Spectral Super-Resolution via Coupled Sparse Dictionary Learning

Konstantina Fotiadou, Grigorios Tsagkatakis, and Panagiotis Tsakalides

***Abstract*—High-spectral resolution imaging systems provide critical information enabling a better identification and characterization of the objects in a scene of interest. Nevertheless, multiple factors may impair spectral resolution, as in the case of modern snapshot spectral imagers that associate each "hyperpixel" with a specific spectral band. In this work, we propose 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. The key contribution of this work is a novel coupled sparse dictionary learning model which considers coupled feature spaces, composed of low and high spectral resolution hypercubes, in order to address the spectral super-resolution problem. We formulate our spectral coupled dictionary learning technique within the context of the Alternating Direction Method of Multipliers, optimizing each variable via closed-form expressions. 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.**

I. Introduction

HIGH-resolution remote sensing architectures including Synthetical Aperture Radars [1] and Hyperspectral Imaging (HSI) [2] offer valuable insights regarding the composition of a scene and singicifantly facilitate tasks like object and material recognition [3], spectral unmixing [4]–[6], and region clustering [7]–[11], among others. To achieve this goal, high spatial and spectral resolution imaging systems must capture massive amounts of measurements, encoding the dynamics of spatial and spectral variations of a scene. However, achieving high spatial, spectral, and temporal resolution is extremely challenging, due to multiple architectural constraints and conflicting objectives.

A representative example of this predicament is illustrated in the case of Spectrally Resolvable Detector Array (SRDA) architectures [12], a new generation of *snapshot* spectral imagers which seek to acquire 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 along the spatial dimensions of the detector. This allows for extremely lightweight cameras, capable of simultaneously acquiring tens of spectral bands. Despite the dramatic reduction 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 [13], [14].

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

In addition to SRDA, traditional HSI architectures that rely of filter wheels for examples also face similar problems since the number of acquired spectral bands is directly related to the size and complexity of the system design. As a consequence, snapshot spectral imagers are limited to a relatively small number of spectral bands. In our work, instead of introducing additional hardware components, we propose a novel computational imaging framework to address the aforementioned limitations.

Formally, this paper employs the concept of *spectral super-resolution*, where low and high spectral resolution training examples are used within a computational learning framework to increase the spectral resolution of existing systems. The proposed *Spectral Coupled Dictionary Learning* (SCDL) algorithm capitalizes on the *Sparse Representations* framework [15] and extents it by introducing a *Coupled Dictionary Learning* process, for estimating responses from spectral bands that were not explicitly acquired by the detectors. Furthermore, we solve the SCDL problem by formulating the spectral super-resolution problem within the highly efficient Alternating Direction Method of Multipliers optimization framework.

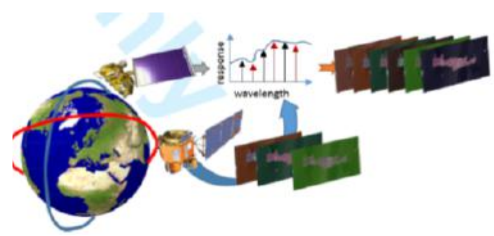


Fig. 1: *A case of spectral super resolution in Earth Observation. A model built on high-low resolution pairs from two instruments is introduced for increasing spectral resolution.*

The particular algorithmic framework can be considered in a wide range of remote sensing applications for Earth Observation. For instance, acquired imagery from low spectral resolution satellites, e.g., MODIS, could be enhanced using images acquired over the same region from higher resolution spectrometers aboard newer platforms, e.g., the EO-1 Hyperion, as shown in Fig. 1. Additionally, such a scheme could be considered for easing communication requirements by training with high-resolution data during the commission phase and by reducing the required bandwidth during normal operation.

The key contributions of this work include:

- 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 sparse dictionary learning architecture, relying on the alternating direction method of multipliers, for efficient identification of the dictionaries;
- the systematic evaluation of the proposed spectral resolution enhancement approach on a variety of challenging real 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 rest of this paper is structured as follows. Section II provides an overview of the related state-of-the-art. Section III presents the spectral super-resolution scheme of multispectral and hyperspectral imagery considered in this work, whereas Section IV exposes the coupled spectral dictionary learning formulation. Section V reports the experimental results, while conclusions and extensions of this work are presented in Section VII.

II. Related Work

In the following section, we overview several representative approaches that address the problems of spatial and spectral resolution enhancement of hyperspectral imagery, as well as techniques for learning coupled 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. This work is an extension of an earlier approach [16], where a sparsity-based architecture was proposed, employing independent dictionaries that model the low and high spectral resolution feature spaces.

A. Hyperspectral Resolution Enhancement

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 [17]. State-of-the-art spatial resolution enhancement approaches may be classified into two representative categories, namely, pan-sharpening and spatio-spectral fusion techniques. 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 [18]–[20]. This is achieved either by replacing the component containing the spatial structure from the HSI image with the panchromatic image [21], or by decomposing the panchromatic image and by re-sampling it to multispectral bands [22]. 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 imager. This requirement restricts the acquisition set-up and it does not consider post-acquisition enhancement.

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. Bieniarz *et al.* [23] enhanced the 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 was proposed in [24], based on a sparse representation in the spatial domaim and a spectral unmixing in the spectral domain.

A significant class of methods considers transferring information between different feature spaces. For instance, Yang *et al.* [25] solved 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 [26] a coupled dictionary learning scheme based on bilevel optimization was proposed and applied on 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. In contrast, He *at al.* [27] propose a beta process based coupled dictionary learning approach, by learning sparse representations with the same sparsity measure, but with different values in the coupled feature spaces.

Recently, Guo *et al.* [28] 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. Contrary to the aforementioned technique, in Section IV we propose a novel scheme that efficiently learns coupled feature spaces, overcoming the limitations arising from independent dictionary learning.

Erturk *et al.* [29] proposed 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. Approaches based on Sparse Representations (SR) have also been considered for spatio-spectral fusion. In [30], Dong *et al.* proposed a non-negative sparsity-based hyperspectral super-resolution technique, combining a low-resolution hyperspectral image with a high-resolution RGB image, where a single dictionary learning scheme is employed for modeling the relations between the low-spectral resolution HSI and the corresponding high resolution RGB images. Additionally, the authors in [31] proposed a Bayesian sparse coding scheme, utilizing a Bayesian non-parametric dictionary learning, in order to enhance the spatial variation of multi- and hyperspectral imagery.

In contrast to spatial super-resolution, enhancing the spectral dimension of HSI scenes has drawn little attention. The work most closely related to the proposed method was presented in [32] where Charles and Rozell introduce a sparsity-based spectral super-resolution approach of hyperspectral images by learning a dictionary of spectral signatures that decomposes the spectral response of each “hyper-pixel”. Specifically, they enhance the spectral dimension of multispectral to hyperspectral level by learning an approximation to the data manifold.

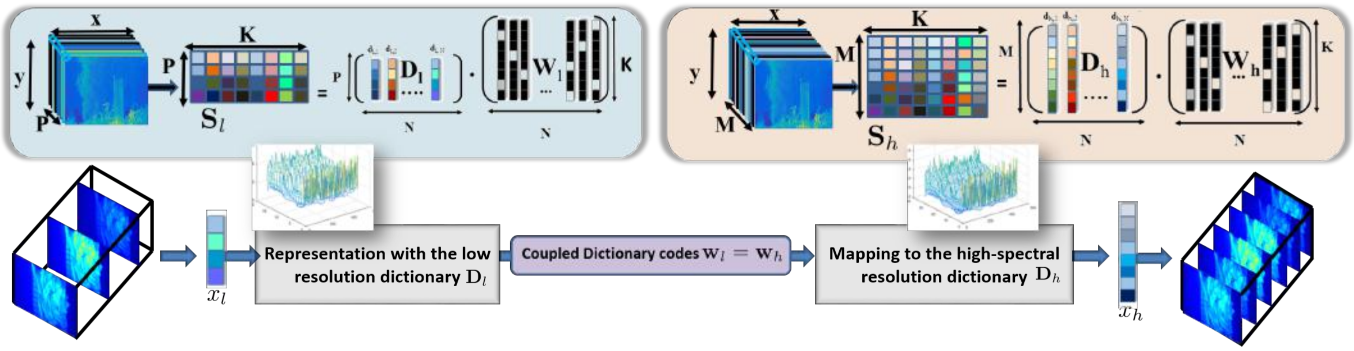


Fig. 2: *SCDL System Block Diagram: The system takes as input a hypercube acquired with a limited number of spectral bands and produces an estimate of an extended spatio-spectral hypercube. During the training phase, multiple high and low-spectral resolution "hyper-pixels" are extracted from training hypercubes. Given these "hyper-pixels" pairs, a coupled sparse 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 subsequently combined with the high resolution dictionary for producing the final estimates.*

As an extension, the same authors introduced in [33] a re-weighted ℓ_1 spatial filtering technique that improves spectral super-resolution.

Another spectral resolution enhancement technique is demonstrated in [34], where 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-spectral resolution spectra are expressed as the linear combination between a blurring matrix and the measured spectra. This problem is solved via a sparse decomposition technique.

As a last point, we must note that over the last years, multiple techniques exploiting the low-rank Matrix Completion framework, a generalization of the SR framework, have been introduced for super-resolving low spatial resolution HSI scenes. For example, a novel approach was proposed in [13], where the authors estimate a high spatial and spectral resolution hypercube from undersampled snapshot mosaic imagery [35]. Although we consider such datasets, the proposed method is applicable to arbitrary low-high resolution pairs.

III. 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 [15]. According to the SR framework, various spectral resolution "hyper-pixels" can be represented as sparse linear combinations of elements from learned over-complete dictionaries. Traditional approaches consider a set of low and high-spectral resolution hyperspectral image pairs and assume that these images are generated by the same statistical process under 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. Each low-spectral resolution "hyper-pixel" $\mathbf{s}_l \in \mathbb{R}^P$ can thus be expressed as a sparse linear combination, encoded in $\mathbf{w} \in \mathbb{R}^N$, 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:

$$\mathbf{s}_l = \mathbf{D}_l \mathbf{w}. \tag{1}$$

Recovery of the sparse coding vector $\mathbf{w} \in \mathbb{R}^N$ is accomplished by solving the following minimization problem:

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

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 alleviated 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. The optimization problem is therefore formulated as:

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

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:

$$\mathbf{s}_h = \mathbf{D}_h \mathbf{w}^\star. \tag{4}$$

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 the sufficient sparsity measure for the sparse coding vector $\mathbf{w}$ and 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. A straightforward strategy to create these dictionaries is to randomly sample multiple registered "hyper-pixels" extracted from corresponding low and high-spectral resolution training scenes and to use this random selection as the sparsifying dictionary. This strategy however 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 like the K-SVD [36] can be considered [16].

IV. Coupled Sparse Dictionary Learning

The proposed SCDL algorithm relies on generating coupled sparse dictionaries which jointly encoding two coupled feature spaces, the observation low-spectral resolution $\mathbf{S}_l \in \mathbb{R}^{P\times K}$, and the latent high-spectral resolution $\mathbf{S}_h \in \mathbb{R}^{M\times K}$, where the signals have sparse representations in terms of the trained dictionaries. 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:

$$\begin{aligned}&\underset{\mathbf{D}_h,\mathbf{D}_\ell,\mathbf{W}_h,\mathbf{W}_\ell}{\text{argmin}} \; ||\mathbf{S}_h-\mathbf{D}_h\mathbf{W}_h||_F^2+||\mathbf{S}_l-\mathbf{D}_\ell\mathbf{W}_\ell||_F^2+\\ &\lambda_h||\mathbf{W}_h||_1+\lambda_\ell\;||\mathbf{W}_\ell||_1, \;\; \text{subject to } \mathbf{W}_h=\mathbf{W}_\ell,\\ &||\mathbf{D}_h(:,i)||_2\le 1,\; ||\mathbf{D}_\ell(:,i)||_2\le 1 \end{aligned} \tag{5}$$

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:

$$\begin{aligned}&\underset{\mathbf{D},\mathbf{W}}{\text{argmin}}||\bar{\mathbf{S}}-\bar{\mathbf{D}}\mathbf{W}||_F+\lambda||\mathbf{W}||_1\\ &\text{subject to } ||\bar{\mathbf{D}}(:,j)||_2^2\le 1, \quad j=\{1,...,K\},\end{aligned} \tag{6}$$

where $\bar{\mathbf{S}} = \begin{bmatrix}\mathbf{S}_h\\ \mathbf{S}_l\end{bmatrix}$, $\bar{\mathbf{D}} = \begin{bmatrix}\mathbf{D}_h\\ \mathbf{D}_l\end{bmatrix}$, 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 (6) is converted into a standard, single sparse decomposition problem, that can be efficiently solved via existing dictionary learning algorithms, such as the K-SVD [36]. However, such a strategy is optimal only in the concatenated feature space, and not in the individual feature spaces of $\mathbf{S}_h$ and $\mathbf{S}_l$. 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 random collection of signal-pairs or on 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 exists 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 propose a computationally efficient *coupled dictionary learning* technique, based on the Alternating Direction Method of Multipliers (ADMM) [37]–[40] 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, $\mathbf{S}_l = \{\mathbf{s}_l\}_{i=1}^N$, 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 $\mathbf{D}_l$ with their corresponding sparse codes $\mathbf{W}_h$ and $\mathbf{W}_l$, by solving the following sparse matrix decomposition problem:

$$\begin{aligned}(\mathbf{D}_h,\mathbf{W}_h)&=\text{argmin}||\mathbf{D}_h\mathbf{W}_h-\mathbf{S}_h||_F+\lambda_h||\mathbf{W}_h||_1\\ (\mathbf{D}_l,\mathbf{W}_l)&=\text{argmin}||\mathbf{D}_l\mathbf{W}_l-\mathbf{S}_l||_F+\lambda_l||\mathbf{W}_l||_1,\\ ||\mathbf{D}_h(:,j)||_2^2&\le 1,\; ||\mathbf{D}_l(:,j)||_2^2\le 1, \text{ and } \mathbf{W}_h=\mathbf{W}_l\end{aligned} \tag{7}$$

To apply the ADMM scheme in our spectral dictionary learning procedure, we reformulate the ℓ_1-minimization problem in (7) as

$$\begin{aligned}&\min_{\mathbf{D}_h,\mathbf{W}_h,\mathbf{D}_l,\mathbf{W}_l}||\mathbf{S}_h-\mathbf{D}_h\mathbf{W}_h||_F^2+||\mathbf{S}_l-\mathbf{D}_l\mathbf{W}_l||_F^2\\ &\qquad\qquad+\lambda_l||\mathbf{Q}||_1+\lambda_h||\mathbf{P}||_1\\ &\text{subject to } \mathbf{P}-\mathbf{W}_h=0, \mathbf{Q}-\mathbf{W}_l=0, \mathbf{W}_h-\mathbf{W}_l=0,\\ &\qquad\qquad ||\mathbf{D}_h(:,i)||_2\le 1, ||\mathbf{D}_l(:,i)||_2\le 1\end{aligned} \tag{8}$$

The ADMM scheme takes into account the separate structure of each variable posed in (8), relying on the minimization of its augmented Lagrangian function:

$$\begin{aligned}&L(\mathbf{D}_h,\mathbf{D}_l,\mathbf{W}_h,\mathbf{W}_l,\mathbf{P},\mathbf{Q},Y_1,Y_2,Y_3)=\frac{1}{2}||\mathbf{D}_h\mathbf{W}_h-\mathbf{S}_h||_F^2+\\ &\frac{1}{2}||\mathbf{D}_\ell\mathbf{W}_\ell-\mathbf{S}_\ell||_F^2+\lambda_h||\mathbf{P}||_1+\lambda_\ell||\mathbf{Q}||_1+<Y_1,\mathbf{P}-\mathbf{W}_h>\\ &+<Y_2,\mathbf{Q}-\mathbf{W}_\ell>+<Y_3,\mathbf{W}_h-\mathbf{W}_\ell>+\frac{c_1}{2}||\mathbf{P}-\mathbf{W}_h||_F^2+\\ &\frac{c_2}{2}||\mathbf{Q}-\mathbf{W}_\ell||_F^2+\frac{c_3}{2}||\mathbf{W}_h-\mathbf{W}_\ell||_F^2\end{aligned} \tag{9}$$

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 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:

$$\begin{aligned}\mathbf{W}_h&=\underset{\mathbf{W}_h}{\text{argmin}}L=\nabla_{\mathbf{W}_h}L\\ \mathbf{W}_l&=\underset{\mathbf{W}_l}{\text{argmin}}L=\nabla_{\mathbf{W}_l}L\end{aligned} \tag{10}$$

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

$$\begin{aligned} &\mathbf{W}_h = (\mathbf{D}_h^T \cdot \mathbf{D}_h + c_1 \cdot I + c_3 \cdot I)^{-1} \cdot (\mathbf{D}_h^T \cdot \mathbf{S}_h + \\ &Y_1 - Y_3 + c_1 \cdot \mathbf{P} + c_3 \cdot \mathbf{W}_l) \\ &\mathbf{W}_l = (\mathbf{D}_l^T \cdot \mathbf{D}_l + c_2 \cdot I + c_3 \cdot I)^{-1} \cdot (\mathbf{D}_l^T \cdot \mathbf{S}_l + \\ &Y_2 + Y_3 + c_2 \cdot \mathbf{Q} + c_3 \cdot \mathbf{W}_h) \end{aligned} \tag{11}$$

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

$$\begin{aligned} &\nabla_{\mathbf{P}} \Big(\lambda_h ||\mathbf{P}||_1 + < Y_1, \mathbf{P} - \mathbf{W}_h > + \frac{c_1}{2} ||\mathbf{P} - \mathbf{W}_h||_F^2 \Big) \\ &\nabla_{\mathbf{Q}} \Big(\lambda_l ||\mathbf{Q}||_1 + < Y_2, \mathbf{Q} - \mathbf{W}_l > + \frac{c_2}{2} ||\mathbf{Q} - \mathbf{W}_l||_F^2 \Big) \end{aligned} \tag{12}$$

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

$$\begin{aligned} \mathbf{P} &= S_{\lambda_h} \Big(\Big| \mathbf{W}_h - \frac{Y_1}{c_1} \Big| \Big) \\ \mathbf{Q} &= S_{\lambda_l} \Big(\Big| \mathbf{W}_l - \frac{Y_2}{c_2} \Big| \Big), \end{aligned} \tag{13}$$

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

$$S_\lambda(x) = sign(x) \cdot max(|x| - \lambda, 0), \tag{14}$$

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

- *Sub-problems* D_h and D_l
  For a fixed set of $\mathbf{W}_h$, $\mathbf{W}_l$, $\mathbf{P}$ and $\mathbf{Q}$, the dictionaries $\mathbf{D}_h$ and $\mathbf{D}_l$ can be updated as:

$$\mathbf{D}_h = \underset{\mathbf{D}_h}{\operatorname{argmin}} L = \nabla_{\mathbf{D}_h} L \tag{15}$$

$$\mathbf{D}_l = \underset{\mathbf{D}_l}{\operatorname{argmin}} L = \nabla_{\mathbf{D}_l} L \Leftrightarrow$$

$$\begin{aligned} &\nabla_{\mathbf{D}_h} \Big(\frac{1}{2} ||\mathbf{S}_h - \mathbf{D}_h \mathbf{W}_h||_F^2 \Big) = -\mathbf{W}_h^T (\mathbf{S}_h - \mathbf{D}_h \mathbf{W}_h) \\ &\nabla_{\mathbf{D}_l} \Big(\frac{1}{2} ||\mathbf{S}_l - \mathbf{D}_l \mathbf{W}_l||_F^2 \Big) = -\mathbf{W}_l^T (\mathbf{S}_l - \mathbf{D}_l \mathbf{W}_l) \end{aligned} \tag{16}$$

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

$$\begin{aligned} \phi_h &= \mathbf{W}_h(j,:) \cdot \mathbf{W}_h(j,:)^T \\ \phi_l &= \mathbf{W}_l(j,:) \cdot \mathbf{W}_l(j,:)^T, \end{aligned} \tag{17}$$

and

$$\begin{aligned} \mathbf{D}_h^{(k+1)}(:,j) &= \mathbf{D}_h(:,j)^{(k)}(:,j) + \frac{\mathbf{S}_h \cdot \mathbf{W}_h(j,:)}{\phi_h + \delta} \\ \mathbf{D}_l^{(k+1)}(:,j) &= \mathbf{D}_l(:,j)^{(k)}(:,j) + \frac{\mathbf{S}_l \cdot \mathbf{W}_l(j,:)}{\phi_l + \delta} \end{aligned} \tag{18}$$

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

$$\begin{aligned} Y_1^{(k+1)} &= Y_1^{(k)} + c_1(\mathbf{P} - \mathbf{W}_h) \\ Y_2^{(k+1)} &= Y_2^{(k)} + c_2(\mathbf{Q} - \mathbf{W}_l) \\ Y_3^{(k+1)} &= Y_3^{(k)} + c_3(\mathbf{W}_h - \mathbf{W}_l) \end{aligned} \tag{19}$$

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 $\mathbf{W}_l$ via Eq. (11)
2) Update $\mathbf{P}$ and $\mathbf{Q}$ via Eq. (13)
3) **for** $j = 1, \cdots, N$ do
   - Update ϕ_h and ϕ_l via Eq. (17)
   - Update the two dictionaries $\mathbf{D}_h$ and $\mathbf{D}_l$ column by column via Eq. (18)

   end
   - Normalize $\mathbf{D}_h$ and $\mathbf{D}_l$ between $[0, 1]$
   - Update Lagrange multiplier matrices Y_1, Y_2 and Y_3 via Eq. (19)

end

V. 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 hypercube. The performance is quantified using the following challenging datasets: (a) the multispectral CAVE indoors image data-base [41], (b) the outdoors snapshot spectral dataset acquired by a snapshot spectral camera equipped with the IMEC's Spectrally Resolvable Detector Array [42]–[44], and (c) the EO-1 NASA's Hyperion satellite hyperspectral Earth Observation scenes [45].

The CAVE database includes 32 multispectral images acquired indoors under controlled illumination conditions. The acquired images have a spatial resolution of 512×512 pixels, resolving 31 spectral bands in the 400 to 700nm range. For the EO-1 satellite data, we conducted experiments on data acquired by NASA's Hyperion hyperspectral instrument. Due to its high spectral coverage, Hyperion scenes have been widely utilized in the remote sensing community for classification and spectral unmixing purposes. We considered hyperspectral scenes of the Hawaii island, acquired on August 30, 2015, and utilized 67 spectral bands in the visible and near infrared spectrum range, from 436.9 to 833.83nm.

Finally, we utilized hyperspectral data acquired by IMEC's snapshot mosaic sensors. These flexible sensors optically multiplex the three-dimensional spatio-spectral information on a two-dimensional CMOS detector array, where a layer of Faby-Perot spectral filters is deposited on top of the detector array. The hyperspectral data is initially acquired in the form of two-dimensional mosaic images. In order to generate the 3D hypercubes, the spectral components are properly rearranged into separate spectral bands. In our experiments, we utilize the 5×5 snapshot mosaic hyperspectral sensor, revealing 25 bands in the VNIR spectrum range from 600 to 875nm.

A. Implementation and evaluation metrics

Regarding the dictionary training phase, three pairs of low and high spectral resolution dictionaries were prepared, one for each sensor data, while for all three cases, we utilized 10 training hypercubes, from which 100.000 training hyper-pixels were randomly extracted. In order to generate the corresponding low-spectral resolution hypercubes, the high-spectral resolution training hypercubes were downsampled along the spectral dimension. We experimented with spectral sub-sampling factors of $2, 3$, and 4, corresponding to 16, 11, and 8 input spectral bands for Columbia; 34, 23, and 17 for Hyperion; and 13, 9, and 7 for the IMEC dataset. The number of the representative dictionary atoms that we utilized in the proposed SCDL coupled dictionary learning scheme was set to 512, balancing the computational cost with the robustness of the representation.

To validate the quality of the reconstructed hypercubes, we employ the *Peak Signal to Noise Ratio* (PSNR) [46] given by:

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

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

$$\text{MSE}(x,y,\lambda) = \frac{\sum_{x,y,\lambda}\left[\mathbf{S}_{h_{(x,y,\lambda)}} - \mathbf{S}_{l_{(x,y,\lambda)}}\right]^2}{n_x, n_y, \lambda}, \quad (20)$$

where x and y denote the spatial dimensions of the input and the synthesized images $\mathbf{S}_l$ and $\mathbf{S}_h$.

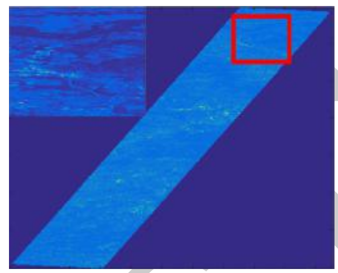
Additionally, each recovered spectral band is compared against the corresponding ground truth spectral band in terms of the *Structural Similarity Index Metric* [47], a psychophysically modeled error metric defined as:

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c1)\cdot(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)\cdot(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (21)$$

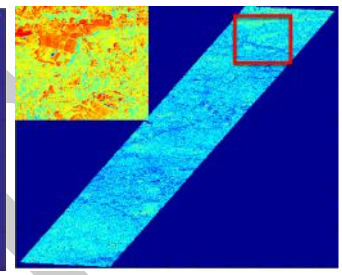
where μ and σ stand for the mean value and the standard deviation, respectively. The reported figures for PNSR and SSIM correspond to the average value over all spectral bands.

B. Experimental Results

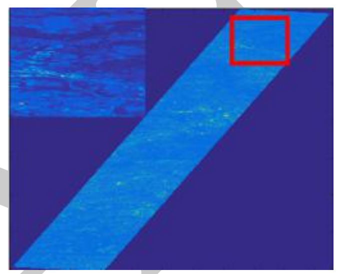
In order to validate the merits of the proposed spectral super-resolution scheme, we first compare the synthesized 3D hypercubes against the ground truth cubes, and against several state-of-art techniques, namely: the simplistic scheme of cubic interpolation among the available spectral bands, the sparse-based scheme of spectral resolution enhancement of hyperspectral imagery using K-SVD dictionary learning [16], and the ℓ_1 spatial filtering approach [32], [33]. In order to achieve a fair comparison with the K-SVD dictionary learning technique, we utilize the same number of atoms for dictionary learning and the same sparsity constraints, while for the reweighted ℓ_1 spatial filtering scheme (RWL1-SF) [32] we fix the parameters on their proposed default settings.



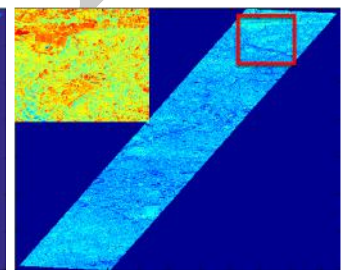
(a) *Ground Truth* 20^{th}*, SSIM: 0.94*



(b) *Ground Truth* 54^{th}*, SSIM: 0.91*



(c) 20^{th} *Band* $\times 4$*, SSIM:0.99*

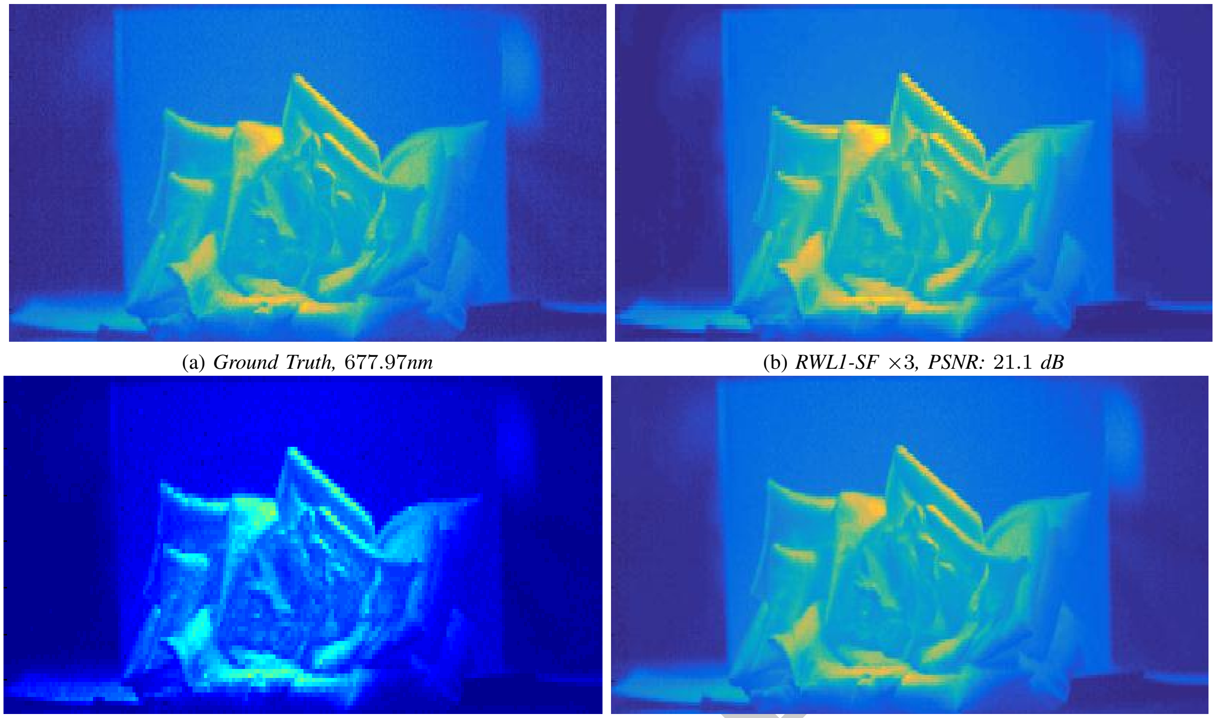


(d) 54^{th} *Band* $\times 4$*, SSIM:0.98*

Fig. 3: *Hyperion spectral bands reconstruction: The full spectrum is composed of 67 bands in the VIS-NIR region, while the sub-sampling factor is set to 4, thus the full resolution hypercube is estimated from 17 input spectral bands. We observe that under real life conditions, the proposed scheme produces a significant quality improvement operating in satellite hyperspectral imagery.*

Figures 3, 4, and 5 showcase representative bands from reconstructed hypercubes obtained by the SCDL method, applied on Hyperion's Hawaii hypespectral scene, on CAVE's synthetic multispectral flower scene, and on IMEC's 5×5 snapshot mosaic roof imagery, respectively. In Figure 3, we subsample the hyperspectral scene by a spectral factor of 4 and we reconstruct the full spectrum composed of 67 spectral observations. In Figures 4 and 5 we downscale the hypercubes by a spectral factor of 2 and we synthesize the full spectrum composed of 25 and 39 spectral bands, respectively. We observe that the reconstructed spectral bands present high similarity and faithfully preserve important image features. For example, the spatial features of the images, like the high frequency content of the flower data in Figure 4, are correctly synthesized while in Figure 5, one can easily notice how different image regions, corresponding to different material, are reliably estimated. The PSNR errors of the SCDL scheme for the recovery of the complete three-dimensional hypercubes are **46.8** db for the Hyperion, **31.6** db for IMEC and **35.8** db for CAVE.

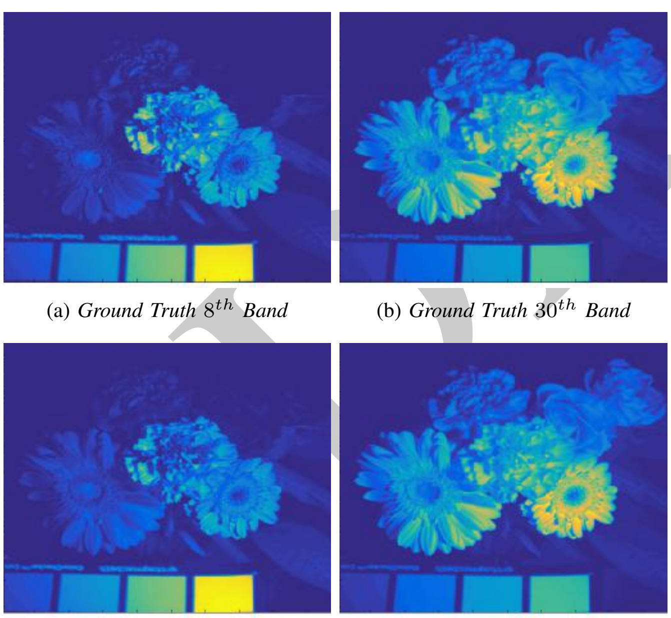
In order to appreciate the quality of reconstruction, we present in Figure 6comparative results with state-of-the-art



(a) *Ground Truth,* 677.97*nm* (b) *RWL1-SF* $\times 3$*, PSNR:* 21.1 *dB*

(c) *SSR-K-SVD* $\times 3$*, PSNR:* 29.4 *dB* (d) *SCDL* $\times 3$*, PSNR:* ***35.1*** *dB*

Fig. 8: *In this experiment we consider the 5th band of the rose hyperspectral scene at* 677.97*nm, acquired by a snapshot spectral camera equipped with IMEC's* 5×5 *snapshot mosaic hyperspectral sensor. The sub-sampling factor is set to* $\times 3$*, while we recover the full hypercube from only* 9 *input spectral observations. Results indicate the SCDL scheme outperforms the comparable state-of-the art approaches, both visually and in terms of error.*



(a) *Ground Truth* 8^{th} *Band* (b) *Ground Truth* 30^{th} *Band*

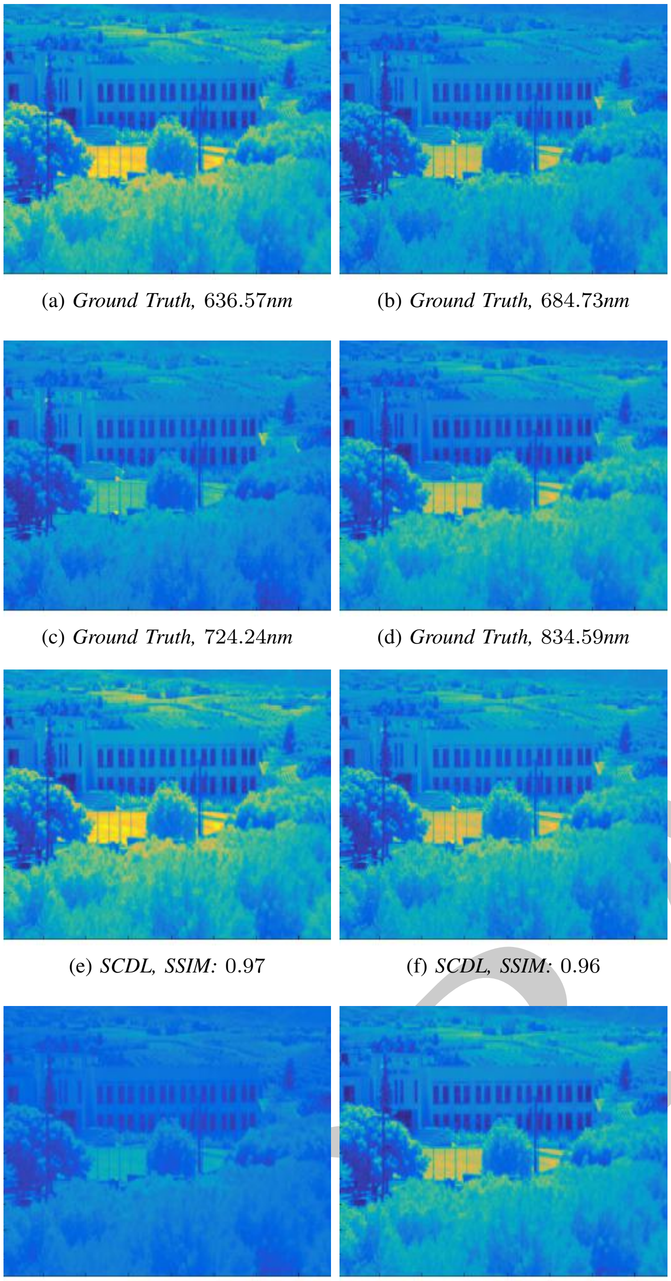
(c) *SCDL* $\times 2$ *SSIM: 0.99* (d) *SCDL* $\times 2$*, SSIM: 0.99*

Fig. 4: *CAVE flower multispectral scene recovery: (a,b) Ground truth spectral bands from the* 31 *band hypercube. (c,d) SCDL reconstructed spectral bands for sub-sampling factor of* 4*, i.e.,* 8 *bands.*

methods applied on the Egyptian statue scene from the CAVE data-base. We set the downsampling factor to $\times 4$, and thus we recover the full spectrum composed of 31 spectral bands, from only 8 input spectral observations. We observe that both the Cubic Interpolation and the reweighted spatial filtering techniques introduce artifacts around the head of the statue and at the text written on the color plate. The K-SVD spectral super-resolution approach performs better compared to cubic and RWL1-SF, but not as well as the proposed SCDL scheme, which exhibits the highest accuracy with the ground three-dimensional hypercube, both visually and quantitatively in terms of the achieved PSNR error metric: the PSNR recovery of the three-dimensional hypercube using the K-SVD spectral resolution enhancement scheme is 41.8 dB, while the PSNR of the SCDL scheme is $\mathbf{45.2}$ dB.

Figure 7 illustrates the performance of the compared techniques when applied on the roof hyperspectral scene, acquired by IMEC's 5×5 snapshot mosaic sensor. Specifically, we depict the spectral band acquired at 669.8nm. In this scenario, the sub-sampling factor is set to 2 and we estimate the 25-band full spectrum from 13 spectral observations. In contrast to the rest of the methods that produce false colouring effects and noisy recoveries, the SCDL algorithm depicts high similarity with the original ground truth spectral data, both visually and quantitatively, in terms of the PSNR error metric.

An indicative set of reconstruction is depicted in Figure 8, where the performance of different methods is evaluated on the *rose* scene dataset. In this experiment, the full 25 spectral bands spectrum is recovered from 9 spectral inputs, and we illustrate the spectral band acquired at 677.97nm. Visual observation reveals that the K-SVD technique introduces severe artifacts at the high spatial frequency regions like the edges of the flower. On the other hand, although the RWL1-SF approach produces a smooth representation of the



(a) *Ground Truth,* 636.57*nm* (b) *Ground Truth,* 684.73*nm*

(c) *Ground Truth,* 724.24*nm* (d) *Ground Truth,* 834.59*nm*

(e) *SCDL, SSIM:* 0.97 (f) *SCDL, SSIM:* 0.96

(g) *SCDL, SSIM:* 0.94 (h) *SCDL, SSIM:* 0.92

Fig. 5: *Roof hyperspectral scene: In this experiment we investigate the performance of our SCDL scheme when applied on hyperspectral data acquired by IMEC's snapshot* 5×5 *sensor. Top row: Original spectral bands. Bottom row: Proposed system's reconstructed spectral bands. The full spectrum is composed of* 25 *bands, while the sub-sampling factor is set to* 2.

spectral representation of the scene, color (spectral) effects are introduced. The SCDL approach provides an accurate and smooth approximation of the ground truth, revealing significant details over all regions of this challenging scene.

Finally, PSNR errors of the comparable techniques applied on several test scenes from CAVE's and IMEC's HSI datasets are provided in Tables I and II, respectively. The results suggest that the proposed spectral super-resolution scheme outperforms all other competing techniques on both test datasets.

(a) *Ground Truth* 23^{th} *Band*

(b) *RWL1-SF* $\times 4$*, PSNR*=26.1 *dB*

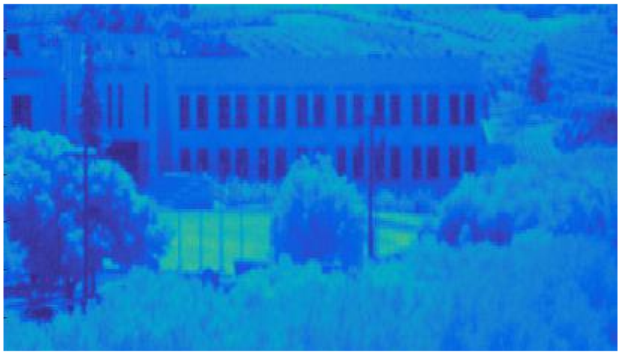
(c) *KSVD-SSR* $\times 4$*, PSNR*=41.8 *dB*

(d) *SCDL* $\times 4$*, PSNR*=45.2 *dB*

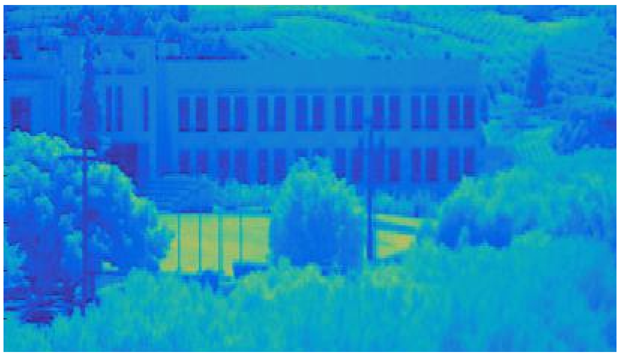
Fig. 6: *CAVE data recovery, Egyptian Statue scene: Comparison with the state-of-the-art approaches for sub-sampling factor* $\times 4$.

C. Sensitivity Analysis

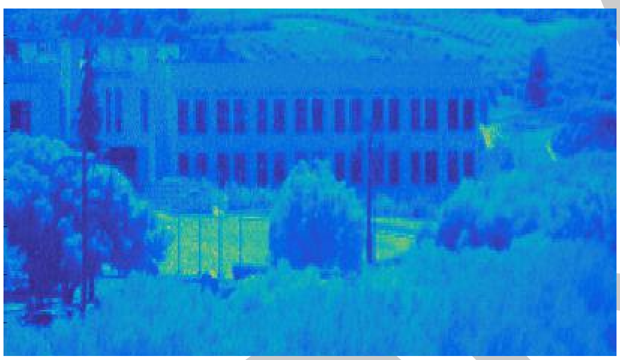
To understand the sensitivity of the algorithm, we evaluate the reconstruction performance of the coupled trained dictionaries over a varying number of training examples. In Figure 9, we provide the PSNR values for the reconstruction of the



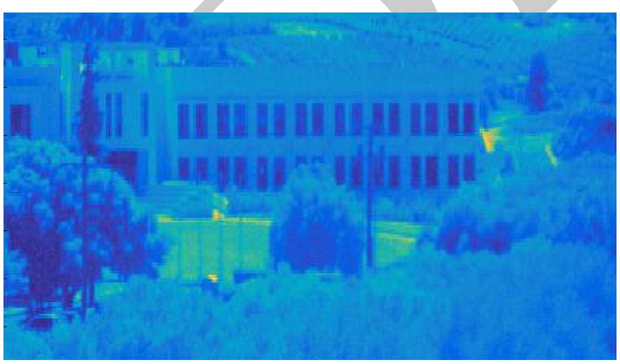
(a) *Ground Truth Band,* $669.81nm$



(b) *RWL1-SF* $\times 2$*, PSNR:* 21.3 *dB*



(c) *SSR-K-SVD* $\times 2$*, PSNR:* 34.7 *dB*



(d) *SCDL* $\times 2$*, PSNR:* ***37.1*** *dB*

Fig. 7: *Roof hyperspectral scene: Comparison with the state-of-the-art. In this experiment we investigate the performance of the SCDL scheme, when applied on snapshot spectral imaging data for down-sampling factor* $\times 2$*, while we recover the full spectrum composed of* 25 *bands from* 13 *input spectral observations.*

Egyptian-statue scene, when the sub-sampling factor is set to $\times 4$, as a function of the number of training examples, i.e., $[10.000, 20.000, 30.000, 50.000, 70.000, 90.000]$. Results indicate that the performance of the SCDL method monotonically increases, based on the amount of the input training examples, however the performance gains are reduced with very large datasets.

Regarding the testing phase, our algorithm by design contains a single parameter, λ, which is responsible for balancing

TABLE I: *CAVE multispectral database: Quantitative results of the proposed SCDL scheme with the state-of-the-art in terms of PSNR error with down-sampling factors of* $\times 2$*,* $\times 3$ *and* $\times 4$*.*

Image	Scale	Cubic	RWL1-SF	SSR-KSVD	SCDL
ballons	2	35.4	25.2	38.0	**40.6**
	3	36.2	18.7	36.8	**37.9**
	4	28.8	22.7	31.6	**37.4**
pompoms	2	32.4	23.0	34.4	**36.5**
	3	32.0	18.9	33.0	**34.9**
	4	27.5	16.6	30.4	**32.4**
flowers	2	37.0	24.1	37.2	**40.1**
	3	35.1	16.0	34.8	**36.2**
	4	33.4	19.5	34.8	**35.8**
egyptian statue	2	43.5	25.0	47.4	**48.7**
	3	42.4	23.8	43.5	**44.5**
	4	36.7	26.0	41.8	**45.2**
stuffed toys	2	34.6	23.1	38.6	**40.6**
	3	32.0	18.8	32.1	**32.6**
	4	27.5	19.1	31.5	**33.4**
average	2	36.6	24.1	36.5	**39.0**
	3	36.9	19.2	36.1	**37.2**
	4	30.8	20.7	34.0	**36.8**

TABLE II: *IMEC's hyperspectral scenes: Quantitative performance evaluation of the SCDL scheme with the state-of-the-art in terms of PSNR error with down-sampling factors of* $\times 2$*,* $\times 3$ *and* $\times 4$

Image	Scale	Cubic	RWL1-SF	SSR-KSVD	SCDL
roof	2	33.5	21.3	34.7	**37.0**
	3	31.9	18.0	32.5	**33.6**
	4	28.4	16.3	30.0	**31.6**
window	2	30.7	23.7	38.5	**40.3**
	3	28.9	22.1	31.9	**32.0**
	4	26.0	20.4	33.6	**34.6**
teddy-bear	2	28.8	20.4	37.5	**41.1**
	3	26.9	17.5	29.6	**32.4**
	4	22.8	14.5	31.3	**32.6**
keys	2	30.5	24.6	37.8	**38.1**
	3	28.4	24.6	30.5	**31.3**
	4	24.4	22.8	30.5	**31.3**
croissant	2	28.7	20.9	36.7	**39.8**
	3	26.8	15.0	31.0	**31.7**
	4	22.5	13.7	31.7	**33.1**
average	2	30.4	26.3	37.0	**39.3**
	3	28.6	19.4	31.1	**32.2**
	4	24.8	17.5	31.4	**32.6**

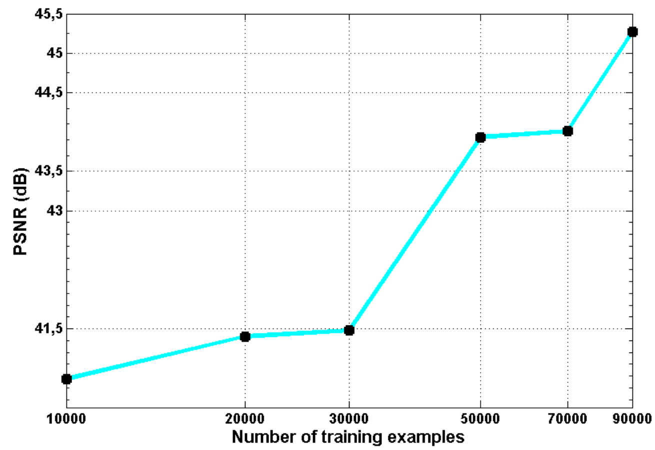


Fig. 9: *CAVE's Egyptian Statue 3D multispectral scene* ($512 \times 512 \times 31$)*: PSNR as a function of the number of training examples. The sub-sampling factor is set to* $\times 4$*.*

the sparsity of our solution with respect to the fidelity of the reconstruction and is the same for both dictionaries.

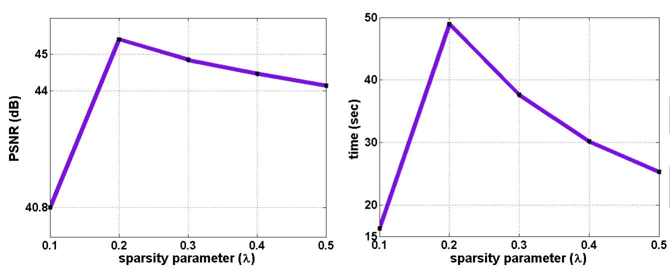


Fig. 10: *CAVE egyptian-statue 3D multispectral scene* ($512 \times 512 \times 31$)*: (Left) PSNR performance at different sparsity levels. (x4) sub-sampling factor. (Right) Run time comparisons of the three-dimensional hypercube reconstruction as a function of the sparsity parameters.*

Additionally, in Figure 10 we investigate the performance of our method in terms of the PSNR metric, over a grid of sparsity parameters λ, i.e $\lambda = [0.1, 0.2, 0.3, 0.4, 0.5]$, when applied to CAVE's multispectral Egyptian-statue 3D scene. In this simulation, the sub-sampling factor was set to 4. We observe that the amount of sparsity in the representation has a strong impact on the quality of the recovery of the three-dimensional hypercube. In our experiments, we empirically set the sparsity parameter λ equal to 0.2, achieving both high quality approximation with short execution time.

VI. Conclusion

In this work, we proposed a novel spectral super-resolution architecture for multi- and hyperspectral imagery, employing the mathematical framework of *Sparse Representations* though a *Coupled Sparse Dictionary Learning* algorithm for encoding the relations between high and low-spectral resolution scenes. To achieve this goal, an efficient formulation is proposed based on the Alternating Direction Method of Multipliers. Experimental results suggest that high quality reconstruction of both remote sensing and terrestrial data is attainable by the proposed scheme. Furthermore, the proposed scheme can be extended to handle arbitrary low-to-high resolution enhancements by simple modifications of the joint dictionary learning process, and offers the capability of addressing additional sources related to HSI image degradation. Experiments results applied on a variety of spectral image datasets, demonstrate that the proposed SCDL coupled dictionary learning scheme surpasses traditional methods based on single dictionary learning.

Acknowledgment

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

References

[1] Andreas Reigber, Rolf Scheiber, Marc Jager, Pau Prats-Iraola, Irena Hajnsek, Thomas Jagdhuber, Konstantinos P Papathanassiou, Matteo Nannini, Esteban Aguilera, Stefan Baumgartner, et al., "Very-high-resolution airborne synthetic aperture radar imaging: Signal processing and applications," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 759–783, 2013.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

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

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

[9] 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.

[10] 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.

[11] 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.

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

[13] 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.

[14] 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.

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

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

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

[35] Bo Fan, Gregory Ely, Shuchin Aeron, and Eric L Miller, "Exploiting algebraic and structural complexity for single snapshot computed tomography hyperspectral imaging systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 6, pp. 990–1002, 2015.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

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

[41] Fumihito Yasuma, Tomoo Mitsunaga, Daisuke Iso, and Shree K Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2241–2253, 2010.

[42] Bert Geelen, Nicolaas Tack, and Andy Lambrechts, "A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic," in *Spie Moems-Mems*. International Society for Optics and Photonics, 2014, pp. 89740L–89740L.

[43] Bert Geelen, Carolina Blanch, Pilar Gonzalez, Nicolaas Tack, and Andy Lambrechts, "A tiny vis-nir snapshot multispectral camera," in *SPIE OPTO*. International Society for Optics and Photonics, 2015, pp. 937414–937414.

[44] Andy Lambrechts, Pilar Gonzalez, Bert Geelen, Philippe Soussan, Klaas Tack, and Murali Jayapala, "A cmos-compatible, integrated approach to hyper-and multispectral imaging," in *2014 IEEE International Electron Devices Meeting*. IEEE, 2014, pp. 10–5.

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

[46] 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.

[47] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

Appendix

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

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

$$\mathbf{W}_h = \underset{\mathbf{W}_h}{\operatorname{argmin}} L = \nabla_{\mathbf{W}_h} L \Leftrightarrow$$

$$\nabla_{\mathbf{W}_h}\Big(\frac{1}{2}||\mathbf{S}_h - \mathbf{D}_h\mathbf{W}_h||_2^2 + < Y_1, \mathbf{P} - \mathbf{W}_h > +$$
$$< Y_3, \mathbf{W}_h - \mathbf{W}_l > + \frac{c_1}{2}||\mathbf{P} - \mathbf{W}_h||_2^2 + \frac{c_3}{2}||\mathbf{W}_h - \mathbf{W}_l||_2^2\Big)$$
$$= -\mathbf{D}_h^T \cdot (\mathbf{S_h} - \mathbf{D}_h\mathbf{W}_h) - Y_1 + Y_3 - c_1 \cdot (\mathbf{P} - \mathbf{W}_h) +$$
$$c_3 \cdot (\mathbf{W}_h - \mathbf{W}_l)$$

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

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

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

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

Setting $\nabla_{\mathbf{W}_l} L = 0 \;\; \Leftrightarrow$

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

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

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

  - For $\mathbf{P} > 0$,
$$\nabla_{\mathbf{P}} L = \lambda_h \cdot I + c_1 \cdot (\mathbf{P} - \mathbf{W}_h) + Y_1$$
Setting $\nabla_{\mathbf{P}} L = 0$,
$$\mathbf{P} = \mathbf{W}_h - \frac{1}{c_1} \cdot (Y_1 + \lambda_h \cdot I)$$
  - For $\mathbf{P} < 0$,
$$\nabla_{\mathbf{P}} L = -\lambda_h \cdot I + c_1 \cdot (\mathbf{P} - \mathbf{W}_h) + Y_1$$
Setting $\nabla_{\mathbf{P}} L = 0$,
$$\mathbf{P} = \mathbf{W}_h - \frac{1}{c_1} \cdot (Y_1 - \lambda_h \cdot I)$$

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* $\mathbf{Q}$

$$\mathbf{Q}^\star = \underset{\mathbf{Q}}{\text{argmin}} L = \nabla_{\mathbf{Q}} L \Leftrightarrow$$

  - For $\mathbf{Q} > 0$,
$$\nabla_{\mathbf{Q}} L = \lambda_l \cdot I + Y_2 + c_2 \cdot (\mathbf{Q} - \mathbf{W}_l)$$
Setting $\nabla_{\mathbf{Q}} L = 0$,
$$\mathbf{Q} = \mathbf{W}_l - \frac{1}{c_2} \cdot (Y_2 + \lambda_l \cdot I)$$
  - For $\mathbf{Q} < 0$,
$$\nabla_{\mathbf{Q}} L = \lambda_l \cdot I + Y_2 + c_2 \cdot (\mathbf{Q} - \mathbf{W}_l)$$
Setting $\nabla_{\mathbf{Q}} L = 0$,
$$\mathbf{Q} = \mathbf{W}_l - \frac{1}{c_2} \cdot (Y_2 - \lambda_l \cdot I)$$

Combining,
$$\mathbf{Q} > 0 \Leftrightarrow \mathbf{W}_l - \frac{1}{c_2} \cdot Y_2 > \frac{1}{c_2} \cdot \lambda_l \cdot I$$
$$\mathbf{Q} < 0 \Leftrightarrow \mathbf{W}_l - \frac{1}{c_2} \cdot Y_2 < -\frac{1}{c_2} \cdot \lambda_l \cdot I$$
we have,
$$|\mathbf{W}_l - \frac{1}{c_2} \cdot Y_2| \leq \frac{1}{c_2} \cdot \lambda_l \cdot I,$$
Consequently,
$$\mathbf{Q}^\star = S_{\lambda_l}\Big(|\mathbf{W}_l - \frac{Y_2}{c_2}|\Big),$$
where S_{λ_l} denotes the soft-thresholding operator, defined as:
$$S_{\lambda_l}(x) = sign(x) \cdot max(|x| - \lambda_l, 0)$$

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

$$\mathbf{D}_h^\star = \underset{\mathbf{D}_h}{\text{argmin}} L = \nabla_{\mathbf{D}_h} L \Leftrightarrow$$
$$\nabla_{\mathbf{D}_h} L = \mathbf{W}_h^T \cdot (\mathbf{S}_h - \mathbf{D}_h \cdot \mathbf{W}_h)$$
Setting, $\nabla_{\mathbf{D}_h} L = 0$,
$$\mathbf{D}_h = \frac{\mathbf{S}_h \cdot \mathbf{W}_h^T}{\phi_h + \delta},$$
where $\phi_h = \mathbf{W}_h \cdot \mathbf{W}_h^T$

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

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

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

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

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