Compensation Details-Based Injection Model for Remote Sensing Image Fusion

Yong Yang, *Senior Member, IEEE*, Lei Wu, Shuying Huang, *Member, IEEE*, Jiancheng Sun, *Senior Member, IEEE*, Weiguo Wan, and Jiahua Wu

***Abstract*— Remote sensing image fusion has a potential spectral distortion problem due to the global/local spectral and spatial correlations between panchromatic (PAN) and multispectral (MS) images. To overcome this problem, in this letter, a compensation details-based injection (CDI) fusion model is presented from a new perspective of compensatory learning. In contrast to the traditional method, the two categories of details, namely, the PAN details and the CD, are considered to compensate for the spatial and spectral differences between low-resolution MS (LRMS) and high-resolution MS images. To obtain the CD, a robust sparse representation was employed to calculate the difference between the PAN and MS images during the fusion. The CD combined with the PAN details extracted by a multiscale-guided filter are then injected into the upsampled LRMS image to achieve a fused image. Extensive experiments were undertaken on several image data sets, and the results demonstrate the effectiveness of the proposed CDI method.**

***Index Terms*— Compensation details (CD), detail injection scheme, remote sensing image fusion, robust sparse representation (RSR).**

I. Introduction

IN MANY practical applications such as land-use classification, map updating, water-quality evaluation [1], image segmentation, and object detection [2], [3], a high spatial resolution multispectral (MS) image is essential. However, due to the signal-to-noise ratio limits and transmission bottlenecks [4], commercial optical satellites such as IKONOS, QuickBird, and WorldView-2 can only provide a high spatial resolution panchromatic (PAN) image and low spatial resolution MS image, simultaneously. The process of remote sensing image fusion can make full use of spatial and spectral information by combining a PAN image and its corresponding low-resolution MS (LRMS) image.

Manuscript received October 23, 2017; revised December 16, 2017 and January 30, 2018; accepted February 25, 2018. Date of publication March 16, 2018; date of current version April 20, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61662026 and Grant 61462031, in part by the Natural Science Foundation of Jiangxi Province under Grant 20161ACB21015, and in part by the Project of the Education Department of Jiangxi Province under Grant KJLD14031. *(Corresponding author: Shuying Huang.)*

Y. Yang, L. Wu, and J. Wu are with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: greatyangy@126.com; jxxywulei@126.com; mrwjh2017@163.com).

S. Huang and J. Sun are with the School of Software and Communication Engineering, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: shuyinghuang2010@126.com; sunjc73@gmail.com).

W. Wan is with the Division of Computer Science and Engineering, Chonbuk National University, Jeonju 561756, South Korea (e-mail: weiguo@jbnu.ac.kr).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2018.2810219

Over more than two decades, a large number of image fusion methods have been developed. Among these methods, a classical remote sensing fusion algorithm was developed based on the extraction of spatial details from the PAN image, followed by injection of these extracted details into the upsampled LRMS image. As reported in [5], these methods, which are based on component substitution (CS) or multiresolution analysis (MRA), can be modeled with a detail injection scheme. The CS method is based on space transformation and a CS strategy. The popular CS-based algorithms include intensity–hue–saturation (IHS) transform [6], principal component analysis [7], the Gram–Schmidt method [8], and an IHS transform coupled with the Fourier domain filtering method [9]. These approaches can achieve good performance in spatial resolution but usually cause serious spectral distortion. An alternative technique is based on MRA, which is generally obtained with linear decomposition methods, such as à trous wavelet transform (ATWT) [10], nonsubsampled contourlet transform [11], and nonsubsampled Shearlet transform (NSST) [12]. The MRA-based methods can effectively preserve the spectral information in the fused image, but are generally not satisfactory in terms of spatial enhancement.

In recent years, because sparse representation (SR) reconstruction has super-resolution capability and robustness, it has been widely used to improve the quality of the fused image by combining the CS- or MRA-based methods (see [12]–[14]). In [14], a robust SR (RSR) model is proposed to improve the robustness of the SR model to non-Gaussian noise or sparse but strong "outliers." As we know, all of the aforementioned CS-, MRA-, or SR-based methods depend on the detail injection models. However, the current detail injection methods estimate the difference between the PAN image and the I component of the MS image or the approximate version of the PAN image to obtain the desired details. In this way, the potential spectral distortion is inevitably produced because of the global/local spectral and spatial correlations between the PAN and MS images. To overcome this problem, a new injection model called a compensation details-based injection (CDI) model is presented for remote sensing image fusion in this letter. The proposed model reconstructs the CD to compensate for the global/local difference between the PAN and MS images. To the best of our knowledge, the compensatory scheme is the first work proposed in this field. The main contributions of this letter are twofold: 1) we define a CDI model for remote sensing image fusion, in which the PAN details and the CD are sufficiently considered and 2) RSR is introduced to extract the CD, which are

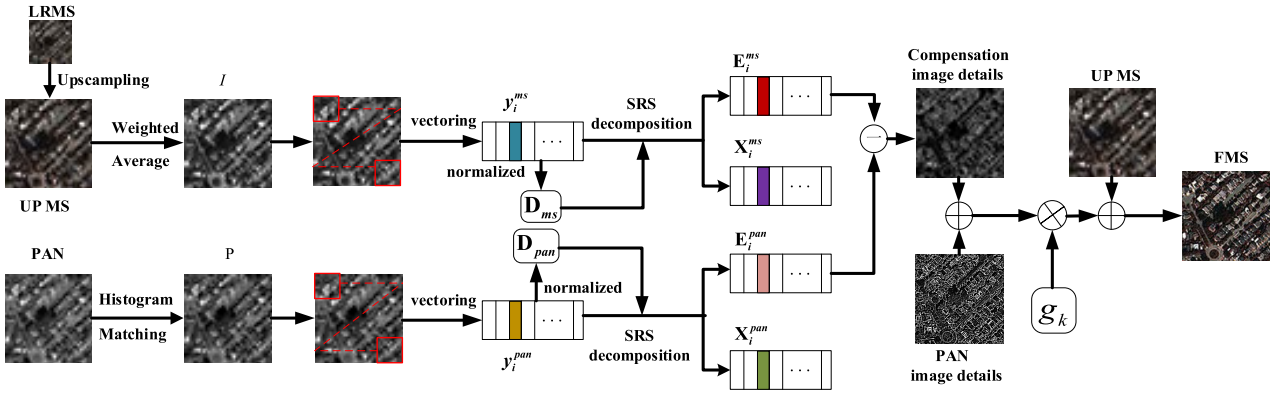

Fig. 1. Flowchart of the proposed approach. UPMS is the upsampled LRMS image. $\mathbf{D}_{\text{pan}}$ and $\mathbf{D}_{\text{ms}}$ are the normalization dictionaries of the PAN image and the intensity component I of the MS image, respectively. $\mathbf{y}_i^{\text{ms}}$, $\mathbf{y}_i^{\text{pan}}$, $\mathbf{X}_i^{\text{ms}}$, $\mathbf{X}_i^{\text{pan}}$, $\mathbf{E}_i^{\text{ms}}$, and $\mathbf{E}_i^{\text{pan}}$ are the ith vector of the image patches, reconstruction matrixes, and the reconstruction error matrixes of the PAN image and I component. The PAN image details were obtained by guided filter. g_k is the modulation coefficient. FMS is the fused result.

employed to compensate for the spectral distortion in the fused image.

II. Robust Sparse Representation Model

RSR can be regarded as the improved SR model [14]. Let $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N]$ be an observed data matrix of size $d \times N$ and $\mathbf{y}_i \in \mathbb{R}^d$ is a column vector. Suppose it is partially corrupted by errors or noise $\mathbf{E} \in \mathbb{R}^{d\times N}$; then, given a dictionary $\mathbf{D} \in \mathbb{R}^{d\times M}$ with M prototype atoms, the RSR model is defined as follows:

$$\min_{\mathbf{X},\mathbf{E}} \|\mathbf{X}\|_0 + \lambda\|\mathbf{E}\|_{2,0} \quad \text{s.t. } \mathbf{Y} = \mathbf{D}\mathbf{X} + \mathbf{E} \tag{1}$$

where the matrix $\mathbf{X} \in \mathbb{R}^{M\times N}$ denotes the sought-after matrix of coefficients and $\|\mathbf{X}\|_0$ denotes its l_0-norm, which is the number of nonzero entries in $\mathbf{X}$. The matrix $\mathbf{E} \in \mathbb{R}^{M\times N}$ is used to address sample-specific corruptions and outliers, and $\|\mathbf{E}\|_{2,0}$ signifies its $l_{2,0}$-norm, which is the number of nonzero columns in $\mathbf{E}$. The parameter $\lambda > 0$ is used to balance the effects of the two components in (1).

As suggested in [14], the defined RSR model can be relaxed to create the following convex optimization problem:

$$\min_{\mathbf{X},\mathbf{E}} \|\mathbf{X}\|_1 + \lambda\|\mathbf{E}\|_{2,1} \quad \text{s.t. } \mathbf{Y} = \mathbf{D}\mathbf{X} + \mathbf{E} \tag{2}$$

where $\|\mathbf{X}\|_1$ denotes the l_1-norm of matrix $\mathbf{X}$ and is defined as $\|\mathbf{X}\|_1 = \sum_j \sum_i |\mathbf{X}(i, j)|$. $\|\mathbf{E}\|_{2,1}$ is the $l_{2,1}$-norm of the matrix $\mathbf{E}$ and is defined as $\|\mathbf{E}\|_{2,1} = \sum_j (\sum_i (\mathbf{E}(i, j))^2)^{1/2}$. $\mathbf{X}(i, j)$ and $\mathbf{E}(i, j)$ are the (i, j)th entries in matrixes $\mathbf{X}$ and $\mathbf{E}$, respectively. The optimization problem in (2) is convex and can be solved with a linearized alternating direction method with adaptive penalty (LADMAP) [14].

III. Proposed Method

In this section, we propose to reduce spectral distortion with the help of the CD and to enhance the spatial resolution by the injection of the PAN image details. The process was realized using the proposed CDI model, which will be described in detail, and the framework of the proposed method is illustrated in Fig. 1.

A. CDI Model

In remote sensing fusion, the goal is to make the fused image [denoted by fused multispectral (FMS)] as similar as possible to the ideal FMS image ($\text{FMS}^{\text{ideal}}$) [13]. Mathematically, we have

$$\widehat{\text{FMS}}_k = \arg\min_{\text{FMS}_k} \left\|\text{FMS}_k^{\text{ideal}} - \text{FMS}_k\right\|_p, \quad k = 1, \ldots, B \tag{3}$$

$$\text{FMS}_k = \text{LMS}_k + \text{HRD}, \quad k = 1, \ldots, B \tag{4}$$

where $\|\cdot\|_p$ is the l_p-norm, B is the number of spectral bands of the LRMS images, LMS_k is the kth band of the upsampled LRMS image, and HRD is the high-frequency detail information. Substituting (4) in (3), (3) can be converted into the following equation:

$$\widehat{\text{HRD}} = \arg\min_{\text{HRD}} \left\|(\text{FMS}_k^{\text{ideal}} - \text{LMS}_k) - \text{HRD}\right\|_p, \quad k = 1, \ldots, B. \tag{5}$$

Equation (5) states that the HRD component should approximate the spatial information of the $\text{FMS}^{\text{ideal}}$ image, which is actually unavailable. HRD can be obtained by extracting the details from the PAN image. If the HRD contains all of the details in the PAN image, and the MS image is highly correlated with PAN image, the HRD component should equate to the spatial information of the $\text{FMS}^{\text{ideal}}$. However, in fact, the remote sensing images, which are produced using commercial optical satellites, have the intrinsic properties [15], such as the inconsistent spectral wavelength between the MS and PAN images, the uneven distribution of the land-cover type in the image, and the object occultation and contrast inversion. These properties indicate that the MS image is lowly correlated with the PAN image and the difference information must exist between the PAN and MS images. The difference information results in global/local instability or dissimilarity between the PAN and MS images. In this letter, the difference information that exists in the MS image and is absent in the PAN image is referred to as CD. Thus, the HRD in (5) can be defined mathematically as follows:

$$\text{HRD} = \text{CD} + \text{PAN}^{\text{detail}} \tag{6}$$

where $\mathrm{PAN}^{\mathrm{detail}}$ represents the PAN image details. From (6), we can determine that it is essential to reconstruct the CD, which are used to compensate for the lack of PAN image details. At the same time, another category of detail, the PAN image details are extracted to enhance the spatial resolution of the LRMS image, and the two categories of detail should be adaptively injected into the upsampled LRMS image. Therefore, the proposed CDI model can be defined as follows:

$$\mathrm{FMS}_k = \mathrm{LMS}_k + \mathrm{g}_k(\mathrm{CD} + \mathrm{PAN}^{\mathrm{detail}}), \quad k = 1, \ldots, B \tag{7}$$

where g_k is the modulation coefficient [4].

B. Obtaining the Compensation Details With RSR

In view of the analysis in Section III-A, the CD reflect the difference between the PAN and MS images, and the quality of the extracted CD plays an important role in the success of the proposed method. In this section, we proposed to employ RSR to decompose the source images, because sparse reconstruction based on dictionary learning can make the reconstructed information adapt to the available source image, and the reconstructed result will be highly relevant to the input image. Furthermore, RSR is an improved SR model that can divide the original image into two subimages: a blur version subimage of the original image and a subimage that contains the details. They are also called the reconstruction matrix, **DX**, and the reconstruction error matrix, **E** [14]. The detailed process of RSR can be described as follows. First, through a sliding-window technique with overlapping areas of size $p_x \times p_y$, from top to bottom, the PAN image and the intensity component of the MS image are divided into V patches: $\{y_i^{\mathrm{pan}} | i = 0, 1, \ldots, V-1\}$ and $\{y_i^{\mathrm{ms}} | i = 0, 1, \ldots, V-1\}$. Then, based on the literature [14], the normalization dictionaries $\mathbf{D}_{\mathrm{pan}}$ and $\mathbf{D}_{\mathrm{ms}}$ can be obtained by the following formulas:

$$\mathbf{D}_{\mathrm{pan}} = \left[\frac{\mathbf{y}_0^{\mathrm{pan}}}{\left\|\mathbf{y}_0^{\mathrm{pan}}\right\|_2}, \frac{\mathbf{y}_1^{\mathrm{pan}}}{\left\|\mathbf{y}_1^{\mathrm{pan}}\right\|_2}, \ldots, \frac{\mathbf{y}_{V-1}^{\mathrm{pan}}}{\left\|\mathbf{y}_{V-1}^{\mathrm{pan}}\right\|_2} \right] \tag{8}$$

$$\mathbf{D}_{\mathrm{ms}} = \left[\frac{\mathbf{y}_0^{\mathrm{ms}}}{\left\|\mathbf{y}_0^{\mathrm{ms}}\right\|_2}, \frac{\mathbf{y}_1^{\mathrm{ms}}}{\left\|\mathbf{y}_1^{\mathrm{ms}}\right\|_2}, \ldots, \frac{\mathbf{y}_{\mathrm{V\text{-}1}}^{\mathrm{ms}}}{\left\|\mathbf{y}_{\mathrm{V\text{-}1}}^{\mathrm{ms}}\right\|_2} \right]. \tag{9}$$

Subsequently, $\mathbf{X}_{\mathrm{pan}}$, $\mathbf{X}_{\mathrm{ms}}$, $\mathbf{E}_{\mathrm{pan}}$, and $\mathbf{E}_{\mathrm{ms}}$ can be calculated using the LADMAP algorithm. In this letter, the aim is to obtain the $\mathbf{E}_{\mathrm{pan}}$ and $\mathbf{E}_{\mathrm{ms}}$ by fixing one item in X and E to calculate another item. The process of iterating over the loop is given as follows:

$$\mathbf{E}_{\mathrm{pan}}^{j+1}(:, i) = \begin{cases} \dfrac{\| \mathbf{G}_{\mathrm{pan}}(:, i) \|_2 - \lambda/\mu^j}{\| \mathbf{G}_{\mathrm{pan}}(:, i) \|_2} \mathbf{G}_{\mathrm{pan}}(:, i), \\ \qquad \text{if } \| \mathbf{G}_{\mathrm{pan}}(:, i) \|_2 \geq \lambda/\mu^j \\ 0, \qquad \text{otherwise} \end{cases} \tag{10}$$

$$\mathbf{E}_{\mathrm{ms}}^{j+1}(:, i) = \begin{cases} \dfrac{\| \mathbf{G}_{\mathrm{ms}}(:, i) \|_2 - \lambda/\mu^j}{\| \mathbf{G}_{\mathrm{ms}}(:, i) \|_2} \mathbf{G}_{\mathrm{ms}}(:, i), \\ \qquad \text{if } \| \mathbf{G}_{\mathrm{ms}}(:, i) \|_2 \geq \lambda/\mu^j \\ 0, \qquad \text{otherwise} \end{cases} \tag{11}$$

where j is the number of iterations. $\mathbf{G} = \mathbf{Y} - \mathbf{D}\mathbf{X}^j + (L^j/\mu^j)$. $\mathbf{E}(:, i)$ and $\mathbf{G}(:, i)$ denote the ith column of the matrix $\mathbf{E}$ and $\mathbf{G}$, respectively. $\mathbf{X}$ can be calculated as follows:

$$\mathbf{X}^{j+1} = \mathbf{S}_{\frac{1}{\eta\mu^j}} \left(\mathbf{X}^j - \frac{1}{\eta}\mathbf{D}^T \left(\mathbf{D}\mathbf{X}^j - \mathbf{Y} + \mathbf{E}^{j+1} - \frac{L^j}{\mu^j} \right) \right) \tag{12}$$

where $S_\tau(x)$ is the threshold function. η is the l_2-norm of the dictionary $\mathbf{D}$. $\mathbf{D}^T$ is the transposed matrix of the dictionary $\mathbf{D}$.

After achieving the reconstruction error matrixes $\mathbf{E}_{\mathrm{pan}}$ and $\mathbf{E}_{\mathrm{ms}}$, the clear and unclear edges and textures from the PAN and MS images are finally obtained. The CD are clear in $\mathbf{E}_{\mathrm{ms}}$ but are unclear in $\mathbf{E}_{\mathrm{pan}}$, according to the analysis of the CD in Section III-A. Considering that the MS image has low spatial resolution and the PAN image has high spatial resolution, we proposed to calculate the difference between the reconstruction error matrixes of PAN and MS images to obtain the CD with $\mathbf{E}_{\mathrm{pan}}$ and $\mathbf{E}_{\mathrm{ms}}$. Therefore, the process of CD can be obtained by defining the following formula:

$$\begin{aligned} \mathrm{CD}(i, j) = \mathrm{abs}(&\mathbf{E}_{\mathrm{pan}}(i, j) - \mathbf{E}_{\mathrm{ms}}(i, j)) \\ &\text{s.t. } \mathbf{E}_{\mathrm{pan}}(i, j) - \mathbf{E}_{\mathrm{ms}}(i, j) < 0 \end{aligned} \tag{13}$$

where abs$(\cdot)$ is the function used to calculate the absolute value and(i, j) indicates the coordinates of the pixel.

C. Extracting the PAN Image Details by Guided Filter

In this section, the guided filter [16] is adopted to extract the PAN image details, because it can acquire the variation tendency of the guidance image and preserve the major information of the input image, simultaneously. Meanwhile, it makes the extracted PAN image details similar to the MS image in the geometric construction. The process of the PAN image details can be defined as follows:

$$P_l = \mathrm{GF}(P_{l-1}, I) \tag{14}$$

$$\mathrm{PAN}_l^{\mathrm{detail}} = P_{l-1} - P_l \tag{15}$$

$$\mathrm{PAN}^{\mathrm{detail}} = \sum_{l=1}^{L} \mathrm{PAN}_l^{\mathrm{detail}} \tag{16}$$

where GF$(\cdot)$ represents the filtering operator of the guided filter. I is the intensity component of the MS image, used as the guidance image. P_{l-1} and P_l are the input and output, respectively. When $l = 1$, P_{l-1} is the histogram-matched PAN image. $\mathrm{PAN}_l^{\mathrm{detail}}$ is the high-resolution detail information of the PAN image at the l level. As seen earlier, through Sections III-B and III-C, the CD and $\mathrm{PAN}_l^{\mathrm{detail}}$ can be obtained. Therefore, the fused image can ultimately be achieved by (7).

IV. Experimental Results and Analysis

A. Experimental Setup

In this letter, we conducted degraded and real experiments on two data sets collected by the QuickBird and WorldView-2 satellites. The spatial resolution of the MS and PAN images is equal to 0.7 and 2.8 m for QuickBird, and 0.5 and 2 m for WorldView-2 [17], respectively. The degraded PAN and MS images are obtained by downsampling with a factor of four from the original PAN and MS

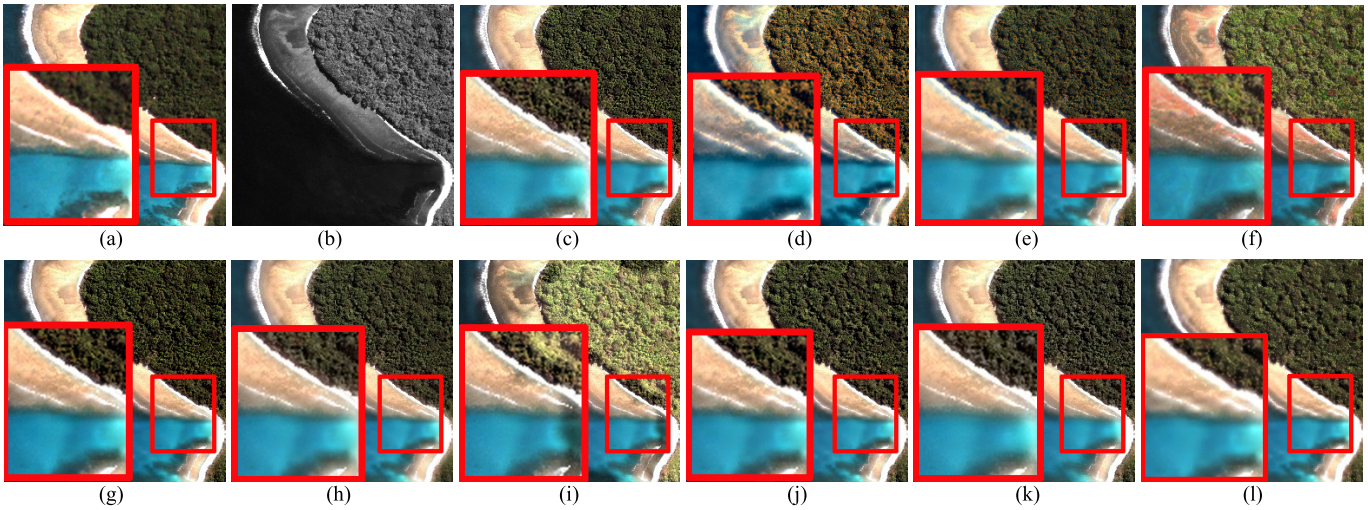

Fig. 2. QuickBird image fusion results. (a) MS image. (b) PAN image. (c) Ehler method. (d) GSA method. (e) CBD method. (f) IAIHS method. (g) BFLP method. (h) MM method. (i) NSST_SR method. (j) IMG method. (k) ATWT method. (l) Proposed method.

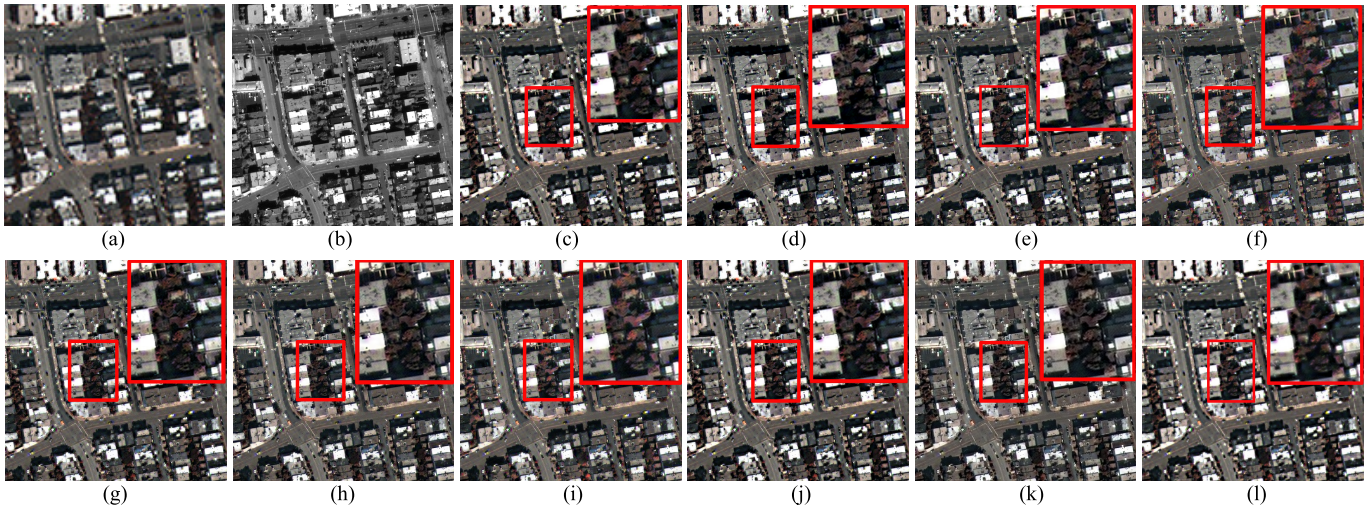

Fig. 3. WorldView-2 image fusion results. (a) MS image. (b) PAN image. (c) Ehler method. (d) GSA method. (e) CBD method. (f) IAIHS method. (g) BFLP method. (h) MM method. (i) NSST_SR method. (j) IMG method. (k) ATWT method. (l) Proposed method.

images, respectively. The original MS image is regarded as the reference image. The low-resolution spectral bands are upsampled (4×4) and interpolated to the scale of the PAN image. Two types of quality assessment approaches, with and without a reference image, were used, including a subjective evaluation and a quantitative assessment [4]. Experiments on degraded data were evaluated using the correlation coefficient (CC), the universal image quality indices (UIQIs), the root-mean-square error (RMSE), the relative average spectral error (RASE), and the erreur relative global adimensionnelle de synthese (ERGAS). The real data from the data sets were evaluated using the index with no reference (QNR).

The proposed method was compared with nine popular fusion methods, i.e., the Ehler fusion method (Ehler) [18], the Gram Schmidt adaptive (GSA) method [19], the ATWT method [10], the improved adaptive IHS (IAIHS) method [20], the bilateral filter luminance proportional (BFLP) method [21], the matting model (MM)-based method [22], the NSST and NSST_SR methods [12], a method based on AIHS and multiscale-guided filter [IHS multiscale guided (IMG)] [4], and the context-based decision (CBD) method [23]. All of the comparison methods used in this letter are open source codes offered by the corresponding authors. In the proposed method, the parameters $\lambda = 10^{-9}$ and $\tau = 10^{-10}$ and the two-level decomposition for guided filter were set as a reference [4], with the parameters$\rho = 1.1$, $\varepsilon = 0.05$, $\mu = 10^{-6}$, and $\mu_{\max} = 10^{10}$ and the size of the patches of the training sample were 8×8 as a reference [14]

B. Experimental Results

In this section, the degraded experiment is shown in Fig. 2. Fig. 2(a) is the original MS image that was used as the reference image to compare with the fused images. Fig. 2(b) shows the degraded PAN image. The corresponding fusion results are given in Fig. 2(c)–(l). The real experiment is shown in Fig. 3. Fig. 3(a) is the upsampled MS (UPMS) image. Fig. 3(b) shows the degraded PAN image. The corresponding fusion results are shown in Fig. 3(c)–(l). The objective assessments of the fusion results of Figs. 2 and 3 are shown in Table I.

TABLE I
QUANTITATIVE EVALUATION OF THE EXPERIMENTAL RESULTS SHOWN IN FIGS. 2 AND 3

	QuickBird					WorldView-2		
Method	CC	UIQI	RASE	RMSE	ERGAS	D_λ	D_S	QNR
Ehler	0.9595	0.9675	21.2631	19.346	5.4475	0.0056	0.0533	0.9414
GSA	0.9228	0.9449	31.685	28.828	8.982	**0.0027**	0.0654	0.9320
CBD	0.9776	0.9807	16.381	14.904	4.3321	0.0046	0.0564	0.9392
IAIHS	0.8812	0.8931	36.365	33.086	9.1263	0.0047	0.0704	0.9252
BFLP	0.9718	0.9717	20.491	18.643	5.7948	0.0040	0.0552	0.9410
MM	0.9737	0.976	16.743	15.233	4.2563	0.0066	0.0643	0.9295
NSST_SR	0.6757	0.8028	60.648	55.180	13.035	0.0035	0.0748	0.9219
IMG	0.9689	0.9740	18.662	16.979	4.8199	0.0048	0.0553	0.9401
ATWT	0.9596	0.9649	20.710	18.843	5.2726	0.0106	0.0683	0.9218
Proposed	**0.9784**	**0.9817**	**15.715**	**14.298**	**3.8635**	0.0033	**0.0490**	**0.9479**

As can be seen from Fig. 2, the fusion results obtained by the GSA, IAIHS, and NSST_SR methods suffer from serious spectral distortion. The result of the CBD method exhibits serious spectral distortion in the forest area. The result of the Ehler method has good spatial quality, but has some spectral distortion in the forest area. The results of the BFLP and MM methods have some color distortion in the forest area. The result of the IMG method can effectively preserve the spatial quality, but the spectral information is not sufficient. There is obvious spectral distortion in the forest area of the ATWT method result. The proposed method achieves the best effect in the visual inspection. In Fig. 3, it is clear that the fusion results obtained by the GSA, IAIHS, and NSST_SR methods suffer from obvious spectral distortion. Compared to the above-mentioned methods, the ATWT, CBD, and MM methods can obtain relatively better spectral effect, but they still exhibit various degrees of spectral distortion in red areas. There is not much difference among the results of the Ehler, BFLP, IMG methods, and the proposed method in visual contrast, and they achieve the best performance with respect to the spatial and spectral information. In addition to subjective evaluation, from Table I, we can also observe that the proposed method obtains the best values of all indices except for D_λ (our method is the second best) in all the fusion methods.

V. CONCLUSION

In this letter, we consider the intrinsic properties of the remote sensing images and propose to obtain the CD to reduce the potential spectral distortion. A new CDI model for remote sensing image fusion was thus developed by injecting two categories of detail to achieve the spectral and spatial compensation in the fused image. In the CDI model, the CD were obtained with RSR, and the spatial details were extracted with a multiscale-guided filter. Then, the two categories of detail were injected into the upsampled LRMS image to obtain the fused image. Experimental results on the degraded and real data from QuickBird and WorldView-2 satellites indicate that the proposed method can achieve better performance than some related and popular existing fusion methods.

REFERENCES

[1] N.-B. Chang, K. Bai, S. Imen, C.-F. Chen, and W. Gao, "Multisensor satellite image fusion and networking for all-weather environmental monitoring," *IEEE Syst. J.*, to be published, doi: 10.1109/JSYST.2016.2565900.

[2] Q. Wang, G. Zhu, and Y. Yuan, "Multi-spectral dataset and its application in saliency detection," *Comput. Vis. Image Understand.*, vol. 117, no. 12, pp. 1748–1754, Dec. 2013.

[3] Q. Wang, P. Yan, Y. Yuan, and X. Li, "Multi-spectral saliency detection," *Pattern Recognit. Lett.*, vol. 34, no. 1, pp. 34–41, 2013.

[4] Y. Yang, W. Wan, S. Huang, F. Yuan, S. Yang, and Y. Que, "Remote sensing image fusion based on adaptive IHS and multiscale guided filter," *IEEE Access*, vol. 4, pp. 4573–4582, 2016.

[5] G. Vivone *et al.*, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.

[6] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, Sep. 2001.

[7] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A comparative analysis of image fusion methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1391–1402, Jun. 2005.

[8] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," U.S. Patent 6 011 875, Jan. 4, 2000.

[9] J. Zhang, "Multi-source remote sensing data fusion: Status and trends," *Int. J. Image Data Fusion*, vol. 1, no. 1, pp. 5–24, Nov. 2010.

[10] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1204–1211, May 1999.

[11] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, Oct. 2006.

[12] A.-U. Moonon, J. Hu, and S. Li, "Remote sensing image fusion method based on nonsubsampled shearlet transform and sparse representation," *Sens. Imag. Int. J.*, vol. 16, no. 1, pp. 1–18, Nov. 2015.

[13] M. Ghahremani and H. Ghassemian, "Remote sensing image fusion using ripplet transform and compressed sensing," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 502–506, Mar. 2015.

[14] Q. Zhang and M. D. Levine, "Robust multi-focus image fusion using multi-task sparse representation and spatial context," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2045–2058, May 2016.

[15] H. Wang, W. Jiang, C. Lei, S. Qin, and J. Wang, "A robust image fusion method based on local spectral and spatial correlation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 2, pp. 454–458, Feb. 2014.

[16] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.

[17] *QuickBird Datasets*. Accessed: Jul. 12, 2016. [Online]. Available: http://www.glcf.umiacs.umd.edu/data/

[18] Y. Ling, M. Ehlers, E. L. Usery, and M. Madden, "FFT-enhanced IHS transform method for fusing high-resolution satellite images," *ISPRS J. Photogram. Remote Sens.*, vol. 61, no. 6, pp. 381–392, Feb. 2007.

[19] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS+Pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.

[20] Y. Leung, J. Liu, and J. Zhang, "An improved adaptive intensity-hue-saturation method for the fusion of remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp. 985–989, May 2014.

[21] N. H. Kaplan and I. Erer, "Bilateral filtering-based enhanced pansharpening of multispectral satellite images," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 11, pp. 1941–1945, Nov. 2014.

[22] X. Kang, S. Li, and J. A. Benediktsson, "Pansharpening with matting model," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5088–5099, Aug. 2014.

[23] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and PAN imagery," *Photogram. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.