# LEARNING SPECTRAL-SPATIAL PRIOR VIA 3DDNCNN FOR HYPERSPECTRAL IMAGE DECONVOLUTION

*Xiuheng Wang* [*], *Jie Chen* [*], *Cédric Richard* [†], *David Brie* [‡]

[*] Center for Intelligent Acoustics and Immersive Communications,
School of Marine Science and Technology, Northwestern Polytechnical University, China
[†] Université Côte d'Azur, CNRS, OCA, France
[‡] CRAN, Université de Lorraine, CNRS, France
Emails: xiuheng.wang@mail.nwpu.edu.cn, dr.jie.chen@nwpu.edu.cn, cedric.richard@unice.fr, david.brie@univ-lorraine.fr

## ABSTRACT

Hyperspectral image (HSI) deconvolution is an ill-posed problem aiming at recovering sharp images with tens or hundreds of spectral channels from blurred and noisy observations. In order to successfully conduct the deconvolution, proper priors are required to regularize the optimization problem. However, handcrafting a good regularizer may not be trivial and complex regularizers lead to difficulties in solving the optimization problem. In this paper, we use the alternating direction method of multipliers (ADMM) to decompose the optimization problem into iterative subproblems where the prior only appears in a denoising subproblem. Then a 3D denoising convolutional neural network (3DDnCNN) is designed and trained with data for solving this problem. In this way, the hyperspectral image deconvolution is then solved with a framework that integrates the optimization techniques and deep learning. Experimental results demonstrate the superiority of the proposed method with several blurring settings in both quantitative and qualitative comparisons.

***Index Terms*—** Hyperspectral image deconvolution, ADMM, spectral-spatial prior, 3D convolution, deep learning.

## 1. INTRODUCTION

Hyperspectral imaging captures a series of images of the same scene over many continuous narrow spectral bands. The high spectral resolution provided by hyperspectral images enables researchers to conduct analyses that cannot be done with conventional imaging techniques. Hyperpectral imaging has been used in many applications such as remote sensing [1], medical science [2], atmospheric monitoring [3]. However, observed hyperspectral images are sometimes blurred during the imaging process, leading to degraded performance in subsequent analyses. Thus, it is desirable to restore images by deconvolution (inversion of the degradation process) techniques before further processing.

Multichannel images contain abundant spectral information across neighboring wavelengths, making image restoration more complex than ordinary 2D images [4, 5]. Deconvolution of multichannel (multispectral) images involves Wiener filter [4, 6], Kalman filter, and regularized least-squares [7]. For hyperspectral deconvolution, in [8] the authors use adaptive 3D Wiener filter, and in [4] the authors use filter-based linear methods for astronomic hyperspectral images. 2D Fast Fourier Transforms (FFTs) [9] and Fourier-wavelet technique [10] are also considered for hyperspectral image restoration, to compute solution efficiently in Fourier and wavelet domains respectively. Since the deconvolution problem is ill-posed, it is important to incorporate prior information on images to regularize the solution. To this end, the fast hyperspectral restoration algorithm in [11] performs deconvolution under positivity constraints while accounting for spatial and spectral correlation. In [12], an online deconvolution algorithm is devised by considering sequentially collected data by push-broom devices.

Similar to conventional image deconvolution, properly defining priors and designing regularizers play an important role in improving the performance and enhancing the stability of the deblurring processing. However, it is a non-trivial task to handcraft a powerful regularizer, and complex regularizers may introduce extra difficulties in solving optimization problems. The variable splitting technique is one of the efficient ways to solve optimization problems with a data fidelity term and a regularization term, especially for the case with non-differentiable regularizers such as $\ell_1$-norm and TV-norm regularizations. Within this framework, various methods have been proposed based on the ADMM [13] or the half quadratic splitting method [14] to solve a variety of 2D image inverse problems like denoising, deblurring and superresolution [15–18].

Recently, learning the prior from the data is used for 2D image processing and shows its advantageous over using the handcrafted regularizers. Benefiting from the variable splitting technique, this approach plugs the data-based image denoising methods as a module in the optimization iterations to solve various inverse problems [17,18]. This efficient strategy has not been employed in hyperspectral image deconvolution problems, though similar difficulties of designing regularizers are encountered therein. In this work, we propose a hyperspectral deconvolution technique that uses spectral-spatial priors learnt from data by a deep neural network. The optimization problem of multichannel image deconvolution is addressed by the ADMM algorithm. A new effective 3DDnCNN is trained as a denoiser to learn priors of hyperspectral images. Experimental results show the effectiveness of the proposed strategy.

## 2. PROBLEM FORMULATION

We denote a degraded hyperspectral image and its latent clean counterpart by $\mathbf{Y} \in \mathbb{R}^{P\times Q\times N}$ and $\mathbf{X} \in \mathbb{R}^{P\times Q\times N}$ respectively, where $P$, $Q$, and N are the numbers of rows, columns and spectral dimension of the image. The degraded image and the clean image of the $i$th spectral band are denoted by $\mathbf{Y}_i \in \mathbb{R}^{P\times Q}$ and $\mathbf{X}_i \in \mathbb{R}^{P\times Q}$. For ease of mathematical formulation, the columns of $\mathbf{Y}_i$ and $\mathbf{X}_i$

This work was funded in part by NSFC 61811530283, 111 project B18041, the seed Foundation of Innovation and Creation for Graduate Students in Northwestern Polytechnical University. The work of C. Richard was funded in part by the CoopIntEER program CNRS-NSFC (DIALOG project). Corresponding author: J. Chen.

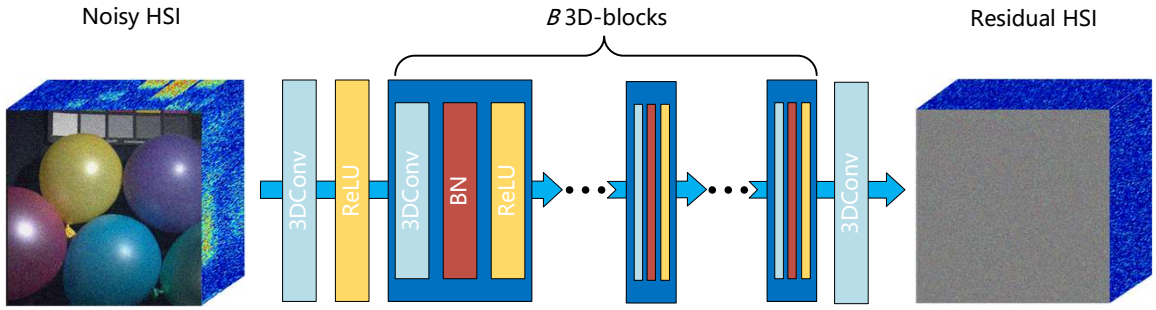


**Fig. 1**. Architecture of the proposed 3DDnCNN for hyperspectral image denoising.

are stacked to form two vectors $\mathbf{y}_i \in \mathbb{R}^{PQ\times 1}$ and $\mathbf{x}_i \in \mathbb{R}^{PQ\times 1}$. Following the linear degradation model in [11] and supposing that the convolution is separable over spectral bands, vectors $\mathbf{y}_i$ and $\mathbf{x}_i$ are related by

$$\mathbf{y}_i = \mathbf{H}_i\mathbf{x}_i + \mathbf{n}_i, \quad \text{for } i = 1, \cdots, N. \tag{1}$$

where $\mathbf{H}_i$ is the 2D blurring matrix, $\mathbf{n}_i$ is the additive independent and identically distributed (i.i.d.) Gaussian noise with standard deviation $\sigma$ for all channels. Under the assumption that $\{\mathbf{H}_i\}_{i=1}^N$ are the same over all channels, i.e., $\mathbf{H}_i = \mathbf{H}$ for $i = 1, \cdots, N$, we can estimate $\mathbf{x}_i$ by seeking the minimum of the following objective function:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \sum_{i=1}^{N} \frac{1}{2}\|\mathbf{y}_i - \mathbf{H}\mathbf{x}_i\|^2 + \lambda\Phi(\mathbf{x}) \tag{2}$$

where $\mathbf{x} = \text{col}\{\mathbf{x}_i\}_{i=1}^N$ with $\text{col}\{\cdots\}$ stack its vector arguments to from a connected vector, the first squared-error term is the data fidelity term, and $\Phi(\mathbf{x})$ is the regularizer that enforces desirable properties of the solution with $\lambda \geq 0$ being the regularization parameter. In hyperspectral image processing, spatial and spectral correlations are often encoded in $\Phi(\mathbf{x})$.

## 3. PROPOSED METHOD AND NETWORK DESIGN

Designing a good regularizor $\Phi(\mathbf{x})$ along with efficient solving method is not trivial task. Instead, we propose to learn priors from hyperspectral data and incorporate it into model-based optimization to tackle the regularized inverse problem in (2). More specifically, using the variable splitting technique, we transform problem (2) into two sub-problems, namely, a simple quadratic problem and a 3D-image denoising problem. These sub-problems are iteratively solved, using a linear method and a deep neural network, respectively, until the convergence criterion is met.

### 3.1. Variable splitting based on ADMM

ADMM is adopted to decouple the data fidelity term and the regularization term in (2). By introducing an auxiliary variable $\mathbf{z}$, problem (2) can be written in the equivalent form:

$$\begin{aligned} \hat{\mathbf{x}} &= \arg\min_{\mathbf{x}} \sum_{i=1}^{N} \frac{1}{2}\|\mathbf{y}_i - \mathbf{H}\mathbf{x}_i\|^2 + \lambda\Phi(\mathbf{z}) \\ &\text{s.t.} \quad \mathbf{z} = \mathbf{x} \quad \text{with } \mathbf{x} = \text{col}\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}. \end{aligned} \tag{3}$$

The associated augmented Lagrangian function is given by

$$\begin{aligned} \mathcal{L}_\rho(\mathbf{x}, \mathbf{z}, \mathbf{v}) = &\arg\min_{\mathbf{x}} \sum_{i=1}^{N} \frac{1}{2}\|\mathbf{y}_i - \mathbf{H}\mathbf{x}_i\|^2 + \lambda\Phi(\mathbf{z}) \\ &+ \mathbf{v}^T(\mathbf{x} - \mathbf{z}) + \frac{\rho}{2}\|\mathbf{x} - \mathbf{z}\|^2 \end{aligned} \tag{4}$$

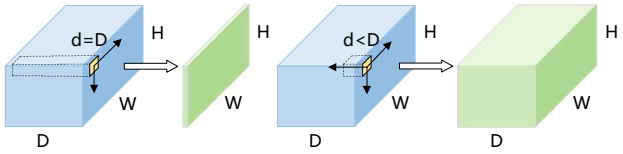


**Fig. 2**. Difference between 2D and 3D convolution.

where $\mathbf{v}$ is the dual variable, and $\rho \geq 0$ is the penalty parameter. Scaling $\mathbf{v}$ as $\mathbf{u} = \frac{1}{\rho}\mathbf{v}$, (4) can be iteratively solved by repeating the following steps:

$$\mathbf{x}_{k+1} = \arg\min_{\mathbf{x}} \sum_{i=1}^{N} \frac{1}{2}\|\mathbf{y}_i - \mathbf{H}\mathbf{x}_i\|^2 + \frac{\rho}{2}\|\mathbf{x} - \tilde{\mathbf{x}}_k\|^2 \tag{5a}$$

$$\mathbf{z}_{k+1} = \arg\min_{\mathbf{z}} \lambda\Phi(\mathbf{z}) + \frac{\rho}{2}\|\tilde{\mathbf{z}}_k - \mathbf{z}\|^2 \tag{5b}$$

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{x}_{k+1} - \mathbf{z}_{k+1} \tag{5c}$$

where $\tilde{\mathbf{x}}_k = \mathbf{z}_k - \mathbf{u}_k$ and $\tilde{\mathbf{z}}_k = \mathbf{x}_{k+1} + \mathbf{u}_k$. In this way, the data fidelity term and the regularization term in (2) are decoupled into sub-problems (5a) and (5b). Note that (5a) is a least square problem, which can be solved analytically with the solution:

$$\mathbf{x}_{k+1} = \text{col}\{(\mathbf{H}^T\mathbf{H} + \rho\mathbf{I})^{-1}(\mathbf{H}^T\mathbf{y}_i + \rho\tilde{\mathbf{x}}_{k,i})\}_{i=1}^N \tag{6}$$

and (5b) can be reformulated as

$$\mathbf{z}_{k+1} = \arg\min_{\mathbf{z}} \frac{1}{2(\sqrt{\lambda/\rho})^2}\|\tilde{\mathbf{z}}_k - \mathbf{z}\|^2 + \Phi(\mathbf{z}) \tag{7}$$

which is actually performed in the 3D image domain. From a Bayesian viewpoint, (7) can be regarded as a denoising problem, removing Gaussian noise with noise-level $\sqrt{\lambda/\rho}$ from the noisy HSI $\tilde{\mathbf{z}}_k$ to obtain the clean HSI $\mathbf{z}_{k+1}$. In other words, a denoising operator can be used for the regularization term $\Phi(\mathbf{x})$.

### 3.2. Learning spectral-spatial prior via 3DDnCNN

As discussed above, instead of using handcrafted regularizers, we propose to train a CNN-based denoiser for (7) by directly learning the spectral-spatial prior from hyperspectral image datasets.

A 3D denoising convolutional neural network is specifically designed for this task. Unlike 2D convolution resulting in spectral information distortion, 3D convolution extracts spatial feature of neighboring pixels and spectral feature of adjacent bands simultaneously without reducing spectral resolution. 3D convolution also involves fewer parameters and is more appropriate for hyperspectral image processing due to the difficulty in capturing a large number of hyperspectral data (see Fig. 2).

**Algorithm 1** HSI deconvolution with prior learnt from 3DDnCNN.

**Input:** Network parameters $\Theta$, blurred observation $\mathbf{y}$, number of spectral bands $N$, regularization parameter $\lambda$, penalty factor $\rho$, number of iterations $K$.

**Output:** Deblurred HSI $\mathbf{x}$.

Initialize $\mathbf{x} = \mathbf{x_0}$, auxiliary variable $\mathbf{z_0} = \mathbf{x_0}$, scaled dual variable $\mathbf{u_0} = \mathbf{0}$, $k = 0$.

**while** Stopping criteria are not met and $k \leq K$ **do**

$\tilde{\mathbf{x}}_k = \mathbf{z}_k - \mathbf{u}_k$

**for** $i = 1$ to $N$ **do**

$\mathbf{x}_{k+1,i} = (\mathbf{H}^T\mathbf{H} + \mu\mathbf{I})^{-1}(\mathbf{H}^T\mathbf{y}_i + \rho\tilde{\mathbf{x}}_{k,i})$

**end for**

$\tilde{\mathbf{x}}_k = \mathbf{x}_{k+1} + \mathbf{u}_k$

$\mathbf{z}_{k+1} = \mathcal{F}(\tilde{\mathbf{z}}_k, \Theta)$

$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{x}_{k+1} - \mathbf{z}_{k+1}$

$k = k + 1$

**end while**

The structure of the proposed 3DDnCNN is illustrated in Fig. 1. Let us define a 3D-block by the composition of a 3D convolution layer (3DConv), a batch normalization (BN) layer and a Rectifier Linear Unit (ReLU) activation function layer. Besides the input and output layers, a 3D convolution layer (3DConv), a ReLU activation function layer, $B$ 3D-blocks and a last 3D convolution layer are sequentially connected to form the proposed network. Batch normalization is also used to speed up the training process as well as to boost the denoising performance [19]. The last convolutional layer contains one 3D-filter while the others are composed of 32 3D-filters. The kernel size of each 3D-filter is 3×3×3, which implies that the depth of the kernel in spectral dimension and size of the kernel in spatial dimension are 3 and 3×3 respectively.

The input of the proposed 3DDnCNN is the noisy hyperspectral image $\tilde{\mathbf{z}} = \mathbf{z} + \mathbf{v}$, where $\mathbf{v}$ is the Gaussian noise. Inspired by 2D image denoising algorithm [19], we adopt the residual learning to predict the residual error $\hat{\mathbf{v}} \approx \mathbf{v}$ in our denoising network, then we can achieve the estimated clean image by $\hat{\mathbf{z}} = \tilde{\mathbf{z}} - \hat{\mathbf{v}}$. For training the network, we use the following loss function:

$$\ell(\Theta) = \sum_{m=1}^{M} \|\mathcal{F}(\tilde{\mathbf{z}}_m; \Theta) - (\tilde{\mathbf{z}}_m - \mathbf{z}_m)\|_1 \tag{8}$$

where $\{(\tilde{\mathbf{z}}_m, \mathbf{z}_m)\}_{m=1}^{M}$ represents $M$ generated noisy-clean HSI (patch) pairs used to train the network function $\mathcal{F}$ parameterized by $\Theta$. Note that $\ell_1$-norm is used as we find it leads to better performance than $\ell_2$-norm in denoising.

After the 3DDnCNN is trained with massive noisy-clean HSI pairs, it is incorporated into the ADMM-based framework as a denoiser, yielding Algorithm 1.

## 4. EXPERIMENTAL RESULTS

In this section, we validate the proposed method to show its effectiveness. We used two public hyperspectral image databases, namely, CAVE [20] database with 32 indoor HSIs recorded under controlled illuminations, and Harvard [21] database with 50 indoor and outdoor HSIs captured under daylight illumination. In the CAVE database, the images consist of $512 \times 512$ pixels, with 31 spectral bands of 10 nm, covering the visible spectrum 400-700 nm. The images from the Harvard database have a spatial resolution of $1392 \times 1040$ pixels and 31 channels, ranging from 420 nm to 720 nm at a wavelength interval of 10 nm. The top left $1024 \times 1024$ pixels were extracted in our experiments.

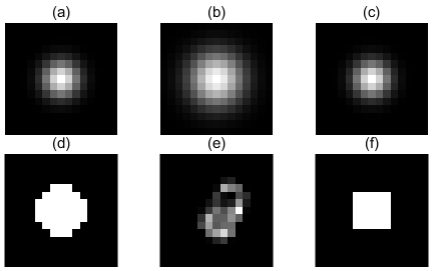


**Fig. 3**. Blurring kernels used in the experiments.

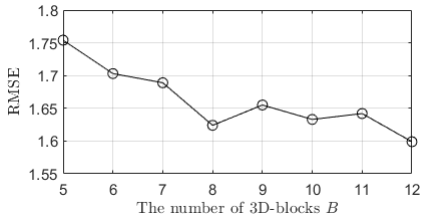


**Fig. 4**. The denoising performance of 3DDnCNN with different $B$.

The following settings of blur kernels and i.i.d. white Gaussian noise with stand deviation $\sigma$ were used to generate degraded images in the experiment.

(a) 15×15 Gaussian kernel with bandwidth $\sigma_k = 1.6$, and $\sigma = 0.01$.

(b) 15×15 Gaussian kernel with bandwidth $\sigma_k = 2.4$, and $\sigma = 0.01$.

(c) 15×15 Gaussian kernel with bandwidth $\sigma_k = 1.6$, and $\sigma = 0.03$.

(d) Circle kernel with diameter of 7, and $\sigma = 0.01$.

(e) Motion kernel from [22] of size 15×15, and $\sigma = 0.01$.

(f) Square kernel with side length of 5, and $\sigma = 0.01$.

See Fig. 3 for these kernels.

For training and evaluating the proposed 3DDnCNN, the first 20 images of the CAVE database were used as the training set and the others were used as the test set. In the Harvard database, the first 30 hyperspectral images were used as the training set while the others were used as the test set. We implemented our denoising network based on Pytorch and Adam optimizer [23] with an initial learning rate 0.0002 and a mini-batch of 64 to minimize the loss function (8) in 500 epochs. The weights were initialized by the method in [24]. In every epoch of the training phase, each original hyperspectral image was randomly cropped into patches of size 64× 64, and each patch was randomly flipped and rotated for data augmentation.

In our proposed method, the regularization parameter $\lambda$ was set to $9.6\times10^{-5}$ and the penalty factor $\rho$ was set to 0.06 except for setting (c) $\lambda = 1.28\times10^{-3}$ and $\rho = 0.8$ due to its different noise level. The influence of the number of 3D-blocks $B$ is examined in Fig. 4, where the average root-mean-square error (RMSE) value based on the test set in the CAVE database was used to evaluate the denoising performance. As shown in Fig. 4, large values of $B$ generally leads better results. In our deconvolution experiment, we set $B = 8$ as a larger value by considering the computational cost and memory demand.

Once the denoiser 3DDnCNN was trained, it was plugged into the ADMM framework. The number of iterations $K$ in Algorithm 1 was set to 20 which was sufficient to ensure the convergence. To evaluate the quality of the deblurred images, four quantitative metrics including the root mean-square error (RMSE), peak-signal-to-noise-ratio (PSNR), spectral angle mapper (SAM) [25] and structural similarity (SSIM) [26]. We compared the proposed method

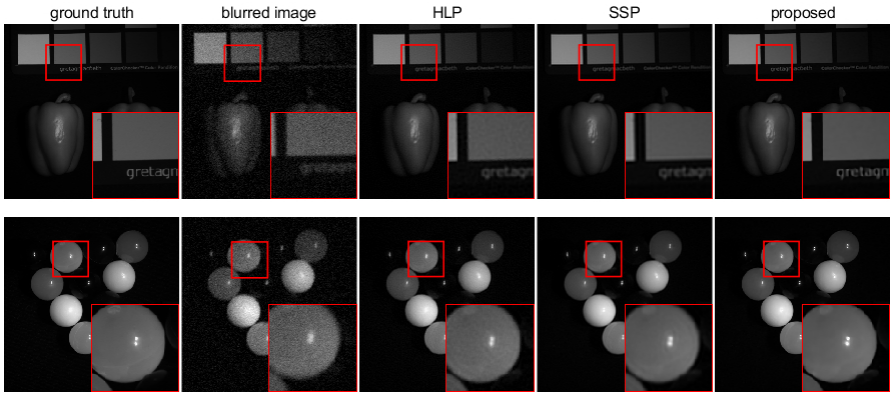


**Fig. 5**. Visual results for comparison of the methods.

**Table 1**. Average RMSE, PSNR, SAM, SSIM of different methods on CAVE and Harvard database in 6 blur scenarios.

| Scenarios | Methods | CAVE database | | | | Harvard database | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | RMSE | PSNR | SAM | SSIM | RMSE | PSNR | SAM | SSIM |
| (a) | HLP | 3.781 | 37.427 | 11.19 | 0.9257 | 3.562 | 38.181 | 8.41 | 0.9133 |
| | SSP | 4.299 | 36.536 | 8.24 | 0.9474 | 3.687 | 38.657 | 5.06 | 0.9353 |
| | proposed | **2.559** | **41.173** | **6.27** | **0.9602** | **2.722** | **41.116** | **4.84** | **0.9486** |
| (b) | HLP | 4.947 | 35.340 | 10.98 | 0.9058 | 4.356 | 36.885 | 8.07 | 0.8903 |
| | SSP | 5.285 | 34.697 | 8.49 | **0.9255** | 4.431 | 37.225 | 5.10 | 0.9110 |
| | proposed | **4.774** | **35.961** | **7.77** | 0.9199 | **3.766** | **38.852** | **4.77** | **0.9202** |
| (c) | HLP | 7.778 | 30.042 | 25.83 | 0.6140 | 7.729 | 30.515 | 23.24 | 0.5947 |
| | SSP | 4.606 | 35.655 | 12.23 | 0.9161 | 4.060 | 37.079 | 8.22 | 0.9046 |
| | proposed | **3.898** | **37.424** | **6.46** | **0.9395** | **3.737** | **38.958** | **4.58** | **0.9165** |
| (d) | HLP | 4.486 | 36.102 | 11.52 | 0.9116 | 4.004 | 37.354 | 8.58 | 0.8992 |
| | SSP | 4.788 | 35.609 | 8.41 | 0.9350 | 3.992 | 38.085 | 5.08 | 0.9232 |
| | proposed | **3.017** | **39.887** | **6.80** | **0.9477** | **2.987** | **40.500** | **4.82** | **0.9384** |
| (e) | HLP | 4.312 | 36.183 | 12.97 | 0.9035 | 3.891 | 37.248 | 9.88 | 0.8945 |
| | SSP | 4.473 | 36.151 | 8.51 | 0.9427 | 3.807 | 38.372 | 5.19 | 0.9317 |
| | proposed | **2.024** | **43.042** | **6.30** | **0.9638** | **2.306** | **42.258** | **4.77** | **0.9568** |
| (f) | HLP | 3.933 | 37.054 | 11.77 | 0.9211 | 3.642 | 37.928 | 8.86 | 0.9090 |
| | SSP | 4.305 | 36.539 | 8.33 | 0.9463 | 3.646 | 38.771 | 5.08 | 0.9347 |
| | proposed | **2.078** | **42.718** | **6.29** | **0.9669** | **2.458** | **41.807** | **4.87** | **0.9541** |

with two deconvolution methods that can be used for hyperspectral images, namely the algorithms with well-designed regularizers in [27] and [11]. The first method (denoted by HLP), considering the spatial priors, i.e., the hyper-Laplacian priors of images. The second method (denoted by SSP), considering both the spatial and spectral priors of hyperspectral images.

The average quantitative results of these methods on the test sets are reported in Table 1. The proposed method achieved satisfactory results in various blur scenarios. For visual comparison, we take scenario (a) for example. Fig. 5 illustrates blurred image, deblurred images, ground truth of *real and fake peppers* and *superballs* (two images from CAVE dataset) at the 15h and 24th bands respectively. Fig. 6 shows the spectra difference between recovered data and the clean data (*sponges* from CAVE dataset) on several pixels. We observe that our proposed method exhibited better spatial and spectral visual results.

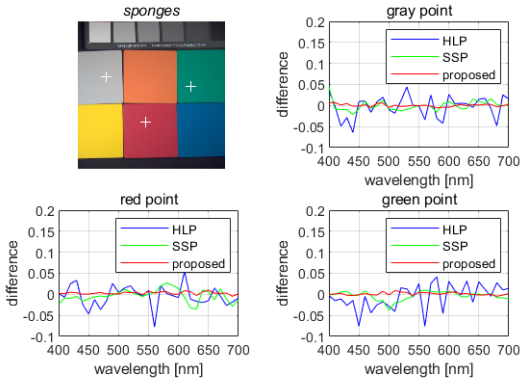


**Fig. 6**. Spectral difference between restored data and clean data on several pixels.

## 5. CONCLUSION

This paper presented hyperspectral image deconvolution method based on the ADMM framework. Instead of using handcraftted priors, we designed a denoiser based on 3DDnCNN to learn the spectral-spatial information of hyperspectral images from data, and plugged it to ADMM based optimization. Experimental results on two public databases demonstrated that proposed method can effectively handle scenarios with various blurring setting. In the future, we will further improve the structure of 3DDnCNN and extend the proposed method to blind hyperspectral deconvolution.

## 6. REFERENCES


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

[2] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *J. Biomed. Opt.*, vol. 19, no. 1, p. 010901, 2014.

[3] C. J. Keith, K. S. Repasky, R. L. Lawrence, S. C. Jay, and J. L. Carlsten, "Monitoring effects of a controlled subsurface carbon dioxide release on vegetation using a hyperspectral imager," *Int. J. Greenh. Gas Control*, vol. 3, no. 5, pp. 626–632, 2009.

[4] S. Bongard, F. Soulez, É. Thiébaut, and É. Pecontal, "3d deconvolution of hyper-spectral astronomical data," *Mon. Not. Roy. Astron. Soc.*, vol. 418, no. 1, pp. 258–270, 2011.

[5] P. Sarder and A. Nehorai, "Deconvolution methods for 3-d fluorescence microscopy images," *IEEE Signal Process. Mag.*, vol. 23, no. 3, pp. 32–45, 2006.

[6] N. P. Galatsanos and R. T. Chin, "Digital restoration of multichannel images," *IEEE Trans. Audio, Speech, Language Process.*, vol. 37, no. 3, pp. 415–421, 1989.

[7] N. P. Galatsanos, A. K. Katsaggelos, R. T. Chin, and A. D. Hillery, "Least squares restoration of multichannel images," *IEEE Trans. Signal Process.*, vol. 39, no. 10, pp. 2222–2236, 1991.

[8] J.-M. Gaucel, M. Guillaume, and S. Bourennane, "Adaptive-3d-wiener for hyperspectral image restoration: Influence on detection strategy," in *Proc. EUSIPCO*, 2006, pp. 1–5.

[9] E. Thiébaut, "Introduction toimage reconstruction and inverse problems," in *Proc. NATO Adv. Study Inst. Opt. Astrophys.*, 2005, pp. 1–26.

[10] R. Neelamani, H. Choi, and R. Baraniuk, "Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 418–433, 2004.

[11] S. Henrot, C. Soussen, and D. Brie, "Fast positive deconvolution of hyperspectral images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 828–833, 2012.

[12] Y. Song, E.-H. Djermoune, J. Chen, C. Richard, and D. Brie, "Online deconvolution for industrial hyperspectral imaging systems," *SIAM Journal on Imaging Sciences*, vol. 12, no. 1, pp. 54–86, 2019.

[13] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[14] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Trans. Image Process.*, vol. 4, no. 7, pp. 932–946, 1995.

[15] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *Proc. GlobalSIP*, 2013, pp. 945–948.

[16] A. Brifman, Y. Romano, and M. Elad, "Turning a denoiser into a super-resolver using plug and play priors," in *Proc. ICIP*, 2016, pp. 1404–1408.

[17] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep cnn denoiser prior for image restoration," in *Proc. CVPR.*, 2017, pp. 3929–3938.

[18] J. Rick Chang, C.-L. Li, B. Poczos, B. Vijaya Kumar, and A. C. Sankaranarayanan, "One network to solve them all–solving linear inverse problems using deep projection models," in *Proc. ICCV.*, 2017, pp. 5888–5897.

[19] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, 2017.

[20] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2241–2253, 2010.

[21] A. Chakrabarti and T. Zickler, "Statistics of real-world hyperspectral images," in *Proc. CVPR.*, 2011, pp. 193–200.

[22] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. CVPR.*, 2009, pp. 1964–1971.

[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. ICCV.*, 2015, pp. 1026–1034.

[25] R. H. Yuhas, A. F. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm," 1992.

[26] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli *et al.*, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[27] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-laplacian priors," in *Proc. NIPS.*, 2009, pp. 1033–1041.