

Blind Multi-Spectral Image Pan-Sharpener

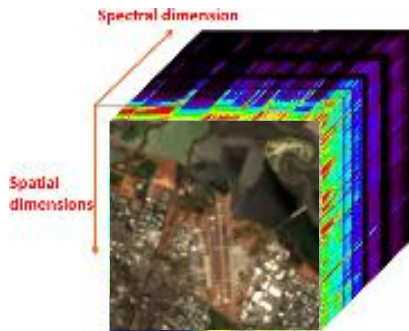
Lantao Yu, Dehong Liu, Hassan Mansour,
Petros T. Boufounos and Yanting Ma

Rice University
Houston, Texas, USA

Mitsubishi Electric Research Laboratories (MERL)
Cambridge, Massachusetts, USA
www.merl.com

Background:

- Multi-spectral imagery (MS) covers a wide range of spectrum but is with low spatial resolution.
- Panchromatic imagery (PAN) is with high spatial resolution, but is likely **NOT well aligned** with the MS.



Covering the same area,
but not well aligned



Problem:

Given low-resolution MS and not well aligned high-resolution PAN, how can we enhance the resolution of MS?

Limitations of Existing methods:

Model-Based:

- The blur kernel estimation is often flawed.
- The cross-channel relationship is not well-exploited.

Learning-Based:

- Difficult to gain enough training data, especially well-aligned data.
- The trained model from one sensor platform's data may not perform well for another sensor platform' data.

[1] C. Bajaj and T. Wang, "Blind hyperspectral-multispectral image fusion via graph laplacian regularization," arXiv:1902.08224, 2019.

[2] M. Simões, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspacebased regularization," TGRS 2014.

[3] X. Fu, Z. Lin, Y. Huang, and X. Ding, "A variational pan-sharpening with local gradient constraints," CVPR 2019.

[4] S. Lohit, D. Liu, H. Mansour, and P. Boufounos, "Unrolled projected gradient descent for multi- spectral image fusion," ICASSP 2019.

Idea:

Simultaneous registration and pan-sharpening via cross-channel prior for the PAN-MS relationship and total generalized variation for the blur kernel.

Mathematical formulation:

$$\min_{\mathbf{Z}, \mathbf{u}} \frac{1}{2} \|\mathbf{X} - \mathbf{D}\mathbf{B}(\mathbf{u})\mathbf{Z}\|_{\mathbb{F}}^2 + \mathbf{R}_1(\mathbf{Z}, \mathbf{Y}) + \mathbf{R}_2(\mathbf{u})$$

$$\mathbf{X} \in \mathbb{R}^{hw \times N}$$

Measured low-resolution MS image with N spectral bands

$$\mathbf{Y} \in \mathbb{R}^{HW \times 1}$$

High-resolution PAN

$$\mathbf{Z} \in \mathbb{R}^{HW \times N}$$

Well-aligned and high-resolution MS of consistent sharpness with PAN

$$\mathbf{D} \in \mathbb{R}^{hw \times HW}$$

Down-sampling operator

$$\mathbf{u} \in \mathbb{R}^{n^2 \times 1}$$

Kernel coefficients in vectorized form

$$\mathbf{B}(\mathbf{u}) \in \mathbb{R}^{HW \times HW}$$

Toeplitz matrix of the blur kernel \mathbf{u}

Motivation:

- Cross-Channel (PAN-MS) Image Prior should be described on high-frequency domain.
- High-frequency components across bands roughly follow a local affine function.

$$\mathbf{R}_1(\mathbf{Z}, \mathbf{Y}) = \frac{\lambda}{2} \sum_{i,j} \sum_{k \in \omega_j} \left([\mathcal{L}(\mathbf{Z}_i)]_{j,k} - a_{i,j} [\mathcal{L}(\mathbf{Y})]_{j,k} - c_{i,j} \right)^2$$

\mathbf{Z}_i	the i^{th} band of the target high-resolution MS image \mathbf{Z}
ω_j	the j^{th} square window of size $(2r + 1) \times (2r + 1)$ in a $H \times W$ image
k	the k^{th} element within the window, $k = 1, 2, \dots, (2r + 1)^2$
$a_{i,j}, c_{i,j}$	constant coefficients of the linear affine transform in window ω_j within \mathbf{Z}_i
$\mathcal{L}(\cdot)$	Laplacian operator $\mathcal{L}(\mathbf{Z}_i) = \mathbf{Z}_i \circledast \mathbf{S}$
λ	scalar

$$\mathbf{S} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

[1] Xueyang Fu, Zihuang Lin, Yue Huang, and Xinghao Ding, "A variational pan-sharpening with local gradient constraints," CVPR 2019.

[2] Kaiming He, Jian Sun, and Xiaoou Tang, "Guided image filtering," PAMI 2012.

Motivation:

- Blur kernel should be non-negative, smooth, sparse, and normalized to unit sum.
- Current ℓ_1 -based regularizer on the gradient of kernel coefficients often force small gradient to be 0.

$$\mathbf{R}_2(\mathbf{u}) = \min_{\mathbf{p}} \{ \alpha_1 \|\nabla \mathbf{u} - \mathbf{p}\|_{2,1} + \alpha_2 \|\mathcal{E}(\mathbf{p})\|_{2,1} \} + \mathbf{I}_{\mathcal{S}}(\mathbf{u})$$

$$\mathbf{u} \in \mathbb{R}^{n^2 \times 1}$$

$$\nabla \mathbf{u} = [\nabla_h \mathbf{u} \quad \nabla_v \mathbf{u}] \in \mathbb{R}^{n^2 \times 2}$$

$$\mathbf{p} \in \mathbb{R}^{n^2 \times 2}$$

$$\mathcal{E}(\mathbf{p}) = \left[\nabla_h \mathbf{p}_1 \quad \frac{\nabla_v \mathbf{p}_1 + \nabla_h \mathbf{p}_2}{2} \quad \frac{\nabla_v \mathbf{p}_1 + \nabla_h \mathbf{p}_2}{2} \quad \nabla_v \mathbf{p}_2 \right] \in \mathbb{R}^{n^2 \times 4}$$

$$\|\mathbf{X}\|_{2,1} = \sum_{i=1}^n \sqrt{\sum_{j=1}^m x_{i,j}^2}$$

$$\alpha_1, \alpha_2$$

$$\mathcal{S} = \left\{ \mathbf{s} \in \mathbb{R}^{n^2 \times 1} \mid s_i \geq 0, \sum_i s_i = 1 \right\}$$

$$\mathbf{I}_{\mathcal{S}}(\cdot)$$

Kernel coefficients in vectorized form

Horizontal and vertical gradients

Ancillary variable for the gradients of \mathbf{u}

First order derivative of \mathbf{p}

$\ell_{2,1}$ norm

Scalars

Simplex

Indicator function

$$\Phi(\mathbf{Z}, \mathbf{u}, \mathbf{p}, \mathbf{A}, \mathbf{C}, \mathbf{x}, \mathbf{y}, \mathbf{z}, \Lambda_1, \Lambda_2, \Lambda_3) =$$

$$\sum_{i=1}^N \left[\frac{1}{2} \|\mathbf{X}_i - \mathbf{DB}(\mathbf{z})\mathbf{Z}_i\|_2^2 + \frac{\lambda}{2} \sum_j \sum_{k \in \omega_j} ([\mathcal{L}(\mathbf{Z}_i)]_{j,k} - a_{i,j}[\mathcal{L}(\mathbf{Y})]_{j,k} - c_{i,j})^2 \right] +$$

$$\alpha_1 \|\mathbf{x}\|_{2,1} + \frac{\alpha_1 \mu_1}{2} \|\mathbf{x} - (\nabla \mathbf{u} - \mathbf{p}) - \Lambda_1\|_F^2 +$$

$$\alpha_2 \|\mathbf{y}\|_{2,1} + \frac{\alpha_2 \mu_2}{2} \|\mathbf{y} - \mathcal{E}(\mathbf{p}) - \Lambda_2\|_F^2 +$$

$$\mathbf{I}_{\mathcal{S}}(\mathbf{z}) + \frac{\mu_3}{2} \|\mathbf{z} - \mathbf{u} - \Lambda_3\|_2^2$$

$$\text{s.t. } \mathbf{x} = \nabla \mathbf{u} - \mathbf{p} \quad \mathbf{y} = \mathcal{E}(\mathbf{p}) \quad \mathbf{z} = \mathbf{u}$$

$$\mu_1, \mu_2, \mu_3 > 0$$

1. Initialize $\mathbf{x}^t, \mathbf{y}^t, \mathbf{z}^t, \mathbf{u}^t, \mathbf{p}^t, \mathbf{Z}^t$

2. Solve $\mathbf{x}^{t+1} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_{2,1} + \frac{\mu_1}{2} \|\mathbf{x} - (\nabla \mathbf{u}^t - \mathbf{p}^t) - \Lambda_1^t\|_F^2$

3. Solve $\mathbf{y}^{t+1} = \underset{\mathbf{y}}{\operatorname{argmin}} \|\mathbf{y}\|_{2,1} + \frac{\mu_2}{2} \|\mathbf{y} - \mathcal{E}(\mathbf{p}^t) - \Lambda_2^t\|_F^2$

4. Solve $\mathbf{z}^{t+1} = \underset{\mathbf{z}}{\operatorname{argmin}} \sum_{i=1}^N \frac{1}{2} \|\mathbf{DB}(\mathbf{z})\mathbf{Z}_i - \mathbf{X}_i\|_2^2 + \frac{\mu_3}{2} \|\mathbf{z} - \mathbf{u}^t - \Lambda_3^t\|_2^2 + \mathbf{I}_S(\mathbf{z})$

5. Solve $(\mathbf{u}^{t+1}, \mathbf{p}^{t+1}) = \underset{\mathbf{u}, \mathbf{p}}{\operatorname{argmin}} \frac{\alpha_1 \mu_1}{2} \|\mathbf{x}^t - (\nabla \mathbf{u} - \mathbf{p}) - \Lambda_1^t\|_F^2 + \frac{\alpha_2 \mu_2}{2} \|\mathbf{y}^t - \mathcal{E}(\mathbf{p}) - \Lambda_2^t\|_F^2$

6. Solve $(a_{i,j}, c_{i,j}) = \underset{a_{i,j}, c_{i,j}}{\operatorname{argmin}} \sum_j \sum_{k \in \omega_j} \left([\mathcal{L}(\mathbf{Z}_i^t)]_{j,k} - a_{i,j} [\mathcal{L}(\mathbf{Y})]_{j,k} - c_{i,j} \right)^2$

7. Solve $\mathbf{z}_i^{t+1} = \underset{\mathbf{z}_i}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{DB}(\mathbf{u}^{t+1})\mathbf{z}_i - \mathbf{X}_i\|_2^2 + \frac{\lambda}{2} \|\mathcal{L}(\mathbf{z}_i) - \hat{\mathbf{L}}_i^z\|_2^2$ where $\hat{\mathbf{L}}_i^z = \bar{\mathbf{A}}_i \cdot \mathcal{L}(\mathbf{Y}) + \bar{\mathbf{C}}_i$

8. Update $\Lambda_1^{t+1} = \Lambda_1^t + \mu(\nabla \mathbf{u}^{t+1} - \mathbf{p}^{t+1} - \mathbf{x}^{t+1})$
 $\Lambda_2^{t+1} = \Lambda_2^t + \mu(\mathcal{E}(\mathbf{p}^{t+1}) - \mathbf{y}^{t+1})$
 $\Lambda_3^{t+1} = \Lambda_3^t + \mu(\mathbf{u}^{t+1} - \mathbf{z}^{t+1})$

9. Update $\mathbf{x}^t, \mathbf{y}^t, \mathbf{z}^t, \mathbf{u}^t, \mathbf{p}^t, \mathbf{Z}^t = \mathbf{x}^{t+1}, \mathbf{y}^{t+1}, \mathbf{z}^{t+1}, \mathbf{u}^{t+1}, \mathbf{p}^{t+1}, \mathbf{Z}^{t+1}$

10. Iterate until $\|\mathbf{Z}^{t+1} - \mathbf{Z}^t\|_F / \|\mathbf{Z}^t\|_F$ is smaller than a threshold, or t is larger than a threshold

Solution of \mathbf{x} , \mathbf{y} -subproblem

$$\mathbf{x}^{t+1} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_{2,1} + \frac{\mu_1}{2} \|\mathbf{x} - (\nabla \mathbf{u}^t - \mathbf{p}^t) - \mathbf{\Lambda}_1^t\|_F^2$$

$$\mathbf{y}^{t+1} = \underset{\mathbf{y}}{\operatorname{argmin}} \|\mathbf{y}\|_{2,1} + \frac{\mu_2}{2} \|\mathbf{y} - \mathcal{E}(\mathbf{p}^t) - \mathbf{\Lambda}_2^t\|_F^2$$

Soft Thresholding the l^{th} row of \mathbf{x}^{t+1} and \mathbf{y}^{t+1}

$$\mathbf{x}^{t+1}(l) = \operatorname{shrink}_2(\nabla \mathbf{u}(l) - \mathbf{p}^t(l) + \mathbf{\Lambda}_1^t(l), \frac{1}{\mu_1})$$

$$\mathbf{y}^{t+1}(l) = \operatorname{shrink}_2(\mathcal{E}(\mathbf{p}^t)(l) + \mathbf{\Lambda}_2^t(l), \frac{1}{\mu_2})$$

$$\operatorname{shrink}_2(\mathbf{e}, t) = \max(\|\mathbf{e}\|_2 - t, 0) \frac{\mathbf{e}}{\|\mathbf{e}\|_2}$$

$$\min_{\mathbf{z}} \sum_{i=1}^N \frac{1}{2} \|\mathbf{D}\mathbf{B}(\mathbf{z})\mathbf{Z}_i - \mathbf{X}_i\|_2^2 + \frac{\mu_3}{2} \|\mathbf{z} - \mathbf{u}^t - \mathbf{\Lambda}_3^t\|_2^2$$

1. Reformulate the subproblem as:

$$\min_{\mathbf{z}} \sum_{i=1}^N \frac{1}{2} \|\mathbf{D}\mathcal{C}(\mathbf{Z}_i)\mathbf{z} - \mathbf{X}_i\|_2^2 + \frac{\mu_3}{2} \|\mathbf{z} - \mathbf{u} - \mathbf{\Lambda}_3\|_2^2$$

since $\mathbf{B}(\mathbf{z})\mathbf{Z}_i = \mathbf{z} \circledast \mathbf{Z}_i = \mathbf{Z}_i \circledast \mathbf{z} = \mathcal{C}(\mathbf{Z}_i)\mathbf{z}$.

\mathcal{C} : a Toeplitz matrix corresponding to the convolution

2. Solve $\min_{\mathbf{z}} \sum_{i=1}^N \frac{1}{2} \|\mathbf{D}\mathcal{C}(\mathbf{Z}_i)\mathbf{z} - \mathbf{X}_i\|_2^2 + \frac{\mu_3}{2} \|\mathbf{z} - \mathbf{u} - \mathbf{\Lambda}_3\|_2^2$ via conjugated gradients

3. Project the above solution onto Simplex \mathbb{S}

[1] Jonathan Richard Shewchuk et al., "An introduction to the conjugate gradient method without the agonizing pain," 1994.

[2] Weiran Wang and Miguel A Carreira-Perpinán, "Projection onto the probability simplex: An efficient algorithm with a simple proof, and an application," arXiv preprint arXiv:1309.1541, 2013.

Solution of \mathbf{u} , \mathbf{p} -subproblem

$$(\mathbf{u}^{t+1}, \mathbf{p}^{t+1}) = \underset{\mathbf{u}, \mathbf{p}}{\operatorname{argmin}} \frac{\alpha_1 \mu_1}{2} \|\mathbf{x}^t - (\nabla \mathbf{u} - \mathbf{p}) - \Lambda_1^t\|_F^2 + \frac{\alpha_2 \mu_2}{2} \|\mathbf{y}^t - \mathcal{E}(\mathbf{p}) - \Lambda_2^t\|_F^2$$

1. Let $\mathbf{q} = [\mathbf{u}^\top \ \mathbf{p}_1^\top \ \mathbf{p}_2^\top]^\top$

2. Enforce first-order necessary condition, we get:

$$\Sigma \mathbf{q} = \mathbf{b}$$

Σ :diagonal block-Toeplitz matrix

\mathbf{q} :can be computed by FFT and inverse FFT

[1] Weihong Guo, Jing Qin, and Wotao Yin, "A new detail- preserving regularization scheme," SIAM journal on imaging sciences, 2014.

$$\min_{a_{i,j}, c_{i,j}} \sum_j \sum_{k \in \omega_j} \left([\mathcal{L}(\mathbf{Z}_i^t)]_{j,k} - a_{i,j} [\mathcal{L}(\mathbf{Y})]_{j,k} - c_{i,j} \right)^2 + \epsilon a_{i,j}^2$$

Guided Imag Filtering

guide image: $\mathcal{L}(\mathbf{Y})$

input image: $\mathcal{L}(\mathbf{Z}_i^t)$

[1] Kaiming He, Jian Sun, and Xiaoou Tang, "Guided image filtering," PAMI 2012.

$$\min_{\mathbf{Z}_i} \frac{1}{2} \|\mathbf{DB}(\mathbf{u}^{t+1})\mathbf{Z}_i - \mathbf{X}_i\|_2^2 + \frac{\lambda}{2} \|\mathcal{L}(\mathbf{Z}_i) - \hat{\mathbf{L}}_i^z\|_2^2$$

where $\hat{\mathbf{L}}^z = \bar{\mathbf{A}}_i \cdot \mathbf{LY} + \bar{\mathbf{C}}_i$

$$\mathcal{L}(\mathbf{Z}_i) = \mathbf{LZ}_i$$

$$\mathcal{L}(\mathbf{Y}) = \mathbf{LY}$$

\mathbf{L} : Toeplitz matrix of Laplacian filter

$$\mathbf{Z}_i = (\mathbf{B}^\top \mathbf{D}^\top \mathbf{DB} + \lambda \mathbf{L}^\top \mathbf{L})^{-1} (\mathbf{B}^\top \mathbf{D}^\top \mathbf{X} + \lambda \mathbf{L}^\top \hat{\mathbf{L}}^z)$$

Conjugate Gradients

or a Fast Algorithm accelerated by FFT

[1] Jonathan Richard Shewchuk et al., "An introduction to the conjugate gradient method without the agonizing pain," 1994.

[2] Ningning Zhao, Qi Wei, Adrian Basarab, Nicolas Dobigeon, Denis Kouamé, and Jean-Yves Tournet, "Fast single image super-resolution using a new analytical solution for l2 - l2 problems," IEEE Transactions on Image Processing, 2016.

Motivation:

- A reliable initialization of the blur kernel can avoid being trapped by bad local minima.
- Assume the stacked PAN as the perfect target MS in terms of position.

$$\min_{\mathbf{u}, \mathbf{p}} \sum_i^{N_0} \frac{1}{2} \|\mathbf{D}\mathcal{C}(\mathbf{Y})\mathbf{u} - \mathbf{X}_i\|_2^2 + \alpha_1 \|\nabla\mathbf{u} - \mathbf{p}\|_{2,1} + \alpha_2 \|\mathcal{E}(\mathbf{p})\|_{2,1} + \mathbf{I}_S(\mathbf{u})$$

N_0 : number of MS bands whose electro-magnetic spectrum overlaps with PAN

Verification of Local Laplacian Prior in Guided Image Upsampling

$$\min_{\mathbf{Z}} \frac{1}{2} \|\mathbf{X} - \mathbf{DZ}\|_{\mathbf{F}}^2 + \mathbf{R}_1(\mathbf{Z}, \mathbf{Y})$$

Experimental Setting:

1. The blur kernel is a δ -function
2. The input MS image is from downsampling the ground-truth MS image by a factor of 2
3. PAN image is already well-aligned with MS
4. Metric: Average PSNR, computed by averaging the PSNR in each channel
5. Dataset: *Pavia University*

Local Laplacian Prior (Ours) : 37.57 dB

Local Gradient Constraints(LGC): 37.33 dB

[1] Xueyang Fu, Zihuang Lin, Yue Huang, and Xinghao Ding, "A variational pan-sharpening with local gradient constraints," CVPR 2019

Verification of Second-Order Generalized Total Variation in Blur Kernel Estimation

Experimental Setting:

1. Ground Truth Blur Kernel: $\mathbf{K}(i, j) = e^{-[(i-x)^2 + (j-y)^2]/(2\sigma^2)}$
 $-r \leq i \leq r, -r \leq j \leq r$
 $n = 19, x = 1.33, y = 0.42, \sigma = 2$

2. Test Image: PAN of *West of Sichuan* from IKONOS

3. Metric: $\epsilon_r = \|\mathbf{K} - \hat{\mathbf{K}}\|_F / \|\mathbf{K}\|_F$

4. Solve $\min_{\mathbf{u}} \frac{1}{2} \|\mathbf{E}\mathbf{u} - \mathbf{f}\|_2^2 + \alpha \|\nabla \mathbf{u}\|_{2,1} + \mathbf{I}_S(\mathbf{u})$ for Isotropic Total Variation

5. Solve $\min_{\mathbf{u}, \mathbf{p}} \frac{1}{2} \|\mathbf{E}\mathbf{u} - \mathbf{f}\|_2^2 + \alpha_1 \|\nabla \mathbf{u} - \mathbf{p}\|_{2,1} + \alpha_2 \|\mathcal{E}(\mathbf{p})\|_{2,1} + \mathbf{I}_S(\mathbf{u})$

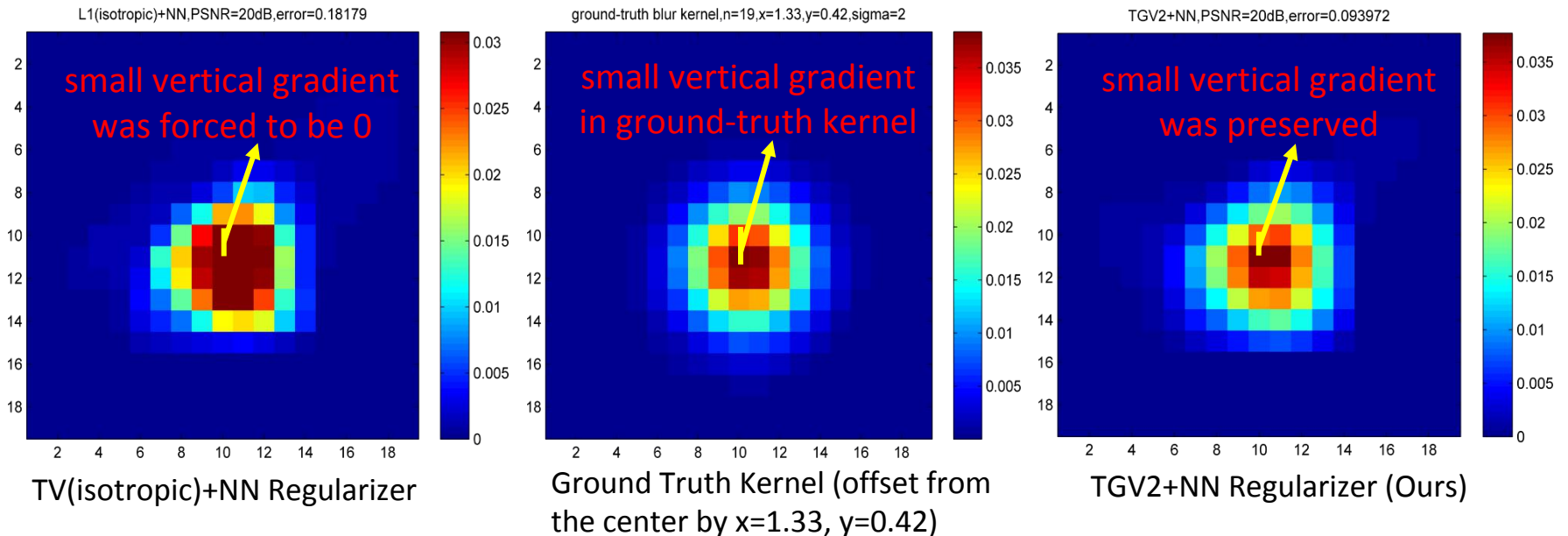
for Second-Order Generalized Total Variation

\mathbf{f} : the blurred, downsampled, noisy version of PAN

\mathbf{E} : the measurement matrix of \mathbf{u} for generating \mathbf{f}

Table 1: Relative Errors corresponding to Different Regularizers and Different Noise Levels

PSNR/dB	Isotropic Total Variation	Second-Order Generalized Total Variation
10	0.2904	0.1607
20	0.1818	0.0940
30	0.1008	0.0520
40	0.0502	0.0288



INPUT



low-res MS

high-res PAN

(only RGB channels
are shown)

OUTPUT



Pan-Sharpended MS

Ground-Truth



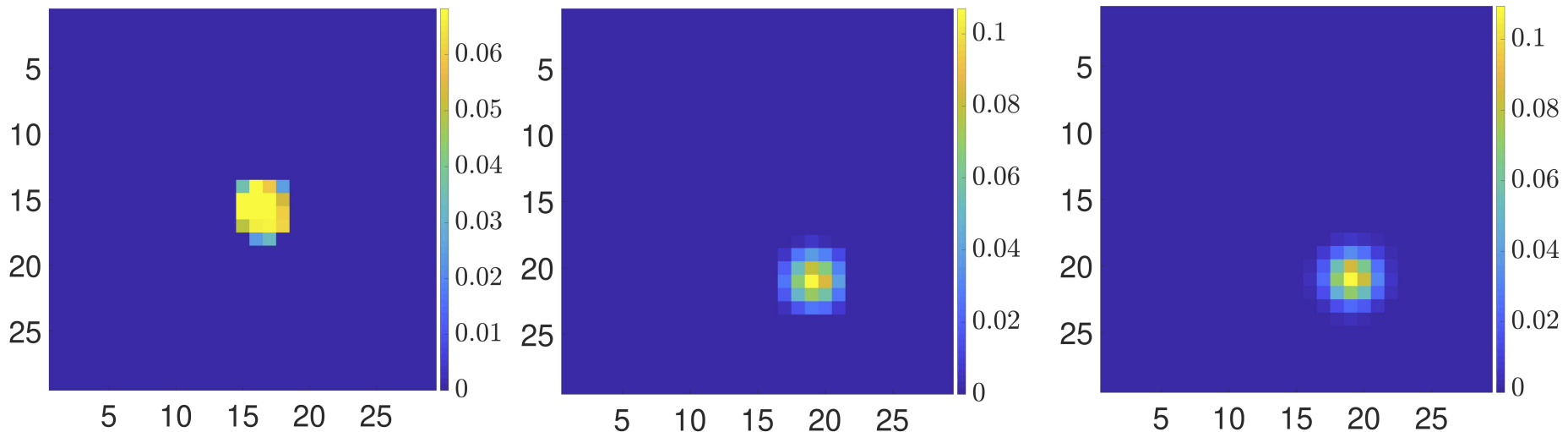
Simulated True
high-res MS

Table 2: Quantitative analysis of blind pan-sharpening results

Approach	BHMIFGLR	HySure	Ours
Exp. 1/Exp. 2	31.72/21.38	30.71/30.70	37.40/37.40

Exp. 1: offset $x=0.87$, $y=0.11$

Exp. 2: offset $x=5.87$, $y=4.11$



Comparison of estimated kernels in Exp. 2 using
(left) BHMIFGLR, (middle) Ours, and (right) Ground-Truth

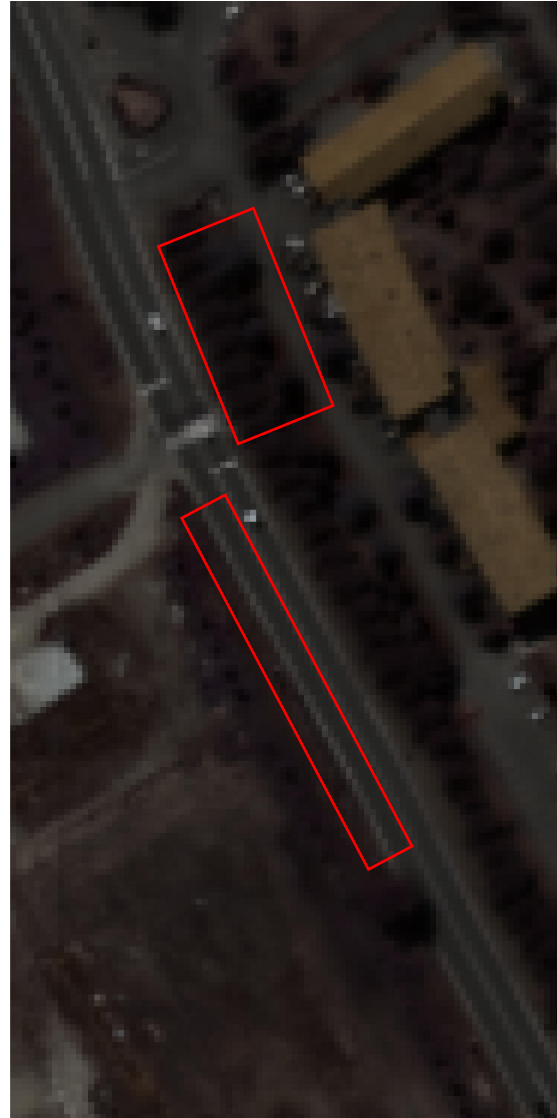
[1] Chandrajit Bajaj and Tianming Wang, "Blind hyperspectral- multispectral image fusion via graph laplacian regularization," arXiv preprint arXiv:1902.08224, 2019.

[2] Miguel Simões, José Bioucas-Dias, Luis B Almeida, and Jocelyn Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," TGRS 2014.

Fusion Results



Reconstructed RGB via BHMIFGLR

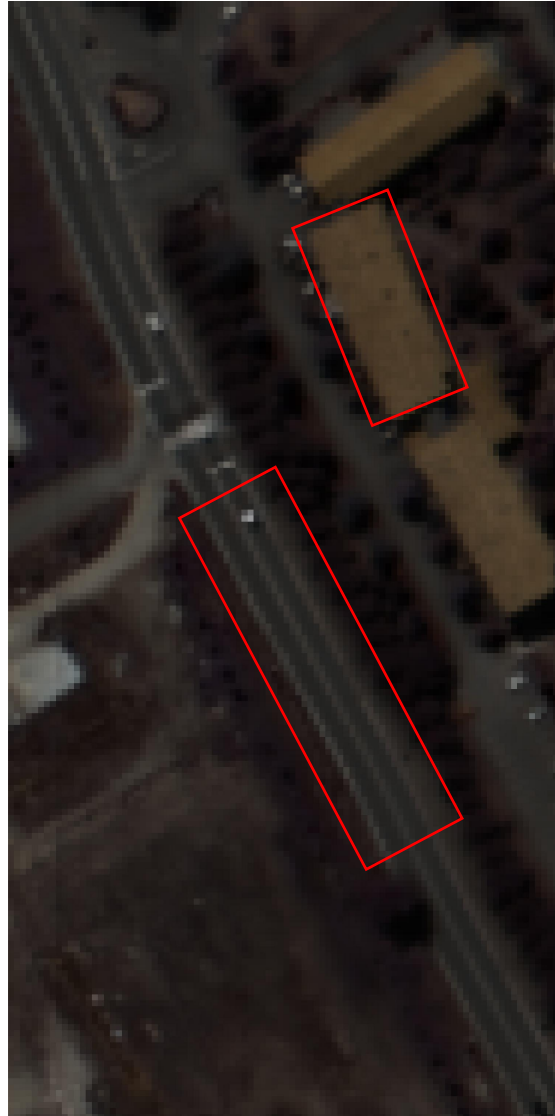
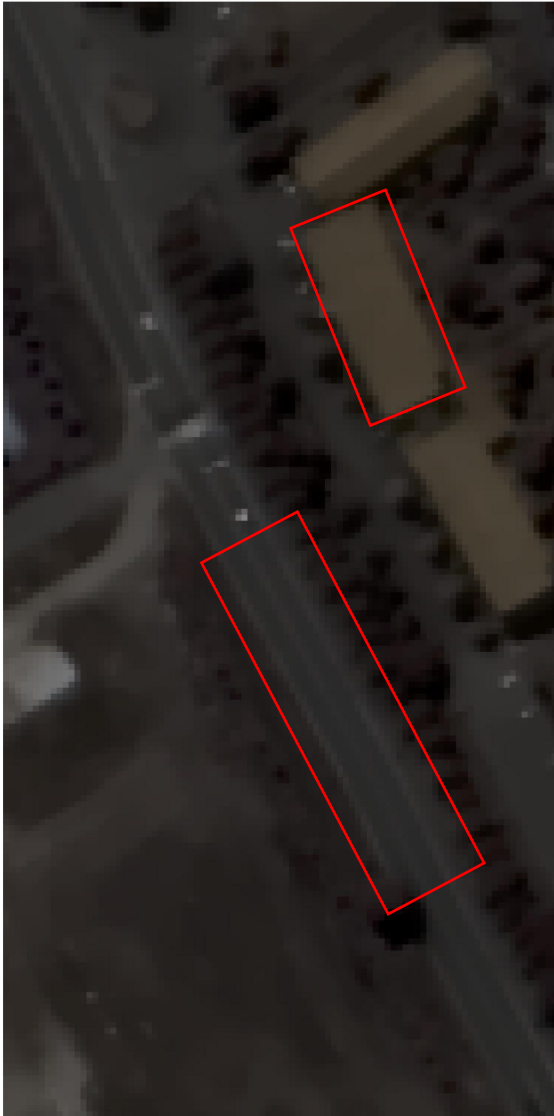


Reconstructed RGB via Ours



Ground-Truth

Fusion Results



Reconstructed RGB via HySure

Reconstructed RGB via Ours

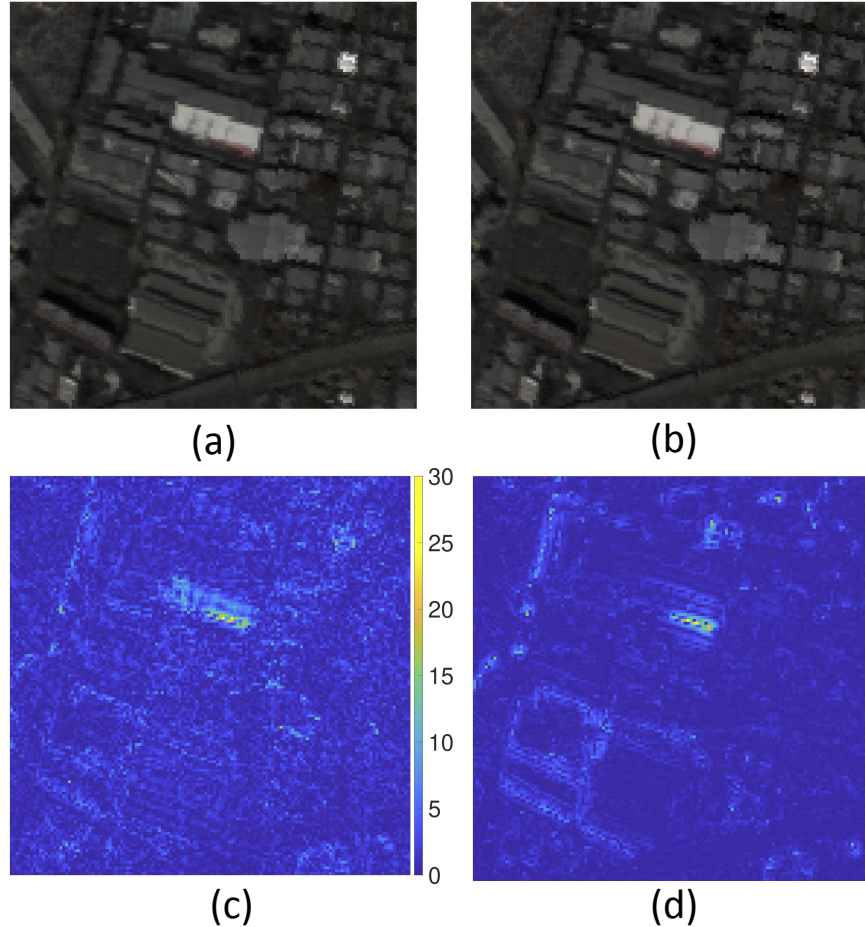
Ground-Truth

Table 3: Average PSNR Comparisons between our approach and a *deep-learning based approach* using test Images: Moffett, Cuprite, Los Angles(L.A.) and Cambria Fire (C.F.) from AVIRIS Data.

Test Images	Moffett	Cuprite	L.A.	C.F.	Mean
Ours	39.94	41.17	38.53	38.91	39.64
UPGD	38.17	39.02	37.77	39.33	38.57

[1] Suhas Lohit, Dehong Liu, Hassan Mansour, and Petros T Boufounos,
“Unrolled projected gradient descent for multi-spectral image fusion,” ICASSP 2019

Fusion Results



Comparison of fused MS images in RGB channels using (a) UPGD and (b) Ours. (c) and (d) are the green channel residual images of (a) and (b) compared to the ground truth.

Our approach outperforms UPGD, especially in smooth areas.

Take-Away Message

- The cross channel-relationship should focus on the **high-frequency** components of MS and PAN image.
- ℓ_1 -based regularizer on the blur kernel is limited because it will force small gradients of the kernel coefficients be 0. We can use **higher-order generalized total variation** to improve the performance.