

## Pan-sharpening with a Hyper-Laplacian Penalty

Yiyong Jiang, Xinghao Ding, Delu Zeng, Yue Huang, John Paisley<sup>†</sup>  
Fujian Key Laboratory of Sensing and Computing for Smart City, Xiamen University  
<sup>†</sup>Department of Electrical Engineering, Columbia University

### Abstract

*Pan-sharpening is the task of fusing spectral information in low resolution multispectral images with spatial information in a corresponding high resolution panchromatic image. In such approaches, there is a trade-off between spectral and spatial quality, as well as computational efficiency. We present a method for pan-sharpening in which a sparsity-promoting objective function preserves both spatial and spectral content, and is efficient to optimize. Our objective incorporates the  $\ell_{1/2}$ -norm in a way that can leverage recent computationally efficient methods, and  $\ell_1$  for which the alternating direction method of multipliers can be used. Additionally, our objective penalizes image gradients to enforce high resolution fidelity, and exploits the Fourier domain for further computational efficiency. Visual quality metrics demonstrate that our proposed objective function can achieve higher spatial and spectral resolution than several previous well-known methods with competitive computational efficiency.*

### 1. Introduction

Multispectral (MS) data provided by satellite optical sensors are useful for many practical applications such as environmental monitoring, object positioning and classification. Because of physical constraints, most remote sensors measure a panchromatic (PAN) image (i.e., gray-scale image) that is high-resolution, and several low resolution multispectral (LRMS) images containing information about RGB colors and the non-visible spectrum. Pan-sharpening is the problem of fusing this low resolution spectral information with the spatial structure in the PAN image to output an approximation of the unmeasured high resolution multispectral (HRMS) images [2]. The trade-offs encountered by such methods are spectral versus structural preservation, as well as computational efficiency.

Various pan-sharpening methods have been developed, the most common being based on a projection-substitution approach where the PAN image is assumed to be equivalent to a linear combination of the HRMS images [18]. Many

of these methods are appealing for their straightforward implementation and fast computation [23, 21, 12], but exhibit spectral distortion as a trade-off [29].

To address the spectral distortion problem, methods based on the concept of *Amélioration de la Résolution Spatiale par Injection de Structures* (ARSIS) have been proposed [22, 19, 20], in which multi-resolution tools such as wavelets and Laplacian pyramids are used to extract details from the PAN image and inject them into the MS images. A suite of model-based fusion methods have also been proposed to address the spectral distortion issue. These methods treat the fusion problem as an image restoration model, with several additional regularization schemes [25, 4, 11, 7, 3]. Another model-based line of work has considered dictionary learning [18, 9, 15], which requires substantial computational resources.

Pan-sharpening methods often use the  $\ell_2$ -norm to enforce spatial resolution, or switch to  $\ell_1$  when sparsity is desired. When used to penalize image gradients, it has been shown that according to empirical image statistics, such Gaussian ( $\ell_2$ ) or Laplacian ( $\ell_1$ ) assumptions are not as appropriate as hyper-Laplacian assumptions [17], which correspond to  $\ell_p$ -norms with  $0 < p < 1$ . Nevertheless, as with a variety of signal processing applications,  $\ell_1$  is often used as the closest convex relaxation of the sparser non-convex  $\ell_p$  norms, with compressed sensing being a prominent example. Indeed, in the compressed sensing problem it has been shown that the  $\ell_1$  convex relaxation often shares the same solution with the desired  $\ell_0$  norm [10].

In this paper we consider the hyper-Laplacian penalty for the pan-sharpening problem as part of a larger objective function. We apply the  $\ell_{1/2}$  penalty on the gradients of the reconstruction error to enforce structural preservation. Using a recently developed efficient learning algorithm for  $\ell_p$  penalties when  $p = 1/2$  [17], we demonstrate that the statistically more appropriate hyper-Laplacian penalty does indeed translate to an improvement in performance for pan-sharpening. In other words, though the solutions are only locally optimal, they are consistently better than the global optimal solutions afforded by a less statistically appropriate convex  $\ell_1$  penalty.



Figure 1. Examples from 208 images used in our experiments.

We formulate our objective function to consider the following aspects:

1. *Spectral preservation* using the assumption that the LRMS images are decimated from the HRMS images by convolution with a blurring kernel.
2. *Structural preservation* by using an  $\ell_{1/2}$ -norm on the gradient of the error between the PAN image and a linear combination of the HRMS reconstructions. Additionally, we use anisotropic total variation as a penalty on the reconstructed HRMS images.

We define our objective function in Section 2. In Section 3, we show how the alternating direction method of multipliers (ADMM) [5, 13], Fourier transform, and closed form  $\ell_{1/2}$  algorithm allow for efficient local optimization of our non-convex problem. We then demonstrate the superior performance of the  $\ell_{1/2}$ -norm in the context of pan-sharpening, as well as the importance of the other components of our objective function on a large set of MS images in Section 4.

## 2. Pan-sharpening with sparse gradients

Images obtained by remote sensors contain both a high resolution panchromatic (PAN) image (i.e., black and white) and low resolution multispectral (LRMS) images consisting of  $B$  spectral bands (for example, red, green, blue and a near infrared, in which case  $B = 4$ ). Pan-sharpening aims to obtain HRMS images by fusing the PAN image with the LRMS images. This is generally an ill-posed problem, and so further constraints are necessary on the desired properties of the reconstructed HRMS images.

We view the LRMS images as decimated version of the desired HRMS images with additive noise. Let the decimated image of each spectral band be  $M \times N$  in size and the

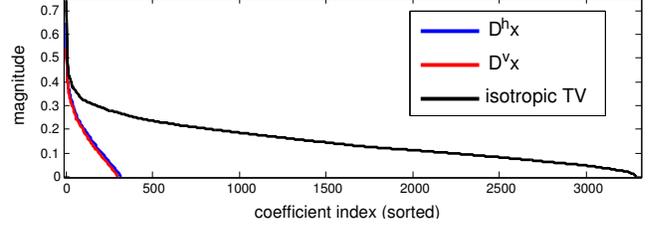


Figure 2. Magnitude of (blue and red) anisotropic TV directional coefficients and (black) isotropic TV coefficients in decreasing order for the image in Figure 4.

desired high resolution version be, e.g.,  $4M \times 4N$ , giving a size reduction ratio of 16:1 (as in our experiments). For notation, we vectorize these images, letting  $y_b \in \mathbb{R}^{16MN}$  correspond to LRMS image at band  $b$ ,  $x_b \in \mathbb{R}^{16MN}$  the corresponding unknown HRMS image, and  $y_P \in \mathbb{R}^{16MN}$  the measured PAN image. The objective function we define for learning the set  $\{x_b\}$  is of the general form

$$\mathcal{L} = f(\{x_b\}, \{y_b\}, y_P) + \lambda \sum_{b=1}^B \|x_b\|_{aTV}. \quad (1)$$

The function  $f(\{x_b\}, \{y_b\}, y_P)$  is a data fidelity term that we define in Section 2.1 with the goal of spectral and spatial preservation and computational efficiency.

The anisotropic total variation term  $\|x_b\|_{aTV}$  is often used to encourage a low-noise reconstruction that doesn't penalize high frequency edge information;  $\lambda > 0$  is the corresponding regularization parameter that controls the trade-off with  $f$ . Using the vectorized notation, the anisotropic TV is written

$$\|x\|_{aTV} = \sum_i \|D_i x\|_1 \quad (2)$$

where  $D_i$  is a  $2 \times 16MN$  matrix that has two nonzero entries in each row corresponding to finite difference in the vertical and horizontal directions, and the summation ranges over the pixel indexes.

We use anisotropic TV instead of isotropic TV since it tends to perform better [6]. Furthermore, the sparsity with anisotropic TV is greater, which is better for focusing on image edges [14]. Figure 2 presents a comparison of the sparsity along the two dimensions of  $D_i x$  (denoted  $D_i^h x$  and  $D_i^v x$ ) and the isotropic TV measure on one image used in our experiments (shown in Figure 4). As is evident, anisotropic TV is significantly more sparse than the isotropic TV, which leads to a reduction in the resolution of the LRMS images necessary to achieve comparable performance.

### 2.1. The data fidelity term

The function  $f$  in Equation (1) enforces the consistency of the reconstructed  $x_b$  to the measured data  $y_b$  and  $y_P$ . We

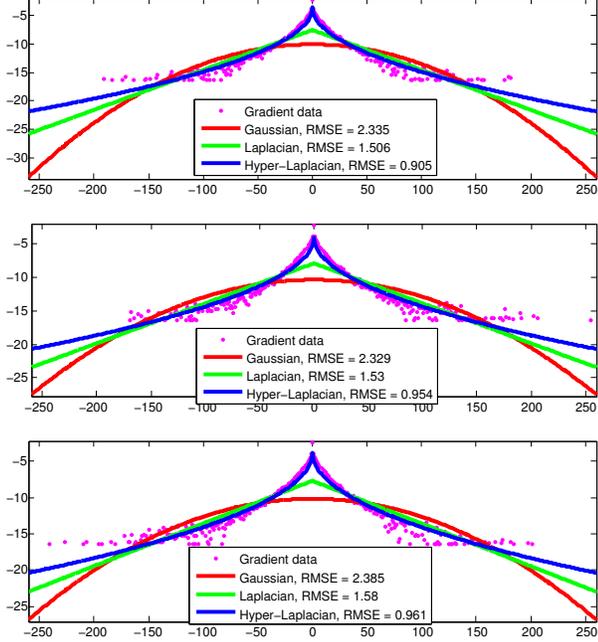


Figure 3. Fitting curves to empirical image gradient data. (magenta) The empirical gradient data, (red) fitting a Gaussian ( $\ell_2$ ) penalty, (green) a Laplacian ( $\ell_1$ ) penalty and (blue) a hyper-Laplacian ( $\ell_p$ ) penalty with  $p = 1/2$ . As is evident and motivated in [17], the gradients of image data require a sparse penalty, but one with a heavier tail than  $\ell_1$ .

break this into the sum of two terms,  $f = \frac{v_1}{2} f_1 + \frac{v_2}{2} f_2$ , intended to preserve the spectral and spatial information in  $y_b$  and  $y_P$  respectively.

### 2.1.1 Spectral preservation

We define the spectral penalty term to be

$$f_1(\{x_b\}, \{y_b\}) = \sum_{b=1}^B \|k * x_b - y_b\|_2^2. \quad (3)$$

This requires the LRMS image  $y_b$  be approximately a decimated version of HRMS image  $x_b$  via convolution with a blurring kernel  $k$ . This preserves the spectral information in the observed LRMS images. For  $k$ , [18] uses an averaging kernel while [16, 24] estimate  $k$  on a per-satellite basis. We use the first in our experiments, but note that both gave comparable results.

### 2.1.2 Structural preservation

For the structure-preserving portion of  $f$ , we define

$$f_2(\{x_b\}, y_P) = \sum_i \|G_i(\sum_{b=1}^B \omega_b x_b - y_P)\|_{1/2}. \quad (4)$$

The weight vector  $\omega$  represents the PAN image as an average of the HRMS images. The matrix  $G_i$  denotes a differential operator along the horizontal, vertical and two diagonal directions, which we will show has advantages over the

two-directional gradient. This term enforces structural consistency between the PAN image and the linear combination of HRMS reconstructions. The corresponding term in other algorithms often use a squared error penalty, which can lead to spectral distortion [2].

It has been observed that the gradient of real-world images is better fit by a heavy-tailed distribution such as a hyper-Laplacian (which has density  $p(x) \propto e^{-\kappa|x|^p}$ ,  $0 < p < 1$ ) [17]. To test this, we collected 208 PAN images with known HRMS images and rescaled these images to 0-255 (examples shown in Figure 1). For each, we fit various distributions to the histogram of  $G_i(\sum_b \omega_b x_b - y_P)$ ; specifically a Gaussian ( $\ell_2$ ), Laplacian ( $\ell_1$ ) and hyper-Laplacian with  $p = 1/2$  ( $\ell_{1/2}$ ). As shown in Figure 3 on three typical examples, the hyper-Laplacian fits these residuals the best. We therefore believe that the  $\ell_{1/2}$ -norm is more reasonable than the  $\ell_2$ -norm [3] or  $\ell_1$ -norm [7] for the structural fidelity term. We also compare their performance in Section 4.

## 3. Optimization

We want to minimize the following objective function with respect to the HRMS images  $x_b$  for each spectral band,

$$\mathcal{L} = \underbrace{\frac{\lambda}{2} \sum_{b=1}^B \sum_i \|D_i x_b\|_1}_{\text{structural preservation}} + \underbrace{\frac{v_2}{2} \sum_i \left\| G_i \left( \sum_b \omega_b x_b - y_P \right) \right\|_{1/2} + \frac{v_1}{2} \sum_{b=1}^B \|k * x_b - y_b\|_2^2}_{\text{spectral preservation}}. \quad (5)$$

The motivation for these terms was discussed in the previous section. We next discuss an algorithm for finding a local minimum of this non-convex objective function. For fast closed form updates, our strategy uses ADMM [5] separately on the  $\ell_1$  and  $\ell_{1/2}$  terms, which modifies this objective function by adding augmented Lagrangian terms.

### 3.1. Augmented Lagrangian form

We split both the structural fidelity terms and the anisotropic TV coefficients for the  $i$ th pixel by defining

$$\alpha_i := G_i(\sum_b \omega_b x_b - y_P), \quad \beta_{i,b} := D_i x_b, \quad (6)$$

respectively, and then relaxing the equality via an augmented Lagrangian. Following an intermediate step, this results in the following objective function,

$$\begin{aligned} \mathcal{L} = & \frac{v_1}{2} \sum_b \|k * x_b - y_b\|_2^2 + \frac{v_2}{2} \sum_i \|\alpha_i\|_{1/2} \\ & + \frac{\eta}{2} \sum_i \|G_i(\sum_b \omega_b x_b - y_P) - \alpha_i + e_i\|_2^2 \\ & + \frac{\ell}{2} \sum_{i,b} \|D_i x_b - \beta_{i,b} + u_{i,b}\|_2^2 + \frac{\lambda}{2} \sum_{i,b} \|\beta_{i,b}\|_1 \\ & + \text{const.} \end{aligned} \quad (7)$$

By the ADMM theory, optimizing this augmented objective using the algorithm in Section 3.2 will find a local optimal solution in which the equality constraints in Equation

---

**Algorithm 1** Outline of optimizing  $\mathcal{L}$  in (7)

---

**Iterate** the following three sub-problems to convergence

**Output** HRMS images  $x_b$  for each band

- (P1) Sec. 3.2.1: Optimize each  $\beta_{i,b}$  (total variation)
  - (P2) Sec. 3.2.2: Optimize each  $\alpha_i$  (hyper-Laplacian)
  - (P3) Sec. 3.2.3: Optimize each  $x_b$  (reconstruction)
- 

(6) are satisfied [5]. Optimizing Equation (7) reduces to iterating between three sub-problems that can be optimized individually using the most recent solutions from the other sub-problems. We sketch these three sub-problems in Algorithm 1 and present their respective updates below.

## 3.2. Algorithm

### 3.2.1 Update for P1: Total variation

We solve for each  $\beta_i$  for each pixel using the generalized shrinkage operation, followed by an update of the Lagrange multiplier [13],

$$\beta_{i,b} = \max\{\|D_i x_b + u_{i,b}\|_2 - \lambda/\rho, 0\} \cdot \frac{D_i x_b + u_{i,b}}{\|D_i x_b + u_{i,b}\|_2}$$
$$u_{i,b} \leftarrow u_{i,b} + D_i x_b - \beta_{i,b}. \quad (8)$$

Recall that  $\beta_{i,b}$  corresponds to the 2-dimensional TV coefficients for pixel  $i$  in band  $b$ , with differences in one direction vertically and horizontally. These coefficients have been split from  $D_i x_b$  using ADMM, but converge to one another as the algorithm iterates [5].

### 3.2.2 Update for P2: Hyper-Laplacian

As detailed in [17], we can optimize the four-dimensional  $\alpha_i$  element-wise in closed form by first solving for the roots of the cubic polynomial. The general form of this polynomial is

$$\alpha^3 - 2\alpha^2(v+e) + \alpha(v+e)^2 - \frac{\text{sign}(v+e)}{(4\eta/v_2)^2} = 0, \quad (9)$$

where we let  $v := G(\sum_b \omega_b x_b - y_p)$ . In this equation,  $\alpha$ ,  $e$  and  $v$  are each indexed by subscripts  $i$  to indicate the  $i$ th pixel, and also  $d = 1, \dots, 4$  to indicate the direction of the derivative in the corresponding rows of  $G$ ; there are thus four independent problems to solve, one for each dimension. For each problem, there are three roots and the best one can be selected quickly by following the discussion in [17]. After updating the dimensions of  $\alpha_i$ , we update the Lagrange multiplier vector

$$e_i \leftarrow e_i + G_i(\sum_b \omega_b x_b - y_p) - \alpha_i. \quad (10)$$

Recall that  $\alpha_i$  is split from  $G_i(\sum_b \omega_b x_b - y_p)$  and from the ADMM theory the two will converge to each other as the algorithm iterates.

### 3.2.3 Update for P3: Reconstruction

In the final sub-problem, we reconstruct the HRMS image  $x_b$  for each spectral band by solving the corresponding least squares problem efficiently in the Fourier domain. Below, we define  $K$  to be the  $16MN \times 16MN$  matrix form of the blurring kernel  $k$  constructed from its point spread function. The matrices  $G$ ,  $D$  and vectors  $\alpha$ ,  $\beta_b$ ,  $e$  and  $u_b$  are also defined by stacking their respective pixel-level components. Differentiating the objective in (7) with respect to  $x_b$  results in the normal equations for the solution of  $x_b$ ,

$$\left(v_1 K^T K + \eta \omega_b^2 G^T G + \rho D^T D\right) x_b = v_1 K^T y_b + \quad (11)$$

$$\eta \omega_b G^T \left(G(\sum_{i \neq b} \omega_i x_i - y_p) + \alpha - e\right) + \rho D^T (\beta_b - u_b)$$

We need to solve for  $x_b$ , but direct calculation is prohibited by the size of the left matrices. However, because these matrices are circulant we can solve for  $x_b$  in the Fourier domain [13]. We let  $\theta_b = \mathcal{F} x_b$  be the Fourier transform of  $x_b$ , replace  $x_b$  with  $\mathcal{F}^T \theta_b$  and take the Fourier transform of each side of Equation (11). The circulant property of  $K^T K$ ,  $G^T G$  and  $D^T D$  means they share the Fourier matrix  $\mathcal{F}$  as eigenvectors and have unique eigenvalues,  $\Lambda_1$ ,  $\Lambda_2$  and  $\Lambda_3$  respectively. As a result, we can transform Equation (11) into solving  $16MN$  one-dimensional problems in the Fourier domain for each band. Let  $\xi_b$  be the right-hand side of Equation (11). Then

$$\theta_{i,b} = (\mathcal{F}_i \xi_b) / (v_1 \Lambda_{1,i} + \eta \omega_b^2 \Lambda_{2,i} + \rho \Lambda_{3,i}), \quad (12)$$

where  $\mathcal{F}_i$  is the  $i$ th Fourier basis function. We then invert the  $16MN$ -dimensional vector  $\theta_b$  via the inverse Fourier transform to obtain the reconstruction  $x_b$ . The FFT can perform both of these computations very quickly.

## 4. Experiments

We perform several experiments with our proposed pan-sharpening objective function and compare with several current methods. We also demonstrate the value of each part of our objective by comparing with several variations on our proposed method. All experiments were performed on a PC with two Intel CPUs (3.60GHz), 16GB RAM and 64-bit Windows-7 operating system, using Matlab R2012b software. We used a  $5 \times 5$  average blurring kernel for  $k$  since we assume the scenario where we lack satellite statistics. We consider problems with four spectral bands (RGB and near IR), and set the weight parameter  $\omega_b = 0.25$  for each band.

### 4.1. Simulation

We first show results on experiments where we generate LRMS images from ground truth HRMS to evaluate the

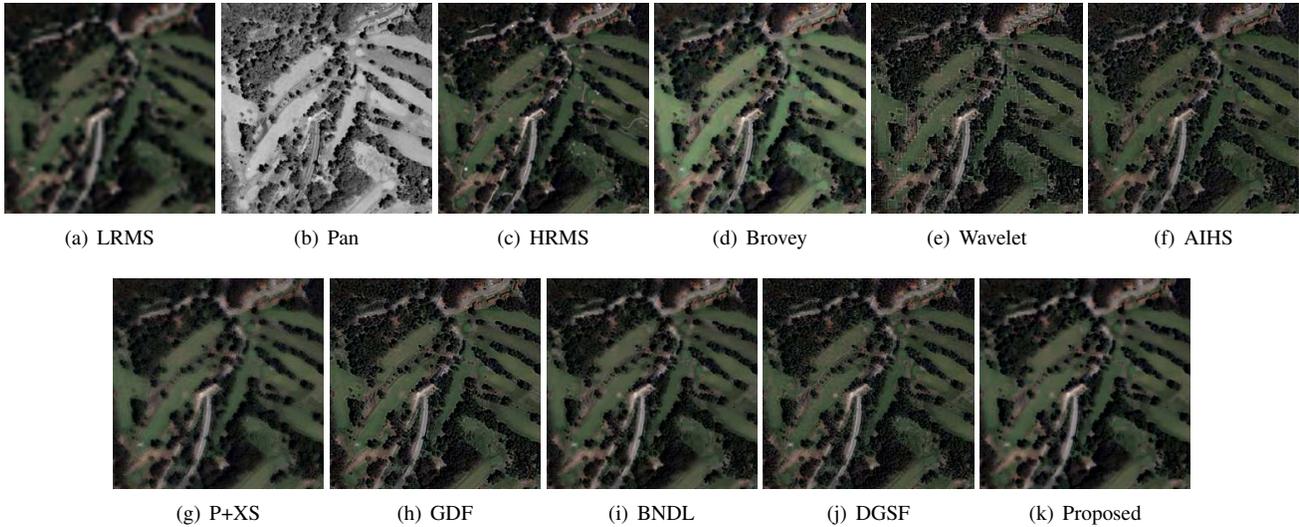


Figure 4. Comparisons of different methods on a QuickBird image (RGB shown,  $256 \times 256$ )

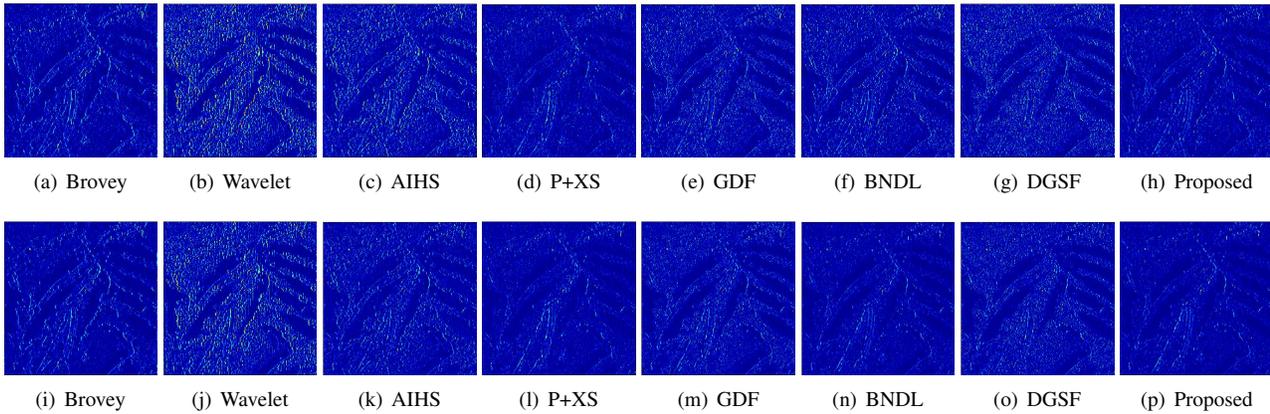


Figure 5. The residual images of (**top**) the blue band and (**bottom**) the green band of Figure 4 (blue color indicates small differences and red color large differences)

fusion results of our algorithm. In these experiments, we manually degrade the high resolution PAN image (0.6m) and MS images (2.4m) using a  $3 \times 3$  Gaussian low-pass filter and bicubic downsampling by four to yield a PAN image with a resolution of 2.4m and MS images with resolution of 9.6m. We use the original 2.4m resolution MS images as the HRMS images for reference as ground truth. We then fuse the 9.6m LRMS images with the 2.4m PAN image and compare the results with the ground truth.

**A QuickBird image.** For our first experiment, we use an image from the QuickBird satellite. We compare the performance with several classical methods: Brovey [12], Wavelet [20], AIHS [21], P+XS [4], as well as some newly proposed methods: Guidedfilter-based (GDF) [11], a Bayesian non-parametric dictionary learning (BNDL) method [9] and Dynamic Gradient Sparsity (DGSF) [7]. To create a fair com-

parison, we tuned parameters of each approach to achieve their own best results.

Figure 4 shows the fusion results and Figure 5 the residuals on the blue and green bands. We can see that Brovey suffers color distortion although its edges are distinct. Wavelet preserves spectral information well, though there are staircase effects and blurry edges. P+XS is more blurry than other methods. GDF and DGSF show details clearly, but their residuals appear worse.

In pan-sharpening spectral quality of the fused MS images is also very important, but is more difficult to judge visually. Therefore we quantitatively evaluate both the spatial and spectral quality of our fusion results using the following five standard quality metrics: relative dimensionless global error in synthesis (ERGAS), spectral angle mapper (SAM) [2], universal image quality index (QAVE) [26], relative average spectral error (RASE) [8] and correlation coefficient

Table 1. Quality metrics of different methods on 208 satellite images.

Algorithm	ERGAS	SAM	QAVE	RASE	CC	RMSE	FCC	Q4
Brovey	19.60±7.57	7.00±1.63	0.57±0.09	50.1±12.1	0.93±0.02	0.137±0.040	<b>0.97±0.01</b>	0.63±0.09
Wavelet	7.71±1.52	8.70±2.06	0.66±0.11	31.0±7.0	0.92±0.03	0.078±0.013	0.92±0.02	0.72±0.06
AIHS	7.96±1.77	7.57±1.75	0.68±0.11	31.2±7.6	0.93±0.03	0.078±0.014	<b>0.97±0.01</b>	0.70±0.07
P+XS	8.18±2.01	10.61±3.08	0.65±0.12	30.4±6.7	0.93±0.03	0.077±0.015	0.96±0.01	0.73±0.07
GDF	7.06±1.43	8.64±2.46	0.69±0.11	28.8±7.1	0.93±0.02	0.072±0.013	<b>0.97±0.01</b>	0.75±0.08
BNDL	6.72±1.49	7.68±1.91	0.69±0.09	27.1±7.0	0.94±0.03	0.068±0.013	<b>0.97±0.01</b>	0.76±0.08
DGSF	6.72±1.39	7.27±1.87	0.72±0.11	26.9±6.6	0.94±0.02	0.067±0.013	0.96±0.02	0.77±0.08
Proposed	<b>5.88±1.31</b>	<b>6.80±1.43</b>	<b>0.76±0.07</b>	<b>24.2±6.1</b>	<b>0.95±0.02</b>	<b>0.061±0.013</b>	<b>0.97±0.01</b>	<b>0.80±0.04</b>
<b>ideal value</b>	0	0	1	0	1	0	1	1

(CC) [2]. We show these performance measures for this QuickBird problem in Table 2. We highlight the best result in bold and in the last row of the table indicate the target value.

It can be seen that the proposed method (Prop) performs very well compared with other methods. In addition, we compare with several variations on our objective function: Prop-1 uses  $\sum_b G_i(x_b - y_P)$  instead of  $G_i(\sum_b \omega_b x_b - y_P)$  in the structural fidelity term. Prop- $\ell_1$  and Prop- $\ell_2$  indicate that we replace the  $\ell_{1/2}$ -norm with  $\ell_1$  and  $\ell_2$  respectively. “Prop w/o G” indicates that  $G_i$  is replaced with the identity matrix, meaning the penalty is directly on the reconstruction error, not its gradient. We also consider replacing the  $G_i$  with “ $G2_i$ ,” written as Prop-G2, in which we only consider the horizontal and vertical directions of the gradient instead of the four-directional gradient.

As is evident, the hyper-Laplacian penalty can improve the pan-sharpened images compared with  $\ell_2$  and  $\ell_1$ . While  $\ell_1$  is sparse as well, the  $\ell_{1/2}$  penalty has heavy tails that better characterizes empirical image statistics. Also evident in these comparisons is that penalizing gradient information is important, and considering gradients in more directions can

do better. We see an advantage of focusing on the high frequency edge information in the PAN image. These results on variations of our penalty are consistent with quantitative evaluations on other images.

**Large-scale evaluation.** We also considered 208 satellite images of sizes between  $256 \times 256$  and  $512 \times 512$ , with examples shown in Figure 1. These satellite datasets are cropped from QuickBird, WorldView-2, WorldView-3<sup>1</sup> and Pléiades<sup>2</sup>. They contain four bands (RGB and near IR) with 8-bit or 16-bit data format. The images capture many interesting objects, such as urban roads, buildings, vegetation, rivers, seaside, etc. We rescale these images to 0-1 before processing. For evaluation, in addition to the metrics used in Table 2, we use other well-known metrics: root mean squared error (RMSE) [26], filtered correlation coefficients (FCC) [28] and Q4 [27]. We tune the parameters for each algorithm on a satellite-by-satellite basis, using the best values averaged across all images from a particular satellite. We show the mean and standard deviation of each measure across the 208 images in Table 1. Again we find that the proposed method is competitive with the other methods.

## 4.2. Real data from IKONOS

We also experiment with real data acquired by the IKONOS satellite<sup>3</sup>. We focus on four satellite images from the Sichuan region of China. In this real-data setting there are no HRMS images for reference, and so we use the down-sampling/upsampling strategy described in [1] using bicubic interpolation for quantitative evaluation.

We show quantitative performance measurements in Tables 3 to 6, where we compare with several methods: Brovey [12], Wavelet [20], AIHS [21], P+XS [4], GDF [11], BNDL [9], DGSF [7] and MBF [3]. We see that the proposed method performs best on all experiments using the measures ERGAS, RASE, CC. For two experiments, MBF obtains the best result using the SAM and QAVE measures.

<sup>1</sup><http://www.digitalglobe.com/product-samples>

<sup>2</sup><http://www.mapmart.com/Samples.aspx>

<sup>3</sup><http://glcf.umd.edu/data/ikonos/>

Table 2. Quality metrics corresponding to Figure 4

Algorithm	ERGAS	SAM	QAVE	RASE	CC
Brovey	18.85	7.842	0.765	48.266	0.916
Wavelet	9.546	10.32	0.852	24.958	0.862
AIHS	7.531	7.969	0.903	20.111	0.898
P+XS	5.175	5.754	0.918	17.920	0.941
GDF	6.250	7.140	0.906	18.889	0.932
BNDL	5.568	6.254	0.930	15.902	0.941
DGSF	6.265	7.004	0.918	17.732	0.934
Prop	<b>4.842</b>	<b>5.515</b>	<b>0.952</b>	<b>15.370</b>	<b>0.952</b>
Prop-1	6.033	6.608	0.925	16.795	0.933
Prop- $\ell_1$	4.950	5.592	0.930	15.405	0.950
Prop- $\ell_2$	5.291	5.970	0.927	15.802	0.945
Prop w/o G	5.364	6.163	0.928	17.551	0.944
Prop-G2	5.008	5.682	0.930	15.430	0.950
<b>ideal value</b>	0	0	1	0	1

Table 3. IKONOS China-Sichuan 58204-0000000.20001116

	ERGAS	SAM	QAVE	RASE	CC
Brovey	11.25	5.271	0.749	32.82	0.874
Wavelet	3.945	6.084	0.846	16.76	0.891
AIHS	3.757	5.371	0.873	15.78	0.906
P+XS	5.729	8.487	0.830	20.53	0.881
GDF	3.970	6.005	0.852	17.05	0.898
BNDL	3.501	5.360	0.877	14.90	<b>0.916</b>
DGSF	3.689	5.631	0.867	15.84	0.906
MBF	3.737	<b>4.119</b>	<b>0.901</b>	14.76	0.904
Prop	<b>3.425</b>	5.093	0.883	<b>14.51</b>	<b>0.917</b>

Table 4. IKONOS China-Sichuan 58205-0000000.20001003

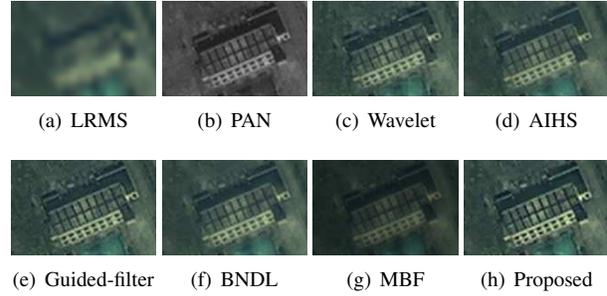
	ERGAS	SAM	QAVE	RASE	CC
Brovey	3.850	3.737	0.910	15.49	0.785
Wavelet	3.512	4.952	0.903	13.67	0.835
AIHS	2.762	3.815	0.930	11.15	0.892
P+XS	4.100	5.229	0.897	17.46	0.797
GDF	2.946	4.234	0.919	15.57	0.888
BNDL	2.682	3.837	0.935	10.95	0.905
DGSF	3.025	4.145	0.931	12.39	0.882
MBF	2.737	3.421	0.934	10.86	0.912
Prop	<b>2.464</b>	<b>3.381</b>	<b>0.938</b>	<b>10.36</b>	<b>0.922</b>

Table 5. IKONOS China-Sichuan 58207-0000000.20000831

	ERGAS	SAM	QAVE	RASE	CC
Brovey	4.221	4.049	0.886	19.57	0.836
Wavelet	3.635	5.134	0.879	14.12	0.852
AIHS	3.089	4.161	0.904	12.46	0.904
P+XS	3.421	4.322	0.900	14.58	0.880
GDF	3.218	4.509	0.899	12.92	0.886
BNDL	2.874	4.031	<b>0.913</b>	11.62	0.910
DGSF	3.006	4.141	<b>0.913</b>	11.77	0.901
MBF	3.242	4.288	0.888	12.39	0.888
Prop	<b>2.838</b>	<b>3.833</b>	<b>0.914</b>	<b>11.42</b>	<b>0.912</b>

Table 6. IKONOS China-Sichuan 58208-0000000.20001108

	ERGAS	SAM	QAVE	RASE	CC
Brovey	9.130	5.536	0.727	28.50	0.895
Wavelet	4.473	6.601	0.787	18.75	0.899
AIHS	4.470	5.695	0.809	18.51	0.912
P+XS	7.090	11.69	0.724	29.33	0.881
GDF	4.422	6.404	0.783	18.99	0.904
BNDL	3.984	5.867	0.811	16.85	0.920
DGSF	4.010	5.958	0.804	16.94	0.919
MBF	4.814	<b>5.173</b>	<b>0.819</b>	18.43	0.885
Prop	<b>3.909</b>	5.682	0.809	<b>16.36</b>	<b>0.921</b>
ideal	0	0	1	0	1

Figure 6. Pan-sharpening results for a zoomed-in ( $106 \times 162$ ) portion of IKONOS China-Sichuan 58205-0000000.20001003.

For qualitative comparisons, we show visual results for two satellite images in Figures 8 and 9. In Figure 6 we show results on a zoomed-in region of a third satellite image for further comparison. In these examples we see that some algorithms trade sharp edges for spectral distortion, while others better match color with the result of fuzziness and blurred edges. These results were consistent with other satellite images we considered.

### 4.3. Computational efficiency

In Figure 7 we compare the computational efficiency with the other algorithms considered in Tables 3 to 6 using MS images between  $256 \times 256$  and  $2048 \times 2048$  in size. The runtimes are calculated as the time necessary for approximate convergence using the same stopping criteria for all algorithms, meaning the number of iterations may vary per algorithm (not significantly according to our observations). Not shown in this figure is the runtime for BNDL, which takes several times longer than the other algorithms. Also not shown are running times for the classical algorithms Brovey, Wavelet, and AIHS, which are substantially faster (between 0.1ms to 3s), which is an attractive feature of these algorithms. When compared with more complex methods, our algorithm has a competitive-to-better running time. We observe that sub-problems P1 and P2 of our algorithm can be easily parallelized to speed up the algorithm.

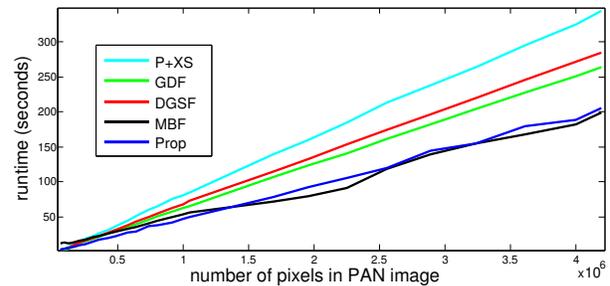


Figure 7. Runtime comparison for different size of MS images.

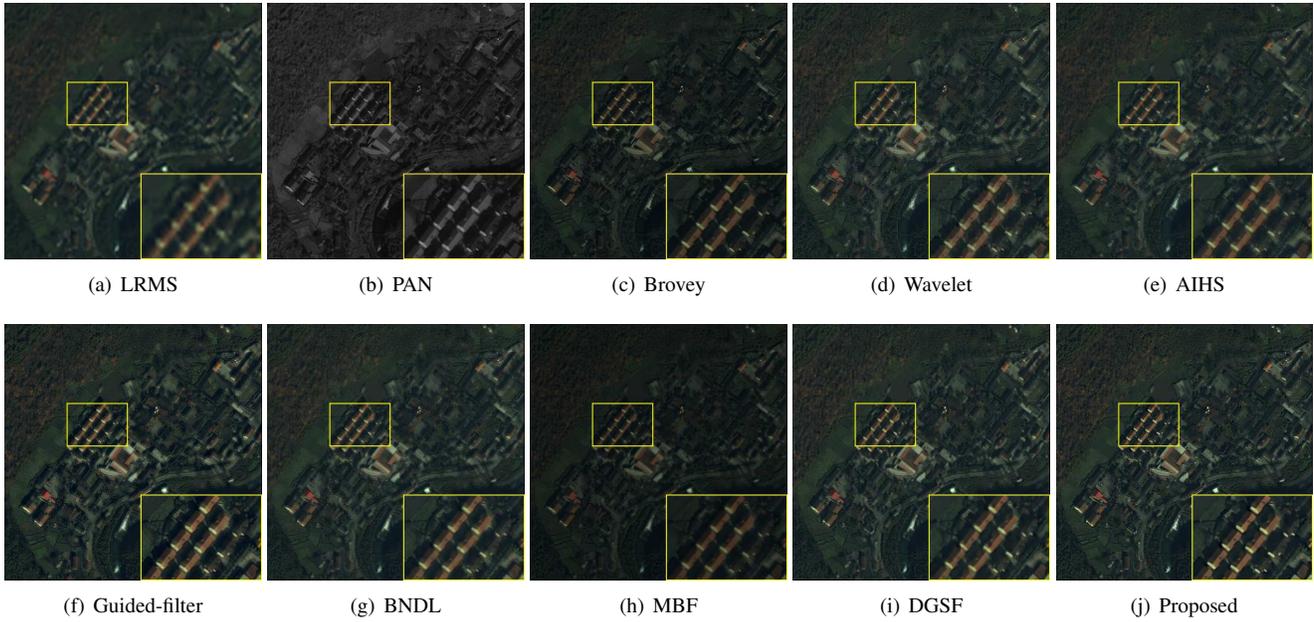


Figure 8. Pan-sharpening results for IKONOS China-Sichuan 58208-0000000.20001108 (RGB shown,  $512 \times 512$ ), along with a magnified area of each image.

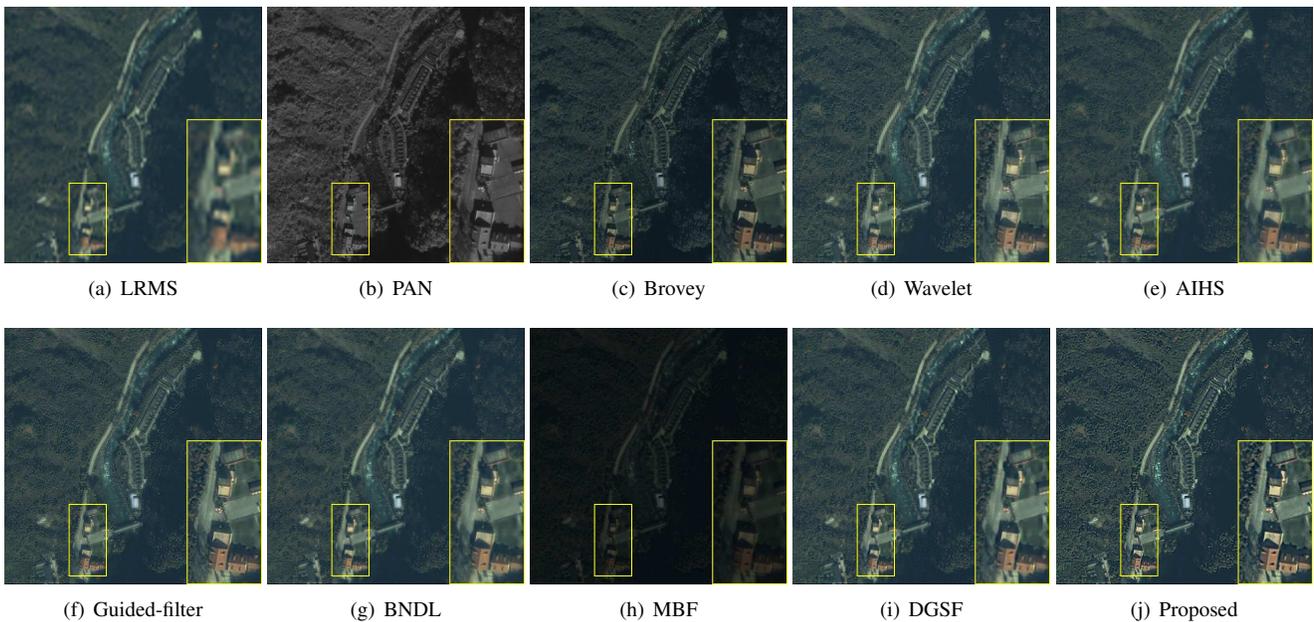


Figure 9. Pan-sharpening results for IKONOS China-Sichuan 58204-0000000.20001116 (RGB shown,  $512 \times 512$ ), along with a magnified area of each image.

## 5. Conclusion

We have proposed a pan-sharpening method that incorporates the  $\ell_{1/2}$  hyper-Laplacian penalty for image gradients. In particular, we showed how the statistically more appropriate non-convex  $\ell_{1/2}$  penalty can outperform convex  $\ell_1$  relaxations, despite having weaker convergence guarantees. Using ADMM and the Fourier domain, we derived a

computationally competitive algorithm for optimization.

**Acknowledgement.** The project was supported by the National Natural Science Foundation of China under grants 30900328, 61172179, 61103121, 81301278, 61571382 and 61571005, the Fundamental Research Funds for the Central Universities under grants 20720150169 and 20720150093, the Natural Science Foundation of Fujian Province of China under grant 2012J05160, and the Research Fund for the Doctoral Program of Higher Education under grant 20120121120043.

## References

- [1] H. Aanæs, J. Sveinsson, A. Nielsen, T. Bovith, and J. Benediktsson. Model-based satellite image fusion. *IEEE Trans. Geosci. Remote Sens.*, 46(5):1336–1346, 2008. [6](#)
- [2] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce. Comparison of pansharpening algorithms: Outcome of the 2006 grs-s data fusion contest. *IEEE Trans. Geosci. Remote Sens.*, 45(10):3012–3021, 2007. [1](#), [3](#), [5](#), [6](#)
- [3] H. A. Aly and G. Sharma. A regularized model-based optimization framework for pan-sharpening. *IEEE Trans. On Image Processing*, 23(6):2596–2608, 2014. [1](#), [3](#), [6](#)
- [4] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rougé. A variational model for p+ xs image fusion. *International J. Computer Vision*, 69(1):43–58, 2006. [1](#), [5](#), [6](#)
- [5] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011. [2](#), [3](#), [4](#)
- [6] T. Chan, S. Esedoglu, F. Park, and A. Yip. Recent developments in total variation image restoration. *Mathematical Models of Computer Vision*, 17, 2005. [2](#)
- [7] C. Chen, Y. Li, W. Liu, and J. Huang. Image fusion with local spectral consistency and dynamic gradient sparsity. In *Computer Vision and Pattern Recognition*, 2014. [1](#), [3](#), [5](#), [6](#)
- [8] M. Choi. A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter. *IEEE Trans. Geosci. Remote Sens.*, 44(6):1672–1682, 2006. [5](#)
- [9] X. Ding, Y. Jiang, Y. Huang, and J. Paisley. Pan-sharpening with a Bayesian nonparametric dictionary learning model. In *International Conference on Artificial Intelligence and Statistics*, 2014. [1](#), [5](#), [6](#)
- [10] D. Donoho. Compressed sensing. *IEEE Trans. on Information Theory*, 52(4):1289–1306, 2006. [1](#)
- [11] F. Fang, F. Li, C. Shen, and G. Zhang. A variational approach for pansharpening. *IEEE Trans. Image Process.*, 22(7):2822–2834, 2013. [1](#), [5](#), [6](#)
- [12] A. R. Gillespie, A. B. Kahle, and R. E. Walker. Color enhancement of highly correlated images. I. Decorrelation and HSI contrast stretches. *Remote Sensing of Environment*, 20(3):209–235, 1986. [1](#), [5](#), [6](#)
- [13] T. Goldstein and S. Osher. The split bregman method for L1 regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009. [2](#), [4](#)
- [14] M. Grasmair and F. Lenzen. Anisotropic total variation filtering. *Applied Mathematics & Optimization*, 62(3):323–339, 2010. [2](#)
- [15] C. Jiang, H. Zhang, H. Shen, and L. Zhang. A practical compressed sensing-based pan-sharpening method. *IEEE Geosci. Remote Sens. Lett.*, 9(4):629–633, 2012. [1](#)
- [16] Y. Jiang, L. Chen, W. Wang, X. Ding, and Y. Huang. A compressed sensing-based pan-sharpening using joint data fidelity and blind blurring kernel estimation. In *IEEE International Conference on Image Processing*, 2014. [3](#)
- [17] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-Laplacian priors. In *Advances in Neural Information Processing Systems*, 2009. [1](#), [3](#), [4](#)
- [18] S. Li and B. Yang. A new pansharpening method using a compressed sensing technique. *IEEE Trans. Geosci. Remote Sens.*, 49(2):738–746, 2011. [1](#), [3](#)
- [19] P. Massip, P. Blanc, and L. Wald. A method to better account for modulation transfer functions in arsis-based pansharpening methods. *IEEE Trans. Geosci. Remote Sens.*, 50(3):800–808, 2012. [1](#)
- [20] X. Otazu and M. Gonzalez-Ausicana. Introduction of sensor spectral response into image fusion methods: Application to wavelet-based methods. *IEEE Trans. Geosci. Remote Sens.*, 43(10):2376–2385, 2005. [1](#), [5](#), [6](#)
- [21] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman. An adaptive ihs pansharpening method. *IEEE Geosci. Remote Sens. Lett.*, 7(4):746–750, 2010. [1](#), [5](#), [6](#)
- [22] T. Ranchin and L. Wald. Fusion of high spatial and spectral resolution images: The arsis concept and its implementation. *Photogramm. Eng. Remote Sens.*, 66(1):49–61, 2000. [1](#)
- [23] T. M. Tu, P. S. Huang, C. L. Hung, and C. P. Chang. A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery. *IEEE Geosci. Remote Sens. Lett.*, 1(4):309–312, 2004. [1](#)
- [24] G. Vivone, M. Simoes, M. D. Mura, R. Restaino, J. M. Bioucas-Dias, G. A. Licciardi, and J. Chanussot. Pansharpening based on semiblind deconvolution. *IEEE Geosci. Remote Sensing*, 53(4):1997–2010, 2015. [3](#)
- [25] W. Wang, L. Jiao, and S. Yang. Fusion of multispectral and panchromatic images via sparse representation and local autoregressive model. *Inf. Fusion*, 20:73–87, 2014. [1](#)
- [26] Z. Wang and A. C. Bovik. Modern image quality assessment. *Synthesis Lectures on Image, Video, and Multimedia Processing*, 2(1):1–156, 2006. [5](#), [6](#)
- [27] Z. Wang and A. C. Bovik. A universal image quality index. *IEEE Signal Processing Letters*, 9(3):81–84, 2002. [6](#)
- [28] J. Zhou, D. L. Civco, and J. A. Silander. A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *International J. Remote Sensing*, 19(4):743–757, 1998. [6](#)
- [29] X. X. Zhu and R. Bamler. A sparse image fusion algorithm with application to pansharpening. *IEEE Trans. Geosci. Remote Sens.*, 51(5):2827–2836, 2013. [1](#)