92**, (**0.032**)
	Sentinel to Hyperion	30.34, (0.253)	38.12, (0.232)	38.35, (0.050)	**41.15**, (**0.035**)
Cube 5 (EO1H1120822017023110K7)	MODIS to Hyperion	33.84, (0.080)	31.64, (0.236)	40.58, (0.023)	**48.59**, (**0.017**)
	Sentinel to Hyperion	38.72, (0.150)	36.18, (0.277)	46.50, (**0.020**)	**49.40**, (0.030)
Average	MODIS to Hyperion	29.75, (0.127)	33.82, (0.187)	37.31, (0.038)	**43.61**, (**0.024**)
	Sentinel to Hyperion	32.80, (0.200)	37.37, (0.221)	40.41, (0.030)	**43.80**, (**0.026**)



(a) *Ground truth* 12^{th} *band* (b) *SCDL, PSNR:39.75 dB, SAM: 0.034* (c) *Difference between SCDL and Ground Truth*

Fig. 10: *Real Data Scenario (North Greece Region): In this experiment we trained the coupled dictionaries using Sentinel-2 and Landsat-8 data scanned in the same regions. During the testing phase, we consider as input* 9 *spectral bands of the Landsat-8 multispectral sensor, and we recover the* 13 *spectral observations of the Sentinel-2 satellite. In order to verify our reconstruction, the output high-spectral resolution hypercube is compared with the ground truth Sentinel-2 cube. Figure (c) demonstrates the difference between the reconstructed* 12^{th} *spectral band and the ground truth* 12^{th} *band. We observe that the recovered band achieves high similarity with the ground truth spectral band, both visually and in terms of the evaluation metrics (PSNR and SAM).*

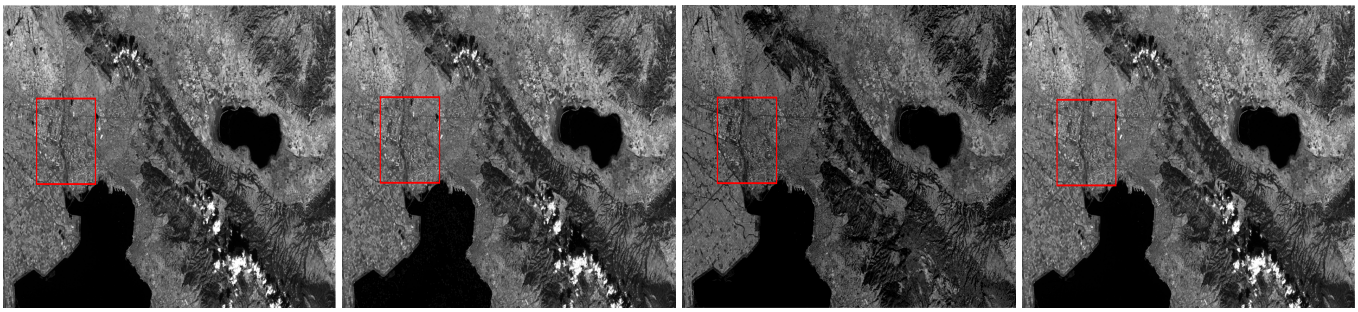
a higher quality without introducing noise effects. Quantitatively, the proposed SCDL scheme outperforms the other techniques in terms of the evaluation metrics, achieving a PSNR value of **39.75** dB, in contrast to the RWL1-SF and the K-SVD that achieve 23.50 dB and 33.57 dB, respectively. Similarly, in terms of the SAM error metric, the proposed technique achieves the smallest value of **0.014** as compared to RWL1-SF and K-SVD that achieve 0.224 and 0.032, respectively.

Moreover, in Figure 12 we depict the difference between the reconstruction by each algorithm and the actual 12th spectral band. As we may observe, both RWL1-SF (Figure 12-b) and K-SVD (Figure 12-c) achieve reconstructions that are noticeably different from the ground truth. In contrast, the difference image attained by our SCDL technique, as depicted in Figure 12-d, demonstrates a high quality recovery that matches accurately with the ground truth 12^{th} spectral band. Finally, Figure 13 demonstrates the SAM images of the compared techniques for the North Greece Region hyper-cube. We observe that both visually and in terms of the SAM index, the proposed ADMM coupled dictionary learning scheme, outperforms the state-of-the-art approaches. In terms of visual perception, we may notice that our approach provides a high quality SAM image of the testing hypercube, without introducing noise artifacts. In constrast, both the K-SVD and RWL1-SF techniques, provide lower-quality SAM images.

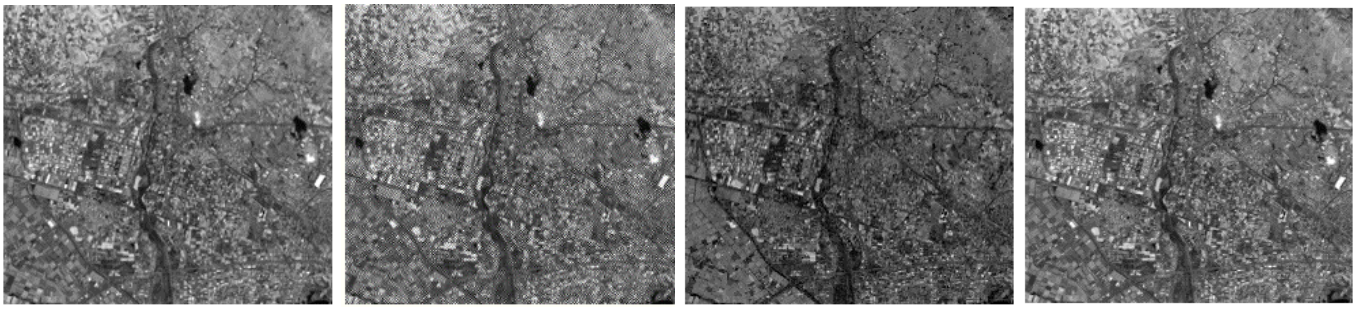
6.2.1 Impact on Hyperspectral Image Denoising

In this paragraph, we examine the performance of the proposed SDCL scheme on the fundamental problem of hyperspectral image denoising. In further detail, we considered the real data acquisition scenario, enhancing the spectral dimensions of Landsat-8 to the spectral resolution of Sentinel-2 sensor. The major difference with the previous paragraphs is that we have also considered the noise existence. Specifically, we assumed that the noise distribution of the low-spectral resolution part, adheres to a normal distribution, $N \sim (0, 0.05)$. Consequently, we have added zero-mean $(\mu = 0)$ Gaussian noise, with 0.05 standard deviation ($\sigma^2 = 0.05$), to Landsat's-8 training examples. The main objective is to recover the 13 spectral bands of the Sentinel-2 sensor using as input the 9 noisy spectral observations of Landsat-8 instrument. In order to achieve this goal, we have prepared one pair of dictionaries that represent the low-resolution and noisy part, and the corresponding high-resolution observations.

Figure 14 stands as an example of the proposed SCDL scheme applied on the hyperspectral image denoising problem. Specifically, we illustrate the ground truth and the noisy 5th observation of Landsat-8 sensor, along with the synthesized 5th spectral band of Sentinel-2. As we may notice, the proposed coupled dictionary learning algorithm is able to provide a high-quality and spectrally super-resolved data-cube, from the noisy low-spatial and spectral

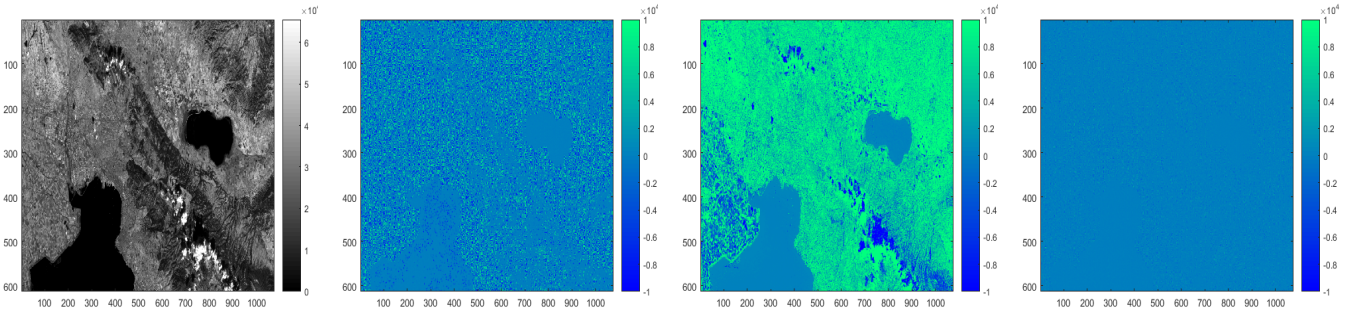


(a) *Ground truth* 12^{th} *band* (b) *RWL1-SF, 23.50 dB, SAM: 0.224* (c) *K-SVD, PSNR:33.57 dB, SAM: 0.032* (d) *SCDL, PSNR:**39.75** dB, SAM: 0.014*



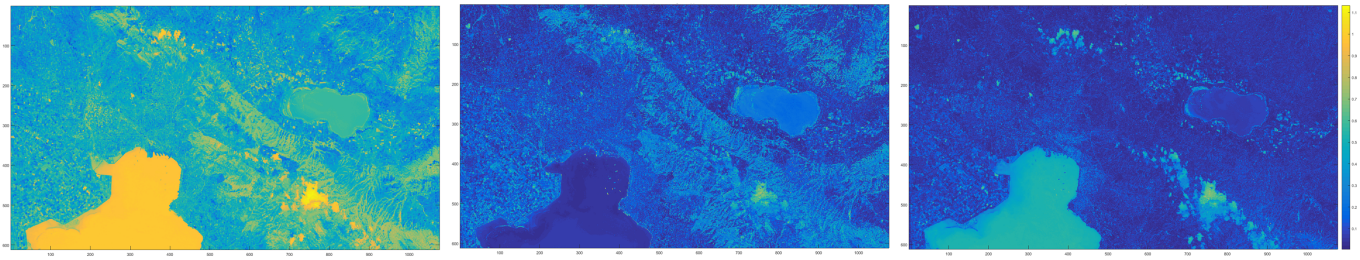
(e) *Zoomed Region -* 12^{th} *Band* (f) *Zoomed Region - RWL1-SF* (g) *Zoomed Region - SSR-K-SVD* (h) *Zoomed Region - SCDL*

Fig. 11: *Real Data Acquisition Scenario (North Greece Region): Comparison with state-of-the-art. In this experiment we depict an example of the proposed algorithm when it is applied on real satellite data. We evaluate the reconstruction quality of the proposed SCDL scheme compared with the state-of-the-art, in the scenario where we recover the* 13 *spectral bands of Sentinel-2, considering as input* 9 *spectral observations of Landsat-8. We observe that the proposed algorithm outperforms the state-of-the-art techniques both in terms of the evaluation indexes and visually. In the zoomed red-square regions depicted on the second row, we illustrate the subtle differences among the various techniques.*



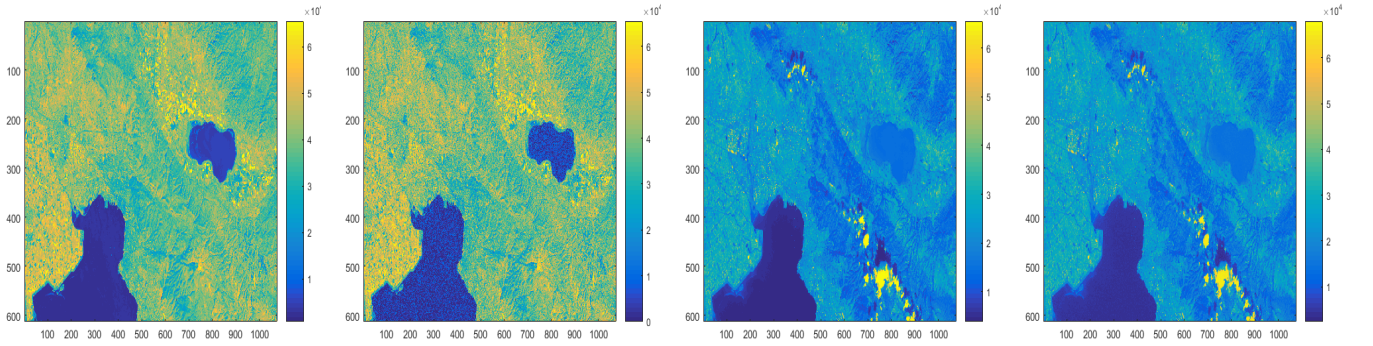
(a) *Ground Truth* 12^{th} *Band* (b) *Difference between RWL1-SF and Ground Truth* (c) *Difference between SSR-K-SVD and Ground Truth* (d) *Difference between SCDL and Ground Truth*

Fig. 12: *Real Data Scenario (North Greece Region): We provide the differences between the* 12*th band reconstructed by the methods and the ground truth* 12^{th} *spectral band of the Sentinel-2 sensor. We notice, that both the RWL1-SF and the K-SVD algorithms produce well-structured difference images, implying a low-quality recovery. In contrast, the proposed technique's difference image with respect to the ground truth denotes an accurate reconstruction of the* 12^{th} *spectral band.*



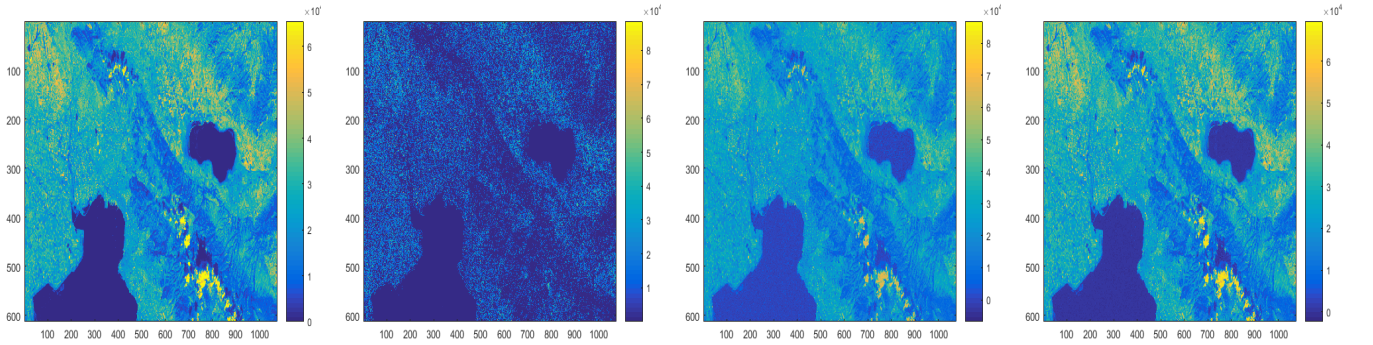
(a) *RWSF, SAM:**0.224*** (b) *SSR-K-SVD, SAM:*0.032 (c) *SCDL, SAM:**0.014***

Fig. 13: *Real Data Scenario (North Greece Region): In this simulation, we illustrate the Spectral Angle Mapper (SAM) images for the comparable techniques. As we may observe, both the RWLW-SF and the K-SVD algorithms provide more degraded reconstructions, in comparison with the proposed algorithm. Additionally, quantitatively, in terms of the error metric the proposed algorithm achieves a lower SAM value, compared to the other two state-of-the-art approaches.*



(a) *Ground truth* 5^{th} *Landsat's-8 band* (b) *Noisy* 5^{th} *Landsat's-8 band* (c) *Ground Truth* 5*th Sentinel's-2 bana* (d) *Recovered* 5*th Sentinel's-2 band*

Fig. 14: *Hyperspectral Image Denoising: In this experiment, we provide the ground* 5*th spectral band of the Landsat-8, influenced by the existence of zero-mean, and* $\sigma^2 = 0.05$ *Gaussian noise. (From left to right:) We illustrate the accurate Landsat's-8* 5*th spectral band, the corresponding noisy observation, the ground truth Sentinel's-2* 5*th spectral band and the proposed system's reconstruction. Both in term of visual perception, and quantitatively, the proposed scheme achieves a faithful reconstruction of the high resolution data-cube even on the extreme scenario when the input low-spectral resolution part is degraded by noise.*



(a) 12^{th} *Sentinel's-2 band* (b) *RWL1-SF, SAM: 0.398* (c) *K-SVD, SAM: 0.091* (d) *SCDL, SAM:* $\mathbf{0.030}$

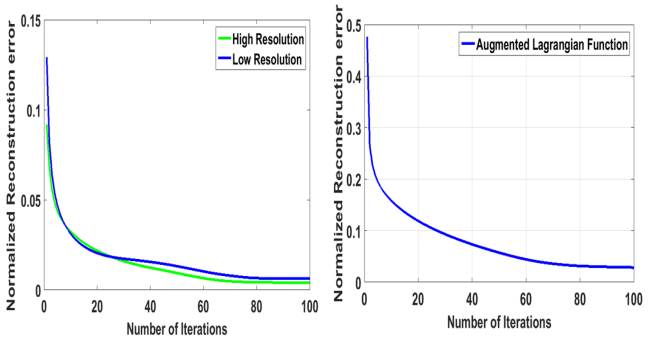
Fig. 15: *Impact of Hyperspectral Image Denoising (Real Data Acquisition - North Greece Region): In this simulation, we examine the reconstruction quality of the comparable techniques when the low-resolution part of the Landsat-8 sensor is degraded by Gaussian noise. For this purpose, we trained coupled dictionaries using Sentinel-2 and noisy Landsat-8 data, scanned in the same spatial locations. To verify our recovery, the output, high-spectral resolution and denoised hypercube is compared with the ground truth Sentinel-2 data-cube. We notice that the proposed SCDL scheme outperforms the comparable literature approaches both in visual perception and quantitatively. Consequently, given the appropriate training, the proposed algorithm is able to confront the challenging problem of hyperspectral image denoising.*

resolution input.

Additionally, in figure 15, we depict the comparison with the state-of-the art algorithms, applied to the scenario where the low-spectral resolution part is influenced by the presence of Gaussian noise. Specifically, we illustrate the 12th band of the multispectral testing scene that was utilized on the previous examples (*i.e* Figures 10,11). As we may notice, the proposed SCDL scheme outperforms both the K-SVD dictionary learning and the RWL1-SF approaches, both visually, but also in terms of quantitative metrics. For instance, the RWL1-SF approach achieves a SAM value of 0.398, the K-SVD reconstruction reaches the SAM value of 0.091, while the proposed SCDL algorithm achieves the smallest of $\mathbf{0.030}$. Although initial simulation results indicate that the proposed SCDL scheme is also capable of addressing the hyperspectral denoising problem, a detailed analysis of the denoising perspectives, including consideration of different types of degradation, and subsequent comparison with state-of-the-art denoising techniques is beyond the scope of this paper, and thus it is left as future work.

6.3 Convergence

In this Section, we investigate the empirical convergence of the proposed algorithm when it is applied on the MODIS to Hyperion spectral super-resolution case. Specifically, we examine the convergence behaviour of both the high and low-spectral resolution dictionaries, $\mathbf{D}_h, \mathbf{D}_\ell$, as well as the convergence of the Augmented Lagrangian function $\mathcal{L}$. Figure 16 depicts the normalized



(a) *Convergence of* $\mathbf{D}_h$ *and* $\mathbf{D}_\ell$ (b) *Convergence of the Augmented Lagrangian* $\mathcal{L}$

Fig. 16: *Convergence Behaviour of the proposed Dictionary Learning Algorithm: (Left) Convergence of the two dictionaries. (Right) Convergence of the Augmented Lagrangian function. In all cases, the proposed algorithm converges into a stationary point.*

reconstruction errors for the two dictionaries and the Augmented

Lagrangian function as a function of the number of iterations. We note that both dictionaries converge after approximately 20 iterations while the Augmented Lagrangian function needs around 60 iterations for convergence. We observed a similar convergence behavior for both the Sentinel-2 to Hyperion and the Landsat-8 to Sentinel-2 spectral super-resolution scenarios.

6.4 Sensitivity to Parameters

In the following paragraphs, we investigate the impact of the parameters selection in our algorithm's performance. First, we examine the number of training examples for the coupled dictionary learning procedure versus the reconstruction quality. Then, we investigate the impact of the dictionary size (number of dictionary atoms) to the reconstruction quality. Finally, we investigate the proper selection of the sparsity regularization parameter (λ) to both the execution time of our algorithm and the reconstruction quality, in terms of PSNR.

To study the sensitivity of the proposed algorithm, we evaluated the reconstruction performance of the coupled trained dictionaries as a function of the number of training examples. In

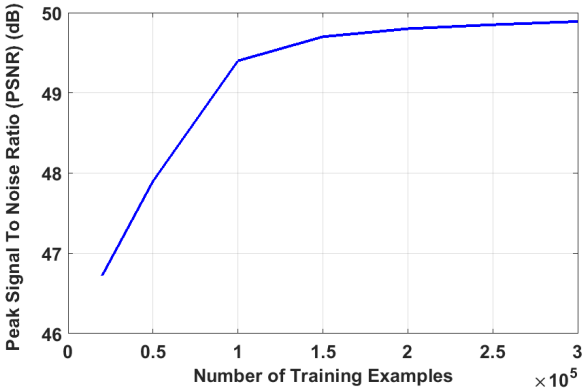


Fig. 17: *Cube 5 hyperspectral scene: Sentinel to Hyperion spectral super-resolution scenario. In this simulation, we illustrate the PSNR of the proposed SCDL algorithm as a function of the number of training examples. We observe that after approximately* 10^5 *training examples, SCDL reaches a stable plateau.*

Figure 17, we provide the PSNR values for the reconstruction of Cube 5 (EO1H1120822017023110K7) hyperspectral scene, as a function of the number of training examples. In this simulation, we use Sentinel's spectral profiles, and thus we synthesize the full-spectrum from 7 spectral observations. Specifically, we investigated the performance of the proposed scheme when we start from a relatively small size of training examples, *i.e.* 0.5×10^5, and we gradually increased the training size to 3×10^5 examples. Results indicate that the performance of the SCDL method monotonically increases as a function of the number of the input training examples, as expected. However, experimentally we observed that increasing the training size over approximately 1.5×10^5 examples, the PSNR reaches a stationary plateau. We note that after approximately 10^5 training examples, SCDL achieves its highest PSNR value of **49.50** dB. Consequently, for the spectral super-resolution problem, using a larger number than 10^5 training examples in the coupled dictionary learning process offers marginal improvement to the reconstruction quality. For this reason, we fixed the training size to 10^5 examples.

Figure 18 demonstrates the PSNR performance for various spectral bands in the Landsat-8 to Sentinel-2 spectral super-resolution scenario. Note that for all spectral bands, the proposed SCDL scheme achieves higher reconstruction quality, as compared with the state-of-the-art techniques. On the other hand,

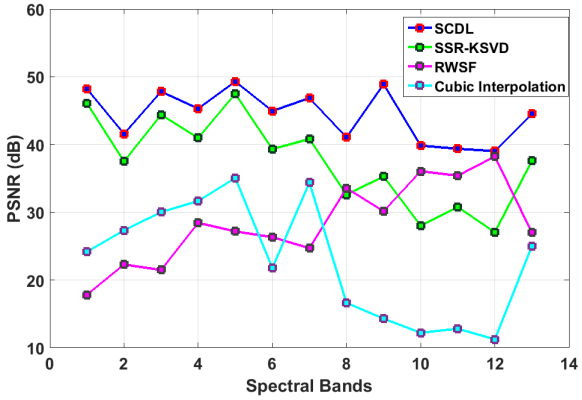
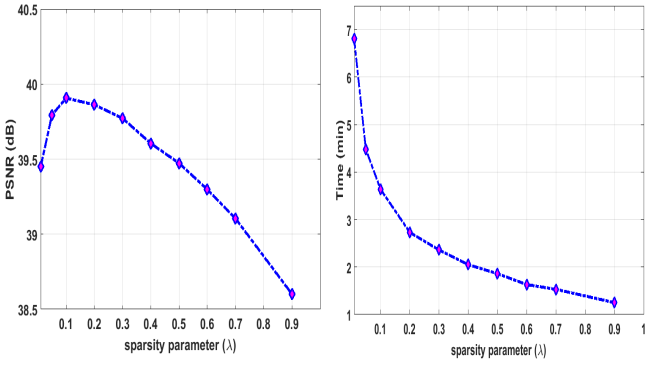


Fig. 18: *Landsat-8 to Sentinel-2, real data scenario: Evaluation of the reconstruction quality in terms of the PSNR metric for 14 spectral band indexes. The proposed SCDL scheme achieves consistently higher PSNR values, compared with state-of-the-art approaches.*

Figure 19 investigates the impact of the sparsity parameter λ, on both the reconstruction quality and the execution time. This experiment was implemented using the Landsat-8 to Sentinel-2 spectral resolution enhancement scenario. Specifically, we used as a test scene, the hypercube that is depicted in Figures 10 and 11 (North Greece Region). In order to select the proper sparsity regularization parameter for the reconstruction process, we have experimented with several λ values, ranging from 0.1 to 0.9. Specifically, we observed that for $\lambda = 0.1$ our system achieves the highest reconstruction performance in terms of the PSNR error metric. As we gradually increase the value of λ, the execution time decreases dramatically, but unfortunately the reconstruction quality degrades. The highest PSNR value is achieved for $\lambda = 0.1$. For this value of λ, the proposed scheme reconstructs the high-spectral resolution north Greece hypercube of spatial dimensions (612×1076) in about 2.8 minutes when the sparse-based RWSF-L1 approach requires approximately 3.2 hours. Consequently, the proposed SCDL scheme outperforms the state-of-the-art both in terms of reconstruction performance and execution time. As a result, the optimal selection of the sparsity parameter is crucial for the every sparse-based algorithm. We note that in order to achieve a fair comparison with the other two sparse-based algorithms (KSVD and RWSF-L1), we used the same sparse regularization term.

Finally, in Figure 20, we demonstrate the impact of the dictionary size on the reconstruction performance. In this experiment, we utilized the hypercube that is depicted in Figure 5, and thus we reconstructed the full spectrum composed of 96 spectral bands from only 7 spectral observations. Specifically, for a fixed sparsity regularization parameter, $\lambda = 0.2$, we have investigated the impact of using different dictionary sizes consisting of 512, 1024, 2048, and 4096 atoms. It is important to note, that as we increase the number of dictionary atoms (> 1024), the reconstruction quality decreases. However, in these cases as well, the proposed recovery outperforms the state-of-the-art. The results indicate that the optimal reconstruction performance is achieved when we use 1024 dictionary atoms. Similar behaviour was observed for all



(a) *Sparsity parameter vs. PSNR* (b) *Sparsity parameter vs. Execution Time*

Fig. 19: *(From Left to right) In this experiment, we demonstrate the impact of the sparsity regularization parameter* (λ) *on the reconstruction performance and on the execution time. As the value of* λ *increases, the execution time decreases and the reconstruction quality drops. For* $\lambda = 0.2$*, our system achieves optimal performance, balancing between a high-quality reconstruction and a short execution time.*

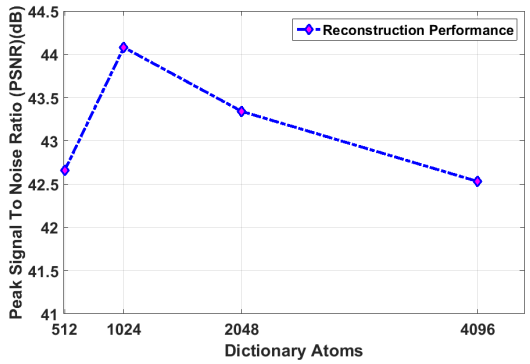


Fig. 20: *Reconstruction performance as a function of the dictionary size: In this experiment, we used the hypercube that is depicted in Figure 5 and we reconstructed the full spectrum composed of* 96 *spectral bands from only* 7 *spectral observations. The best performance of* ***44.04*** *dB is achieved when we use* 1024 *dictionary atoms.*

investigated spectral resolution enhancement scenarios.

6.5 Spectral Coupled Dictionary Learning for Hyperspectral Image Understanding

In the following paragraphs, we examine the impact of the proposed spectral super-resolution framework on the fundamental understanding process of hyperspectral unmixing.

Separating a remote sensing image into its elementary components is crucial for multiple applications, including precision agriculture, and weather and climate forecasting among others [70]. Consequently, spectral unmixing provides a well-organized mapping of the pure materials that exist in the raw data, by identifying the spectral signatures of these materials (*i.e* endmembers) along with their relative contributions (*i.e* abundancies). The endmembers are assumed to represent the pure materials that are depicted on the hyperspectral scene, while the abundances, represent the amount of each endmember that is presented in the pixel.

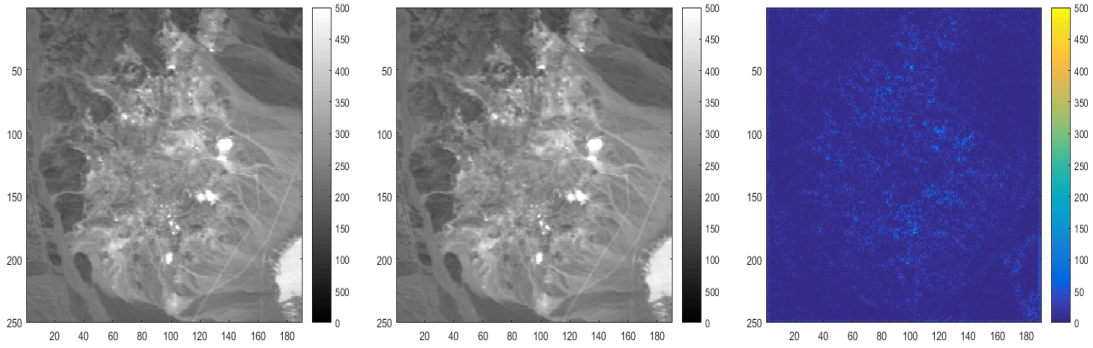
In order to investigate the hyperspectral unmixing performance after applying the proposed spectral super-resolution technique, we exploited a novel linear spectral unmixing algorithm, namely: *Fast Unmixing Algorithm* (FUN) [71]. According to FUN algorithm, each captured pixel in a hyperspectral scene, can be represented as a linear combination between a definite number of endmembers weighted by an abundace factor. The specific technique provides effectively, and with low computational complexity, simultaneous extraction of the endmembers along with their abundancies, by exploiting a modified GramSchmidt technique.

Regarding the evaluation setup, we have experimented with the Cuprite scene, acquired by the AVIRIS hyperspectral instrument [72]. The specific hyperspectral scene has been widely utilized for both unmixing and classification purposes. In order to demonstrate that the proposed spectral super-resolution architecture provides improvement in the hyperspectral unmixing problem, we adhered to the following experimental scenario: first, we enhanced the spectral dimension of the Cuprite scene using the proposed SCDL scheme, and then we applied the FUN algorithm for the extraction of the endmembers and abundancies. The number of reconstructed endmembers for both the ground truth Cuprite, and the reconstructed from the degraded input, was 12. As a result, we evaluated the unmixing performance in the scenarios when we have the high-spectral versus the low-spectral resolution hypercubes.

In greater detail, the high resolution data-cube was created by removing the noisy (B1-B2 and B221-B224), and water-vapor absorption B104-B113 and B148-B167 spectral bands, resulting into 188 spectral observations. For the construction of the low-resolution counterpart, we considered the spectral response functions of the Landsat-8 satellite. In order to provide the correspondence between the two instruments, we constructed the spectral profiles matrix, by finding their overlapping wavelength values. Specifically, the overlapping spectral bands among the two instruments are: (B8,B10,B21,B24,B31,B54,B133,B188). The spectral calibration data for the AVIRIS acquisition scenario form a matrix of size 188×8. Consequently, we reconstruct the 188 high resolution bands of AVIRIS from only 8 spectral observations of Landsat-8 instrument.

Regarding the dictionary training phase, we prepared one pair of dictionaries, corresponding to the low (8-bands) and high (188-bands) feature spaces. Additionally, we utilized 5 training hypercubes from which 50.000 training hyper-pixels were randomly extracted. The number of selected dictionary atoms was set to 512, while the proposed SCDL scheme converged into approximately 30 iterations. Quantitatively, the proposed SCDL scheme reconstructs the Cuprite hyperspectral scene, achieving a PSNR value of **44.41** dB, and a SAM error metric of **0.0368**. In Figure 21, we provide the recovered 20th spectral band of AVIRIS Cuprite scene, using the proposed SCDL scheme, the ground truth 20th spectral and their absolute difference. As we may observe, the proposed SCDL scheme also corresponds with high efficiency in the specific scenario of spectral degradation.

Concluding, in Figure 22 we evaluate the unmixing performance, by comparing the extracted abundancies using the SCDL super-resolved hypercubes, versus the abundancies that were synthesized directly from the low-spectral resolution Cuprite hypercube. As a baseline, we utilize the 12 publicly available Cuprite ground truth abundancies [72]. In all cases, we observe that the estimated abundancies using the synthesized high-spectral resolution SCDL hypercube provide an accurate reconstruction of the ground truth abundancies. Consequently, the proposed SCDL recovery provides a high-quality, super-resolved data-cube, able to facilitate the hyperspectral unmixing procedure.



(a) *Ground Truth* $20th^{th}$ *Band* (b) *SCDL Recovered 20th Band* (c) *Difference between SCDL and Ground Truth*

Fig. 21: *In this figure we provide the SCDL recovered* 20*th AVIRIS Cuprite band, along with the ground truth* 20^{th} *spectral band, and the absolute difference. We observe, that the proposed technique's difference image with respect to the ground truth denotes an accurate reconstruction of the* 20^{th} *spectral band.*

7 Conclusion

In this work, we proposed a novel spectral super-resolution architecture for multi- and hyperspectral imagery, employing the mathematical framework of *Sparse Representations* through a *Coupled Dictionary Learning* algorithm for encoding the relations between high and low-spectral resolution scenes. To achieve this goal, an efficient formulation was proposed based on the Alternating Direction Method of Multipliers. Experimental results on both synthetic and real remote sensing data support our claim that high quality reconstruction of multi- and hyperspectral data is obtained using our method. Additionally, the proposed spectral super-resolution algorithm is able to provide vital facilitation in sub-sequent hyperspectral image understanding techniques, such as spectral unmixing and classification. Finally, our scheme can be extended to handle arbitrary low-to-high resolution enhancements by simple modifications of the joint dictionary learning process, and it offers the capability of addressing additional sources of HSI image degradation.

Acknowledgment

This work was partially funded by the DEDALE project, contract no. 665044, within the H2020 Framework Program of the European Commission.

References

[1] Antonio Plaza, Jon Atli Benediktsson, Joseph W Boardman, Jason Brazile, Lorenzo Bruzzone, Gustavo Camps-Valls, Jocelyn Chanussot, Mathieu Fauvel, Paolo Gamba, Anthony Gualtieri, et al., "Recent advances in techniques for hyperspectral image processing," *Remote sensing of environment*, vol. 113, pp. S110–S122, 2009.

[2] Jinha Jung and Melba M Crawford, "Extraction of features from lidar waveform data for characterizing forest structure," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 492–496, 2012.

[3] José M Bioucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gader, and Jocelyn 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, 2012.

[4] Junhwa Chi and Melba M Crawford, "Spectral unmixing-based crop residue estimation using hyperspectral remote sensing data: A case study at purdue university," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2531–2539, 2014.

[5] Paris V Giampouras, Konstantinos E Themelis, Athanasios A Rontogiannis, and Konstantinos D Koutroumbas, "Simultaneously sparse and low-rank abundance matrix estimation for hyperspectral image unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4775–4789, 2016.

[6] Chein-I Chang, *Hyperspectral imaging: techniques for spectral detection and classification*, vol. 1, Springer Science & Business Media, 2003.

[7] Dimitris Manolakis and Gary Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 29–43, 2002.

[8] Li Ma, Melba M Crawford, and Jinwen Tian, "Local manifold learning-based-nearest-neighbor for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 11, pp. 4099–4109, 2010.

[9] Xiuping Jia, Bor-Chen Kuo, and Melba M Crawford, "Feature mining for hyperspectral image classification," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 676–697, 2013.

[10] Spyridoula D Xenaki, Konstantinos D Koutroumbas, Athanasios A Rontogiannis, and Olga A Sykioti, "A new sparsity-aware feature selection method for hyperspectral image clustering," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2015, pp. 445–448.

[11] "Quickbird," https://www.geoimage.com.au/satellite/quickbird.

[12] "Ikonos," https://www.satimagingcorp.com/satellite-sensors/ikonos/.

[13] USGS, "Hyperion," https://eo1.usgs.gov/sensors/hyperion, 2010 (accessed December 7, 2014).

[14] "Modis," https://modis.gsfc.nasa.gov/.

[15] "Smap," https://smap.jpl.nasa.gov/.

[16] Nathan Hagen and Michael W Kudenov, "Review of snapshot spectral imaging technologies," *Optical Engineering*, vol. 52, no. 9, pp. 090901–090901, 2013.

[17] Grigorios Tsagkatakis, Murali Jayapala, Bert Geelen, and Panagiotis Tsakalides, "Non-negative matrix completion for the enhancement of snapshot mosaic multispectral imagery," *Electronic Imaging*, vol. 2016, no. 12, pp. 1–6, 2016.

[18] Grigorios Tsagkatakis and Panagiotis Tsakalides, "Recovery of quantized compressed sensing measurements," in *SPIE/IS&T Electronic Imaging*. International Society for Optics and Photonics, 2015, pp. 940106–940106.

[19] "Sentinel-2," https://sentinel.esa.int/web/sentinel/missions/sentinel-2.

[20] Michael Elad, "Prologue," in *Sparse and Redundant Representations*, pp. 3–15. Springer, 2010.

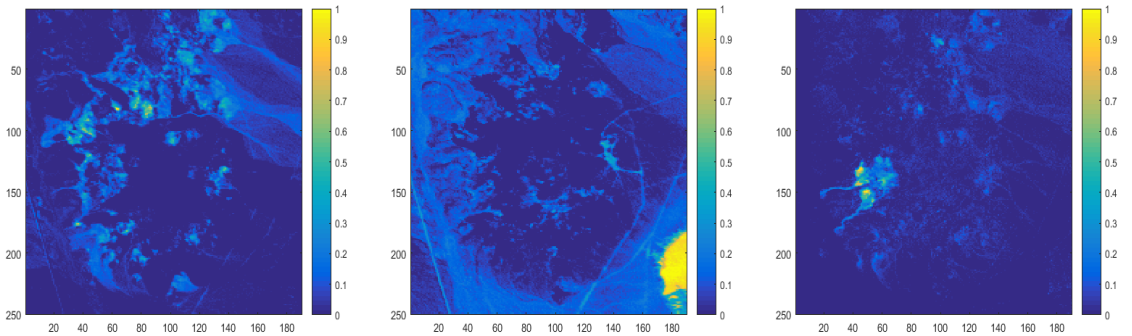
[21] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[22] Yuling Jiao, Qinian Jin, Xiliang Lu, and Weijie Wang, "Alternating direction method of multipliers for linear inverse problems," *arXiv preprint arXiv:1601.02773*, 2016.

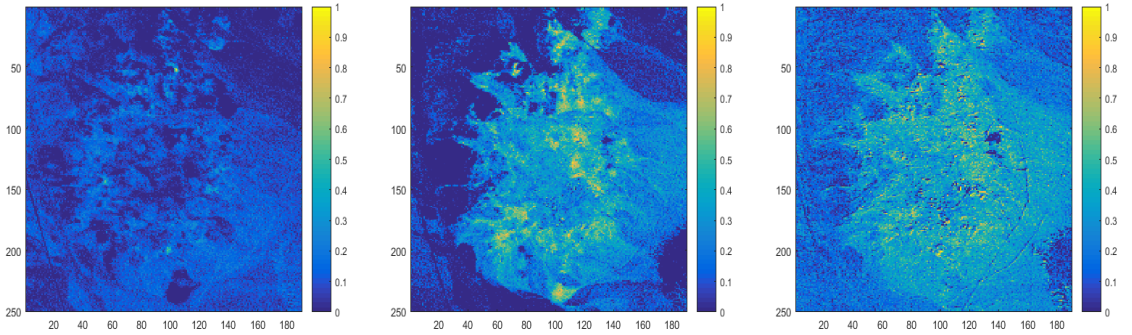
[23] José M Bioucas-Dias, Antonio Plaza, Gustavo Camps-Valls, Paul Scheunders, Nasser Nasrabadi, and Jocelyn Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 2, pp. 6–36, 2013.

[24] Cle Pohl and John L Van Genderen, "Review article multisensor image fusion in remote sensing: concepts, methods and applications," *International journal of remote sensing*, vol. 19, no. 5, pp. 823–854, 1998.

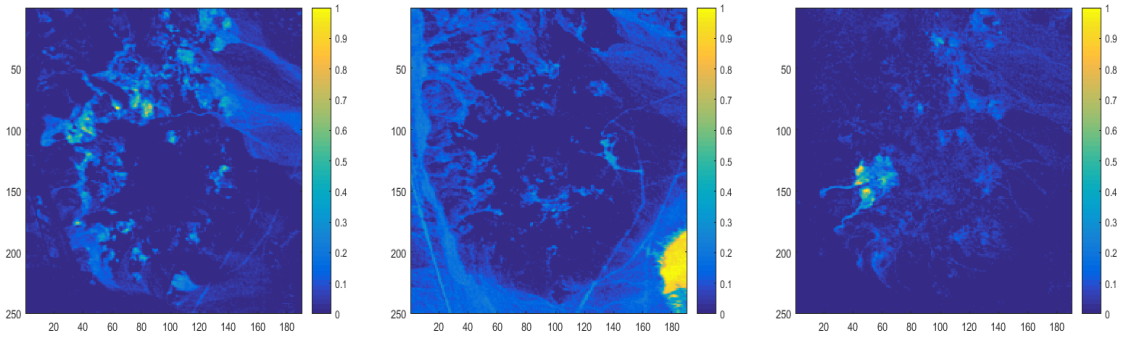
[25] Gemine Vivone, Luciano Alparone, Jocelyn Chanussot, Mauro Dalla Mura, Andrea Garzelli, Giorgio A Licciardi, Rocco Restaino, and



(a) *Ground Truth Endmember #3* (b) *Ground Truth Endmember #5* (c) *Ground Truth Endmember #6*



(d) *Endmember #3 from Low-Resolution* (e) *Endmember #5 from Low-Resolution* (f) *Endmember #6 from Low-Resolution*



(g) *Endmember #3 from SCDL* (h) *Endmember #5 from SCDL* (i) *Endmember #6 from SCDL*

Fig. 22: *In this figure we provide the unmxing results of the AVIRIS Cuprite hyperspectral scene. Specifically, we illustrate the abundance maps for the extracted #3, #5 and #6 endmembers. From top to bottom: In the first row, we depict the ground abundance maps, while in the second row the unmixing results from the low-resolution Cuprite hypercube are presented. Finally, in the third row, we illustrate the unmixing results from the enhanced Cuprite hypercube. As we may notice, the reconstructed abundances using as input the proposed super-resolved hypercube, present high similarity with the accurate, ground truth abundances. Consequently, the proposed SCDL scheme is able to facilitate the crucial problem of hyperspectral image unmixing.*

Lucien Wald, "A critical comparison among pansharpening algorithms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2565–2586, 2015.

[26] Xiao Xiang Zhu and Richard Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 2827–2836, 2013.

[27] Michael T Eismann and Russell C Hardie, "Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 455–465, 2005.

[28] Xiao Xiang Zhu, Claas Grohnfeldt, and Richard Bamler, "Exploiting joint sparsity for pansharpening: The j-sparsefi algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2664–2681, 2016.

[29] Shutao Li and Bin Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 2, pp. 738–746, 2011.

[30] Gemine Vivone, Luciano Alparone, Jocelyn Chanussot, Mauro Dalla Mura, Andrea Garzelli, Giorgio A Licciardi, Rocco Restaino, and Lucien Wald, "A critical comparison among pansharpening algorithms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2565–2586, 2015.

[31] Massimo Selva, Bruno Aiazzi, Francesco Butera, Leandro Chiarantini, and Stefano Baronti, "Hyper-sharpening: A first approach on sim-ga data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 3008–3024, 2015.

[32] Naoto Yokoya, Takehisa Yairi, and Akira 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, 2012.

[33] Qi Wei, José Bioucas-Dias, Nicolas Dobigeon, and Jean-Yves Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3658–3668, 2015.

[34] Naoto Yokoya, Claas Grohnfeldt, and Jocelyn Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent

literature," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 2, pp. 29–56, 2017.

[35] Cheng Jiang, Hongyan Zhang, Huanfeng Shen, and Liangpei Zhang, "Two-step sparse coding for the pan-sharpening of remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 5, pp. 1792–1805, 2014.

[36] Huanfeng Shen, Xiangchao Meng, and Liangpei Zhang, "An integrated framework for the spatio–temporal–spectral fusion of remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7135–7148, 2016.

[37] Kai Zhang, Min Wang, and Shuyuan Yang, "Multispectral and hyperspectral image fusion based on group spectral embedding and low-rank factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 3, pp. 1363–1371, 2017.

[38] Yancong Wei, Qiangqiang Yuan, Huanfeng Shen, and Liangpei Zhang, "Boosting the accuracy of multispectral image pansharpening by learning a deep residual network," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1795–1799, 2017.

[39] Xudong Kang, Shutao Li, and Jón Atli Benediktsson, "Pansharpening with matting model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 5088–5099, 2014.

[40] Xiyan He, Laurent Condat, José M Bioucas-Dias, Jocelyn Chanussot, and Junshi Xia, "A new pansharpening method based on spatial and spectral sparsity priors," *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp. 4160–4174, 2014.

[41] Min Guo, Hongyan Zhang, Jiayi Li, Liangpei Zhang, and Huanfeng Shen, "An online coupled dictionary learning approach for remote sensing image fusion," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1284–1294, 2014.

[42] Silvano Galliani, Charis Lanaras, Dimitrios Marmanis, Emmanuel Baltsavias, and Konrad Schindler, "Learned spectral super-resolution," *arXiv preprint arXiv:1703.09470*, 2017.

[43] Jing Hu, Yunsong Li, and Weiying Xie, "Hyperspectral image super-resolution by spectral difference learning and spatial error correction," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1825–1829, 2017.

[44] Shaohui Mei, Xin Yuan, Jingyu Ji, Yifan Zhang, Shuai Wan, and Qian Du, "Hyperspectral image spatial super-resolution via 3d full convolutional neural network," *Remote Sensing*, vol. 9, no. 11, pp. 1139, 2017.

[45] Jakub Bieniarz, Rupert Müller, Xiao Xiang Zhu, and Peter Reinartz, "Hyperspectral image resolution enhancement based on joint sparsity spectral unmixing," in *2014 IEEE Geoscience and Remote Sensing Symposium*. IEEE, 2014, pp. 2645–2648.

[46] Yongqiang Zhao, Chen Yi, Jingxiang Yang, and Jonathan Cheung-Wai Chan, "Coupled hyperspectral super-resolution and unmixing," in *2014 IEEE Geoscience and Remote Sensing Symposium*. IEEE, 2014, pp. 2641–2644.

[47] Alp Ertürk, Mehmet Kemal Güllü, Davut Çeşmeci, Deniz Gerçek, and Sarp Ertürk, "Spatial resolution enhancement of hyperspectral images using unmixing and binary particle swarm optimization," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 12, pp. 2100–2104, 2014.

[48] Weisheng Dong, Fazuo Fu, Guangming Shi, Xun Cao, Jinjian Wu, Guangyu Li, and Xin Li, "Hyperspectral image super-resolution via non-negative structured sparse representation," *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2337–2352, 2016.

[49] Naveed Akhtar, Faisal Shafait, and Ajmal Mian, "Bayesian sparse representation for hyperspectral image super resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3631–3640.

[50] Haitao Yin, Shutao Li, and Leyuan Fang, "Simultaneous image fusion and super-resolution using sparse representation," *Information Fusion*, vol. 14, no. 3, pp. 229–240, 2013.

[51] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma, "Image super-resolution via sparse representation," *IEEE transactions on image processing*, vol. 19, no. 11, pp. 2861–2873, 2010.

[52] Jianchao Yang, Zhaowen Wang, Zhe Lin, Xianbiao Shu, and Thomas Huang, "Bilevel sparse coding for coupled feature spaces," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 2360–2367.

[53] Li He, Hairong Qi, and Russell Zaretzki, "Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 345–352.

[54] Konstantina Fotiadou, Grigorios Tsagkatakis, Bruno Moraes, Filipe Abdalla, and Panagiotis Tsakalides, "Denoising galaxy spectra with coupled dictionary learning," in *2017 25th European Signal Processing Conference (EUSIPCO)*.

[55] Adam S Charles, Bruno A Olshausen, and Christopher J Rozell, "Learning sparse codes for hyperspectral imagery," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 963–978, 2011.

[56] Adam S Charles and Christopher J Rozell, "Spectral superresolution of hyperspectral imagery using reweighted spatial filtering," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 3, pp. 602–606, 2014.

[57] AS Charles, CJ Rozell, and NB Tufillaro, "Sparsity based spectral super-resolution and applications to ocean water color," *IGARSS 2014 Technical Abstract*, 2014.

[58] Konstantina Fotiadou, Grigorios Tsagkatakis, and Panagiotis Tsakalides, "Multi-source image enhancement via coupled dictionary learning," in *Signal Processing with Adaptive Sparse Structured Representations Workshop (SPARS), 2017*.

[59] Konstantina Fotiadou, Grigorios Tsagkatakis, and Panagiotis Tsakalides, "Spectral resolution enhancement of hyperspectral images via sparse representations," in *Proceedings) Proc. Computational Imaging, IS&T Int. Symposium on Electronic Imaging*, 2016.

[60] Lu Yuan, Jian Sun, Long Quan, and Heung-Yeung Shum, "Image deblurring with blurred/noisy image pairs," in *ACM Transactions on Graphics (TOG)*. ACM, 2007, vol. 26, p. 1.

[61] Konstantina Fotiadou, Grigorios Tsagkatakis, and Panagiotis Tsakalides, "Low light image enhancement via sparse representations," in *International Conference Image Analysis and Recognition*. Springer, 2014, pp. 84–93.

[62] Michal Aharon, Michael Elad, and Alfred Bruckstein, "K-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[63] Yusheng Li, Xinchang Xie, and Zhouwang Yang, "Alternating direction method of multipliers for solving dictionary learning models," *Communications in Mathematics and Statistics*, vol. 3, no. 1, pp. 37–55, 2015.

[64] Mingyi Hong, Zhi-Quan Luo, and Meisam Razaviyayn, "Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems," *SIAM Journal on Optimization*, vol. 26, no. 1, pp. 337–364, 2016.

[65] "Landsat-8," https://landsat.usgs.gov/landsat-8.

[66] Arvind Kumar Singh, HV Kumar, GR Kadambi, JK Kishore, J Shuttleworth, and J Manikandan, "Quality metrics evaluation of hyperspectral images," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, no. 8, pp. 1221, 2014.

[67] Roberta H Yuhas, Alexander FH Goetz, and Joe W Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm," 1992.

[68] O Abilio De Carvalho and Paulo Roberto Meneses, "Spectral correlation mapper (scm): an improvement on the spectral angle mapper (sam)," in *Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18*. JPL Publication Pasadena, CA, 2000, vol. 9.

[69] Jack O'Sullivan, Paul R Hoy, and Harvey N Rutt, "An extended spectral angle map for hyperspectral and multispectral imaging," in *CLEO: Science and Innovations*. Optical Society of America, 2011, p. CMG2.

[70] José M Bioucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gader, and Jocelyn 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, 2012.

[71] Raúl Guerra, Lucana Santos, Sebastián López, and Roberto Sarmiento, "A new fast algorithm for linearly unmixing hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 12, pp. 6752–6765, 2015.

[72] "Aviris," https://aviris.jpl.nasa.gov/.

APPENDIX

Derivations of the individual sub-problems for the SCDL-ADMM based dictionary learning scheme, as described in Section IV.

- *Sub-problem* $\mathbf{W}_h$

$\mathbf{W}_h^\star = \underset{\mathbf{W}_h}{\operatorname{argmin}}\mathcal{L} \Leftrightarrow$
$\nabla_{\mathbf{W}_h}\Big(\frac{1}{2}||\mathbf{S}_h - \mathbf{D}_h\mathbf{W}_h||_2^2 + < Y_1, \mathbf{P} - \mathbf{W}_h > +$
$< Y_3, \mathbf{W}_h - \mathbf{W}_\ell > + \frac{c_1}{2}||\mathbf{P} - \mathbf{W}_h||_2^2 + \frac{c_3}{2}||\mathbf{W}_h - \mathbf{W}_\ell||_2^2\Big)$
$= -\mathbf{D}_h^T \cdot (\mathbf{S_h} - \mathbf{D}_h\mathbf{W}_h) - Y_1 + Y_3 - c_1 \cdot (\mathbf{P} - \mathbf{W}_h) +$
$c_3 \cdot (\mathbf{W}_h - \mathbf{W}_\ell)$

Setting $\nabla_{\mathbf{W}_h}\mathcal{L} = 0 \Leftrightarrow$

$-\mathbf{D}_h \cdot \mathbf{S}_h + \mathbf{D}_h^T \cdot \mathbf{D}_h \cdot \mathbf{W}_h - Y_1 + Y_3 - c_1 \cdot \mathbf{P} + c_1 \cdot \mathbf{W}_h +$
$c_3 \cdot \mathbf{W}_h - c_3 \cdot \mathbf{W}_\ell = 0 \Leftrightarrow$
$(\mathbf{D}_h^T \cdot \mathbf{D}_h + c_1 \cdot I + c_3 \cdot I) \cdot \mathbf{W}_h = \mathbf{D}_h^T \cdot \mathbf{S}_h + Y_1 - Y_3 +$
$c_1 \cdot \mathbf{P} + c_3 \cdot \mathbf{W}_\ell \Leftrightarrow$
$\mathbf{W}_h^\star = (\mathbf{D}_h^T \cdot \mathbf{D}_h + c_1 \cdot I + c_3 \cdot I)^{-1} \cdot (\mathbf{D}_h^T \cdot \mathbf{S}_h + Y_1 - Y_3 +$
$c_1 \cdot \mathbf{P} + c_3 \cdot \mathbf{W}_\ell)$

- *Sub-problem* $\mathbf{W}_\ell$

$\mathbf{W}_\ell = \underset{\mathbf{W}_\ell}{\operatorname{argmin}}\mathcal{L} \Leftrightarrow$
$\nabla_{\mathbf{W}_l}\Big(\frac{1}{2} \cdot ||\mathbf{S}_\ell - \mathbf{D}_\ell\mathbf{W}_\ell||_2^2 + < Y_2, \mathbf{Q} - \mathbf{W}_\ell > +$
$< Y_3, \mathbf{W}_h - \mathbf{W}_\ell > + \frac{c_2}{2}||\mathbf{Q} - \mathbf{W}_\ell||_2^2 + \frac{c_3}{2}||\mathbf{W}_h - \mathbf{W}_\ell||_2^2\Big)$
$= -\mathbf{D}_\ell^T \cdot (\mathbf{S}_\ell - \mathbf{D}_\ell\mathbf{W}_\ell) - Y_2 - Y_3 - c_2 \cdot (\mathbf{Q} - \mathbf{W}_\ell) +$
$c_3 \cdot (\mathbf{W}_h - \mathbf{W}_\ell)$

Setting $\nabla_{\mathbf{W}_\ell}\mathcal{L} = 0 \Leftrightarrow$

$-\mathbf{D}_\ell \cdot \mathbf{S}_\ell + \mathbf{D}_\ell^T \cdot \mathbf{D}_\ell \cdot \mathbf{W}_\ell - Y_2 - Y_3 - c_2 \cdot \mathbf{Q} + c_2 \cdot \mathbf{W}_\ell -$
$c_3 \cdot \mathbf{W}_h = 0 \Leftrightarrow$
$(\mathbf{D}_\ell^T \cdot \mathbf{D}_\ell + c_2 \cdot I + c_3 \cdot I) \cdot \mathbf{W}_\ell = \mathbf{D}_\ell^T \cdot \mathbf{S}_\ell + Y_2 + Y_3 +$
$c_2 \cdot \mathbf{Q} + c_3 \cdot \mathbf{W}_h - c_3 \cdot \mathbf{W}_\ell \Leftrightarrow$
$\mathbf{W}_\ell^\star = (\mathbf{D}_\ell^T \cdot \mathbf{D}_\ell + c_2 \cdot I + c_3 \cdot I)^{-1} \cdot (\mathbf{D}_\ell^T \cdot \mathbf{S}_\ell + Y_2 + Y_3 +$
$c_2 \cdot \mathbf{Q} + c_3 \cdot \mathbf{W}_h)$

- *Sub-problem* $\mathbf{P}$

$\mathbf{P}^\star = \underset{\mathbf{P}}{\operatorname{argmin}}\mathcal{L} \Leftrightarrow$
$\nabla_{\mathbf{P}}(\lambda_h||\mathbf{P}||_1 + < Y_1, \mathbf{P} - \mathbf{W}_h > + \frac{c_1}{2}||\mathbf{P} - \mathbf{W}_h||_2^2)$

  - For $\mathbf{P} > 0$,

$$\nabla_{\mathbf{P}}\mathcal{L} = \lambda_h \cdot I + c_1 \cdot (\mathbf{P} - \mathbf{W}_h) + Y_1$$

Setting $\nabla_{\mathbf{P}}\mathcal{L} = 0$,

$$\mathbf{P}^\star = \mathbf{W}_h - \frac{1}{c_1} \cdot (Y_1 + \lambda_h \cdot I)$$

  - For $\mathbf{P} < 0$,

$$\nabla_{\mathbf{P}}\mathcal{L} = -\lambda_h \cdot I + c_1 \cdot (\mathbf{P} - \mathbf{W}_h) + Y_1$$

Setting $\nabla_{\mathbf{P}}\mathcal{L} = 0$,

$$\mathbf{P}^\star = \mathbf{W}_h - \frac{1}{c_1} \cdot (Y_1 - \lambda_h \cdot I)$$

Combining

$$\mathbf{P} > 0 \Leftrightarrow \mathbf{W}_h - \frac{1}{c_1} \cdot Y_1 > \frac{1}{c_1} \cdot \lambda_h \cdot I$$
$$\mathbf{P} < 0 \Leftrightarrow \mathbf{W}_h - \frac{1}{c_1} \cdot Y_1 < -\frac{1}{c_1} \cdot \lambda_h \cdot I$$

we have

$$|\mathbf{W}_h - \frac{1}{c_1} \cdot Y_1| \leq \frac{1}{c_1} \cdot \lambda_h \cdot I.$$

Consequently,

$$\mathbf{P}^\star = S_{\lambda_h}\Big(|\mathbf{W}_h - \frac{Y_1}{c_1}|\Big),$$

where S_{λ_h} denotes the soft-thresholding operator, defined as

$$S_{\lambda_h}(x) = sign(x) \cdot max(|x| - \lambda_h, 0)$$

- *Sub-problem* $\mathbf{Q}$

$$\mathbf{Q}^\star = \underset{\mathbf{Q}}{\operatorname{argmin}}\mathcal{L} \Leftrightarrow$$

  - For $\mathbf{Q} > 0$,

$$\nabla_{\mathbf{Q}}\mathcal{L} = \lambda_\ell \cdot I + Y_2 + c_2 \cdot (\mathbf{Q} - \mathbf{W}_\ell)$$

Setting $\nabla_{\mathbf{Q}}\mathcal{L} = 0$,

$$\mathbf{Q} = \mathbf{W}_\ell - \frac{1}{c_2} \cdot (Y_2 + \lambda_\ell \cdot I)$$

  - For $\mathbf{Q} < 0$,

$$\nabla_{\mathbf{Q}}\mathcal{L} = \lambda_\ell \cdot I + Y_2 + c_2 \cdot (\mathbf{Q} - \mathbf{W}_\ell)$$

Setting $\nabla_{\mathbf{Q}}\mathcal{L} = 0$,

$$\mathbf{Q} = \mathbf{W}_\ell - \frac{1}{c_2} \cdot (Y_2 - \lambda_\ell \cdot I)$$

Combining

$$\mathbf{Q} > 0 \Leftrightarrow \mathbf{W}_\ell - \frac{1}{c_2} \cdot Y_2 > \frac{1}{c_2} \cdot \lambda_\ell \cdot I$$
$$\mathbf{Q} < 0 \Leftrightarrow \mathbf{W}_\ell - \frac{1}{c_2} \cdot Y_2 < -\frac{1}{c_2} \cdot \lambda_\ell \cdot I$$

we have

$$|\mathbf{W}_\ell - \frac{1}{c_2} \cdot Y_2| \leq \frac{1}{c_2} \cdot \lambda_\ell \cdot I.$$

Consequently,

$$\mathbf{Q}^\star = S_{\lambda_\ell}\Big(|\mathbf{W}_\ell - \frac{Y_2}{c_2}|\Big),$$

where S_{λ_ℓ} denotes the soft-thresholding operator, defined as

$$S_{\lambda_\ell}(x) = sign(x) \cdot max(|x| - \lambda_\ell, 0)$$

- *Sub-problem* $\mathbf{D}_h$

$$\mathbf{D}_h^\star = \underset{\mathbf{D}_h}{\operatorname{argmin}}\mathcal{L} \Leftrightarrow$$
$$\nabla_{\mathbf{D}_h}\mathcal{L} = -(\mathbf{S}_h - \mathbf{D}_h \cdot \mathbf{W}_h) \cdot \mathbf{W}_h^T$$

Setting, $\nabla_{\mathbf{D}_h}\mathcal{L} = 0$,

$$\mathbf{D}_h = \frac{\mathbf{S}_h \cdot \mathbf{W}_h^T}{\phi_h + \delta},$$

where $\phi_h = \mathbf{W}_h \cdot \mathbf{W}_h^T$.

- *Sub-problem* $\mathbf{D}_\ell$

$$\mathbf{D}_\ell^\star = \underset{\mathbf{D}_\ell}{\operatorname{argmin}} \mathcal{L} \Leftrightarrow$$
$$\nabla_{\mathbf{D}_\ell}\mathcal{L} = -(\mathbf{S}_\ell - \mathbf{D}_\ell \cdot \mathbf{W}_\ell) \cdot \mathbf{W}_\ell^T$$

Setting, $\nabla_{\mathbf{D}_h}\mathcal{L} = 0$,

$$\mathbf{D}_\ell = \frac{\mathbf{S}_\ell \cdot \mathbf{W}_\ell^T}{\phi_\ell + \delta},$$

where $\phi_h = \mathbf{W}_\ell \cdot \mathbf{W}_\ell^T$.