Multi-Source Image Enhancement via Coupled Dictionary Learning

Konstantina Fotiadou*, Grigorios Tsagkatakis*, Panagiotis Tsakalides*
Institute of Computer Science
Foundation for Research and Technology (FORTH)
Emails: kfot@ics.forth.gr, greg@ics.forth.gr, tsakalid@ics.forth.gr

Motivation. Hyperspectral remote sensing imagery provides valuable insights regarding the composition of a scene and significantly facilitates tasks such as object and material recognition, spectral unmixing, and region clustering [1], [2]. However, current remote sensing imaging architectures are unable to concurrently acquire high spatial and spectral resolution imagery due to fact that the three-dimensional hyperspectral data must be acquired using a single, a 1D array, or a 2D plane detector.

Traditional "push-broom" sensors obtain a high spectral resolution profile of a low spatial resolution area (single line) during each exposure and generate the complete hyperspectral cube by progressive scanning of the scene. Modern Snapshot Spectral Imaging architectures sample the full spatio-spectral cube in a single exposure, without any need for successive frame acquisition, by associating each pixel with a specific spectral band. Even for the same type of sensor, a different point in the spatio-spectral operational curve might be selected depending on the application.

We report herein on a novel machine learning method for post-acquisition enhancement of multi and hyperspectral imagery. Example applications include imagery acquired from legacy low spectral resolution satellites, which could be enhanced using images of the same region, acquired by high resolution spectrometers aboard newer platforms. Respectively, the limited spatial information acquired by hyperspectral instruments, could be enhanced using high-spatial resolution imagery extracted by sensors with higher spatial resolution.

Proposed approach. In contrast to traditional hyperspectral super-resolution approaches that focus either on the spatial [3] or the spectral resolution enhancement [4], we propose a novel technique which addresses the spatio-spectral enhancement, where pairs of low and high spatio-spectral resolution training examples are used within a sparsifying dictionary learning framework. The proposed multi-instrument Coupled Dictionary Learning (CDL) technique capitalizes on the Sparse Representations framework [5] and extents it by introducing an efficient multi-source dictionary learning scheme, for estimating spatial and spectral information that was not explicitly acquired by the detectors. The CDL algorithm relies on generating coupled dictionary pairs that jointly encode two feature spaces, the low-spectral/high-spatial and the corresponding high-spectral/low-spatial resolution ones, where signals admit sparse representations.

Multi-source dictionary learning can be formulated as the concurrent identification of two dictionary matrices $\mathbf{D_X}, \mathbf{D_Y}$, corresponding to the feature spaces $\mathbf{S_X}$ and $\mathbf{S_Y}$, such that both $\mathbf{s_x} \in \mathbf{S_X}$ and $\mathbf{s_y} \in \mathbf{S_Y}$ share exactly the same sparse coding in terms of $\mathbf{D_X}$ and $\mathbf{D_Y}$, respectively. A straightforward approach is to convert the joint dictionary learning into a standard single dictionary learning problem by concatenating the individual feature spaces and utilize a common sparse representation, able to reconstruct both spaces. This problem can be efficiently solved via the K-SVD algorithm [6]. However, such a strategy is optimal only in the concatenated feature space, and not in the individual signal resolutions of $\mathbf{S_X}$ and $\mathbf{S_Y}$. In other words, when presented only with examples from $\mathbf{S_Y}$, the corresponding generated dictionary $\mathbf{D_Y^*}$ admits different space coding compared to the concatenated case. 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 a computationally efficient CDL technique, based on the Alternating Direction Method of Multipliers (ADMM) [7]. The main task of the ADMM coupled dictionary learning is to recover both the dictionaries with their corresponding sparse codes, by solving the following sparse decomposition problem:

$$\begin{aligned}&\min_{\mathbf{D_X},\mathbf{W_X},\mathbf{D_Y},\mathbf{W_Y}} ||\mathbf{S_X}-\mathbf{D_X}\mathbf{W_X}||_F^2 + ||\mathbf{S_Y}-\mathbf{D_Y}\mathbf{W_Y}||_F^2 \\ &+\lambda_{\mathbf{Y}}||\mathbf{Q}||_1 + \lambda_{\mathbf{X}}||\mathbf{P}||_1, \text{ subject to } \mathbf{P}=\mathbf{W_X}, \mathbf{Q}=\mathbf{W_Y}, \\ &\mathbf{W_X}=\mathbf{W_Y}, \ ||\mathbf{D_X}(:,i)||_2 \le 1, ||\mathbf{D_Y}(:,i)||_2 \le 1, \end{aligned} \tag{1}$$

where $\lambda_{\mathbf{X}}$, $\lambda_{\mathbf{Y}}$ are the sparsity balancing terms. The ADMM scheme takes into account the separate structure of each variable in (1), relying on the minimization of its augmented Lagrangian function:

$$\begin{aligned}&\mathcal{L}(\mathbf{D_X},\mathbf{D_Y},\mathbf{W_X},\mathbf{W_Y},\mathbf{P},\mathbf{Q},\mathbf{Y}_1,\mathbf{Y}_2,\mathbf{Y}_3) = ||\mathbf{D_X}\mathbf{W_X}-\mathbf{S_X}||_F^2 + \\ &||\mathbf{D_Y}\mathbf{W_Y}-\mathbf{S_Y}||_F^2 + \lambda_{\mathbf{X}}||\mathbf{P}||_1 + \lambda_{\mathbf{Y}}||\mathbf{Q}||_1 + <Y_1, \mathbf{P}-\mathbf{W_X}> + \\ &<Y_2, \mathbf{Q}-\mathbf{W_Y}> + <Y_3, \mathbf{W_X}-\mathbf{W_Y}> + \frac{c_1}{2}||\mathbf{P}-\mathbf{W_X}||_F^2 + \\ &\frac{c_2}{2}||\mathbf{Q}-\mathbf{W_Y}||_F^2 + \frac{c_3}{2}||\mathbf{W_X}-\mathbf{W_Y}||_F^2\end{aligned} \tag{2}$$

where $\mathbf{Y}_1$, $\mathbf{Y}_2$, and $\mathbf{Y}_3$ denote the Lagrange multiplier matrices, while c_1, c_2, c_3 denote the non-negative 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. Once the dictionaries are learned, a similar algorithmic strategy to the one described in [4] is employed for estimating the high resolution hyper-pixels. An illustrative block-diagram is depicted in Figure 1 which demonstrates how acquired data from two different sensors can be enhanced utilizing the proposed multi-instrument CDL approach.

Experimental Validation. To validate the merits of the proposed scheme, we report its performance when applied for the spectral super-resolution of hyperspectral data acquired by NASA's AVIRIS sensor [8], resolving 224 spectral bands between 400 and 2500 nm. Specifically, 20.000 randomly selected training hyper-pixels, from 5 training AVIRIS hyper-cubes are considered as the high resolution measurements. The low-spectral resolution hypercubes were produced via a uniform spectral response sampling corresponding to Landsat TM spectral bands [9]. Specifically, we have experimented with spectral sub-sampling factors of $2, 4, 8$, and 16, corresponding to 112, 56, 28 and 14 input spectral bands. Additionally, we investigated the extreme scenario of only 6 input spectral observations. Figures 2 and 3 illustrate representative bands from the reconstructed hypercubes obtained by the proposed enhancement technique.

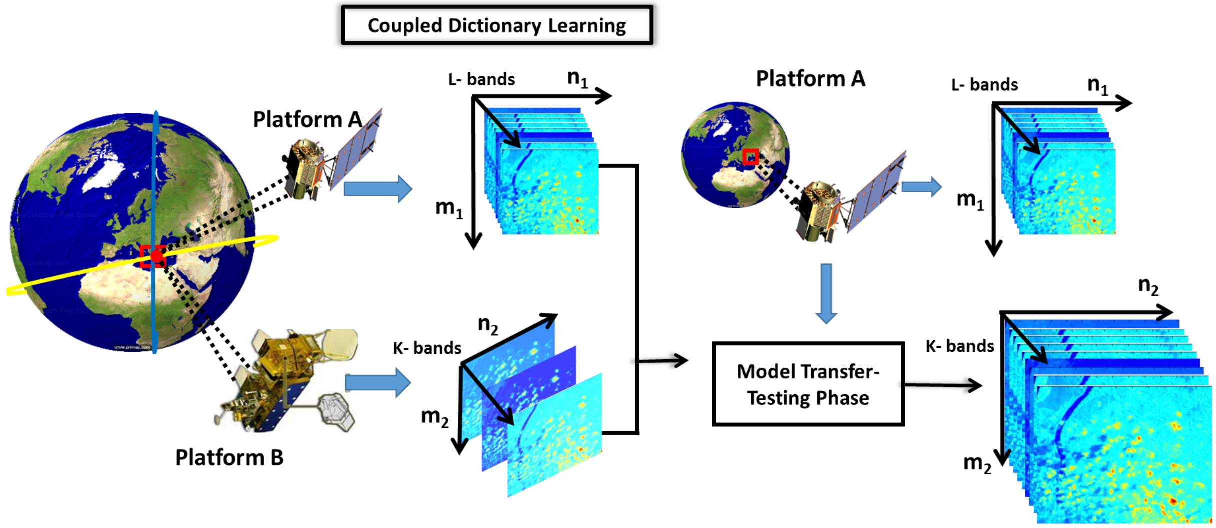
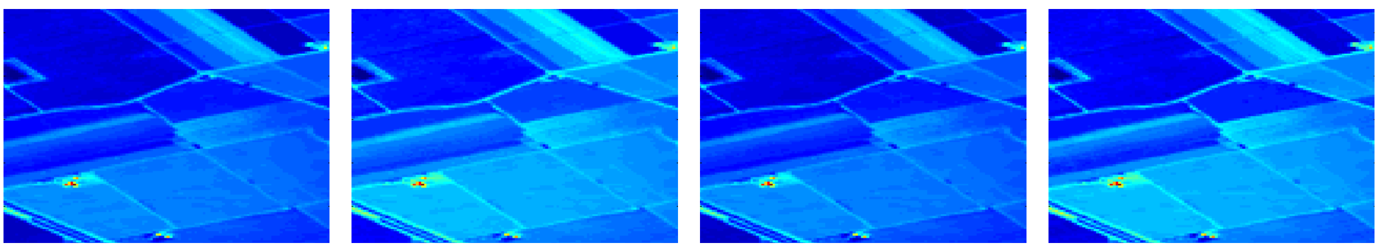
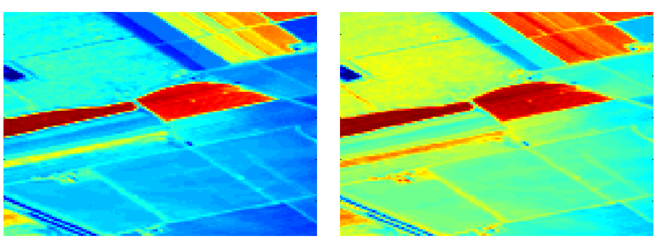


Fig. 1: Proposed Block Diagram: Our algorithm takes as input a 3D data-cube acquired with a limited spatial resolution from Platform A, and utilizes the high-spatial resolution trained model from Platform B, to increase spatial resolution. Respectively, the high-spectral resolution learned model from Platform B is utilized as prior knowledge to enhance the limited spectral information of Platform A.

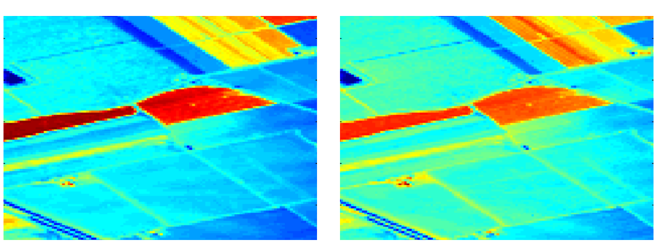


(a) Ground Truth 10^{th} (b) (x4) Recovered, PSNR: 48.24 db (c) (x8) Recovered, PSNR: 32.96 db (d) (x16) Recovered, PSNR: 23.32 db

Fig. 2: AVIRIS scene reconstruction: The full spectrum is composed of 224 bands between 400 and 2500 nm. In this experiment, we evaluate the performance of the proposed technique against multiple spectral down-sampling factors: (x2),(x4),(x8),(x16). The reconstruction quality of the full spectrum is evaluated via the Peak Signal to Noise Ratio (PSNR) [10]. We observe that under real life conditions, the proposed scheme produces a significant quality improvement operating in satellite hyperspectral imagery.



(a) Ground Truth 50^{th} (b) Ground Truth 70^{th}



(c) Recovered 50^{th} Band (d) Recovered 70^{th} Band

Fig. 3: In this experiment, the high resolution hypercube is estimated from only 6 input spectral observations. The PSNR metric for the recovery of the 3D hyper-cube is $\mathbf{20.36}$ db. We observe that the reconstructed spectral bands present a faithful representation with the ground truth observations.

References

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

[2] 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.

[3] Hongyan Zhang, Liangpei Zhang, and Huanfeng Shen, "A super-resolution reconstruction algorithm for hyperspectral images," *Signal Processing*, vol. 92, no. 9, pp. 2082–2096, 2012.

[4] 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.

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

[6] 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.

[7] 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.

[8] "Aviris," http://aviris.jpl.nasa.gov/, 2010 (accessed December 7, 2014).

[9] 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.

[10] Alain Hore and Djemel Ziou, "Image quality metrics: Psnr vs. ssim," in *Pattern recognition (icpr), 2010 20th international conference on*. IEEE, 2010, pp. 2366–2369.