High-Quality Bayesian Pansharpening

Tingting Wang, Faming Fang, Fang Li, and Guixu Zhang

***Abstract*— Pansharpening is a process of acquiring a multi-spectral image with high spatial resolution by fusing a low resolution multi-spectral image with a corresponding high resolution panchromatic image. In this paper, a new pansharpening method based on the Bayesian theory is proposed. The algorithm is mainly based on three assumptions: 1) the geometric information contained in the pan-sharpened image is coincident with that contained in the panchromatic image; 2) the pan-sharpened image and the original multi-spectral image should share the same spectral information; and 3) in each pan-sharpened image channel, the neighboring pixels not around the edges are similar. We build our posterior probability model according to above-mentioned assumptions and solve it by the alternating direction method of multipliers. The experiments at reduced and full resolution show that the proposed method outperforms the other state-of-the-art pansharpening methods. Besides, we verify that the new algorithm is effective in preserving spectral and spatial information with high reliability. Further experiments also show that the proposed method can be successfully extended to hyper-spectral image fusion.**

***Index Terms*— Pansharpening, multi-spectral image, panchromatic image, Bayesian theory, optimization model, alternating direction method of multipliers.**

I. Introduction

SATELLITE capture systems, such as Spot, GeoEye, QuickBird and IKONOS, only provide two types of images, i.e., low spatial resolution multi-spectral (LRMS) image and high spatial resolution panchromatic (PAN) image. Multi-spectral (MS) image usually covers a more precise spectral range for the spectral resolution can be up to eight bands captured in the visible and near-infrared wavelengths. For example, the MS image provided by QuickBird satellite has four bands: red, green, blue and near-infrared spectrum. In contrast, PAN image contains little spectral information. To obtain MS images with high spatial resolution, various approaches have been raised by combining a MS image with the PAN image acquired over the same area. In a more technical sense, this process is usually referred to as pansharpening [1], a branch of traditional image fusion [2]. A specific realization flowchart of pansharpening on QuickBird satellite image is shown in Fig. 1.

As a preliminary step for enhancing images in many remote sensing tasks, pansharpening has important applications varying from military aspect to scientific research, such as change detection [3], target recognition [4], vegetation mapping and disaster warning [5]. In recent years, a fast growing number of commercial products using high-resolution images (e.g., Google Earth and Auto Navi Map) have brought the increasingly demand for pan-sharpened data. Since pansharpening can make the best use of the acquired data provided by commercial optical satellites, the research is significant.

The classic pansharpening methods can be broadly divided into four classes [6]: component substitution (CS), multi-resolution analysis (MRA), Bayesian approach, and variational approach. CS methods project the MS image into another space to separate the spatial and spectral information [7], then substitute the PAN image for the components containing spatial structure [8]. The result images can be obtained by bringing the data back to the original space through the inverse transformation. Among these CS methods, the intensity-hue-saturation (IHS) [9]–[11], the principal component analysis (PCA) [12], [13], Brovey [14]–[16] and Gram-Schmidt (GS) [17], [18] methods are most popular.

Different from CS methods which make use of per-pixel channel transformations, MRA methods are based on the injection of spatial details obtained by a multi-resolution decomposition of the PAN image into the MS image bands after resampling [8]. Representative MRA methods contain decimated wavelet transform (DWT) [19], undecimated wavelet transform (UDWT) [20], "à-trous" wavelet transform (ATWT) [21], Laplacian pyramid (LP) [22], and nonseparable transforms, either based on wavelets (e.g., contourlet [23]) or not (e.g., curvelet [24]). Hybrid methods which combine CS with multi-scale decomposition have also been proposed, such as guided filter PCA (GFPCA) [25].

Bayesian model has been successfully applied to image processing, such as deblurring [26], image segmentation [27], super resolution [28], image interpolation [29], and face sketch synthesis [30]. In pan-sharpening field, the Bayesian approach is committed to finding a suitable statistical model to jointly characterize the pansharpened result and the available LRMS and PAN images. Since Fasbender *et al.* [31] proposed the pioneering pansharpening method based on Bayesian estimation theory, many contributions have flourished in the literature to cope with the usual illposedness of the pansharpening inverse problems. Some works take advantage of compressive

Manuscript received October 8, 2017; revised April 20, 2018 and June 9, 2018; accepted August 19, 2018. Date of publication August 23, 2018; date of current version September 19, 2018. This work was sponsored in part by "Chenguang Program" supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission under Grant 17CG25, in part by the Key Project of the National Natural Science Foundation of China under Grant 61731009, and in part by the National Natural Science Foundation of China under Grant 61871185 and Grant 61501188. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Lei Zhang. *(Corresponding author: Guixu Zhang.)*

T. Wang, F. Fang, and G. Zhang are with the Department of Computer Science and Technology, East China Normal University, Shanghai 200062, China (e-mail: ttwang@stu.ecnu.edu.cn; fmfang@cs.ecun.edu.cn; gxzhang@cs.ecnu.edu.cn).

F. Li is with the Department of Mathematics, East China Normal University, Shanghai 200241, China (e-mail: fli@math.ecnu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2018.2866954

1057-7149 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

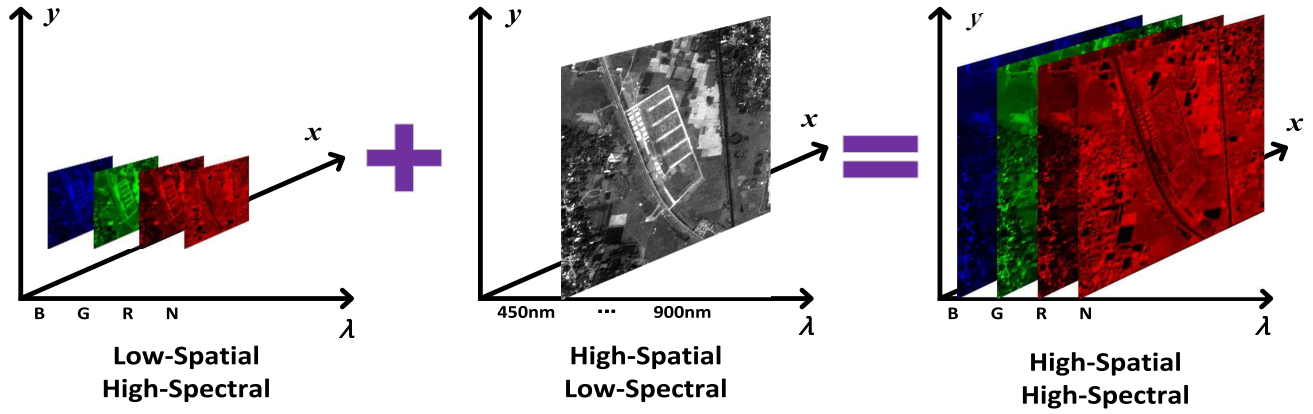


Fig. 1. The basic idea of pansharpening (here we take a four-band MS image as an example).

sensing theory [32] or sparse signal representations [33], and others introduce total variation penalization terms [34], [35].

The variational approach can be regarded as particular case of the Bayesian one [6], in which the pansharpening problem is transformed into an optimization problem by injecting the formulation with some suitable prior knowledge. In 2006, Ballester *et al*. [36] first proposed a variational model named P+XS based on the linear combination assumption that the geometric information is contained in the PAN image and the upsampled MS image is the desired result after blurring, which is an innovation in remote sensing image fusion. Due to the difficulty of estimating an accurate blur kernel, VWP (variational wavelet pansharpening) and AVWP (alternate VWP) [37] were presented for sharping high dimensional images by replacing this term with a spectral ratio constraint and ensuring the approximation/detail wavelet coefficients of fused image are closed to those of the LRMS/PAN image. Fang *et al*. [38] proposed a new variational pansharpening method based on some effective assumptions and discussed the existence of the minimizer of the energy functional. Aly and Sharma [39] minimized an objective function comprised of the sum of squared residual errors to jointly estimating the HRMS images. Chen *et al*. [40] did not make any assumption about the upsampled MS image but assumed that the fused image after downsampling should be close to the original MS image. They formulated their fusion as a convex optimization problem which minimizes a linear combination of a least-squares fitting term and a dynamic gradient sparsity regularizer.

In recent years, inspired by the impressive performance of machine learning methods in the field of computer vision, some learning-based pan-sharpening models are proposed. Ding *et al*. [41] presented a pan-sharpening method that utilizes a Bayesian nonparametric dictionary learning model to give an underlying sparse representation for image reconstruction. Masi *et al*. [42] adapted a simple and effective three-layer architecture originally designed for super-resolution to the pan-sharpening problem. In [43], a deep residual network is exploited to make the full use of the high nonlinearity of the deep learning models. Yang *et al*. [44] proposed a deep network architecture for pansharpening. This trained network, called PanNet, achieves plausible image reconstruction results.

Related researches about multi-sensor image fusion similar to pansharpening have attracted increasing attention of researchers in the remote sensing community. such as hyperspectral (HS) and MS image fusion and infrared and visible image fusion. The former focuses on fusing HS images with corresponding high spacial resolution MS images. Bayesian inference offers a convenient way to tackle the ill-posed fusion problem [45], [46]. The fusion of HS and MS images based on spectral unmixing [47], [48] and Sparse representation [49] has also been explored. The latter aims at fusing thermal infrared (IR) and visible images obtained from the same scene to enhance the performance in terms of human visual perception and target recognition [50]. To tackle this issue, some works based on image decomposition [51] and total variation minimization [52] have been proposed.

In this paper, a high quality pansharpening model is proposed by applying Bayesian theory. Our model is mainly based on three reasonable assumptions which are formulated into three probability terms within Bayesian framework. In each term, we introduce multi-order gradients to capture more fine and complete information from input images. The first term can keep sharp edges of the HRMS image, and the second term protects spectral information from degradation. Moreover, we add the third term to keep the similarity of neighboring pixels. To optimize the entire posterior probability, the alternating direction method of multipliers (ADMM) [53], [54] is applied. The optimum can be obtained with an acceptable time cost. In summary, we make the following contributions:

- We introduce the multi-order gradients into our model which can significantly improve the pansharpening performance. To the best of our knowledge, it is the first work that uses multi-order gradients for pansharpening.
- We design three novel probability terms based on corrsponding reasonable assumptions to jointly characterize the HRMS and the available LRMS and PAN images, and propose an efficient pansharpening model by combine them into Bayesian framework.
- We demonstrate that our approach produces better results than other state-of-the-art techniques. Furthermore, we show our method can be easily extended to hyperspectral image fusion.

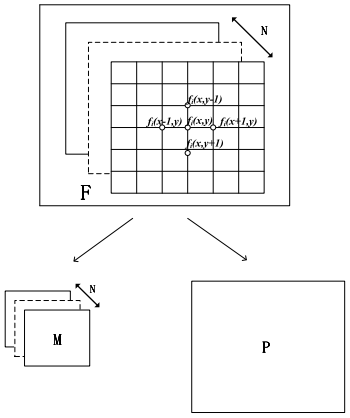


Fig. 2. The relationship network of $\boldsymbol{F}$, $\boldsymbol{M}$ and $\boldsymbol{P}$.

The rest of this paper is organized as follows. In section II, we present the proposed model based on three basic assumptions. In section III, we apply ADMM algorithm to minimize the energy function. The experimental results are shown in section IV. Finally, conclusion will be presented in section V.

II. The Proposed Model

To explain ideas, we assume that a PAN image can be modeled as a function $\boldsymbol{P} : \Omega \to \mathbb{R}$, where Ω is a subset of $\mathbb{R}^2$, and the LRMS image is given by a function $\boldsymbol{M} = \{\boldsymbol{m}_1, \boldsymbol{m}_2, \ldots, \boldsymbol{m}_N\} : \Omega \to \mathbb{R}^N$, where N is the number of bands of the MS image. Then, let $\boldsymbol{F} = \{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_N\} : \Omega \to \mathbb{R}^N$ be the pan-sharpened high resolution MS (HRMS) image.

According to the principle of pansharpening, the relationship among $\boldsymbol{F}$, $\boldsymbol{M}$ and $\boldsymbol{P}$ can be illustrated as a Bayesian network model as shown in Fig. 2, from which we have the posterior probability:

$$g(\boldsymbol{F}|\boldsymbol{M}, \boldsymbol{P}) \propto g(\boldsymbol{P}, \boldsymbol{M}|\boldsymbol{F})g(\boldsymbol{F}), \tag{1}$$

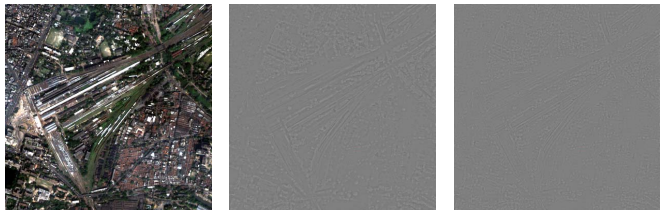
where $g(\boldsymbol{P}, \boldsymbol{M}|\boldsymbol{F})$ represents the likelihood and $g(\boldsymbol{F})$ denotes the prior on the required image $\boldsymbol{F}$. Based on the D-Separation theory [55], P and $\boldsymbol{M}$ are independent events when $\boldsymbol{F}$ is known. Therefore, equation (1) can be rewritten as follows.

$$g(\boldsymbol{F}|\boldsymbol{M}, \boldsymbol{P}) \propto g(\boldsymbol{P}|\boldsymbol{F})g(\boldsymbol{M}|\boldsymbol{F})g(\boldsymbol{F}). \tag{2}$$

In what follows, two kinds of gradient operator, ∇_2 and ∇_3, will be used. Here ∇_2 denotes the general gradient operator towards horizontal and vertical directions, respectively. Different from ∇_2, ∇_3 adds one more direction along the bands besides horizontal and vertical directions. We now define these terms in (2) based on three certain properties.

A. Definition of the Probability Terms

1) Term $g(\boldsymbol{P}|\boldsymbol{F})$: In the same scene, the spatial range of the PAN image almost covers the whole range of that of the MS image [40]. Thus, the spatial information of HRMS image $\boldsymbol{F}$ is contained in the PAN image. It was reasonable to make the assumption that a linear combination of all bands of $\boldsymbol{F}$ should be closed to $\boldsymbol{P}$ in spatial information. And an image's spatial information can be generally expressed by the measure of gradient field [56]. Thus the assumption can be represented as the following equation:



(a) (b) (c)

Fig. 3. (a): the ground truth image. The error image ϵ_1 using (b): the simple assumption (3), and (c): the proposed multi-order gradient assumption (4). (For simplicity, the multi-order gradient is set as $\{\nabla_2, \frac{1}{\sqrt{2}}\nabla_2^2\}$ here.)

$$\nabla_2 \boldsymbol{P} = \sum_{i=1}^{N} \alpha_i \nabla_2 \boldsymbol{f}_i + \epsilon_1, \tag{3}$$

where α_i represents the coefficients related to the MTF function [57] of the sensor, and ϵ_1 is the noise or error which assumed to follow the Gaussian distribution $\mathcal{N}(0, \sigma_1^2)$.

However, as discussed in [26] and [58], this model is weak and failed to capture the spatial randomness property of image noise. To illustrate, in Fig. 3(b), we show the noise map generated by our algorithm using the simple assumption (3). One can see that the noise map is clearly structured and not spatially random. In fact, since ϵ_1 is a random sequence following Gaussian distribution, it has been proven that $\nabla_2\epsilon_1$ also follows Gaussian distribution $\mathcal{N}(0, \sigma^2)$ with standard deviation $\sigma = \sqrt{2}\sigma_1$ [59]. Therefore (3) can be reset as:

$$\nabla_2^* \boldsymbol{P} = \sum_{i=1}^{n} \alpha_i \nabla_2^* \boldsymbol{f}_i + \nabla_2^* \epsilon_1, \tag{4}$$

where the operator $\nabla_2^* = \{\nabla_2, \frac{1}{\sqrt{2}}\nabla_2^2, \ldots, \frac{1}{\sqrt{2}^{n-1}}\nabla_2^n, \ldots\}$ is named as multi-order gradients. Fig. 3(c) shows the noise map generated by our algorithm using the proposed multi-order gradients definition (4). Obviously, Fig. 3(c) contains less image structure compared to Fig. 3(b), which is a significant improvement.

From (4), we build the likelihood as:

$$\begin{aligned} g(\boldsymbol{P}|\boldsymbol{F}) &= g(\nabla_2^*\epsilon_1) = \prod_n \mathcal{N}(\nabla_2^n \epsilon_n \mid 0, \sqrt{2}^{n-1}\sigma_1) \\ &= \prod_n \mathcal{N}(\sum_{i=1}^{N} \alpha_i \nabla_2^n \boldsymbol{f}_i \mid \nabla_2^n \boldsymbol{P}, \sqrt{2}^{n-1}\sigma_1) \\ &\propto \mathcal{N}(\sum_{i=1}^{N} \alpha_i \nabla_2^* \boldsymbol{f}_i \mid \nabla_2^* \boldsymbol{P}, \sigma_1), \end{aligned} \tag{5}$$

where $n = 1, 2, \ldots$. Thus, we have deduced that $\nabla_2^*\epsilon_1$ also follows the Gaussian distribution.

2) Term $g(\boldsymbol{M}|\boldsymbol{F})$: The HRMS image should contain the same spectral information as LRMS image. And for each band, the low resolution pixels are generated by high resolution ones with low-pass filtering followed by down-sampling. Denote the low-pass filter and down-sampling operator as H and D, respectively. The relationship of $\boldsymbol{M}$ and $\boldsymbol{F}$ can be formulated as:

$$\boldsymbol{M} = DH\boldsymbol{F} + \epsilon_2, \tag{6}$$

and we assume that the noise ϵ_2 obeys gaussian distribution $\mathcal{N}(\epsilon_2 \mid 0, \sigma_2^2)$.

Equation (6) only reflects the relationship between LRMS and HRMS images in intensity. Actually, $DH\boldsymbol{F}$ and $\boldsymbol{M}$ should be same in both spectral and spatial information. Therefore, in order to preserve more detail information, equation (6) can be reconstructed as a more reasonable expression:

$$\nabla_3^* \boldsymbol{M} = \nabla_3^*(DH\boldsymbol{F}) + \nabla_3^* \epsilon_2, \tag{7}$$

where operator $\nabla_3^* = \{I, \frac{1}{\sqrt{2}}\nabla_3, \ldots, \frac{1}{\sqrt{2^n}}\nabla_3^n, \ldots\}$ is multi-order gradients, too. Compared with the former equation (6), (7) guarantees the spatial information consistency apart from the spectral information by adding gradient restriction like (4).

According to above analysis, $\nabla_3^n \epsilon_2$ also follows Gaussian distribution. Therefore, the likelihood $g(\boldsymbol{M}|\boldsymbol{F})$ can be written as follows:

$$\begin{aligned} g(\boldsymbol{M}|\boldsymbol{F}) &= \prod_n \mathcal{N}(\nabla_3^n \epsilon_2 \mid 0, \sqrt{2}^n \sigma_2) \\ &= \prod_n \mathcal{N}(\nabla_3^n \boldsymbol{F} \mid \nabla_3^n \boldsymbol{M}, \sqrt{2}^n \sigma_2) \\ &\propto \mathcal{N}(\nabla_3^* \boldsymbol{F} \mid \nabla_3^* \boldsymbol{M}, \sigma_2), \end{aligned} \tag{8}$$

where $n = 0, 1, 2, \ldots$.

3) Term $g(\boldsymbol{F})$: In a normal image, a pixel $\boldsymbol{f}_i(x, y)$ (see Fig. 2) not around the edges should be similar with neighboring pixels in each band of pan-sharpened image. We assume that the similarity will not be transmitted, i.e., memoryless, and the relationship between neighboring pixels can be represented as a Markov random field model. Note that pixels around the edges only take up a very small proportion, this assumption is acceptable in probability theory. As the exponential distribution has a good property of memoryless, we thus model the prior of $\boldsymbol{F}$ as exponential distribution:

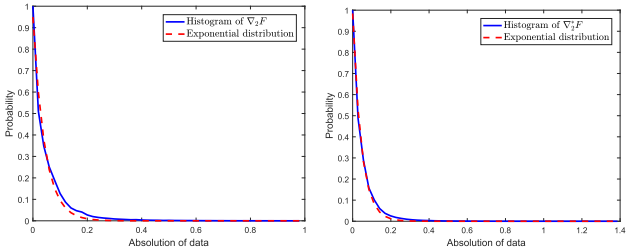
$$g(\boldsymbol{F}) = \prod_{i=1}^{N} e^{-\tau\{\sum[\boldsymbol{f}_i(x+1,y)-\boldsymbol{f}_i(x,y)]+\sum[\boldsymbol{f}_i(x,y+1)-\boldsymbol{f}_i(x,y)]\}}, \tag{9}$$

where τ is the rate parameter. Here we choose two of the four-neighborhood pixels to express the similarity, and the sum of differences between pixels can be simplified into the l_1 norm of first-order gradient:

$$g(\boldsymbol{F}) = e^{-\tau\|\nabla_2 \boldsymbol{F}\|_1}, \tag{10}$$

where $\|\cdot\|_p$ is the l_p-norm.

To verify the reliability of above standpoint, we take some real images (contains approximately 10^8 pixels) from QuickBird and IKONOS satellites, and calculate the histogram of $\nabla_2 \boldsymbol{F}$. The statistical result is shown in Fig. 4(a). One can see that the histogram is well approximated by an exponential distribution (with root-mean square error MSE = 3.6012×10^{-4}). On the other hand, Fig. 4(b) shows the histogram of $\nabla_2^* \boldsymbol{F}$ and the corresponding fitted exponential distribution (with MSE = 7.0299×10^{-5}, much less than that of $\nabla_2 \boldsymbol{F}$). The higher fitting degree indicates the more reasonability of multi-order gradient. Accordingly, we replace the first-order gradient with multi-order one. Thus (10) can be improved as follows:



(a) (b)

Fig. 4. Histograms of $\nabla_2 \boldsymbol{F}$ and $\nabla_2^* \boldsymbol{F}$ with corresponding fitted exponential distributions (for simplicity, $\nabla_2^* \boldsymbol{F}$ is set as $\{\nabla_2, \frac{1}{\sqrt{2}}\nabla_2^2\}$ here). The errors of fitting are: (a) MSE=3.6012×10^{-4}, (b) MSE=7.0299×10^{-5}.

$$g(\boldsymbol{F}) = e^{-\tau\|\nabla_2^* \boldsymbol{F}\|_1}. \tag{11}$$

B. Posterior Probability and Energy Function

In conclusion, by taking all likelihood and prior definitions into (2), the joint probability function can be expressed as:

$$\begin{aligned} g(\boldsymbol{F}|\boldsymbol{M}, \boldsymbol{P}) \propto \mathcal{N}(\sum_{i=1}^{N} \alpha_i \nabla_2^* \boldsymbol{f}_i \mid \nabla_2^* \boldsymbol{P}, \sigma_1) \\ \cdot \mathcal{N}(\nabla_2^* \boldsymbol{F} \mid \nabla_3^* \boldsymbol{M}, \sigma_2) \cdot e^{-\tau\|\nabla_2^* \boldsymbol{F}\|_1} \end{aligned} \tag{12}$$

Our aim is to pursue an optimal $\boldsymbol{F}$ to maximize the probability $g(\boldsymbol{F}|\boldsymbol{M}, \boldsymbol{P})$. By taking negative logarithm to equation (12), i.e., $E(\boldsymbol{F}) = -\log(g(\boldsymbol{F}|\boldsymbol{M}, \boldsymbol{P}))$, we get our energy minimization problem expressed as:

$$\begin{aligned} E(\boldsymbol{F}) = \frac{1}{2}\|\nabla_2^* \boldsymbol{P} - \nabla_2^* \sum_{i=1}^{N} \alpha_i \boldsymbol{f}_i\|^2 \\ + \frac{\beta}{2}\|\nabla_3^* \boldsymbol{M} - \nabla_3^* DH\boldsymbol{F}\|^2 + \gamma \|\nabla_2^* \boldsymbol{F}\|_1, \end{aligned} \tag{13}$$

where $\beta = \sigma_1^2/\sigma_2^2$ and $\gamma = \tau\sigma_1^2$.

III. Numerical Algorithm

In this section, we will solve (13) to find the optimum. First of all, we show some analysis and presetting about the high-order gradient operator.

Remark 1: Due to the fact that high-order gradient ∇^n ($n \geq 3$) contains little spatial information but greatly increase the computational complexity, based on the analysis in [26], we set the multi-order gradients with a maximum number of two in our numerical algorithm as a trade-off. That is,

$$\nabla_2^* = \{\nabla_2, \frac{1}{\sqrt{2}}\nabla_2^2\} \quad \text{and} \quad \nabla_3^* = \{I, \frac{1}{\sqrt{2}}\nabla_3, \frac{1}{\sqrt{2}^2}\nabla_3^2\}.$$

In addtion, large amount of the experiments also show that the simplification satisfies the requirements on both speed and quality of pansharpening.

Accordingly, we have following rather simple property:

Proposition 1: The $\nabla_3^{*T}\nabla_3^*$ and $\nabla_2^{*T}\nabla_2^*$ can be calculated by,

$$\begin{cases} \nabla_3^{*T}\nabla_3^* &= I + \frac{1}{2}\nabla_3^T\nabla_3 + \frac{1}{4}(\nabla_3^T\nabla_3)^2, \\ \nabla_2^{*T}\nabla_2^* &= \nabla_2^T\nabla_2 + \frac{1}{2}(\nabla_2^T\nabla_2)^2. \end{cases}$$

On the other hand, there are some methods which can be used to solve (13), such as split-Bregman method [60] and primal-dual hybid gradient method (PDHG) [61], [62]. In this paper we apply a classical method called ADMM, a powerful algorithm for solving structured convex optimization problems [54].

By replacing $\boldsymbol{f}_i$ as $\boldsymbol{c}_i$, $H\boldsymbol{F}$ as $\boldsymbol{B}_1$, $D\boldsymbol{B}_1$ as $\boldsymbol{B}_2$, and $\nabla_2^*\boldsymbol{F}$ as $\boldsymbol{B}_3$, (13) can be rewritten as:

$$\begin{aligned} E(\boldsymbol{F}) = &\frac{1}{2}\|\nabla_2^*\boldsymbol{P} - \nabla_2^*\sum_{i=1}^{N}\alpha_i\boldsymbol{c}_i\|^2 \\ &+ \frac{\beta}{2}\|\nabla_3^*\boldsymbol{M} - \nabla_3^*\boldsymbol{B}_2\|^2 + \gamma\|\boldsymbol{B}_3\|_1 \\ &\text{subject to} \begin{cases} f_i = c_i \quad (i = 1, 2, \ldots, N) \\ H\boldsymbol{F} = \boldsymbol{B}_1 \\ D\boldsymbol{B}_1 = \boldsymbol{B}_2 \\ \nabla_2^*\boldsymbol{F} = \boldsymbol{B}_3 \end{cases} \end{aligned} \tag{14}$$

The augmented Lagrangian function of (14) is:

$$\begin{aligned} L = &\frac{1}{2}\|\nabla_2^*\boldsymbol{P} - \nabla_2^*\sum_{i=1}^{N}\alpha_i\boldsymbol{c}_i\|^2 \\ &+ \frac{\beta}{2}\|\nabla_3^*\boldsymbol{M} - \nabla_3^*\boldsymbol{B}_2\|^2 + \gamma\|\boldsymbol{B}_3\|_1 \\ &+ \langle\boldsymbol{\Lambda}_1, \boldsymbol{C} - \boldsymbol{F}\rangle + \langle\boldsymbol{\Lambda}_2, \boldsymbol{B}_1 - H\boldsymbol{F}\rangle + \langle\boldsymbol{\Lambda}_3, \boldsymbol{B}_2 - D\boldsymbol{B}_1\rangle \\ &+ \langle\boldsymbol{\Lambda}_4, \boldsymbol{B}_3 - \nabla_2^*\boldsymbol{F}\rangle + \frac{\mu}{2}\|\boldsymbol{C} - \boldsymbol{F}\|_2^2 + \frac{\mu}{2}\|\boldsymbol{B}_1 - H\boldsymbol{F}\|_2^2 \\ &+ \frac{\mu}{2}\|\boldsymbol{B}_2 - D\boldsymbol{B}_1\|_2^2 + \frac{\mu}{2}\|\boldsymbol{B}_3 - \nabla_2^*\boldsymbol{F}\|_2^2, \end{aligned} \tag{15}$$

where μ is a parameter, $\boldsymbol{\Lambda}_i (i = 1, \ldots, 4)$ is Lagrange multiplier and $\boldsymbol{C} = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_4\}$. Our purpose is to minimize $\boldsymbol{F}$, $\boldsymbol{c}_i$, $\boldsymbol{B}_1$, $\boldsymbol{B}_2$, $\boldsymbol{B}_3$ and maximize $\boldsymbol{\Lambda}_i (i = 1, \ldots, 4)$. Although our model looks a little complicated for there are many unknowns, the optimum could be easily obtained after some iterative steps. We will give the iterative solution for each variable in the following.

A. Updating $\boldsymbol{c}_i$

Firstly, we fix $\boldsymbol{F}^t$, $\boldsymbol{\Lambda}_i^t (i = 1, \ldots, 4)$, $\boldsymbol{c}_j^t (j \neq i)$ and optimize $\boldsymbol{c}_i$. Here t is the iteration step. By simply setting the partial derivation $\delta L/\delta \boldsymbol{c}_i$ to zero, we get:

$$\begin{aligned} \alpha_i\nabla_2^{*T}(\nabla_2^*\sum_{j=1}^{N}\alpha_j\boldsymbol{c}_j^t - \nabla_2^*\boldsymbol{P}) \\ + (\boldsymbol{\Lambda}_1)_i^t + \mu(\boldsymbol{c}_i - \boldsymbol{f}_i^t) = 0. \end{aligned}$$

It can be rewritten as:

$$(lhs)\boldsymbol{c}_i = rhs^t, \tag{16}$$

where

$$\begin{aligned} lhs &= \alpha_i^2\nabla_2^{*T}\nabla_2^* + \mu\boldsymbol{1}, \\ rhs^t &= \alpha_i\nabla_2^{*T}\nabla_2^*(\boldsymbol{P} - \sum_{j\neq i}^{N}\alpha_j\boldsymbol{c}_j^t) + \mu\boldsymbol{f}_i^t - (\boldsymbol{\Lambda}_1)_i^t. \end{aligned}$$

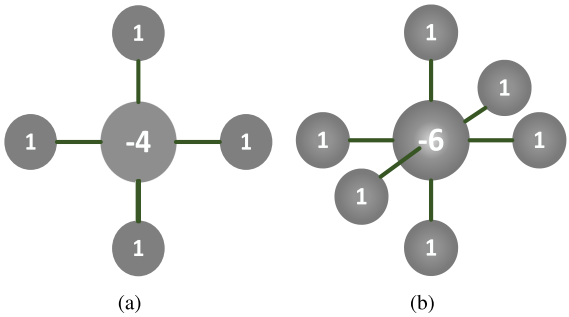


Fig. 5. (a) The 2D laplacian operator. (b) The 3D laplacian operator.

Due to the existence of gradient operators in (16), we operate in the frequency domain to obtain the closed-form solution by using fast Fourier transform (FFT) to make the computation efficient. Denoting the FFT operator and its inverse as $\mathcal{F}$ and $\mathcal{F}^{-1}$, respectively, we get:

$$\boldsymbol{c}_i^{t+1} = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}(rhs^t)}{\mathcal{F}(lhs)}\right\}. \tag{17}$$

The gradient operator $\nabla_2 = [\nabla_{2x}, \nabla_{2y}]$ can be computed by the forward difference between variables representing neighboring pixels, i.e.,

$$\begin{aligned} \nabla_{2x}\boldsymbol{I}(x, y, :) &= \boldsymbol{I}(x+1, y, :) - \boldsymbol{I}(x, y, :), \\ \nabla_{2y}\boldsymbol{I}(x, y, :) &= \boldsymbol{I}(x, y+1, :) - \boldsymbol{I}(x, y, :) \end{aligned}$$

where $\boldsymbol{I}$ is a single- or multi-band image. Besides,

$$\begin{aligned} \nabla_2^T\boldsymbol{I}(x, y, :) = &(\nabla_{2x}\boldsymbol{I}(x, y, :) - \nabla_{2x}\boldsymbol{I}(x-1, y, :)) \\ &+ \big(\nabla_{2y}\boldsymbol{I}(x, y, :) - \nabla_{2y}\boldsymbol{I}(x, y-1, :)\big). \end{aligned}$$

As is well-known, the operator $\nabla_2^T\nabla_2$ can be expressed as a negative Laplacian operator, i.e., $\nabla_2^T\nabla_2 = -\Delta_2$ where Δ_2 is 2D Laplacian operator. And the 2D Laplacian operator is shown in Fig. 5(a).

B. Updating $\boldsymbol{F}$

Secondly, by fixing $\boldsymbol{C}^{t+1}$, $\boldsymbol{B}_i^t (i = 1, 2, 3)$, $\boldsymbol{\Lambda}_i^t (i = 1, \ldots, 4)$, we have:

$$\begin{aligned} -\boldsymbol{\Lambda}_1^t - H^T\Lambda_2^t - \nabla_2^{*T}\Lambda_4^t - \mu(\boldsymbol{C}^{t+1} - \boldsymbol{F}) - \mu H^T \\ \times(\boldsymbol{B}_1^t - H\boldsymbol{F}) - \mu\nabla_2^{*T}(\boldsymbol{B}_3^t - \nabla_2^*\boldsymbol{F}) = 0, \end{aligned} \tag{18}$$

where H is a Gaussian low-pass filter. Accordingly, we exploit the FFT to solve (18) and get $\boldsymbol{F}^{t+1}$, as shown in (19) at the bottom of this page, where $\circ$ represents the element-wise multiplication operator.

C. Updating $\boldsymbol{B}_1$

Thirdly, we optimize $\boldsymbol{B}_1$ by fixing $\boldsymbol{C}^{t+1}$, $\boldsymbol{F}^{t+1}$, $\boldsymbol{B}_i^t (i = 2, 3)$, $\boldsymbol{\Lambda}_i^t (i = 1, \ldots, 4)$. Let the derivative of L w.r.t. $\boldsymbol{B}_1$ be zero, we get:

$$(\mu\boldsymbol{1} + \mu D^T D)\boldsymbol{B}_1 = D^T(\boldsymbol{\Lambda}_3^t + \mu\boldsymbol{B}_2^t) + \mu H\boldsymbol{F}^{t+1} - \boldsymbol{\Lambda}_2^t. \tag{20}$$

$$\boldsymbol{F}^{t+1} = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}[\boldsymbol{\Lambda}_1^t + H^T(\boldsymbol{\Lambda}_2^t + \mu\boldsymbol{B}_1^t) + \nabla_2^{*T}(\boldsymbol{\Lambda}_4^t + \mu\boldsymbol{B}_3^t) + \mu\boldsymbol{C}^{t+1}]}{\mu\boldsymbol{1} + \overline{\mathcal{F}(H)} \circ \mathcal{F}(H) + \mathcal{F}(\nabla_2^{*T}\nabla_2^*)}\right\}, \tag{19}$$

Since the operator $D^T D$ is a diagonal matrix, $(\mu \mathbf{1} + \mu D^T D)^{-1}$ can be easily computed. Then, the closed-form solution of $\boldsymbol{B}_1$ is,

$$\boldsymbol{B}_1^{t+1} = \left(\mu \mathbf{1} + \mu D^T D\right)^{-1} \left(D^T (\boldsymbol{\Lambda}_3^t + \mu \boldsymbol{B}_2^t) + \mu H \boldsymbol{F}^{t+1} - \boldsymbol{\Lambda}_2^t\right). \tag{21}$$

D. Updating $\boldsymbol{B}_2$

Then with $\boldsymbol{C}^{t+1}$, $\boldsymbol{F}^{t+1}$, $\boldsymbol{B}_1^{t+1}$, $\boldsymbol{B}_3^t$, $\boldsymbol{\Lambda}_i^t (i = 1, \ldots, 4)$ fixed, we optimize $\boldsymbol{B}_2$:

$$-\beta \nabla_3^{*T} \nabla_3^* (\boldsymbol{M} - \boldsymbol{B}_2) + \boldsymbol{\Lambda}_3^t + \mu (\boldsymbol{B}_2 - D \boldsymbol{B}_1^{t+1}) = 0. \tag{22}$$

By using FFT, the closed-form solution of $\boldsymbol{B}_2$ can be easily obtained as follows:

$$\boldsymbol{B}_2^{t+1} = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}[\beta \nabla_3^{*T} \nabla_3^* \boldsymbol{M} - \boldsymbol{\Lambda}_3 \mathbf{1} + \mu D \boldsymbol{B}_1^{t+1}]}{\mu \mathbf{1} + \beta \mathcal{F}(\nabla_3^{*T} \nabla_3^*)} \right\}. \tag{23}$$

For an multi-band image $\boldsymbol{I}$, we have $\nabla_3 \boldsymbol{I} = \{\nabla_{3x} \boldsymbol{I}, \nabla_{3y} \boldsymbol{I}, \nabla_{3z} \boldsymbol{I}\}$ and the specific computations are similar to the 2-D case. Besides, the $\nabla_3^T \nabla_3$ operator is a standard negative Laplacian operator with three-dimension which is shown in Fig. 5(b).

E. Updating $\boldsymbol{B}_3$

Different from the above-mentioned variables, $\boldsymbol{B}_3$ has a straightforward solution by shrinkage operator. By fixing the values of other variables, we have:

$$\begin{aligned} \min_{\boldsymbol{B}_3} \gamma \|\boldsymbol{B}_3\|_1 + &< \boldsymbol{\Lambda}_4^t, \boldsymbol{B}_3 - \nabla_2^* \boldsymbol{F}^{t+1} > + \frac{\mu}{2} \|\boldsymbol{B}_3 - \nabla_2^* \boldsymbol{F}^{t+1}\|_2^2 \\ &= \min_{\boldsymbol{B}_3} \frac{\gamma}{\mu} \|\boldsymbol{B}_3\|_1 + \frac{1}{2} \|\boldsymbol{B}_3 - (\nabla_2^* \boldsymbol{F}^{t+1} - \frac{1}{\mu} \boldsymbol{\Lambda}_4^t)\|_2^2 \end{aligned} \tag{24}$$

The solution of subproblem (24) can be given directly by the following soft-thresholding formula:

$$\boldsymbol{B}_3^{t+1} = \text{shrink}(\nabla_2^* \boldsymbol{F}^{t+1} - \frac{1}{\mu} \boldsymbol{\Lambda}_4^t, \frac{\gamma}{\mu}), \tag{25}$$

where

$$\text{shrink}(x, \zeta) = \frac{x}{|x|} \cdot \max(|x| - \zeta, 0), \tag{26}$$

F. Updating $\boldsymbol{\Lambda}_i (i = 1, \ldots, 4)$

The maximum of $\boldsymbol{\Lambda}_i (i = 1, \ldots, 4)$ can be easily obtained as described below:

$$\boldsymbol{\Lambda}_1^{t+1} = \boldsymbol{\Lambda}_1^t + \mu (\boldsymbol{C}^{t+1} - \boldsymbol{F}^{t+1}), \tag{27}$$

$$\boldsymbol{\Lambda}_2^{t+1} = \boldsymbol{\Lambda}_2^t + \mu (\boldsymbol{B}_1^{t+1} - H \boldsymbol{F}^{t+1}) \tag{28}$$

$$\boldsymbol{\Lambda}_3^{t+1} = \boldsymbol{\Lambda}_3^t + \mu (\boldsymbol{B}_2^{t+1} - D \boldsymbol{B}_1^{t+1}) \tag{29}$$

$$\boldsymbol{\Lambda}_4^{t+1} = \boldsymbol{\Lambda}_4^t + \mu (\boldsymbol{B}_3^{t+1} - \nabla_2^* \boldsymbol{F}^{t+1}) \tag{30}$$

G. Summary

Overall, taking all above analyses into account, we can summarize the complete numerical procedure for the proposed method. The detailed description is shown in Algorithm 1.

Algorithm 1 The Overall Procedure for the Proposed Model

Input: PAN image $\boldsymbol{P}$, LRMS image $\boldsymbol{M}$.
Initialize:
set $\boldsymbol{F}^0 = \boldsymbol{C}^0 = \boldsymbol{B}_1^0 = \mathbf{0}$, $\boldsymbol{B}_2^0 = \mathbf{0}$, $\boldsymbol{B}_3^0 = \mathbf{0}$
$\boldsymbol{\Lambda}_1^0 = \boldsymbol{\Lambda}_2^0 = \mathbf{1}$, $\boldsymbol{\Lambda}_3^0 = \mathbf{1}$, $\boldsymbol{\Lambda}_4^0 = \mathbf{1}$;
Repeat:

$$\boldsymbol{c}_i^{t+1} = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}[\alpha_i \nabla_2^{*T} \nabla_2^* (\boldsymbol{P} - \sum_{j \neq i}^N \alpha_j \boldsymbol{c}_j^t) + \mu \boldsymbol{f}_i^t - (\boldsymbol{\Lambda}_1)_i^t]}{\mu \mathbf{1} + \alpha_i^2 \mathcal{F}(\nabla_2^{*T} \nabla_2^*)} \right\},$$

$$\boldsymbol{F}^{t+1} = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}[\boldsymbol{\Lambda}_1^t + H^T (\boldsymbol{\Lambda}_2^t + \mu \boldsymbol{B}_1^t) + \nabla_2^{*T} (\boldsymbol{\Lambda}_4^t + \mu \boldsymbol{B}_3^t) + \mu \boldsymbol{C}^{t+1}]}{\mu \mathbf{1} + \overline{\mathcal{F}(H)} \circ \mathcal{F}(H) + \mathcal{F}(\nabla_2^{*T} \nabla_2^*)} \right\},$$

$$\boldsymbol{B}_1^{t+1} = \left(\mu \mathbf{1} + \mu D^T D\right)^{-1} \left(D^T (\boldsymbol{\Lambda}_3^t + \mu \boldsymbol{B}_2^t) + \mu H \boldsymbol{F}^{t+1} - \boldsymbol{\Lambda}_2^t\right),$$

$$\boldsymbol{B}_2^{t+1} = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}[\beta \nabla_3^{*T} \nabla_3^* \boldsymbol{M} - \lambda_3 \mathbf{1} + \mu D \boldsymbol{B}_1^{t+1}]}{\mu \mathbf{1} + \beta \mathcal{F}(\nabla_3^{*T} \nabla_3^*)} \right\},$$

$$\boldsymbol{B}_3^{t+1} = \text{shrink}(\nabla_2^* \boldsymbol{F}^{t+1} - \tfrac{1}{\mu} \boldsymbol{\Lambda}_4^t, \tfrac{\gamma}{\mu}),$$

$$\boldsymbol{\Lambda}_1^{t+1} = \boldsymbol{\Lambda}_1^t + \mu (\boldsymbol{C}^{t+1} - \boldsymbol{F}^{t+1}),$$

$$\boldsymbol{\Lambda}_2^{t+1} = \boldsymbol{\Lambda}_2^t + \mu (\boldsymbol{B}_1^{t+1} - H \boldsymbol{F}^{t+1}),$$

$$\boldsymbol{\Lambda}_3^{t+1} = \boldsymbol{\Lambda}_3^t + \mu (\boldsymbol{B}_2^{t+1} - D \boldsymbol{B}_1^{t+1}),$$

$$\boldsymbol{\Lambda}_4^{t+1} = \boldsymbol{\Lambda}_4^t + \mu (\boldsymbol{B}_3^{t+1} - \nabla_2^* \boldsymbol{F}^{t+1}),$$

update iteration: $t = t + 1$,
Until $\frac{\|\boldsymbol{F}^{t+1} - \boldsymbol{F}^t\|}{\|\boldsymbol{F}^t\|} < \varsigma$.
Output: HRMS image $\boldsymbol{F}$.

IV. Experiment Results and Analysis

In this section, some experiment results will be presented to illustrate the effectiveness of the proposed method. First, we compare our method with previous works in visual and quantitative aspects at reduced and full resolution, respectively. Then we test the sensitivity to change in scale ratio between MS and PAN images. Finally our method is extended to hyper-spectral image pansharpening.

The experiments on MS pansharpening are validated on two datasets (see: http://glcf.umd.edu/) acquired by different sensors: a) Quickbird sensor which provides LRMS image at 2.44-2.88 m resolution and PAN image at 0.61-0.72 m resolution; b) IKONOS sensor, the resolution cell of which is 4 m × 4 m for the LRMS bands and 1 m × 1 m for the PAN band. Note that both the two sensors work in the visible and near-infrared spectrum range with four bands (blue, green, red, and near infrared) in MS images and the scale ratio $r = 4$. All results are shown by using an RGB representation.

We compare the proposed algorithm with some relevant state-of-the-art techniques including A-IHS [63], P+XS [36], VWP [37], MTF-GLP-HPM [57], [64], Indusion [65], SIRF [40], PNN [42] and PanNet [44]. The parameters for each method are tuned individually according to the authors' suggestions and the best set is selected for each method, respectively. When no authors' specifications are provided, the interpolation of the LRMS image for generating the input HRMS image is carried out by using bicubic interpolation. Besides, we train the PNN and PanNet using 10000 PAN/LRMS/HRMS patch pairs of size 64 × 64/16 × 16 from datasets of QuickBird and IKONOS satellites. The parameters of our algorithm are selected as follows: $\mu \in [5, 15]$, $\beta \in [0.5, 2]$. The value of γ depends on the noise amount, and is set to 0.005 in our experiments. The value of α_i depends on the MTF [57] of MS sensor. Besides, the parameter ς in stop criterion is set as 10^{-4}. By the way, we note that all the following experiments are implemented

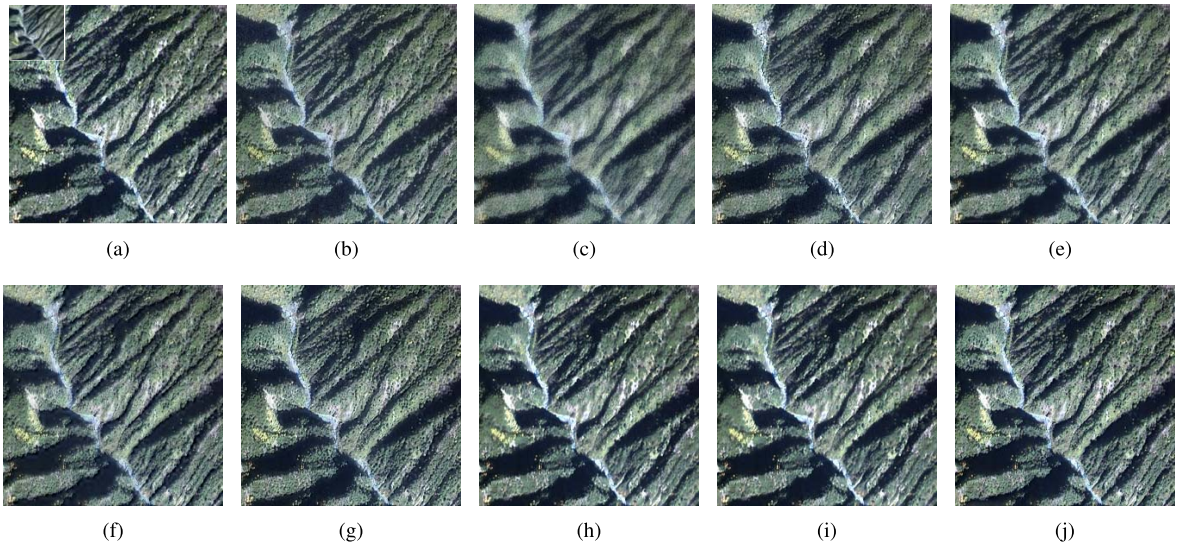


Fig. 6. Pansharpening results comparison(sourse:IKONOS). The LRMS image is on the upper-left corner of the ground-truth. (a) Ground-truth/LRMS. (b) A-IHS. (c) P+XS. (d) VWP. (e) MTF-GLP-HPM. (f) Indusion. (g) SIRF. (h) PNN. (i) PanNet. (j) Ours.

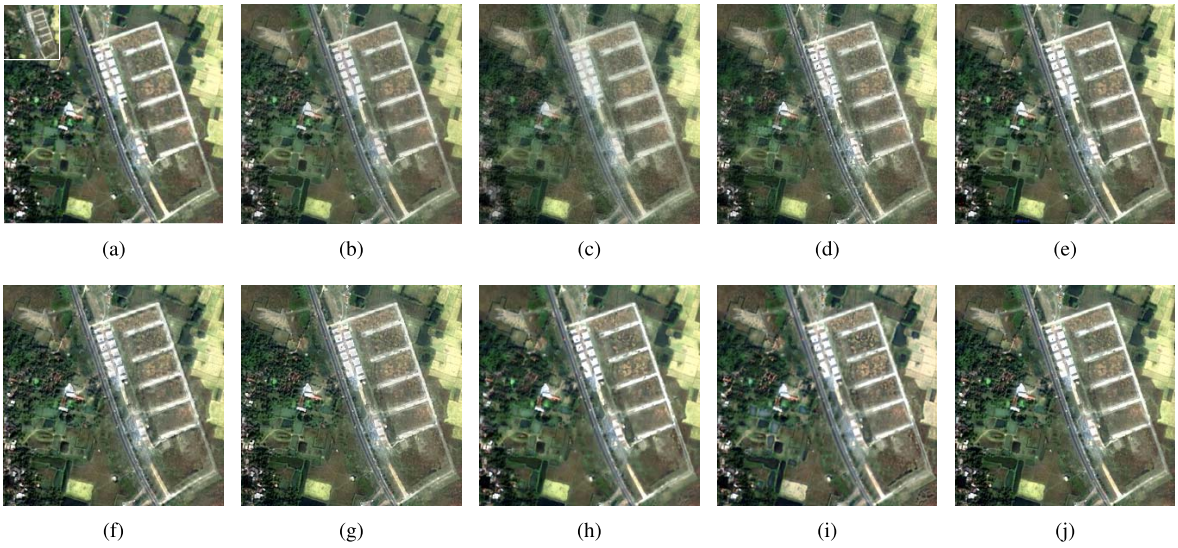


Fig. 7. Pansharpening results comparison(sourse:QuickBird). The LRMS image is on the upper-left corner of the ground-truth. (a) Ground-truth/LRMS. (b) A-IHS. (c) P+XS. (d) VWP. (e) MTF-GLP-HPM. (f) Indusion. (g) SIRF. (h) PNN. (i) PanNet. (j) Ours.

using MATLAB 2015a on a desktop computer with 3.4GHZ Intel core i7-4770 CPU, and it costs about 14s for PAN images with 256×256 pixels and scale ratio $r = 4$.

A. Comparison at Reduced Resolution

In the first part, we do some simulated experiments at reduced resolution based on the Wald's protocol [66]. Firstly, the original MS images are treated as the reference HRMS ones, i.e., the ground truth. Then, the PAN images and corresponding LRMS images at reduced resolution are generated in the same way as illustrated in [8]. More in detail, the degradation is obtained by applying to the available MS and PAN images an LPF and a decimation operator characterized by a sampling factor equal to the resolution ratio r.

In the following, the simulated PAN and LRMS images will be exploited to yield the resulting HRMS images by our method and eight previous works [36], [37], [40], [42], [44], [57], [63]–[65], and the results will be compared with each other in visual and quantitative aspects. Clearly, a better result should be closer to the ground truth.

1) Visual Comparison: We now show some pan-sharpened results as well as the ground truth images in Figs. 6–8. From these result images we can see that each methods can obtain images with better visual effect than LRMS image. The results by P+XS [36], VWP [37] and Indusion are relatively inferior at the first sight as the overall image details are blurred and the intensities of the images have been slightly changed. Lots of artifacts exist in the results of Indusion [65] and SIRF [40], especially near the sharp edges, though they preserves spectral information well. Images produced by A-IHS [63] and MTF-GLP-HPM [57], [64] suffer from slight color distortion although its edges are distinct. The results by deep learning methods [42], [44] and the proposed methods seem much better than others with no spectral distortion or obvious artifacts.

For better visualization, Fig. 9 shows the corresponding error images, the pixel values of which are difference between

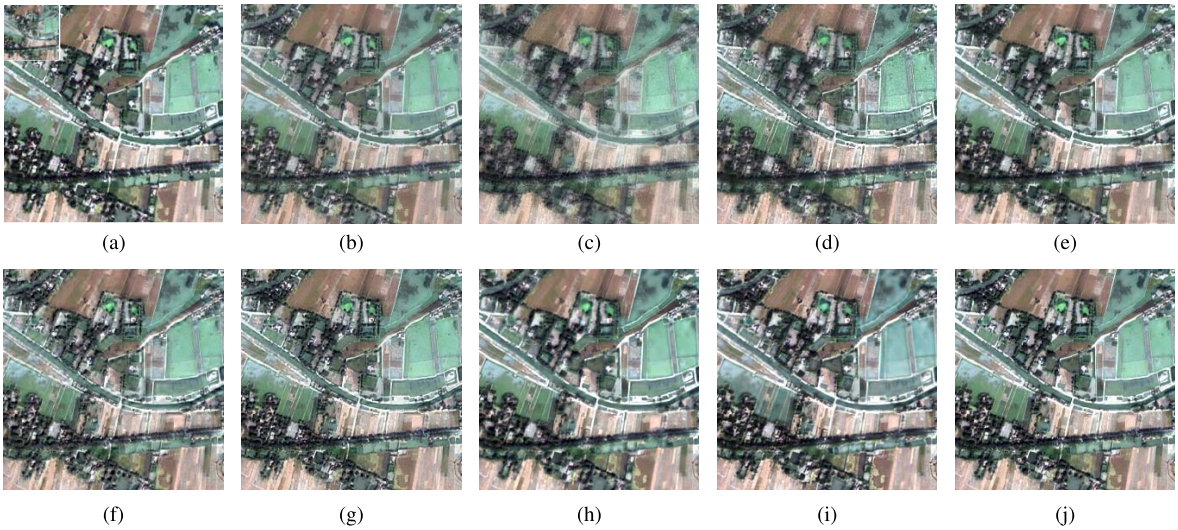


Fig. 8. Pansharpening results comparison(sourse:QuickBird). The LRMS image is on the upper-left corner of the ground-truth. (a) Ground-truth/LRMS. (b) A-IHS. (c) P+XS. (d) VWP. (e) MTF-GLP-HPM. (f) Indusion. (g) SIRF. (h) PNN. (i) PanNet. (j) Ours.

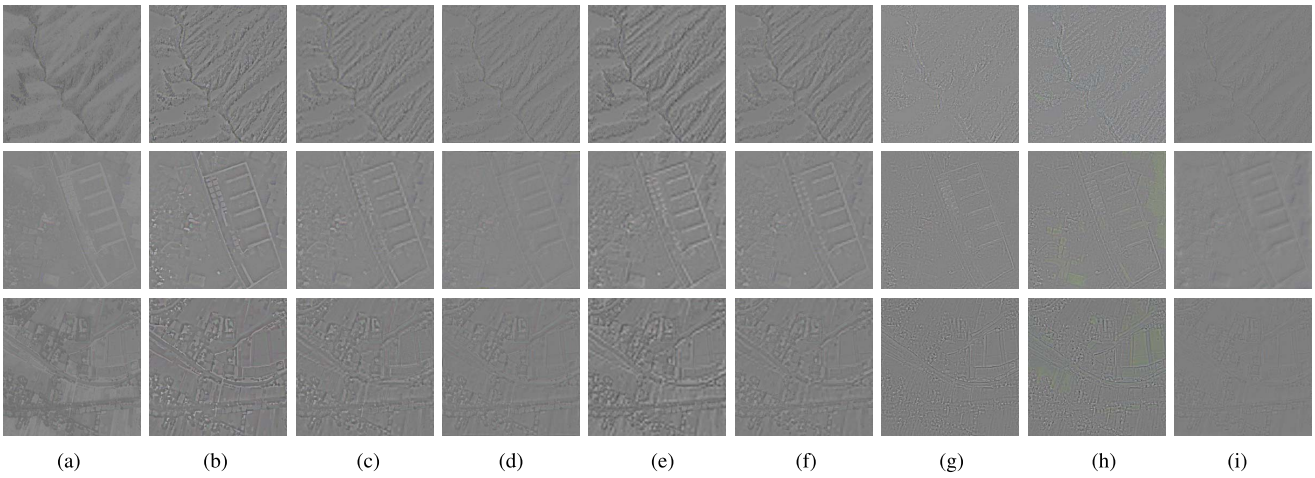


Fig. 9. The corresponding error images of results in Figs. 6-8. (a) A-IHS. (b) P+XS. (c) VWP. (d) MTF-GLP-HPM. (e) Indusion. (f) SIRF. (g) PNN. (h) PanNet. (i) Ours.

the pan-sharpened HRMS images and the corresponding ground-truth and are rescaled. From these error images, any deficiencies can be clearly observed, which is consistent with above analysis. For example, owing to severe spatial and spectral distortion, the performance of P+XS [36] and Indusion [65] is poorer than others. VWP [37] and SIRF [40] have some blurriness near edges. A-IHS [63] suffers from relatively severe spectral distortion, while MTF-GLP-HPM [57], [64] has slight spectral and spatial distortion. The errors of PNN [42] and PanNet [44] are minor and spread over the whole image, while errors of our proposed method are also minor but sparser distributed, which means that our pansharpened result is closer to the ground-truth.

2) Quantitative Analysis: There are many metrics for evaluating the quality of the pan-sharpened product, with respect to an available reference HRMS image. Since each metric has its own advantages and limitations, we choose nine common metrics for comprehensive assessment. Four global indexes are used including the Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS) [67], universal image quality index (Q-index) [68], Root Mean Squared Error (RMSE) and Peak Signal to Noise Ratio (PSNR). Three metrics account for spectral distortion, including Spectral Angle Mapper (SAM) [69], A vector extension of the Q-index (Q4) [70] and Relative Average Spectral Error(RASE) [71]. The Spatial Correlation Coefficient (SCC) [72] and Correlation Coefficient (CC) [67] are introduced for evaluating spatial distortion. We declare that the block size for calculating Q4 is 32 in our experiments.

The quantitative results of aforementioned three contrast tests, i.e., Figs. 6-8, are presented in Tables I-III, respectively. The ideal value for each metric is shown in the last row and the best value for each measure has been highlighted. We can use these data to better analyse previous works and the proposed method. P+XS and Indusion perform worse than others with respect to most measures. Beyond that, A-IHS has a relative poor rank in terms of metrics evaluating spectral distortion, such as SAM and RASE, while VWP is not good as to metrics evaluating spatial distortion. The performance of SIRF and PNN was mediocre. MTF-GLP-HPM, PanNet and ours usually obtain top three results using most measures. Viewed as a whole, our method is better than all previous methods

TABLE I
PERFORMANCE COMPARISON ON THE IMAGES IN FIG. 6

Method	SAM	ERGAS	SCC	Q4	RMSE	RASE	Q-index	CC	PSNR
A-IHS	8.2224	3.7262	0.9334	0.5775	0.0616	15.7976	0.9123	0.9914	24.2137
P+XS	7.5108	4.4625	0.8987	0.4484	0.0732	18.7959	0.8270	0.9357	22.7042
VWP	7.6982	3.5859	0.9265	0.5197	0.0591	15.1700	0.8984	0.9783	24.5658
MTF-GLP-HPM	6.9933	3.0657	**0.9485**	0.5990	**0.0493**	12.6444	0.9269	0.9796	26.1476
Indusion	7.0413	4.2355	0.9013	0.5015	0.0684	17.5499	0.8732	0.9865	23.3001
SIRF	8.8793	3.6377	0.9125	0.5736	0.0569	14.5965	0.9113	0.9839	24.9003
PNN	6.4530	4.1631	0.9209	0.7970	0.0677	16.9286	0.9468	0.8932	23.3846
PanNet	**6.4514**	3.6198	0.9316	**0.8247**	0.0592	14.7954	**0.9631**	0.9099	24.5544
OURS	8.9496	**3.0200**	0.9346	0.6345	**0.0493**	**12.6416**	0.9388	**0.9964**	**26.1495**
Reference	0	0	1	1	0	0	1	1	$+\infty$

TABLE II
PERFORMANCE COMPARISON ON THE IMAGES IN FIG. 7

Method	SAM	ERGAS	SCC	Q4	RMSE	RASE	Q-index	CC	PSNR
A-IHS	6.2663	3.4633	0.8897	**0.8052**	0.0633	14.5445	0.9113	0.9762	23.9707
P+XS	6.5066	4.3208	0.8349	0.6607	0.0768	17.6442	0.8541	0.8708	22.2927
VWP	6.1165	3.7546	0.8616	0.7158	0.0680	15.6230	0.8883	0.9345	23.3494
MTF-GLP-HPM	6.2525	3.7784	0.8525	0.7859	0.0678	15.5779	0.9212	0.8685	23.3775
Indusion	6.3805	4.2945	0.8228	0.7110	0.0760	17.4659	0.8689	0.9711	22.3809
SIRF	6.0379	3.5801	0.8595	0.7564	0.0647	14.8683	0.9027	0.9629	23.7795
PNN	5.3540	3.4983	0.8717	0.7301	0.0584	13.4277	0.9190	0.7284	24.6646
PanNet	**4.8781**	3.2430	0.8778	0.7560	**0.0548**	**12.5844**	0.9244	0.7321	**25.2280**
OURS	5.9717	**3.0520**	**0.8928**	0.8003	0.0562	12.9101	**0.9332**	**0.9788**	25.0061
Reference	0	0	1	1	0	0	1	1	$+\infty$

TABLE III
PERFORMANCE COMPARISON ON THE IMAGES IN FIG. 8

Method	SAM	ERGAS	SCC	Q4	RMSE	RASE	Q-index	CC	PSNR
A-IHS	7.4707	4.6755	0.8714	0.7639	0.0928	18.4376	0.8709	0.9824	20.6370
P+XS	8.1870	5.4322	0.8041	0.6559	0.1082	21.4932	0.7942	0.8657	19.3050
VWP	7.4911	4.7694	0.8279	0.7216	0.0948	18.8146	0.8611	0.9351	20.4611
MTF-GLP-HPM	7.4186	4.2115	0.8686	0.7702	0.0834	16.5522	0.9042	0.9703	21.5739
Indusion	7.3866	5.1863	0.8090	0.6632	0.1035	20.5547	0.8456	0.9767	19.6928
SIRF	7.1396	4.4283	0.8455	0.7355	0.0879	17.4415	0.8893	0.9669	21.1194
PNN	5.8884	4.1492	0.8537	0.7006	0.0835	16.5750	0.8951	0.7003	21.5620
PanNet	**5.2564**	3.9321	0.8576	0.7349	0.0791	15.7003	0.9031	0.6905	22.0329
OURS	6.8081	**3.8258**	**0.8799**	**0.7851**	**0.0757**	**15.0223**	**0.9217**	**0.9846**	**22.4163**
Reference	0	0	1	1	0	0	1	1	$+\infty$

TABLE IV
AVERAGE PERFORMANCE COMPARISON ON LARGE IMAGES

Method	SAM	ERGAS	SCC	Q4	RMSE	RASE	Q-index	CC	PSNR
A-IHS	6.3409	3.0509	0.8760	0.7045	0.0560	11.5676	0.8242	0.9867	24.4729
P+XS	6.4487	3.6317	0.8285	0.6032	0.0719	13.8677	0.8136	0.9061	22.8946
VWP	6.1814	3.1565	0.8460	0.6524	0.0626	12.0019	0.8254	0.9440	24.1200
MTF-GLP-HPM	6.9170	4.9649	0.8719	0.5902	0.1100	13.2456	0.8277	0.9593	24.3734
Indusion	**6.1784**	3.4189	0.8260	0.6405	0.0686	13.1847	0.8366	0.9830	23.3185
SIRF	6.4222	3.0255	0.8549	0.6855	0.0591	11.3423	0.8309	0.9761	24.6107
PNN	6.5355	3.3908	0.8598	0.6775	0.0597	12.5053	0.8752	0.8847	24.5360
PanNet	6.5104	3.2192	0.8722	0.6922	0.0557	11.6622	**0.8883**	0.8886	25.1271
OURS	6.2559	**2.6372**	**0.8807**	**0.7321**	**0.0518**	**9.9063**	0.8371	**0.9883**	**25.8033**
Reference	0	0	1	1	0	0	1	1	$+\infty$

in terms of most metrics, which is enough to demonstrate the effectiveness of our methods.

3) Large Scale Images: In order to verify the validity of the above analysis, we do experiments on some large-scale images with the number of pixels from 6.4×10^6 to 1.6×10^7 for original MS images. These images, containing vegetation (e.g., forest, crops), bodies of water (e.g., lake, river) and urban area (e.g., house, road), are typically partitioned into tiles for efficient processing. Here we give the average metric results of these images which are listed in Table IV. Clearly, our method is consistently better than all previous methods. Once again, these results demonstrate the success of our method.

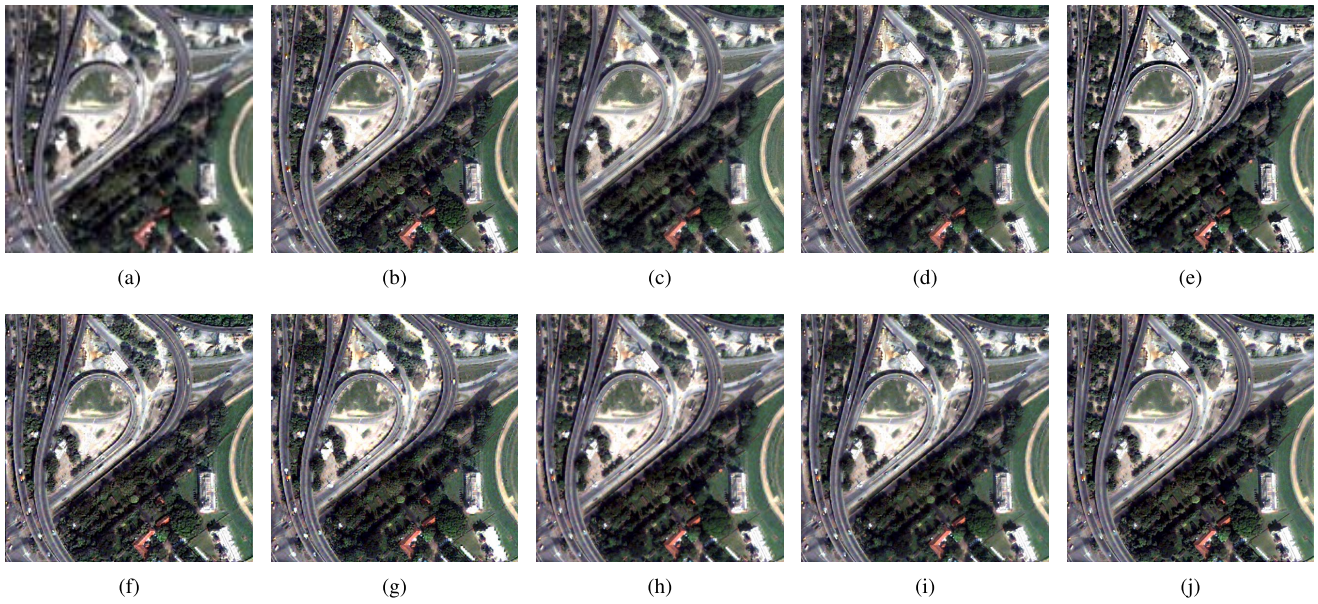


Fig. 10. Pansharpening results comparison (sourse:Quickbird) on the real-world images. (a) MS. (b) A-IHS. (c) P+XS. (d) VWP. (e) MTF-GLP-HPM. (f) Indusion. (g) SIRF. (h) PNN. (i) PanNet. (j) Ours.

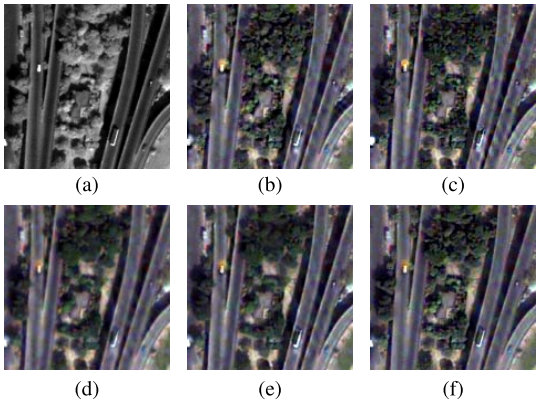


Fig. 11. The zoomed-in areas of PAN image and results of Indusion, SIRF, PNN, PanNet and ours in Fig. 10. (a) PAN. (b) Indusion. (c) SIRF. (d) PNN. (e) PanNet. (f) Ours.

TABLE V
AVERAGE PERFORMANCE ON THE 22 IKONOS IMAGES

Method	D_λ	D_S	QNR
A-IHS	0.1310	0.1754	0.7208
P+XS	0.1208	0.1253	0.7723
VWP	0.1039	0.1031	0.8058
MTF-GLP-HPM	0.1001	0.1141	0.7982
Indusion	0.0952	0.0567	0.8543
SIRF	0.0827	0.0828	0.8455
PNN	0.0526	0.0981	0.8557
PanNet	0.0438	0.0765	0.8836
OURS	**0.0414**	**0.0493**	**0.9127**
Reference	0	0	1

B. Comparison at Full Resolution

In the second part, we evaluate different pansharpening algorithms at full resolution which are also acquired by QuickBird and IKONOS satellites. The PAN and LRMS images are at their capture resolution and we do not add any artificial transformation on them.

1) Visual Comparison: The results are shown in Fig. 10. As is shown, many results obtained in the analysis performed at reduced resolution are in line with those obtained at full resolution. The results by VWP and P+XS are so blurry that many details are missing. As for images yield by Indusion and A-IHS, we can see that color in the vegetation area tends to change, which is a sign of color distortion. Although SIRF and MTF-GLP-HPM preserve details well, they have many artifacts near edges. PNN suffers from slight blurriness. By comparison, PanNet and our method perform better, with no obvious artifacts or color distortion.

A closer look at the fused images of Indusion, SIRF, PNN, PanNet and ours (top five methods in the following quantitative evaluation) are shown in Fig. 11, which presents the zoomed-in areas. One can clearly see that there are color artifacts near the road edges in all images. Artifacts of our proposed method are relatively slighter.

2) Quantitative Analysis: In order to perform quantitative assessment at full resolution, the Quality with No Reference (QNR) index [73] was proposed. The QNR index is defined as

$$QNR = (1 - D_\lambda)^\alpha (1 - D_S)^\beta \tag{31}$$

The two separate values D_λ and D_S quantify the spectral and the spatial distortion, respectively. We consider 22 images from IKONOS satellite and 16 images from QuickBird satellite, all with size 512×512 for PAN images. Tables V and VI report the average values of the D_λ, D_S and QNR. Again we find that the proposed method outperforms the other methods.

C. Sensitivity to Change in Scale Ratio r

All experiments we mentioned above are validated on PAN images and corresponding LRMS images with the scale ratio $r = 4$. What will the effects be if the scale ratio changes? In order to verify the reliability of our method, we do one more comparison experiment. With different scale ratio r, we compute PSNR of each method to measure their performance.

TABLE VI
AVERAGE PERFORMANCE ON THE 16 QUICKBIRD IMAGES

Method	D_λ	D_S	QNR
A-IHS	0.1254	0.1045	0.7831
P+XS	0.1256	0.0902	0.8044
VWP	0.1165	0.0679	0.8428
MTF-GLP-HPM	0.0962	0.0857	0.8153
Indusion	0.1084	0.0504	0.8752
SIRF	0.0785	**0.0441**	0.8926
PNN	0.0470	0.0503	0.9051
PanNet	0.0479	0.0556	0.8993
OURS	**0.0349**	0.0498	**0.9171**
Reference	0	0	1

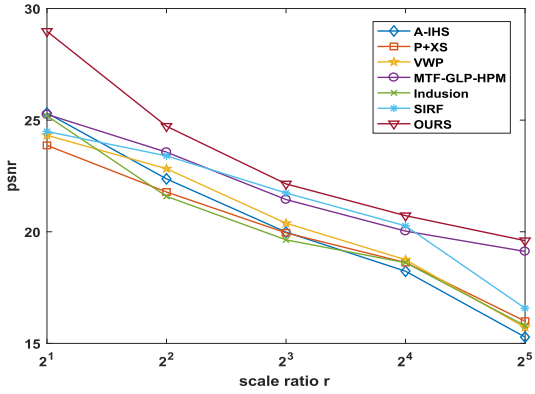


Fig. 12. PSNR of results with different scale ratio.

The result is shown in Fig. 12, in which the deep learning based model PanNet [44] and PNN [42] are not involved for it is time-consuming to retrain the two models with different scale ratio. We can obviously see that although PSNR decreases with the increasing of scale ratio for all methods, the proposed method always outperforms the others.

D. Extension to Hyper-Spectral (HS) Images

The proposed method can be easily extended to HS pansharpening without modifications, in which the MS image is replaced by the HS image. The number of bands and the scale ratio are dependent on the input data. Some experimental tests were conducted at reduced resolution with Moffett field datasets according to the Wald's protocol [66]. The dimensions of the PAN are 185×395 with a spatial resolution of 20m whereas the size of the HS image is 37×75 with a spatial resolution of 100m, which means the scale ratio is 5. The HS image contains 224 bands covering the spectral range 0.4-2.5 nm. Note that the running time is about 1000s. We compared our method with GS [17] and GFPCA [25], two methods which originally designed for MS pansharpening but can be also extended to HS pansharpening, and CNMF [47], method designed for HS pansharpening. And we select four metrics to evaluate these methods. The result is shown in Table VII. With no doubt, CNMF [47] performs best with respect to all metrics for its pertinence. Our performance is closed to the best one, and among those methods which are applicable for both MS and HS pansharpening, the proposed method is superior to others with slight advantage.

TABLE VII
PERFORMANCE COMPARISON ON THE MOFFETT FIELD DATASETS

Method	CC	SAM	RMSE	ERGAS
GS	0.9172	12.9589	420.5469	7.2204
GFPCA	0.9161	11.3363	404.0979	7.0619
CNMF	**0.9558**	**9.0831**	**311.2472**	**5.3755**
OURS	0.9174	9.3825	395.6495	6.9321
Reference	1	0	0	0

V. CONCLUSION

In this paper, a novel and powerful pansharpening algorithm is proposed based on Bayesian theory. The model turns three reasonable assumptions into a combination of three probability terms, the negative log function of which is a minimization problem. Then ADMM is used to find the optimum of the problem, with sufficient implementation details. To evaluate the effectiveness of our method, several kinds of experiments were conducted at full and reduced resolusion. The results show that the proposed method outperforms other state-of-the-art fusion methods in terms of both spectral and spatial qualities, with higher reliability. Further experiment shows that the proposed method can be easily extended to HS pansharpening.

Our method may benefit some applications in remote sensing because of the high accuracy and high reliability property. Of course, our method is not perfect. For example, three parameters need to be tuned manually in our algorithm. In the future, we will focus on auto-tuning parameters and try to improve Bayesian model to pansharpening task.

REFERENCES

[1] L. Wald, "Some terms of reference in data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1190–1193, May 1999.
[2] H. Aanæs, J. R. Sveinsson, A. A. Nielsen, T. Bovith, and J. A. Benediktsson, "Model-based satellite image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1336–1346, May 2008.
[3] C. Souza, Jr., L. Firestone, L. M. Silva, and D. Roberts, "Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models," *Remote Sens. Environ.*, vol. 87, no. 4, pp. 494–506, Nov. 2003.
[4] A. Mohammadzadeh, A. Tavakoli, and M. J. V. Zoej, "Road extraction based on fuzzy logic and mathematical morphology from pan-sharpened IKONOS images," *Photogramm. Rec.*, vol. 21, no. 113, pp. 44–60, Mar. 2006.
[5] Y. Zhang, "Pan-sharpening for improved information extraction," in *Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences*. London, U.K.: Taylor & Fransis, 2008, pp. 185–202.
[6] L. Loncan *et al.*, "Hyperspectral pansharpening: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.
[7] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, "Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1301–1312, May 2008.
[8] G. Vivone *et al.*, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.
[9] W. J. Carper, T. M. Lillesand, and R. W. Kiefer, "The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data," *Photogramm. Eng. Remote Sens.*, vol. 56, no. 4, pp. 459–467, Apr. 1990.
[10] R. Haydn, G. W. Dalke, J. Henkel, and J. E. Bare, "Application of the IHS color transform to the processing of multisensor data and image enhancement," *Proc. Nat. Acad. Sci. USA*, vol. 79, no. 13, pp. 571–577, 1982.

[11] 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.*, vol. 1, no. 4, pp. 309–312, Oct. 2004.

[12] M. Ehlers, "Multisensor image fusion techniques in remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 46, no. 1, pp. 19–30, 1991.

[13] V. K. Shettigara, "A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set," *Photogram. Eng. Remote Sens.*, vol. 58, no. 5, pp. 561–567, 1992.

[14] P. S. Chavez, Jr., S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic," *AAPG Bull.*, vol. 57, no. 3, pp. 265–303, 1991.

[15] Q. Du, N. H. Younan, R. King, and V. P. Shah, "On the performance evaluation of pan-sharpening techniques," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 518–522, Oct. 2007.

[16] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly correlated images. II. Channel ratio and 'chromaticity' transformation techniques," *Remote Sens. Environ.*, vol. 22, no. 3, pp. 343–365, 1987.

[17] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," U.S. Patent 6 011 875, Jan. 4, 2000.

[18] C. Padwick, M. Deskevich, F. Pacifici, and S. Smallwood, "WorldView-2 pan-sharpening," in *Proc. ASPRS Annu. Conf.*, San Diego, CA, USA, 2010. [Online]. Available: http://dergipark.gov.tr/download/article-file/429667

[19] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[20] G. P. Nason and B. W. Silverman, "The stationary wavelet transform and some statistical applications," in *Wavelets and Statistics* (Lecture Notes in Statistics). New York, NY, USA: Springer, 1995, pp. 281–299. [Online]. Available: https://link.springer.com/chapter/10.1007%2F978-1-4612-2544-7_17, doi: 10.1007/978-1-4612-2544-7_17.

[21] M. Shensa, "The discrete wavelet transform: Wedding the a trous and Mallat algorithms," *IEEE Trans. Signal Process.*, vol. 40, no. 10, pp. 2464–2482, Oct. 1992.

[22] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, no. 4, pp. 532–540, Apr. 1983.

[23] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.

[24] J. L. Starck, J. Fadili, and F. Murtagh, "The undecimated wavelet decomposition and its reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 297–309, Feb. 2007.

[25] L. Wenzhi *et al.*, "Processing of multiresolution thermal hyperspectral and digital color data: Outcome of the 2014 IEEE GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2984–2996, Jun. 2015.

[26] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 15–19, 2008.

[27] M. Mignotte, "A label field fusion Bayesian model and its penalized maximum rand estimator for image segmentation," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1610–1624, Jun. 2010.

[28] H. Zhang, Y. Zhang, H. Li, and T. S. Huang, "Generative Bayesian image super resolution with natural image prior," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4054–4067, Sep. 2012.

[29] X. Liu, D. Zhao, J. Zhou, W. Gao, and H. Sun, "Image interpolation via graph-based Bayesian label propagation," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1084–1096, Mar. 2014.

[30] N. Wang, X. Gao, L. Sun, and L. Jie, "Bayesian face sketch synthesis," *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1264–1274, Mar. 2017.

[31] D. Fasbender, J. Radoux, and P. Bogaert, "Bayesian data fusion for adaptable image pansharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1847–1857, Jun. 2008.

[32] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[33] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[34] F. Palsson, J. R. Sveinsson, and M. O. Ulfarsson, "A new pansharpening algorithm based on total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 318–322, Jan. 2014.

[35] X. He, L. Condat, J. M. Bioucas-Dias, J. Chanussot, and J. Xia, "A new pansharpening method based on spatial and spectral sparsity priors," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 4160–4174, Sep. 2014.

[36] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rougé, "A variational model for P+XS image fusion," *Int. J. Comput. Vis.*, vol. 69, no. 1, pp. 43–58, 2006.

[37] M. Möller, T. Wittman, A. L. Bertozzi, and M. Burger, "A variational approach for sharpening high dimensional images," *SIAM J. Imag. Sci.*, vol. 5, no. 1, pp. 150–178, 2012.

[38] F. Fang, F. Li, C. Shen, and G. Zhang, "A variational approach for pan-sharpening," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2822–2834, Jul. 2013.

[39] H. A. Aly and G. Sharma, "A regularized model-based optimization framework for pan-sharpening," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2596–2608, Jun. 2014.

[40] C. Chen, Y. Li, W. Liu, and J. Huang, "Image fusion with local spectral consistency and dynamic gradient sparsity," in *Proc. Comput. Vis. Pattern Recognit.*, 2014, pp. 2760–2765.

[41] X. Ding, Y. Jiang, Y. Huang, and J. Paisley, "Pan-sharpening with a Bayesian nonparametric dictionary learning model," in *Proc. 17th Int. Conf. Artif. Intell. Statist.*, 2014, pp. 176–184.

[42] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, p. 594, Jul. 2016.

[43] Y. Wei, Q. Yuan, H. Shen, and L. Zhang, "Boosting the accuracy of multispectral image pansharpening by learning a deep residual network," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1795–1799, Oct. 2017.

[44] J. Yang, X. Fu, Y. Hu, Y. Huang, X. Ding, and J. Paisley, "PanNet: A deep network architecture for pan-sharpening," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 1753–1761.

[45] R. C. Hardie, M. T. Eismann, and G. L. Wilson, "MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1174–1184, Sep. 2004.

[46] Q. Wei, N. Dobigeon, and J. Y. Tourneret, "Bayesian fusion of multi-band images," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1117–1127, Sep. 2015.

[47] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.

[48] N. Yokoya, J. Chanussot, and A. Iwasaki, "Hyperspectral and multispectral data fusion based on nonlinear unmixing," in *Proc. 4th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens.*, 2012, pp. 1–4.

[49] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J. Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3658–3668, Jul. 2015.

[50] S. G. Kong *et al.*, "Multiscale fusion of visible and thermal IR images for illumination-invariant face recognition," *Int. J. Comput. Vis.*, vol. 71, no. 2, pp. 215–233, 2007.

[51] G. Cui, H. Feng, Z. Xu, Q. Li, and Y. Chen, "Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition," *Opt. Commun.*, vol. 341, pp. 199–209, Apr. 2015.

[52] J. Ma, C. Chen, C. Li, and J. Huang, "Infrared and visible image fusion via gradient transfer and total variation minimization," *Inf. Fusion*, vol. 31, pp. 100–109, Sep. 2016.

[53] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989. [Online]. Available: https://www.iitgn.ac.in/sites/default/files/library_files/2017/23012017.pdf

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

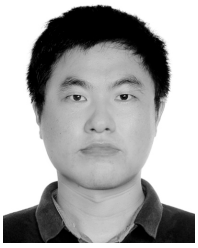
[55] C. R. Hadlock, "Causality: Models, reasoning, and inference," *J. Amer. Statist. Assoc.*, vol. 75, no. 100, pp. 1095–1096, 2005.

[56] G. Piella, "Image fusion for enhanced visualization: A variational approach," *Int. J. Comput. Vis.*, vol. 83, no. 1, pp. 1–11, Jun. 2009.

[57] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and PAN imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.

[58] S. Baronti, B. Aiazzi, M. Selva, A. Garzelli, and L. Alparone, "A theoretical analysis of the effects of aliasing and misregistration on pansharpened imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 446–453, Jun. 2011.
[59] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*. New York, NY, USA: Springer, 2006.
[60] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 323–343, 2009.
[61] M. Zhu and T. F. Chan, "An efficient primal-dual hybrid gradient algorithm for total variation image restoration," Dept. Center Appl. Math., UCLA, Los Angeles, CA, USA, CAM Rep. 08-34, 2008.
[62] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, no. 1, pp. 120–145, 2011.
[63] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An adaptive IHS pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 746–750, Oct. 2010.
[64] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "An MTF-based spectral distortion minimizing model for pan-sharpening of very high resolution multispectral images of urban areas," in *Proc. Workshop Remote Sens. Data Fusion Over Urban Areas*, 2003, pp. 90–94.
[65] M. M. Khan, J. Chanussot, L. Condat, and A. Montanvert, "Indusion: Fusion of multispectral and panchromatic images using the induction scaling technique," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 98–102, Jan. 2008.
[66] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691–699, 1997.
[67] 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.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.
[68] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
[69] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*, 1992, pp. 147–149.
[70] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.
[71] M. Choi, "A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1672–1682, Jun. 2006.
[72] J. Zhou, D. L. Civco, and J. A. Silander, "A wavelet transform method to merge Landsat TM and SPOT panchromatic data," *Int. J. Remote Sens.*, vol. 19, no. 4, pp. 743–757, 1998.
[73] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, "Multispectral and panchromatic data fusion assessment without reference," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200, Feb. 2008.

Tingting Wang received the B.Sc. degree in microelectronics from East China Normal University, Shanghai, China, in 2012, where she is currently pursuing the Ph.D. degree with the Department of Computer Science. Her main research area is image processing using mathematical methods and deep learning methods.



Faming Fang received the Ph.D. degree in computer science from East China Normal University, Shanghai, China, in 2013. He is currently an Associate Professor with the Department of Computer Science, East China Normal University. His main research area is image processing using the variational methods and PDEs.

Fang Li received the M.Sc. degree in mathematics from the South West China Normal University in 2004 and the Ph.D. degree in mathematics from the East China Normal University in 2007. Since 2007, she has been with the Department of Mathematics, East China Normal University, where she is currently a Professor. Her research interests include variational methods, optimization, and deep learning in image processing.

Guixu Zhang received the Ph.D. degree from the Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, China, in 1998. He is currently a Professor with the Department of Computer Science and Technology, East China Normal University, Shanghai, China. His research interests include hyperspectral remote sensing, image processing, and artificial intelligence.