Fusion of Hyperspectral and Panchromatic Images
Using an Average Filter and a Guided filter

Jiahui Qu, a,b Yunsong Li,a,b* Wenqian Dong a

a Xidian University, State Key Lab. of Integrated Service Networks, School of Telecommunications Engineering, No.2, South Taibai Street, Hi-Tech Development Zone, Xi’an, China, 710071
b Xidian University, Joint Laboratory of High Speed Multi-source Image Coding and Processing, School of Telecommunications Engineering, No.2, South Taibai Street, Hi-Tech Development Zone, Xi’an, China, 710071

Abstract. The fusion of hyperspectral and panchromatic images aims to generate a fused image with high spatial and high spectral resolutions. This paper proposes a novel hyperspectral pansharpening method using an average filter and a guided filter. Based on the traditional component substitution methods, we propose a new and simple method to extract the spatial information of the HS image by average filtering at first. Then to solve the significant spectral distortion, a guided filter is utilized to obtain more detailed spatial information from the PAN image which has been sharpened. The appropriate injection gains matrix is generated by selecting the optimal value of the tradeoff coefficient. The spatial detail is finally injected into each band of the interpolated HS image to achieve the fused image. Experimental results demonstrate that the proposed method can obtain more spatial information and preserve more spectral information in both subjective and objective evaluations.

Keywords: Hyperspectral (HS) image, panchromatic (PAN) image, guided filter, average filter, component substitution (CS).

* Corresponding Author, E-mail: ysl@mail.xidian.edu.cn

1 Introduction

Image fusion is a process which can synthesize and extract the information of two or more different images to obtain improved information by using a certain algorithm [1]. Remote sensing image fusion is an important part of image fusion. Remote sensing image fusion aims to combine the information of different spectral and spatial resolutions images to achieve more
useful information [2]. Low spatial resolution hyperspectral (HS) image and high spatial resolution panchromatic (PAN) image are provided by various sensors. The HS image is a low spatial resolution image but has high spectral resolution. The PAN image is a high spatial resolution image with low spectral resolution. Hence the fusion of hyperspectral and panchromatic images is a meaningful technology because it can produce a fused image with the high spectral resolution of the former and the high spatial resolution of the latter.

A large number of methods have been proposed for the fusion of hyperspectral and panchromatic images [3]. They can be broadly separated into five classes: component substitution (CS) algorithms, multiresolution analysis (MRA) algorithms, matrix factorization algorithms, Bayesian algorithms, and hybrid algorithms [4]. CS algorithms and MRA algorithms are traditional fusion methods. CS approaches convert the HS image into another data space where spectral and spatial information are separated. Then the spatial component is replaced by the PAN image [5]. The CS techniques contain algorithms such as the generalized intensity-hue-saturation (GIHS) [6], the principal component analysis (PCA) [7], [8], the Gram-Schmidt (GS) [9], and the GS Adaptive (GSA) [10] method. Those CS methods have three main advantages: 1. High fidelity of the spatial details [11], 2. Fast and simple implementation [6], and 3. Good robustness [11]. But the CS techniques also have a serious shortcoming. The methods can generate significant spectral distortion [12].
MRA approaches inject the spatial information of the PAN image into the HS image. MRA approaches include algorithms, such as the smoothing filter-based intensity modulation (SFIM) [13], the MTF-Generalized Laplacian Pyramid (MTF-GLP) [14], the MTF-GLP with High Pass Modulation (MTF-GLP-HPM) [15], and the decimated wavelet transform (DWT) [16] method. The advantages of the MRA methods are spectral consistency and temporal coherence [17]. The main shortcoming is the complicated implementation [3].

Matrix factorization approaches, Bayesian approaches, and hybrid approaches are proposed recently. Matrix factorization algorithms and Bayesian algorithms are model based methods. They perform well but are accomplished with high computational cost, e.g., the coupled nonnegative matrix factorization (CNMF) [18], the Bayesian HySure [19], the Bayesian Sparsity promoted Gaussian prior (Bayesian Sparse) [20] [21], and the Bayesian Naive Gaussian prior (Bayesian Naive) [22] method. Hybrid methods use concepts from different methods integrated into one method, such as the guided filter PCA (GFPCA) [23] method. The GFPCA method can preserve the spectral information well. However, the GFPCA method produces lots of blurs because the spatial information is not sufficiently integrated in the fused product.

This paper proposes a novel hyperspectral image fusion method with an average filter and a guided filter (AFGF) to solve the problems mentioned above. Based on the model of component substitution methods, we propose a simple and effective approach to retrieve the spatial
information of the HS image by average filtering at first. Then to overcome the serious spectral distortion of component substitution methods, the guided filter is used for extracting more detailed spatial information from the PAN image, which has been sharpened. Finally, the appropriate injection gains matrix is presented to reduce the spectral and spatial distortion. Experimental results illustrate that the method using the enhanced PAN image can achieve better effects. Experimental results also demonstrate that the proposed fusion method performs superior in terms of subjective and objective evaluations.

2 Related Work

2.1 Component Substitution Pansharpening Methods

Component substitution (CS) is a popular and classical pansharpening method. CS technique projects the HS image into another space to separate spectral and spatial information \([12]\). Then the transformed spatial information is substituted by the histogram-matched PAN image. Finally, the fused image is obtained by applying the inverse spectral transformation \([3]\). Many CS pansharpening methods are extended from multispectral images to hyperspectral images.

A general formulation of the CS method is given by \([24]\)

\[
HS^F_k = HS_k + g_k(P - I)
\]

where \(HS^F_k\) and \(HS_k\) are the \(k\)th band of the final fused image and the interpolated HS image, \(P\)
is the PAN image, \( g_k \) is the injection gain, and \( I \) represents the spatial information of the HS image that is defined as

\[
I = \sum_{i=1}^{m_\lambda} \alpha_i HS_i
\]  

(2)

where \( m_\lambda \) is the number of the HS image bands, and \( \alpha_i \) is the weight that measures the proportional spectral overlap between each HS band and the PAN image [12]. The weighting factor \( \alpha_i \) is generally calculated as \( 1/m_\lambda \).

In CS methods, the weight \( \alpha_i \) and the injection gain \( g_k \) of the generalized intensity-hue-saturation (GIHS) [6] approach are defined as \( \alpha_i = 1/m_\lambda \) and \( g_k = 1 \). In [9], the weight \( \alpha_i \) and the injection gain \( g_k \) of the Gram-Schmidt (GS) orthogonalization method are defined as \( \alpha_i = 1/m_\lambda \) and \( g_k = \text{cov}(HS_k, I)/\text{var}(I) \), where \( \text{cov}(A, B) \) denotes the covariance operation and \( \text{var}(C) \) denotes the variance operation. In [10], an enhanced GS Adaptive (GSA) method is proposed, in which the weight \( \alpha_i \) is obtained by minimizing the mean-square error (mse) between the estimated spatial information \( I \) and the decimated PAN image. In addition, in [26], an adaptive IHS method is introduced, in which the weighting factor \( \alpha_i \) is calculated by solving the following optimization problem

\[
\min_{\alpha_1, \ldots, \alpha_n} \left\| P - \sum_{i=1}^{m_\lambda} \alpha_i HS_i \right\|^2 \quad \text{s.t. } \alpha_i \geq 0, \ldots, \alpha_n \geq 0
\]  

(3)

In [26], the injection gain \( g_k \) is a weighting matrix which is produced by the edges of the PAN
image and is given by

$$g_k = \exp\left(-\frac{\xi}{\|\nabla P\|^2 + \eta}\right)$$  \hspace{1cm} (4)

where $\xi$ and $\eta$ are the tradeoff parameters, and $\nabla P$ is the gradient of the PAN image.

2.2 Guided Filter

A guided filter was proposed by He et al. [27]. The guided filter is an edge-preserving smoothing filter which is derived from a local linear model. It has numerous image processing applications, such as detail enhancement, HDR compression, image matting, dehazing, etc. It can transfer structures from the guidance image $G$ to the output image $O$. The guided filter can also preserve the gradient and avoid the reversal artifacts. It can be applied to the image fusion because the guided filter has the advantages mentioned above.

In [27], the guided filter assumes that the output image $O$ is a linear transformation of the guidance image $G$ in a local window $w_k$ centered at pixel $k$, i.e.

$$O_i = a_k G_i + b_k, \forall i \in w_k$$  \hspace{1cm} (5)

where $w_k$ is a window of size $(2r+1) \times (2r+1)$ and the coefficients $a_k$ and $b_k$ are assumed to be constant in $w_k$. $(a_k, b_k)$ can be estimated by minimizing the squared difference between the output image $O$ and the input image $I$ in the window $w_k$:

$$E(a_k, b_k) = \sum_{i \in w_k} ((a_k G_i + b_k - I_i)^2 + \lambda a_k^2)$$  \hspace{1cm} (6)
where $\lambda$ is a regularization parameter. The constants $a_k$ and $b_k$ can be directly estimated by linear ridge regression [28]

\[
a_k = \frac{1}{|w|} \sum_{i \in w_k} G_i I_i - \mu_k \overline{I}_k}{\sigma_k^2 + \lambda}
\]

\[
b_k = \overline{I}_k - a_k \mu_k
\]

where $|w|$ is the number of pixels in $w_k$, $\mu_k$ and $\sigma_k^2$ are the mean and variance of $G$ in $w_k$, and $\overline{I}_k$ is the mean of $I$ in $w_k$. The value of output image $O_i$ will change when it is computed in different widows because several different local windows centered at the pixel $k$ contain pixel $i$. So the values of coefficients $a_k$ and $b_k$ are averaged. The final guided filter is computed as follows:

\[
O_i = a_i G_i + b_i
\]

where $a_i = \frac{1}{|w|} \sum_{k \in w_i} a_k$ and $b_i = \frac{1}{|w|} \sum_{k \in w_i} b_k$. In this Paper, we use the following formula to represent the guided filter operation

\[
O = f(I, G)
\]

here, $f$ represents the guided filter function, $G$ is the guidance image, $I$ is the input image, and $O$ is the output image.

3 Proposed Method

Fig.1 shows the main processes of the proposed hyperspectral pansharpening technique with
average and guided filters. The proposed approach mainly consists of three steps. First, the spatial information of the HS image is extracted by average filtering. Then, a guided filter is applied to obtain more detailed spatial information from the PAN image which has been sharpened. Finally, the injection gains matrix is generated.

The traditional component substitution methods rely on substituting the spatial information component of the HS image with the PAN image. They generate serious spectral distortion. In order to overcome this problem, the proposed method utilizes the guided filter to transfer the spatial information from the enhanced PAN image to the HS image, without causing spectral distortion. The injection gains matrix of the proposed method is also generated to restrain the spectral and spatial distortion. The proposed method is based on the component substitution approach. To obtain the spatial information of the HS image $I$, the common method is to solve an optimization problem. We present a simple method which makes use of the average filter to reduce the amount of calculation.

3.1 Extracting the Spatial Information of HS Image using an Average Filter

According to traditional component substitution approaches, the first step of the proposed method obtains the spatial information of the HS image. Firstly, the low resolution HS image is interpolated to match the scale of the PAN image. Let us denote this upsampled image as $HS$. 

8
Then, all bands of the $HS$ image are decomposed into two-scale representations using an average filter. The base layer of the $k$th band can be described as

$$B_{HS_k} = HS_k \ast W$$

where $B_{HS}$ represents the base layer of the $HS$ image, $B_{HS_k}$ represents the base layer of the $k$th band, $HS_k$ represents the $k$th band of the $HS$ image, and $W$ represents the average filter. The detail layer is obtained by subtracting the base layer from the $HS$ image. This process can be depicted as

$$D_{HS} = HS - B_{HS}$$
where $D_{HS}$ represents the detail layer of the $HS$ image. The detail layer contains the details of each band. Finally, the spatial information of HS image is obtained by adding the detail layer of all bands together as follows:

$$I = \sum_{k=1}^{m} D_{HS}^k$$

where $I$ is the spatial information of the HS image, $m$ is the number of bands, and $D_{HS}^k$ is the detail layer of the $k$th band.

### 3.2 Transferring Spatial Details Using a Guided Filter

In this step, we transfer the spatial details from the PAN image to the spatial information of HS image $I$ using a guided filter to reduce the spectral distortion. As shown in Fig. 1, the PAN image is sharpened to enhance the details first. The detail enhancement of the PAN image can be accomplished using the LOG (Laplacian of Gaussian) operator as follows. First, a Gaussian filter is applied to the PAN image to reduce noise.

$$PAN_g = PAN * G$$

where $G$ is a Gaussian filter. Then we make use of the Laplace operator to sharpen and enhance the image $PAN_g$. This procedure can be described as

$$PAN_s = PAN_g + \nabla^2 PAN_g$$

where $PAN_s$ represents the enhanced PAN image, and $\nabla^2 PAN_g$ represents the Laplace operator.
The spatial detail of the $PAN_s$ image is clearer than that of the $PAN$ image. Next, the guided filter is applied to the spatial information of the HS image $I$ with the sharpened image $PAN_s$ serving as the guidance image.

$$I_{GF} = f(I, PAN_s)$$

(16)

here, $I_{GF}$ is the obtained spatial details, $I$ is the input image, and $f$ represents the guided filter function.

3.3 Generating the Injection Gains Matrix

This step generates the injection gains matrix. The injection gains matrix $g_k$ of the proposed method is generated with reducing the spectral and spatial distortion as follows. The proportions between each pair of the HS image bands should remain unchanged to preserve the spectral information. Thus, the injection gains $g_k$ can be defined as

$$g_k \propto \frac{HS_k}{(1/m_\lambda)\sum_{k=1}^{m_\lambda} HS_k}$$

(17)

where $m_\lambda$ is the number of bands, and $HS_k$ is the $k$th band of the interpolated image $HS$. Then, we define a tradeoff coefficient $\beta$ to restrain the spatial distortion:

$$g_k = \beta \cdot \frac{HS_k}{(1/m_\lambda)\sum_{k=1}^{m_\lambda} HS_k}$$

(18)

The optimal value of the tradeoff coefficient $\beta$ is selected by spatial data analysis, which is
shown in section 4. Finally, the fused image $H S^F$ is obtained by injecting the spatial details $I_{GF}$ into the interpolated image $H S$ for each band

$$H S^F_k = H S_k + g_k I_{GF} \quad (19)$$

4 Experimental Results and Analysis

In this part, the experimental results of the proposed hyperspectral pansharpening technique with an average filter and a guided filter (AFGF) are compared with six state-of-the-art image fusion methods. There are the principal component analysis (PCA) method [7], [8], the Gram-Schmidt adaptive (GSA) method [10], the hybrid method based on a guided filter and PCA (GFPCA) method [23], the coupled nonnegative matrix factorization (CNMF) method [18], the MTF-GLP with High Pass Modulation (MTF-GLP-HPM) method [15], and the Bayesian Sparsity promoted Gaussian prior (Bayesian Sparse) [20] [21] method. Several quality measures are used for assessing the performance of different hyperspectral image fusion methods.

4.1 Quality Measures

In the experiments, four well-known indexes are adopted for evaluating the property of the image fusion methods, i.e., the cross correlation (CC) [3], the spectral angle mapper (SAM) [29], the root mean squared error (RMSE) [3] and the erreur relative global adimensionnelle de synthesè (ERGAS) [30]. Those quality measures can reflect the similarity between the obtained
fused HS image F and the reference HS image R.

1) Cross correlation (CC)

The CC which is a spatial index measures the geometric distortion between the obtained fused HS image F and the reference HS image R. It is defined as

\[
CC(F, R) = \frac{1}{m} \sum_{k=1}^{m} CCS(F_k, R_k)
\]

where \(m\) is the number of the HS image bands, \(F_k\) and \(R_k\) are the \(k\)th band of the fused image and the reference image, and CCS is the cross correlation of a single-band image, which is defined as

\[
CCS(M, N) = \frac{\sum_{i=1}^{n} (M_i - \mu_M)(N_i - \mu_N)}{\sqrt{\sum_{i=1}^{n} (M_i - \mu_M)^2 \sum_{i=1}^{n} (N_i - \mu_N)^2}}
\]

where \(M_i\) and \(N_i\) are the pixel values of one band of the fused image and the reference image, \(\mu_M\) and \(\mu_N\) are the mean, and \(n\) is the number of pixels for one band. The ideal value of the CC is 1.

2) Spectral angle mapper (SAM)

The SAM which is a spectral index reflects the spectral distortion between the obtained fused HS image F and the reference HS image R. It is defined as

\[
SAM(F, R) = \frac{1}{n} \sum_{i=1}^{n} SAM(F_i, R_i)
\]
where \( n \) is the number of pixels for one band, \( F_i \) and \( R_i \) are the spectral vectors of the fused image and the reference image, and given the spectral vectors \( A, B \in \mathbb{R}^{m_\lambda}, \)

\[
SAM(A, B) = \arccos \left( \frac{\langle A, B \rangle}{\|A\|_2 \|B\|_2} \right)
\]

where \( \langle A, B \rangle = A^T B \) is the inner product between \( A \) and \( B \). The ideal value of the SAM is 0.

3) Root mean squared error (RMSE)

The RESM is a global index measure of the difference between the obtained fused image \( F \) and the reference image \( R \). It is defined as

\[
RMSE(F, R) = \frac{\|F - R\|_F}{\sqrt{n^* m_\lambda}}
\]

where \( \|A\|_F = \sqrt{\text{trace}(A^T A)} \) is the Frobenius norm of \( A \), \( n \) is the number of pixels for one band, and \( m_\lambda \) is the number of the HS image bands. The smaller the RMSE value, the more similar is the fused image to the reference image. The optimal value of the RMSE is 0.

4) erreur relative global adimensionnelle de synthèse (ERGAS)

The ERGAS, which is also a global index, reflects the entire quality of the obtained fused image \( F \). It is defined as

\[
ERGAS(F, R) = 100d \sqrt{\frac{1}{m_\lambda} \sum_{k=1}^{m_\lambda} \left( \frac{\text{RMSE}_k}{\mu_k} \right)^2}
\]

Where \( d \) is the ratio between the pixel sizes of the PAN image and the HS
image, \( RMSE_k = \|F_k - R_k\|_F / \sqrt{n} \), \( m_k \) is the number of the HS image bands, and \( \mu_k \) is the mean of the \( k \)th band of the reference image \( R \). The optimal value of the ERGAS is 0.

4.2 Analysis of the Influence of Parameter \( \beta \)

In the proposed method, the parameter \( \beta \) controls the quantity of the injected spatial information and restraints the spatial distortion of fused HS image. Thus, the tradeoff coefficient \( \beta \) is controlled by the indexes CC, RMSE and ERGAS which can reflect the degree of spatial distortion. Larger CC and smaller RMSE, ERGAS values indicate better fused result. In order to analyze the influence of parameter \( \beta \), the proposed method is performed on the Moffett Field dataset with different \( \beta \) settings. Table 1 shows the objective performance of the proposed method with different tradeoff coefficient \( \beta \) settings. As shown in Table 1, when the tradeoff coefficient \( \beta \) is 0.2, the fused results achieve the best performance. We have also tested the fusion performance of the proposed method on some hyperspectral images, and found that \( \beta = 0.2 \) also acquires the optimal effects there. Therefore, in the proposed method, the tradeoff coefficient \( \beta \) is set to 0.2.

<table>
<thead>
<tr>
<th>( \beta )</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>0.9344</td>
<td>0.9591</td>
<td><strong>0.9660</strong></td>
<td>0.9617</td>
<td>0.9505</td>
<td>0.9355</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0344</td>
<td>0.0276</td>
<td><strong>0.0255</strong></td>
<td>0.0293</td>
<td>0.0371</td>
<td>0.0469</td>
</tr>
<tr>
<td>ERGAS</td>
<td>5.8745</td>
<td>4.6899</td>
<td><strong>4.2779</strong></td>
<td>4.8319</td>
<td>6.0891</td>
<td>7.7109</td>
</tr>
</tbody>
</table>
4.3 Experimental Results

In order to assess the performance of the proposed average filtering and guided filtering based fusion method (AFGF), we test two kinds of semi-synthetic datasets, which are the hyperspectral remote sensing images and the hyperspectral images of natural scenes. According to [27], since the guided filter transfers the detailed structures in the proposed method, a suitable filter size \( r \) which should not be too large or too small and a small blur degree \( \lambda \) are preferred. According to the image pixel size, the parameters are set to \( r = 58 \) and \( \lambda = 10^{-6} \). In the experiments, the pixel values of every test images are normalized to the range of 0-1.0 to reduce the quantity of calculation.

1) Hyperspectral remote sensing images

The first experiment is performed on a semi-synthetic dataset derived from the Moffett field dataset. This hyperspectral remote sensing dataset was acquired by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) operated by NASA/JPL [4]. A mixed urban and rural scene is contained in this dataset. The HS images are characterized by 224 bands and acquired in the spectral range of 0.4-2.5\( \mu \)m. Fig. 2(a) shows a HS image which is used as the reference image. According to the Wald’s protocol [31], the simulated HS image is obtained by blurring and downsampling the reference image with a ratio of 5, and the simulated PAN image is obtained by means of averaging the bands of the visible range of the reference image. Fig. 2(b) shows the
simulated PAN image. The HS image is $37 \times 79$ with 100m spatial resolution and the PAN image is $185 \times 395$ with 20m spatial resolution.

![Fig. 2](image_url)

**Fig. 2** Moffett field dataset fusion results. (a) Reference HS image. (b) Synthetic PAN image. (c) GSA method. (d) PCA method. (e) MTF-GLP-HPM method. (f) GFPCA method. (g) CNMF method. (h) Bayesian Sparse method. (i) AFGF method without sharpening the PAN. (j) AFGF method.

Fig. 2 shows the Moffett field dataset fusion results obtained by each method. From Fig. 2(c)-(d), it can be observed that the results produced by GSA and PCA methods have high fidelity in rendering the spatial information, but generate significant spectral distortion. Moreover, the GSA and PCA methods may add some extra gaps, especially at the top area of the fused images. The result of the MTF-GLP-HPM method has better spectral performance, but
brings in serious spatial distortion in some regions. Fig. 2(f) shows that the GFPCA method generates lots of blurs in some areas and the spatial details are not sufficient. Fig. 2(g) displays that the CNMF method obtains better spectral performance but the spatial information added is limited in some details. From Fig. 2(h), it can be seen that the Bayesian Sparse method works well, but the river at the bottom area of the fused image has insufficient spatial details. Fig. 2(i) shows the result of the AFGF method which the PAN image is not sharpened in the section 3.2. A comparison of Fig. 2(i) and Fig. 2(j) reveals that the result of the AFGF method has better performance in spatial aspects. Thus, it is necessary to sharpen the PAN image in the proposed frame. By comparison, the result shown in Fig. 2(j) proves that the proposed method performs superior in both spectral and spatial aspects.

<table>
<thead>
<tr>
<th>Indexes</th>
<th>GSA</th>
<th>PCA</th>
<th>MTF-GLP-HPM</th>
<th>GFPCA</th>
<th>CNMF</th>
<th>Bayesian Sparse</th>
<th>AFGF not sharpened</th>
<th>AFGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>0.9530</td>
<td>0.9066</td>
<td>0.8919</td>
<td>0.9130</td>
<td>0.9393</td>
<td>0.9523</td>
<td>0.9489</td>
<td>0.9660</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0324</td>
<td>0.0444</td>
<td>0.1339</td>
<td>0.0424</td>
<td>0.0347</td>
<td>0.0313</td>
<td>0.0311</td>
<td><strong>0.0255</strong></td>
</tr>
<tr>
<td>ERGAS</td>
<td>5.3450</td>
<td>7.2843</td>
<td>28.8234</td>
<td>7.0149</td>
<td>5.8157</td>
<td>5.3132</td>
<td>5.3177</td>
<td><strong>4.2779</strong></td>
</tr>
</tbody>
</table>

The objective quantitative assessments of different methods are shown in Table 2. It is obvious that, for the Moffett field dataset, the CC index of the proposed method is the biggest, and the RMSE and ERGAS values of this approach are the smallest. The SAM value ranks as the
second. These results demonstrate that the AFGF method shows the best fusion performance.

Fig. 3 Hyperion dataset fusion results. (a) Reference HS image. (b) Synthetic PAN image. (c) GSA method. (d) PCA method. (e) MTF-GLP-HPM method. (f) GFPCA method. (g) CNMF method. (h) Bayesian Sparse method. (i) AFGF method without sharpening the PAN. (j) AFGF method.

Another semi-synthetic dataset derived from the Hyperion instrument is utilized to assess the fusion performance of the proposed approach. This hyperspectral dataset has been acquired by the EO-1 spacecraft operated by NASA [4]. A class of Earth observation data is provided in this dataset. These hyperspectral images consist of 242 bands. They cover the range of 0.4-2.5 μm. Fig. 3(a) shows the reference HS image. According to the Wald’s protocol [31], the synthetic HS image is provided by blurring and downsampling the reference HS image with a ratio of 3, and the synthetic PAN image is acquired by averaging the bands of the visible range. Fig. 3(b) shows the synthetic PAN image. The HS image is 58×72 with 30m spatial resolution.
and the PAN image is $174 \times 216$ with 10m spatial resolution.

The Hyperion dataset fusion results achieved by different methods are shown in Fig. 3. As can be observed from the Fig. 3(c) and (d), the results achieved by the GSA, and PCA methods generate serious spectral distortion. Fig. 3(e) displays that the MTF-GLP-HPM method obtains better performance in spatial aspect, but has slightly spectral distortion. As shown in Fig. 3(f), the GFPCA method is significantly blurred. Fig. 3(g) and (h) describe that the CNMF and Bayesian Sparse methods work well, but certain areas of the fused images have deficient spatial details. Similar to the first experiment, the fused result of the AFGF method for which the PAN image is not sharpened is shown in Fig. 3(i). By comparing it with Fig. 3(j), it can be seen that the AFGF method using the enhanced PAN image contains more spatial detail and is less spectrally distorted. In total, the proposed method is more similar to the reference HS image in spatial and spectral aspects.

Table 3. Objective quantitative assessments of the Hyperion dataset fusion results shown in Fig. 3.

<table>
<thead>
<tr>
<th>Indexes</th>
<th>GSA</th>
<th>PCA</th>
<th>MTF-GLP-HPM</th>
<th>GFPCA</th>
<th>CNMF</th>
<th>Bayesian Sparse</th>
<th>AFGF not sharpened</th>
<th>AFGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>0.9507</td>
<td>0.9502</td>
<td>0.9420</td>
<td>0.8880</td>
<td>0.8969</td>
<td>0.8795</td>
<td>0.9403</td>
<td><strong>0.9508</strong></td>
</tr>
<tr>
<td>SAM</td>
<td>2.4700</td>
<td>2.4891</td>
<td>2.4803</td>
<td>3.2906</td>
<td>2.4830</td>
<td>3.5371</td>
<td>2.4121</td>
<td><strong>2.4123</strong></td>
</tr>
<tr>
<td>RMSE</td>
<td><strong>0.0061</strong></td>
<td>0.0106</td>
<td>0.0068</td>
<td>0.0159</td>
<td>0.0118</td>
<td>0.0129</td>
<td>0.0105</td>
<td>0.0069</td>
</tr>
<tr>
<td>ERGAS</td>
<td>2.7031</td>
<td>3.4075</td>
<td>2.6705</td>
<td>5.1613</td>
<td>3.8485</td>
<td>4.2760</td>
<td>4.0567</td>
<td><strong>2.6472</strong></td>
</tr>
</tbody>
</table>

Table 3 reveals the objective evaluation results of the Hyperion dataset. From Table 3, it
can be seen that, for the AFGF method, the indexes CC, SAM, and ERGAS obtain the best values. The RMSE value appears second best. The results prove the superior performance of the proposed method.

2) Hyperspectral images of natural scenes

The third experiment is accomplished on a semi-synthetic dataset which is the Manchester university dataset [32]. This dataset contains 50 Hyperspectral images of natural scenes and were provided by David H. Foster et al. These hyperspectral images of natural scenes which were almost all taken under a clear sky have been acquired during the summers of 2002 and 2003. Mixed rural scenes and urban scenes are contained in this dataset. These images are characterized by 33 bands in the spectral range of 400-720nm.

Fig. 4(a) shows a HS image from this database, which serves as the reference image. The original HS image is $1344 \times 1024$. The number of rows and columns of the HS image is $600 \times 800$ for this experiment. According to the Wald’s protocol [31], the simulated HS image is acquired by applying a blurring to the reference image and then downsampling it with a ratio of 5. The simulated PAN image is provided by pansharpening the spectral response. The simulated HS image is $160 \times 120$. The simulated PAN image is $800 \times 600$. 

21
Fig. 4 Hyperspectral image of natural scenes fusion results. (a) Reference HS image. (b) GSA method. (c) PCA method. (d) MTF-GLP-HPM method. (e) GFPCA method. (f) CNMF method. (g) Bayesian Sparse method. (h) AFGF method without sharpening the PAN. (i) AFGF method.

The hyperspectral image of natural scenes fusion results are shown in Fig. 4(b)-(i). By comparison, the results of the GSA, PCA, GFPCA and CNMF methods produce spectral distortion, especially in the area of the red ball. Besides, the result of the CNMF method shown in Fig. 4(f) is a bit blurry. As shown in Fig. 4(d), it can be seen that the MTF-GLP-HPM method has serious spectral and spatial distortion. The fused result of the Bayesian Sparse method obtains better spectral performance, but the spatial information of some areas shows slightly fuzzy. Fig. 4(h) produced without a sharpened PAN image is less sharp than Fig. 4(i). It proves...
once again that sharpening the PAN image within the proposed frame is significant. The AFGF method achieves better result in both spatial details and spectral information aspects. Meanwhile the better performance of the AFGF method is apparent from the objective evaluation results in Table 4. For the proposed method, the indexes SAM, RMSE and ERGAS are the smallest and the CC is the largest. In total, for the hyperspectral image of natural scenes, the proposed method achieves superior effects.

Table 4. Objective quantitative assessments of the hyperspectral image of natural scenes fusion results shown in Fig. 4.

<table>
<thead>
<tr>
<th>Indexes</th>
<th>GSA</th>
<th>PCA</th>
<th>MTF-GLP-HPM</th>
<th>GFPCA</th>
<th>CNMF</th>
<th>Bayesian Sparse</th>
<th>AFGF not sharpened</th>
<th>AFGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>0.9963</td>
<td>0.9939</td>
<td>0.7697</td>
<td>0.9894</td>
<td>0.9928</td>
<td>0.9953</td>
<td>0.9963</td>
<td><strong>0.9969</strong></td>
</tr>
<tr>
<td>SAM</td>
<td>2.2476</td>
<td>2.4590</td>
<td>4.1117</td>
<td>3.7124</td>
<td>3.2902</td>
<td>2.0441</td>
<td><strong>1.9333</strong></td>
<td><strong>1.9333</strong></td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0103</td>
<td>0.0152</td>
<td>0.3026</td>
<td>0.0139</td>
<td>0.0090</td>
<td>0.0104</td>
<td>0.0091</td>
<td><strong>0.0078</strong></td>
</tr>
<tr>
<td>ERGAS</td>
<td>1.6775</td>
<td>2.0699</td>
<td>128.8305</td>
<td>2.9613</td>
<td>2.5918</td>
<td>1.9240</td>
<td>1.7001</td>
<td><strong>1.5658</strong></td>
</tr>
</tbody>
</table>

5 Conclusions

The intent of this paper is to introduce a new hyperspectral image fusion method which uses an average filter and a guided filter. This method is based on the component substitution approach. To reduce the amount of calculation, we first propose a simple method that utilizes the average filter to obtain the spatial information of the HS image \( I \). Subsequently, the PAN image is sharpened to enhance the spatial detail. In order to avoid the spectral distortion, a guided filter is
used in transferring the spatial information from the enhanced PAN image to the spatial information of the HS image $I$. Then, the injection gains matrix is generated to reduce spectral distortion and restraint spatial distortion. Experimental results demonstrate that the proposed method is a more effective method compared with the state-of-the-art fusion methods both in subjective and objective evaluations. The proposed algorithm has been applied to hyperspectral pansharpening. In the future, whether the proposed method can be applied to the fusion of infrared and visual images can be further researched.

Acknowledgments

The authors would like to thank the editors, and the anonymous reviewers for their insightful comments and suggestions which have greatly improved this paper. This work was supported by the National Science Foundation of China under Grants 61222101, 61272120, 61301287, 61301291 and 61350110239.

References


[18] N. Yokoya, T. Yairi, and A. Iwasaki, “Coupled nonnegative matrix factorization unmixing for


Highlights:

• An average filtering and guided filtering based fusion method is proposed.
• We utilize the average filter to extract the spatial information of the HS image.
• The PAN image is sharpened to enhance the spatial detail.
• The guided filter is utilized to transfer spatial information.
• The proposed method is effective for hyperspectral pansharpening.