Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-018-0866-4

ORIGINAL RESEARCH

# Pansharpening using a guided image filter based on dual-scale detail extraction

**Lihua Jian[1] · Xiaomin Yang[1] · Wei Wu[1] · Awais Ahmad[2] · Arun Kumar Sangaiah[3] · Gwanggil Jeon[4]**

Received: 29 December 2017 / Accepted: 14 May 2018


**Abstract**
High spatial resolution multispectral (HMS) images can provide sufficient information for researchers to analyze the potential disasters in the living environment. However, an original multispectral (MS) image is with low-space-resolution and high-spectrum-resolution, while an original panchromatic (PAN) image has the opposite property. Pansharpening aims at obtaining HMS image by retaining spectrum of the MS image and injecting details of the PAN image simultaneously. In this paper, we present a new pansharpening method. First, we use a bilateral filter (BF) to obtain the low-frequency-component (LFC) of PAN and MS images, respectively. Then the high-frequency-component (HFC) of PAN and MS images are readily obtained. Second, an adaptive intensity-hue-saturation (AIHS) based method is applied to generate the HFC of intensity. Finally, a dual-scale guided image filter (GIF) is utilized to calculate the difference between HFC of intensity and PAN to get the detail images. And then, these detail images are injected into the original MS image to achieve the HMS image. The proposed method is applied into testing various satellite data sets, and performs better effect on both visual quality and objective indictors than the existing methods.

**Keywords** Pansharpening · Bilateral filter · AIHS transform · Guided image filter

## 1 Introduction

Satellite images are widely used in many applications, such as drawing maps of Earth's surface structure and content, monitoring environment, detecting agricultural disaster, etc. However, because of some limitations when extracting different portions of electromagnetic spectrum, satellite sensor cannot provide both high space and spectral resolution images simultaneously (Zhang 2004; Simone et al. 2002; Wu et al. 2016). Therefore, many methods about fusing the multispectral (MS) image and the panchromatic (PAN) image have been proposed, which is named pansharpening. Pansharpening can improve visual perception of satellite images, since it could merge the complementary information, which come from two different resolution of satellite images. Additionally, many applications about pansharpening are to solve the extreme imaging environment, such as street view imagery in smart city (Gebru et al. 2017; Javidnia and Corcoran 2017; Chen et al. 2016; Tao et al. 2018), and satellite maps visualization (Miao et al. 2016; Ferster and Coops 2016) in smart environment. At present, the remote sensing data has become very useful by using Internet of Things (IoT) systems when predicting the weather.

Generally, there are two ways to improve the resolution of the MS image. One is to reinvent optic devices and update hardware equipments. However, this approach is not only high cost but also difficult to realize. Therefore, more research interests have been turned to the intelligent algorithms like pansharpening. During this process, balancing the space detail enhancement and the spectral consistency of the fused image is the key step to obtain the high resolution multispectral (HMS) image.

In recent years, numerous pansharpening methods have been presented to increase the space resolution of the MS images. Most of these methods (Liu and Liang 2016;

✉ Xiaomin Yang
arielyang@scu.edu.cn

[1] College of Electronics and Information Engineering, Sichuan University, Chengdu 610064, Sichuan, China

[2] Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, South Korea

[3] School of Computing Science and Engineering, VIT University, Vellore, India

[4] Department of Embedded Systems Engineering, Incheon National University, Incheon, South Korea

Published online: 24 May 2018

Ghassemian 2016; Rahmani et al. 2010; Alparone et al. 2007; Jameel et al. 2016) are commonly divided into two steps: (1) estimating the detail information from the PAN images; (2) sharpening the corresponding MS images by transferring the detail information, which can be extracted by subtracting the approximate image of PAN from the original PAN image. Component-substitution (CS) methods (Rahmani et al. 2010; Tu et al. 2001; Leung et al. 2014; Aiazzi et al. 2007) and multi-resolution-analysis (MRA) methods (Nunez et al. 1999; Upla et al. 2013; Jin et al. 2014; Choi et al. 2005; Jian et al. 2018) are two classical methods to calculate the approximation of the PAN image.

The CS methods are frequently used because they are simple and efficient. Methods belong to this class are based $w_i$ on transformation of the MS image into a new space, in which spectral information and spectral features can be independent of each other. In such conditions, the spatial resolution of the MS image can be enhanced by component substitution with the PAN information. Some well-known methods are intensity-hue-saturation (IHS) transform (Tu et al. 2001), various versions of IHS (Rahmani et al. 2010; Leung et al. 2014), the principle component analysis (PCA) (Sun et al. 2006), and the Gram-Schmidt (GS) (Aiazzi et al. 2007). A general expression of CS methods can be written as follows:

$$HMS_i = MS_i + w_i(PAN - APP),\ i = 1, \ldots, N \tag{1}$$

where $HMS_i$ indicates the pansharpening result. is the injection gain coefficient. $MS_i$ denotes the $i$th band of the MS image. $N$ is the number of the bands of the MS image. $APP$ is the approximation of the PAN image, which can be defined as:

$$APP = \sum_{i=1}^{N} \alpha_i MS_i \tag{2}$$

where $\alpha_i$ is the $i$th band weight coefficient of mapping transformation.

As the most representative CS method, the IHS-based method (Tu et al. 2001) ($w_i = 1, \alpha_i = 1/N,\ N = 3$) directly regards the difference between intensity component ($APP$) and the PAN image as the missing high-frequency detail of the MS image. However, results obtained by this method often produce serious spectral distortion. Hence, to eliminate the spectral distortion of the fused image, Rahmani et al. (Rahmani et al. 2010) propose an adaptive IHS (AIHS)-based method, which uses the edges of the PAN image to calculate the linear combination coefficients ($\alpha_i$) of the MS bands. The following section will illustrate how to calculate the adaptive coefficients $\alpha_i$.

Another CS technique is the PCA-based method (Sun et al. 2006). Being similar with the IHS-based method, PCA rotates the multidimensional image space to the two-dimensional plane (new vector space). Then, different components of the MS image are reordered by computing the eigenvectors of the covariance matrix of the new vector space. Finally, the first principle component of the MS image is replaced by the PAN image. However, the spectral feature of the first principle component is not exact with the PAN image. Hence, it will lose some spectral information of the MS image, leading to spectral distortion of the fused image.

The GS-based method is also a typical CS method. This method uses the knowledge of linear algebra and multivariate statistics. According to (Aiazzi et al. 2007), the main operations of the GS-based method can be summarized as: (1) up-sampling the MS image at the same scale of the PAN image; (2) subtracting the corresponding mean values from each MS image band, and the operation for the PAN image; (3) generating the low resolution version intensity ($I_L$) by using the common IHS-transform; (4) obtaining the difference (detail information) between $I_L$ and the PAN image by the histogram matching; (5) calculating the weight coefficients about the detail information injected in the MS image. Actually, the GSA-based method is an extension of the GS method just adaptively selecting the weight coefficients of the IHS-transform. Although these CS methods can reduce computation load, the results sometimes suffer from spectral distortion.

To solve spectral distortion, various pansharpening methods based on multi-resolution-analysis (MRA) are presented. The MRA-based methods, typically include Laplacian pyramid (LP) (Aiazzi et al. 1999), wavelet transform (WT) (Nunez et al. 1999; Jin et al. 2014), and contourlet transform (CT) (Upla et al. 2013; Choi et al. 2005), etc. The missing detail information of the MS image can be obtained by computing the difference between the PAN image and its low-pass version, which can be obtained by an iterative wavelet decomposition. However, results of these methods might generate spatial blurring phenomena, which leads to reducing the quality of the fused image.

According to Eq. (1), the larger correlation between the PAN image and its approximation will introduce the lower distortion of the fused image (Thomas et al. 2008). However, for the PAN image and its approximation, the major difference lies in their high-frequency parts (Song et al. 2016). Therefore, we can remove the influence of low-frequency parts in processing, while Eq. (1) can be rewritten as:

$$HMS_i = MS_i + w_i\left(PAN^H - APP^H\right),\ i = 1, ..., N \tag{3}$$

where $HMS_i$ indicates the pansharpening result. $MS_i$ denotes the $i$th band of the MS image. $PAN^H$ is the high-frequency (HFC) of the PAN image. $w_i$ is the injection gain coefficient. $N$ is the number of the bands of the MS image. $APP^H$ is the HFC of the PAN approximation, which can be given by:

$$APP^H = \sum_{i=1}^{N} \alpha_i MS_i^H \tag{4}$$

where $\alpha_i$ is the $i$th-band weight coefficient. $MS_i^H$ is the HFC of the MS image. Therefore, there still exist two improvements for the existing pansharpening methods. (1) how to efficiently eliminate the low-frequency-component (LFC) of PAN and MS images. (2) how to narrow the discrepancy between $APP^H$ and $PAN^H$ to precisely extract the detail images.

In this paper, a new CS-based pansharpening method is proposed. (1) We present bilateral filter (BF) to obtain the LFC of PAN and MS images to eliminate the influence of the low-frequency parts. As mentioned above, detail extraction depends on the difference of the HFC of the PAN image and its approximation (Song et al. 2016). BF can preserve the edge information of the original images when smoothing an image. (2) To narrow the discrepancy between the HFC of intensity channel and PAN images means enhancing the correlation between $APP^H$ and $PAN^H$ (Thomas et al. 2008). we use the adaptive intensity-hue-saturation (AIHS)-based method to create the $APP^H$. The AIHS-based method can adaptively find the best weight coefficients ($\alpha_i$) to obtain the approximation ($APP^H$). (3) We further use the dual-scale guided image filter (GIF) to let the $APP^H$ approximate the $PAN^H$ as close as possible, since the result of GIF is an approximation of the input image. This process could further minimize the discrepancy between $PAN^H$ and $APP^H$ to obtain the precise detail images. (4) The detail images are transferred into the MS image to get the HMS image. Experimental results show that the method we proposed can efficiently reduce spectral distortion and avoid the local artifacts.

Our method provides following three contributions:

1. BF is used to calculate the LFC of PAN and MS images.
2. AIHS-based method is used to estimate the approximation of the HFC of PAN image.
3. dual-scale GIF is performed to obtain the missing detail of the MS image.

The structure of this paper is as follows. In Sect. 2, BF, GIF, and AIHS-based method will be introduced briefly; In Sect. 3, the proposed method is illustrated in particular; The experimental results are discussed in Sect. 4; At last, Sect. 5 summarizes our work.

## 2 Related work

### 2.1 Bilateral filter

Bilateral filter (BF) is a nonlinear filter, which possesses the property of preserving the edges when smoothing an image (Tomasi and Manduchi 1998). The filter can be applied to various image processing areas, including multi-scale decomposition of images and image denoising. Compared with other edge-preserving filters, such as weighted least squares (WLS) filter, the BF can preserve the edges in space and range domain simultaneously. As shown in Fig. 1, the space and range kernel can be obtained by computing the Gaussian weights among these pixels and the Gaussian weights among the intensity values of these pixels, respectively. Then, the two Gaussian kernels will be multiplied to obtain the bilateral kernel. Finally, the input image is smoothed by using the bilateral kernel.

Thus, it can achieve a balance between the blurring and sharpening by adjusting the space and range kernels, since the space kernel determines the spatial information contributed to the resultant image while the range kernel controls the edge information of the result. In our work, we

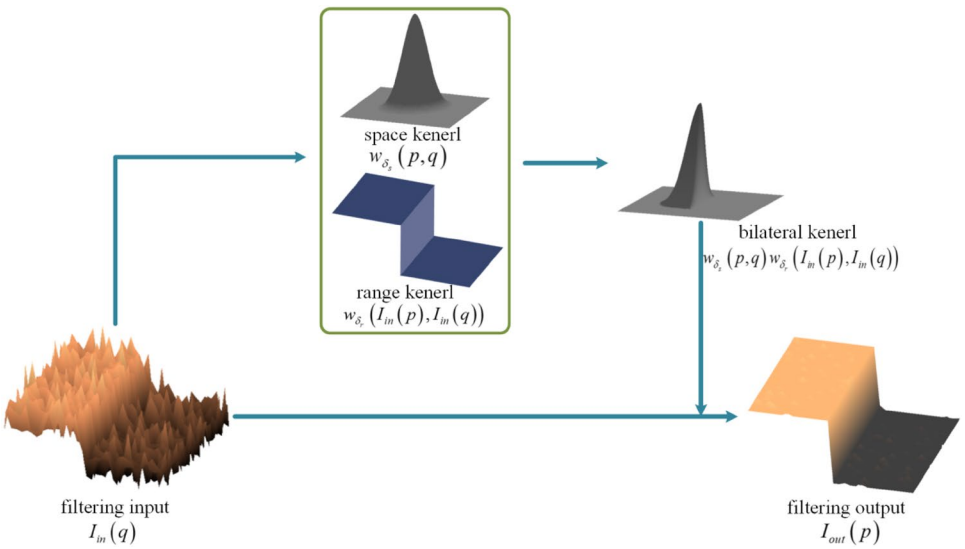


**Fig. 1** Diagram of the bilateral filter

adopt BF as a low-pass filter to obtain the LFC of PAN and MS images. The BF is described in (Tomasi and Manduchi 1998), whose mathematical equation is expressed as:

$$I_{out}(p) = \frac{1}{U}\sum_{q\in S} w_{\delta_s}(p,q)w_{\delta_r}\left(I_{in}(p), I_{in}(q)\right)I_{in}(q) \tag{5}$$

where $I_{out}(p)$ is the output pixel intensity at the central pixel $p$. $U = \sum_{q\in S} w_{\delta_s}(p,q)w_{\delta_r}\left(I_{in}(p), I_{in}(q)\right)$ is a normalization factor. $w_{\delta_s}(p,q)$ is the space weight while the $w_{\delta_r}\left(I_{in}(p), I_{in}(q)\right)$ is the range weight. $I_{in}$ denotes the input image. $I_{in}(p)$ denotes the pixel intensity at the central pixel $p$, while $I_{in}(q)$ represents intensities at the neighborhood pixels $S$. $\delta_s$ and $\delta_r$ are the standard deviations of the space Gaussian kernel and the range Gaussian kernel, respectively. The higher values of $\delta_s$ and $\delta_r$ indicate a more blurred effect. where the Gaussian kernel $w_\delta$ is given by:

$$w_\delta(x,y) = \frac{1}{\delta\sqrt{2\pi}} exp\left(-\frac{\|x-y\|^2}{2\delta^2}\right) \tag{6}$$

In this paper, the operator of bilateral filtering can be defined as follows:

$$I_{out} = BF\left(I_{in}, \delta_s, \delta_r\right) \tag{7}$$

where $BF(\cdot)$ represents bilateral filter function. By changing the parameters $\delta_s$ and $\delta_r$, we can obtain different blurry versions of the input image.

## 2.2 Guided image filter

The guided image filter (GIF) is not only an edge-preserving smoothing filter but also a local linear model based filter, which is first proposed by He et al. (2013). This filter can be applied into many image processing occasions, for example, image fusion and detail enhancement. GIF assumes that the output image is a linear representation of the guidance image in a local window. To narrow the discrepancy between the HFC of PAN and intensity images, we assume that the HFC of PAN should be locally and linearly represented by the HFC of intensity image in a local window. Therefore, the GIF is very suitable for this filtering process, which is a compromise between two constraints. One is the linear relationship of output image and guidance image (see Eq. 8). The other is that the filtering output should be as similar as possible to the input image (see Eq. 9). Balancing these two constraints can efficiently suppress the spectral distortion.

The diagram of the guided image filter is shown in Fig. 2. It can be seen that the process of the GIF is similar with a low-pass filter, which is filtering an image by a guidance image. The filtering output ($I_{out}$) is approximation of the filtering input ($I_{in}$) with removing noise ($n(q)$). At the same time, the guided image filter can well preserve the edge or structure information in the guidance image ($G$).

Mathematically, given an input image $I_{in}$ and a guidance image $G$, the output image $I_{out}$ can be linearly transformed by $G$ at pixel $p$ in a window $n_p$. This can be expressed as:

$$I_{out}(q) = a_p G(q) + b_p \ \forall q \in n_p \tag{8}$$

where $n_p$ is size of $(2r+1)\times(2r+1)$ (a square window of a filter radius $r$). The filter radius $r$ controls the blurring degree, and the small $r$ value will lead to a sharpened result. $a_p$ and $b_p$ are the linear coefficients, which are constants in window $n_p$. Through minimizing the squared difference $E\left(a_p, b_p\right)$ between the $I_{out}$ and $I_{in}$, these constants can be estimated.

$$E\left(a_p, b_p\right) = \sum_{q\in n_p}\left(\left(a_p G(q) + b_p - I_{in}(q)\right)^2 + \gamma a_p^2\right) \tag{9}$$

where $\gamma$ is a regulation factor, which is an important parameter to obtain edge-preserving effect. More specifically, if the variance of a local window much smaller than the parameter $\gamma$, the pixels in this window will be smoothed, whereas those

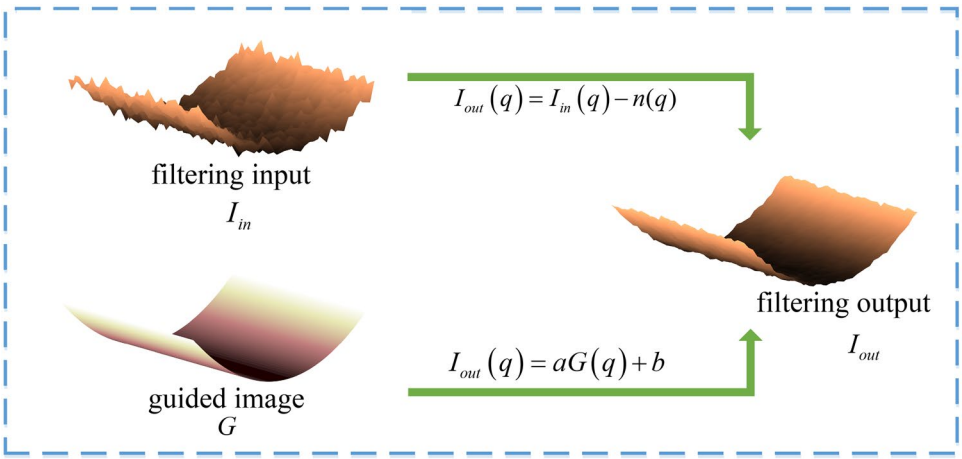


**Fig. 2** Diagram of the guided image filter

pixels are preserved. $a_p$ and $b_p$ can be directly calculated by linear regression as follows:

$$a_p = \frac{\frac{1}{|n|}\sum_{q\in n_p} I_{out}(q)I_{in}(q) - \mu_p \overline{I_{in_p}}}{\delta_p^2 + \gamma} \tag{10}$$

$$b_p = \overline{I_{in_p}} - a_p \mu_p \tag{11}$$

Here, in a local window $n_p$ of $G$, $\delta_p^2$ denotes the variance while $\mu_p$ denotes the mean, $|n|$ represents the number of pixels. And $\overline{I_{in_p}}$ indicates the mean of $I_{in}$ in window $n_p$.

To avoid the various values of output image $I_{out}$ computed in different windows, all the obtained values of $a_p$ and $b_p$ are first averaged. Thus, the final guided image filter can be rewritten as follows:

$$I_{out}(q) = \overline{a_p} G(q) + \overline{b_p} \tag{12}$$

where $\overline{a_p} = (1/|n|)\sum_{q\in n_p} a_p$ and $\overline{b_p} = (1/|n|)\sum_{q\in n_p} b_p$.

In this paper, the below equation is defined to express the guided filtering procedure

$$I_{out} = GIF\left(I_{in}, G, r, \gamma\right) \tag{13}$$

where $GIF(\cdot)$ denotes the guided image filter function.

### 2.3 AIHS-based method

The traditional intensity-hue-saturation (IHS)-based fusion methods are often written as:

$$F_i = MS_i + w_i(P - I) \tag{14}$$

where $F_i$ denotes the final HMS image. $MS_i$ presents the *i-th* band of the MS image. $w_i$ is the gain coefficient for each band. $P$ is the PAN image. $I$ represents the intensity channel calculated through a linear transform of the MS image as follows:

$$I = \sum_{i=1}^{n} \alpha_i MS_i \tag{15}$$

where $n$ denotes the number of spectral bands of the MS image. $\alpha_i$ represents the weight coefficients which can be computed by kinds of algorithms. In order to make $I$ component be close to the approximation of the PAN image ($P$), Rahmani et al. (2010) presented a solving method by optimizing the following problem (Kalpoma et al. 2013).

$$\min_{\alpha_1,\ldots,\alpha_n} \left\| P - \sum_{i=1}^{n} \alpha_i MS_i \right\|^2 \; s.t.\; \alpha_1 \geqslant 0, \ldots, \alpha_n \geqslant 0 \tag{16}$$

where Eq. (16) can determine the image coefficients $\alpha_i$ adaptively.

## 3 Proposed method

Figure 3 shows the main scheme of the proposed method. The proposed method consists of three steps: (1) obtaining the HFC of MS and PAN images; (2) estimating the approximation of the HFC of PAN image; (3) extracting detail images to reconstruct the fused image.

Given the original MS and PAN images (we assume that the original MS image and PAN image are pretreated well, see the parameter setup in Sect. 4.1). First, the HFC of MS and PAN images are obtained by subtracting the corresponding LFC of MS and PAN images from the original images, respectively. Second, the AIHS-based method is used to generate the approximate image of the HFC of PAN image. Finally, the dual-scale GIF is utilized to estimate detail images via calculating the difference between the HFC of PAN and intensity images. Then we synthesize the original MS image with these detail images to reconstruct the HMS image. More details will be explained in the following sections.

### 3.1 Obtaining the HFC of MS and PAN images

We use BF as the low-pass filter rather than other edge-preserving filters through a series of comparative experiments. Since BF achieves better results in retaining edge details, which are tuned to perception of human. According to the Eq. (7), the LFC of MS and PAN images can be calculated by:

$$MS_i^L = BF\left(MS_i, \delta_s, \delta_r\right) \tag{17}$$

$$PAN^L = BF\left(PAN, \delta_s, \delta_r\right) \tag{18}$$

where $BF(\cdot)$ denotes BF function. $MS_i$ and $PAN$ represent the original images. $\delta_s, \delta_r$ are the space standard deviation and range standard deviation of BF, respectively. $MS_i^L$ and $PAN^L$ denote the LFC of MS and PAN images, respectively.

Thus, the HFC of MS and PAN images can be easily obtained from the difference between original images and the corresponding $MS_i^L, PAN^L$, respectively. Thereby, we can get the $MS_i^H, PAN^H$ as follows:

$$MS_i^H = MS_i - MS_i^L \tag{19}$$

$$PAN^H = PAN - PAN^L \tag{20}$$

### 3.2 Estimating the approximation of the HFC of PAN

The existing methods for extracting the detail image are to compute the difference between intensity band and PAN

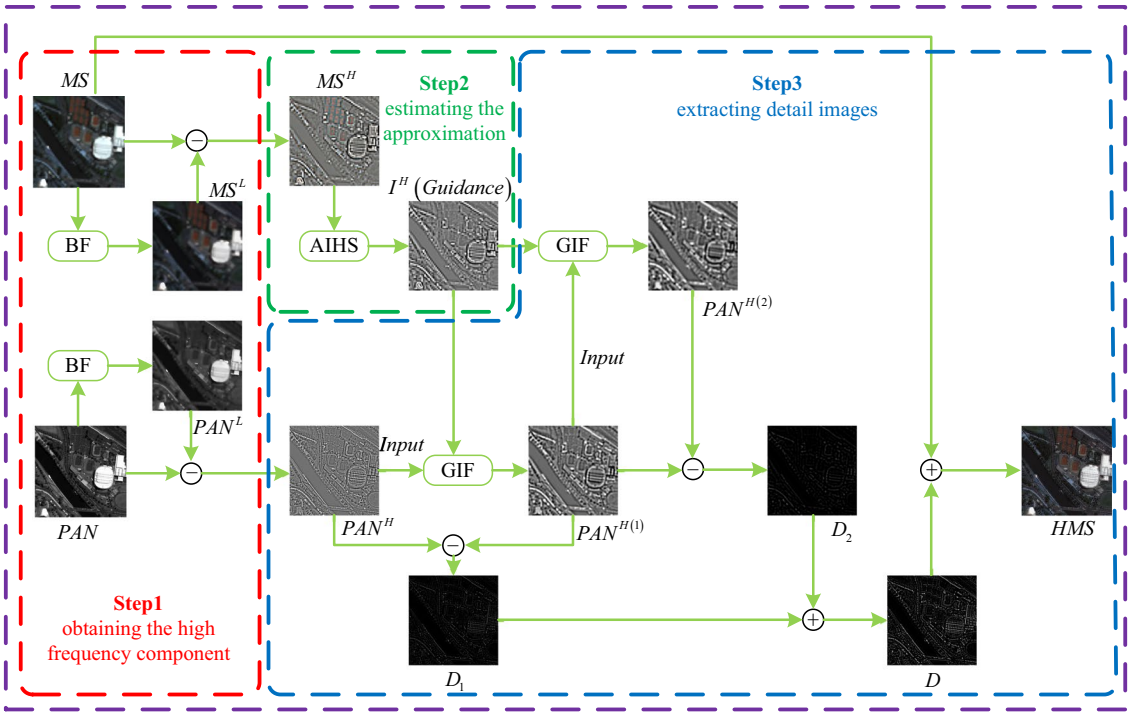


**Fig. 3** Scheme of the proposed method

image. However, the detail image will be introduced with LFC leading to spectral distortion. To avoid the undesired impact of LFC, we should directly estimate detail images from the HFC of intensity and PAN images.

According to Eq. (16), a coefficient-adaption IHS (AIHS)-based method was adopted, in which the HFC of intensity could be calculated closely to the approximation of the HFC of the PAN image by the following optimization function:

$$\min_{\alpha_1,\ldots,\alpha_N} \left\| PAN^H - \sum_{i=1}^{N} \alpha_i MS_i^H \right\|^2 \qquad s.t.\ \alpha_1 \geqslant 0, \ldots, \alpha_N \geqslant 0 \tag{21}$$

where $MS_i^H$ denotes the HFC of $i$th band of multispectral image. $PAN^H$ denotes the HFC of the PAN image. $\alpha_i$ is the coefficient of the $i$th band of the multispectral image.

Therefore, $I^H$ can be represented as the best approximate of $PAN^H$ with these adaptive coefficients $\alpha_1, \ldots, \alpha_N$.

$$I^H = \sum_{i=1}^{N} \alpha_i MS_i^H \tag{22}$$

where $I^H$ represents the HFC of the intensity.

### 3.3 Generating fused image

To further narrow the discrepancy between $PAN^H$ image and $I^H$ image efficiently, a novel method is proposed to use GIF to process the $PAN^H$ image. Actually, the GIF can minimize the difference between $PAN^H$ and $I^H$ images through local linear approximation. According to Eq. (13), this guided filtering first linearly represents $PAN^H$ by the guidance image $I^H$ in a local window, and then minimizes the cost function with regression coefficients obtained by linear regression, which shows a similar effect with the least square convergence. Therefore, we can get the low-resolution version of $PAN^H$ as follows:

$$PAN^{H(i)} = GIF\left(PAN^{H(i-1)}, I^H, r, \gamma\right), \quad i = 1, 2, \ldots, K \tag{23}$$

where $GIF(\cdot)$ denotes the GIF function. The $K$ is the scale factor. In this paper, the smooth degree can be determined by the filter size $r$ and regularization parameter $\gamma$. Obviously, the$PAN^{H(i)}$ image is more blurred than the $PAN^{H(i-1)}$ image. Thus, we can get the detail image through the inequality relationships, they are conveyed as follows:

$$D_i = PAN^{H(i-1)} - PAN^{H(i)},\ i = 1, 2, \ldots, K \tag{24}$$

where $PAN^{H(i-1)}$ is the original $PAN^H$ image when $i = 1$. $D_i$ is the spatial details between the two neighboring images of $PAN^{H(i)}$.

After obtaining the final detail images, we can transfer these detail images into the original MS image to get the final HMS image. The step consists of two parts. First, we should sum these detail images to get the total detail image.

$$D = \sum_{i=1}^{K} D_i \tag{25}$$

Second, put the total detail image into the original MS image to generate the high-resolution MS image as follows:

$$HMS_i = MS_i + D \tag{26}$$

## 4 Experimental results and analysis

To validate the effectiveness of the proposed method, test images are derived from three forms of satellite dataset, which contain WorldView-2, Pléiades, and QuickBird. The MS images (see Fig. 4) of Pléiades and QuickBird satellites contain four spectral bands, i.e., RGB and NIR (red band, green band ,blue band, and near infrared band). The WorldView-2 consists of eight spectral bands, that is, RGB, red-edge band (RE), yellow band (Y), coastal band (C), near infrared band (NIR1), and rear infrared band (NIR2), only four bands (R, G, B, NIR1) are used in our experiments. In the experiments, three pairs of PAN and MS images selected from these satellites are utilized. For diversity of the images, the scenes of these images include mountain area, vegetation area, and urban area.

### 4.1 Parameters setup

In our experiments, both original MS and PAN images captured by satellites are size of $256 \times 256$ pixels and $1024 \times 1024$ pixels, respectively. First, the original MS and PAN images are downsampled in a lower resolution by bicubic interpolation with a downsampling factor 1/4. The degraded MS ($64 \times 64$ pixels) and PAN ($256 \times 256$ pixels) images are served as the input images. Due to no real HMS image for reference, the downsampled MS image is pan-sharpened back to the original size as the output HMS image, and the original MS image is regarded as a reference HMS image for visual comparison and objective measurements.

In our experiments, for bilateral filter, we set a fixed range standard deviation $\delta_r$, determining the space standard deviation $\delta_s$ to enhance the ERGAS value and decrease the SAM

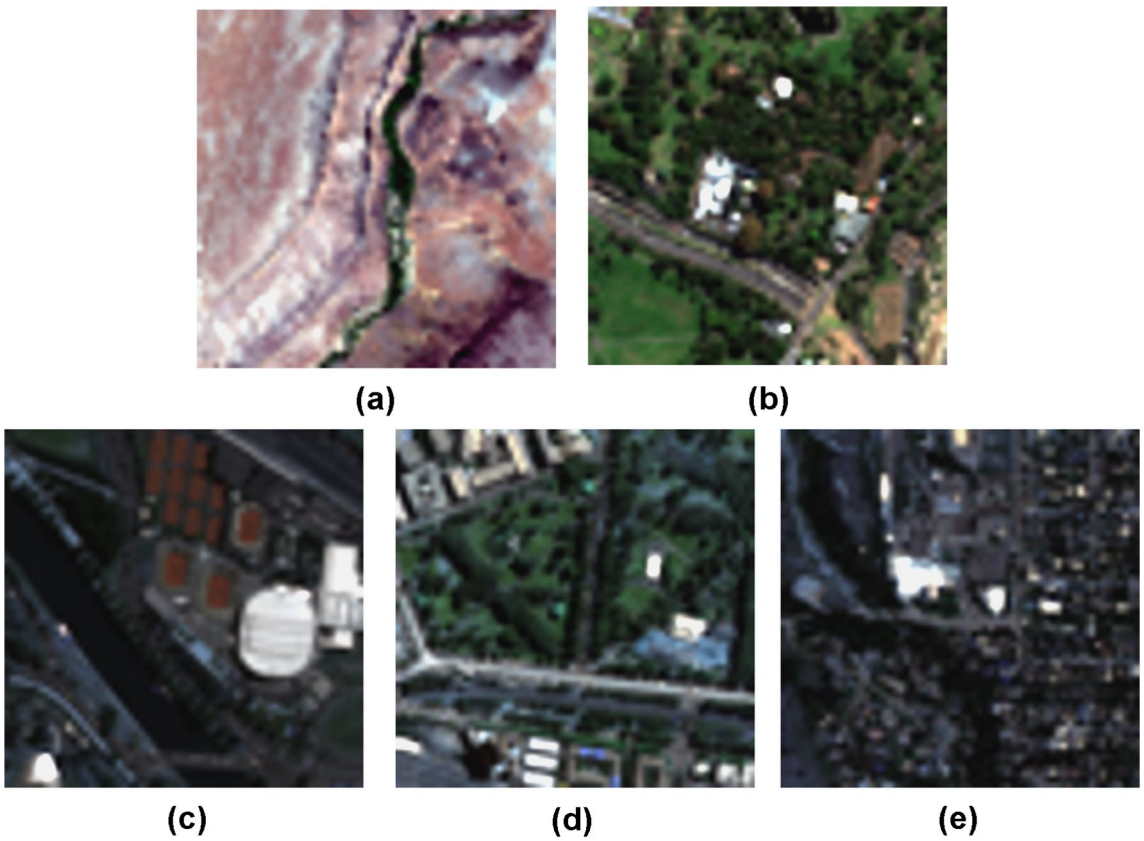


**Fig. 4** The satellite images (degraded images) used in this paper. **a, b** WorldView-2 image. **c, d** Pléiades image. **e** QuickBird image

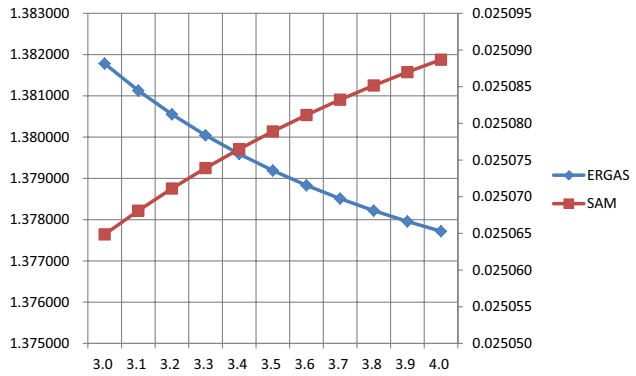


**Fig. 5** The selection of the parameter $\delta_s$

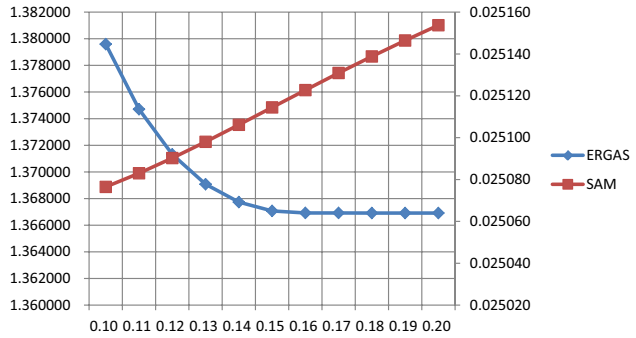


**Fig. 6** The selection of the parameter $\delta_r$

value, simultaneously. Since the small ERGAS value indicates the good overall performance of the result, and the small SAM value shows the less spectral angle distortion. As shown in Fig. 5, $\delta_s = 3.4$ is quite reasonable for the tradeoff between ERGAS and SAM values. For the range standard deviation $\delta_r$, we can also take the control variate method to select a proper parameter. From the Fig. 6, the ERGAS and SAM values are in a balance when we set $\delta_r = 0.12$.

In the detail images extracted step, there are two key parameters for GIF. One is the scale of guided filtering the $PAN^H$ image, which is set $K = 2$(dual-scale). The parameters with respect to GIF are filter radius $r = 2$, and regularization factor $\gamma = 0.01$. According to (He et al. 2013), the first experimental figure shows us the selection of the filter radius size $r$ and the regularization parameter $\gamma$. From the experiment and the description of this paper, the filter radius controls the filtering window size, and the large filtering window will weaken the linear relation between the guidance image and the filter output. Therefore, we select a small filter radius $r = 2$ to ensure this linear relation and prevent the filter output (should be approximated with the filter input) from over-smoothing. The small regulation parameter determines the small textures preserved in an image. Considering these two factors, we select the two small parameters [first experimental figure in He et al. (2013)]. The parameters are set as same values for each test image.

Moreover, the proposed method will be compared with six traditional pansharpening methods, i.e., the AIHS method (Leung et al. 2014), the BFLP method (Kaplan and Erer 2014), the GSA method (Aiazzi et al. 2007), the PCA method, the BDSD method (Garzelli et al. 2008) and AWLS method (Song et al. 2016). The parameters of these methods are given by authors, since we can search these available methods online.

## 4.2 Quality evaluation

To quantitatively compare various existing methods, we select six global metrics (with reference image) and three metrics (without reference image) to evaluate objective performance of different pansharpening methods. We use F to denote the fused result, and R to denote the reference image. These metrics are defined as follows:

1. CC (correlation coefficient) (Alparone et al. 2007) measures the relevant degree between F and R. The F is closer to R when CC value is larger. The highest value for UIQI is 1.
2. SAM (spectral angle mapper) (Yuhas et al. 1992) computes the absolute value of the spectral angle between the vectors of F and R. A smaller SAM value means less spectral angle distortion. The smallest value for RMSE is 0.
3. RMSE (root mean squared error) denotes the average squared deviation between F and R. A small RMSE represents good performance.
4. UIQI (universal image quality indexes) (Wang and Bovik 2002) evaluates the similarity of F to R. The performance becomes better when increasing the UIQI value.
5. ERGAS (relative dimensionless global error in synthesis) (Wald 2000) assesses the whole performance of F. A smaller ERGAS value indicates a better performance.
6. RASE (relative average spectral error) (Yang et al. 2016) shows the spectral average performance. The higher spectral quality has a lower RASE value.
7. QNR (quality with no reference) (Alparone et al. 2008) consists of a spatial distortion index $D_s$ and a spectral distortion index $D_\lambda$. The QNR value is determined by the spatial and spectral indexes. These three objective metrics are computed without a reference image. A small spatial distortion and spectral distortion portends a satisfying fusion result. However, the higher QNR value means the better fused effect.

### 4.3 Results and comparison

Experiments are first performed on WorldView-2 satellite image captured in mountain area. Figure 7a, b shows a MS image and a PAN image, respectively. Figure 7c shows the reference HMS image. Figure 7d–j shows the resultant HMS images generated by various methods. As shown in Fig. 7h, i, the results of PCA- and BDSD- based methods suffer serious spectral distortion. Obviously, it can be seen that the vegetation color in the close-up region of PCA-based method is close to black, and the color of rock surface is lighter than the reference HMS image (see Fig. 7h). The AIHS-based method produces a better effect in terms of spectral information, but generates few local artifacts (see Fig. 7d). And the result obtained by GSA-based method also has obvious spectral distortion in vegetation areas (see close-up view in Fig. 7g). The BFLP, AWLS, and the proposed method have highly similar visual effect, which is hard to distinguish (see Fig. 7e–j). Table 1 gives the objective quality metrics of seven different pansharpening methods. It can be seen that the proposed method achieves the best in most selected metrics except the metrics with no reference image. The BDSD-based method has good objective metrics in terms of QNR, $D_s$ and $D_\lambda$, since it directly fuses the high resolution panchromatic image and the up-sampled multispectral image (Garzelli et al. 2008). Generally, the objective evaluation is consistent with the subjective evaluation in terms of the proposed method. Therefore, we can conclude that the proposed method has the ability to obtain a good result, which is closer to the reference HMS image.

Another example is also conducted on WorldView-2 satellite image derived from urban areas. As shown in Fig. 8, the BFLP-, the GSA-, the PCA- and the BDSD- based methods are confronted with various degrees of spectral

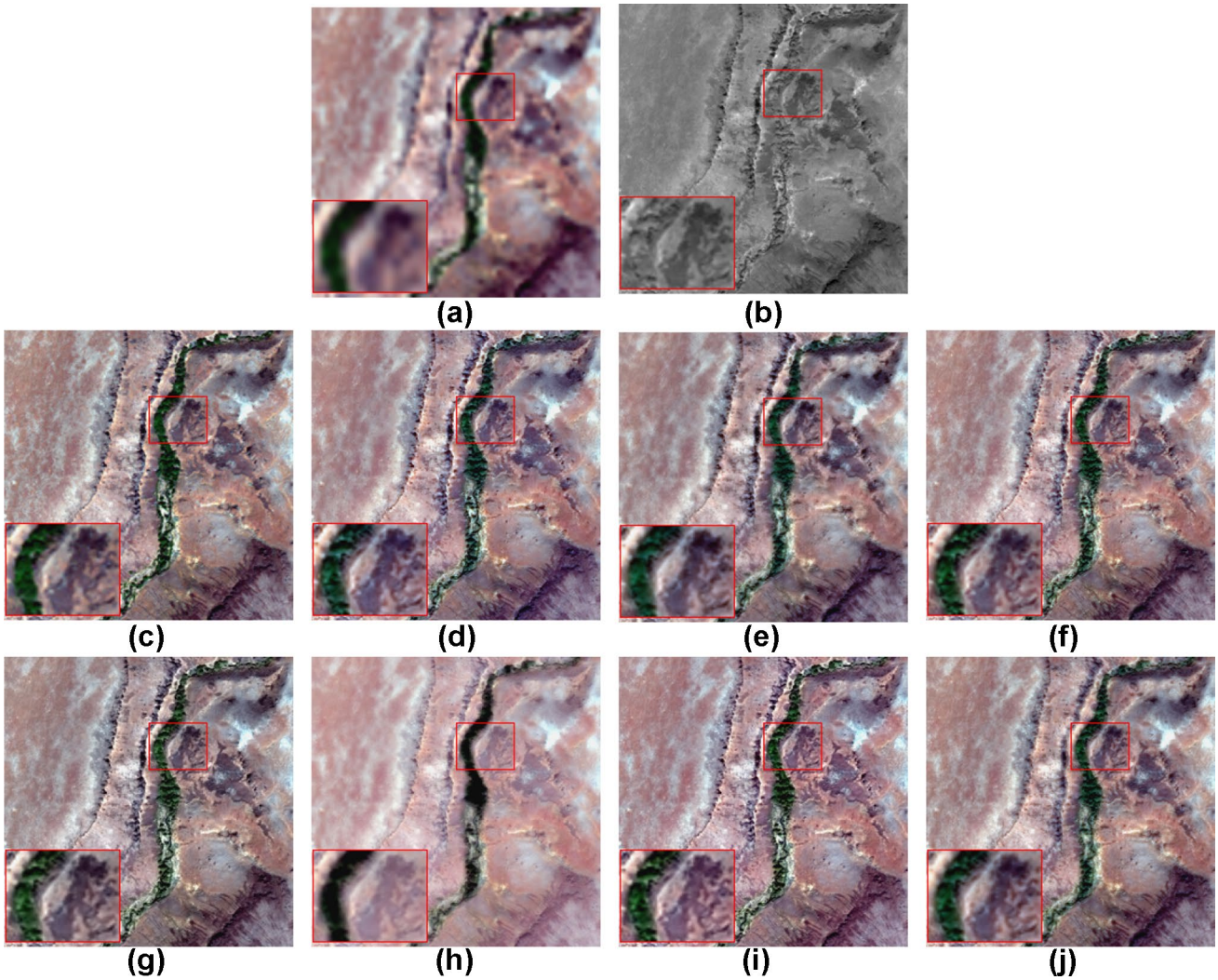


**Fig. 7** Pansharpening results of WorldView-2 satellite image: **a** MS image. **b** PAN image. **c** Reference image. **d** AIHS method. **e** BFLP method. **f** AWLS method. **g** GSA method. **h** PCA method. **i** BDSD method. **j** Proposed method

**Table 1** Objective quality metrics for Fig. 7

| Methods | CC (1) | ERGAS (0) | UIQI (1) | SAM (0) | RASE (0) | RMSE (0) | $D_s$ (0) | $D_\lambda$ (0) | QNR (1) |
|---|---|---|---|---|---|---|---|---|---|
| AIHS | 0.944 | 1.444 | 0.939 | 0.026 | 6.379 | 4.468 | 0.043 | 0.025 | 0.932 |
| BFLP | 0.948 | 1.522 | 0.947 | 0.028 | 6.877 | 4.817 | 0.037 | 0.066 | 0.899 |
| AWLS | 0.951 | 1.390 | 0.949 | 0.025 | 6.169 | 4.321 | 0.039 | 0.026 | 0.936 |
| GSA | 0.951 | 1.444 | 0.949 | 0.025 | 6.314 | 4.423 | 0.040 | 0.054 | 0.908 |
| PCA | 0.901 | 2.223 | 0.842 | 0.033 | 9.791 | 6.858 | 0.156 | 0.044 | 0.807 |
| BDSD | 0.935 | 1.782 | 0.929 | 0.031 | 7.821 | 5.478 | **0.019** | **0.005** | **0.976** |
| Proposed | **0.952** | **1.386** | **0.951** | **0.025** | **6.151** | **4.308** | 0.031 | 0.033 | 0.937 |

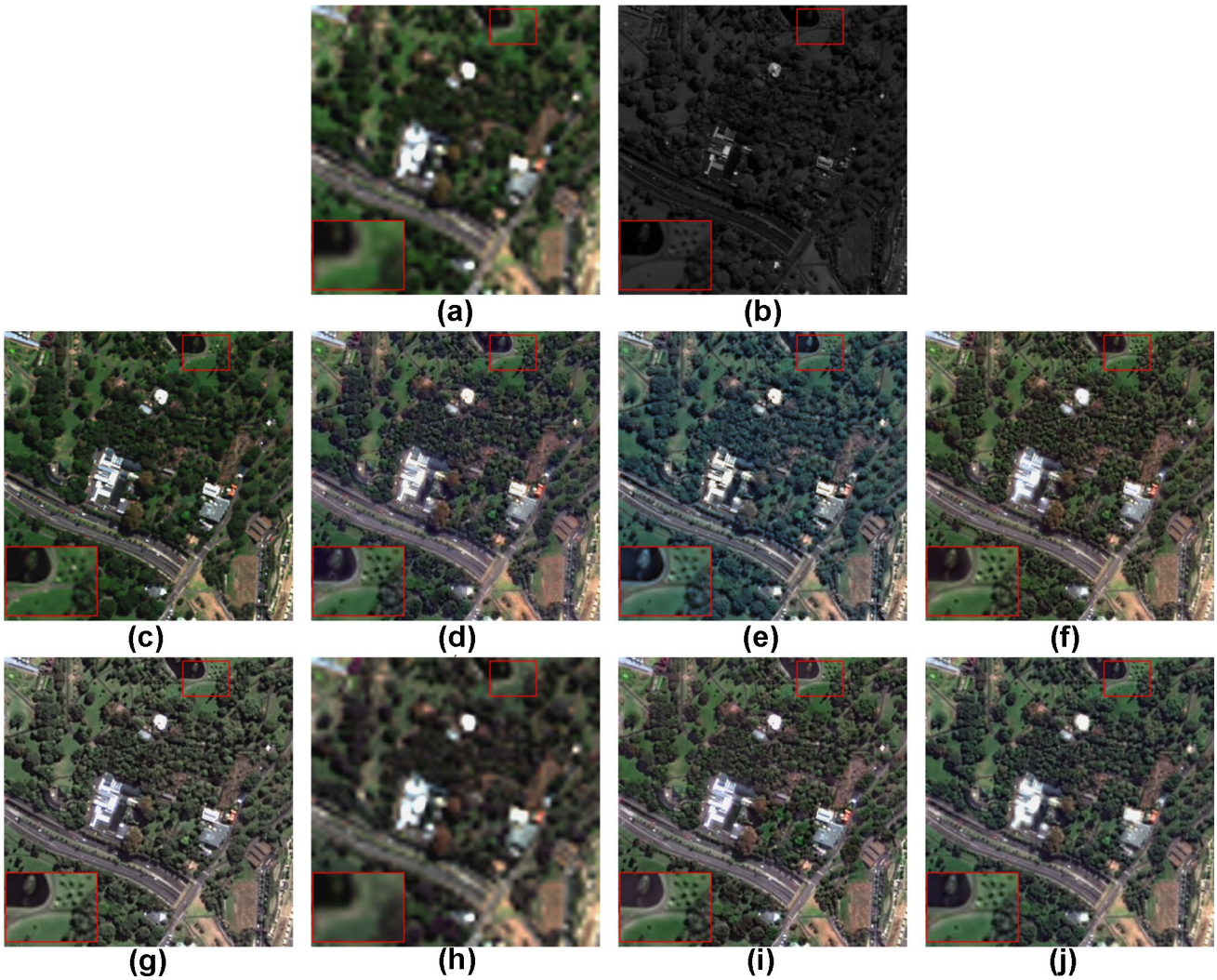


**Fig. 8** Pansharpening results of WorldView-2 satellite image: **a** MS image. **b** PAN image. **c** Reference image. **d** AIHS method. **e** BFLP method. **f** AWLS method. **g** GSA method. **h** PCA method. **i** BDSD method. **j** Proposed method

distortion (see Fig. 8e, g, h, i). Compared with the grass areas in close-up view of Fig. 8c, the BFLP- ,the GSA- based and the BDSD- based methods (see the close-up view of Fig. 8e, g, i) produce unnatural color, which is lighter than the reference HMS image. The PCA-based method performs poor performance on both spatial resolution and spectral resolution (see Fig. 8h). The spectral information can be well preserved by using the AIHS- ,the AWLS- and the proposed methods. However, the spatial detail information performs a little blur for the AIHS-based method (see Fig. 8d). The overall performance of the result obtained by using the proposed method is the closest to the reference image (see

Fig. 8j). As shown in Table 2, the objective quality is consistent with the subjective evaluation. The proposed method shows good objective effect in the aspect of CC, ERGAS, UIQI, RASE, RMSE, and $D_s$ indexes. Although the AWLS-based method has a little larger values for the $D_\lambda$, and QNR, our proposed method still ranks in the top three. Moreover, the SAM value is related with the spectral type of an image, therefore, the method obtaining the best value is not fixed. This example shows that our method can produce a satisfing result on pansharpening the WorldView-2 satellete images.

Experiments are also done on another type of satellite image derived from the Pléiades. As shown in Fig. 9,

**Table 2** Objective quality metrics for Fig. 8

| Methods | CC (1) | ERGAS (0) | UIQI (1) | SAM (0) | RASE (0) | RMSE (0) | $D_s$ (0) | $D_\lambda$ (0) | QNR (1) |
|---|---|---|---|---|---|---|---|---|---|
| AIHS | 0.865 | 4.642 | 0.849 | 0.096 | 21.768 | 65.075 | 0.037 | 0.012 | 0.952 |
| BFLP | 0.857 | 4.945 | 0.842 | **0.092** | 21.680 | 64.809 | 0.123 | 0.097 | 0.792 |
| AWLS | 0.874 | 4.625 | 0.866 | 0.098 | 21.339 | 63.790 | 0.032 | **0.010** | **0.958** |
| GSA | 0.849 | 5.601 | 0.843 | 0.103 | 24.841 | 74.260 | 0.137 | 0.098 | 0.779 |
| PCA | 0.787 | 6.295 | 0.735 | 0.176 | 35.816 | 107.070 | 0.266 | 0.151 | 0.623 |
| BDSD | 0.859 | 5.365 | 0.856 | 0.112 | 24.939 | 74.552 | 0.044 | **0.010** | 0.947 |
| Proposed | **0.875** | **4.534** | **0.866** | 0.098 | **21.030** | **62.867** | **0.026** | 0.026 | 0.949 |

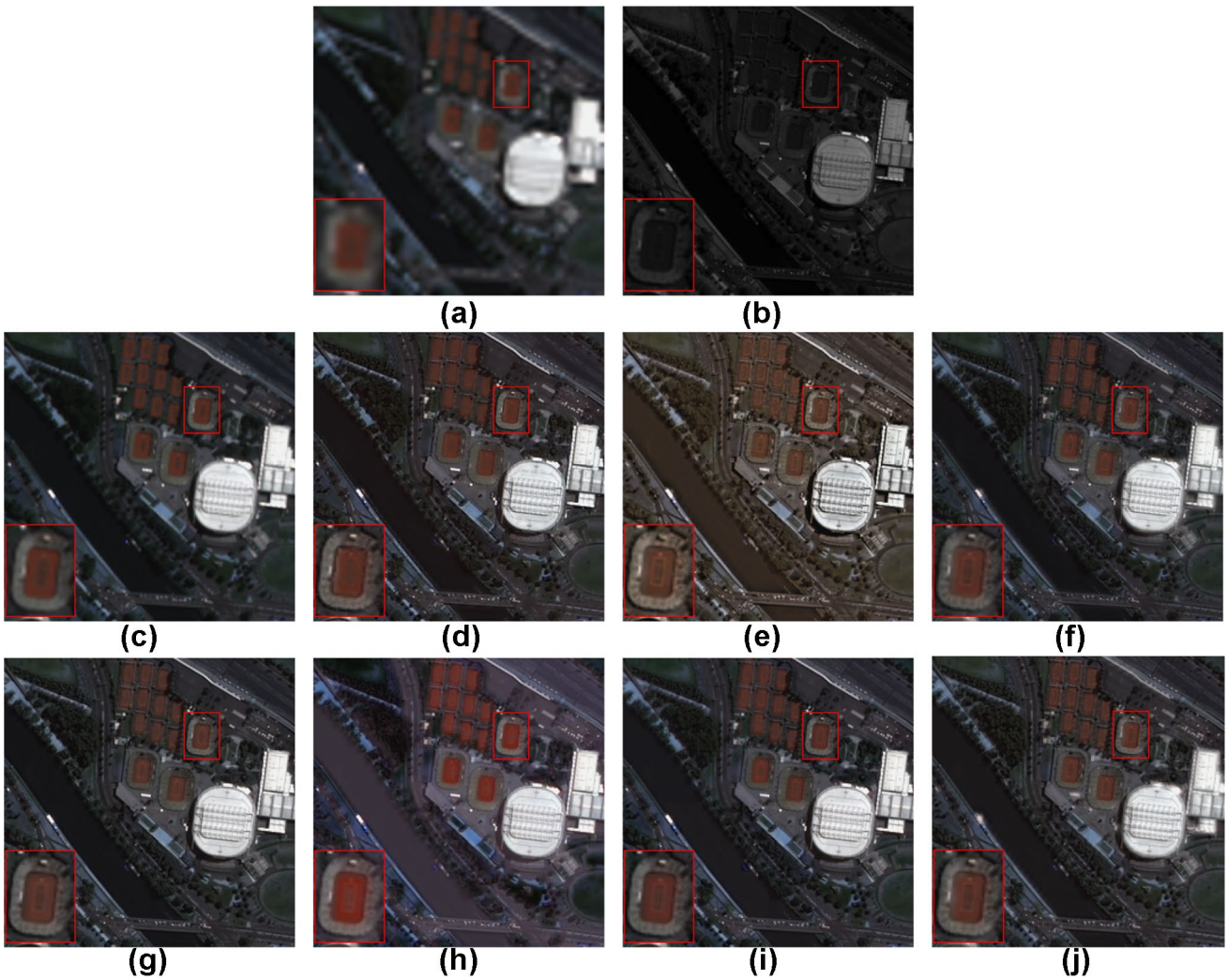


**Fig. 9** Pansharpening results of Pléiades satellite image: **a** MS image. **b** PAN image. **c** Reference image. **d** AIHS method. **e** BFLP method. **f** AWLS method. **g** GSA method. **h** PCA method. **i** BDSD method. **j** Proposed method

the BFLP- and the PCA- based methods produce obvious spectral distortion (see Fig. 9e, h). It can be seen that the AIHS-based method preserves the spectral information of the MS image to some extent. However, the spatial detail information of the PAN image cannot be transferred well into the sharpened image (see the close-up view in Fig. 9d). The remainder methods including the AWLS-, the GSA-, the BDSD- based, and the proposed methods show almost the same performance in terms of spectral and space information (see Fig. 9f, g, i, j). Table 3 compares the objective quality metrics among the different methods. It is observed that the proposed method shows the preferable effect in terms of CC, ERGAS, UIQI, RASE, RMSE, and $D_s$ metrics. Although the SAM metric is not the best one, it ranks only second to the AWLS-based method. As the same explaination above, the BDSD-based method can obtain good objective metrics in terms of $D_\lambda$ and QNR without reference HMS image. To sum up, this example demonstrates that our method is superior than other methods on pansharpening the Pléiades satellete images.

Furthermore, Fig. 10a, b shows the MS and PAN images of the Pléiades satellite, respectively. The visual results of BFLP-, GSA- and BDSD- based methods are confronted with different degrees of spectrum-distortion in the vegetation area (see Fig. 10e, g, i). From the magnified regions marked by a red rectangle, we can see that the PCA-based method (see Fig. 10h) even could not enhance spatial details for this example at all. The results (see Fig. 10d, f, j) obtained by AIHS-, AWLS-based methods and the proposed method perform very close to the reference MS image. The objective quality metrics for Fig. 10 listed in Table 4 are consistent with visual comparison. Although the BFLP-based method generates a slightly better SAM value than the proposed method, our method has better objective performance in sense of the most quality metrics, such as, CC, ERGAS, UIQI, RASE, RMSE, $D_s$ and QNR. For the $D_\lambda$, it varies among these different methods, since it is related with the spectral properties of the image itself. Generally, we can concluded that the presented method outperforms traditional methods in term of visual and objective evaluation.

To further validate the pansharpening method we proposed, we select QuickBird images with respect to urban areas. As shown in Fig. 11, it can be seen that the GSA- and PCA-based methods seriously distort the spectral information (see Fig. 11g, h), especially the GSA-based method is obvious without any spectral information. The BDSD-based method is fail in enhancing spatial details of left-bottom area in this example (see close-up view of Fig. 11i). The BFLP-based method produces result whose color is totally different with the reference image (see Fig. 11e). The AIHS-, AWLS-based methods and the proposed method (see Fig. 11d, f, j) are most similar with each other. The objective quality metric values for this example are given in Table 5. Although the proposed method achieves not better performance in the term of SAM index (a little lower than the AIHS-based method), it reveals best effect for remaining indexes, i.e., CC, ERGAS, UIQI, RASE, and RMSE. Therefore, the proposed method demonstrates the best performance for QuickBird images.

## 5 Conclusion

A novel pansharpening method by using guided image filter based on dual-scale detail extraction is proposed. Our method adopts the bilateral filter to remove influence of the low frequency parts of the input images before estimating the approximation. In order to enlarge the correlation between the approximation and the $PAN^H$ image, we use the AIHS-based method to obtain the approximation. Specially, we extract detail images with a dual-scale guided image filter, which can obtain more missing details of the MS image. Compared with other existing pansharpening methods, our method performs better on both subjective and objective indications.

**Table 3** Objective quality metrics for Fig. 9

| Methods | CC (1) | ERGAS (0) | UIQI (1) | SAM (0) | RASE (0) | RMSE (0) | c (0) | $D_\lambda$ (0) | QNR (1) |
|---|---|---|---|---|---|---|---|---|---|
| AIHS | 0.933 | 6.107 | 0.932 | 0.063 | 23.881 | 73.493 | 0.018 | 0.050 | 0.933 |
| BFLP | 0.895 | 9.342 | 0.872 | 0.068 | 37.835 | 116.436 | 0.124 | 0.123 | 0.769 |
| AWLS | 0.960 | 4.696 | 0.960 | **0.061** | 18.659 | 57.421 | 0.014 | 0.032 | 0.954 |
| GSA | 0.944 | 6.207 | 0.940 | 0.074 | 24.100 | 74.167 | 0.069 | 0.081 | 0.855 |
| PCA | 0.918 | 8.913 | 0.797 | 0.085 | 34.550 | 106.326 | 0.203 | 0.150 | 0.677 |
| BDSD | 0.941 | 6.126 | 0.939 | 0.088 | 24.668 | 75.913 | 0.034 | **0.008** | **0.958** |
| Proposed | **0.962** | **4.599** | **0.962** | 0.063 | **18.241** | **56.135** | **0.012** | 0.048 | 0.941 |

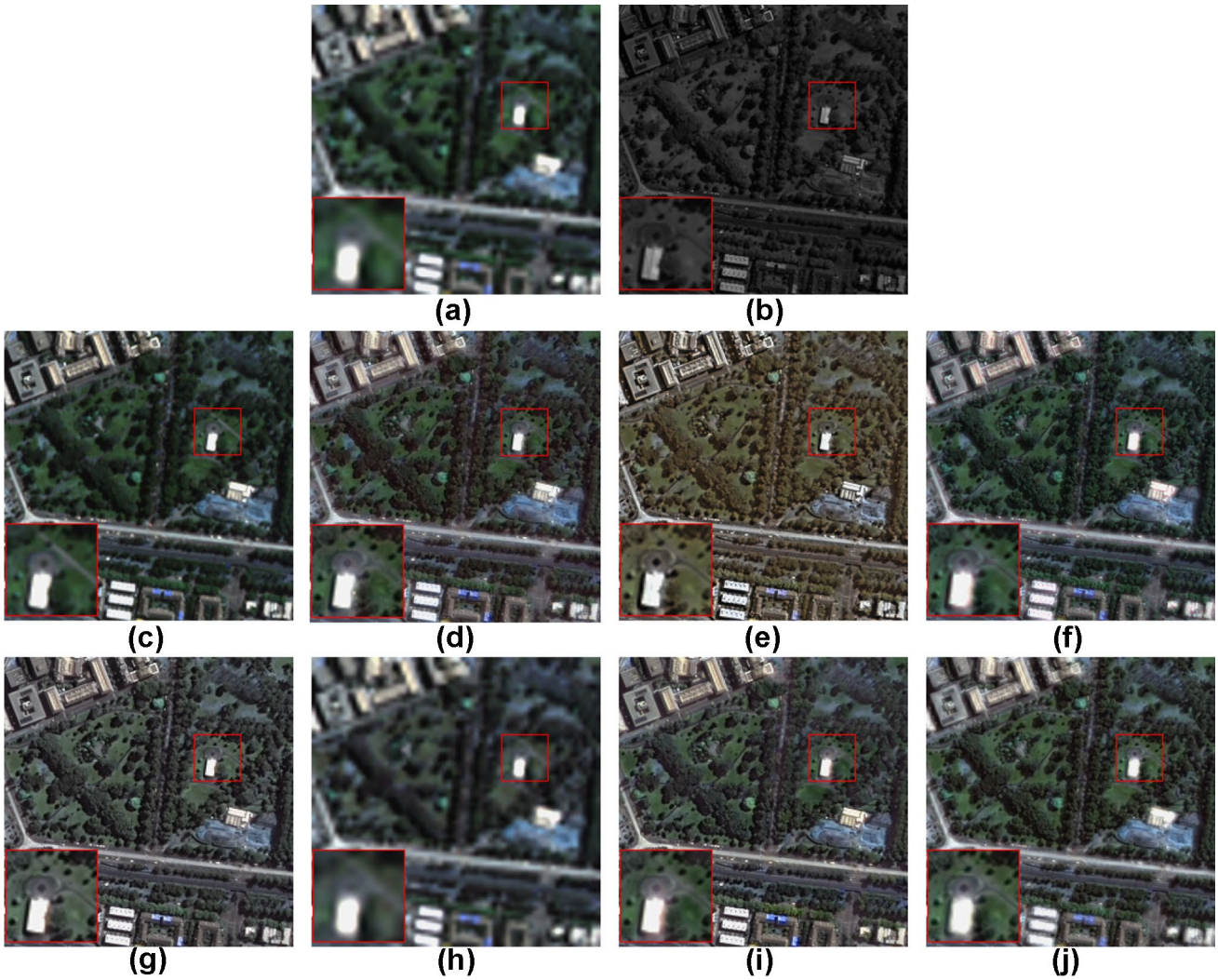


**Fig. 10** Pansharpening results of Pléiades satellite image: **a** MS image. **b** PAN image. **c** Reference image. **d** AIHS method. **e** BFLP method. **f** AWLS method. **g** GSA method. **h** PCA method. **i** BDSD method. **j** Proposed method

**Table 4** Objective quality metrics for Fig. 10

| Methods | CC (1) | ERGAS (0) | UIQI (1) | SAM (0) | RASE (0) | RMSE (0) | $D_s$ (0) | $D_\lambda$ (0) | QNR (1) |
|---|---|---|---|---|---|---|---|---|---|
| AIHS | 0.914 | 4.388 | 0.904 | 0.076 | 19.607 | 48.436 | 0.046 | **0.006** | 0.948 |
| BFLP | 0.907 | 4.868 | 0.896 | **0.071** | 20.929 | 51.703 | 0.166 | 0.123 | 0.731 |
| AWLS | 0.926 | 4.207 | 0.922 | 0.077 | 18.746 | 46.309 | 0.038 | 0.011 | 0.952 |
| GSA | 0.859 | 7.007 | 0.836 | 0.109 | 28.171 | 69.592 | 0.144 | 0.105 | 0.766 |
| PCA | 0.820 | 6.810 | 0.789 | 0.195 | 39.324 | 97.145 | 0.162 | 0.172 | 0.694 |
| BDSD | 0.874 | 6.358 | 0.861 | 0.114 | 27.729 | 68.500 | 0.055 | 0.013 | 0.933 |
| Proposed | **0.931** | **3.997** | **0.928** | 0.075 | **17.993** | **44.450** | **0.028** | 0.018 | **0.954** |

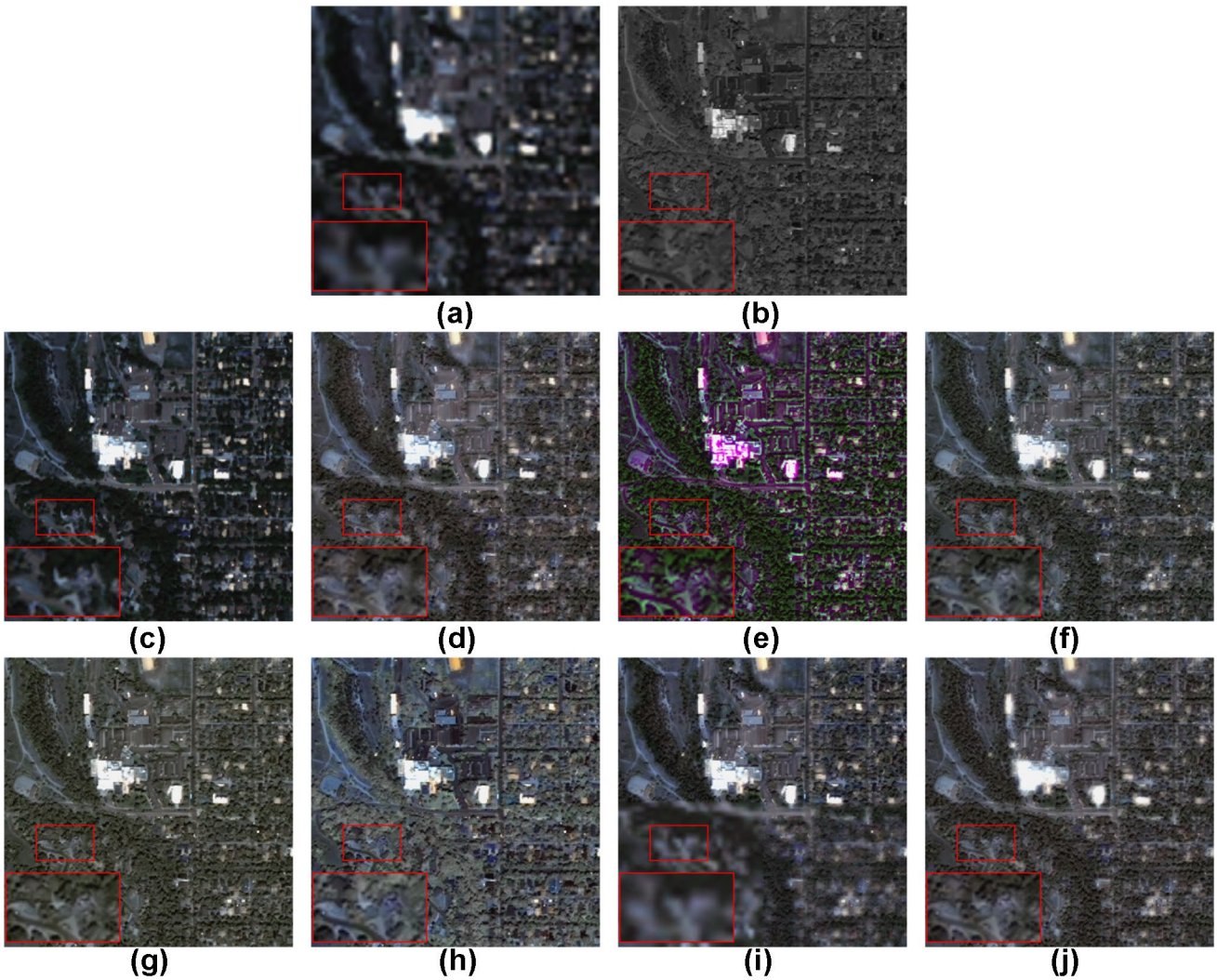


**Fig. 11** Pansharpening results of QuickBird satellite image: **a** MS image. **b** PAN image. **c** Reference image. **d** AIHS method. **e** BFLP method. **f** AWLS method. **g** GSA method. **h** PCA method. **i** BDSD method. **j** Proposed method

**Table 5** Objective quality metrics for Fig. 11

| Methods | CC (1) | ERGA S(0) | UIQI (1) | SAM (0) | RASE (0) | RMSE (0) | $D_s$ (0) | $D_\lambda$ (0) | QNR (1) |
|---|---|---|---|---|---|---|---|---|---|
| AIHS | 0.803 | 6.537 | 0.785 | **0.107** | 24.465 | 13.637 | 0.113 | 0.012 | 0.876 |
| BFLP | 0.727 | 21.329 | 0.673 | 0.440 | 85.199 | 47.489 | 0.153 | 0.264 | 0.623 |
| AWLS | 0.817 | 6.423 | 0.809 | 0.111 | 24.232 | 13.506 | 0.149 | 0.029 | 0.826 |
| GSA | 0.745 | 8.976 | 0.734 | 0.140 | 31.200 | 17.391 | 0.246 | 0.115 | 0.667 |
| PCA | 0.604 | 9.163 | 0.527 | 0.174 | 30.873 | 17.208 | 0.187 | 0.115 | 0.720 |
| BDSD | 0.809 | 7.114 | 0.804 | 0.123 | 27.814 | 15.503 | **0.081** | **0.007** | **0.912** |
| Proposed | **0.833** | **6.115** | **0.825** | 0.110 | **23.028** | **12.836** | 0.129 | 0.032 | 0.844 |

**Acknowledgements** The research in our paper is sponsored by National Natural Science Foundation of China (Nos. 61701327, 61711540303, 61473198), also is supported by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) Fund, Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology (CICAEET) Fund.

## References

Aiazzi B, Alparone L, Barducci A, Baronti S, Pippi I (1999) Multispectral fusion of multisensor image data by the generalized Laplacian pyramid. In: Geoscience and Remote Sensing

Symposium, 1999. IGARSS'99 Proceedings. IEEE 1999 International, 1999. IEEE, pp 1183–1185
Aiazzi B, Baronti S, Selva M (2007) Improving component substitution pansharpening through multivariate regression of MS $+ $ Pan data. IEEE Trans Geosci Remote Sens 45(10):3230–3239
Alparone L, Wald L, Chanussot J, Thomas C, Gamba P, Bruce LM (2007) Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest. IEEE Trans Geosci Remote Sens 45(10):3012–3021
Alparone L, Aiazzi B, Baronti S, Garzelli A, Nencini F, Selva M (2008) Multispectral and panchromatic data fusion assessment without reference. Photogram Eng Remote Sensing 74(2):193–200
Chen H-T, Eddy D, Chen R-L, Chou C-L (2016) Speed-adaptive street view image generation using driving video recorder. In: Multimedia and Expo (ICME), 2016 IEEE International Conference on, 2016. IEEE, pp 1–6
Choi M, Kim RY, Nam M-R, Kim HO (2005) Fusion of multispectral and panchromatic satellite images using the curvelet transform. IEEE Geosci Remote Sens Lett 2(2):136–140
Ferster CJ, Coops NC (2016) Integrating volunteered smartphone data with multispectral remote sensing to estimate forest fuels. Int J Digital Earth 9(2):171–196
Garzelli A, Nencini F, Capobianco L (2008) Optimal MMSE pan sharpening of very high resolution multispectral images. IEEE Trans Geosci Remote Sens 46(1):228–236
Gebru T, Krause J, Wang Y, Chen D, Deng J, Aiden EL, Fei-Fei L (2017) Using deep learning and google street view to estimate the demographic makeup of the us. arXiv preprint arXiv:170206683
Ghassemian H (2016) A review of remote sensing image fusion methods. Information Fusion 32:75–89
He K, Sun J, Tang X (2013) Guided image filtering. IEEE Trans Pattern Anal Mach Intell 35(6):1397–1409
Jameel A, Riaz MM, Ghafoor A (2016) Guided filter and IHS-based pan-sharpening. IEEE Sens J 16(1):192–194
Javidnia H, Corcoran P (2017) Real-time automotive street-scene mapping through fusion of improved stereo depth and fast feature detection algorithms. In: Consumer Electronics (ICCE), 2017 IEEE International Conference on, 2017. IEEE, pp 225–228
Jian L, Yang X, Zhou Z, Zhou K, Liu K (2018) Multi-scale image fusion through rolling guidance filter. Future Gen Comput Syst 83:310–325
Jin B, Kim G, Cho NI (2014) Wavelet-domain satellite image fusion based on a generalized fusion equation. J Appl Remote Sens 8(1):080599
Kalpoma KA, Kawano K, Kudoh J-i (2013) IKONOS image fusion process using steepest descent method with bi-linear interpolation. Int J Remote Sens 34(2):505–518
Kaplan NH, Erer I (2014) Bilateral filtering-based enhanced pansharpening of multispectral satellite images. IEEE Geosci Remote Sens Lett 11(11):1941–1945
Leung Y, Liu J, Zhang J (2014) An improved adaptive intensity–hue–saturation method for the fusion of remote sensing images. IEEE Geosci Remote Sens Lett 11(5):985–989
Liu J, Liang S (2016) Pan-sharpening using a guided filter. Int J Remote Sens 37(8):1777–1800
Miao Z, Shi W, Samat A, Lisini G, Gamba P (2016) Information fusion for urban road extraction from VHR optical satellite images. IEEE J Select Topics Appl Earth Obs Remote Sens 9(5):1817–1829
Nunez J, Otazu X, Fors O, Prades A, Pala V, Arbiol R (1999) Multiresolution-based image fusion with additive wavelet decomposition. IEEE Trans Geosci Remote Sens 37(3):1204–1211
Rahmani S, Strait M, Merkurjev D, Moeller M, Wittman T (2010) An adaptive IHS pan-sharpening method. IEEE Geosci Remote Sens Lett 7(4):746–750
Simone G, Farina A, Morabito FC, Serpico SB, Bruzzone L (2002) Image fusion techniques for remote sensing applications. Inf Fusion 3(1):3–15
Song Y, Wu W, Liu Z, Yang X, Liu K, Lu W (2016) An adaptive pansharpening method by using weighted least squares filter. IEEE Geosci Remote Sens Lett 13(1):18–22
Sun J, Jiang Y, Zeng S (2006) A study of PCA image fusion techniques on remote sensing. In: International conference on space information technology, 2006. International Society for Optics and Photonics, p 59853X
Tao F, Yang X, Wu W, Liu K, Zhou Z, Liu Y (2018) Retinex-based image enhancement framework by using region covariance filter. Soft Comput 22(5):1399–1420
Thomas C, Ranchin T, Wald L, Chanussot J (2008) 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
Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: Computer Vision, 1998. Sixth International Conference on, 1998. IEEE, pp 839–846
Tu T, Su S-C, Shyu H-C, Huang PS (2001) A new look at IHS-like image fusion methods, Information
Upla KP, Gajjar PP, Joshi MV (2013) Pan-sharpening based on Non-subsampled Contourlet Transform detail extraction. In: Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2013 Fourth National Conference on, 2013. IEEE, pp 1–4
Wald L (2000) Quality of high resolution synthesised images: is there a simple criterion? In: Third conference" Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", 2000. SEE/URISCA, pp 99–103
Wang Z, Bovik AC (2002) A universal image quality index. IEEE Signal Process Lett 9(3):81–84
Wu W, Yang X, Liu K, Liu Y, Yan B (2016) A new framework for remote sensing image super-resolution: sparse representation-based method by processing dictionaries with multi-type features. J Syst Architect 64:63–75
Yang Y, Wan W, Huang S, Yuan F, Yang S, Que Y (2016) Remote sensing image fusion based on adaptive IHS and multiscale guided filter. IEEE Access 4:4573–4582
Yuhas RH, Goetz AF, Boardman JW (1992) Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm
Zhang Y (2004) Understanding image fusion. Photogramm Eng Remote Sens 70(6):657–661