




Multispectral and panchromatic image fusion based on spatial consistency

Jiayuan Li , Qingwu Hu and Mingyao Ai

School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China

ABSTRACT

Multispectral (MS) and panchromatic (Pan) image fusion, which is used to obtain both high spatial- and spectral-resolution images, plays an important role in many remote-sensing applications such as environmental monitoring, agriculture, and mineral exploration. This article presents an image fusion framework based on the spatial distribution consistency. First, a YUV transform is adopted to separate the luminance component from the colour components of the original MS image. Then, the relationships between the ideal high-resolution multispectral (HRMS) colour components and the Pan band are established based on the spatial distribution consistency, and finally an inverse transform is employed to obtain the fused image. In this article, two types of relationship models are presented. The first model stems from the physical meaning of the assumption and uses a local linear model to describe it. The second model directly uses its algebraic meaning to design the objective cost function and obtains the global optimal solution. The proposed two models are compared with 15 other widely used methods on six real remote-sensing image data sets. Experimental results show that the proposed method outperforms the compared state-of-the-art approaches.



ARTICLE HISTORY

Received 9 March 2017

Accepted 11 October 2017

1. Introduction

With the rapid development of remote-sensing technology, it is often difficult to meet the needs of information extraction using single remote-sensing datum, which requires a variety of remote-sensing data for information fusion. For example, land-use classification, change detection, and map updates all require both high spectral- and spatial-resolution images. However, being limited by the imaging technology, satellites such as IKONOS, QuickBird, GeoEye, and WorldView-2 provide observation data that are generally composed of high spatial-resolution panchromatic (Pan) band and low spatial-resolution multispectral (MS) image (Jiang et al. 2014). Therefore, the advantages of different remote-sensing data should be synthesized to obtain both accurate spatial information and spectral information. Image fusion technology is an effective way to solve this problem, which has become an active topic in remote-sensing image processing.

CONTACT Qingwu Hu  huqw@whu.edu.cn  School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China

© 2017 Informa UK Limited, trading as Taylor & Francis Group

MS image and Pan band fusion, also called pansharpening, is a technique to merge the low spatial-resolution MS bands with the high spatial-resolution Pan band of the same satellite to produce a high spatial-resolution multispectral (HRMS) image. Several detailed surveys on pansharpening can be found in Wang et al. (2005), Thomas et al. (2008), Amro et al. (2011), and Vivone et al. (2015). Vivone et al. (2015) made a critical comparison among the widely used pansharpening approaches and grouped them into two categories, i.e. the component substitution (CS) methods and the multi-resolution analysis (MRA) methods. The main idea of the CS methods is to transform the MS image into other colour spaces and substitute a component such as luminance with the Pan image. CS methods generally consist of five steps: up-sampling, positive transformation, brightness matching, CS, and inverse transform (Amro et al. 2011). Approaches belonging to this category include band-dependent spatial detail (BDSD) (Garzelli, Nencini, and Capobianco 2008), Brovey transformation (Gillespie, Kahle, and Walker 1987), generalized intensity-hue-saturation (GIHS) (Tu et al. 2001), Gram–Schmidt (GS) spectral sharpening (Laben and Brower 2000), adaptive GS (GSA) (Aiazzi, Baronti, and Selva 2007), intensity-hue-saturation (IHS) (Carper, Lillesand, and Kiefer 1990), principal component analysis (PCA) (Chavez, Sides, and Anderson 1991), and partial replacement adaptive component substitution (PRACS) (Choi, Yu, and Kim 2011). In addition, several recently developed CS methods can be found in the literature (Xu et al. 2014). CS methods usually provide good visual effects and spatial quality. However, they easily distort the spectral information. The core idea of MRA methods is to inject the spatial details that are decomposed from the Pan image into the resampled MS bands. The spatial details can be obtained by several approaches: a trous wavelet transform (ATWT) (Vivone et al. 2014), ATWT using Model 2 (ATWT-M2) (Ranchin and Wald 2000), ATWT using Model 3 (ATWT-M3) (Ranchin and Wald 2000), additive wavelet luminance proportional (AWLP) (Otazu et al. 2005), decimated wavelet transform (DWT) (Mallat 1989), high-pass filtering (HPF) (Chavez, Sides, and Anderson 1991), decimated wavelet transform using an additive injection model (Indusion) (Khan et al. 2008), generalized Laplacian pyramid (GLP) (Aiazzi et al. 2002) with modulation transfer function (MTF) (MTF-GLP) (Aiazzi et al. 2006), and smoothing filter-based intensity modulation (SFIM) (Liu 2000; Wald and Ranchin 2002).

Wang, Shi, and Atkinson (2016) reviewed another group of algorithms for pansharpening, i.e. geostatistical solutions, and stated that geostatistical solutions have the advantage of preserving the spectral information. Downscaling co-Kriging (DSCK) (Pardo-Igúzquiza, Chica-Olmo, and Atkinson 2006) is a pansharpening algorithm for Landsat ETM+ images, which considers the low spatial-resolution MS image as the primary variable and the Pan band as the secondary variable. Spatially adaptive downscaling co-Kriging (SADSK) (Pardo-Igúzquiza et al. 2011) is an extension of DSCK, which adds a spatially adaptive filtering scheme to change the co-Kriging weights. Tang, Atkinson, and Zhang (2015) presented a post-processing step for DSCK, which considered multiple-point statistics to increase the pansharpening accuracy. Wang, Shi, and Atkinson (2016) proposed an area-to-point regression Kriging (ATPRK) method and its extension, adaptive ATPRK (AATPRK). ATPRK treats each HRMS band as a linear transformation of the Pan band and considers the residuals via area-to-point Kriging (ATPK) residual downscaling in the regression model for spectral compensation.

Variational methods have also received increasing attention. These methods usually build the objective cost function based on certain assumptions and obtain the optimal solution by minimizing the energy function. P+XS (Ballester et al. 2006) considers the up-sampled MS image as the blurred HRMS image and builds the objective function based on Brovey linear combination assumptions. Fang et al. (2013) proposed an iterative method to alternately solve the blur model parameters and the pansharpened image since the blurred kernel function of the P+XS (Panchromatic and MultiSpectral images) method is difficult to estimate. Alternate variational wavelet pansharpening (AVWP) (Möller et al. 2012) utilizes spectral ratio constraint to keep the spectral information, and obtains the spatial information by minimizing the P+XS energy function terms. These methods can preserve the spectral information well; however, the spatial information may not be well-preserved, which may cause severe block effects on the image edges. Chen et al. (2014) assumed that the fused image should be as consistent as possible with the up-sampled MS image, and took the local spectral continuity and gradient sparsity into account to obtain the global optimal solution. This method considers the similarity of spectral continuity and inter-band spatial gradient simultaneously, and utilizes an optimization method based on iterative threshold contraction. More recently, deep learning has also been applied for pansharpening. Huang et al. (2015) proposed the deep neural network (DNN)-based image fusion method with pre-training and fine-tuning stages. Masi et al. (2016) introduced a three-layer convolutional neural networks (CNN) proposed for super-resolution into the pansharpening task. Compared with the traditional CS and MRA methods, geostatistical solutions and variational methods have a common disadvantage in that the computational complexity is extremely high. Such high computational complexity may prevent their usages in real applications because pansharpening is usually a preprocessing step where tens of seconds of processing time for a 512×512 pixels image is unacceptable.

This article presents a spatial consistency assumption, which supposes that the spatial distribution of the HRMS image and the Pan band should be the same. That is, if the values of the adjacent pixels in the Pan image are similar, the values of the adjacent pixels in the MS image should be similar as well. Although Thomas et al. (2008) had mentioned spatial heterogeneity, it is undeniable that spatial consistency also exists in most of the satellite images, especially in high-resolution satellite images. In fact, spatial consistency has been widely used in image-fusion methods, such as in Fang et al. (2013), Chen et al. (2014), and Wang, Shi, and Atkinson (2016). They assumed that there was a linear combination or other relationships between the intensity or gradient of the MS image and the Pan band.

According to the physical and algebraic meaning of the spatial consistency assumption, two modelling methods are presented to express the relationship between the Pan and the MS image. The first method is based on a local linear model, which is computationally efficient. The second one finds the optimal solution of the global energy function, with a very good accuracy. Our local linear method is similar to ATPRK. However, there are several differences between them. First, ATPRK uses a linear transformation to model the relationship between each MS band and the Pan band. Differently, our method uses a more accurate model, i.e. the local linear model. Second, our article presents a spatial consistency assumption and verifies this assumption, which is the reason why we apply the local linear model. Third, ATPRK considers the

residuals via ATPK residual downscaling in their model, which will largely increase the computational complexity. For example, ATPRK takes 127 s on pansharpening of a 125×125 pixels MS image and a 500×500 pixels Pan band; in contrast, we solve the local linear model efficiently, similar to the guided filter. We compare the proposed two methods with 15 other widely used methods on six real remote-sensing image data sets. The experimental results show that the proposed methods outperform the compared state-of-the-art approaches.

2. Theory and method

2.1. The characteristics and verification of spatial consistency

Suppose that the MS image and the Pan band of the same satellite are accurately registered. We assume that the pixels at the same location in the HRMS image and the Pan band have similar spatial distribution, which is called spatial consistency. The definition of spatial consistency is as follows: if the values of the adjacent pixels in the Pan image are similar, the values of the adjacent pixels in the MS image tend to be similar as well. Spatial consistency has the following two properties.

- Local extreme position in the MS image corresponds to the one in the Pan image.
- Change trend of the adjacent pixels of the MS image is consistent with the one of the Pan image.

Figure 1 visually illustrates the spatial consistency prior, where each waveform represents the spatial distribution of a line of pixels. As can be seen, waveforms 1–3 have the same extreme positions and are consistent local monotonicity, so they conform to the spatial distribution consistency. In contrast, waveforms 4 and 5 do not satisfy the above-mentioned properties, so they are not subject to spatial consistency.

To verify the correctness of the spatial consistency, this article collected 300 satellite images with their corresponding Pan bands from Geoeye-1, Landsat 8 OLI, QuickBird, SPOT 6, Worldview-2, and Worldview 3 as the verification data sets. We upscale (upscale factor is set to 2 for Landsat 8 OLI and 4 for other satellites) the collected satellite

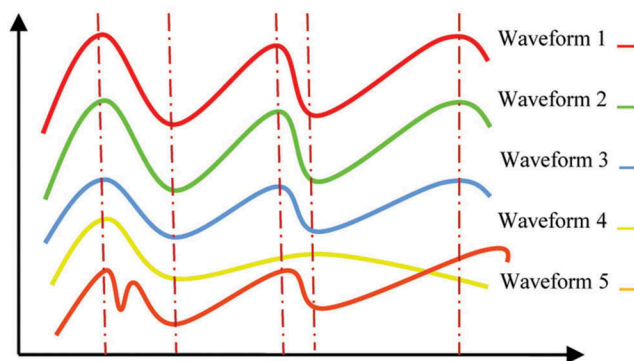


Figure 1. Illustration of spatial consistency.

images and Pan bands to obtain the MS images and the Pan images, respectively, and treat the original satellite images as the ideal HRMS images. First, we calculate the correlations among red (R), green (G), blue (B), and Pan bands. (The positive correlation between bands is the necessary condition of spatial consistency, i.e. if the spatial consistency is correct, the R, G, B, and Pan bands have positive correlations.) The experimental results are shown in Table 1. As can be seen, there are strong positive correlations among the R, G, B, and Pan bands (the average correlation coefficient is 0.897). Second, we randomly select 10 lines of pixels in each image and compare their pixel intensity values. We find that the spatial distributions of the R, G, B, and Pan bands are almost the same, which conform to the above-mentioned properties. Therefore, the spatial consistency is reliable and effective. Figure 2 shows the spatial consistency on some image rows. As shown, although the pixel intensity values of the R, G, B, and Pan bands in the same position may be different, the extreme positions and local monotonicity of the four bands are consistent.

2.2. The fusion framework based on spatial consistency

The proposed method is based on the YUV colour space, where Y represents the luminance component, and U and V represent the colour components. The reason why we use YUV colour space is just that YUV colour space can separate the luminance component from the colour components of an image. Note that other colour spaces

Table 1. Correlations between R, G, B, and Pan bands.

Image band	Pan &R	Pan &G	Pan &B
Correlation Coefficient	0.905	0.913	0.874
Standard Deviation	0.074	0.065	0.067

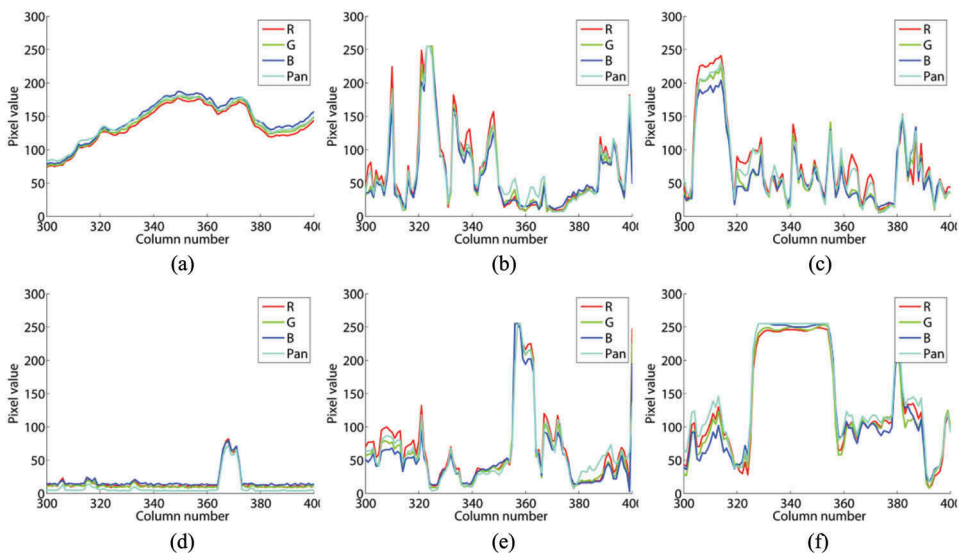


Figure 2. The spatial consistency between R, G, B, and the Pan band.

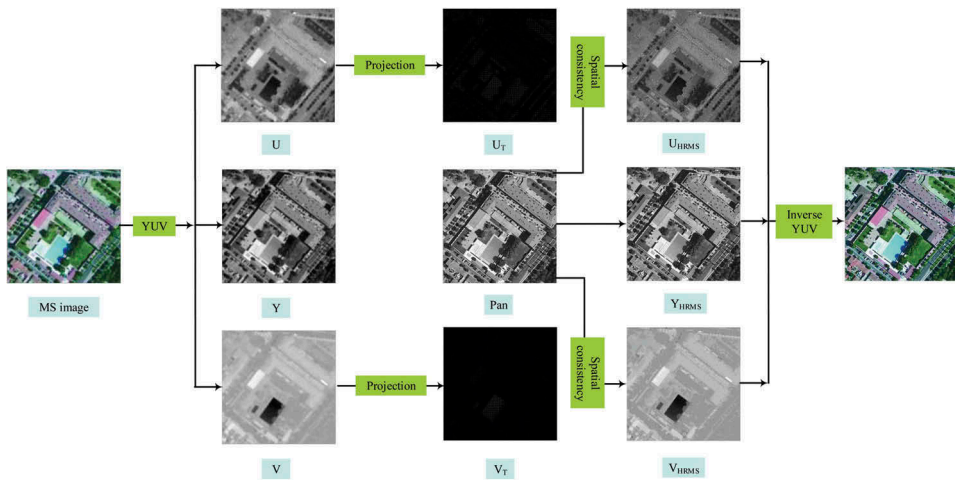


Figure 3. The flow chart of the proposed framework. (a) MS projection and (b) partial unknown values.

may also be applied. The major contribution of this article is the verification and modelling of the spatial consistency. The framework of the proposed method is given in Figure 3. First, we perform a YUV transform on the original MS image to obtain the colour channels U and V. Then, U and V are projected into channels U_T and V_T according to the spatial resolution ratio between the MS image and the Pan band, where U_T and V_T are the colour channels of the HRMS image with unknown values. Figure 4 (a) is the projected result of the original MS image according to spatial ratio 4. In this figure, the non-zero pixels (the interval is 4) come from the original MS image, and the zero pixels are the unknown values to be estimated. Based on the spatial consistency constraint, the remaining unknown values (e.g. the pixels with value 0 in Figure 4(b)) can be estimated according to the known values in U_T/V_T and the corresponding values of the Pan band (e.g. the pixels with value 1 in Figure 4(b)); then, the colour channels U_{HRMS} and V_{HRMS} of the HRMS images without unknown values can be estimated. Finally, the fused HRMS image is obtained by performing a YUV inverse transform.

In traditional CS methods, the colour components of the HRMS image are usually obtained by considering upscaling operation on the original MS image, which makes the colour components blurred, and results in spectral information loss. In contrast, the proposed method uses the spatial consistency constraint to exactly estimate the colour components of the HRMS image, which largely increases the spectral accuracy. Therefore, the core of our method is to establish the relationship between the HRMS image and the Pan image based on the spatial consistency assumption. We will provide two relationship modelling methods next.

2.3. Local linear model

Based on the guided filter (He, Sun, and Tang 2013) and the guided filter-based fusion (Li, Kang, and Hu 2013), we use a local linear model to represent the spatial consistency between the MS image and the Pan band. The local linear model can well interpret the

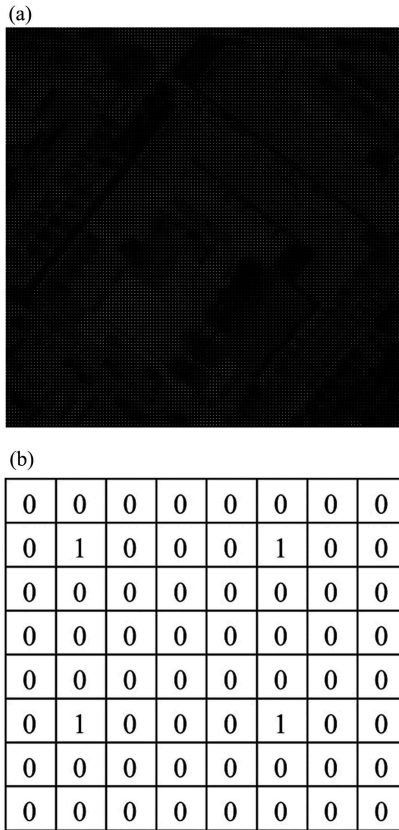


Figure 4. Unknowns of the proposed framework.

physical meaning of the spatial consistency and totally satisfy the above-mentioned two properties. The model describes the linear relationship between the ideal HRMS image (take channel U_T as an example) and the Pan band P , within a small window ω_k centred at pixel k :

$$\rho_{U_T}(x) = a_k \rho_P(x) + b_k, \forall x \in \omega_k, \quad (1)$$

where (a_k, b_k) are the linear coefficients, and $\rho_{U_T}(x)$ and $\rho_P(x)$ are the intensity values of U_T and P at pixel x , respectively.

Equation (1) is a typical least squares problem. To solve the linear coefficient (a_k, b_k) , we need to minimize the differences between the observed values and the calculated values of the channel U_T :

$$E(a_k, b_k) = \sum_{x \in \omega_k} ((a_k \rho_P(x) + b_k - \rho_{U_T}(x))^2), \quad (2)$$

where E is the objective cost function. To prevent a_k from becoming too large, a regularization term is added:

$$E(a_k, b_k) = \sum_{x \in \omega_k} ((a_k \rho_P(x) + b_k - \rho_{U_T}(x))^2 + \varepsilon a_k^2), \quad (3)$$

where ε is a regularization parameter. Finally, the objective cost function becomes a linear ridge regression model (Draper and Smith 2014), whose solution is

$$\left. \begin{aligned} a_k &= \frac{\frac{1}{|\omega_k|} \sum_{x \in \omega_k} \rho_P(x) \rho_{U_T}(x) - \mu_k \bar{\rho}_{U_T}(k)}{\sigma_k^2 + \varepsilon} \\ b_k &= \bar{\rho}_{U_T}(k) - a_k \mu_k \end{aligned} \right\}, \quad (4)$$

where μ_k and σ_k^2 are the mean and variance of P in the local window ω_k , respectively; $|\omega_k|$ refers to the number of pixels in window ω_k ; $\bar{\rho}_{U_T}(k) = \frac{1}{|\omega_k|} \sum_{x \in \omega_k} \rho_{U_T}(x)$ is the mean of U_T in window ω_k . **Algorithm 1** summarizes our local linear fusion method.

Algorithm 1: pansharpening based on the spatial consistency

Input: the MS image and the Pan band

Output: pansharpened image

1. Transform the original MS image using YUV to extract the colour channels U and V .
 2. Project U and V to obtain the HRMS image bands U_T and V_T according to the spatial resolution ratio.
 3. Build the relationship between the HRMS image and the Pan band via the local linear model/global cost function, and estimate the unknowns in the HRMS image.
 4. Perform an inverse YUV transform on the estimated HRMS to obtain the pansharpened image in the RGB colour space.
-

2.4. Global objective function model

Suppose x and y are adjacent pixels. According to spatial consistency, if the values of x and y in the Pan image are similar, they should be similar as well in the ideal HRMS image. Thus, from the algebraic meaning of spatial consistency, the above-mentioned relationship can be approximately expressed by minimizing the differences between the pixel value $\rho_{U_T}(x)$ in the HRMS image and its weighted average of the adjacent pixels:

$$E(U_T) = \sum_x (\rho_{U_T}(x) - \sum_{y \in N(x)} w_{xy} \rho_{U_T}(y))^2, \quad (5)$$

where $E(U_T)$ is an objective cost function, $N(x)$ is a local window of the pixel x , and w_{xy} is the weight function whose sum is 1. If the pixel values $\rho_P(x)$ and $\rho_P(y)$ in the Pan image are similar, w_{xy} is assigned a large value; in contrast, if the difference between $\rho_P(x)$ and $\rho_P(y)$ is large, w_{xy} is assigned a small value. The following equation is used as the weight function:

$$w_{xy} \propto e^{-\frac{(\rho_P(x) - \rho_P(y))^2}{2\sigma_x^2}}, \quad (6)$$

where σ_x^2 is the variance of $N(x)$.

Equation (5) is the least square solution of equation (7):

$$\rho_{U_T}(x) = \sum_{y \in N(x)} w_{xy} \rho_{U_T}(y). \quad (7)$$

The error equation matrix can be expressed as

$$\Delta \mathbf{V} = \mathbf{A} \mathbf{X} - \mathbf{L}, \quad (8)$$

where \mathbf{X} is the unknown column vector; \mathbf{A} is the coefficient matrix, where the values of the diagonal elements are 1 and the values of the non-diagonal elements are $-w_{xy}$; and \mathbf{L} is the observation vector, which is composed of the original MS-projected image pixel values, as in Figure 2(a). In equation (8), \mathbf{A} and \mathbf{L} are known, so the solution of \mathbf{X} is

$$\mathbf{X} = (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{L}). \quad (9)$$

In this global objective function model, all of the pixel values are involved in the cost-minimization procedure. The least squares method can minimize the cost function and reach a global optimal solution. **Algorithm 1** summarizes our global optimization fusion method.

3. Results

3.1. Data sets and experiment environment

Three hundred satellite images with Pan bands collected from Geoeye-1, Landsat 8 OLI, QuickBird, SPOT 6, Worldview-2, and Worldview 3¹ are used as the test sets to evaluate the proposed method. Detailed information about these images can be found in Table 2. To obtain the reference images, we consider an upscaling operation according to the spatial resolution ratios (upsampling factor is set to 2 for Landsat 8 OLI and 4 for other satellites) between the MS images and the Pan bands. This strategy has been widely applied for pansharpening evaluation, such as in Fang et al. (2013), Chen et al. (2014), Masi et al. (2016), and Wang, Shi, and Atkinson (2016).

The proposed two methods, the local linear model (noted by Ours1) and the global cost model (noted by Ours2), are compared with 15 widely used approaches summarized in Vivone et al. (2015) to illustrate the benefits of the proposed spatial consistency prior. These compared methods are BDSF, Brovey, GIHS, GS, GSA, PCA, PRACS, ATWT, ATWT-M2, ATWT-M3, AWLP, HPF, Indusion, MTF-GLP, and SFIM. The source codes of these 15 methods are publicly available.² There are three main parameters in the proposed methods, namely the size of window ω_k , the regularization parameter ε , and the size of window $N(x)$. In the following experiments, ω_k is set to 64×64 pixels, ε is set to 0.01, and $N(x)$ is set to 5×5 pixels. The parameters of the other compared methods are set according to their author's recommendations. All of the tests are completed on a single-core I5 CPU computer with Matlab 2014.

Table 2. Detailed information of the collected six data sets.

Satellite sensor	Resolution (m)	Image size (pixel)	Number of images
Geoeye-1	Pan: 0.41; MS: 1.65	512 × 512	50
Landsat 8 OLI	Pan:15; MS: 30	512 × 512	50
Quickbird	Pan: 0.55; MS: 2.16	512 × 512	50
SPOT-6	Pan:1.5; MS: 6	512 × 512	50
Worldview-2	Pan: 0.46; MS: 1.85	512 × 512	50
Worldview-3	Pan: 0.31; MS: 1.24	512 × 512	50

3.2. Visual comparison

We first compare the visual appearance of these methods on six images that are selected from the collected six data sets. The results are shown in Figures 5–10. For better visualization, a 256×256 pixels sub-area cropped from each original result image is shown. As can be seen, all methods can enhance the spatial resolution of the original MS images, namely all fusion results are visually clearer than the upscaled MS images. The performance of BDSD may be relative to the image sensors. For example, the results of BDSD in Figures 5–6 are slightly blurred, the results in Figures 7–8 are sharpened, and the results in Figures 9–10 suffer from severe image artefacts. The Brovey, GIHS, GS, GSA, and PCA approaches suffer from strong spectral distortions. These methods usually have low contrasts. The image colours of grass, trees, and white buildings are changed, which indicates spectral distortion. For instance, the trees in Figure 8 are brightened, whereas the white buildings in Figure 9 are darkened by Brovey, GIHS, GS, and PCA; GSA turns the green playground in Figure 7 to black. Although PRACS can preserve high spectral quality, it may cause speckle noise, such as in Figure 6. In the MRA group, ATWT-M2 and ATWT-M3 obtain over-smooth results (the results shown in the figures are severely blurred) and fail to restore the detailed information. In other words, the spatial information is poor in the results of ATWT-M2 and ATWT-M3. The ATWT, AWLP, HPF, Indusion, MTF-GLP, and SFIM methods usually obtain results with over-enhanced details and suffer from severe image artefacts. For example, the trees in Figure 7 seem to be divided into small parts by these methods. In addition, AWLP and SFIM may also cause speckle noise such as in Figure 7. Compared with these methods, Ours1 and Ours2 achieve better visual appearance on these images. Both spatial and spectral properties are preserved; as a result, their fused images are comparable to the reference image.

To provide more intuitionistic visual expression, Figure 11 shows the error maps (the brighter the pixel, the larger the errors) between the fusion results and the reference image of Figure 7 (the error range is [0–50] pixels, where errors larger than 50 pixels are shown by 50). The spectral distortions, blurred edges, and image artefacts can be clearly observed in these error maps. For example, the spectral distortions of the BDSD, Brovey, GIHS, GS, GSA, and PCA methods are severe; the errors of white buildings or trees are generally larger than 50 pixels. PRACS achieves the best performance among the CS group. It has small errors except in the playground regions. Compared with the CS methods, the MRA methods suffer from less spectral distortions. However, the MRA methods suffer from information detail loss or image artefacts. For instance, the edges in the error maps of ATWT-M2, ATWT-M3, and Indusion are much less sharper than the other methods, namely their results preserve less-detailed spatial information; image artefacts can be easily observed in the error maps of ATWT, AWLP, HPF, MTF-GLP, and SFIM. Although some spectral distortions are still produced by the two proposed methods (Ours1 and Ours2), their overall performances are much better than the other methods.

3.3. Quantitative analysis

Five spectral and one spatial metrics are selected to evaluate the pansharpener quality of each method. The spectral metrics include relative average spectral error (RASE) (Ranchin and Wald 2000), relative global-dimensional synthesis error (ERGAS)

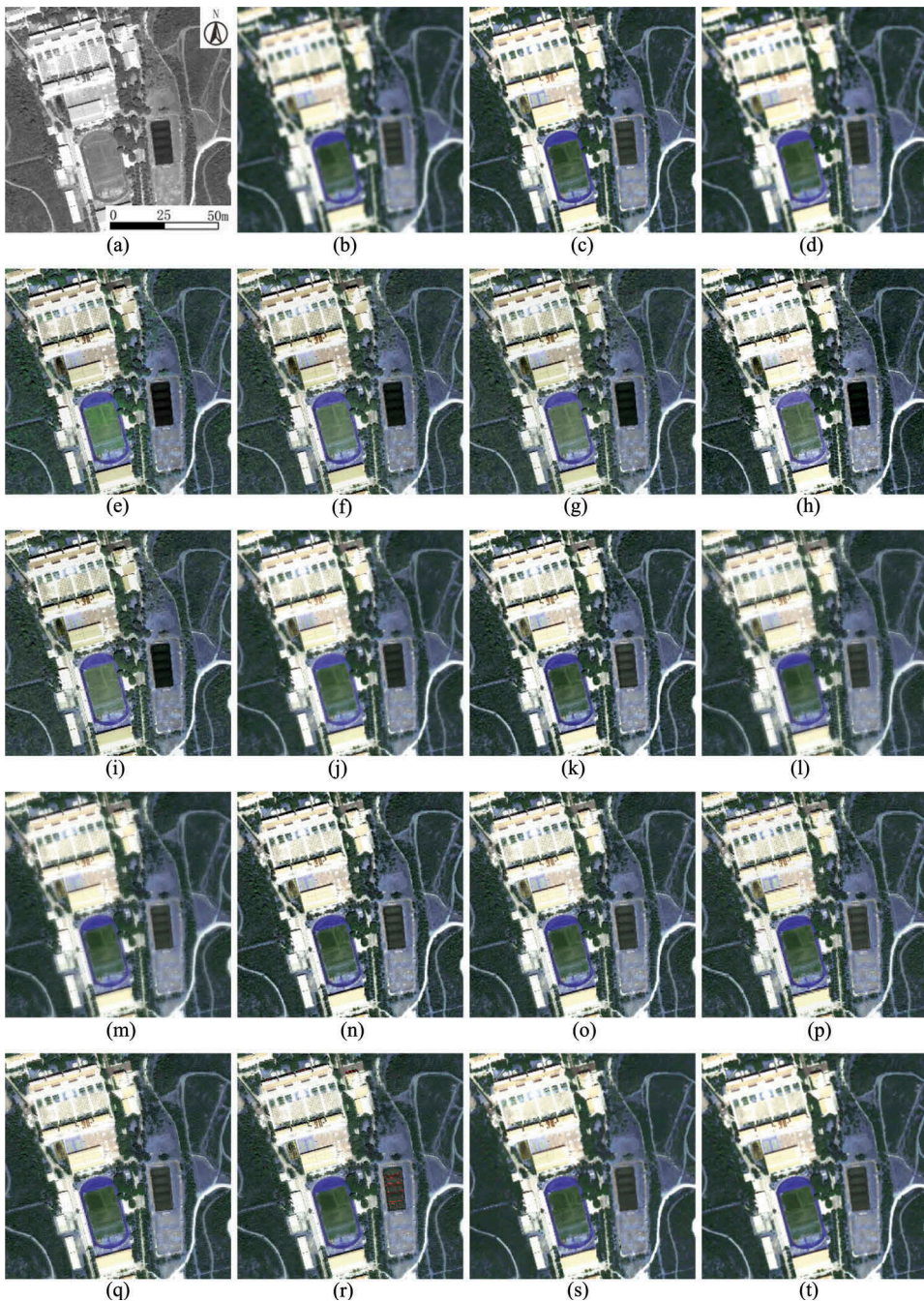


Figure 5. Pansharpening results for a sample image from the Geosy-1 data set. (a) Pan band. (b) Upscaled MS image. (c) Reference image. (d) BDSD. (e) Brovey. (f) GIHS. (g) GS. (h) GSA. (i) PCA. (j) PRACS. (k) ATWT. (l) ATWT-M2. (m) ATWT-M3. (n) AWLP. (o) HPF. (p) Indusion. (q) MTF-GLP. (r) SFIM. (s) Ours1. (t) Ours2.

(Ranchin and Wald 2000), universal image quality index (QAVE) (Zhou and Bovik 2002), Spectral Angle Mapper (SAM) (Yuhas, Goetz, and Boardman 1992), and spectral information divergence (SID) (Chang 2000). The spatial metric is called coherence (Wang,

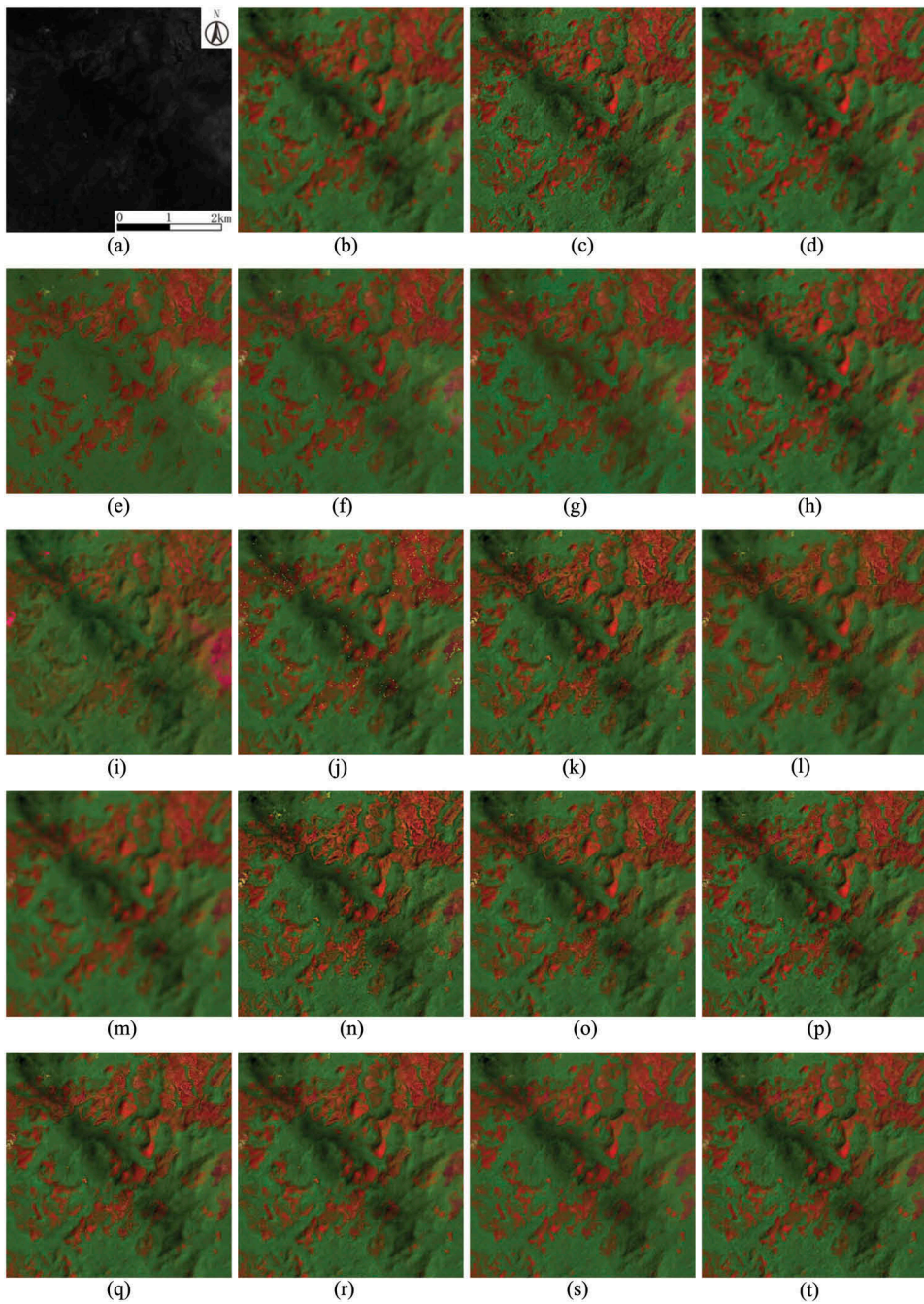


Figure 6. Pansharpening results for a sample image from the Landsat 8 OLI data set. (a) Pan band. (b) Upscaled MS image. (c) Reference image. (d) BSDS. (e) Brovey. (f) GIHS. (g) GS. (h) GSA. (i) PCA. (j) PRACS. (k) ATWT. (l) ATWT-M2. (m) ATWT-M3. (n) AWLP. (o) HPF. (p) Indusion. (q) MTF-GLP. (r) SFIM. (s) Ours1. (t) Ours2.

Shi, and Atkinson 2016), which is quantified by one minus the correlation coefficient (r). Coherence is an index measuring the relation between the pansharpening results and the reference image. In addition, the root mean square error (RMSE)

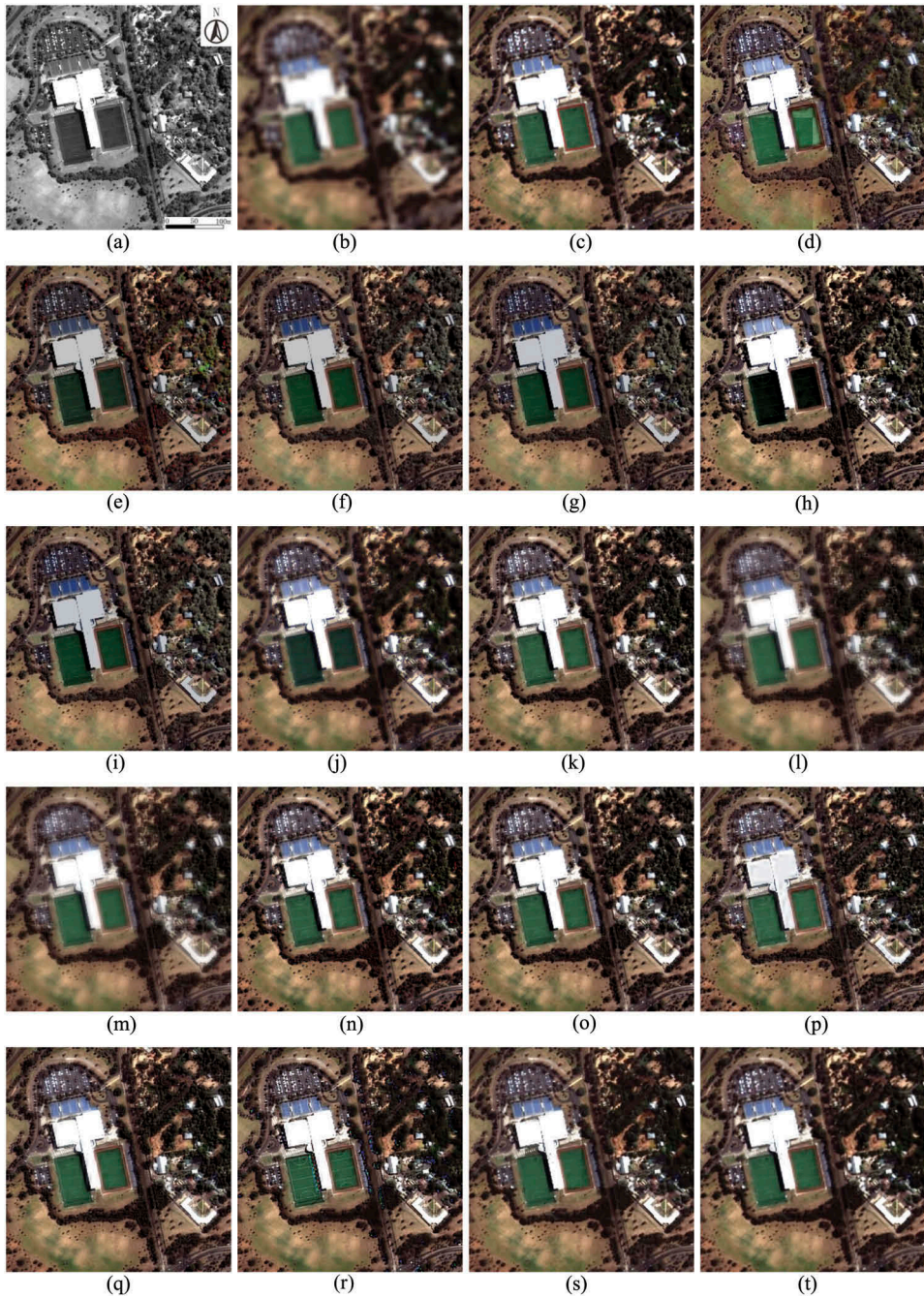


Figure 7. Pansharpener results for a sample image from the SPOT-6 data set. (a) Pan band. (b) Upscaled MS image. (c) Reference image. (d) BSDS. (e) Brovey. (f) GIHS. (g) GS. (h) GSA. (i) PCA. (j) PRACS. (k) ATWT. (l) ATWT-M2. (m) ATWT-M3. (n) AWLP. (o) HPF. (p) Indusion. (q) MTF-GLP. (r) SFIM. (s) Ours1. (t) Ours2.

and the peak signal to noise ratio (PSNR) are also used to evaluate the overall performance. The quantitative assessment results for the 17 methods on the six data sets are reported in Tables 3–8.

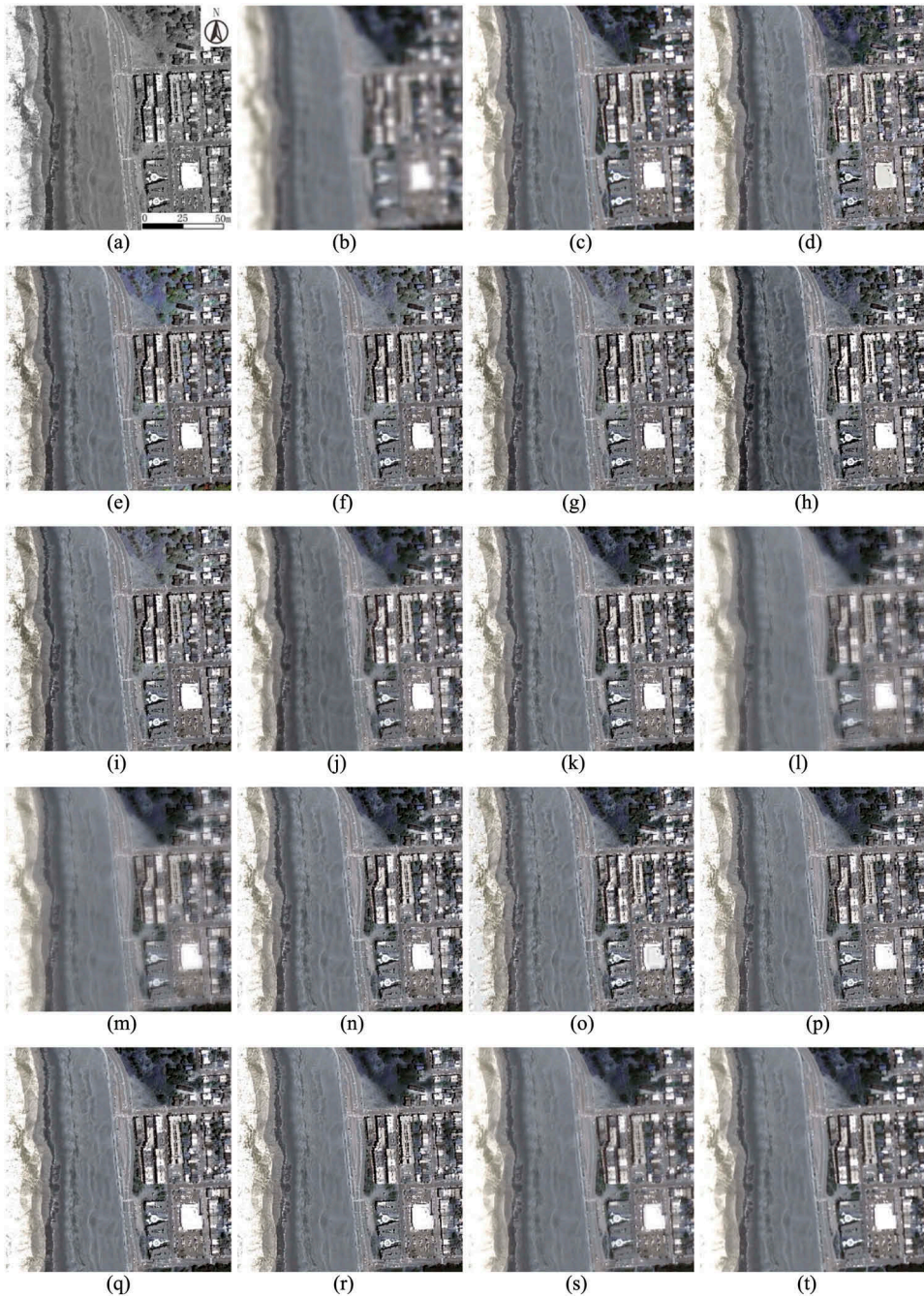


Figure 8. Pansharpening results for a sample image from the QuickBird data set. (a) Pan band. (b) Upscaled MS image. (c) Reference image. (d) BDSD. (e) Brovey. (f) GIHS. (g) GS. (h) GSA. (i) PCA. (j) PRACS. (k) ATWT. (l) ATWT-M2. (m) ATWT-M3. (n) AWLP. (o) HPF. (p) Indusion. (q) MTF-GLP. (r) SFIM. (s) Ours1. (t) Ours2.

The ideal reference value for each evaluation metric is also provided for convenience of inter-comparison in the first row of each table. From these tables, the following factors are observed. (1) CS methods generally suffer from severe spectral distortions,



Figure 9. Pansharpening results for a sample image from the Worldview-2 data set. (a) Pan band. (b) Upscaled MS image. (c) Reference image. (d) BSDS. (e) Brovey. (f) GIHS. (g) GS. (h) GSA. (i) PCA. (j) PRACS. (k) ATWT. (l) ATWT-M2. (m) ATWT-M3. (n) AWLP. (o) HPF. (p) Indusion. (q) MTF-GLP. (r) SFIM. (s) Ours1. (t) Ours2.

while having good spatial coherence. PRACS performs the best in the CS family. It even achieves the best results among all of the 17 methods on the SPOT-6 data set. (2) MRA methods usually preserve good spectral properties and poor spatial coherence; the

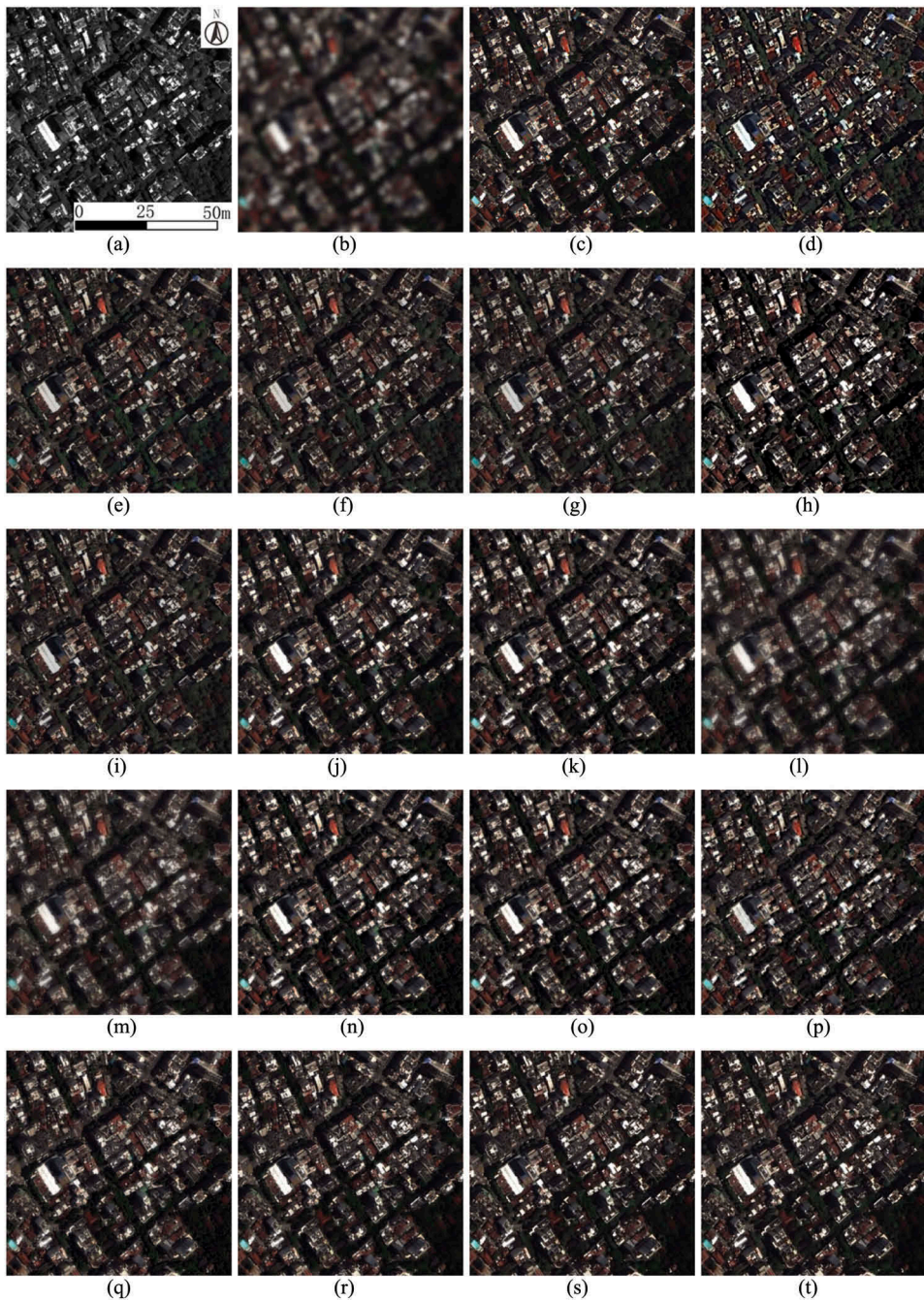


Figure 10. Pansharpener results for a sample image from the Worldview-3 data set. (a) Pan band. (b) Upscaled MS image. (c) Reference image. (d) BSDS. (e) Brovey. (f) GIHS. (g) GS. (h) GSA. (i) PCA. (j) PRACS. (k) ATWT. (l) ATWT-M2. (m) ATWT-M3. (n) AWLP. (o) HPF. (p) Indusion. (q) MTF-GLP. (r) SFIM. (s) Ours1. (t) Ours2.

overall performance of the MRA methods (except for ATWT-M2 and ATWT-M3) is usually better than the CS methods (except for PRACS). Although ATWT-M2 and ATWT-M3 achieve good RMSEs on several data sets, their performance is poor due to the severe

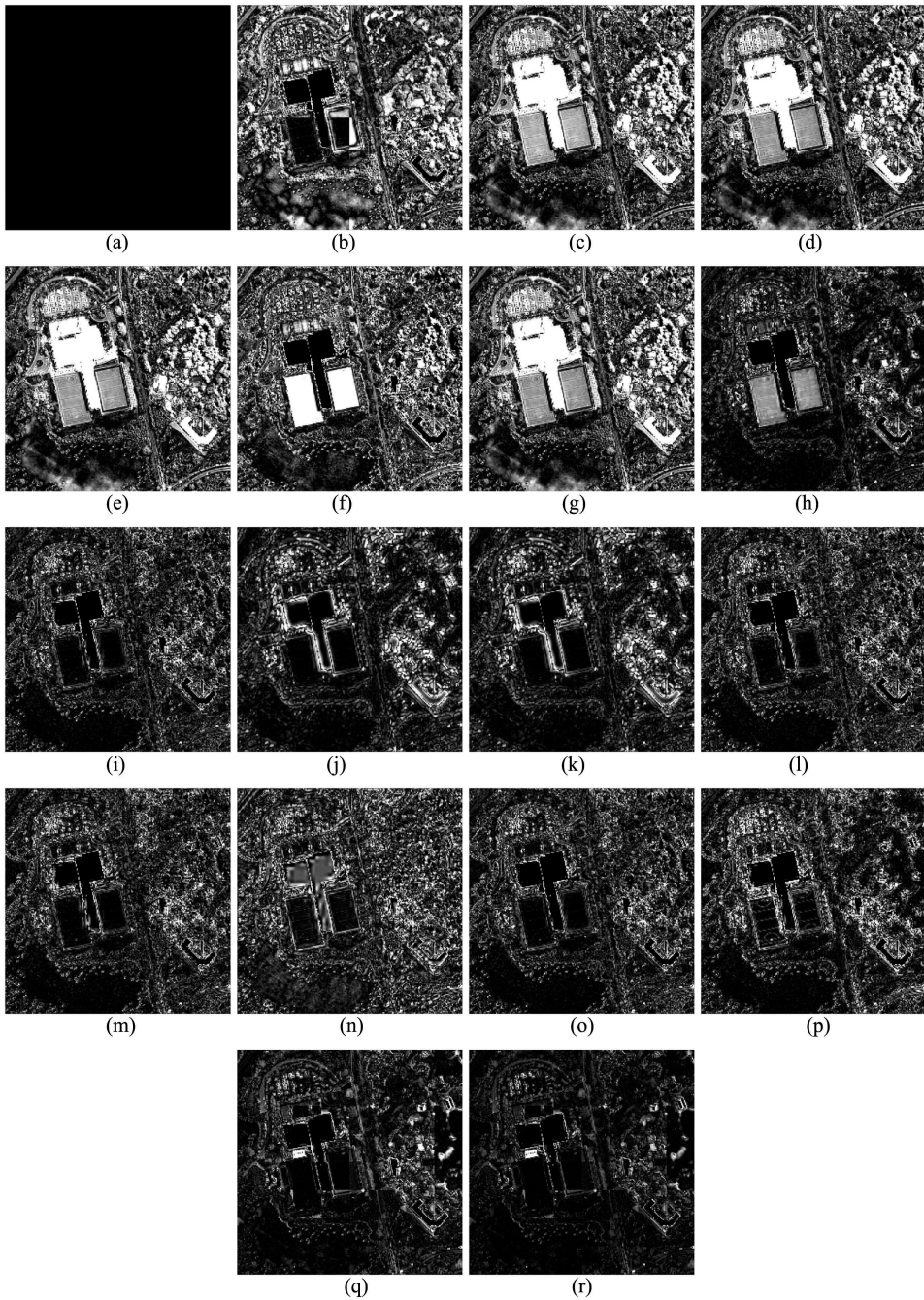


Figure 11. Error maps between the fusion results and the reference image in Figure 7. (a) Reference. (b) BDSD. (c) Brovey. (d) GIHS. (e) GS. (f) GSA. (g) PCA. (h) PRACS. (i) ATWT. (j) ATWT-M2. (k) ATWT-M3. (l) AWLP. (m) HPF. (n) Indusion. (o) MTF-GLP. (p) SFIM. (q) Ours1. (r) Ours2.

blurry effects. (3) The proposed two methods achieve impressive results due to the consideration of the spatial distribution characteristic. Ours2 achieves the best overall performance on five of the six data sets. Table 9 provides the average performance of

Table 3. Quantitative assessment of the fusion methods on the Geoeye-1 data set.

Method	RMSE	RASE	PSNR	ERGAS	QAVE	SAM (°)	SID	1-r
Reference	0	0	$+\infty$	0	1	0	0	0
BDSB	25.88	23.48	19.97	5.84	0.503	4.08	0.088	0.017
Brovey	22.55	20.69	21.18	5.13	0.527	3.75	0.072	0.020
GIHS	22.25	20.34	21.30	5.14	0.536	3.92	0.082	0.013
GS	22.26	20.34	21.29	5.12	0.531	3.90	0.080	0.013
GSA	24.19	21.99	20.55	5.52	0.516	4.10	0.093	0.028
PCA	22.23	20.30	21.31	5.11	0.530	3.90	0.080	0.013
PRACS	21.78	19.58	21.48	4.88	0.518	3.65	0.060	0.019
ATWT	20.75	18.92	21.91	4.75	0.533	3.71	0.077	0.022
ATWT-M2	17.37	16.02	23.47	3.99	0.539	3.39	0.061	0.014
ATWT-M3	17.38	16.01	23.46	3.98	0.539	3.38	0.059	0.015
AWLP	20.99	19.27	21.80	4.75	0.520	4.08	0.095	0.016
HPF	20.89	19.05	21.85	4.78	0.534	3.66	0.074	0.021
Indusion	21.85	19.93	21.45	5.01	0.530	3.97	0.089	0.014
MTF-GLP	20.86	19.02	21.86	4.77	0.533	3.75	0.079	0.022
SFIM	20.96	19.16	21.82	4.82	0.529	3.54	0.068	0.018
Ours1	18.56	17.09	22.89	4.29	0.542	3.48	0.064	0.014
Ours2	16.06	14.64	24.15	3.66	0.545	3.37	0.058	0.014

The bold values indicate the most accurate result in each term.

Table 4. Quantitative assessment of the fusion methods on the Landsat 8 OLI data set.

Method	RMSE	RASE	PSNR	ERGAS	QAVE	SAM (°)	SID	1-r
Reference	0	0	$+\infty$	0	1	0	0	0
BDSB	6.87	4.73	34.17	2.36	0.655	1.12	0.016	0.009
Brovey	8.37	5.65	31.43	2.83	0.654	1.08	0.014	0.008
GIHS	8.30	5.61	31.50	2.81	0.654	1.10	0.011	0.008
GS	8.32	5.62	31.48	2.80	0.651	1.12	0.011	0.009
GSA	10.30	7.04	30.08	3.50	0.643	1.67	0.021	0.013
PCA	8.31	5.62	31.49	2.80	0.651	1.12	0.012	0.009
PRACS	8.26	5.59	31.62	2.78	0.648	1.21	0.024	0.012
ATWT	7.23	4.83	32.31	2.41	0.653	1.14	0.015	0.012
ATWT-M2	5.82	3.99	35.28	1.99	0.650	1.10	0.013	0.011
ATWT-M3	5.91	4.04	35.02	2.02	0.649	1.11	0.014	0.012
AWLP	7.28	4.86	32.24	2.42	0.651	1.17	0.017	0.011
HPF	7.36	4.92	32.24	2.46	0.653	1.12	0.015	0.012
Indusion	9.01	6.10	30.94	3.04	0.645	1.17	0.014	0.009
MTF-GLP	7.14	4.78	32.44	2.38	0.653	1.12	0.015	0.012
SFIM	7.33	4.90	32.26	2.45	0.653	1.09	0.011	0.012
Ours1	6.01	4.05	34.73	2.03	0.655	1.09	0.014	0.009
Ours2	5.47	3.70	35.33	1.85	0.655	1.08	0.014	0.009

The bold values indicate the most accurate result in each term.

each method on these six data sets. The methods in the table are ranked by the RMSE index. Note that the ATWT-M2 and ATWT-M3 methods are not included due to their severe blurry effects. As can be seen, PRACS is the best in the CS family, which ranks the third of all of the 17 methods; ATWT, AWLP, and MTF-GLP perform better than the other MRA methods; Ours2 and Ours1 outperform the other 15 widely used methods; Ours2 is better than Ours1.

3.4. Running time comparison

We compare the running time of the proposed two methods with several widely used approaches, including GIHS, PRACS, ATWT, AWLP, and MTF-GLP, on the SPOT-6 data set.

Table 5. Quantitative assessment of the fusion methods on the SPOT-6 data set.

Method	RMSE	RASE	PSNR	ERGAS	QAVE	SAM (°)	SID	1-r
Reference	0	0	$+\infty$	0	1	0	0	0
BDSB	35.37	36.58	17.28	9.12	0.475	7.48	0.206	0.007
Brovey	32.23	33.40	18.01	8.38	0.483	4.92	0.094	0.015
GIHS	31.98	33.13	18.08	8.41	0.517	5.34	0.086	0.007
GS	32.00	33.15	18.07	8.42	0.516	5.43	0.088	0.008
GSA	31.47	32.46	18.29	8.27	0.462	11.35	0.255	0.026
PCA	31.99	33.14	18.07	8.42	0.515	5.43	0.088	0.008
PRACS	19.66	20.27	22.33	5.14	0.518	4.17	0.049	0.010
ATWT	21.39	22.05	21.58	5.62	0.523	5.90	0.136	0.020
ATWT-M2	22.65	23.39	21.12	5.93	0.529	4.23	0.047	0.015
ATWT-M3	22.29	23.02	21.27	5.84	0.530	4.22	0.046	0.016
AWLP	21.51	22.19	21.52	5.59	0.502	5.97	0.091	0.014
HPF	22.45	23.14	21.16	5.89	0.529	5.29	0.108	0.019
Indusion	26.34	27.14	19.79	6.89	0.516	6.05	0.145	0.010
MTF-GLP	21.31	21.97	21.60	5.59	0.521	6.15	0.147	0.021
SFIM	23.40	24.18	20.78	6.22	0.531	4.78	0.096	0.016
Ours1	22.17	22.72	21.33	5.86	0.549	4.23	0.047	0.013
Ours2	20.50	21.19	21.97	5.46	0.549	4.17	0.049	0.013

The bold values indicate the most accurate result in each term.

Table 6. Quantitative assessment of the fusion methods on the QuickBird data set.

Method	RMSE	RASE	PSNR	ERGAS	QAVE	SAM (°)	SID	1-r
Reference	0	0	$+\infty$	0	1	0	0	0
BDSB	30.65	30.07	18.94	7.41	0.479	5.99	0.107	0.014
Brovey	36.53	36.70	17.87	9.13	0.449	5.68	0.076	0.017
GIHS	36.33	36.49	17.91	9.10	0.467	6.12	0.103	0.014
GS	36.37	36.53	17.90	9.09	0.449	6.09	0.103	0.016
GSA	35.49	34.64	17.54	8.62	0.445	9.38	0.122	0.020
PCA	36.40	36.56	17.89	9.10	0.448	6.10	0.103	0.016
PRACS	19.32	18.93	22.73	4.77	0.487	4.50	0.064	0.015
ATWT	22.54	22.40	21.73	5.58	0.479	6.67	0.102	0.020
ATWT-M2	17.61	17.58	23.73	4.40	0.486	4.68	0.077	0.017
ATWT-M3	17.49	17.49	23.84	4.38	0.489	4.76	0.077	0.017
AWLP	22.66	22.54	21.68	5.61	0.468	6.72	0.091	0.017
HPF	22.88	22.73	21.59	5.66	0.481	6.38	0.099	0.019
Indusion	26.59	26.39	20.25	6.58	0.475	6.38	0.101	0.016
MTF-GLP	22.71	22.59	21.69	5.62	0.478	6.77	0.103	0.020
SFIM	26.01	26.34	20.61	6.61	0.471	5.86	0.098	0.021
Ours1	18.58	18.74	23.42	4.76	0.501	4.41	0.077	0.015
Ours2	17.27	17.16	24.32	4.21	0.510	4.31	0.077	0.015

The bold values indicate the most accurate result in each term.

The results are shown in Table 10. GIHS is the most efficient, which is an order of magnitude faster than the others. Ours1 is faster than PRACS and the other compared MRA methods. Ours2 is a global energy function minimization method, which achieves the best accuracy performance at the cost of heavy running time. However, Ours2 is much faster than the ATPRK method, whose reported running time is hundreds of seconds for a 256×256 image on an Intel Core i7 Processor at 3.40 GHz. It mainly profits from the fact that the global cost function minimization is a linear least squares problem, which can avoid the iterative procedure and greatly reduce computational complexity.

Table 7. Quantitative assessment of the fusion methods on the Worldview-2 data set.

Method	RMSE	RASE	PSNR	ERGAS	QAVE	SAM (°)	SID	1-r
Reference	0	0	$+\infty$	0	1	0	0	0
BDSB	24.09	35.89	20.58	8.78	0.340	8.92	0.165	0.010
Brovey	25.79	39.38	20.04	9.86	0.332	6.91	0.068	0.011
GIHS	25.70	39.29	20.08	9.83	0.347	7.54	0.094	0.015
GS	25.71	39.29	20.08	9.83	0.344	7.60	0.094	0.014
GSA	21.32	32.45	21.67	8.10	0.313	14.67	0.312	0.025
PCA	25.69	39.27	20.09	9.83	0.344	7.61	0.094	0.014
PRACS	18.70	28.63	22.96	7.23	0.367	6.57	0.067	0.017
ATWT	17.72	27.12	23.46	6.79	0.360	8.44	0.143	0.021
ATWT-M2	23.10	35.16	21.06	8.80	0.351	6.67	0.069	0.018
ATWT-M3	22.16	33.72	21.42	8.44	0.357	6.66	0.070	0.019
AWLP	17.65	27.00	23.47	6.77	0.356	7.83	0.082	0.018
HPF	17.56	26.90	23.56	6.74	0.366	7.55	0.115	0.019
Indusion	19.53	29.56	22.42	7.38	0.351	8.95	0.142	0.016
MTF-GLP	17.62	26.95	23.47	6.74	0.359	8.78	0.148	0.022
SFIM	16.83	25.73	23.82	6.48	0.355	7.01	0.100	0.019
Ours1	18.05	27.57	23.26	6.92	0.369	6.67	0.072	0.015
Ours2	15.84	24.23	24.40	6.08	0.370	6.63	0.074	0.015

The bold values indicate the most accurate result in each term.

Table 8. Quantitative assessment of the fusion methods on the Worldview-3 data set.

Method	RMSE	RASE	PSNR	ERGAS	QAVE	SAM (°)	SID	1-r
Reference	0	0	$+\infty$	0	1	0	0	0
BDSB	12.02	27.74	27.27	6.76	0.429	5.92	0.081	0.007
Brovey	12.83	30.08	26.32	7.49	0.381	7.18	0.040	0.009
GIHS	12.83	30.05	26.33	7.48	0.391	7.68	0.084	0.008
GS	12.81	30.02	26.33	7.47	0.380	7.65	0.086	0.009
GSA	10.59	24.68	28.17	6.14	0.395	9.72	0.146	0.015
PCA	12.79	29.96	26.35	7.45	0.380	7.65	0.086	0.009
PRACS	8.69	19.96	30.08	5.03	0.440	4.12	0.020	0.012
ATWT	8.84	20.54	29.83	5.14	0.432	6.06	0.074	0.013
ATWT-M2	12.59	29.16	27.53	7.29	0.422	4.18	0.017	0.011
ATWT-M3	11.65	26.96	28.14	6.73	0.421	4.25	0.020	0.012
AWLP	8.81	20.48	29.86	5.13	0.430	5.73	0.038	0.011
HPF	8.97	20.92	29.81	5.23	0.438	5.54	0.061	0.014
Indusion	10.05	23.39	28.71	5.83	0.420	6.46	0.078	0.010
MTF-GLP	8.97	20.62	29.45	5.15	0.431	6.34	0.078	0.014
SFIM	10.39	24.62	28.05	6.13	0.425	5.52	0.074	0.020
Ours1	8.35	19.15	30.46	4.85	0.433	4.26	0.018	0.010
Ours2	8.05	17.91	30.37	4.54	0.430	4.28	0.020	0.011

The bold values indicate the most accurate result in each term.

4. Discussion

This article proposes a spatial distribution consistency assumption between the ideal HRMS image and the Pan band, and develops their relationship model. However, the spatial consistency and the spatial heterogeneity between the Pan band and the MS image exist at the same time (Thomas et al. 2008). The proposed constraint requires only the extreme positions and local change trend to be consistent. Actually, the magnitude of the specific change is different, so a certain extent of 'the Pan and MS images spatial heterogeneity characteristics' is taken into account. Balancing the consistency and heterogeneity needs further study and research.

Table 9. Average quantitative assessment of the fusion methods on the six data sets.

Method	RMSE	RASE	PSNR	ERGAS	QAVE	SAM (°)	SID	1-r
Reference	0	0	$+\infty$	0	1	0	0	0
Ours2	13.87	16.47	26.76	4.30	0.510	3.97	0.052	0.013
Ours1	15.29	18.22	26.02	4.79	0.508	4.02	0.051	0.013
PRACS	16.07	18.81	25.20	4.97	0.496	4.04	0.047	0.014
ATWT	16.41	19.31	25.14	5.05	0.497	5.32	0.091	0.018
MTF-GLP	16.44	19.32	25.09	5.04	0.496	5.49	0.095	0.019
AWLP	16.48	19.39	25.10	5.05	0.488	5.25	0.069	0.015
HPF	16.69	19.61	25.04	5.13	0.500	4.92	0.079	0.017
SFIM	17.49	20.82	24.56	5.45	0.494	4.63	0.076	0.018
Indusion	18.90	22.09	23.93	5.79	0.490	5.50	0.095	0.013
GSA	22.21	25.54	22.72	6.69	0.462	8.48	0.158	0.021
BDSD	22.48	26.42	23.04	6.71	0.480	5.59	0.111	0.011
GIHS	22.90	27.49	22.53	7.19	0.485	5.28	0.077	0.011
PCA	22.90	27.48	22.53	7.12	0.478	5.30	0.077	0.012
GS	22.91	27.49	22.53	7.12	0.479	5.30	0.077	0.012
Brovay	23.05	27.65	22.48	7.14	0.471	4.92	0.061	0.013

The bold values indicate the most accurate result in each term.

Table 10. Running time comparisons.

Method	GIHS	PRACS	ATWT	AWLP	MTF-GLP	Ours1	Ours2
Time (s)	0.02	0.41	0.55	0.47	0.44	0.21	31.26

Another important issue is the universality of the spatial consistency. Different objects have different reflectances and absorptions, so the intensity values of different bands are quite different. For example, vegetation is absorbed in the R band and reflected in the G band, making the intensity values of the G band greater than the ones of the R band. However, the intensity value differences between the R and G bands do not contradict with the assumption of spatial consistency. The intensity value differences at the same position do not affect the consistency of the spatial characteristics. The same object at different image bands has a consistent change trend, i.e. local trends and extreme positions are the same. Figure 12 provides an example to show the intensity of the statistical characteristics of vegetation in the R and G bands.

In the figure, the first row displays four remote-sensing images of different vegetation regions, and the red line in each image indicates the selected pixels. The green lines in the last row of Figure 12 are the curves of the pixel values in the G band, and the red lines are the pixel curves in the R band. As can be seen, the intensity values of the G band are generally larger than the values of the R band. This heterogeneity agrees with the fact that 'vegetation is absorbed in the R band and reflected in the G band'. However, it can also be found that the local extreme positions in the R and G bands are almost at the same locations, and the local monotonicity of the two curves is also consistent. Obviously, spatial heterogeneity and spatial consistency exist simultaneously in remote-sensing images. Spatial heterogeneity reflects the reflection and absorption characteristic differences of the same object in different bands. The incident angles caused by the different spatial positions and terrains are not the same, resulting in 'synonyms spectrum' phenomenon. In different spatial positions, the change trends of the same object in different bands are the same and consistent with the relevant spatial terrain, which is the spatial consistency proposed in this article. The utility of spatial

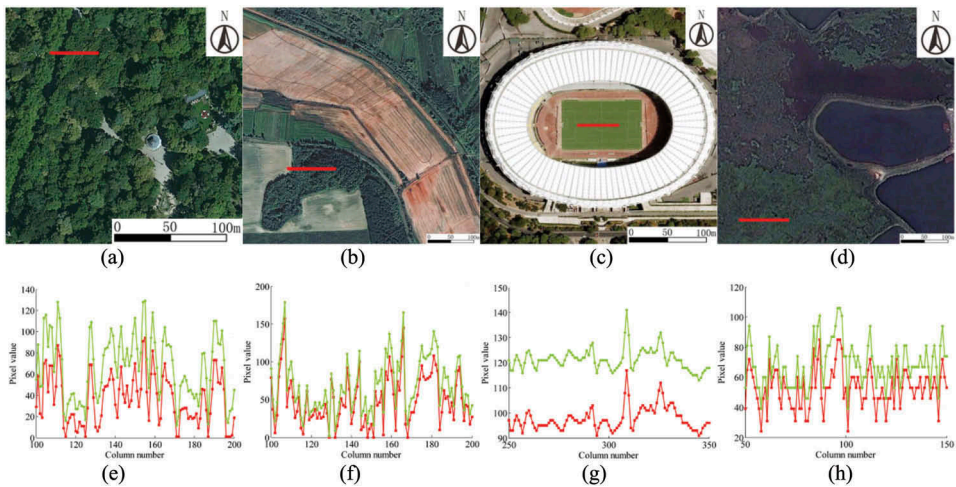


Figure 12. Spatial consistency in different vegetation regions.

consistency characteristics may reduce the fragmentation problem of image classification caused by the synonyms spectrum. The same object in different locations can be classified through its intensity features and texture features; meanwhile, the fragmented things can be merged by introducing an intensity-scale consistency between different bands.

5. Conclusions

In this article, we propose a spatial distribution consistency prior for remote-sensing image pansharpening. This prior assumes that the spatial distributions of the ideal HRMS image pixel values and the Pan image pixel values are consistent. According to the physical and algebraic meaning of the spatial consistency assumption, two modelling methods, namely local linear transformation and global cost function, are presented to express the relationship between the Pan and the MS images. The local linear method makes a trade-off between accuracy and efficiency, which is less accurate than the global cost method but faster than it. Both methods are validated on six real satellite image data sets, which are composed of 300 images, and are compared with 15 other widely used approaches. The experimental results show that the proposed two methods achieve very impressive performance.

Notes

1. <https://www.digitalglobe.com/product-samples>.
2. <http://www.openremotesensing.net/index.php/codes/11-pansharpening/2-pansharpening>.

Acknowledgement

The authors would like to express their gratitude to the editors and the reviewers for their constructive and helpful comments for substantial improvement of this article.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Natural Science Foundation of China [Nos. 41271452 and 41701528].

ORCID

Jiayuan Li  <http://orcid.org/0000-0002-9850-1668>

References

- Aiazzi, B., L. Alparone, S. Baronti, A. Garzelli, and M. Selva. 2006. "MTF-tailored Multiscale Fusion of High-Resolution MS and Pan Imagery." *Photogrammetric Engineering & Remote Sensing* 72 (5): 591–596. doi:10.14358/PERS.72.5.591.
- Aiazzi, B., L. Alparone, S. Baronti, and A. Garzelli. 2002. "Context-Driven Fusion of High Spatial and Spectral Resolution Images Based on Oversampled Multiresolution Analysis." *IEEE Transactions on Geoscience and Remote Sensing* 40 (10): 2300–2312. doi:10.1109/TGRS.2002.803623.
- Aiazzi, B., S. Baronti, and M. Selva. 2007. "Improving Component Substitution Pansharpening through Multivariate Regression of MS \pm Pan Data." *IEEE Transactions on Geoscience and Remote Sensing* 45 (10): 3230–3239. doi:10.1109/TGRS.2007.901007.
- Amro, I., J. Mateos, M. Vega, R. Molina, and A. K. Katsaggelos. 2011. "A Survey of Classical Methods and New Trends in Pansharpening of Multispectral Images." *EURASIP Journal on Advances in Signal Processing* 2011 (1): 79. doi:10.1186/1687-6180-2011-79.
- Ballester, C., V. Caselles, L. Igual, J. Verdera, and R. Bernard. 2006. "A Variational Model for P+ XS Image Fusion." *International Journal of Computer Vision* 69 (1): 43–58. doi:10.1007/s11263-006-6852-x.
- Carper, W., T. Lillesand, and R. Kiefer. 1990. "The Use of Intensity-Hue-Saturation Transformations for Merging SPOT Panchromatic and Multispectral Image Data." *Photogrammetric Engineering and Remote Sensing* 56 (4): 459–467.
- Chang, C.-I. 2000. "An Information-Theoretic Approach to Spectral Variability, Similarity, and Discrimination for Hyperspectral Image Analysis." *IEEE Transactions on Information Theory* 46 (5): 1927–1932. doi:10.1109/18.857802.
- Chavez, P., S. C. Sides, and J. A. Anderson. 1991. "Comparison of Three Different Methods to Merge Multiresolution and Multispectral Data- Landsat TM and SPOT Panchromatic." *Photogrammetric Engineering and Remote Sensing* 57 (3): 295–303.
- Chen, C., Y. Li, W. Liu, and J. Huang. 2014. "Image Fusion with Local Spectral Consistency and Dynamic Gradient Sparsity." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- Choi, J., Y. Kiyun, and Y. Kim. 2011. "A New Adaptive Component-Substitution-Based Satellite Image Fusion by Using Partial Replacement." *IEEE Transactions on Geoscience and Remote Sensing* 49 (1): 295–309. doi:10.1109/TGRS.2010.2051674.
- Draper, N. R., and H. Smith. 2014. *Applied Regression Analysis*. New York, NY: John Wiley & Sons.
- Fang, F., L. Fang, C. Shen, and G. Zhang. 2013. "A Variational Approach for Pan-Sharpener." *IEEE Transactions on Image Processing* 22 (7): 2822–2834. doi:10.1109/TIP.2013.2258355.
- Garzelli, A., F. Nencini, and L. Capobianco. 2008. "Optimal MMSE Pan Sharpening of Very High Resolution Multispectral Images." *IEEE Transactions on Geoscience and Remote Sensing* 46 (1): 228–236. doi:10.1109/TGRS.2007.907604.
- Gillespie, A. R., A. B. Kahle, and R. E. Walker. 1987. "Color Enhancement of Highly Correlated Images. II. Channel Ratio and "Chromaticity" Transformation Techniques." *Remote Sensing of Environment* 22 (3): 343–365. doi:10.1016/0034-4257(87)90088-5.

- He, K., J. Sun, and X. Tang. 2013. "Guided Image Filtering." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35 (6): 1397–1409. doi:10.1109/TPAMI.2012.213.
- Huang, W., L. Xiao, Z. Wei, H. Liu, and S. Tang. 2015. "A New Pan-Sharpener Method with Deep Neural Networks." *IEEE Geoscience and Remote Sensing Letters* 12 (5): 1037–1041. doi:10.1109/LGRS.2014.2376034.
- Jiang, C., H. Zhang, H. Shen, and L. Zhang. 2014. "Two-Step Sparse Coding for the Pan-Sharpener of Remote Sensing Images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (5): 1792–1805. doi:10.1109/JSTARS.2013.2283236.
- Khan, M. M., J. Chanussot, L. Condat, and A. Montanvert. 2008. "Indusion: Fusion of Multispectral and Panchromatic Images Using the Induction Scaling Technique." *IEEE Geoscience and Remote Sensing Letters* 5 (1): 98–102. doi:10.1109/LGRS.2007.909934.
- Laben, C. A., and B. V. Brower. 2000. "Process for Enhancing the Spatial Resolution of Multispectral Imagery Using Pan-Sharpener." In *Google Eastman Kodak, US Patent* (6011875).
- Li, S., X. Kang, and H. Jianwen. 2013. "Image Fusion with Guided Filtering." *IEEE Transactions on Image Processing* 22 (7): 2864–2875. doi:10.1109/TIP.2013.2244222.
- Liu, J. G. 2000. "Smoothing Filter-Based Intensity Modulation: A Spectral Preserve Image Fusion Technique for Improving Spatial Details." *International Journal of Remote Sensing* 21 (18): 3461–3472. doi:10.1080/014311600750037499.
- Mallat, S. G. 1989. "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11 (7): 674–693. doi:10.1109/34.192463.
- Masi, G., D. Cozzolino, L. Verdoliva, and G. Scarpa. 2016. "Pansharpening by Convolutional Neural Networks." *Remote Sensing* 8 (7): 594. doi:10.3390/rs8070594.
- Möller, M., T. Wittman, A. L. Bertozzi, and M. Burger. 2012. "A Variational Approach for Sharpening High Dimensional Images." *SIAM Journal on Imaging Sciences* 5 (1): 150–178. doi:10.1137/100810356.
- Otazu, X., M. González-Audicana, O. Fors, and N. Jorge. 2005. "Introduction of Sensor Spectral Response into Image Fusion Methods. Application to Wavelet-Based Methods." *IEEE Transactions on Geoscience and Remote Sensing* 43 (10): 2376–2385. doi:10.1109/TGRS.2005.856106.
- Pardo-Igúzquiza, E., V. F. Rodríguez-Galiano, M. Chica-Olmo, and P. M. Atkinson. 2011. "Image Fusion by Spatially Adaptive Filtering Using Downscaling Cokriging." *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (3): 337–346. doi:10.1016/j.isprsjprs.2011.01.001.
- Pardo-Igúzquiza, E., M. Chica-Olmo, and P. M. Atkinson. 2006. "Downscaling Cokriging for Image Sharpening." *Remote Sensing of Environment* 102 (1): 86–98. doi:10.1016/j.rse.2006.02.014.
- Ranchin, T., and L. Wald. 2000. "Fusion of High Spatial and Spectral Resolution Images: The ARSIS Concept and Its Implementation." *Photogrammetric Engineering and Remote Sensing* 66 (1): 49–61.
- Tang, Y., P. M. Atkinson, and J. Zhang. 2015. "Downscaling Remotely Sensed Imagery Using Area-To-Point Cokriging and Multiple-Point Geostatistical Simulation." *ISPRS Journal of Photogrammetry and Remote Sensing* 101: 174–185. doi:10.1016/j.isprsjprs.2014.12.016.
- Thomas, C., T. Ranchin, L. Wald, and J. Chanussot. 2008. "Synthesis of Multispectral Images to High Spatial Resolution: A Critical Review of Fusion Methods Based on Remote Sensing Physics." *IEEE Transactions on Geoscience and Remote Sensing* 46 (5): 1301–1312. doi:10.1109/TGRS.2007.912448.
- Tu, T.-M., S.-C. Su, H.-C. Shyu, and P. S. Huang. 2001. "A New Look at IHS-like Image Fusion Methods." *Information Fusion* 2 (3): 177–186. doi:10.1016/S1566-2535(01)00036-7.
- Vivone, G., L. Alparone, J. Chanussot, M. D. Mura, A. Garzelli, G. A. Licciardi, R. Restaino, and L. Wald. 2015. "A Critical Comparison among Pansharpening Algorithms." *IEEE Transactions on Geoscience and Remote Sensing* 53 (5): 2565–2586. doi:10.1109/TGRS.2014.2361734.
- Vivone, G., R. Restaino, M. D. Mura, G. Licciardi, and J. Chanussot. 2014. "Contrast and Error-Based Fusion Schemes for Multispectral Image Pansharpening." *IEEE Geoscience and Remote Sensing Letters* 11 (5): 930–934. doi:10.1109/LGRS.2013.2281996.
- Wald, L., and T. Ranchin. 2002. "Liu'Smoothing Filter-Based Intensity Modulation: A Spectral Preserve Image Fusion Technique for Improving Spatial Details." *International Journal of Remote Sensing* 23 (3): 593–597. doi:10.1080/01431160110088772.

- Wang, Q., W. Shi, and P. M. Atkinson. 2016. "Area-To-Point Regression Kriging for Pan-Sharpener." *ISPRS Journal of Photogrammetry and Remote Sensing* 114: 151–165. doi:[10.1016/j.isprsjprs.2016.02.006](https://doi.org/10.1016/j.isprsjprs.2016.02.006).
- Wang, Z., D. Ziou, C. Armenakis, L. Deren, and L. Qingquan. 2005. "A Comparative Analysis of Image Fusion Methods." *IEEE Transactions on Geoscience and Remote Sensing* 43 (6): 1391–1402. doi:[10.1109/TGRS.2005.846874](https://doi.org/10.1109/TGRS.2005.846874).
- Xu, Q., L. Bo, Y. Zhang, and L. Ding. 2014. "High-Fidelity Component Substitution Pansharpening by the Fitting of Substitution Data." *IEEE Transactions on Geoscience and Remote Sensing* 52 (11): 7380–7392. doi:[10.1109/TGRS.2014.2311815](https://doi.org/10.1109/TGRS.2014.2311815).
- Yuhas, R. H., A. F. H. Goetz, and J. W. Boardman. 1992. "Discrimination among Semi-Arid Landscape Endmembers Using the Spectral Angle Mapper (SAM) Algorithm." *Summaries of the Third Annual JPL Airborne Geoscience Workshop, JPL Publ. 92-14 1*: 147–149.
- Zhou, W., and A. C. Bovik. 2002. "A Universal Image Quality Index." *IEEE Signal Processing Letters* 9 (3): 81–84. doi:[10.1109/97.995823](https://doi.org/10.1109/97.995823).