Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Gradient transferred pansharpening method based on cosparse analysis model

Chang Han
Nong Sang
Hongyan Zhang
Liangpei Zhang

Chang Han, Nong Sang, Hongyan Zhang, Liangpei Zhang, "Gradient transferred pansharpening method based on cosparse analysis model," *J. Appl. Remote Sens.* **11**(2), 025009 (2017), doi: 10.1117/1.JRS.11.025009.

Gradient transferred pansharpening method based on cosparse analysis model

Chang Han,[a,b,c] **Nong Sang,**[a,*] **Hongyan Zhang,**[c] **and Liangpei Zhang**[c]
[a]Huazhong University of Science and Technology, School of Automation, National Key Laboratory of Science and Technology on Multispectral Information Processing, Wuhan, China
[b]Wuhan Business University, Department of Mechatronics Engineering, Wuhan, China
[c]Wuhan University, The State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan, China

Abstract. The remote sensing image pansharpening problem under cosparse analysis framework is addressed. To preserve the spatial information of the high-resolution (HR) panchromatic (PAN) image, a gradient transfer strategy is proposed by introducing a gradient consistency constraint to the cosparse analysis-based remote sensing image pansharpening model. Thus, by learning the image gradient information from the HR PAN image, the spatial details of the fused image can be effectively enhanced. In the proposed method, to save running time, the cosparse analysis operator is trained offline in advance with a set of training samples. Both simulated and full-scale, real-data experiments were conducted, and the experimental results confirm that the proposed method outperforms the state-of-the-art remote sensing image fusion methods, in terms of both the visual evaluation and quantitative measurements. *© 2017 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JRS.11.025009]

Keywords: gradient transfer; pansharpening; remote sensing; cosparse analysis model.

Paper 16536 received Jul. 16, 2016; accepted for publication Apr. 18, 2017; published online May 11, 2017.

1 Introduction

Remote sensing images with high spatial and spectral resolutions have been found to be able to enhance the performance in many computer vision tasks,[1,2] such as classification, change detection, map updating, disaster monitoring, and so on. However, most of the earth observation satellites have two separate channels: a high-resolution (HR) panchromatic (PAN) channel and several multispectral (MS) channels with a lower spatial resolution (LR). Unfortunately, this fact is impeding their widespread application. In this regard, a feasible approach is to take advantage of the complementary spatial/spectral resolution characteristics of the two channels to obtain HR MS remote sensing images, with the HR details being extracted from the PAN image and then injected into the MS bands. This data fusion technique is also known as "image fusion."

During the last decades, a large number of image fusion methods have been developed. According to the most common classification system, these image fusion approaches can be divided into three categories: (1) component substitution (CS), (2) multiresolution analysis (MRA), and (3) Bayesian estimation theory-based methods.[2] CS methods include intensity hue saturation transformation,[3] principal component analysis,[4] the Gram–Schmidt transform-based methods,[5] and so on, which are usually effective in rendering the spatial details in the fused image. However, spectral distortion may be produced in these approaches because of the local dissimilarities between the PAN and MS images originating from the spectral mismatch between the two channels of the optical sensors. Moreover, band-dependent spatial detail (BDSD) with local parameter estimation,[6,7] a data-dependent self-adaptive approach, is an adaptive CS method. Recently, Restaino et al.[8] proposed a new context-adaptive approach in which the detail injection is performed through injection coefficients, whose values are estimated over

---

*Address all correspondence to: Nong Sang, E-mail: nsang@hust.edu.cn image segments achieved through a binary partition tree segmentation algorithm. In Refs. 9–13, MRA-based methods were exploited, where the high-frequency channels of the PAN image are added to the MS image to provide a sharpened MS image. For instance, à trous wavelet transform-based pansharpening (AWLP) utilizes shift-invariant redundant transforms to obtain the HR MS images.[14,15] Aiazzi et al.[16] investigated the behavior of CS and MRA, the two main classes of pansharpening methods, in the presence of temporal and/or instrumental misalignments between the MS and PAN data sets. In recent years, a new image fusion branch based on Bayesian estimation theory has attracted the attention of more and more researchers.[17–21] This approach reconstructs the (unknown) HR image from its coarse measurements depending on penalization terms and sparse representation or compressive sensing theory. Li and Yang[17] were the first to address the image fusion problem from the perspective of sparse representation. Zhu and Blamer[22] proposed sparse fusion of images (SparseFI), which explores the same sparse coefficient vector of the corresponding HR/LR MS image patches over the coupled HR/LR dictionaries to construct the HR MS images. Jiang et al.[23] proposed a two-step sparse coding approach with patch normalization (PN-TSSC) for image fusion. Following this, Zhu et al.[24] proposed a sophisticated sparse image fusion algorithm that is based on the SparseFI method, named "jointly sparse fusion of image" (J-SparseFI). The main contribution of this method is that the band correlation is taken into consideration. The J-SparseFI algorithm overcomes the mismatch problem of the coupled dictionary because the spectral response of the PAN image cannot fully cover the range of the MS image spectral response.

More recently, as a new sparse representation-based approach, the cosparse analysis model has been attracting researchers' attention.[25–27] In the approach proposed in our previous work,[27] the analysis operator for each band is trained with the geometric analysis operator learning (GOAL) algorithm, offering advanced performance compared with the state-of-the-art image fusion methods. However, the GOAL image fusion (GOAL-IF) method adopts an online dictionary self-training approach and image patch strategy, which takes much more time than the traditional methods.

The contributions of this paper are threefold. First, the paper addresses the remote sensing image fusion problem from the sparse analysis regularization aspect. Under the cosparse analysis framework, a gradient transfer strategy is proposed to preserve the spatial information. Second, the analysis operator is offline trained in advance with a set of training samples. That extends the application potential of the method. Last, the proposed method greatly reduces the computational burden compared with the other sparse representation-based methods.

2 Gradient Transferred Pansharpening Method Based on Cosparse Analysis Model

2.1 *Pansharpening Problem and Analysis Sparsity Priors*

In the proposed approach, we treat the problem of image fusion as a linear inverse problem. Basically, a strategy based on image patches is adopted, and the image patch size of the HR MS image is $\sqrt{m} \times \sqrt{m} \times B$ (where B is the number of bands). However, most of the patch-based methods lead to poor results as they do not take into account the global support of the whole image during the reconstruction process. Hence, we consider the correlation of the whole image in the patch-based strategy during the image reconstruction stage. The goal of the image fusion is to reconstruct an HR MS image $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r) \in \Re^{p\times r}$ (where $p = Bm$ and $\mathbf{x}_r \in \Re^{p\times 1}$ is a columned image patch), which can be divided into r image patches from a set of measurements $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_1) \in \Re^{q\times 1}$ (where $q = B\frac{m}{\rho^2} + m$, $\mathbf{y}_r \in \Re^{q\times 1}$ is the columned observation, and ρ is the spatial resolution ratio between the PAN image and the LR MS image) that are possibly corrupted by noise, i.e., an LR MS image and a PAN image, with $q < p$

$$\mathbf{Y} = \mathbf{M}\mathbf{X} + \nu, \tag{1}$$

where $\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{\mathrm{MS}} \\ \beta\mathbf{Y}^{\mathrm{PAN}} \end{bmatrix} \in \Re^{q\times r}$ are the observations, which consist of $\mathbf{Y}_{MS}$ LR MS images and $\mathbf{Y}^{\mathrm{PAN}}$ PAN images. $\mathbf{M} \in \Re^{q\times p}$ is the observation matrix modeling the modulation transfer function (MTF)-shaped blur filter and the decimation operator for the HR MS image, as well as the spectral response for the PAN image. The factor β is used as a trade-off parameter to balance the relative contribution of the LR MS image and the HR PAN image to the final result of the image fusion.

Consequently, reconstructing $\mathbf{X}$ in Eq. (1) is highly ill-posed. Using additional information about the signal's structure can help to tackle this linear inverse problem. One prior assumption that has proved useful is that the signals of interest allow a sparse representation. The vectors, whose entries are equal to zero or have a sufficiently small magnitude, are called "sparse." The cosparse analysis model[28,29] assumes that applying a suitable analysis operator $\mathbf{\Omega} \in \Re^{n\times p}$ with $n \geq p$ to a signal $\mathbf{X} \in \Re^{p\times r}$ results in a sparse matrix $\mathbf{\Omega X} \in \Re^{n\times r}$, which consists of sparse vectors. The simplest sparsity measurement is using the ℓ_0-pseudonorm $\|\mathbf{\Omega X}\|_0$. Therefore, the analysis sparse model can be exploited to tackle the linear inverse problem by solving

$$\arg \min_{\mathbf{X}\in\mathbb{R}^{p\times r}} \|\mathbf{\Omega X}\|_0 \quad \text{subject to } \|\mathbf{MX}-\mathbf{Y}\|_2^2 \leq \nu. \tag{2}$$

Clearly, the above equation is a regularization problem imposing a sparse analysis prior.

However, minimizing Eq. (2) with sparse analysis regularization is known to be nondeterministic polynomial-hard. Several workarounds have been proposed to alleviate this problem. One feasible approach is convex relaxation, which involves replacing the ℓ_0-pseudonorm with the ℓ_1-norm. Therefore, any solution $\mathbf{X}$ of Eq. (2) using ℓ_1-regularization can be written as

$$\arg \min_{\mathbf{X}\in\mathbb{R}^{p\times r}} \|\mathbf{\Omega X}\|_1 \quad \text{subject to } \|\mathbf{MX}-\mathbf{Y}\|_2^2 \leq \nu. \tag{3}$$

The most crucial factor for the success of the analysis approach is the choice of an appropriate analysis operator. However, an analysis operator that is learned from a set of training samples is known to perform better.

2.2 *Gradient Transfer for Remote Sensing Images*

For the pansharpening task, the goal is to generate a fused image that simultaneously preserves the spectral radiation information and the appearance information in the two images. Moreover, the PAN image is typically characterized by the pixel intensities. To fuse the detailed spatial information, a straightforward scenario is to constrain the fused image to have a similar pixel intensity distribution to the PAN image. However, the intensity of a pixel in the same physical location may be different for the MS image and the PAN image as they differ significantly in resolution.

The MS and PAN images may also present some local dissimilarities, such as object occultation, contrast inversion, or moving objects, due to the different spectral bands of the sensors or the different acquisition times. These effects are due to environmental physics.[30] However, the main details in the scene can be essentially characterized by the gradients in the image. Therefore, we suggest that the pansharpened image will have a pixel gradient distribution that is similar to that of the PAN image, and the detailed spatial information of the PAN image can be transferred into each band of the pansharpened image. Therefore, the gradient consistency constraint can be rewritten as follows:

$$\arg \min_{\mathbf{X}^b\in\mathbb{R}^{p\times r}} \|\nabla X^b - \nabla Y^{\mathrm{PAN}}\|, \tag{4}$$

where $\nabla\mathbf{Y}^{\mathrm{PAN}}$ and ∇X^b represent the gradients of the PAN and b'th band images, respectively. The HR PAN image can be approximated as a linear combination of the desired HR MS image.[17] The linear model can be modeled as

$$\mathbf{Y}^{\mathrm{PAN}} = \sum_b \theta_b X^b + \mathbf{v}_2 = \mathbf{M}_2 X + \mathbf{v}_2, \tag{5}$$

where Y^{PAN} and X^b represent the PAN image and the b'th band of the HR MS images, respectively, and θ_b is the weight of the b'th band. $\mathbf{M}_2$ is the linear combination matrix, which denotes the spectral response of the different remote sensors. $\mathbf{v}_2$ is the additive zero-mean Gaussian noise.[17] The parameters θ_b and $\mathbf{M}_2$ are determined by the specific satellite sensors. Since the pixel gradient is similar between the PAN image and each band, the formula can be expanded as

$$\begin{aligned}\mathbf{Y}^{\mathrm{PAN}} &= \sum_b \theta_b X^b + \mathbf{v}_2\\ &= \theta_1 X^1 + \theta_2 X^2 + \theta_3 X^3 + \theta_4 X^4 + \mathbf{v}_2.\end{aligned} \tag{6}$$

To obtain the gradient of the PAN image, we obtain the derivative for the above equation

$$\nabla \mathbf{Y}^{\mathrm{PAN}} I_1 = \theta_1 \nabla X^1 \quad \nabla \mathbf{Y}^{\mathrm{PAN}} I_2 = \theta_2 \nabla X^2 \quad \nabla \mathbf{Y}^{\mathrm{PAN}} I_3 = \theta_3 \nabla X^3 \quad \nabla \mathbf{Y}^{\mathrm{PAN}} I_4 = \theta_4 \nabla X^4, \tag{7}$$

and then the above equation can be combined as

$$\nabla Y^{\mathrm{PAN}}[I_1, I_2, I_3, I_4] = [\alpha_1, \alpha_1, \alpha_1, \alpha_1] \begin{bmatrix} \nabla X^1 \\ \nabla X^2 \\ \nabla X^3 \\ \nabla X^4 \end{bmatrix}. \tag{8}$$

Supposing $\nabla X = \begin{bmatrix} \nabla X^1 \\ \nabla X^2 \\ \nabla X^3 \\ \nabla X^4 \end{bmatrix}$, $\mathbf{M}_3 = [I_1, I_2, I_3, I_4] \in \Re^{64\times 256}$, $\mathbf{M}_2[\alpha_1, \alpha_2, \alpha_3, \alpha_4]$, and $I_1 = I_2 = I_3 = I_4 \in \Re^{64\times 64}$, then the above equation can be written as

$$\mathbf{M}_3 \nabla Y^{\mathrm{PAN}} = \mathbf{M}_2 \nabla X. \tag{9}$$

This shows the relationship between the gradient of the PAN image and the gradient of the HR MS image. The function of the gradient transformation in Eq. (4) is revised as follows:

$$\arg \min_{\mathbf{X}\in\mathbb{R}^{p\times r}} (\mathrm{M}_3^{\dagger}\mathrm{M}_2 - 1)\|\nabla X\|, \tag{10}$$

where $M_3^{\dagger}$ represents the pseudoinverse of the M_3 matrix. Clearly, it is based on the gradient consistency constraint and total variation minimization.

2.3 *Gradient Transferred Pansharpening Model Based on Cosparse Analysis Model*

Considering the sparse structure of the LR MS image and PAN image, as well as the analysis operator $\mathbf{\Omega}$, which is introduced in the next part, the image fusion model can be formulated as follows:

$$\arg \min_{\mathbf{X}} \|\mathbf{MX} - \mathbf{Y}\|_2^2 + \frac{\lambda}{2}\|\mathbf{\Omega X}\|_1 + \lambda_1 \cdot (\mathrm{M}_3^{\dagger}\mathrm{M}_2 - 1) \cdot J(\mathbf{X}), \tag{11}$$

with $J(\mathbf{X}) = |\nabla \mathbf{X}| = \sum \sqrt{(\nabla_i^h \mathbf{X})^2 + (\nabla_i^v \mathbf{X})^2}$. This can be solved with the alternating direction method of multipliers (ADMM). Assuming $Z = \mathbf{\Omega X}$, then $\mathbf{X}$ is the image consisting of vectored image patches, and the problem is a constrained convex minimization problem with two parameters Z and X, as follows:

$$\arg \min_{\mathbf{X},Z} \|\mathbf{MX} - \mathbf{Y}\|_2^2 + \frac{\lambda}{2}\|Z\|_1 + \lambda_1 \cdot (\mathrm{M}_3^{\dagger}\mathrm{M}_2 - 1) \cdot J(\mathbf{X}) \quad \text{s.t. } Z = \mathbf{\Omega X}. \tag{12}$$

The advantage is that this method can result in global optimization for the whole image rather than local optimization. In the Lagrangian multiplier method, we use the dual parameter B and add a penalty term $\langle B, \Omega\mathbf{X} - Z\rangle$. The new objective $\Psi(\mathbf{X}, Z, \mathbf{B})$ is as follows:

Algorithm 1 Pansharpening method based on the ADMM.

```
Input: $\Omega$, $K_{\max}$, $\mathbf{X}_0$, $\gamma$, $\lambda$, $\varepsilon = 0.001$
1. Initialization: $k = 1$, $\mathbf{X}_k = \mathbf{X}_0$, $\mathbf{B}_k = 0$, $Z_k = \Omega\mathbf{X}_k$.
2. while $\varepsilon \le \|\mathbf{X}_k - \mathbf{X}_{k-1}\|_F$ and $k \le K_{\max}$
3. $\mathbf{X}_{k+1} = (\lambda\mathbf{M}^T\mathbf{M} + \gamma\Omega^T\Omega)^{-1}[\lambda\mathbf{M}^T\mathbf{Y} + \gamma\Omega^T(Z_k - \mathbf{B}_k) - \lambda_1 \cdot (\mathrm{M}_3^{?}\mathrm{M}_2 - 1) \cdot \frac{\partial J(X)}{\partial X}]$;
4. $Z_{k+1} = \mathbf{S}_{\frac{1}{\gamma}}\{\Omega\mathbf{X}_{k+1} + \mathbf{B}_k\}$
5. $\mathbf{B}_{k+1} = \mathbf{B}_k - (Z_{k+1} - \Omega\mathbf{X}_{k+1})$
6. $k = k + 1$
7. end while
8. output: $\mathbf{X} = \mathbf{X}_{k+1}$
```

$$\begin{aligned}\psi(\mathbf{X}, Z, \mathbf{B}) &= \|\mathbf{M}X - Y\|_2^2 + \frac{\lambda}{2}\|Z\|_1 + \lambda_1 \cdot (M_3^{\dagger}M_2 - 1) \cdot J(X) + \gamma\langle B, \Omega X - Z\rangle + \frac{\gamma}{2}\|\Omega\mathbf{X} - Z\|_2^2 \\ &= \|\mathbf{M}X - Y\|_2^2 + \frac{\lambda}{2}\|Z\|_1 + \lambda_1 \cdot (M_3^{\dagger}M_2 - 1) \cdot J(X) - \frac{\gamma}{2}\|B\|_2^2 + \frac{\gamma}{2}\|B + \Omega\mathbf{X} - Z\|_2^2,\end{aligned} \tag{13}$$

where $\Psi(\mathbf{X}, Z, \mathbf{B})$ is convex with respect to Z and X and concave with respect to B. We can thus iteratively update each of the parameters, while keeping the others fixed. The process can be summarized as shown in Algorithm 1.

In this algorithm, the relative variation $\varepsilon = 0.001$ of the estimated fusion images in two consecutive iterations is set as the stopping criterion. $\mathbf{S}_{\frac{1}{\gamma}}$ in line 4, with $\frac{1}{\gamma} > 0$, is the entrywise soft-thresholding operator defined by $\mathbf{S}_{\frac{1}{\gamma}}(\alpha) = \alpha - \frac{1}{\gamma}\mathrm{sgn}(\alpha)$ if $|\alpha| \ge \frac{1}{\gamma}$ and is 0 otherwise.

2.4 *Analysis Operator Learning*

The image fusion framework employed in this work consists of the two separate stages of analysis learning and image fusion. Offline analysis operator learning is undertaken in advance with a set of training samples from the MS and PAN images, as shown in Fig. 1.

Examples of the training samples for the analysis operator learning are shown in Fig. 1. In this data set, two pairs of available IKONOS images and two pairs of QuickBird images with MS and PAN channels are available. The MS images and PAN images have the sizes of $256 \times 256 \times 4$ and 1024×1024 pixels, respectively. To obtain the prominent features of the remote sensing images, the training samples are dealt with using an image patch strategy. Image patches of size

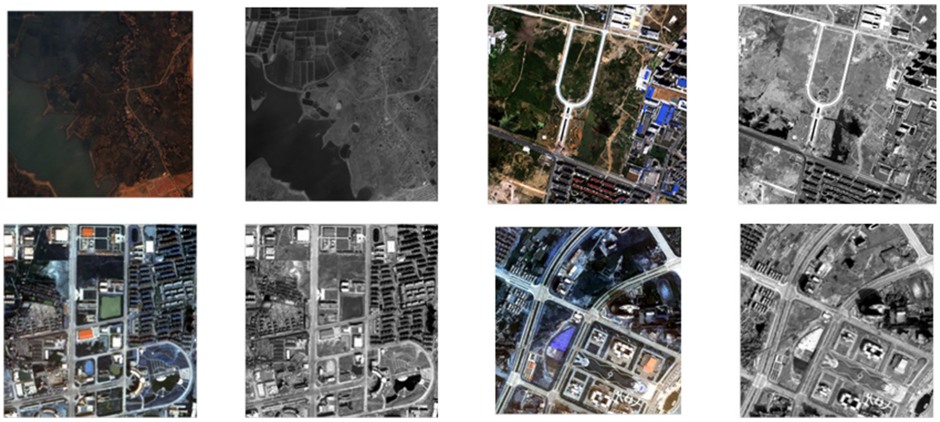


Fig. 1 The training samples for analysis operator learning.

$\sqrt{m} \times \sqrt{m} \times B$ are cropped from the LR MS data set. Therefore, each MS image cell has the size of $\sqrt{m} \times \sqrt{m} \times B$. Each PAN image cell consisting of B PAN image patches with the size of $\sqrt{m} \times \sqrt{m}$ is then columned as training samples, with $\mathbf{s}_i \in \Re^p$ being the i'th training sample. These training samples are then selected randomly as the training sample set $\mathbf{S} \in \Re^{p\times N}$ to learn an analysis operator $\mathbf{\Omega} \in \Re^{n\times p}$. N represents the number of training samples. Here, we set $m = 64$, $B = 4$, $N = 10000$, and $p = 256$.

The goal of analysis operator learning is to find the suitable matrix $\mathbf{\Omega} \in \Re^{n\times p}$ with $n \geq p$, which leads to the analyzed vector $\mathbf{\Omega}\mathrm{s}_i$ being as sparse as possible for the training sample $\mathbf{s}_i$. Therefore, the analysis operator for a given set of image patches $\{\mathbf{s}_i \in \Re^{256\times 1}\}_{i=1}^N$ is formally denoted as

$$\mathbf{\Omega} \in \arg\ \min \sum \|\Omega s_i\|_0. \tag{14}$$

Clearly, this is nonconvex and discontinuous. One tractable approach is to replace it with a smooth log-square function[29]

$$\mathbf{\Omega} \in \arg\ \min \sum_{i}^{N} \sum_{j=1}^{p} \log[1 + \mu(\mathbf{\Omega s})_{i,j}^2]. \tag{15}$$

Furthermore, to avoid a trivial solution in Eq. (14), we impose certain constraints on $\mathbf{\Omega}$. We enforce the full-rank constraint with the below equation to the analysis operator $\mathbf{\Omega}$

$$h(\mathbf{\Omega}) = -\frac{1}{p\ \log(p)} \log\ \det\left(\frac{1}{n}\mathbf{\Omega}^T\mathbf{\Omega}\right). \tag{16}$$

The mutual coherence of the analysis operator can be controlled via the below equation

$$r(\mathbf{\Omega}) = -\sum_{1\leq i\leq j\leq n} \log[1 - (\boldsymbol{\omega}_{\mathbf{i},:}^{\mathbf{T}} \boldsymbol{\omega}_{\mathbf{j},:})^2], \tag{17}$$

where $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_j$ represent the transpose of the i'th row and j'th row of $\mathbf{\Omega}$, respectively. According to Ref. 29, the cost function of the analysis operator learning can be expressed as

$$\mathbf{\Omega} := \underset{\mathbf{\Omega}^T \in \mathbf{OB}(p,n)}{\arg\ \min} \sum_{i}^{N} \mathrm{g}(\mathbf{\Omega s}_i) + k_1 \cdot h(\mathbf{\Omega}) + \mu_1 \cdot r(\mathbf{\Omega}). \tag{18}$$

The conjugate gradient on the oblique manifold $\mathbf{OB}(p, n)$ is employed to solve the optimization problem [Eq. (18)]. The parameters k_1 and μ_1 influence the condition. For more details, we refer the readers to Ref. 29.

3 Experiments and Discussion

3.1 *Experimental Setting*

The results are evaluated both visually and quantitatively. For the quantitative evaluation, the following typical evaluation indices based on Wald's protocol are utilized: the correlation coefficient (CC),[31] the structural similarity metric (SSIM),[32] the root-mean-square error (RMSE), the spectral angle mapper (SAM),[31] the *erreur relative globale adimensionnelle de synthèse* (ERGAS),[33] and the Q_4 quality index.[32,34] The Q_4 index is calculated over a window of M-by-M, which is normally selected as $M = 16$ or $M = 32$. Q_4 is averaged over the whole image to obtain a global quality metric. Q_4 is in the range [0, 1], where 1 represents the ideal fusion, i.e., the fused and the reference images are identical.[34] In this paper, the block size was set as 32×32 for all the methods. To obtain the simulated data, MTF-tailored low-pass filters were used to produce the LR MS and PAN images. The Nyquist cutoff frequencies of QuickBird and IKONOS for the different spectral bands are provided in Refs. 18 and 35. The original MS images were selected as the real HR images to compare with the fused images.

To make the comparison more convincing, we chose two top-performing benchmarks, BDSD and AWLP, from the pansharpening toolbox in Ref. 2. PN-TSSC[23] and J-SparseFI[24] were also used as comparative methods. The parameters for the other methods were set as recommended. The optimal size of the image patch for the LR MS image was set as 7×7, and the overlapping area size was set as 7×4 for the PN-TSSC method. The regularization parameter for PN-TSSC was set as 8192. Meanwhile, the image patch size of the LR image was set to 5×5, which corresponds to 20×20 for the PAN image patch. For the GOAL-IF method, the parameters were set as suggested in Ref. 27.

For the proposed method, k_1 and μ_1 in Eq. (18) were empirically set as $k_1 = 40$ and $\mu_1 = 1000$ to obtain the optimal analysis operator. The value of β in Eq. (1) was set as $\beta = 0.14$ for

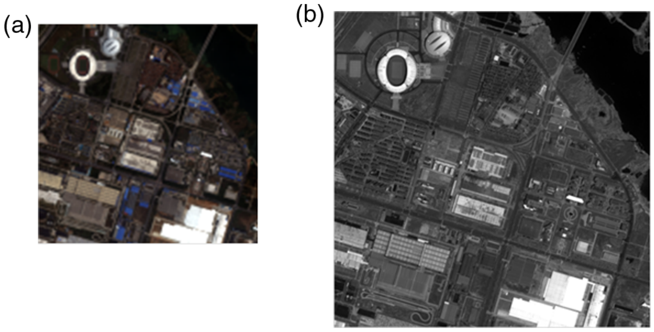


Fig. 2 IKONOS images: (a) degraded MS image at a 16-m spatial resolution, and (b) degraded PAN image at a 4-m spatial resolution.

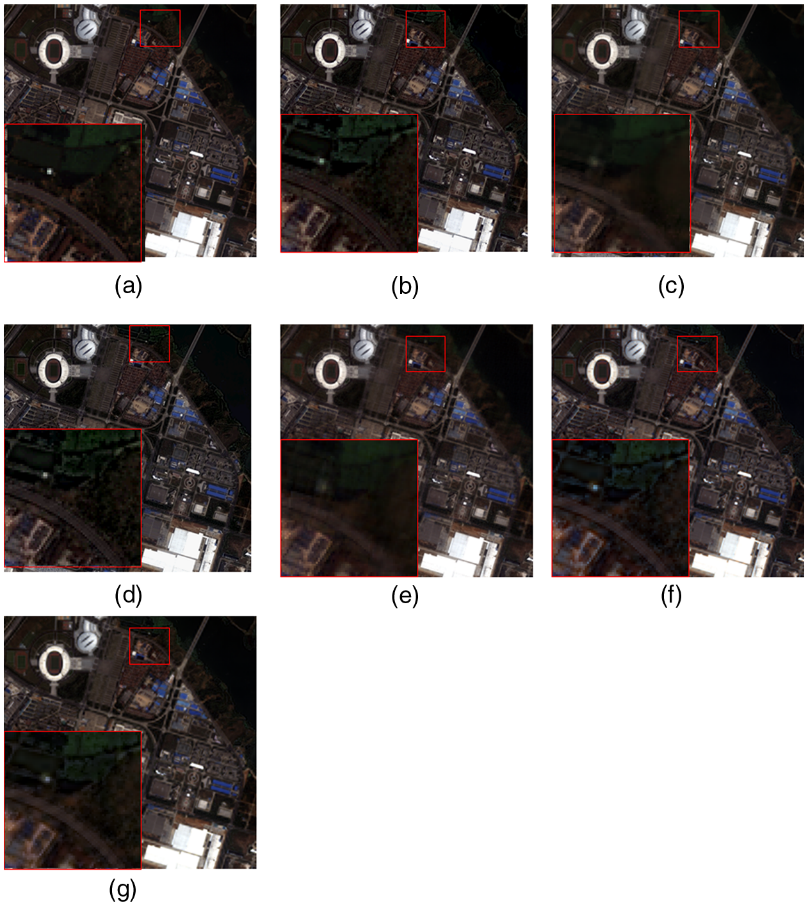


Fig. 3 Results of the image fusion with the degraded IKONOS image: (a) original, (b) AWLP, (c) PN-TSSC, (d) BDSD, (e) J-SparseFI, (f) GOAL-IF, and (g) the proposed method.

the IKONOS data set and $\beta = 0.12$ for the QuickBird and Pléiades data sets. We set $\lambda = 7 \times 10^9$ and $\lambda_1 = 0.31$ in Eq. (11).

3.2 *IKONOS Data Experiment*

The IKONOS system simultaneously offers a four-band MS image with a 4-m resolution and a single-band PAN image with a 1-m resolution. In this study, we utilized a simulated IKONOS LR MS image with a resolution of 16 m and a PAN image with a resolution of 4 m, as shown in Fig. 2. In this simulated experiment, the LR MS image size was $128 \times 128 \times 4$ pixels, and the corresponding PAN image size was 512×512 pixels.

It can be clearly observed from Fig. 3 that, in the pansharpening results obtained by PN-TSSC and J-SparseFI, the two traditional sparse reconstruction-based algorithms, some of the spatial details are missing and the results appear a little blurry in some areas. The result obtained by the AWLP method is acceptable with regard to the spectral characteristic. The BDSD method suffers from color distortion, and there are some obvious outlier pixels. Although GOAL-IF and the proposed method obtain more details than the other methods, the proposed method results in less spectral distortion than the GOAL-IF method. The quantitative assessment results are shown in Table 1, where the evaluation results, in terms of CC, RMSE, ERGAS, SSIM, and Q_4, indicate that the proposed method achieves a better fusion result than the other methods.

3.3 *Pléiades Data Experiment*

To make the experiments more convincing, an experiment with the Pléiades data set was also implemented. This data set was used for the 2006 GRS-S data fusion contest[36] and was provided

Table 1 Quantitative assessment results of the simulated experiment shown in Fig. 3.

		AWLP	BDSD	PN-TSSC	J-SparseFI	GOAL-IF	Proposed
CC	B	0.9581	0.9561	0.9590	0.9597	*0.9604*	**0.9626**
	G	0.9583	0.9593	0.9631	0.9640	*0.9662*	**0.9665**
	R	0.9618	0.9591	0.9626	*0.9639*	0.9566	**0.9683**
	NIR	0.9297	0.9272	0.9303	*0.9334*	0.9317	**0.9376**
	Average	0.9520	0.9504	0.9537	*0.9553*	0.9538	**0.9587**
RMSE	B	70.6731	73.0352	69.6082	68.9156	*68.3721*	**66.6313**
	G	63.4490	62.6783	58.9273	58.0928	*56.4732*	**56.3961**
	R	39.2762	40.9982	38.7256	*37.9625*	42.3549	**35.6994**
	NIR	80.8010	82.8169	80.2390	*78.4048*	79.9656	**76.0650**
	Average	63.5498	64.8821	61.8750	*60.8439*	61.7915	**58.6980**
SSIM	B	0.7673	0.7577	*0.7718*	0.7604	0.7592	**0.7723**
	G	0.7899	0.7919	0.8110	0.8040	*0.8148*	**0.8154**
	R	0.8598	0.8539	*0.8686*	0.8657	0.8442	**0.8797**
	NIR	0.6851	0.6857	0.6861	0.6863	*0.6980*	**0.7048**
	Average	0.7755	0.7722	0.7844	*0.7791*	*0.7791*	**0.7930**
SAM		*2.6884*	2.8834	2.6857	2.7273	2.8332	**2.5933**
ERGAS		3.1144	3.1803	2.6252	*2.5847*	2.6339	**2.5035**
Q_4		0.7198	0.6934	*0.7744*	0.7548	0.7642	**0.7817**

Note: The best results for each quality index are labeled in bold, and the second-best results for each quality index are italicized.

by Centre National d' Etudes Spatiales, the French Space Agency. The image covers an urban area of Toulouse (France), and it has a size of 1024×1024 pixels. The resolution of the four MS bands is 60 cm, and there is no available PAN image, since the corresponding sensor was under development. To obtain the simulated HR PAN data, the green and red channels were averaged, and the result was filtered with a system characterized by the nominal MTF of the PAN sensor. The resolution ratio between the MS and the PAN image was set as 4, and the radiometric resolution was 11 bits. The simulated LR MS image with a resolution of 2.4 m and the PAN image with a resolution of 0.6 m are shown in Fig 4. In this experiment, the LR MS image was $256 \times 256 \times 4$ pixels, and the corresponding PAN image was 1024×1024 pixels.

The results of the different methods are shown in Fig. 5. To facilitate a comparison, details of local regions of the images are exhibited in Fig. 6. By visually comparing the fused images with the original image from Fig. 5, it can be clearly seen that all the methods can effectively fuse the

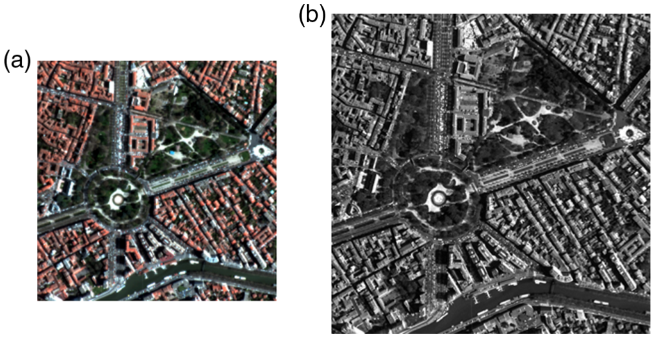


Fig. 4 Pléiades data set: (a) degraded MS image at a 2.4-m spatial resolution and (b) degraded PAN image at a 0.6-m spatial resolution.

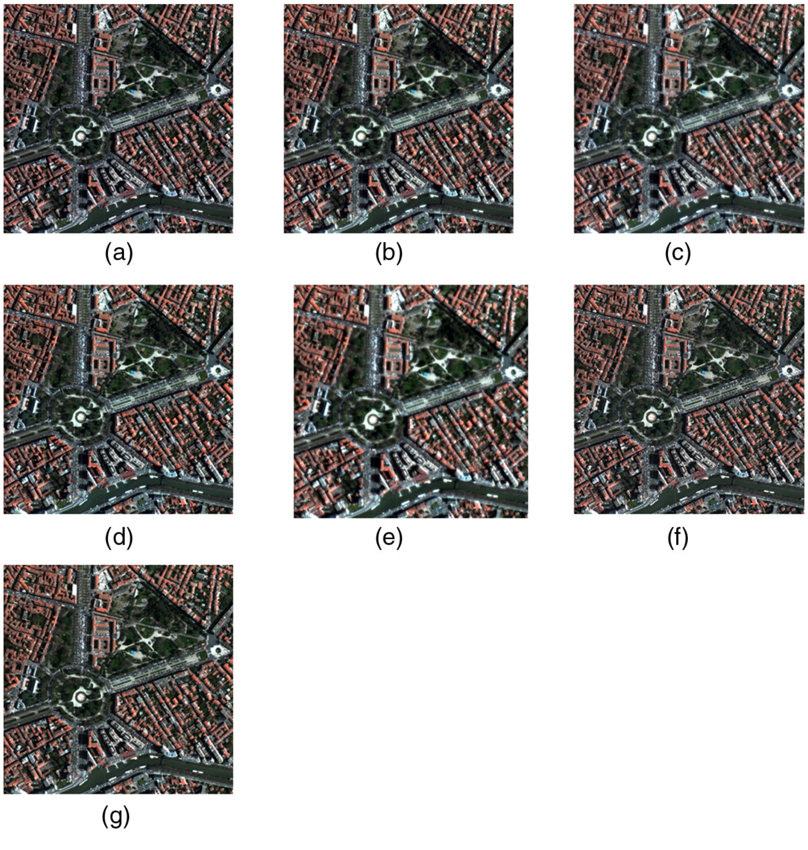


Fig. 5 Results of the image fusion with the simulated Pléiades image: (a) original, (b) AWLP, (c) PN-TSSC, (d) BDSD, (e) J-SparseFI, (f) GOAL-IF, and (g) the proposed method.

LR MS image and the PAN image. All the images have been improved with regard to both the spatial and spectral aspects. However, it can be clearly seen from the local zoomed image that the result obtained by the PN-TSSC method suffers from serious spectral distortion, and the rectangular building marked with the red box almost disappears when compared with the original image. It can also be clearly seen that only a small amount of spatial distortion appears in the image fusion result obtained by the J-SparseFI method. When we observe Figs. 6(b) and 6(d), it can be seen that the right edges of the building are blurred to some degree. Although the building outline can be seen in the result of GOAL-IF in Fig. 6(e), there is some shape distortion. Overall, the image fusion result obtained by the proposed method is the closest to the original image.

The quantitative assessment indices for the fusion results are shown in Table 2, in which the best results for each quality index are labeled in bold. This shows that the proposed method acquires the best evaluation results, in terms of CC, RMSE, and SSIM, which indicates that the fusion result of the proposed method is the most closely correlated to the original MS image. However, the J-SparseFI method obtains a better result for the SAM assessment metric, and BDSD and the proposed method obtain the best results in the ERGAS and Q_4 indices. Overall, the quantitative assessment results agree with the visual evaluation, and the proposed method achieves a better fusion result than the other methods.

3.4 *QuickBird Data Experiment*

The QuickBird data set provides a four-band 2.8-m resolution MS image and a 0.7-m resolution PAN image. The size of the LR MS image in the simulated QuickBird data set experiments was $128 \times 128 \times 4$ pixels, and the corresponding PAN image was 512×512 pixels, with resolutions of 11.2 and 2.8 m, respectively, as shown in Fig. 7.

It can be clearly observed from the area marked with the yellow oval in Fig. 8 that the result obtained by AWLP maintains the spatial information well, but the fused image contains spectral

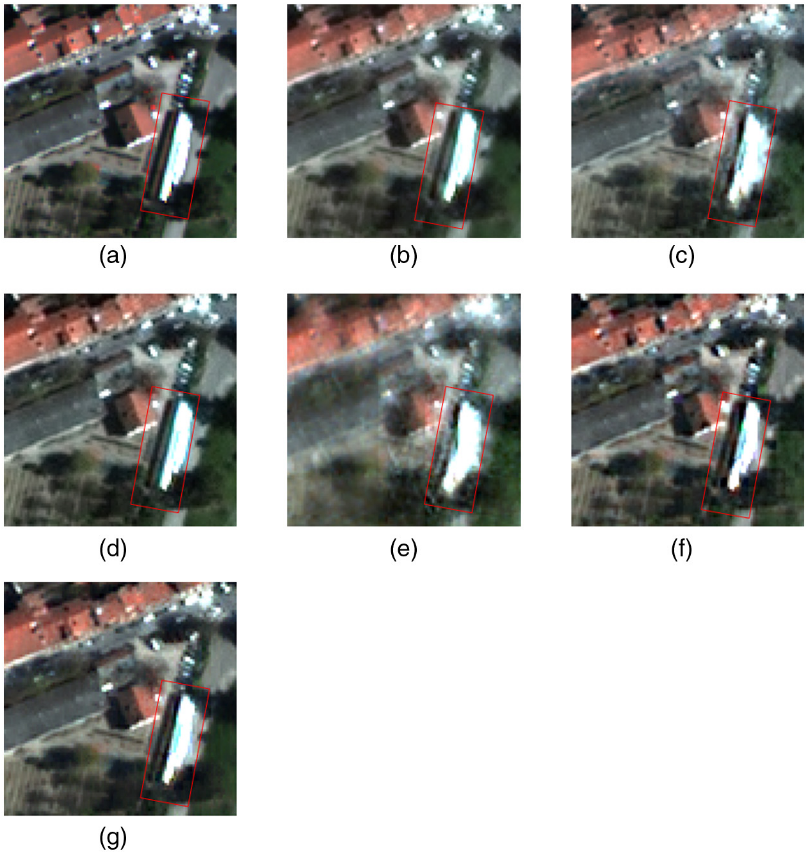

(a) (b) (c) (d) (e) (f) (g)

Fig. 6 Close-ups from Fig. 5: (a) original, (b) AWLP, (c) PN-TSSC, (d) BDSD, (e) J-SparseFI, (f) GOAL-IF, and (g) the proposed method.

Table 2 Quantitative assessment results of the simulated experiment shown in Fig. 5.

		AWLP	BDSD	PN-TSSC	J-SparseFI	GOAL-IF	Proposed
CC	B	0.9311	0.9055	0.9217	0.9282	*0.9624*	**0.9745**
	G	0.9479	0.9034	0.9222	0.9320	*0.9648*	**0.9794**
	R	0.9365	0.9017	0.9283	0.9392	*0.9653*	**0.9819**
	NIR	0.8888	*0.9509*	0.8821	0.9440	0.9386	**0.9522**
	Average	0.9261	0.9154	0.9136	0.9359	*0.9578*	**0.9720**
RMSE	B	21.6809	28.4180	26.3716	*18.6852*	19.3270	**14.8853**
	G	19.2945	30.7077	29.6512	*18.9014*	19.6195	**14.6786**
	R	30.9171	44.2816	39.1592	*25.6498*	26.4099	**19.0255**
	NIR	65.2980	52.7596	75.8721	**39.0127**	54.8312	*48.0537*
	Average	34.2976	39.0417	42.7635	*25.5623*	30.0469	**24.1608**
SSIM	B	0.9439	0.9177	0.9602	0.9681	*0.9693*	**0.9810**
	G	0.9555	0.9083	0.9475	0.9664	*0.9716*	**0.9812**
	R	0.9145	0.8584	0.9240	0.9459	*0.9562*	**0.9735**
	NIR	0.7715	0.8462	0.7690	**0.9025**	0.8643	*0.8836*
	Average	0.8963	0.8826	0.9002	*0.9457*	0.9403	**0.9548**
SAM		4.3356	4.0749	5.3176	**2.5387**	4.0963	*3.4975*
ERGAS		3.5219	**2.7811**	5.3368	3.2093	3.7310	*2.9754*
Q_4		0.9426	**0.9650**	0.8442	0.8899	0.9387	*0.9554*

Note: The best results for each quality index are labeled in bold, and the second-best results for each quality index are italicized.

distortion when compared with the original MS image. Meanwhile, the result obtained by BDSD suffers from the most serious spectral distortion. It can also be observed that PN-TSSC and J-SparseFI produce similar results, with only a few artifacts and a small amount of spectral distortion appearing in the pansharpened image. GOAL-IF and the proposed method produce similar results, but some details are smoothed. However, the local zoomed images show that the fusion result obtained by the proposed method contains more spatial details than the results obtained by the other methods in the runway area marked with the red rectangle. The proposed method performs well in both the spatial and spectral aspects in this figure. The quantitative

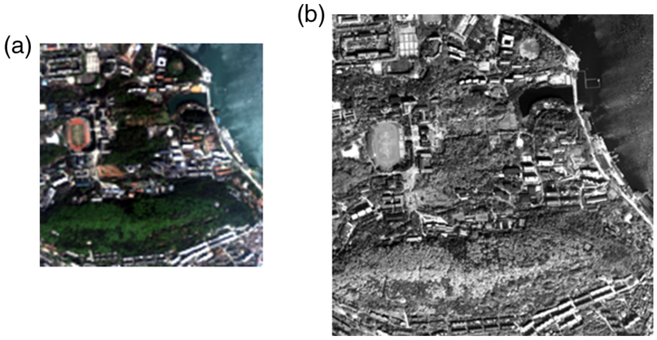


Fig. 7 QuickBird images: (a) degraded MS image at a 11.2-m spatial resolution and (b) degraded PAN image at a 2.8-m spatial resolution.

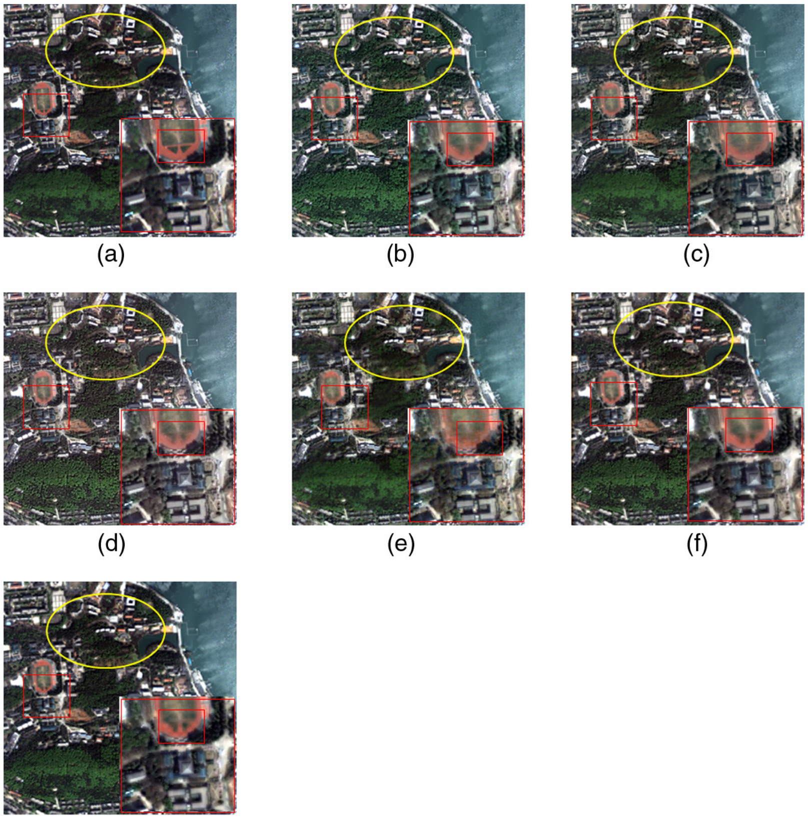

(a) (b) (c) (d) (e) (f) (g)

Fig. 8 Results of the image fusion with the degraded QuickBird image: (a) original, (b) AWLP, (c) PN-TSSC, (d) BDSD, (e) J-SparseFI, (f) GOAL-IF, and (g) the proposed method.

assessment indices for the fusion results are shown in Table 3, confirming that the proposed method obtains better values for most of the selected metrics in this experiment.

3.5 *Full-Scale, Real-Data Experiment*

We also evaluate the proposed method in a full-scale, real-data experiment with an IKONOS image. Figures 9(a) and 9(b) show the original 4-m resolution MS image and the 1-m resolution PAN image, respectively. D_λ represents the spectral distortion and D_s represents the spatial distortion. The quality not requiring a reference (QNR) index[37] is also used to evaluate the fused image.

The results of the real-data experiment are confirmed quantitatively in terms of QNR, D_λ, and D_s. The spectral distortion D_λ is calculated as follows:[37]

$$D_\lambda = \frac{2}{B(B-1)} \sum_{b=1}^{B} \sum_{k=1}^{b-1} |Q(MS_{\mathrm{low},b}, MS_{\mathrm{low},k}) - Q(\widehat{MS_b}, \widehat{MS_k})|, \tag{19}$$

where B is the number of bands of the MS image, $MS_{\mathrm{low},b}$ and $\widehat{MS_b}$ $[b \in (1, \ldots, B)]$ are the b'th low-resolution and sharpened MS bands, respectively, and Q is the universal image quality index. The spatial distortion D_s is calculated as follows:

$$D_s = \frac{1}{B} \sum_{b=1}^{B} |Q(MS_{\mathrm{low},b}, \mathrm{PAN}_{\mathrm{low}}) - Q(\widehat{MS_b}, \mathrm{PAN})|, \tag{20}$$

where PAN and $\mathrm{PAN}_{\mathrm{low}}$ are the full-resolution and downsampled PAN images, respectively. Then, taking into account the spatial and spectral distortion, the QNR can be written as[37]

Table 3 Quantitative assessment results of the simulated experiment shown in Fig. 8.

		AWLP	BDSD	PN-TSSC	J-SparseFI	GOAL-IF	Proposed
CC	B	0.9280	0.9184	0.9239	0.9331	*0.9374*	**0.9447**
	G	0.9169	0.9145	0.9271	0.9318	*0.9385*	**0.9410**
	R	0.9390	0.9158	0.9281	0.9325	*0.9407*	**0.9421**
	NIR	0.9171	0.9190	0.9203	0.9248	*0.9302*	**0.9338**
	Average	0.9253	0.9169	0.9248	0.9306	*0.9367*	**0.9404**
RMSE	B	10.5840	11.6378	10.6183	9.9679	*9.7855*	**9.0170**
	G	23.3733	23.6227	20.3296	19.5810	*18.7901*	**18.2666**
	R	19.4556	24.7230	20.9856	20.2972	*19.1827*	**18.8071**
	NIR	41.0628	40.4951	38.6505	37.3867	*36.1285*	**35.1795**
	Average	23.6189	25.1196	22.6460	21.8082	*20.9717*	**20.3176**
SSIM	B	0.9802	0.9754	0.9829	0.9840	*0.9849*	**0.9866**
	G	0.9311	0.9272	0.9527	0.9565	*0.9589*	**0.9614**
	R	0.9553	0.9192	0.9506	0.9549	*0.9552*	**0.9586**
	NIR	0.8647	0.8653	0.8676	0.8692	*0.8782*	**0.8805**
	Average	0.9328	0.9218	0.9385	0.9412	*0.9443*	**0.9468**
SAM		*2.2300*	2.4252	2.4293	2.4251	**2.2255**	2.2362
ERGAS		2.2532	2.4249	2.1869	2.1117	*2.0410*	**2.0014**
Q_4		0.7923	0.7494	0.8067	0.8016	*0.8091*	**0.8183**

Note: The best results for each quality index are labeled in bold, and the second-best results for each quality index are italicized.

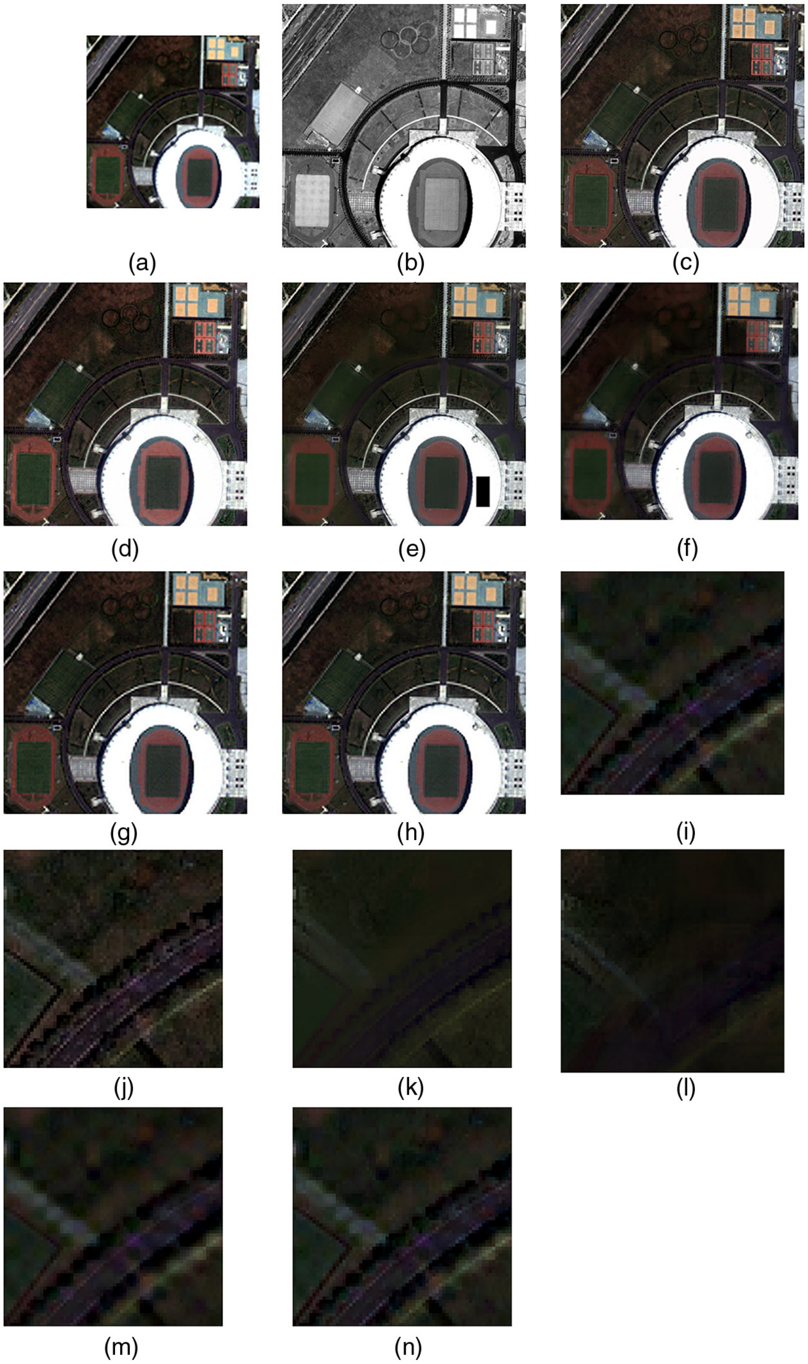
$$\text{QNR} = (1 - D_\lambda)(1 - D_s). \tag{21}$$

The results of AWLP, BDSD, PN-TSSC, and the proposed method are shown in Figs. 9(c)–9(h). Local zoomed images of the results are shown in Figs. 9(i)–9(n). Here, it can be seen that the BDSD method generates spectral distortion, as shown in Figs. 9(d) and 9(j). We can also observe from Fig. 9(e) that the result of the PN-TSSC method has a defect area with incorrect reconstruction. Moreover, the details shown in Fig. 9(l) have been lost, and the objects in the scene are blurred to some degree. The result of the J-SparseFI method is also blurred to some degree in Figs. 9(f) and 9(l). Goal-IF and the proposed method [Figs. 9(g) and 9(h)] generate an HR MS image with satisfactory spectral and spatial preservation.

The quality assessment indices are shown in Table 4. Clearly, the advantage of J-SparseFI, GOAL-IF, and the proposed method is that they obtain very low spatial distortion ($D_s < 0.1$). However, the proposed method generates better D_λ, D_s, and QNR values than PN-TSSC. Overall, the results of the real-data experiment are consistent with the results of the simulated experiment.

3.6 *Parameter Analysis*

In this paper, the fused results are influenced by parameters β, λ, and λ_1 in Eq. (11). Figure 10(a) shows the variation tendency of the SAM/ERGAS indices as parameter β changes. This illustrates that the trade-off between the MS image and PAN image requires an optimal parameter setting. The influence of parameter λ_1 is shown in Fig. 10(b). The tendency of the image fusion



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k) (l) (m) (n)

Fig. 9 The full-scale experiment with a real IKONOS image: (a) original MS image, (b) PAN image, (c) AWLP, (d) BDSD, (e) PN-TSSC, (f) J-SparseFI, (g) GOAL-IF, and (h) the proposed method. Local zoomed images of the results of (i) the AWLP method, (j) the BDSD method, (k) the PN-TSSC method, (l) the J-SparseFI method, (m) the GOAL-IF method, and (n) the proposed method.

performance in terms of the ERGAS and SAM indices demonstrates that the fusion method performs better and better as λ_1 increases. Parameter λ_1 determines the degree of the gradient constraint during the image pansharpening process.

The sensitivity analysis for the input parameter γ of ADMM is shown in Fig. 11. To obtain the tendency of the performance with parameter γ, the other parameters were set as the optimal values. The performance of the image fusion is represented in terms of ERGAS, SAM, and Q_4, respectively. Moreover, $\gamma = 7 \times 10^6$ is the minimum considered in this paper because ADMM will not obtain a solution with less than this value. Therefore, we use $\log \gamma$ to display the trend

Table 4 Quality measures for the different image fusion approaches with real IKONOS data.

	AWLP	BDSD	PN-TSSC	J-SparseFI	GOAL-IF	Proposed
D_λ	0.1408	0.1037	NAN	0.0488	0.0682	**0.0464**
D_s	0.2045	0.1639	NAN	0.1040	0.0781	**0.0654**
QNR	0.6835	0.7497	NAN	0.8523	0.8590	**0.8913**

Note: The best results for each quality index are labeled in bold.

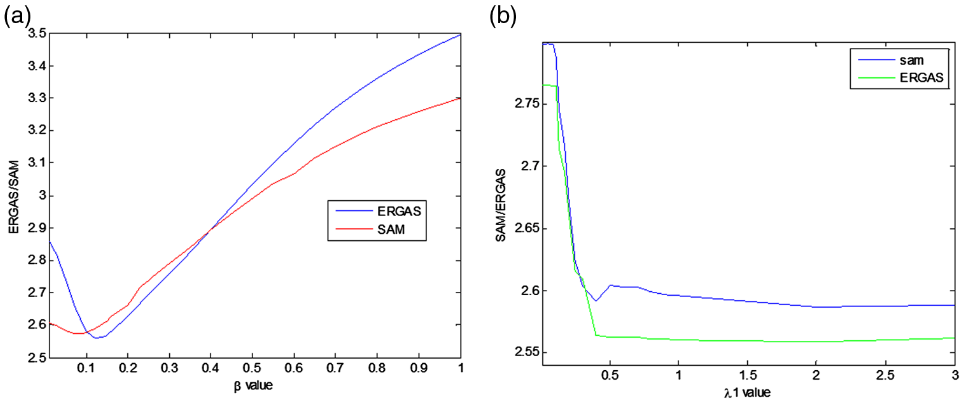


Fig. 10 Variation of the image fusion performance (ERGAS and SAM) with parameters β and λ_1, respectively. (a) Influence of β and (b) influence of λ_1.

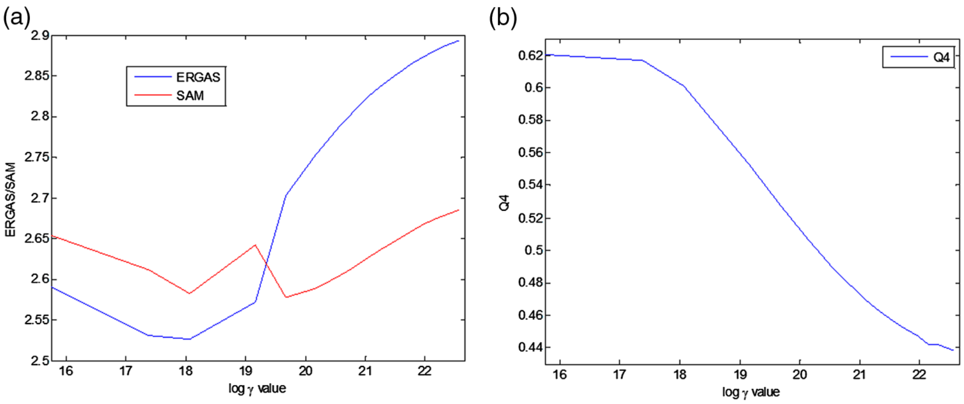


Fig. 11 Variation of the image fusion performance with parameter γ. (a) ERGAS and SAM. (b) Q_4.

Table 5 Quantitative assessment results of the proposed method with different values of λ (λ is from 3×10^3 to 7×10^{10}) for the IKONOS data.

λ	7×10^3	7×10^4	7×10^5	7×10^6	7×10^7	7×10^8	7×10^9	7×10^{10}
Time (s)	1.3×10^4	1.4×10^4	1.38×10^4	2.58×10^4	2.42×10^4	*13.6*	*11.2*	*11.7*
Q_4	0.7093	0.7090	0.7073	0.7071	0.7071	0.7822	0.7817	0.7817
ERGAS	2.8237	2.9034	2.9128	2.9190	2.9191	2.4968	2.5035	2.5035
SAM	3.3015	3.5194	3.5469	3.6396	3.6400	2.5992	2.5933	2.5959
SSIM_{avg}	0.7638	0.7531	0.7517	0.7462	0.7462	0.7927	0.7930	0.7910

Table 6 Time costs of the different image fusion methods for the 512 × 512 IKONOS image.

Method	AWLP	BDSD	PN-TSSC	J-SparseFI	GOAL-IF	Proposed
Time (s)	0.7	0.21	38.2566	19.0887	1287.4158	11.200

Table 7 Time costs of the different image fusion methods for the 512 × 512 QuickBird image.

Method	AWLP	BDSD	PN-TSSC	J-SparseFI	GOAL-IF	Proposed
Time (s)	0.7	0.3	13.5683	11.6302	1036.110	10.3340

chart. It can be clearly observed from Figs. 11(a) and 11(b) that the larger the value of the parameter γ, the worse the performance.

Parameter λ in Eq. (11) determines the weight of the data fidelity term in the loss function. The quantitative assessment results of the proposed method with different values of parameter λ are listed in Table 5. It can be clearly seen that the best results are obtained when λ is above 7×10^8. Moreover, the computational burden is greatly reduced after this point.

3.7 *Time Cost*

All the methods were implemented in MATLAB® 2012a. The personal computer used was a DELL T1500. The central processing unit was a dual-core Intel Core i3 540 at 3.07 GHz. The RAM was 6 GB, and the operating system was Windows 7 64-bit. The running times of all the methods in the experiments with the IKONOS and QuickBird data sets are provided in Tables 6 and 7, respectively.

From Tables 6 and 7, it can be seen that the sparse representation-based image fusion methods (PN-TSSC, J-SparseFI, and GOAL-IF) are time-consuming, which is due to the large computational complexity of the ℓ_1-norm minimization problem. However, the proposed method takes much less time than the other sparse representation-based methods.

4 Conclusions

We have proposed a remote sensing image fusion method based on a cosparse analysis model. To preserve the spatial information, a gradient consistency constraint for the image fusion is proposed under a cosparse analysis framework. The analysis operator is trained offline in advance, and the HR MS image is obtained by solving a cosparse analysis model. With regard to the size of the analysis operator, the image is rearranged with vectored image patches of the same dimension to obtain the global optimization. Both simulated and full-scale, real-data experiments were implemented, and the proposed method was compared with the other state-of-the-art methods. It was found that the proposed method performs well in both spectral and spatial aspects. Furthermore, the proposed method takes much less time than the other sparse representation-based approaches.

References

1. M. Fauvel et al., "Advances in spectral spatial classification of hyperspectral images," *Proc. IEEE* **101**(3), 652–675 (2013).
2. G. Vivone et al., "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.* **53**(5), 2565–2586 (2015).
3. T. M. Tu et al., "A new look at IHS like image fusion methods," *Inf. Fusion* **2**(3), 177–186 (2001).
4. V. P. Shah and N. H. Younan, "An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets," *IEEE Trans. Geosci. Remote Sens.* **46**(5), 1323–1335 (2008).

5. 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, Tech. Rep., Eastman Kodak Company (2000).
6. A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE pan sharpening of very high resolution multispectral images," *IEEE Trans. Geosci. Remote Sens.* **46**(1), 228–236 (2008).
7. A. Garzelli, "Pansharpening of multispectral images based on nonlocal parameter optimization," *IEEE Trans. Geosci. Remote Sens.* **53**(4), 2096–2107 (2015).
8. R. Restaino et al., "Context-adaptive pansharpening based on image segmentation," *IEEE Trans. Geosci. Remote Sens.* **55**(2), 753–766 (2017).
9. X. Otazu et al., "Introduction of sensor spectral response into image fusion methods: application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.* **43**(10), 2376–2385 (2005).
10. K. Amolins, Y. Zhang, and P. Dare, "Wavelet based image fusion techniques—an introduction, review and comparison," *ISPRS J. Photogramm. Remote Sens.* **62**(4), 249–263 (2007).
11. A. Kallel, "MTF-adjusted pansharpening approach based on coupled multiresolution decompositions," *IEEE Trans. Geosci. Remote Sens.* **53**(2), 3124–3145 (2015).
12. F. Palsson et al., "MTF-based deblurring using a Wiener filter for CS and MRA pansharpening methods," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **9**(6), 2255–2269 (2016).
13. L. Alparone et al., "Spatial methods for multispectral pansharpening: multiresolution analysis demystified," *IEEE Trans. Geosci. Remote Sens.* **54**(5), 2563–2576 (2016).
14. J. Nunez et al., "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.* **37**(3), 1204–1211 (1999).
15. B. Aiazzi et al., "Advantages of Laplacian pyramids over 'à trous' wavelet transforms for pansharpening of multispectral images," *Proc. SPIE* **8537**, 853704 (2012).
16. B. Aiazzi et al., "Sensitivity of pansharpening methods to temporal and instrumental changes between multispectral and panchromatic data sets," *IEEE Trans. Geosci. Remote Sens.* **55**(1), 308–319 (2017).
17. S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.* **49**(2), 738–746 (2011).
18. S. Li, H. Yin, and L. Fang, "Remote sensing image fusion via sparse presentations over learned dictionaries," *IEEE Trans. Geosci. Remote Sens.* **51**(9), 4779–4789 (2013).
19. C. Jiang et al., "A practical compressed sensing based pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.* **9**(4), 629–633 (2012).
20. M. Guo et al., "An online coupled dictionary learning approach for remote sensing image fusion," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **7**(4), 1284–1294 (2014).
21. X. He et al., "A new pansharpening method based on spatial and spectral sparsity priors," *IEEE Trans. Image Process.* **23**(9), 4160–4174 (2014).
22. X. X. Zhu and R. Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Trans. Geosci. Remote Sens.* **51**(5), 2827–2836 (2013).
23. C. Jiang et al., "Two-step sparse coding for the pan-sharpening of remote sensing images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **7**(5), 1792–1805 (2014).
24. X. X. Zhu, C. Grohnfeldt, and R. Bamler, "Exploiting joint sparsity for pansharpening: the J-SparseFI algorithm," *IEEE Trans. Geosci. Remote Sens.* **54**(5), 2664–2681 (2016).
25. M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Prob.* **23**(3), 947–968 (2007).
26. S. Nam et al., "The cosparse analysis model and algorithms," *Appl. Comput. Harmon. Anal.* **34**(1), 30–56 (2013).
27. C. Han et al., "A remote sensing image fusion method based on analysis sparse model," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **9**(1), 439–453 (2016).
28. M. Yaghoobi et al., "Constrained overcomplete analysis operator learning for cosparse signal modeling," *IEEE Trans. Signal Process.* **61**(9), 2341–2355 (2013).
29. S. Hawe, M. Kleinsteuber, and K. Diepold, "Analysis operator learning and its application to image reconstruction," *IEEE Trans. Image Process.* **22**(6), 2138–2150 (2013).
30. C. Thomas et al., "Synthesis of multispectral images to high spatial resolution: a critical review of fusion methods based on remote sensing physics," *IEEE Trans. Geosci. Remote Sens.* **46**(5), 1301–1312 (2008).

31. Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.* **9**(3), 81–84 (2002).
32. Z. Wang et al., "Image quality assessment: from error measurement to structural similarity," *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).
33. L. Wald, "Quality of high resolution synthesised images: is there a simple criterion?" in *Proc. 3rd Conf. Fusion Earth Data, Merging Point Measurements, Raster Maps Remotely Sensed Images*, Sophia-Antipolis, France, pp. 99–103 (2000).
34. L. Alparone et al., "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.* **1**(4), 313–317 (2004).
35. M. M. Khan, L. Alparone, and J. Chanussot, "Pan-sharpening quality assessment using the modulation transfer functions of instruments," *IEEE Trans. Geosci. Remote Sens.* **47**(11), 3880–3891 (2009).
36. L. Alparone et al., "Comparison of pansharpening algorithms: outcome of the 2006 GRS-S data fusion contest," *IEEE Trans. Geosci. Remote Sens.* **45**(10), 3012–3021 (2007).
37. L. Alparone et al., "Multispectral and panchromatic data fusion assessment without reference," *Photogramm. Eng. Remote Sens.* **74**(2), 193–200 (2008).

Chang Han received his BS degree in electronic science and technology from Nanchang Hangkong University, Nanchang, China, in 2006, and his MS degree in physical electronics from Fuzhou University, Fuzhou, China, in 2009. He is currently working toward his PhD at the School of Automation, Huazhong University of Science and Technology, Wuhan, China. His research interests include image fusion, sparse representation, pattern recognition, computer vision, and electronics technology.

Nong Sang graduated from Huazhong University of Science and Technology and received his BE degree in computer science and engineering in 1990, his MS degree in pattern recognition and intelligent control in 1993, and his PhD in pattern recognition and intelligent systems in 2000. He is currently a professor at the School of Automation, Huazhong University of Science and Technology, Wuhan, China. His research interests include computational modeling of biological vision perception and applications in computer vision.

Hongyan Zhang received his BS degree in geographic information systems and his PhD in photogrammetry and remote sensing from Wuhan University, China, in 2005 and 2010, respectively. Since 2010, he has been a lecturer at the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. His current research interests focus on image reconstruction and remote sensing image processing.

Liangpei Zhang received his BS degree in physics from Hunan Normal University, ChangSha, China, in 1982, his MS degree in optics from the Xi'an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi'an, China, in 1988, and his PhD in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998. His current research interests focus on remote sensing image processing.