

Hyperspectral Image Super-resolution with Deep Priors and Degradation Model Inversion

Xiuheng Wang*, Jie Chen†, Cédric Richard*

*Université Côte d'Azur, CNRS, OCA, France.

†School of Marine Science and Technology, Northwestern Polytechnical University, China.

7th-13th May



- Problem Formulation
- Related Work and Motivations
- Our Method with Deep Priors and Degradation Model Inversion
- Iterative Numerical Optimization
- Experimental Results
- Conclusions

Problem Formulation

To overcome inherent hardware limitations of hyperspectral imaging systems with their **spatial resolution**, fusion-based hyperspectral image (HSI) super-resolution [Lanaras et al. 2015] is attracting increasing attention.

Fusion-based HSI super-resolution:

Restore a high-resolution (HR) HSI \mathbf{X} by fusing a low-resolution (LR) HSI \mathbf{Y} and a conventional HR RGB image \mathbf{Z} .

This framework is based on the following linear degradation model:

$$\mathbf{Y} = \mathbf{XBS}, \quad \mathbf{Z} = \mathbf{RX}. \quad (1)$$

where \mathbf{B} is the spatial blurring matrix, \mathbf{S} is a downsampling operator, and \mathbf{R} denotes the spectral response function (SRF) of the RGB camera.

Problem Formulation

Based on (1), this problem can be written as the minimization of a constrained objective function of the form:

$$J(\mathbf{X}) = \|\mathbf{Y} - \mathbf{XBS}\|_F^2 + \|\mathbf{Z} - \mathbf{RX}\|_F^2 + \varphi(\mathbf{X}) \quad (2)$$

where $\varphi(\cdot)$ is some regularization functions.

Reconstructing \mathbf{X} from \mathbf{Y} and \mathbf{Z} by minimizing (2) without $\varphi(\cdot)$ is a highly **ill-posed** problem. This justifies the use of $\varphi(\cdot)$ to constrain the solution space by exploiting **the prior information** on \mathbf{X} .

Related Work and Motivations

The classic **model-based** optimization framework considers the **well-defined** prior information to enhance the super-resolution performance.

- Designing explicit $\varphi(\cdot)$ is one efficient way to incorporate such image priors.
 - Sparsity [Akhtar et al. 2014].
 - Spatial continuity [Lanaras et al. 2015].
 - Edge preserving [Dian et al. 2019].
 - ...
- **Drawback:** designing a powerful $\varphi(\cdot)$ is not trivial and may also cause difficulty in finding optimal solutions.

Inspired by the success of **deep learning**, **data-driven** convolutional neural networks (CNN) have been considered [Palsson et al. 2017; Zhang et al. 2020].

- Deep learning methods require less prior information on **X** and achieve significant performance enhancement compared to model-based methods.
- **Drawback:** they need massive data for training and may not be consistent with the physical degradation model involving **Y** and **Z**.

Is it possible to leverage the merits of both model-based and deep learning methods?

To tackle this issue, recent approaches have started to plug the output of an CNN, denoted as $\tilde{\mathbf{X}}$, into the objective function (2) as a *deep prior regularizer*.

- The Frobenius norm $\varphi(\mathbf{X}) = \|\mathbf{X} - \tilde{\mathbf{X}}\|_F^2$ is considered in [Xie et al. 2019; Wang et al. 2021].
- The 2D Total Variation (TV) norm $\varphi(\mathbf{X}) = \|\mathbf{X} - \tilde{\mathbf{X}}\|_{TV}^2 + \|\mathbf{X}\|_{TV}$ is used in [Vella et al. 2021].

Nevertheless, none of these methods simultaneously exploits **the spectral-spatial gradient information** for enhancing fusion process.

Our Method with Deep Priors and Degradation Model Inversion

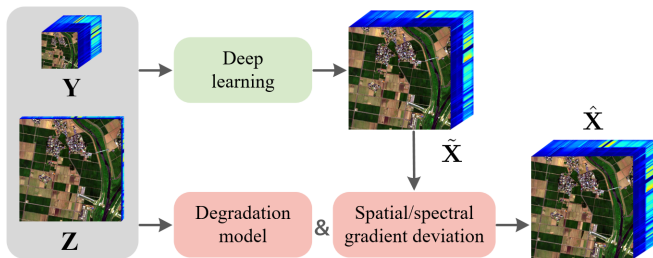


Figure: The scheme of our method.

In this work, as illustrated in Figure, we address this problem with deep priors and model degradation inversion accounting for spatial/spectral gradient deviation of HSIs.

Our Method with Deep Priors and Degradation Model Inversion

Specific, an CNN output $\tilde{\mathbf{X}}$ is used to enhance the result via $\varphi(\mathbf{X})$ by reducing the difference between \mathbf{X} and $\tilde{\mathbf{X}}$ in spectral and spatial gradient domains respectively.

Under this design, the objective function (2) becomes:

$$\begin{aligned} J(\mathbf{X}) &= \|\mathbf{Y} - \mathbf{XBS}\|_F^2 + \|\mathbf{Z} - \mathbf{RX}\|_F^2 + \varphi(\mathbf{X}) \\ \text{with } \varphi(\mathbf{X}) &= \mu\|\mathbf{D}(\mathbf{x} - \tilde{\mathbf{x}})\|^2 + \nu\|\mathbf{E}(\mathbf{x} - \tilde{\mathbf{x}})\|^2 \\ \text{and } \tilde{\mathbf{X}} &= \text{CNN}(\mathbf{Y}, \mathbf{Z}) \end{aligned} \quad (3)$$

- $\|\mathbf{Y} - \mathbf{XBS}\|_F^2$ and $\|\mathbf{Z} - \mathbf{RX}\|_F^2$ guarantee that the candidate solution is consistent with the degradation model (1).
- $\|\mathbf{D}(\mathbf{x} - \tilde{\mathbf{x}})\|^2$ and $\|\mathbf{E}(\mathbf{x} - \tilde{\mathbf{x}})\|^2$ are regularization terms exploiting deep priors, with positive hyper-parameters μ and ν .

Our Method with Deep Priors and Degradation Model Inversion

Matrix \mathbf{D} can be designed by choosing first a convolution kernel \mathcal{D} . We consider the Laplacian filter \mathcal{D}_ℓ for each channel ℓ :

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}. \quad (4)$$

We construct a block-Toeplitz matrix \mathbf{D}_ℓ as in [Henrot et al. 2012]. This leads to matrix \mathbf{D} in (3) with block-diagonal structure:

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{D}_B \end{pmatrix}. \quad (5)$$

Our Method with Deep Priors and Degradation Model Inversion

A typical choice of \mathbf{E} is a first-order derivative filter $\mathcal{E}_0 = [1 - 1]$ along the spectral dimension. The convolution matrix is then given by:

$$\mathbf{E}_0 = \begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -1 & 1 \end{pmatrix}. \quad (6)$$

This yields:

$$\mathbf{E} = \mathbf{E}_0 \otimes \mathbf{I}_N \quad (7)$$

where \otimes denotes the Kronecker product and \mathbf{I}_N is the identity matrix.

Variable splitting based on HQS

HQS is employed to decouple the data fidelity terms and the regularization terms in (3). By introducing an auxiliary variable \mathbf{V} , we have:

$$\hat{\mathbf{X}} = \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{XBS}\|_F^2 + \|\mathbf{Z} - \mathbf{RX}\|_F^2 + \mu\|\mathbf{D}(\mathbf{v} - \tilde{\mathbf{x}})\|^2 + \nu\|\mathbf{E}(\mathbf{v} - \tilde{\mathbf{x}})\|^2 \quad \text{s.t. } \mathbf{V} = \mathbf{X}. \quad (8)$$

The augmented Lagrangian function is given by:

$$\mathcal{L}_\rho(\mathbf{X}, \mathbf{V}) = \|\mathbf{Y} - \mathbf{XBS}\|_F^2 + \|\mathbf{Z} - \mathbf{RX}\|_F^2 + \rho\|\mathbf{X} - \mathbf{V}\|_F^2 + \mu\|\mathbf{D}(\mathbf{v} - \tilde{\mathbf{x}})\|^2 + \nu\|\mathbf{E}(\mathbf{v} - \tilde{\mathbf{x}})\|^2 \quad (9)$$

where ρ is a penalty parameter. HQS method then minimizes (9) via the following steps:

$$\mathbf{X}_{k+1} = \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{XBS}\|_F^2 + \|\mathbf{Z} - \mathbf{RX}\|_F^2 + \rho\|\mathbf{X} - \mathbf{V}_k\|_F^2 \quad (10)$$

$$\mathbf{v}_{k+1} = \min_{\mathbf{v}} \rho\|\mathbf{x}_{k+1} - \mathbf{v}\|^2 + \mu\|\mathbf{D}(\mathbf{v} - \tilde{\mathbf{x}})\|^2 + \nu\|\mathbf{E}(\mathbf{v} - \tilde{\mathbf{x}})\|^2 \quad (11)$$

Optimization w.r.t. \mathbf{X}

To solve sub-problem (10), we set the gradient of the objective function (10) w.r.t. \mathbf{X} to zero. Thus, \mathbf{X}_{k+1} is the solution of the Sylvester equation:

$$\mathbf{C}_1 \mathbf{X}_{k+1} + \mathbf{X}_{k+1} \mathbf{C}_2 = \mathbf{C}_3 \quad (12)$$

where

$$\begin{aligned} \mathbf{C}_1 &= \mathbf{R}^T \mathbf{R} + \mu \mathbf{I}_B \\ \mathbf{C}_2 &= (\mathbf{BS})(\mathbf{BS})^T \\ \mathbf{C}_3 &= \mathbf{R}^T \mathbf{Z} + \mathbf{Y}(\mathbf{BS})^T + \rho \mathbf{V}_k \end{aligned} \quad (13)$$

and \mathbf{I}_B is the identity matrix.

For a fast algorithm solving (12), the interested reader can refer to [Wei et al. 2015].

Optimization w.r.t. \mathbf{v}

We can rewrite the objective function in (11) in 3D image domain with a sum running over spectral channels:

$$\min_{\mathcal{V}} \sum_{\ell=1}^B \left(\rho \|\mathcal{X}_{k+1,\ell} - \mathcal{V}_\ell\|_F^2 + \mu \|\mathcal{D}_\ell *_{2D} (\mathcal{V}_\ell - \tilde{\mathcal{X}}_\ell)\|_F^2 + \nu \|\mathcal{E}_0 *_{1D} (\mathcal{V} - \tilde{\mathcal{X}})\|_F^2 \right) \quad (14)$$

The operator $*_{2D}$ denotes 2D spatial convolution while $*_{1D}$ represents 1D spectral convolution. Using 2D DFT in the spatial domain, we can rewrite (14) as:

$$\min_{\underline{\mathcal{V}}} \sum_{\ell=1}^B \left(\rho \|\underline{\mathcal{X}}_{k+1,\ell} - \underline{\mathcal{V}}_\ell\|_F^2 + \mu \|\underline{\mathcal{D}}_\ell \odot (\underline{\mathcal{V}}_\ell - \underline{\tilde{\mathcal{X}}}_\ell)\|_F^2 + \nu \|\mathcal{E}_0 *_{1D} (\underline{\mathcal{V}} - \underline{\tilde{\mathcal{X}}})\|_F^2 \right) \quad (15)$$

with \odot the Hadamard product and underline symbols denoting Fourier transform.

Optimization w.r.t. \mathbf{v}

The optimization problem in (11) can be decomposed into a set of independent sub-problems with each spatial frequency variable \mathbf{f} :

$$\min_{\mathbf{v}_f} \|\mathbf{x}_{k+1,\mathbf{f}} - \mathbf{v}_f\|^2 + \mu' \|\Delta_{\mathcal{D}}(\mathbf{f})(\mathbf{v}_f - \tilde{\mathbf{x}}_f)\|^2 + \nu' \|\mathbf{E}_0(\mathbf{v}_f - \tilde{\mathbf{x}}_f)\|^2 \quad (16)$$

where $\mu' = \mu/\rho$, $\nu' = \nu/\rho$ and

$$\begin{aligned} \mathbf{v}_f &= \{\underline{\mathcal{V}}_\ell(\mathbf{f}), \ell = 1, \dots, B\}, \quad \tilde{\mathbf{x}}_f = \{\tilde{\underline{\mathcal{X}}}_\ell(\mathbf{f}), \ell = 1, \dots, B\} \\ \mathbf{x}_{k+1,\mathbf{f}} &= \{\underline{\mathcal{X}}_{k+1,\ell}(\mathbf{f}), \ell = 1, \dots, B\}, \quad \Delta_{\mathcal{D}}(\mathbf{f}) = \text{diag} \{\underline{\mathcal{D}}_\ell(\mathbf{f}), \ell = 1, \dots, B\} \end{aligned} \quad (17)$$

For each \mathbf{f} , the solution of (16) can be computed as:

$$\mathbf{v}_f = \mathbf{T}_f^{-1}(\mathbf{x}_{k+1,\mathbf{f}} + \mu' \Delta_{\mathcal{D}}(\mathbf{f})^* \Delta_{\mathcal{D}}(\mathbf{f}) \tilde{\mathbf{x}}_f + \nu' \mathbf{E}_0^* \mathbf{E}_0 \tilde{\mathbf{x}}_f) \quad (18)$$

where $*$ denotes the complex conjugate and $\mathbf{T}_f = (\mathbf{I}_B + \mu' \Delta_{\mathcal{D}}(\mathbf{f})^* \Delta_{\mathcal{D}}(\mathbf{f}) + \nu' \mathbf{E}_0^* \mathbf{E}_0)$.

Finally, we can obtain \mathbf{v}_{k+1} by calculating the inverse 2D DFT in the spatial domain.

Experimental Results

- Datasets: CAVE [Yasuma et al. 2010] and Harvard [Chakrabarti and Zickler 2011].
- Setups:
 - 1 **B**: an 32×32 uniform blurring operator;
 - 2 **S**: a down-sampling operator with the factor 32;
 - 3 **R**: the response of a Nikon D700 camera.
- Methods to calculate $\tilde{\mathbf{X}}$ for each \mathbf{X} :
 - 1 UAL [Zhang et al. 2020] considering deep learning;
 - 2 NSSR [Dong et al. 2016] based on sparse decomposition;
 - 3 LTTR [Dian et al. 2019] using tensor factorization.
- The code is made available at github.com/xiuheng-wang.

Experimental Results

Quantitative comparisons:

Table: Averaged RMSE, PSNR, SAM, ERGAS and SSIM of different methods on the CAVE and Harvard data sets.

Methods	CAVE data set					Harvard data set				
	RMSE	PSNR	ERGAS	SAM	SSIM	RMSE	PSNR	ERGAS	SAM	SSIM
UAL	1.854	44.656	0.196	4.33	0.9910	1.833	45.807	0.323	3.58	0.9832
UAL + Ours	1.587	45.939	0.171	4.08	0.9917	1.784	46.034	0.316	3.54	0.9833
NSSR	2.236	43.439	0.244	5.22	0.9849	1.874	45.540	0.363	3.73	0.9821
NSSR + Ours	2.068	44.044	0.230	5.19	0.9854	1.844	45.649	0.357	3.69	0.9822
LTTR	2.300	43.277	0.249	5.50	0.9848	1.914	45.251	0.375	3.81	0.9813
LTTR + Ours	2.235	43.613	0.243	5.27	0.9851	1.887	45.392	0.374	3.77	0.9915

It can be observed that our algorithm significantly improved the performance of all the baselines.

Experimental Results

Qualitative comparisons:

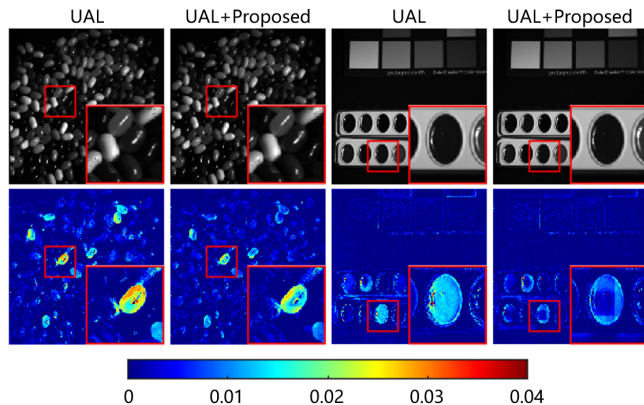









Figure: Reconstructed images and corresponding error maps of two images from the CAVE data set in the 540 nm band.

It confirms that our approach produced smaller reconstruction errors than the UAL.


We introduced an HSI super-resolution based on a deep prior regularizer:

- Deep priors are used to constrain degradation model inversion in the form of a regularizer designed in the spatial and spectral gradient domain.
- Experiments showed the performance improvement achieved with this strategy compared with state-of-the-art methods.

References I

-  Akhtar, Naveed, Faisal Shafait, and Ajmal Mian (2014). “Sparse spatio-spectral representation for hyperspectral image super-resolution”. In: *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Springer, pp. 63–78.
-  Chakrabarti, Ayan and Todd Zickler (2011). “Statistics of real-world hyperspectral images”. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 193–200.
-  Dian, Renwei, Shutao Li, and Leyuan Fang (2019). “Learning a low tensor-train rank representation for hyperspectral image super-resolution”. In: *IEEE Trans. Neural Netw. Learn. Syst.* 30.9, pp. 2672–2683.
-  Dong, Weisheng, Fazuo Fu, Guangming Shi, Xun Cao, Jinjian Wu, Guangyu Li, and Xin Li (2016). “Hyperspectral image super-resolution via non-negative structured sparse representation”. In: *IEEE Trans. Image Process.* 25.5, pp. 2337–2352.
-  Henrot, Simon, Charles Soussen, and David Brie (2012). “Fast positive deconvolution of hyperspectral images”. In: *IEEE Trans. Image Process.* 22.2, pp. 828–833.
-  Lanaras, Charis, Emmanuel Baltsavias, and Konrad Schindler (2015). “Hyperspectral super-resolution by coupled spectral unmixing”. In: *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, pp. 3586–3594.
-  Palsson, Frosti, Johannes R Sveinsson, and Magnus O Ulfarsson (2017). “Multispectral and hyperspectral image fusion using a 3-D-convolutional neural network”. In: *IEEE Geosci. Remote Sens. Lett.* 14.5, pp. 639–643.

References II

-  Vella, Marija, Bowen Zhang, Wei Chen, and João F. C. Mota (2021). “Enhanced Hyperspectral Image Super-Resolution via RGB Fusion and TV-TV Minimization”. In: *Proc. IEEE Int. Conf. Image Process. (ICIP)*, pp. 3837–3841. DOI: 10.1109/ICIP42928.2021.9506715.
-  Wang, Xiuheng, Jie Chen, Qi Wei, and Cédric Richard (2021). “Hyperspectral Image Super-Resolution via Deep Prior Regularization with Parameter Estimation”. In: *IEEE Trans. Circuits Syst, Video Technol.*
-  Wei, Qi, Nicolas Dobigeon, and Jean-Yves Tournieret (2015). “Fast fusion of multi-band images based on solving a Sylvester equation”. In: *IEEE Trans. Image Process.* 24.11, pp. 4109–4121.
-  Xie, Weiyang, Jie Lei, Yuhang Cui, Yunsong Li, and Qian Du (2019). “Hyperspectral pansharpening with deep priors”. In: *IEEE Trans. Neural Netw. Learn. Syst.*
-  Yasuma, Fumihito, Tomoo Mitsunaga, Daisuke Iso, and Shree K Nayar (2010). “Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum”. In: *IEEE Trans. Image Process.* 19.9, pp. 2241–2253.
-  Zhang, Lei, Jiangtao Nie, Wei Wei, Yanning Zhang, Shengcai Liao, and Ling Shao (2020). “Unsupervised Adaptation Learning for Hyperspectral Imagery Super-Resolution”. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 3073–3082.

Thanks for your attention !