

The fusion of panchromatic and multispectral remote sensing images via tensor-based sparse modeling and hyper-Laplacian prior

Liang-Jian Deng^{a,*}, Minyu Feng^b, Xue-Cheng Tai^c

^a*School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan, 611731, China.*

^b*School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, 611731, China.*

^c*Department of Mathematics, Hong Kong Baptist University, Kowloon, Hong Kong, China*

Abstract

In this paper, we propose a tensor-based non-convex sparse modeling approach for the fusion of panchromatic and multispectral remote sensing images, and this kind of fusion is generally called pansharpening. We first upsample the low spatial-resolution multispectral image by a classical interpolation method to get an initial upsampled multispectral image. Based on the hyper-Laplacian distribution of errors between the upsampled multispectral image and the ground-truth high resolution multispectral image on gradient domain, we formulate a ℓ_p ($0 < p < 1$)-norm term to more reasonably describe the relation of these two datasets. In addition, we also model a tensor-based weighted fidelity term for the panchromatic and low resolution multispectral images, aiming to recover more spatial details. Moreover, total variation regularization is also employed to depict the sparsity of the latent high resolution multispectral image on the gradient domain. For the model solving, we design an alternating direction method of multipliers based algorithm to efficiently solve the proposed model. Furthermore, the involved non-convex ℓ_p subproblem is handled by an efficient generalized shrinkage/thresholding algorithm. Finally, extensive experiments on many datasets collected by different sensors demonstrate the effectiveness of our method when compared with several state-of-the-art image fusion approaches.

Keywords: Pansharpening, Tensor-based sparse modeling, hyper-Laplacian, Alternating direction method of multipliers.

1. Introduction

Pansharpening refers to fuse a *high* spatial-resolution panchromatic (PAN) image and a *low* spatial-resolution multispectral (MS) image to generate a high spatial-resolution MS

*Corresponding Author

Email addresses: liangjian.deng@uestc.edu.cn (Liang-Jian Deng), fmy1987@sina.com (Minyu Feng), xuechengtai@hkbu.edu.hk (Xue-Cheng Tai)

image. In general, the PAN and MS images are simultaneously obtained by several sensors installed on optical devices, *e.g.*, satellites. Due to the physical limitation, the acquired MS image generally only holds low spatial-resolution, therefore we need to enhance the spatial resolution of MS image with the help of PAN image, expecting to get better spatial and spectral details. Pansharpening has important applications in our life. For instance, some commercial products such as Google Earth consider pansharpening as a crucial preliminary step for the subsequent applications, *e.g.*, change detection [1], *etc.*

In 2006, a contest organized by the Data Fusion Committee of IEEE Geoscience and Remote Sensing Society [2, 3] was held, from then on, pansharpening attracts more interests in scientific community [4, 5]. More and more literature from different perspectives and methodologies are published. Most of the pansharpening literature can be divided into three categories, *i.e.*, component substitution (CS) methods, multiresolution analysis (MRA) methods and regularization methods.

The CS techniques rely upon the substitution of a component of the image with the PAN image, after a spectral transformation of the MS data, see *e.g.*, the intensity-hue-saturation [6], the principal component analysis [7], and the Gram-Schmidt (GS) spectral sharpening [8]. The approaches in this category first project the upsampled MS image into a new space, then substitute image components by the high spatial-resolution details of PAN image, and finally execute an inverse projection to yield the high spatial-resolution fused image. In general, CS methods have low computational burden, however, the results are affected by spectral distortions.

The MRA approaches are based on injecting the high spatial-resolution details obtained from the PAN image into the upsampled MS image, see *e.g.*, additive wavelet luminance proportional [9], Laplacian pyramid [10], generalized Laplacian pyramid [11], “à-trous” wavelet transform [12], *etc.* Comparing with CS techniques, MRA approaches mainly suffer from spatial distortions but preserve spectral information well.

Recently, regularization based approaches attract more and more attention of researchers, see *e.g.*, [13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29]. Ballester *et al.* in [13] presented a variational regularization pansharpening method depending on two assumptions: 1) PAN image is the linear combination of high spatial-resolution multispectral bands; 2) PAN image can provide high spatial-resolution information to the multispectral image. However, this approach sometimes suffers from spectral distortion for hyperspectral image pansharpening because of the first unrealistic assumption [4]. In [24], the authors proposed a non-convex sparse regularization model for pansharpening by investigating the distribution of PAN and HRMS images, in which a hyper-Laplacian penalty¹ using $\ell_{1/2}$ norm is employed for spatial details preservation. This non-convex model is efficiently solved by alternating direction method of multipliers (ADMM) approach. Besides, pansharpening based on unmixing is also turned out as a quite effective technique, see [30] and [31]. In [30], Yokoya *et al.* presented an image fusion approach from the perspective of coupled

¹The hyper-Laplacian is defined as $f(x) \propto e^{-k|x|^p}$ where $0 < p < 1$, k is a positive value, and $f(x)$ is the probability function. It is utilized to model the heavy-tailed distribution which has been proven as effective priors for a range of image applications.

nonnegative matrix factorization (CNMF) unmixing. This method can produce high-quality fusion results, both spectrally and spatially. It can be applied to the fusion of multispectral and hyperspectral images. In [22], the authors proposed a pansharpening approach based on local spectral consistency and dynamic gradient sparsity. The given model is then solved by FISTA algorithm [32]. Based on [22], Chen *et al.* in [23] further proposed a unified framework of simultaneous image registration and pansharpening, as well as obtained promising results. Besides, tensor-based regularization methods recently have been applied to many image applications, *e.g.*, video rain removal [33], tensor completion [34, 35, 36], hyperspectral image restoration [37], multispectral image denoising [38], missing data recovery of remote sensing image [39], *etc*, since high-dimensional data can naturally be represented as a tensor which can overcome the drawback of the matrix-based model. In particular, our method in this paper belongs to the category of tensor-based regularization methods. Especially, convolutional neural networks (CNNs) based methods recently have been proposed and show very powerful ability for the pansharpening application, see *e.g.*, [40]. These CNNs models are generally based on the assumption which the relationship between HR/LR multispectral image patches is the same as that between the corresponding HR/LR panchromatic image patches.

In this paper, we propose a hyper-Laplacian sparsity-promoting model for the fusion of panchromatic and multispectral images via tensor modeling. The hyper-Laplacian term depicted by $\ell_p(0 < p < 1)$ norm is motivated by analyzing the error distribution of the upsampled multispectral image and ground-truth HRMS image on the gradient domain. The resulting non-convex ℓ_p subproblem under the framework of alternating direction method of multiplier (ADMM) can be efficiently solved by generalized shrinkage/thresholding (GST) algorithm [41]. To recover spatial details, a linear assumption between panchromatic and latent HRMS images is given by tensor-vector multiplication format. Moreover, we employ total variation (TV) regularization to describe the sparsity of latent HRMS image on the gradient domain. In the experiments, we compare our method with some recent state-of-the-art approaches on four datasets that are from different sensors, such as Pléiades, IKONOS, Quickbird and WorldView-2. Furthermore, some discussion on parameters selection, hyper-Laplacian term, *etc.*, are also presented in the experiments section. Finally, visual and quantitative results demonstrate the effectiveness of our approach.

The *main contributions* of this paper are summarized as follows:

- To the best of our knowledge, this is the first work for pansharpening via tensor-based non-convex modeling, which does not require any training phase and extra data. Comparing with matrix-based methods, tensor-based techniques are more powerful to deal with high-dimensional data, *e.g.*, multispectral images, which can depict the correlations between and within different dimensions better.
- By investigating the error distribution of the upsampled multispectral image and ground-truth HRMS image on the gradient domain, we find that the errors obey hyper-Laplacian distribution which can be depicted by $\ell_p(0 < p < 1)$ -norm. This resulting ℓ_p term is a better description than ℓ_1 or ℓ_2 terms (see Fig. 1).

- The formulated ℓ_p term can reduce the model complexity and skip blur kernel that is generally assumed as inaccurate Gaussian one with an empirical variance. To solve the proposed model, an ADMM based algorithm with high efficiency is designed. Moreover, extensive experimental results demonstrate the effectiveness of our method, both visually and quantitatively.

This paper is organized as follows. The proposed approach is detailedly exhibited in Section 2. Section 3 is devoted to the description of experimental results, followed by some analysis of the obtained results. Some conclusions are drawn in Section 4. Finally, an appendix about GST method will be introduced in Section 6.

2. The proposed model and its solution

2.1. Related notations

Since this work is about tensor-based high dimensional image application, thus it is necessary to introduce some basic tensor notations. Here, following [42], we employ low-case letters for vectors, *e.g.*, \mathbf{g} , upper-case letters for matrices, *e.g.*, \mathbf{G} , and calligraphic letters for tensors, *e.g.*, \mathcal{G} . In addition, an N -mode tensor is defined as $\mathcal{G} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, and $\mathbf{g}_{i_1 i_2 \dots i_N}$ represents its (i_1, i_2, \dots, i_N) -th element. In what follows, the definitions of the **inner product**, **Frobenius norm**, **tensor-vector multiplication** and the **mode- n unfolding** of a tensor are defined as follows:

- The **inner product** of two same-sized tensors \mathcal{G} and \mathcal{H} is defined as $\langle \mathcal{G}, \mathcal{H} \rangle := \sum_{i_1, i_2, \dots, i_N} \mathbf{g}_{i_1 i_2 \dots i_N} \cdot \mathbf{h}_{i_1 i_2 \dots i_N}$.
- The **Frobenius norm** of a tensor is here defined as $\|\mathcal{G}\|_F := \sqrt{\langle \mathcal{G}, \mathcal{G} \rangle}$.
- The **tensor-vector** multiplication is here defined as $\mathcal{G}_{\times n} \mathbf{a} := \mathbf{a}^T \mathbf{G}_{(n)} \in \mathbb{R}^{1 \times \prod_{i \neq n} I_i}$, where $\mathbf{G}_{(n)} \in \mathbb{R}^{I_n \times \prod_{i \neq n} I_i}$ and $\mathbf{a} \in \mathbb{R}^{I_n \times 1}$.
- The **mode- n unfolding** of a tensor \mathcal{G} is defined as $\text{Unfold}_n(\mathcal{G}) := \mathbf{G}_{(n)} \in \mathbb{R}^{I_n \times \prod_{i \neq n} I_i}$, which is composed by taking the mode- n vectors of \mathcal{G} as its columns. Similarly, we can also fold $\mathbf{G}_{(n)}$ into the tensor format by $\text{Fold}_n(\mathbf{G}_{(n)})$.

Before introducing the proposed method, it is also necessary to state data notations along this paper:

- Low spatial-resolution multispectral (LRMS) image (3 mode tensor format): $\mathbf{MS} \in \mathbb{R}^{m \times n \times z}$ with z bands. Upsampled multispectral image: $\mathbf{M} \in \mathbb{R}^{M \times N \times z}$ with z bands, where $M = s \cdot m$, $N = s \cdot n$ with the scale factor s . In this work, we generate \mathbf{M} by applying an interpolation method to \mathbf{MS} (see details from the next subsection).
- Ground-truth high spatial-resolution multispectral image (HRMS) (3 mode tensor format): $\mathbf{U} \in \mathbb{R}^{M \times N \times z}$ with z bands. Latent high spatial-resolution multispectral (HRMS) image: $\mathbf{U}_l \in \mathbb{R}^{M \times N \times z}$ with z bands, which is the one we try to compute.
- Panchromatic (PAN) image: $\mathbf{P} \in \mathbb{R}^{M \times N}$.

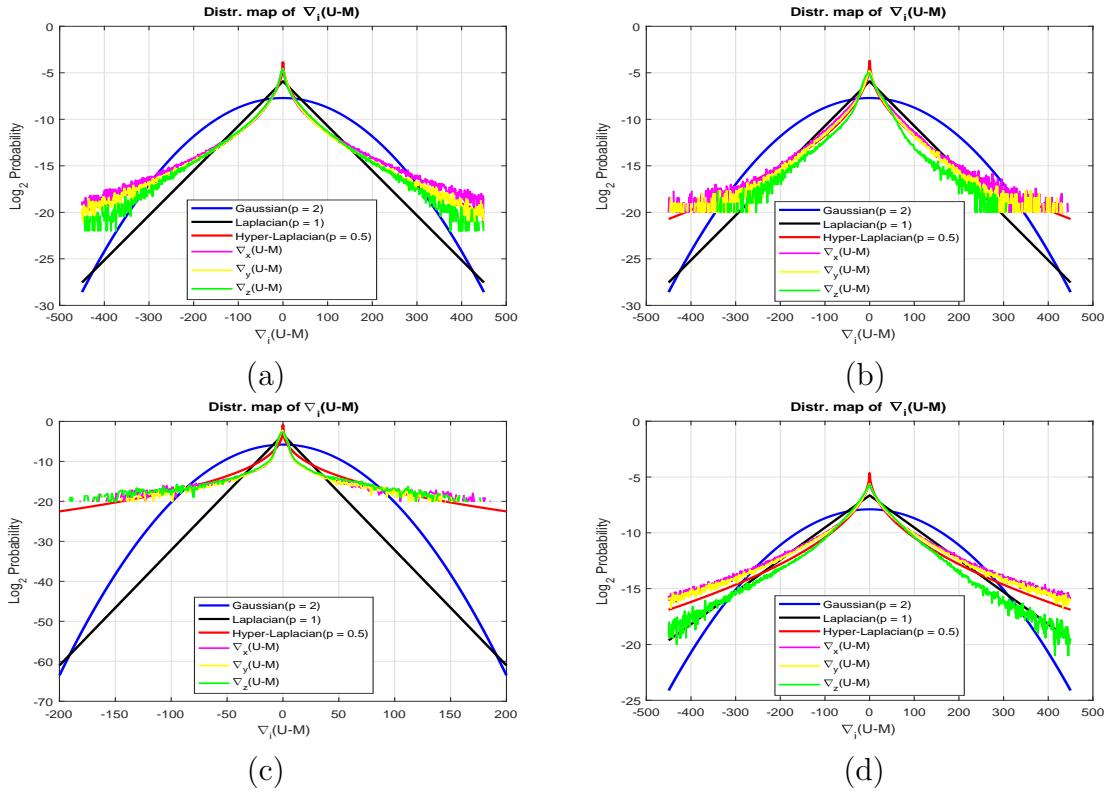


Figure 1: Distributions of $\nabla_i(\mathcal{U} - \mathcal{M})$ ($i = x, y, z$, *i.e.*, 1, 2, 3, respectively in Eq. (1)), which the data respectively comes from Pleiades (a), IKONOS (b), Quickbird (c) and World-View 2 (d) sensors. For each sensor, we choose 4 pairs of LRMS images (All LRMS images are upsampled to \mathcal{M} by a classical interpolation method “interpo23tap”) and ground-truth HRMS images for experiments, which are respectively with the size of $128 \times 128 \times 4$ (or $128 \times 128 \times 8$) and $512 \times 512 \times 4$ (or $512 \times 512 \times 8$).

2.2. The Model

In this work, we propose a non-convex sparse image fusion model and the model is motivated by some extensively statistical investigations between the ground-truth high-resolution multispectral (HRMS) image \mathcal{U} and the upsampled multispectral image \mathcal{M} . Due to the powerful representation ability of tensor, the proposed model is funded by the tensor tool. Comparing with matrix-based image fusion model, the tensor-based model can perfectly represent the high-dimensional data, *e.g.*, multispectral remote sensing images in this work, and depict more latent properties of high-dimensional data better, *e.g.*, the correlation between spectral bands of multispectral remote sensing images.

Next, we will present our tensor-based non-convex model for the fusion of panchromatic and multispectral remote sensing images.

1) **The non-convex ℓ_p term on gradient domain:** For most of image fusion models, they require to know the blur kernel and downsampling matrix in advance. However, the kernel is generally unknown, and people have to empirically assume it as Gaussian distribution with a specific variance, which actually is not fully reasonable. In the work, we try to skip the unknown kernel and investigate the latent distribution between the ground-truth image and the upsampled image. In this subsection, we propose the non-convex term mainly due to the following two reasons. First, we upsample the known LRMS image to get an upsampled multispectral image \mathcal{M} by a classical interpolation method. Here, we choose “interp23tap”² as the interpolation method. This method is quite simple and fast, and has been used in many fusion works, see [6, 11, 9, 43, 5]. Second, we investigate the distributions of $\nabla_i(\mathcal{U} - \mathcal{M})$ on several datasets from different sensors, where ∇_i is the gradient operator, it is clear that the distributions along three dimensions all obey a hyper-Laplacian case (see red lines of Fig. 1), neither Laplacian (black lines of Fig. 1) nor Gaussian (blue lines of Fig. 1) distribution. Due to the two reasons, we can see that it is more reasonable to use an ℓ_p ($0 < p < 1$) norm for the gradient-matching by maximum a posteriori (MAP) rule. Therefore, the first energy term (denoted as $Eng^{(1)}$) is defined as follows,

$$Eng^{(1)} = \sum_{i=1}^3 \alpha_i \|\nabla_i(\mathcal{U}_l - \mathcal{M})\|_p, \quad 0 < p < 1, \quad (1)$$

where $\alpha_i, i = 1, 2, 3$ are positive parameters. In this work, we take $p = 1/2$ since it can excellently depict the distribution between \mathcal{M} and \mathcal{U}_l on the gradient domain (see Fig. 1).

2) **The linear assumption between latent HRMS and panchromatic images:** The panchromatic image \mathbf{P} contains important high spatial-resolution information, thus we may utilize it to increase the spatial details of LRMS image. For the correlation between panchromatic and latent HRMS images, many types of research [44, 17, 24] assume that it obeys a linear relation, which promotes the panchromatic image is equal to the weighted

²“interp23tap” can interpolate one image using a polynomial with 23 coefficients interpolator. Readers can find the corresponding code from the website: <http://openremotesensing.net/knowledgebase/a-critical-comparison-among-pansharpening-algorithms/>

summation of each spectral bands of the latent HRMS image. Therefore, this will lead to the following fidelity term,

$$Eng^{(2)} = \frac{\lambda}{2} \|\mathcal{U}_{l \times 3} \mathbf{w} - \mathbf{p}\|_2^2, \quad (2)$$

where λ is a positive parameter, \mathbf{p} is the vector-form of panchromatic image \mathbf{P} , and $\times 3$ represents tensor-vector multiplication that has been introduced in Section 2.1.

3) The total variation assumption: Total variation (TV) is a very powerful and efficient regularization for image processing, and it has been applied to various of image applications, such as image denoising [45, 46], deblurring [47, 48], segmentation [49], fusion [50, 21], etc. It is a popular and general constraint for images due to the sparsity on gradient domain, thus we also employ it as one regularizer for the latent HRMS image. Besides, by considering that the fused image is the multispectral image but the hyperspectral image, which indicates the spectral curve is not continuous, thus here we only apply TV to the spatial directions but spectral direction, avoiding to influence the spectral relations of original multispectral image. The TV term for the latent HRMS is given as follows,

$$Eng^{(3)} = \sum_{i=1}^2 \omega_i \|\nabla_i \mathcal{U}_l\|_1, \quad (3)$$

which is an anisotropic TV regularizer, and $\omega_i, i = 1, 2$ represents positive parameters for the variations along two spatial directions

The final model: Combining the above three energies together, we finally formulate tensor-based ℓ_p ($p = 1/2$) sparse model for the fusion of remote sensing images, see as follows,

$$\min_{\mathcal{U}_l} \sum_{i=1}^3 \alpha_i \|\nabla_i (\mathcal{U}_l - \mathcal{M})\|_{1/2} + \frac{\lambda}{2} \|\mathcal{U}_{l \times 3} \mathbf{w} - \mathbf{p}\|_2^2 + \sum_{i=1}^2 \omega_i \|\nabla_i \mathcal{U}_l\|_1, \quad (4)$$

which is a non-convex and non-smooth model that involves a $\ell_{1/2}$ term and a ℓ_1 term. In the next subsection, we will present the solving algorithm detailedly for the proposed model.

2.3. The solution

For the solution of the model (4), there are many approaches to solve it, such as alternating direction method of multipliers (ADMM) [51, 52], primal-dual approach [53], etc. Here, we use ADMM which has been applied to many image applications, e.g., image restoration [41], super-resolution [54, 55], to decompose the difficult minimization problem (4) into some simple subproblems. First of all, substituting variables by $\mathcal{T}_i = \nabla_i (\mathcal{U}_l - \mathcal{M})$, $\mathcal{X}_i = \nabla_i \mathcal{U}_l$, $i = 1, 2$, and $\mathcal{V} = \mathcal{U}_l$, the augmented Lagrangian equation $\mathcal{L}(\mathcal{T}_i, \mathcal{V}, \mathcal{X}_i, \mathcal{A}_i, \mathcal{B}_i, \mathcal{C})$ is exhibited as follows

$$\begin{aligned}
\mathcal{L}(\mathcal{T}_i, \mathcal{V}, \mathcal{X}_i, \mathcal{A}_i, \mathcal{B}_i, \mathcal{C}) = & \\
\sum_{i=1}^3 \alpha_i \|\mathcal{T}_i\|_{1/2} + \sum_{i=1}^3 \frac{\eta_i}{2} \|\mathcal{T}_i - \nabla_i(\mathcal{U}_l - \mathcal{M}) + \mathcal{A}_i\|_F^2 & \\
+ \sum_{i=1}^2 \omega_i \|\mathcal{X}_i\|_1 + \sum_{i=1}^2 \frac{\beta_i}{2} \|\mathcal{X}_i - \nabla_i \mathcal{U}_l + \mathcal{B}_i\|_F^2 + \frac{\lambda}{2} \|\mathcal{V}_{\bar{x}3} \mathbf{w} - \mathbf{p}\|_2^2 + \frac{\gamma}{2} \|\mathcal{V} - \mathcal{U}_l + \mathcal{C}\|_F^2. &
\end{aligned} \tag{5}$$

1) The \mathcal{T}_i -subproblem: \mathcal{T}_i has the following minimization problem:

$$\min_{\mathcal{T}_i} \alpha_i \|\mathcal{T}_i\|_{1/2} + \frac{\eta_i}{2} \|\mathcal{T}_i - \nabla_i(\mathcal{U}_l - \mathcal{M}) + \mathcal{A}_i\|_F^2, \tag{6}$$

which results in a non-convex $\ell_{1/2}$ problem that can be solved by some recent excellent algorithms [56, 57, 41, 58]. Here, we choose GST algorithm [41] (see more details from the appendix in Section 6) to solve the non-convex problem because it has a very simple scheme and quite fast speed, which holds the following solution:

$$\mathcal{T}_i^{k+1} = T_{1/2}^{GST} \left(\nabla_i(\mathcal{U}_l^k - \mathcal{M}) - \mathcal{A}_i^k, \frac{\alpha_i}{\eta_i} \right), \quad i = 1, 2, 3. \tag{7}$$

Note that the algorithm summarized in Algorithm 1 needs to compute \mathcal{T}_i , $i = 1, 2, 3$, separately on each image band.

2) The \mathcal{X}_i -subproblem: For \mathcal{X}_i -subproblem, it has the following minimization problem that comes from Eq. (5):

$$\min_{\mathcal{X}_i} \sum_{i=1}^2 \omega_i \|\mathcal{X}_i\|_1 + \sum_{i=1}^2 \frac{\beta_i}{2} \|\mathcal{X}_i - \nabla_i \mathcal{U}_l + \mathcal{B}_i\|_F^2. \tag{8}$$

This convex ℓ_1 problem is quite easy to solve, and it holds a closed-form solution by soft-thresholding strategy [59],

$$\mathcal{X}_i^{k+1} = \text{Shrink} \left(\nabla_i \mathcal{U}_l^k - \mathcal{B}_i^k, \frac{\omega_i}{\beta_i} \right), \tag{9}$$

where $\text{Shrink}(a, b) = \text{sgn}(a) \cdot \max(|a| - b, 0)$, and $\text{sgn}(a)$ is the sign function that $\text{sgn}(a) = -1$ for $a < 0$, $\text{sgn}(a) = 0$ for $a = 0$ and $\text{sgn}(a) = 1$ for $a > 0$.

3) The \mathcal{V} -subproblem: For the \mathcal{V} -subproblem, we need to solve the following minimization problem:

$$\min_{\mathcal{V}} \frac{\lambda}{2} \|\mathcal{V}_{\bar{x}3} \mathbf{w} - \mathbf{p}\|_2^2 + \frac{\gamma}{2} \|\mathcal{V} - \mathcal{U}_l + \mathcal{C}\|_F^2. \tag{10}$$

This tensor based subproblem can be easily transformed into a matrix-vector form according to the rule of tensor-vector multiplication:

$$\min_{\mathbf{V}_{(3)}} \frac{\lambda}{2} \|\mathbf{w}^T \mathbf{V}_{(3)} - \mathbf{p}\|_2^2 + \frac{\gamma}{2} \|\mathbf{V}_{(3)} - \mathbf{U}_{l(3)} + \mathbf{C}_{(3)}\|_F^2, \quad (11)$$

where $\mathbf{V}_{(3)}, \mathbf{U}_{l(3)}, \mathbf{C}_{(3)} \in \mathbb{R}^{z \times MN}$ are the unfold matrices along the mode 3 of \mathbf{V}, \mathbf{U}_l and \mathbf{C} , respectively. $\mathbf{w}^T \in \mathbb{R}^{1 \times z}$ is the weight vector, and $\mathbf{p} \in \mathbb{R}^{1 \times MN}$ represents the vector-form of \mathbf{P} . The minimization problem (11) can be solved by least squares method and have the following closed-form solution,

$$\mathbf{V}_{(3)}^{k+1} = (\lambda \mathbf{w} \mathbf{w}^T + \gamma \mathbf{I})^{-1} (\gamma (\mathbf{U}_{l(3)}^k - \mathbf{C}_{(3)}^k) - \lambda \mathbf{w} \mathbf{p}), \quad (12)$$

where \mathbf{I} is an identity matrix with the size $z \times z$. Note that the computation of $\mathbf{V}_{(3)}^{k+1}$ is quite cheap, since it only involves an inverse computing of a small size matrix³. After getting $\mathbf{V}_{(3)}^{k+1}$, it is easy to fold $\mathbf{V}_{(3)}^{k+1}$ to yield the final tensor format result,

$$\mathbf{V}^{k+1} = \text{Fold}_3 (\mathbf{V}_{(3)}^{k+1}). \quad (13)$$

4) The \mathbf{U}_l -subproblem: According to Eq. (5), \mathbf{U}_l -subproblem is presented as follows

$$\min_{\mathbf{U}_l} \sum_{i=1}^3 \frac{\eta_i}{2} \|\mathcal{T}_i - \nabla_i (\mathbf{U}_l - \mathbf{M}) + \mathbf{A}_i\|_F^2 + \frac{\gamma}{2} \|\mathbf{V} - \mathbf{U}_l + \mathbf{C}\|_F^2 + \sum_{i=1}^2 \frac{\beta_i}{2} \|\mathcal{X}_i - \nabla_i \mathbf{U}_l + \mathbf{B}_i\|_F^2, \quad (14)$$

which can be solved easily by FFT algorithm under the periodic boundary condition,

$$\mathbf{U}_l^{k+1} = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(\mathbf{W})}{\mathcal{F}(\mathbf{N})} \right), \quad (15)$$

where \mathcal{F} and \mathcal{F}^{-1} represent FFT and inverse FFT, respectively, and

$$\mathbf{W} = \sum_{i=1}^3 \eta_i \nabla_i^T (\nabla_i \mathbf{M} + \mathcal{T}_i^{k+1} + \mathbf{A}_i^k) + \gamma (\mathbf{V}^{k+1} + \mathbf{C}^k) + \sum_{i=1}^2 \beta_i \nabla_i^T (\mathcal{X}_i^{k+1} + \mathbf{B}_i^k), \quad (16)$$

and

$$\mathbf{N} = \sum_{i=1}^3 \eta_i \nabla_i^T \nabla_i + \sum_{i=1}^2 \beta_i \nabla_i^T \nabla_i + \gamma \mathcal{O}, \quad (17)$$

where \mathcal{O} is an tensor with the size $M \times N \times z$ whose all elements are 1.

³The matrix size is $z \times z$, where z represents the number of bands of multispectral image, and is generally not large.

5) Update Multipliers: After finishing the computing of each subproblems, we also need to update the Lagrangian multipliers \mathcal{A}_i , \mathcal{B}_i and \mathcal{C}_i on the $k + 1$ -th iteration by:

$$\begin{aligned}\mathcal{A}_i^{k+1} &= \mathcal{A}_i^k + (\mathcal{T}_i^{k+1} - \nabla_i(\mathcal{U}_l^{k+1} - \mathcal{M})) , i = 1, 2, 3 \\ \mathcal{B}_i^{k+1} &= \mathcal{B}_i^k + (\mathcal{X}_i^{k+1} - \nabla_i \mathcal{U}_l^{k+1}) , i = 1, 2 \\ \mathcal{C}^{k+1} &= \mathcal{C}^k + (\mathcal{V}^{k+1} - \mathcal{U}_l^{k+1}) .\end{aligned}\quad (18)$$

The steps **1) - 5)** represent one iteration of ADMM based framework, which decomposes the difficult minimization problem (5) into four simpler subproblems (*i.e.*, \mathcal{T}_i -, \mathcal{X}_i -, \mathcal{V} - and \mathcal{U}_l -subproblems). Note that these four subproblems all have fast and accurate techniques to compute solutions. For instance, the \mathcal{T}_i -subproblem can be solved by fast GST algorithm that is reported in Algorithm 2, and \mathcal{X}_i -subproblem keeps the closed-form solution by the soft-thresholding strategy. Moreover, \mathcal{V} -subproblem only involves a small matrix inverse (see Eq. (12)), thus the closed-form solution can be computed efficiently. Furthermore, \mathcal{U}_l -subproblem can be addressed by the efficient and well-known FFT technique. In particular, the final algorithm for the proposed model (4) is summarized in Algorithm 1.

Algorithm 1: The solution of model (4) by ADMM

Input: Initial image \mathcal{M} , panchromatic image \mathbf{P}

Output: Fused image \mathcal{U}_l

Initialize:

1) $k \leftarrow 0$, $\mathcal{U}_l^0 \leftarrow \mathbf{0}$, $\mathcal{A}_i^0 \leftarrow \mathbf{0}$, $\mathcal{B}_i^0 \leftarrow \mathbf{0}$, $\mathcal{C}^0 \leftarrow \mathbf{0}$

While not converged do

2) $k \leftarrow k + 1$

3) Solve \mathcal{T}_i^{k+1} by **Algorithm 2** in the appendix

4) Solve \mathcal{X}_i^{k+1} by Eq. (9)

4) Solve \mathcal{V}^{k+1} by Eq. (13)

6) Solve \mathcal{U}_l^{k+1} by Eq. (15)

7) Update Lagrangian multipliers \mathcal{A}_i^{k+1} , \mathcal{B}_i^{k+1} and \mathcal{C}^{k+1} by Eq. (18)

Endwhile

Algorithm 1 mainly lists all steps of the ADMM based framework that are utilized to iteratively and alternatively solve the proposed fusion model (4). In Algorithm 1, although it involves many parameters, most of them are not sensitive to different examples, and we will present parameters selection and discussion detailedly in the next section.

In what follows, we will compare the proposed approach with some competitive methods.

3. Results and discussions

In this section, we compare the proposed method with extensive image fusion approaches on several datasets acquired by four different sensors, *i.e.*, Pléiades, IKONOS, Quickbird

and WorldView-2. These compared approaches include: three CS based methods, *i.e.*, Principal Component Analysis (PCA) [7], Gram Schmidt (GS) [8], Nonlinear Intensity-Hue-Saturation (NIHS) [60]; two MRA based methods, *i.e.*, Additive Wavelet Luminance Proportional (AWLP) [9], Generalized Laplacian Pyramid (GLP) [11]; and three regularization based methods, *i.e.*, Dynamic Gradient Sparsity (DGS) [23], Coupled Nonnegative Matrix Factorization (CNMF) [30], Pansharpening based on Hyper-Laplacian Penalty (PHLP) [24]. The source codes of some compared methods are freely available on the website⁴. The scale factors are all set as 4. In particular, the number of bands is 4 for Pléiades, IKONOS and Quickbird datasets (*i.e.*, $z = 4$) and 8 for WorldView-2 dataset (*i.e.*, $z = 8$). All experiments are implemented in MATLAB(R2016a) on a laptop of 8Gb RAM and Intel(R) Core(TM) i7-7500U CPU: @2.70 GHz 2.90GHz.

Although the proposed method involves many parameters, most of them are relatively stable. For the selection of parameters, we empirically set $\gamma = 1 \times 10^{-3}$, $\eta_1 = \eta_2 = 5 \times 10^{-3}$, $\eta_3 = 1 \times 10^{-3}$, $\alpha_1 = \alpha_2 = 5 \times 10^{-3}$, $\omega_i = 1 \times 10^{-2}$, $\beta_i = 1 \times 10^{-2}$, $i = 1, 2$, for all datasets, and set 1) $\alpha_3 = 10$, $\lambda = 0.5$ for IKONOS, Quickbird, and two real datasets, 2) $\alpha_3 = 20$, $\lambda = 0.05$ for Pléiades dataset, and 3) $\alpha_3 = 30$, $\lambda = 0.01$ for WorldView-2 dataset. Note that, fine tuning of parameters for different datasets may lead to better results, but we unify the parameters to illustrate the stability of our method. More discussions on the selection of parameters can be found in Section 3.6.

To simulate LRMS images, we first apply Gaussian blur kernel to ground-truth HRMS image and then downsample it to get simulated LRMS image by a scale factor 4 with a very simple nearest interpolation. In the experiments, all LRMS images are obtained in the same way. Panchromatic images are simulated by different conditions which will be detailedly described in the following subsections.

The spectral quality of the fused multispectral images is quite important, it is however difficult to judge visually. To compare the spectral and spatial performance of different fusion methods, there exist many metrics to evaluate, including universal image quality index (QAVE) [61], Q4 for 4-band image, Q8 for 8-band image, spectral angle mapper (SAM) [2], ERGAS [62], quality without reference (QNR) [63], correlation coefficient (CC) [2], relative average spectral error (RASE) [64], peak signal to noise ratio (PSNR), root mean square error (RMSE), structural similarity index (SSIM) [65], and significance study [66]. However, *some of them represent the similar meaning, thus we only select some representative metrics such as Q4, Q8, SAM, ERGAS, QNR, PSNR and SSIM to evaluate the performance of different methods.* In particular, the larger Q4, Q8, QNR, PSNR, SSIM and the smaller SAM, ERGAS are shown, the better performance is obtained. For fair comparisons and

⁴For PCA, GS, AWLP, GLP and CNMF methods, we can download the source codes from: <http://openremotesensing.net/knowledgebase/a-critical-comparison-among-pansharpening-algorithms/>; For NIHS method, the code can be found from: <http://openremotesensing.net/knowledgebase/nonlinear-ihs-a-promising-method-for-pan-sharpening/>; For DGS method, the source code can be downloaded from: <http://cchen156.web.engr.illinois.edu/SIRF.html>; For PHLP method, we write the code according to the corresponding paper.

unified computing of metrics⁵, here we rescale all computed images into the range of [0, 1], just like the way in [24, 60], and then compute the corresponding metrics directly.

In the next subsections, we will report the visual and quantitative performance for each datasets, as well as give more analysis about the parameters selection, the influence of different energy terms, *etc.*

In particular, we divided the examples into two categories: 1) Simulated PAN image, see the example of IKONOS, Quickbird, and Rio data; The weights for this case are predefined (known) in advance. In particular, the PAN image for Pléiades data is also simulated, but the weight for this case is not predefined (unknow). The estimation way is same as the following real PAN case. 2) Real PAN image, see examples in Section 3.5. The weights need to be estimated by an automatic strategy. For example, the PAN image with the size 512×512 and the LRMS image with the size $128 \times 128 \times 4$, we could first downgrade the PAN image by the MTF filter and a decimation operation (*e.g.*, nearest interpolation in our work) to get the degraded PAN image with the size 128×128 (named as DePAN), then we assume that the DePAN image is obtained based on the linear spectral combination of the spectral bands of the LRMS image: $\text{DePAN} = w_1 \cdot \text{LRMS}(:,:,1) + w_2 \cdot \text{LRMS}(:,:,2) + w_3 \cdot \text{LRMS}(:,:,3) + w_4 \cdot \text{LRMS}(:,:,4)$, which is easy to calculate the weight w_i (see details from [44]).

3.1. IKONOS data

We first take a data that is named “Sichuan” and acquired from IKONOS sensor to test the performance of different methods. This data is with 16-bit and can be downloaded from the website⁶. We first crop a $512 \times 512 \times 4$ part of given multispectral image as ground-truth HRMS image, then generate LRMS image of size $128 \times 128 \times 4$ by the way introduced at the beginning of this section, and finally simulate the panchromatic image of size 512×512 via the linear combination of each band of the ground-truth HRMS image with weights [0.1 0.35 0.45 0.1], which indicates that the simulated PAN image is with the relation of: $\text{PAN} = 0.1 \cdot \text{HRMS}(:,:,1) + 0.35 \cdot \text{HRMS}(:,:,2) + 0.45 \cdot \text{HRMS}(:,:,3) + 0.1 \cdot \text{HRMS}(:,:,4)$.

Fig. 2 shows the visual results of different approaches, where (a), (b) and (c) stand for ground-truth HRMS, panchromatic and LRMS images. Thereinto, the ground-truth HRMS image contains excellent spatial details and spectral information. The panchromatic image also contains favorite spatial details but only with single image band which results in bad spectral performance. Although LRMS image includes good spectral preservation, its spatial details have degenerated. From the figure, it is clear that PCA, GS, PHLP, NIHS, and DGS preserve image details well. However, these methods fail to keep spectral information of the multispectral image and lead to significant spectral distortion. Although AWLP, GLP, and CNMF generate relatively good spatial and spectral preservation, they still can not outperform our method. Tab. 1 reports quantitative comparisons on extensive metrics. From this table, it is easy to see that the proposed method yields the competitive performance for most of all metrics, which demonstrates the effectiveness of our approach.

⁵Some methods rescale image intensity into [0, 1] in the provided code, while others keep the original

Table 1: Quantitative results on **IKONOS** dataset (Sichuan) for different methods (Bold: the best; Underline: the second best).

	PSNR	SSIM	QNR	ERGAS	Q4	SAM
GS	25.2566	0.8346	0.8246	5.2288	0.8579	5.0989
PCA	22.1698	0.7808	0.8059	8.5064	0.8332	5.9192
AWLP	25.6254	0.8151	0.8725	5.1501	0.8943	6.1350
GLP	25.3338	0.7774	0.8843	5.3411	0.8789	6.8087
CNMF	<u>29.8108</u>	0.9318	<u>0.9401</u>	<u>2.3224</u>	<u>0.9461</u>	<u>4.5503</u>
DGS	27.1984	0.8357	0.8945	3.5319	0.8901	7.6174
PHLP	26.2043	0.8743	0.9147	3.3241	0.9157	5.0401
NIHS	26.2449	0.8741	0.9120	3.2922	0.9035	5.3845
Our	30.0673	<u>0.9252</u>	0.9410	2.3085	0.9501	4.0735
Ideal value	$+\infty$	1	1	0	1	0

3.2. Quickbird data

In this section, a data from the Quickbird sensor is employed for the experiments. This data is with 9-bit and can be freely downloaded from the website⁷. We also crop a $512 \times 512 \times 4$ part from the downloaded multispectral image as the ground-truth HRMS and simulate LRMS image by the same way as IKONOS data. The panchromatic image is simulated by the linear weights [0.05 0.45 0.45 0.05] to each band of the ground-truth HRMS image. Our goal is to get the fused high spatial-resolution HRMS image by the simulated LRMS and panchromatic images.

From Fig. 3, the visual results by PCA and NIHS show weak spectral information preservation. For instance, the fused image by NIHS shows bad color contrast and spectral preservation, which are significantly not consistent with the ground-truth HRMS image (*i.e.*, Fig. 3 (a)). Moreover, although GS and AWLP methods exhibit relatively good ability of spectral preservation, they lose many spatial details that may result in the undesired visual display. GLP and the proposed approaches both perform excellently on increasing spatial details and preserving spectral information, but our method could hold better spatial details than GLP approach (*e.g.*, see cars on the road). In Tab. 2, the proposed method obtains the promising quantitative performance for all metrics. From the results of this table, AWLP, GLP and DGS methods show the excellent ability of spectral preservation since their SAM values exhibit competitive performance. In particular, our method also shows promising results.

intensity range.

⁶<http://glcf.umd.edu/data/ikonos/>

⁷<http://www.digitalglobe.com/product-samples>

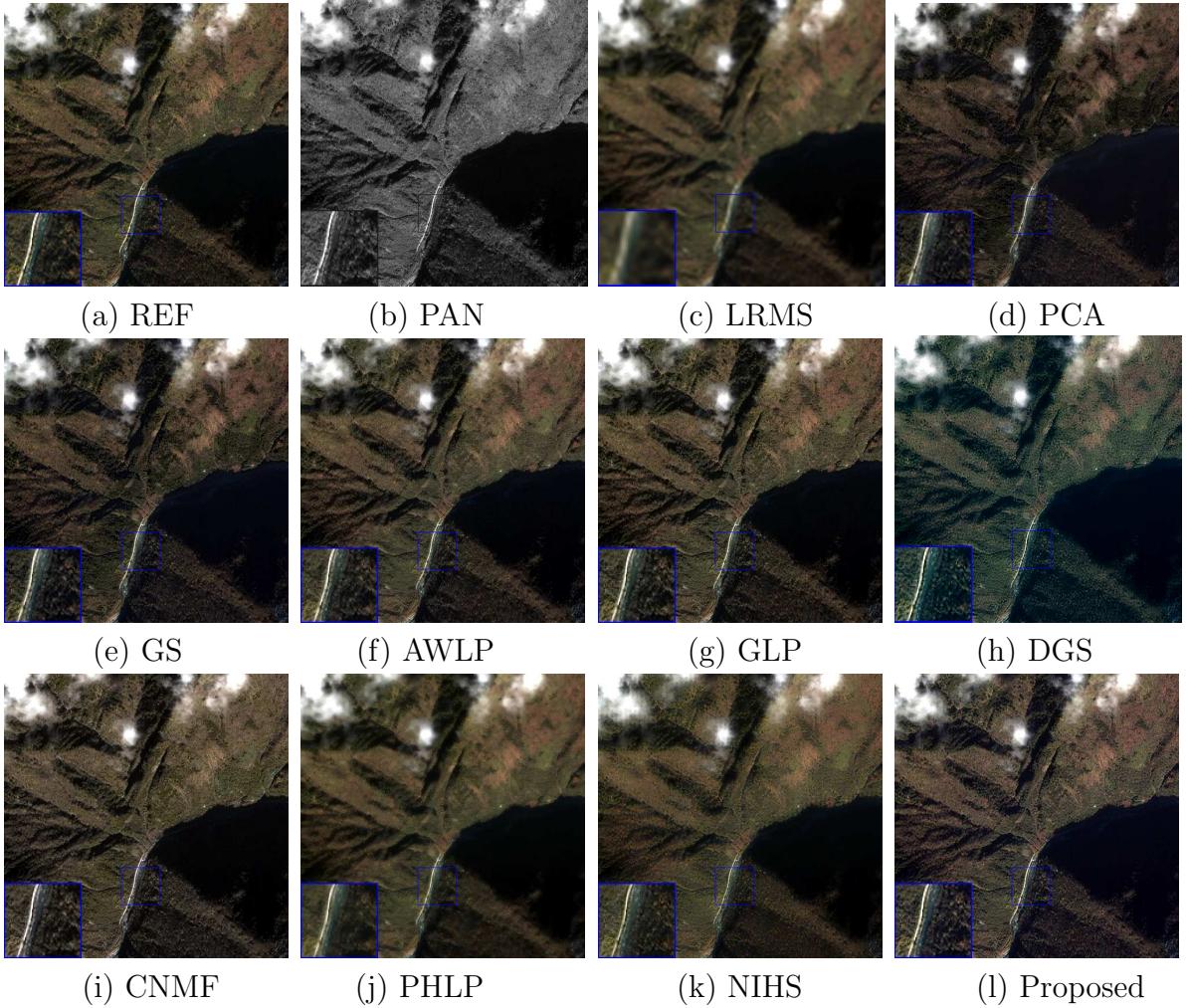


Figure 2: (a) The ground-truth HRMS image (**IKONOS dataset**, $512 \times 512 \times 4$); (b) PAN image; (c) LRMS image; (d)-(l) Visual results by PCA, GS, AWLP, GLP, DGS, CNMF, PHLP, NIHS methods and the proposed method. The weights for this example are predefined as [0.1 0.35 0.45 0.1].

3.3. Pléiades data

In this part, we take a data acquired from Pléiades sensor to compare the performance of different approaches. We can link the website⁸ to download this data, which has four available multispectral bands (with the size of $1024 \times 1024 \times 4$) collected by the aerial platform have 60 cm resolution. For fast computation and comparisons, here we only crop the top-left part of size $512 \times 512 \times 4$ as the ground-truth HRMS. In addition, the LRMS image with spatial resolution of four times lower than that of the ground-truth HRMS image is simulated according to the Wald’s protocol, namely by MTF filtering and decimation, *i.e.*, with size $128 \times 128 \times 4$. Different with previous experiments, *the weights are unknown*

⁸<http://openremotesensing.net/knowledgebase/a-critical-comparison-among-pansharpening-algorithms/>

Table 2: Quantitative results on **Quickbird** dataset for different methods (Bold: the best; Underline: the second best).

	PSNR	SSIM	QNR	ERGAS	Q4	SAM
GS	24.9937	0.9296	0.8097	2.9514	0.8204	2.2152
PCA	24.9388	0.9304	0.8924	2.9649	0.8152	2.1849
AWLP	29.6728	0.9360	<u>0.9197</u>	1.8851	0.8059	<u>1.7772</u>
GLP	<u>30.6675</u>	0.9470	0.9073	<u>1.6694</u>	0.9056	1.6315
CNMF	30.3471	<u>0.9480</u>	0.9151	1.6939	<u>0.9551</u>	2.5169
DGS	28.4666	0.9448	0.8875	2.1591	0.8741	<u>1.5707</u>
PHLP	25.9346	0.8128	0.9053	2.7770	0.8000	2.6729
NIHS	24.2877	0.8355	0.8929	3.3881	0.7644	2.7362
Our	35.3173	0.9765	0.9228	0.9793	0.9695	1.4641
Ideal value	$+\infty$	1	1	0	1	0

Table 3: Quantitative results on **Pléiades** dataset for different methods (Bold: the best; Underline: the second best).

	PSNR	SSIM	QNR	ERGAS	Q4	SAM
GS	24.4189	0.7031	0.75405	6.8177	0.5939	9.5480
PCA	23.4296	0.6326	0.7614	8.0242	0.5846	10.6617
AWLP	28.9008	0.8900	0.7897	4.1960	0.7608	6.5899
GLP	<u>29.8154</u>	<u>0.9148</u>	0.8845	<u>3.7787</u>	0.8973	5.5111
CNMF	29.2537	0.8770	0.9472	4.0243	0.8820	5.8090
DGS	26.3182	0.9052	0.8320	4.9647	<u>0.9105</u>	4.1410
PHLP	25.8028	0.8736	<u>0.9441</u>	5.2505	0.9026	<u>3.9697</u>
NIHS	22.4724	0.8255	0.8642	7.3949	0.8247	5.8551
Our	33.2681	0.9490	0.9345	2.3829	0.9243	3.8409
Ideal value	$+\infty$	1	1	0	1	0

in this experiment, thus we need to estimate the weights by the linear regression of the multispectral image and the spatially degraded panchromatic image [44]. The estimated weights are [0.1015, 0.4225, 0.4586, 0.0069]. In particular, the radiometric intensity is 11-bits.

Fig. 4 shows the visual results of different approaches. We can see that PCA, GS and AWLP methods not only fail to keep high spatial-resolution but also do not preserve spectral information. In addition, GLP and NIHS approaches keep relatively good spectral information, however, they can not obtain the best spatial image details (see blue close-ups). In particular, the proposed method and CNMF method yield similarly and competitively spatial and spectral results. Tab. 3 demonstrates the effectiveness of our approach for most of all metrics.

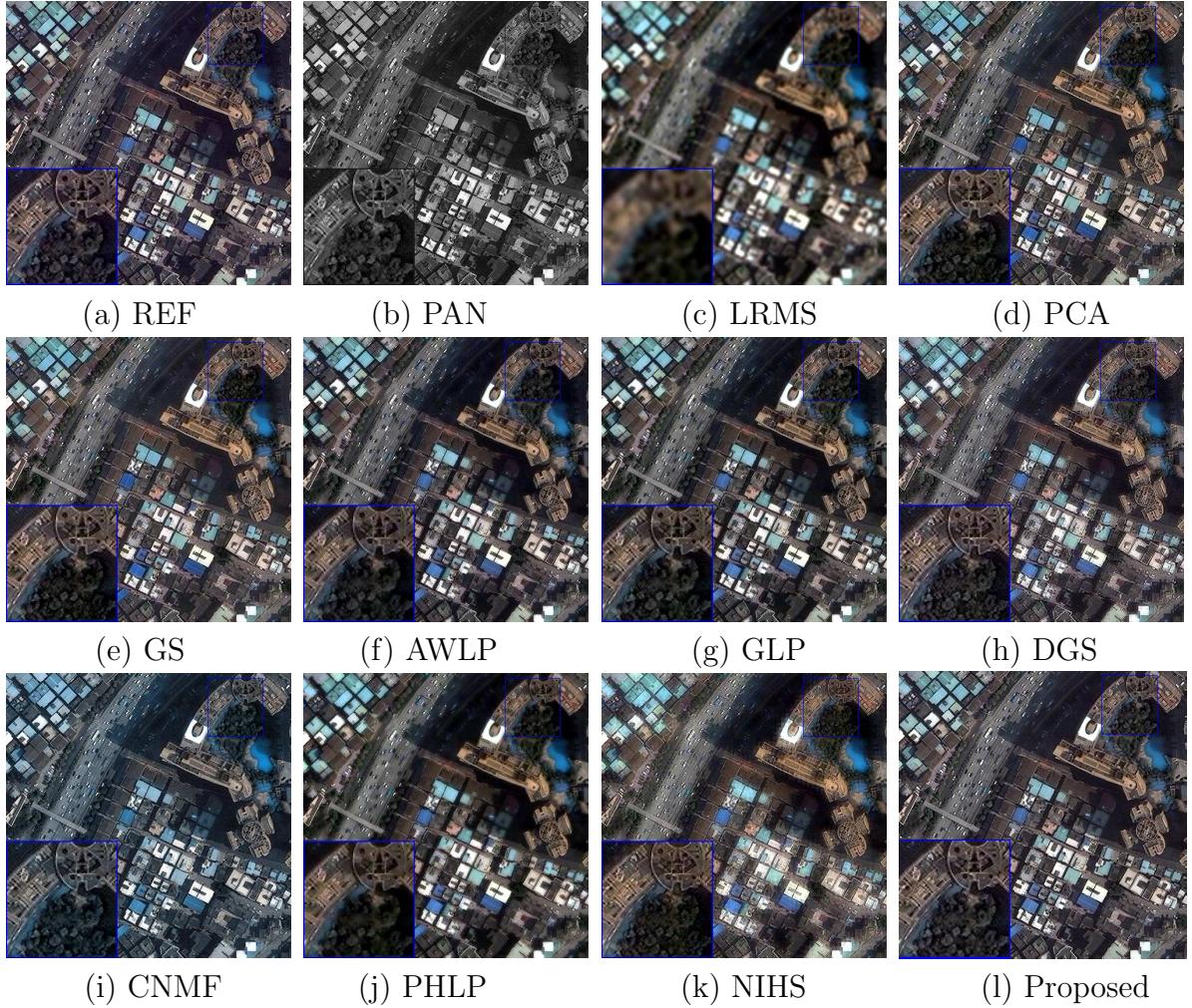


Figure 3: (a) The ground-truth HRMS image (**Quickbird dataset**, $512 \times 512 \times 4$); (b) PAN image; (c) LRMS image; (d)-(l) Visual results by PCA, GS, AWLP, GLP, DGS, CNMF, PHLP, NIHS methods and the proposed method. The weights for this example are predefined as [0.05 0.45 0.45 0.05].

3.4. WorldView-2 data

In this section, we employ a 8-band data (Rio) acquired from WorldView-2 sensor⁹ to test the performance of different methods. This data that is obtained by cropping the original image has eight available multispectral bands (with the size of $512 \times 512 \times 8$). We assume it as the ground-truth HRMS image and generate LRMS image of size $128 \times 128 \times 8$ by the same way as IKONOS data, and finally generate the panchromatic image of size 512×512 via the ground-truth HRMS image with the linear weights [0.05 0.05 0.2 0.2 0.2 0.2 0.05 0.05]. Note that since the provided codes of GLP and PHLP methods cannot support the computing of eight multispectral bands, thus we do not compare these two methods in this

⁹<http://www.digitalglobe.com/product-samples>

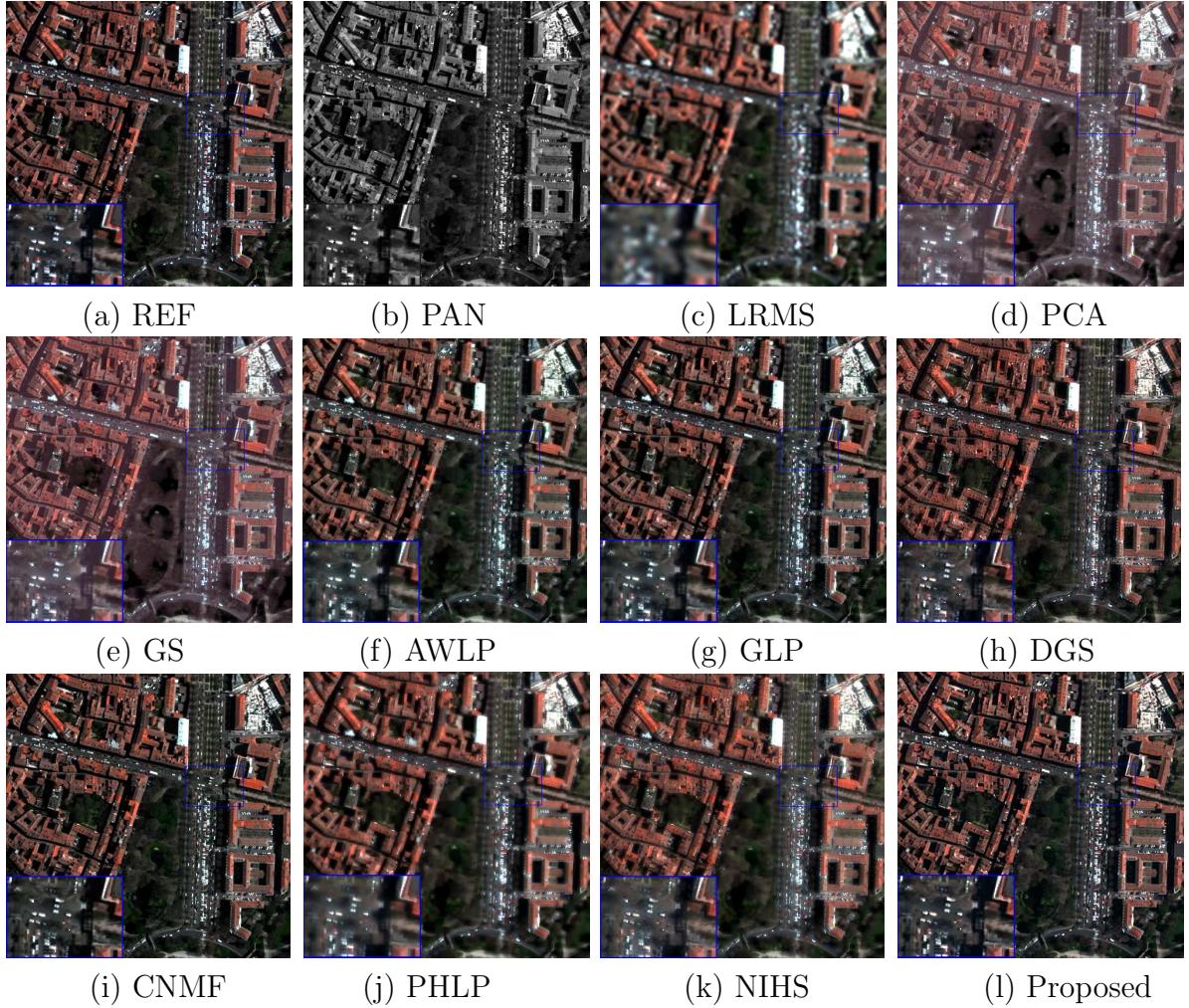


Figure 4: (a) The ground-truth HRMS image (**Pleiades dataset**, $512 \times 512 \times 4$); (b) PAN image; (c) LRMS image; (d)-(l) Visual results by PCA, GS, AWLP, GLP, DGS, CNMF, PHLP, NIHS methods and the proposed method. The weights for this example is not predefined, but automatically estimated by the linear regression of the multispectral image and the spatially degraded panchromatic image. The estimated weights are [0.1015, 0.4225, 0.4586, 0.0069].

experiment.

From Fig. 5, NIHS and CNMF approaches result in significant spectral distortion although they keep relatively good spatial details. PCA and DGS methods perform competitively from the spectral perspective, but the quantitative results shown in Tab. 4 are not so good compared with the proposed method. Additionally, AWLP and the proposed method performs competitively on the visual aspect, and they also obtain competitive quantitative results.

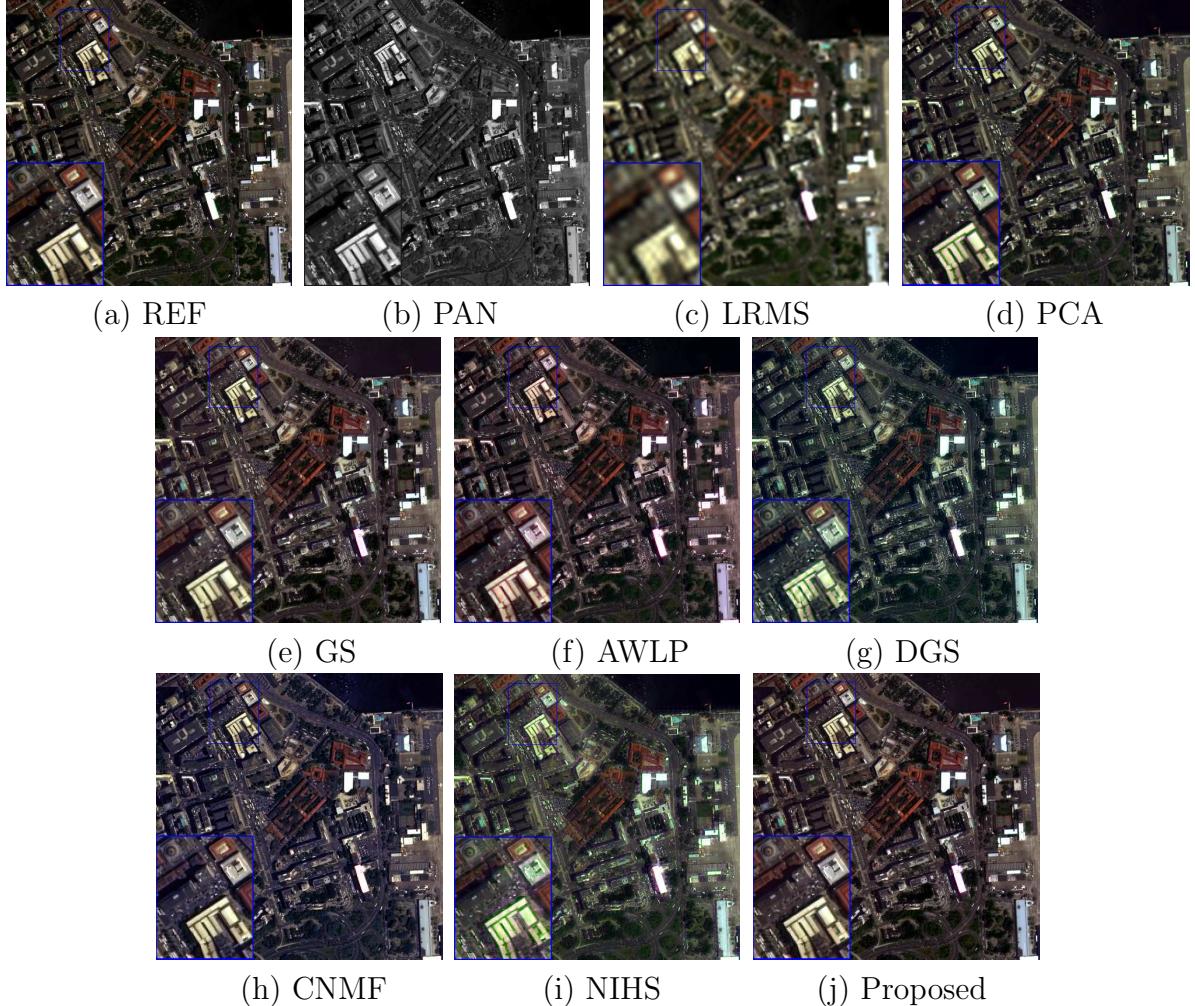


Figure 5: (a) The ground-truth HRMS image (**WorldView-2 dataset**, $512 \times 512 \times 8$); (b) PAN image; (c) LRMS image; (d)-(j) Visual results by PCA, GS, AWLP, DGS, CNMF, NIHS methods and the proposed method. The weights for this example are predefined as [0.05 0.05 0.2 0.2 0.2 0.2 0.05 0.05].

Table 4: Quantitative results on **WorldView-2** dataset (Rio) for different methods (Bold: the best; Underline: the second best).

	PSNR	SSIM	QNR	ERGAS	Q8	SAM
GS	26.9012	0.8292	0.8867	5.3388	0.7706	9.0570
PCA	28.9979	0.9118	0.8543	3.6304	0.8894	6.3350
AWLP	33.2816	0.9461	<u>0.9046</u>	1.7738	<u>0.9154</u>	4.6157
CNMF	31.3842	0.9254	0.7930	2.9088	0.9045	5.7530
DGS	30.9198	0.9260	0.8830	2.8525	0.9056	4.8345
NIHS	29.1831	0.9062	0.8000	3.3903	0.8982	5.3239
Our	33.0662	0.9419	0.9257	<u>1.9590</u>	0.9185	4.3578
Ideal value	$+\infty$	1	1	0	1	0

3.5. Real Data

In this subsection, we take two real datasets, *i.e.*, WorldView-2 (town) and GeoEye-1, to test the performance of some state-of-the-art methods. Since the proposed model needs to estimate the linear weights for each spectral bands, here we still utilize the strategy which has been introduced at the beginning of this section to calculate the weights. The estimated weights for WorldView-2 (town) and GeoEye-1 are about [0.15, 0.31, 0.34, 0.16] and [0.27, 0.08, 0.33, 0.26], respectively. Due to the lack of ground-truth HRMS images, we could only show visual results for the case of real dataset. From Fig. 6, it is clear that the proposed method also obtains competitive visual results, while other methods such that DGS, NIHS, PCA fail to preserve good spatial image details. Especially, the results by AWLP and GLP also show quite competitive performance comparing with the proposed method. By this experiment, it demonstrates the effectiveness of our method.

3.6. More discussions

In this section, we give more analysis for the proposed method based on extensive experiments. For the simplicity, here we only take the Quickbird dataset for all following tests.

- **Results with different $Eng^{(1)}$ term:** One main contribution of the proposed model (4) is the non-convex $\ell_{1/2}$ term which is promoted by investigating the error distributions for different datasets on the gradient domain. Here, we intend to investigate the performance with different $Eng^{(1)}$ terms (*i.e.*, Eq. (1)) to the model (4), which corresponding notations can be found from Tab. 5. All these models are similarly solved by ADMM based algorithm. Fig. 7 presents the visual results by the proposed model with different $Eng^{(1)}$ terms. For these models, Proposed- ℓ_1 and Proposed- ℓ_2 appear significant spectral distortion and fail to obtain excellent spatial details, while Proposed- $g\ell_2$ gets promising spectral preservation but results in a slight blur for spatial details. In addition, Proposed- $g\ell_1$ and Proposed- $g\ell_{1/2}$ (*i.e.*, the final proposed model) approaches perform similarly and competitively on visual results. However, the final

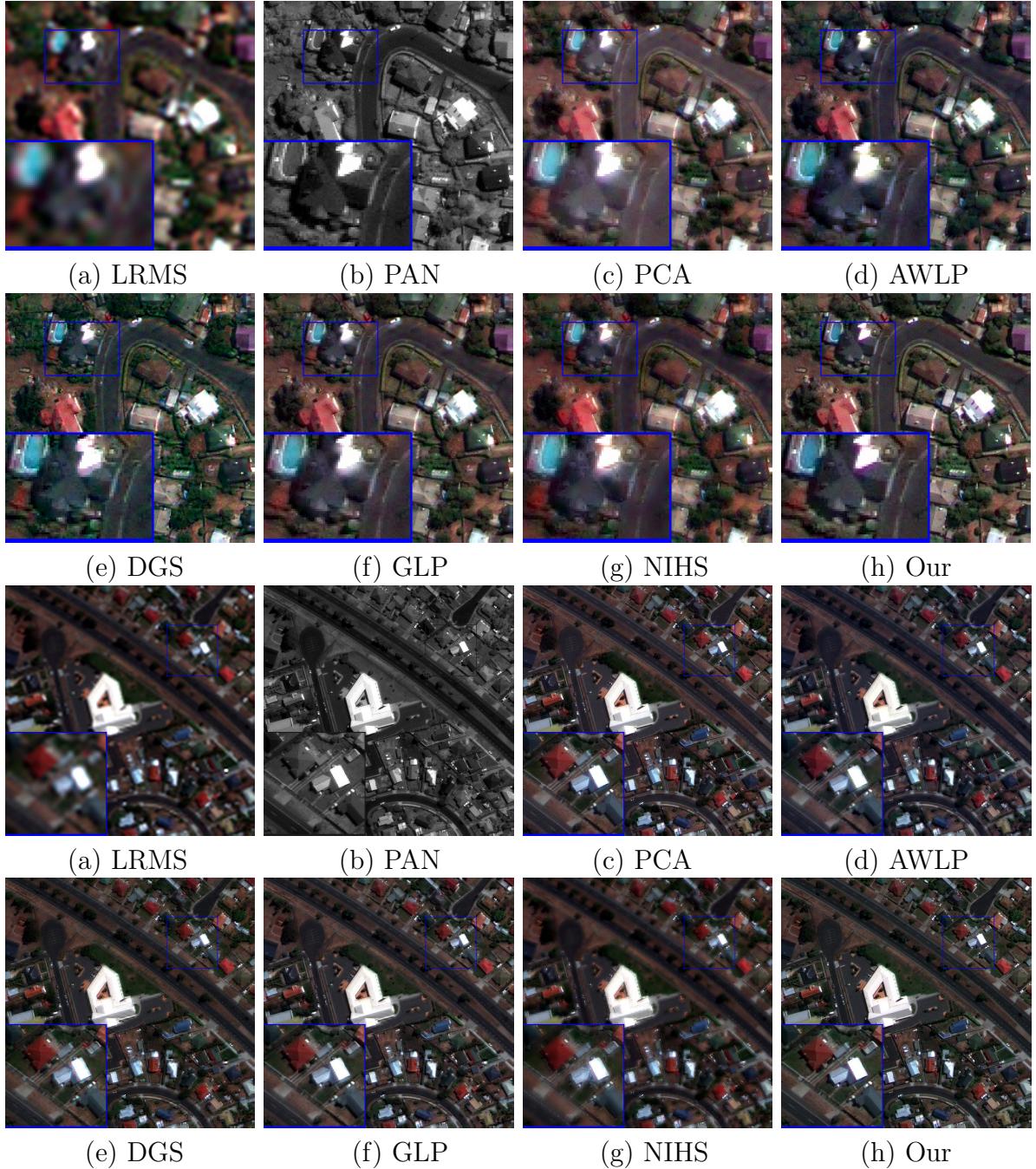


Figure 6: Visual comparisons on two real datasets, *i.e.*, WorldView-2 (town) in the first example and GeoEye-1 in the second example. (a) The upsampled LRMS image by bicubic interpolation; (b) PAN image; (c)-(h) Visual results by PCA, AWLP, DGS, GLP, NIHS and the proposed methods.

proposed model yields the best quantitative results on most of metrics (see Tab. 6). These visual and quantitative results are also consistent with existing common con-

clusions, that is ℓ_2 term may result in more smoothing results and ℓ_1 term can obtain relatively sharp edges and more high-frequency image details.

- **The influence of different parameters for the proposed method:** Although our approach involves many parameters, *e.g.*, γ , α_i ($i = 1, 2, 3$), λ , *etc.*, they are all relatively stable except α_3 and λ for all testing datasets. Therefore, here it is necessary to show the influence of different parameters for the given method. Note that, since there are only slightly quantitative changes for some parameters, here we treat the obtained metrics by $(metric - mean(metric))/std(metric)$ where *mean* and *std* stand for the mean value and standard deviation, respectively, aiming to distinguish the slight difference. In Fig. 8, we take Quickbird dataset as an example. It is clear that the suggested parameters combination at the beginning of this section is the best choice for the Quickbird dataset for almost of all metrics. Note that, fine tuning one parameter may lead to better results for one specific metric, however, it may result in worse results for other metrics. Therefore, to balance the performance of all metrics, we take the suggested parameters. In particular, we fix other parameters but α_3 and λ . The main reason is that we tend to fix most of parameters and only change few of them to make the parameter selection simpler. About α_3 , we may roughly make the following rule: if the intensity difference between each band of a multispectral image is small, we may set a relatively big α_3 ; if the difference is big, the α_3 should be set as a small value. Moreover, we may also make the following rules for λ : 1) if $\|\mathcal{U}_{l \times 3}\mathbf{w} - \mathbf{p}\|_2^2$ is small, the λ should be set as a big value, *e.g.*, the example of accurate weight \mathbf{w} for 4-band image. 2) if $\|\mathcal{U}_{l \times 3}\mathbf{w} - \mathbf{p}\|_2^2$ is big, the λ should be set as a small value, *e.g.*, the example of inaccurate weight \mathbf{w} for 8-band image.
- **The influence of three energy terms in our model (4):** The final proposed model (4) includes three energy terms, here we will exhibit the influence of each term on the final results. From Fig. 9, we know that the second term of our model, *i.e.*, $Eng^{(2)}$ that involves the linear relation between panchromatic and latent HRMS images, plays the most important role to the final results. Moreover, the first term $Eng^{(1)}$ also shows important influence on our results, which means if we discard it, the final quantitative performance will decrease significantly. Furthermore, the third term $Eng^{(3)}$ in our model slightly affects the final results (see Fig. 9), but it is also a necessary part in the final model. In summary, all energy terms in our model (4) are quite essential to the final results, therefore we do not remove any one of them.
- **The differences between the proposed method and [57, 24]:** The difference between the proposed method and [57] mainly is explained below: 1) The method in [57] mainly focuses on using matrix-based modeling to deal with single image (2D) processing, *i.e.*, image deconvolution, while the proposed method mainly focuses on using tensor-based modeling to deal with two images application, *i.e.*, the fusion of panchromatic image(2D) + multispectral image (3D). 2) Our model and algorithm are different from [57]. Therefore, they are different in modeling, algorithm and application. Moreover, the difference between the proposed method and [24] mainly has

the following two aspects: 1) Different modeling which includes: i) The method in [24] considers that the LRMS image y_b should approximate to a decimated version of HRMS image x_b via convolution with a blurring kernel k , i.e., $\|y_b - k * x_b\|_2^2$. The kernel k is unknown and need to be estimated by users, and the inaccurate kernel may affect the final result significantly. Whereas the proposed method does not use the kernel to avoid the issue of kernel estimation. ii) The work in [24] formulates the matrix-based model for each band of the multispectral image, while our method utilizes the tensor-based model to directly describe the multispectral image (not band-by-band), which make the model and algorithm simple by some tensors tools. 2) Different algorithm: The algorithms for the models in [24] and this paper are also quite different. Note that, the proposed method and [57, 24] also have the similarity, that is they all use the similar statistical analysis to investigate and discover the latent hyper-Laplacian prior which leads to l_p norms for the fidelity term in the energy functional.

- **The influence of the keynote parameter p :** In Fig. 10, we test the influence of the keynote parameter p with the test range of $\{0.3, 0.4, 0.5, 0.6, 0.7\}$. From the figure, we know that when $p = 0.5$, it could get the best SAM and Q4 performance for the Quickbird example, while $p = 0.4$ can obtain the best ERGAS result. For convenience, we tend to uniform $p = 0.5$ for all examples, even though $p = 0.5$ cannot get the best performance for all examples and metrics.
- **The influence of the weight \mathbf{w} :** In (2), we assume that the panchromatic image is the linear combination of spectral bands of a latent HRMS image. The weight is a key parameter for the proposed model. In this part, we first let the weight \mathbf{w}_i of one band be small¹⁰, then to see what will happen for the spatial performance of this band. Here, we take band 2 of Quickbird data as an example (with the corresponding weight \mathbf{w}_2). From figure 11, the PSNR and SSIM, two common metrics to evaluate spatial performance, are reduced from 40.20 and 0.9838 (figure 11(b)) to 17.03 and 0.9004 (figure 11(c)), respectively. It indicates that the spatial details if we let the weight be small will not be preserved well and the intensity contrast also becomes worse. More related discussions can be found from [67].

4. Conclusions

In this paper, a non-convex sparse model based on tensor format was proposed for the fusion of panchromatic and multispectral remote sensing images. The given model mainly included three terms, one term is to depict the investigated hyper-Laplacian distribution between the upsampled multispectral image and the latent HRMS image, which results in one ℓ_p ($0 < p < 1$) energy term; another term is to describe the linear relation between the panchromatic image and latent HRMS image, which has been utilized in many types of

¹⁰if there are four bands, $\mathbf{w} = [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4]$

Table 5: Notation of the proposed model (4) with different $Eng^{(1)}$ terms.

Notation	$Eng^{(1)}$
Proposed- ℓ_1	$\sum_{i=1}^3 \alpha_i \ \mathcal{U}_l - \mathcal{M}\ _1$
Proposed- ℓ_2	$\sum_{i=1}^3 \alpha_i \ \mathcal{U}_l - \mathcal{M}\ _2$
Proposed- $g\ell_1$	$\sum_{i=1}^3 \alpha_i \ \nabla_i(\mathcal{U}_l - \mathcal{M})\ _1$
Proposed- $g\ell_2$	$\sum_{i=1}^3 \alpha_i \ \nabla_i(\mathcal{U}_l - \mathcal{M})\ _2$
Proposed (<i>i.e.</i> , Proposed- $g\ell_{1/2}$)	$\sum_{i=1}^3 \alpha_i \ \nabla_i(\mathcal{U}_l - \mathcal{M})\ _{1/2}$

Table 6: Quantitative results of different models for Quickbird dataset.

	PSNR	SSIM	QNR	ERGAS	Q4	SAM
Proposed- ℓ_1	22.8904	0.6633	0.8190	3.9102	0.5944	4.7852
Proposed- ℓ_2	23.0484	0.6437	0.8845	3.8720	0.7053	3.0307
Proposed- $g\ell_1$	<u>35.1150</u>	0.9690	0.9300	<u>0.9977</u>	0.9039	1.5213
Proposed- $g\ell_2$	34.6150	<u>0.9724</u>	0.9105	1.0600	<u>0.9086</u>	<u>1.5020</u>
Proposed	35.2019	0.9760	0.9240	0.9853	0.9167	1.4854

research; and the last term is to model the sparsity of the latent HRMS image on the gradient domain. In addition, the proposed non-convex model could be efficiently solved by a recent GST algorithm based on the framework of ADMM. Extensive experiments demonstrate that our method could outperform recent state-of-the-art fusion approaches, both visually and quantitatively. Furthermore, we also reported more results analysis about the influences of parameters, non-convex term, and energy terms. The analysis demonstrated the effectiveness of the proposed method.

This method also encounters some drawbacks. For instance, although the tensor-based modeling is a promising way for high-dimensional data processing, it is difficult to obtain excellent results for the fusion of panchromatic and *hyperspectral* remote sensing images, since the linear relation between these two images may not hold at this moment. In addition, the related non-convex ℓ_p ($0 < p < 1$) model fails to hold global convergence under the ADMM based framework.

5. Acknowledgment

The first author would like to thank the supports by NSFC (61702083, 61772003) and Fundamental Research Funds for the Central Universities (ZYGX2016KYQD142).

6. Appendix

The paper presents a tensor-based non-convex modeling approach for the fusion of panchromatic and multispectral images. The non-convex sparse term is depicted by the

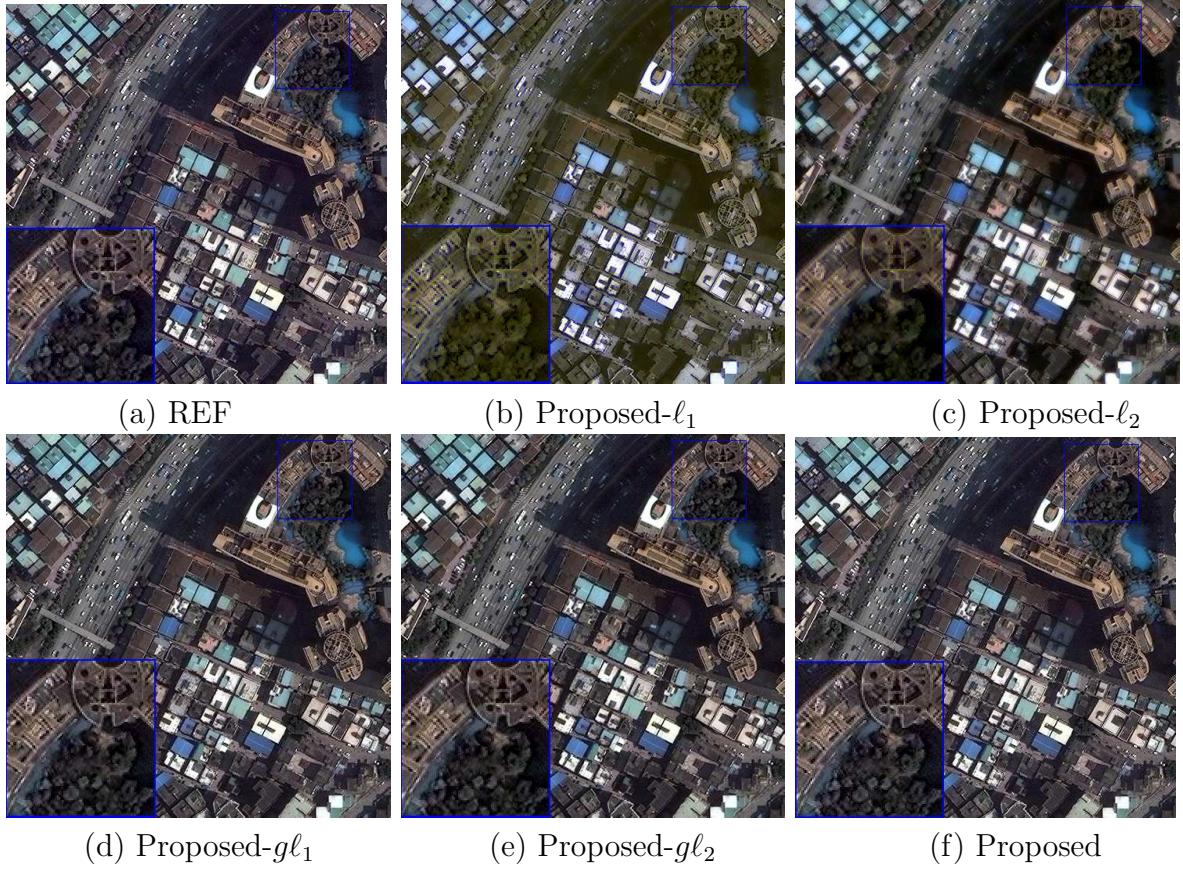


Figure 7: Visual results of different models in Tab. 5 for Quickbird dataset. (a) The ground-truth HRMS image; (b)-(f) The fused images by Proposed- ℓ_1 , Proposed- ℓ_2 , Proposed- $g\ell_1$, Proposed- $g\ell_2$ and the proposed model, respectively.

ℓ_p norm where $0 < p < 1$, which this term comes from the abundant distribution investigation for the errors between upsampled and ground-truth HRMS images. However, due to the non-convexity of ℓ_p ($0 < p < 1$) problem, it is quite difficult to design an efficient algorithm to solve it with guarantees of global convergence. Recently, there have appeared a number of algorithms for solving the non-convex ℓ_p problems for image applications, *e.g.*, compressive sensing [56], image restoration [57, 41], face recognition [58]. Here, we choose an efficient and effective non-convex algorithm [41] for the given non-convex model. In what follows, we will first introduce this algorithm which is related to our final method.

Inspired by soft-thresholding, Zuo *et al.*, [41] proposed a generalized shrinkage/thresholding (GST) algorithm to solve the following non-convex ℓ_p minimization problem,

$$\min_x \frac{1}{2}(x - y)^2 + \lambda|x|^p, \quad 0 < p < 1, \quad (19)$$

where y is the known 1D signal. For the minimization problem (19), if $y > 0$, the solution of (19) will fall into the range of $[0, y]$; otherwise, into the range of $[y, 0]$. Without loss of

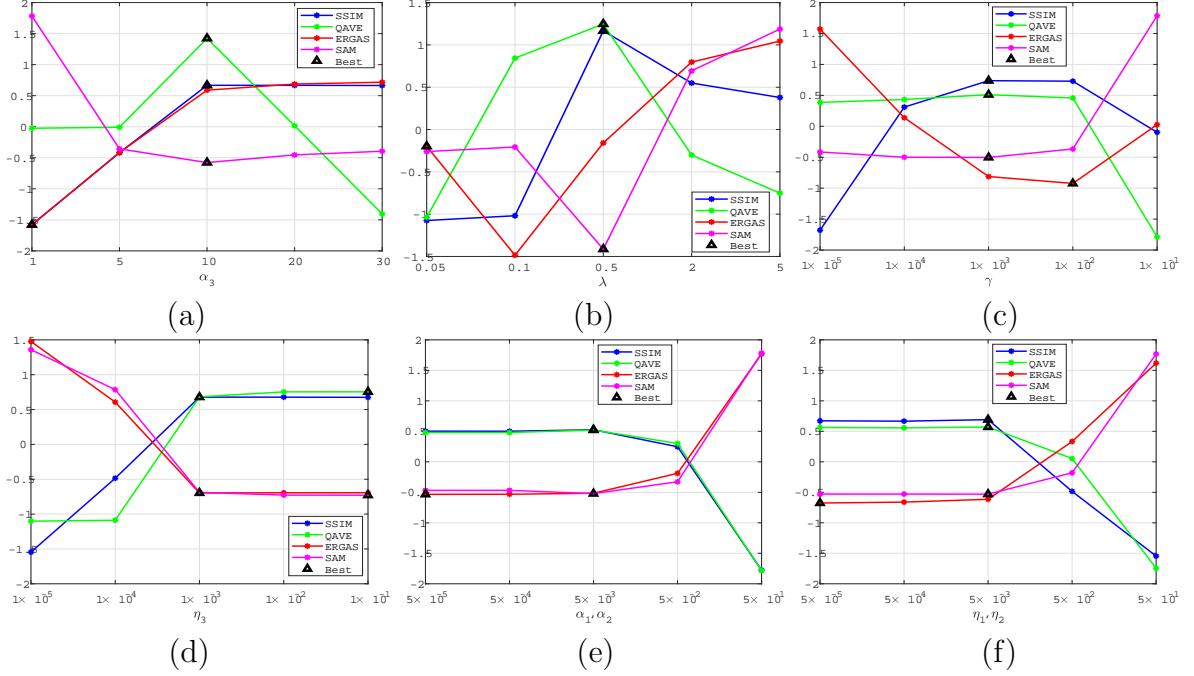


Figure 8: The influence of different parameters for the proposed method on Quickbird dataset. We obtain the results by varying the target parameter and fixing other ones. (a)-(f) are the quantitative results with varying α_3 , λ , γ , η_3 , α_i and η_i ($i = 1, 2$), respectively. Note that, to better distinguish the performance of different approaches, we treat the obtained metrics by $(metric - mean(metric))/std(metric)$ where $mean$ and std represent the mean value and standard deviation, respectively.

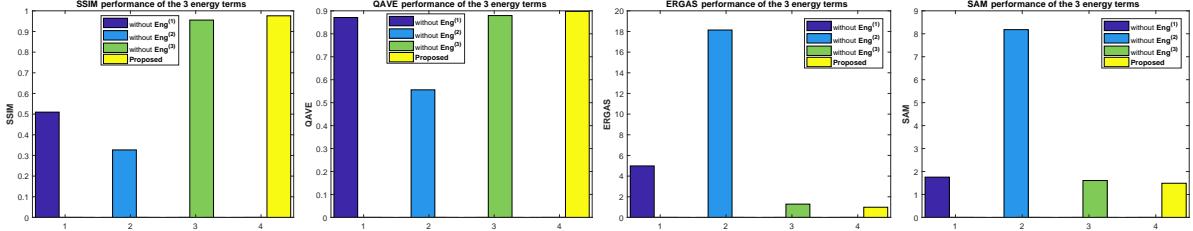


Figure 9: The quantitative performance when removing *one* of the three energy terms, i.e., $Eng^{(1)}$ - $Eng^{(3)}$. It is clear that the proposed model performs best than the models of removing one energy term.

generality, here we only consider the case of $y > 0$. Let $f(x) = \frac{1}{2}(x - y)^2 + \lambda|x|^{p+1}$ which is differentiable in the range of $(0, +\infty)$, thus the first-order derivative of $f(x)$ is as follows,

$$f'(x) = x - y + \lambda p x^{p-1}, \quad (20)$$

By setting $p = 0.5$ and $\lambda = 1$, the plots of $f(x)$ with a typical y is given in Fig. 12. From this figure, it is clear that there exists a specific threshold $\tau_p^{GST}(\lambda)$ after fixing p and λ , if $y < \tau_p^{GST}(\lambda)$, $x = 0$ is the global minimum of (19); otherwise, one nonzero solution will be optimal. Therefore, to generate the final GST formula, it should address two important

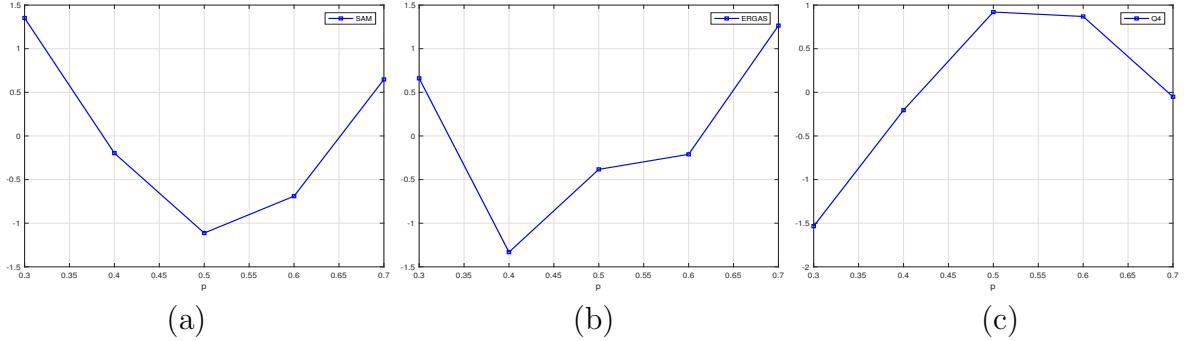


Figure 10: The SAM (a), ERGAS (b) and Q4 (c) influence of the keynote parameter p with the test range of $\{0.3, 0.4, 0.5, 0.6, 0.7\}$ for the Quickbird example. Similar to Fig. 8, to better distinguish the difference of values with quite small changes, we treat the obtained metrics by $(metric - mean(metric))/std(metric)$ where $mean$ and std represent the mean value and standard deviation, respectively.

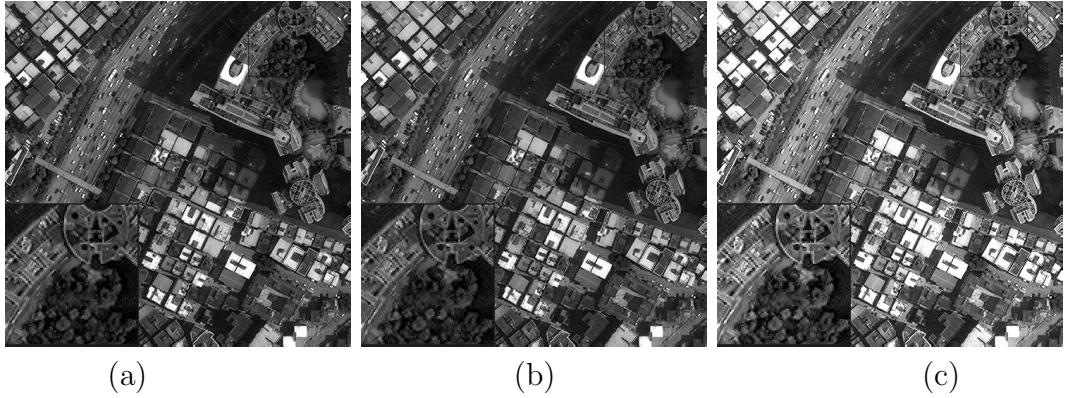


Figure 11: Test results for the weight influence of Quickbird data. (a) The ground truth of band 2 of Quickbird data; (b) The visual result of band 2 using the predefined weight $w = [0.05, 0.45, 0.45, 0.05]$; (c) The visual result of band 2 if we deliberately let w_2 be small, here using $w_2 = 0.2$. It indicates that the spatial details of (c) fail to be preserved well and the intensity contrast also becomes worse. Especially, the PSNR and SSIM decrease from 40.20 and 0.9838 (b) to 17.03 and 0.9004 (c), respectively.

issues, one is to calculate the threshold $\tau_p^{GST}(\lambda)$, the other is to search the nonzero solution by one fast strategy.

As shown in Fig. 12, it is easy to know that there is a specific y value to make $f(x_p^*) = f(0)$, which means:

$$\frac{1}{2}(x_p^* - \tau_p^{GST}(\lambda))^2 + \lambda|x_p^*|^p = \frac{1}{2}(\tau_p^{GST}(\lambda))^2, \quad (21)$$

where y is replaced by $\tau_p^{GST}(\lambda)$ to determine the threshold value. In addition, it also has the relation $f'(x_p^*) = 0$, *i.e.*,

$$x_p^* - \tau_p^{GST}(\lambda) + \lambda p(x_p^*)^{p-1} = 0. \quad (22)$$

Combining Eq. (21) and Eq. (22), we may compute the solution x_p^* in the range of

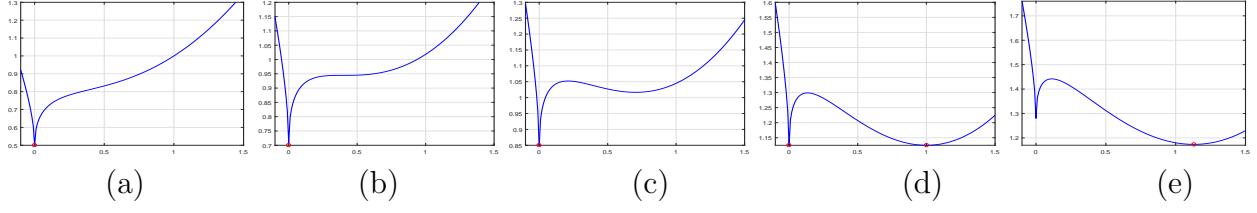


Figure 12: The plot of $f(x) = \frac{1}{2}(x - y)^2 + \lambda|x|^p$ when fixing $p = 0.5$, $\lambda = 1$ and with varying y values. (a) $y = 1$, (b) $y = 1.19$, (c) $y = 1.3$, (d) $y = 1.5$, (e) $y = 1.6$.

$$(x_0^{(\lambda,p)}, +\infty)^{11}, \\ x_p^* = (2\lambda(1-p))^{\frac{1}{2-p}}, \quad (23)$$

and then the threshold value $\tau_p^{GST}(\lambda)$ is

$$\tau_p^{GST}(\lambda) = (2\lambda(1-p))^{\frac{1}{2-p}} + \lambda p(2\lambda(1-p))^{\frac{p-1}{2-p}}. \quad (24)$$

Theorem 1 ([41]): For any $y \in (\tau_p^{GST}(\lambda), +\infty)$, $f(x)$ has one unique minimum $S_p^{GST}(y; \lambda)$ in the range of $(x_p^*, +\infty)$ by solving:

$$S_p^{GST}(y; \lambda) - y + \lambda p(S_p^{GST}(y; \lambda))^{p-1} = 0.$$

Theorem 2 ([41]): For any $y \in (\tau_p^{GST}(\lambda), +\infty)$, let $S_p^{GST}(y; \lambda)$ be the unique minimum of $f(x)$ in the range of $(x_p^*, +\infty)$, we have the following inequality:

$$f(0) > f(S_p^{GST}(y; \lambda)).$$

By the two theorems, the new thresholding function $T_p^{GST}(y; \lambda)$ is defined as follows,

$$T_p^{GST}(y; \lambda) = \begin{cases} 0, & \text{if } |y| \leq \tau_p^{GST}(\lambda), \\ \text{sgn}(y)S_p^{GST}(|y|; \lambda), & \text{if } |y| > \tau_p^{GST}(\lambda), \end{cases}$$

which is similar to soft-thresholding function.

After obtaining the threshold value $T_p^{GST}(y; \lambda)$, the authors in [41] finally propose an iterative GST algorithm, see Algorithm 2, which can solve the non-convex minimization problem efficiently.

¹¹ $x_0^{(\lambda,p)}$ is the solution of $f''(x) = 0$

Algorithm 2 (GST): $T_p^{GST}(y; \lambda) = GST(y, \lambda, p, J)$

Input: y, λ, p, J

- 1) $\tau_p^{GST}(\lambda) = (2\lambda(1-p))^{\frac{1}{2-p}} + \lambda p(2\lambda(1-p))^{\frac{p-1}{2-p}}$
- 2) **If** $|y| \leq \tau_p^{GST}(\lambda)$
- 3) $T_p^{GST}(y; \lambda) = 0$
- 4) **Else**
- 5) $k = 0, x^{(k)} = |y|$
- 6) **for** $k = 1:J$
- 7) $x^{(k+1)} = |y| - \lambda p(x^{(k)})^{p-1}$
- 8) **endfor**
- 9) $T_p^{GST}(y; \lambda) = \text{sgn}(y)x^{(k)}$
- 10) **Endif**

Output: $T_p^{GST}(y; \lambda)$

In Algorithm 2, J is the iteration number, which is suggested as 2 in [41] and is enough to obtain competitive results. The Algorithm 2 is quite related to the proposed method and will be applied to solving the involved non-convex subproblem.

References

- [1] C. Souza, L. Firestone, L. Silva, D. Roberts, Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models, *Remote Sensing of Environment* 87 (2003) 494–506.
- [2] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, L. M. Bruce, Comparison of pansharpening algorithms: Outcome of the 2006 GRSS data fusion contest, *IEEE Transactions on Geoscience and Remote Sensing* 45 (2007) 3012–3021.
- [3] M. Dalla Mura, S. Prasad, F. Pacifici, P. Gamba, J. Chanussot, Challenges and opportunities of multimodality and data fusion in remote sensing, *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)* (2014) 106–110.
- [4] C. Thomas, T. Ranchin, L. Wald, J. Chanussot, Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics, *IEEE Transactions on Geoscience and Remote Sensing* 46 (2008) 1301–1312.
- [5] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G. Licciardi, R. Restaino, L. Wald, A critical comparison among pansharpening algorithms, *IEEE Transactions on Geoscience and Remote Sensing* 53 (2015) 2565–2586.
- [6] P. S. Chavez Jr, S. C. Sides, J. A. Anderson, Comparison of three different methods to merge multiresolution and multispectral data- Landsat TM and SPOT panchromatic, *Photogrammetric Engineering and remote sensing* 57 (3) (1991) 295–303.
- [7] P. S. Chavez Jr, A. Y. Kwarteng, Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis, *Photogrammetric Engineering and Remote Sensing* 55 (3) (1989) 339–348.
- [8] C. A. Laben, B. V. Brower, Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening, uS Patent 6011875 (2000).
- [9] X. Otazu, M. González-Audicana, O. Fors, J. Núñez, Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods, *IEEE Transactions on Geoscience and Remote Sensing* 43 (2005) 2376–2385.
- [10] P. J. Burt, E. H. Adelson, The Laplacian pyramid as a compact image code, *IEEE Transactions on Communications* 31 (4) (1983) 532–540.

- [11] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis, *IEEE Transactions on Geoscience and Remote Sensing* 40 (2002) 2300–2312.
- [12] M. J. Shensa, The discrete wavelet transform: wedding the a trous and Mallat algorithms, *IEEE Transactions on Signal Processing* 40 (10) (1992) 2464–2482.
- [13] C. Ballester, V. Caselles, L. Igual, J. Verdera, B. Rougé, A variational model for P+XS image fusion, *International Journal of Computer Vision* 69 (2006) 43–58.
- [14] M. Möller, T. Wittman, A. L. Bertozzi, A variational approach to hyperspectral image fusion, *The International Society for Optical Engineering (SPIE)* 7334 (2009) 73341E.
- [15] S. Li, B. Yang, A new pan-sharpening method using a compressed sensing technique, *IEEE Transactions on Geoscience and Remote Sensing* 49 (2) (2011) 738–746.
- [16] X. X. Zhu, R. Bamler, A sparse image fusion algorithm with application to pan-sharpening, *IEEE Transactions on Geoscience and Remote Sensing* 51 (5) (2013) 2827–2836.
- [17] F. Fang, F. Li, C. Shen, G. Zhang, A Variational Approach for Pan-Sharpening, *IEEE Transactions on Image Processing* 22 (2013) 2822–2834.
- [18] J. Duran, A. Buades, B. Coll, C. Sbert, A Nonlocal Variational Model for Pansharpening Image Fusion, *SIAM Journal on Imaging Sciences* 7 (2014) 761–796.
- [19] F. Palsson, J. R. Sveinsson, M. O. Ulfarsson, A new pansharpening algorithm based on total variation, *IEEE Geoscience and Remote Sensing Letters* 11 (1) (2014) 318–322.
- [20] H. A. Aly, Sharma, A regularized model-based optimization framework for pan-sharpening, *IEEE Transactions on Image Processing* 23 (2014) 2596–2608.
- [21] X. He, L. Condat, J. M. Bioucas-Dias, J. Chanussot, J. Xia, A New Pansharpening Method Based on Spatial and Spectral Sparsity Priors, *IEEE Transactions on Image Processing* 23 (2014) 4160–4174.
- [22] C. Chen, Y. Li, W. Liu, J. Huang, Image fusion with local spectral consistency and dynamic gradient sparsity, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2014) 4213–4224.
- [23] C. Chen, Y. Li, W. Liu, J. Huang, SIRF: Simultaneous satellite image registration and fusion in a unified framework, *IEEE Transactions on Image Processing* 24 (2015) 4213–4224.
- [24] Y. Jiang, X. Ding, D. Zeng, Y. Huang, J. Paisley, Pan-sharpening with a Hyper-Laplacian Penalty, *International Conference on Computer Vision (ICCV)* (2015) 540–548.
- [25] X. Meng, J. Li, H. Shen, L. Zhang, H. Zhang, Pansharpening with a Guided Filter Based on Three-Layer Decomposition, *Sensors* 16.
- [26] L. J. Deng, G. Vivone, W. H. Guo, M. Dalla Mura, J. Chanussot, A variational pansharpening approach based on reproducible kernel Hilbert space and Heaviside function, *IEEE International Conference on Image Processing (ICIP)*, 2017. To appear, <http://www.escience.cn/people/dengliangjian/paper.dhome>.
- [27] L. J. Deng, G. Vivone, W. H. Guo, M. Dalla Mura, J. Chanussot, A variational pansharpening approach based on reproducible kernel Hilbert space and Heaviside function, *IEEE Transactions on Image Processing* 27 (2018) 4330–4344.
- [28] Z. Y. Zhang, T. Z. Huang, L. J. Deng, J. Huang, X. L. Zhao, C.-C. Zheng, A Framelet-Based Iterative Pan-Sharpening Approach, *Remote Sensing* 10 (2018) 622.
- [29] X. Meng, H. Shen, H. Li, L. Zhang, R. Fu, Review of the pansharpening methods for remote sensing images based on the idea of meta-analysis: Practical discussion and challenges, *Information Fusion* 46 (2019) 102–113.
- [30] N. Yokoya, T. Yairi, A. Iwasaki, Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion, *IEEE Transactions on Geoscience and Remote Sensing* 50 (2012) 528–537.
- [31] R. Restaino, G. Vivone, M. Dalla Mura, J. Chanussot, Fusion of multispectral and panchromatic images based on morphological operators, *IEEE Transactions on Image Processing* 25 (2016) 2882–2895.
- [32] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM Journal on Imaging Sciences* 2 (2009) 183–202.
- [33] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, L.-J. Deng, Y. Wang, A novel tensor-based video rain streaks

- removal approach via utilizing discriminatively intrinsic priors, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2017).
- [34] T.-Y. Ji, T.-Z. Huang, X.-L. Zhao, T.-H. Ma, L.-J. Deng, A non-convex tensor rank approximation for tensor completion, *Applied Mathematical Modelling* 48 (2017) 410–422.
 - [35] T.-X. Jiang, T. Z. Huang, X. L. Zhao, L. J. Deng, A novel nonconvex approach to recover the low-tubal-rank tensor data: when t-SVD meets PSSV, *arXiv preprint arXiv:1712.05870*.
 - [36] T.-X. Jiang, T. Z. Huang, X. L. Zhao, T.-Y. Ji, L. J. Deng, Matrix factorization for low-rank tensor completion using framelet prior, *Information Sciences* 436 (2018) 403–417.
 - [37] H. Fan, Y. Chen, Y. Guo, H. Zhang, G. Kuang, Hyperspectral Image Restoration Using Low-Rank Tensor Recovery, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2017) 1–16.
 - [38] Y. Chang, L. Yan, S. Zhong, Hyper-Laplacian Regularized Unidirectional Low-rank Tensor Recovery for Multispectral Image Denoising, *IEEE Computer Vision and Pattern Recognition (CVPR)*, (2017).
 - [39] M. K. P. Ng, Q. Yuan, L. Yan, J. Sun, An Adaptive Weighted Tensor Completion Method for the Recovery of Remote Sensing Images With Missing Data, *IEEE Transactions on Geoscience and Remote Sensing* 55 (2017) 3367–3381.
 - [40] Q. Yuan, Y. Wei, X. Meng, H. Shen, L. Zhang, A Multiscale and Multidepth Convolutional Neural Network for Remote Sensing Imagery Pan-Sharpening, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11 (2018) 978–989.
 - [41] W. Zuo, D. Meng, L. Zhang, X. Feng, D. Zhang, A generalized iterated shrinkage algorithm for non-convex sparse coding, *IEEE International Conference on Computer Vision (ICCV)* (2013) 217–224.
 - [42] T. G. Kolda, B. W. Bader, Tensor decompositions and applications, *SIAM Review* 51 (3) (2009) 455–500.
 - [43] B. Aiazzi, S. Baronti, M. Selva, L. Alparone, Bi-cubic interpolation for shift-free pan-sharpening, *ISPRS Journal of Photogrammetry and Remote Sensing* 86 (2013) 65–76.
 - [44] B. Aiazzi, S. Baronti, M. Selva, Improving component substitution pansharpening through multivariate regression of MS+PAN data, *IEEE Transactions on Geoscience and Remote Sensing* 45 (10) (2007) 3230–3239.
 - [45] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D* 60 (1992) 259–268.
 - [46] X. L. Zhao, F. Wang, M. K. Ng, A new convex optimization model for multiplicative noise and blur removal, *SIAM Journal on Imaging Sciences* 7 (2014) 456–475.
 - [47] A. Chambolle, An algorithm for total variation minimization and applications, *Journal of Mathematical Imaging and Vision* 20 (2004) 89–97.
 - [48] M. Lysaker, X. C. Tai, Iterative Image Restoration Combining Total Variation Minimization and a Second-Order Functional, *International Journal of Computer Vision* 66 (2006) 5–18.
 - [49] C. Wu, X. C. Tai, Augmented Lagrangian Method, Dual Methods, and Split Bregman Iteration for ROF, Vectorial TV, and High Order Models, *SIAM Journal on Imaging Science* 3 (2010) 300–339.
 - [50] F. Palsson, J. R. Sveinsson, M. O. Ulfarsson, A New Pansharpening Algorithm Based on Total Variation, *IEEE Geoscience and Remote Sensing Letters* 11 (2014) 318–322.
 - [51] B. He, M. Tao, X. Yuan, Alternating Direction Method with Gaussian Back Substitution for Separable Convex Programming, *SIAM Journal on Optimization* 22 (2012) 313–340.
 - [52] X. L. Zhao, F. Wang, T. Z. Huang, M. K. Ng, R. Plemmons, Deblurring and sparse unmixing for hyperspectral images, *IEEE Transactions on Geoscience and Remote Sensing* 51 (2013) 4045–4058.
 - [53] A. Chambolle, T. Pock, A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging, *Journal of Mathematical Imaging and Vision* 40 (2011) 120–145.
 - [54] L. J. Deng, W. Guo, T. Z. Huang, Single image super-resolution by approximate Heaviside functions, *Information Sciences* 348 (2016) 107–123.
 - [55] L. J. Deng, W. Guo, T. Z. Huang, Single image super-resolution via an iterative reproducing kernel Hilbert space method, *IEEE Transactions on Circuits and Systems for Video Technology* 26 (2016) 2001–2014.

- [56] R. Chartrand, Exact reconstruction of sparse signals via nonconvex minimization, *IEEE Signal Processing Letters* 14 (2007) 707–710.
- [57] D. Krishnan, R. a. Fergus, Fast image deconvolution using hyper-Laplacian priors, *Advances in Neural Information Processing Systems (NIPS)* (2009) 1033–1041.
- [58] Q. Lyu, Z. Lin, Y. She, C. Zhang, A comparison of typical ℓ_p minimization algorithms, *Neurocomputing* 119 (2013) 413–424.
- [59] D. L. Donoho, De-noising by soft-thresholding, *IEEE Transactions on Information Theory* 41 (1995) 613–627.
- [60] M. Ghahremani, H. Ghassemian, Nonlinear IHS: A Promising Method for Pan-Sharpening, *IEEE Geoscience and Remote Sensing Letters* 13 (2016) 1606–1610.
- [61] Z. Wang, A. C. Bovik, Modern image quality assessment, *Synthesis Lectures on Image, Video, and Multimedia Processing* 2 (1) (2006) 1–156.
- [62] L. Wald, Data Fusion: Definitions and Architectures Fusion of Images of Different Spatial Resolutions, Paris, France: Les Presses de l’Ecole des Mines, (2002).
- [63] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, M. Selva, Multispectral and panchromatic data fusion assessment without reference, *Photogrammetric Engineering & Remote Sensing* 74 (2008) 193200.
- [64] M. Choi, A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter, *IEEE Transactions on Geoscience and Remote Sensing* 44 (2006) 1672–1682.
- [65] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, *IEEE Trans. Image Processing* 13 (2004) 600–612.
- [66] Z. Liu, E. Blasch, V. John, Statistical comparison of image fusion algorithms: Recommendations, *Information Fusion* 33 (2017) 251–260.
- [67] H. Shen, X. Meng, L. Zhang, An Integrated Framework for the SpatioTemporalSpectral Fusion of Remote Sensing Images, *IEEE Transactions on Geoscience and Remote Sensing* 54 (2016) 7135–7148.