

An Online Coupled Dictionary Learning Approach for Remote Sensing Image Fusion

Min Guo, Hongyan Zhang, *Member, IEEE*, Jiayi Li, *Student Member, IEEE*,
Liangpei Zhang, *Senior Member, IEEE*, and Huanfeng Shen, *Senior Member, IEEE*

Abstract—Most earth observation satellites, such as IKONOS, QuickBird, GeoEye, and WorldView-2, provide a high spatial resolution (HR) panchromatic (Pan) image and a multispectral (MS) image at a lower spatial resolution (LR). Image fusion is an effective way to acquire the HR MS images that are widely used in various applications. In this paper, we propose an online coupled dictionary learning (OCDL) approach for image fusion, in which a superposition strategy is applied to construct the coupled dictionaries. The constructed coupled dictionaries are further developed via an iterative update to ensure that the HR MS image patch can be almost identically reconstructed by multiplying the HR dictionary and the sparse coefficient vector, which is solved by sparsely representing its counterpart LR MS image patch over the LR dictionary. The fusion results from IKONOS and WorldView-2 data show that the proposed fusion method is competitive or even superior to the other state-of-the-art fusion methods.

Index Terms—Coupled dictionary, image fusion, remote sensing imagery, sparse representation (SR).

I. INTRODUCTION

AS A POWERFUL quality improvement technique, data fusion has been gradually improved in recent years. In [1], data fusion is defined as a formal framework which includes expressed means and tools for combining and utilizing data originating from different sources. Accounting for most of the data fusion studies, image fusion is the integration of different information sources by taking advantage of the complementary spatial/spectral resolution characteristics of remote sensing imagery. For most earth observation satellites, such as IKONOS, QuickBird, GeoEye, and WorldView-2, the data provided are composed of a high spatial resolution (HR) panchromatic (Pan) image and a low spatial resolution (LR) multispectral (MS) image. The process of acquiring an HR MS image by blending an HR Pan image and its corresponding LR MS image is referred to as “image pan-sharpening.” In practice, images with high spectral and spatial resolutions are useful in an increasing

number of applications, such as feature detection [2], segmentation/classification [3], [4], and so on.

During the past two decades, a large amount of image fusion methods have been developed [5]–[7]. In [8] and [9], the fusion methods are grouped into three categories: 1) projection-substitution methods, 2) relative spectral contribution methods, and 3) methods that belong to the Amélioration de la Résolution Spatiale par Injection de Structures (ARSIS) concept. Projection-substitution methods, which transform the MS image into another space and exchange one structural component with the Pan image, are widely used and have been integrated into some commercial software packages. Among these methods, the most popular are intensity hue saturation (IHS) transformation [10], [11], principal component analysis (PCA) [12], and the Gram-Schmidt transform-based methods [13]. The relative spectral contribution methods are based on the assumption that the LR Pan image can be written as a linear combination of the original MS image, of which the Brovey transform [14] and the $P + XS$ [15] method are two successful application instances. These two types of methods can produce a noticeable increase in visual impression with a good geometrical quality, but a major drawback comes from the nonignorable spectral distortion. As for the ARSIS concept-based methods, it is assumed that the missing spatial information in the LR MS image can be inferred from the high frequencies of the HR Pan image. To be specific, details extracted from the HR Pan image by certain multi-scale or multi-resolution decomposition algorithms are injected into the LR MS image [8], [16]. The significant advantage of the ARSIS concept-based methods is the preservation of the spectral content of the original MS image. The à trous wavelet pan-sharpening (AWLP) [17] method and the context-based decision (CBD) [18] method are two effective ARSIS concept-based methods, which both lead to good fusion results.

Recently, a new image fusion branch, which transfers the image fusion problem into an image-related inverse problem resolved with the help of sparse representation (SR) and compressed sensing (CS) theory, has emerged and shown impressive fusion performances. Li and Yang [19] were the first to perform the remote sensing image fusion task from the perspective of CS [20] theory. Subsequently, Jiang *et al.* [21] extended the above model by learning a joint dictionary from the LR MS image and Pan image to make it more practical. Nevertheless, these CS-based methods require a large collection of images to train the dictionary, which is computationally expensive. To deal with this problem, Li *et al.* [22] developed a restoration-based remote sensing image fusion method with sparsity regularization, in which the dictionary is adaptively learned with the source image.

Manuscript received November 01, 2013; revised January 27, 2014; accepted February 28, 2014. Date of publication April 03, 2014; date of current version April 18, 2014. This work was supported in part by the National Basic Research Program of China (973 Program) under Grant 2011CB707105, in part by the 863 program under Grant 2013AA12A301, and in part by the National Natural Science Foundation of China under Grant 61201342 and Grant 61261130587. (Corresponding author: H. Zhang.)

M. Guo, H. Zhang, J. Li, and L. Zhang are with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: zhanghongyan@whu.edu.cn).

H. Shen is with the School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China.

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

Digital Object Identifier 10.1109/JSTARS.2014.2310781

With an effective and robust performance, the method still needs to assume a spectral composition model, which is a little complicated to implement. In [23], Zhu and Bamler proposed an image fusion method named sparse fusion of images (SparseFI) which explores the same sparse coefficient vector of the corresponding HR/LR MS image patches over the coupled dictionaries, which are offline constructed from the Pan image and its down-sampled LR version. Due to its ease of implementation and the lack of requirement for external image data, SparseFI has been considered as a promising approach with a broader application range. Recently, a two-step sparse coding strategy for the pan-sharpening of remote sensing images was proposed in [24] on the basis of the SparseFI method.

In this paper, we propose an online coupled dictionary learning (OCDL) approach for image fusion, in which we make full use of the available LR MS image and the HR Pan image to decrease the spectral distortion and preserve the spatial information of the LR MS image. In the proposed OCDL method, a superposition strategy is adopted to produce two intermediate images for the coupled dictionary construction for each band. In order to ensure that the HR MS image patch can be almost identically reconstructed by multiplying the HR dictionary and the same sparse coefficient vector, which is solved by sparsely representing its counterpart LR MS image patch over the LR dictionary, an iterative update method is utilized to update the coupled dictionaries, which can be referred to as an online dictionary learning process. The theoretical analyses and experimental results in this paper indicate that the proposed method can produce competitive fusion results, even if the Pan image has a low correlation with some of the MS bands.

The rest of the paper is structured as follows. Section II briefly describes SR in image processing and the coupled dictionary model for image fusion. Thereafter, the scheme of the proposed algorithm is reported in Section III. In Section IV, experiments with two IKONOS data sets and one WorldView-2 data set verify the effectiveness of the proposed method, with respect to the visual, spatial, and spectral quality. Finally, the conclusions are drawn in Section V.

II. RELATED WORKS

A. SR in Image Processing

Sparsity has recently been the subject of intensive research, and the field of image processing has benefitted a lot from the progress in both theory and practice [25], [26]. In the image processing approach, each signal $\mathbf{x} \in \mathbb{R}^n$ lexicographically stacking the $\sqrt{n} \times \sqrt{n}$ pixels can be sparsely represented by a suitable overcomplete dictionary $\mathbf{D} \in \mathbb{R}^{n \times k}$ ($n \ll k$) [27], each column of which corresponds to a possible image patch (also lexicographically stacking the pixel values in this patch as a vector). That is to say, signal \mathbf{x} can be represented as $\mathbf{x} = \mathbf{D}\boldsymbol{\alpha}$, which simultaneously assumes the sparsity of the coefficient vector $\boldsymbol{\alpha}$. This problem can be formulated as

$$\min \|\boldsymbol{\alpha}\|_0 \quad s.t. \quad \|\mathbf{D}\boldsymbol{\alpha} - \mathbf{x}\|_2 \leq \varepsilon \quad (1)$$

where $\|\boldsymbol{\alpha}\|_0$ denotes the number of nonzero components in $\boldsymbol{\alpha}$.

This optimization problem is NP-hard. It has been shown that the optimization problem can be converted to an ℓ_1 -norm minimization problem if the desired coefficient $\boldsymbol{\alpha}$ is sufficiently sparse [28], which converts (1) to

$$\min \|\boldsymbol{\alpha}\|_1 \quad s.t. \quad \|\mathbf{D}\boldsymbol{\alpha} - \mathbf{x}\|_2 \leq \varepsilon. \quad (2)$$

A large number of solution algorithms have been developed to solve the ℓ_1 -norm optimization problem [29], [30], with one of the classic algorithms being the LASSO algorithm [30].

B. The Coupled Dictionary Model and Its Application in Image Fusion

The coupled dictionary model was designed to solve the cross-style image synthesis problem [31], in which each style for the scene can be mutually transferred by learning the underlying mapping from the example image pairs. Suppose that we have some example image pairs from the coupled feature spaces. For convenience, we assume that the images in one space follow the style $x(S_x)$, and the images in the other space follow style $y(S_y)$. The image cross-style synthesis problem can then be formulated as follows: recover the image in style y when its corresponding description in style x is given.

The working mechanism of this model is that there is a corresponding relationship between the counterpart atoms in the coupled dictionaries, which leads to a mapping function between the sparse coefficient vectors of the image patch pairs in the coupled feature spaces. Clearly, the coupled dictionaries play an important role in this model. In general, the coupled dictionaries are simply generated by randomly sampling raw patches from the training image pairs of the same scene in the coupled spaces, or learned from the above raw patch dictionaries. Once the coupled dictionaries are constructed, each patch of style x is sparsely represented over the dictionary in the space S_x . The commonly used and effective mapping function refers to the assumption that the sparse coefficient vectors in different styles should be the same, with respect to the delicate coupled dictionary construction [22], [23], [32]. Therefore, the associated patch of style y can be reconstructed with the same sparse coefficient vector and the dictionary in the space S_y .

Image resolution enhancement is one of the classic cross-style image synthesis problems, where the coupled dictionaries refer to coupled spaces: the high- and low-resolution signal spaces in the patch-based SR [32]. Image fusion is a common method of image resolution enhancement, and the coupled dictionary model can be used to solve this problem. The SparseFI [23] method has recently been proposed as an application of the coupled dictionary model in image fusion. Since an HR Pan image \mathbf{Y}^h and its down-sampled LR version \mathbf{Y}^l can be directly utilized, we are able to directly construct the coupled dictionaries without an extra image data set. In this way, in the SparseFI method, the coupled dictionaries, which consist of an LR dictionary \mathbf{D}^l and an HR dictionary \mathbf{D}^h , are directly constructed from the Pan image and its down-sampled LR version. To be specific, we down-sample the HR Pan image to the same scale as the LR MS image by using bicubic interpolation, and we then get an LR Pan image. The LR dictionary \mathbf{D}^l is generated by sampling raw patches from the LR Pan image with overlapping areas.

Likewise, the HR dictionary \mathbf{D}^h is generated by tiling the HR Pan image into patches that are R times the size of the LR Pan image patches, so that each HR patch corresponds to an LR patch. However, with a small amount of data, a dictionary constructed directly from the Pan image cannot meet the requirement that the dictionary should be overcomplete to ensure a good SR result. The above limitation can be alleviated as follows. First, the Pan image covers the same scene as the LR MS image, so that the Pan image will exhibit a very high similarity to the MS image in geometric structure. Moreover, the range of the wavelength spectrum of the Pan modality is usually overlapped or partly overlapped by each of the MS bands. As a result, the Pan image is highly correlated with each of the MS bands, and a dictionary atom generated from the Pan image will be highly relevant to the corresponding MS image patch. Based on SR theory, a sparse solution can be obtained when the atom in the dictionary is highly relevant to the signal to be represented.

In the image fusion application, the LR MS image \mathbf{X}^l can be modeled as the degraded version of the unknown HR MS image \mathbf{X}^h , which can be formulated as $x^l = \mathbf{M}\mathbf{x}^h + \mathbf{n}_1$, where \mathbf{M} represents the degradation process and \mathbf{n}_1 is the noise. As we stated before, the HR image patches \mathbf{x}^h can be sparsely represented by a suitable overcomplete HR dictionary \mathbf{D}^h

$$\mathbf{x}^h = \mathbf{D}^h \boldsymbol{\alpha} + \mathbf{n}_2 \quad \text{for some } \boldsymbol{\alpha} \in \mathbb{R}^K \text{ with } \|\boldsymbol{\alpha}\|_0 \ll K. \quad (3)$$

Therefore, the corresponding LR image patch can be expressed as

$$\mathbf{x}^l = \mathbf{M}\mathbf{D}^h \boldsymbol{\alpha} + \mathbf{n}_3. \quad (4)$$

As the atoms in the two coupled dictionaries are in one-to-one correspondence, and the LR atoms from the LR dictionary are subject to the same degradation process as the LR MS image patches, from the HR style to LR style, we can obtain the following relationship:

$$\mathbf{D}^l = \mathbf{M}\mathbf{D}^h. \quad (5)$$

Taking (4) and (5) into consideration, we obtain $\mathbf{x}^l = \mathbf{D}^l \boldsymbol{\alpha} + \mathbf{n}_3$. Therefore, based on (3), it is reasonable, to a certain extent, to assume that the sparse coefficient vectors between the low- and high-resolution patch pair, with respect to their dictionaries, are the same in the image fusion application.

In the SparseFI method, the coupled dictionaries are the same for each LR MS band pan-sharpening process. In the sparse coefficient estimation stage, each one of the LR MS image patches is sparsely represented by the LR dictionary \mathbf{D}^l , and the HR MS image patch can be reconstructed with the same sparse coefficient vector and the HR dictionary \mathbf{D}^h . After all the HR MS image patches are reconstructed, the HR MS image is obtained by averaging the overlapping image patches.

III. THE OCDL APPROACH FOR IMAGE FUSION

In this section, we propose an OCDL approach for image fusion. The concept diagram of the proposed method is set out in

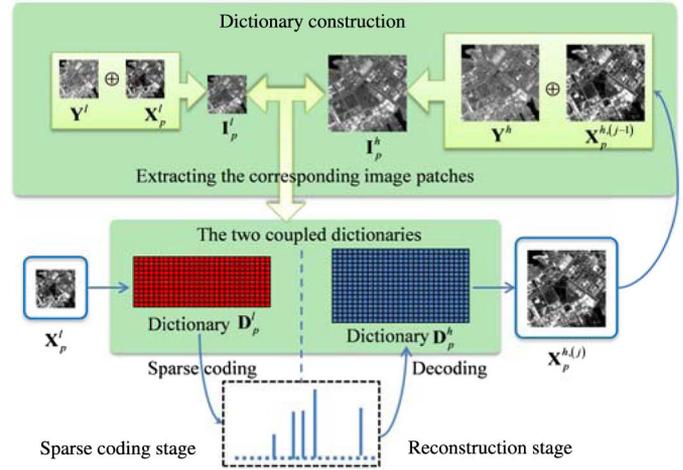


Fig. 1. Proposed OCDL method for image fusion.

Fig. 1. Taking the p th LR MS band as an example, we introduce the proposed method as follows. As shown in Fig. 1, we construct an intermediate image pair, which is composed of an HR intermediate image and its LR version. The former HR image is constructed by the superposition of the HR Pan image and the HR MS band obtained from the last iteration. Meanwhile, the corresponding LR intermediate image construction is offline, which directly imposes the LR MS band on the down-sampled Pan image. Iteration processing is utilized in the proposed approach, and it is believed that the MS information can be fully enhanced in the coupled dictionary reconstruction process, which makes the dictionary more specific and suitable for the LR MS image sharpening task.

Let \mathbf{X}_p^l and $\mathbf{X}_p^{h(l)}$ ($p = 1, \dots, N$) denote the p th band of the LR MS image and the HR fused MS image, respectively, where N stands for the number of bands in the MS image. The Pan image with a resolution R times higher than the LR MS image is represented as \mathbf{Y}^h . For most very high-resolution satellite sensor systems, e.g., IKONOS and QuickBird, the value of the scale ratio R is 4. The proposed OCDL method for image fusion consists of two steps: 1) OCDL and 2) image fusion with the online coupled dictionary.

A. Online Coupled Dictionary Learning

The construction of the coupled dictionaries in the coupled dictionary model is of vital importance. Although the coupled dictionaries generated directly from the Pan image and its down-sampled version can be used for image fusion, the available MS image information is ignored in the dictionary construction stage. In view of this, we have designed a superposition strategy to construct the coupled dictionaries. For the p th MS band, the LR intermediate image is obtained by

$$\mathbf{I}_p^l = \mathbf{Y}^l \oplus \mathbf{X}_p^l \quad (6)$$

where \mathbf{Y}^l is the LR Pan image, which is obtained by down-sampling the HR Pan image \mathbf{Y}^h to the same resolution as the LR MS image \mathbf{X}^l . \oplus represents the superposition operation of the two images. Suppose that there is an HR MS image \mathbf{X}^h , then,

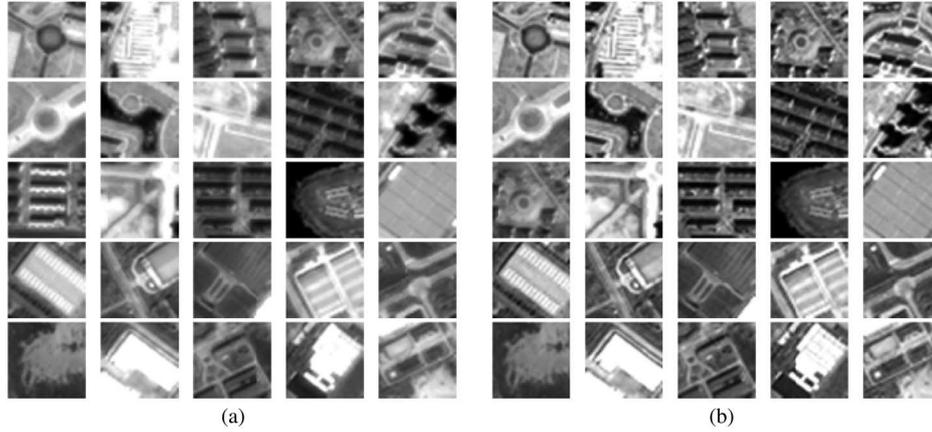


Fig. 2. Some atoms of (a) $\mathbf{D}^{h,(initial)}$ and (b) $\mathbf{D}^{h,(final)}$, which correspond to the HR dictionary in the first and last iterations, respectively. In this example, the patches are of size 36×36 . (a) Initial iteration. (b) Final iteration.

similarly, the HR intermediate image for the p th band can be obtained by

$$\mathbf{I}_p^h = \mathbf{Y}^h \oplus \mathbf{X}_p^h. \quad (7)$$

After we have obtained the two intermediate images, the coupled dictionaries for the p th band are constructed by sampling raw patches from the LR and HR intermediate images, respectively. In order to match the corresponding relationship of the atoms in the coupled dictionaries, the proportional relation of the patch size and the step size between the HR intermediate image and the LR intermediate image should be consistent with the scale ratio R .

Unfortunately, the HR MS image \mathbf{X}^h is unavailable. To deal with this problem, the initial estimation of $\mathbf{X}^{h,(0)}$ is obtained by the interpolation of the LR MS image with bicubic interpolation. However, the integration of the LR MS image destroys the mapping function between the coupled dictionaries, which should be constructed from the corresponding HR/LR image pair. In order to achieve the goal that the SR of the LR MS image patch, in terms of the LR dictionary, can effectively reconstruct its underlying HR MS image patch, with respect to the HR dictionary, an iterative update approach is utilized, where the fusion result from the last iteration is used to construct \mathbf{D}^h for the next iteration. Fig. 2 provides an example of some atoms of the HR dictionary \mathbf{D}^h of the initial and final iterations. It can be seen that the image patch contrast is greatly enhanced, and the detail information is increased in the final HR dictionary, compared to the initial HR dictionary.

The success of the proposed OCDL method is mainly due to the fact that the superposition strategy increases the correlation between the LR MS image and the image used for dictionary construction, when compared with the processing strategy of the SparseFI [23] method. That is to say, the dictionary atoms generated from the intermediate image are more relevant to the corresponding MS image patches than those generated from the original Pan image directly. According to SR theory, a better SR result tends to be obtained when the dictionary atoms are more relevant [25] to the signal to be represented. In view of this, the HR MS image patches reconstructed by multiplying the HR dictionary with the associated sparse coefficient vectors benefit from the excellent LR SR performance. Due to the integration of

the MS image in the dictionary construction, the proposed method relieves the fusion process from the bottleneck of the Pan image and leads to wider application possibilities.

B. Image Fusion With the Online Coupled Dictionary

In this stage, we attempt to infer the HR patch with each LR patch from the input. Let $\mathbf{x}_{p,i}^l = \mathbf{R}_i \mathbf{X}_p^l$ denote the i th patch of \mathbf{X}_p^l , where \mathbf{R}_i is a matrix that extracts the i th block from the image and $\mathbf{x}_{p,i}^h$ represents the reconstructed HR image patch. For each patch $\mathbf{x}_{p,i}^h$, the sparse coefficient vector $\alpha_{p,i}$, with respect to the LR dictionary \mathbf{D}_p^l , is estimated as follows:

$$\min \|\alpha_{p,i}\|_1 \quad s.t. \quad \|\mathbf{D}_p^l \alpha_{p,i} - \mathbf{x}_{p,i}^l\| \leq \varepsilon \quad (8)$$

where the parameter ε restricts the reconstruction error.

In the proposed image fusion method, we also assume that the sparse coefficient vectors between the low- and high-resolution vectors, with respect to their dictionaries, are the same. The HR image patches $\mathbf{x}_{p,i}^h$ can be reconstructed by

$$\mathbf{x}_{p,i}^h = \mathbf{D}_p^h \alpha_{p,i}. \quad (9)$$

After all the optimal sparse coefficients of the patches $\mathbf{x}_{p,i}^l$ are obtained, the p th band of the HR MS image \mathbf{X}_p^h can be reconstructed by the following formulation:

$$\mathbf{X}_p^h = \left(\sum_i \mathbf{R}_i^T \mathbf{R}_i \right)^{-1} \left(\sum_i \mathbf{R}_i^T \mathbf{D}_p^h \alpha_{p,i} \right). \quad (10)$$

In the proposed method, a threshold on the residual error between the estimated fusion images in two consecutive iterations is utilized as the stopping criterion. The proposed OCDL approach is summarized in Algorithm 1. With the following processing scheme, all the LR MS bands can be sharpened.

IV. EXPERIMENTAL RESULTS AND ANALYSES

To demonstrate the effectiveness of the proposed OCDL method, three simulated groups of experiments were carried out with two kinds of satellite sensor data, i.e., IKONOS and WorldView-2. The spectral ranges of the IKONOS and WorldView-2 data are shown in Fig. 3.

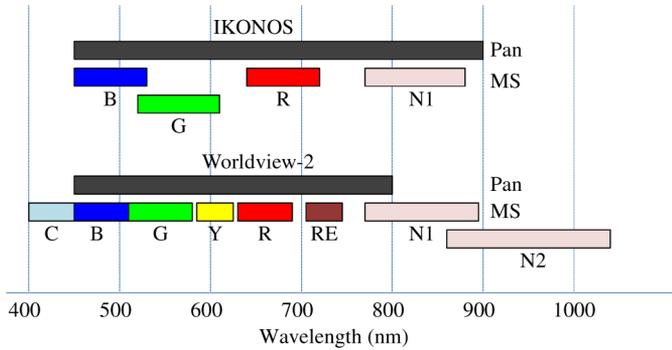


Fig. 3. Spectral ranges of the IKONOS and WorldView-2 data.

(1) The IKONOS data set is from Hubei province, China, and was obtained by IKONOS-2 in 2009. It is composed of four MS bands spanning the visible and near-infrared (NIR) wavelengths: blue (450–530 nm), green (520–610 nm), red (640–720 nm), and NIR (770–880 nm), and a corresponding Pan band with a broader range (450–900 nm). All the data had been radiometrically calibrated from digital counts, orthorectified, i.e., resampled to uniform ground resolutions of 4 and 1 m for MS and Pan, respectively, and packed in 11-bit bytes. Two IKONOS scenes with significantly different land-cover types (i.e., one mainly containing buildings and the other covering a bare area) were utilized for the two experiments.

(2) The WorldView-2 data set is from an urban area of Rio de Janeiro, Brazil, and was provided by DigitalGlobe for the IEEE-IGARSS 2011 Data Fusion Contest. The image consists of one Pan and eight MS bands with 0.5 m and 2 m spatial resolutions, respectively. The spectral ranges of the MS bands span from the visible to the NIR, and include coastal (400–450 nm), blue (450–510 nm), green (510–580 nm), yellow (585–625 nm), red (630–690 nm), red edge (705–745 nm), near-infrared 1 (770–895 nm), and near-infrared 2 (860–1040 nm). The spectral range of the Pan image covers the interval of 450–800 nm.

Algorithm 1 Online coupled dictionary learning (OCDL) approach for image fusion, taking the p th band as an example

Input: the Pan image \mathbf{Y}^h and the p th LR MS band \mathbf{X}_p^l

Initialization: Initialize $j = 1$, and

1) Down-sample the panchromatic image \mathbf{Y}^h to the resolution of the multispectral image and acquire the LR intermediate image. $\mathbf{I}_p^l = \mathbf{Y}^h \downarrow \oplus \mathbf{X}_p^l$

2) Interpolate the multispectral image \mathbf{X}_p^l to the resolution of the panchromatic image and acquire the initial HR intermediate image. $\mathbf{I}_p^{h,(0)} = \mathbf{Y}^h \oplus \mathbf{X}_p^l \uparrow$

3) Set the threshold on the residual error between the estimated fusion images in two consecutive iterations σ .

While $\|\mathbf{X}_p^{h,(j)} - \mathbf{X}_p^{h,(j-1)}\|_F / \|\mathbf{X}_p^{h,(j-1)}\|_F > \sigma$

1) Dictionary construction: build the LR dictionary \mathbf{D}_p^l from \mathbf{I}_p^l , and the HR dictionary $\mathbf{D}_p^{h,(j)}$ from $\mathbf{I}_p^{h,(j)}$, respectively.

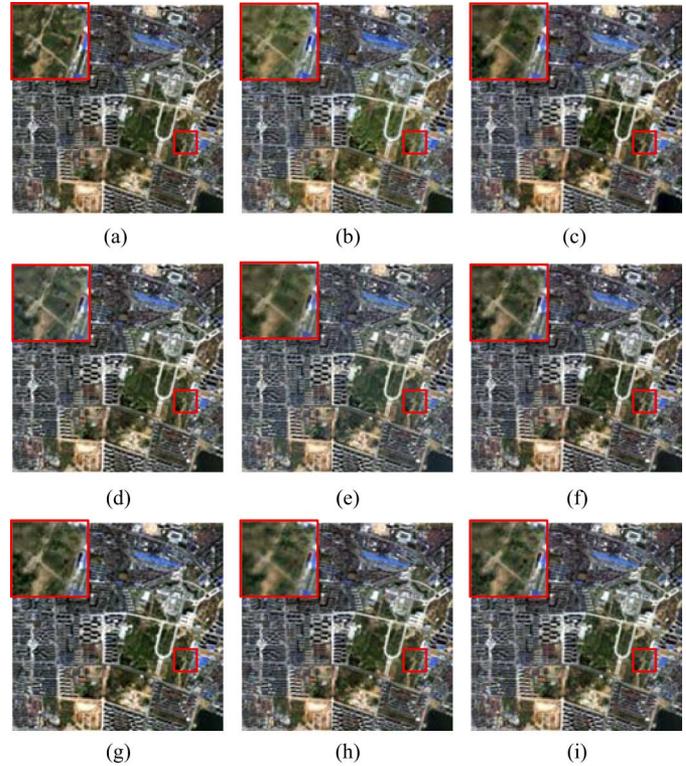


Fig. 4. Simulated image experiments: (a) original HR MS image (RGB, 604 × 604, 4 m); (b) GS; (c) AIHS; (d) AWLP; (e) CBD; (f) CSIF; (g) IFLD; (h) SparseFI; and (i) OCDL.

TABLE I
TIME COSTS FOR DIFFERENT IMAGE FUSION METHODS

Method	GS	AIHS	AWLP	CBD	CSIF	IFLD	SparseFI	OCDL
Time (s)	2	0.9	0.3	76	15042	4942	147	191

2) Sparse coding: compute the representation vector $\alpha_{p,i}$ via (8) for each patch $\mathbf{R}_i \mathbf{X}_p^l$ by approximating the solution of (8).

3) Reconstruct the p th band of the HR multispectral image $\mathbf{X}_p^{h,(j)}$ by (10).

4) Update the HR intermediate image $\mathbf{I}_p^{h,(j)} : \mathbf{I}_p^{h,(j)} = \mathbf{Y}^h + \mathbf{X}_p^{h,(j)}$.

End While

Output: The p th HR MS band \mathbf{X}_p^h .

In this study, three independent simulated image experiments were carried out, in which the original Pan image and MS image were spatially down-sampled by using bicubic interpolation with a factor of R , and image fusion was then conducted with the down-sampled Pan image and the MS image. According to the Wald's protocol that any synthetic image should be as identical as possible to the image that the corresponding sensor would observe with the highest spatial resolution [33], we compared the fused HR MS images with the original HR MS image, which is referred to as the reference. In the experiments, five typical evaluation metrics were adopted. The correlation coefficient (CC) [34]

TABLE II
QUANTITATIVE ASSESSMENT RESULTS OF THE SIMULATED EXPERIMENT SHOWN IN FIG. 4

	EXP	GS	AIHS	AWLP	CBD	CSIF	IFLD	SparseFI	OCDL	
CC	B	0.8296	0.9083	0.9314	0.9348	0.8999	<u>0.9384</u>	0.9351	0.9333	0.9398
	G	0.8234	0.9354	0.9530	0.9513	0.9488	<u>0.9573</u>	0.9549	0.9557	0.9590
	R	0.8313	0.9287	0.9465	0.9491	0.9469	<u>0.9536</u>	0.9524	0.9522	0.9555
	NIR	0.8426	0.9099	0.9331	0.9420	0.9388	<u>0.9471</u>	0.9460	0.9435	0.9478
	Avg	0.8317	0.9206	0.9410	0.9443	0.9336	<u>0.9491</u>	0.9471	0.9462	0.9505
RMSE	B	48.2238	40.0396	33.6096	30.5212	44.8128	34.0957	<u>30.3840</u>	30.8761	29.3075
	G	78.8129	56.3241	49.6935	42.5657	46.3743	52.7009	40.9753	<u>40.7572</u>	39.0502
	R	92.6944	68.4169	62.1554	52.1776	54.0510	60.6583	<u>50.6376</u>	50.7169	48.8710
	NIR	107.2043	82.5989	77.1484	67.9663	68.3605	89.8303	<u>64.2201</u>	65.7750	63.2585
	Avg	81.7338	61.8449	55.6517	48.3077	53.3997	59.3213	<u>46.5542</u>	47.0313	45.1218
SAM	4.4282	4.6448	4.1196	3.6759	4.1472	4.2021	3.7715	3.7692	<u>3.7225</u>	
ERGAS	4.2572	3.2067	2.9043	2.5090	2.7257	3.0903	<u>2.4146</u>	2.4397	2.3433	
Q4	0.6722	0.8415	0.8610	0.9134	0.9017	0.9114	<u>0.9145</u>	0.9066	0.9214	

B, G, R, and NIR represent the results of the blue, green, red, and NIR bands, respectively, and Avg is the average result of the results of R, G, B, and NIR.

TABLE III
QUANTITATIVE ASSESSMENT RESULTS OF THE PROPOSED METHOD WITH DIFFERENT PATCH SIZES, USING THE SAME DATA AS FIG. 4. ϵ IS SET TO 100

Patch size	5×5	7×7	9×9	11×11	13×13
CC _{avg}	0.9266	0.9476	0.9505	0.9491	0.9492
RMSE _{avg}	54.0923	46.2102	45.1218	45.9508	46.0217
SAM	4.0132	3.7853	3.7225	3.7972	3.7923
ERGAS	2.7969	2.3956	2.3433	2.3887	2.3932
Q4	0.8702	0.9150	0.9214	0.9208	0.9209

and root-mean-square error (RMSE) were calculated for each band between the fused MS images and the reference original MS image. Erreur relative globale adimensionnelle de synthèse (ERGAS) [33] and Q4 [35], which are two comprehensive evaluation indexes, provide unique measures of the fusion performance for all the MS bands. Furthermore, the spectral angle mapper (SAM) [34], [35] index was also considered to measure the spectral distortion. Smaller values of RMSE, SAM, and ERGAS tend to be achieved by a better fusion result, as do larger CC and Q4 values. It should be noted that the computation of the Q4 index was performed on nonoverlapping 32×32 blocks, as suggested in [36].

The proposed OCDL method was compared with four commonly used image fusion methods, namely GS [13], adaptive IHS (AIHS) [11], the AWLP method [17], the CBD method [18], and three state-of-the-art sparsity regularization-based methods, CS-based image fusion (CSIF) [19], image fusion via SRs over learned dictionaries (IFLD) [22], and SparseFI [23].

A. Fusion Results With the IKONOS Data

The original IKONOS Pan image and the MS image in the two data sets were spatially down-sampled by using bicubic interpolation with a factor of 4 to yield a 4 m spatial resolution Pan image and a 16 m spatial resolution MS image, respectively. Image fusion was then conducted with the down-sampled Pan image and the MS image. In this study, we used a simulated IKONOS LR MS image with the size of 151×151 , and the corresponding Pan image sized 604×604 . Two free parameters, the patch size and the overlap area size of the SparseFI method and the proposed OCDL method, were utilized for the patch-based

processing. The optimal size of the image patch in the LR MS image was 9×9 , with an overlapping area size of 9×4 pixels in the SparseFI and the proposed method. In the dictionary construction stage, the same patch size and overlapping area size were adopted as in the SR stage, and the dictionary size was 81×1444 , with the LR MS image sized 151×151 . In this paper, LASSO was adopted to pursue the sparse coefficient vector, and the parameter ϵ in (5) was set to 100, where most of the indexes reached their best values. The threshold σ was empirically set to 10^{-4} . The impacts of the patch size and parameter ϵ in the proposed method are discussed later on in this section.

The GS algorithm was implemented in ENVI 4.7 software in mode 1. Two levels of decomposition were utilized for the *à trous* wavelet transform used for the AWLP and CBD methods. As for the CSIF method, a dictionary with the size of $256 \times 10\,000$ was randomly sampled from the 10 HR MS images that have the same spatial resolution as the Pan image. In the IFLD method, the two coupled dictionaries, D^{pan} and D_i^{MS} , were trained to 512 atoms from 50 000 image patch pairs randomly sampled from the LR MS and Pan images.

Fig. 4 shows the HR MS fusion results of the different methods with the first IKONOS data set. By visually comparing the fusion images with the original MS image, it can be seen that all the methods can effectively pan-sharpen the LR MS image. We next discuss the fusion result in both the spectral and spatial views in detail. Among the different fusion results, the most serious spectral distortion exists in the fusion result of the GS method, due to the modification of the low frequencies of the original MS image. The results of the CSIF method and the IFLD method exhibit rich, detailed spatial information; however, there are artifacts in the fusion results, and the reason for this can mainly be attributed to the fact that the Pan image imposes a strong constraint over the reconstruction of the HR MS image. The result of the CSIF method also shows spectral distortion to a certain degree. The AIHS method adopting image-adaptive coefficients for IHS can obtain a more accurate spectral resolution. It can be clearly seen that the AWLP method and the proposed method have more advantages than the other methods in maintaining the spectral information of the original MS image. It can also be seen that the fused image with the SparseFI method

TABLE IV
QUANTITATIVE ASSESSMENT RESULTS OF THE PROPOSED METHOD WITH $\varepsilon = 10^K$ (K IS FROM -4 TO 4), USING THE SAME DATA AS FIG. 4

K	-4	-3	-2	-1	0	1	2	3	4
CC _{avg}	0.9505	0.9505	0.9505	0.9505	0.9505	0.9506	0.9505	0.9492	0.9358
RMSE _{avg}	45.1738	45.1735	45.1725	45.1695	45.1486	45.1237	45.1218	45.6456	51.0215
SAM	3.7436	3.7435	3.7433	3.7426	3.7392	3.7332	3.7225	3.7456	4.1828
ERGAS	2.3463	2.3463	2.3462	2.3461	2.3448	2.3434	2.3433	2.3688	2.6337
Q4	0.9218	0.9218	0.9218	0.9218	0.9218	0.9218	0.9214	0.9183	0.8910

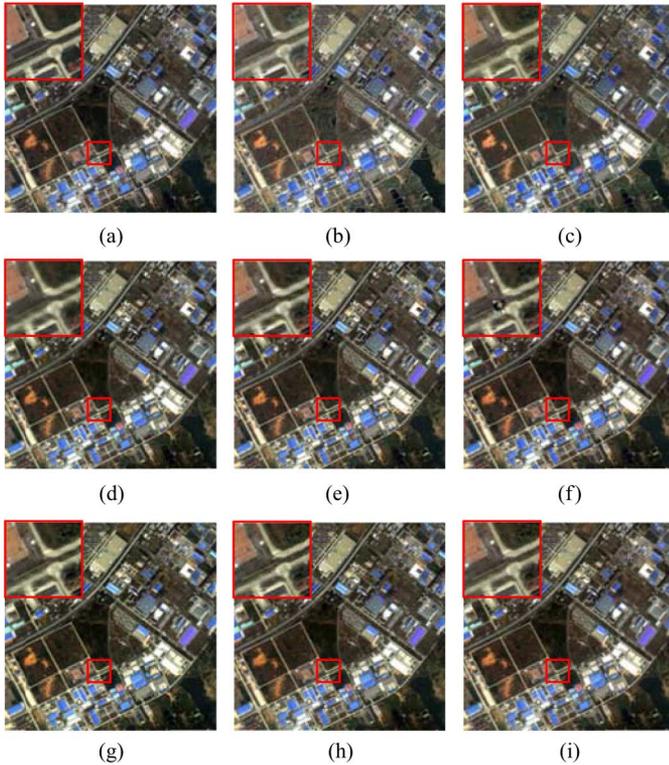


Fig. 5. Simulated image fusion experiments: (a) original HR MS image (RGB, 604×604 , 4 m); (b) GS; (c) AIHS; (d) AWLP; (e) CBD; (f) CSIF; (g) IFLD; (h) SparseFI; and (i) OCDL.

is blurred to some degree, as with the AWLP method. To facilitate a comparison, a detailed region is shown in the top-left corner of each image. Here, it can be observed that the fusion result of the proposed OCDL method is the closest to the original image. Since the sparsity regularization-based fusion methods are usually time-consuming due to the large computational complexity of the ℓ_1 -norm minimization problem, the running times of all the methods compared in the experimental part with the first IKONOS data set are provided in Table I. To sum up, the proposed OCDL method not only provides high-quality spatial details but also decreases the spectral distortion to a great extent, within an acceptable time frame.

The quantitative assessment indexes for the fusion results are calculated with the original HR MS image, as shown in Table II. The best results for each quality index are labeled in bold, and the second-best results for each quality index are underlined. The EXP entry represents the plain (bicubic) resampling of the MS data set at the scale of the Pan [18]. Here, it can be seen that the

TABLE V
TIME COSTS FOR DIFFERENT IMAGE FUSION METHODS USING THE SAME DATA AS FIG. 5

Method	GS	AIHS	AWLP	CBD	CSIF	IFLD	SparseFI	OCDL
Time (s)	2	0.8	0.3	81	14727	4667	148	170

proposed method acquires the best evaluation results for all the bands, in terms of both CC and RMSE, i.e., the fusion result of the proposed method is the most correlated with the original MS image and has the least radiometric distortion. The second-best SAM value is obtained by the proposed method, which implies that the proposed method can effectively preserve the spectral information in the fusion process. For the ERGAS index and Q4 index, the proposed method again demonstrates a better performance than the other fusion methods, as expected. Overall, the quantitative assessment results are consistent with the visual evaluation, and the proposed method achieves the best fusion result.

The impact of the patch size in the proposed method is investigated in Table III, in which the overlapping size is set to 4. It can be observed that the optimal patch size with the IKONOS data set is 9×9 , from the perspective of all the evaluation indexes. With the consideration of the fact that the time cost increases with the increase in the patch size, we use 9×9 in the following experiments.

The impact of parameter ε in the proposed method is also investigated in Table IV. Here, it can be seen that all the indexes share a similar trend with the variation of ε , and the best fusion results tend to be achieved when parameter ε is less than or equal to 100. Overall, it is shown that the fusion results of the proposed method are quite robust with regard to parameter ε .

Fig. 5 shows the HR MS fusion results of the different methods with the second IKONOS data set in which the sizes of LR MS image and Pan image are the same as the IKONOS data set one, and it can be seen that the results are consistent with the first IKONOS data set. For example, the fused image of the GS method achieves a good visual impression, despite the loss of the spectral details in the bare area and some building roofs. It can also be observed that both AWLP and the proposed OCDL method perform well, with well-preserved spectral detail information. To facilitate a comparison, detailed regions are shown in the top-left corner of the images. Here, it can be observed that the fused images of the AWLP method and the SparseFI method are blurred to some degree. Therefore, it can be concluded that the proposed OCDL method performs the best with respect to both the spatial and spectral perspectives. The running times of all the

TABLE VI
QUANTITATIVE ASSESSMENT RESULTS OF THE SIMULATED EXPERIMENT SHOWN IN FIG. 5

	EXP	GS	AIHS	AWLP	CBD	CSIF	IFLD	SparseFI	OCDL
CC	B	0.9307	0.9363	0.9709	<u>0.9720</u>	0.9593	0.9709	0.9657	0.9740
	G	0.9113	0.9511	0.9750	0.9741	0.9750	0.9759	0.9705	<u>0.9762</u>
	R	0.9024	0.9427	0.9670	0.9681	0.9666	<u>0.9687</u>	0.9633	<u>0.9687</u>
	NIR	0.8830	0.9168	0.9461	0.9499	0.9463	<u>0.9548</u>	0.9463	0.9501
	Avg	0.9068	0.9367	0.9647	0.9660	0.9618	<u>0.9676</u>	0.9614	0.9663
RMSE	B	35.8586	35.4085	24.3238	<u>23.0096</u>	29.7417	30.3732	25.4115	23.6811
	G	56.6374	44.4747	33.0991	31.0640	30.8767	49.4524	33.3071	<u>29.8188</u>
	R	68.8379	55.1147	43.4900	39.9565	40.8373	55.7852	43.1808	<u>39.6328</u>
	NIR	75.6418	64.1970	53.3940	53.0273	51.9716	86.4252	52.1018	<u>50.2630</u>
	Avg	59.2439	49.7987	38.5767	36.7644	38.3568	55.5090	38.5003	<u>35.8489</u>
SAM	2.7691	3.4334	2.6800	<u>2.5016</u>	2.7334	3.2232	2.6723	2.5236	2.4524
ERGAS	2.7663	2.2996	1.7961	1.7054	1.7663	2.5676	1.7874	<u>1.6640</u>	1.5274
Q4	0.7470	0.8420	0.8880	<u>0.9140</u>	0.9022	0.9090	0.9061	0.9004	0.9227

TABLE VII

TIME COSTS FOR DIFFERENT IMAGE FUSION METHODS USING THE SAME DATA AS FIG. 6

Method	GS	AIHS	AWLP	CBD	SparseFI	OCDL
Time (s)	3	1.3	0.4	75	279	321

methods with the second IKONOS data set share a similar trend with the first IKONOS data set as shown in Table V due to the same experimental data scale.

Table VI reports the quantitative scores for all the methods. As with the first data set, the observations for all the indexes show similar trends. The proposed method achieves the best SAM index value, which demonstrates the superiority of the proposed method in preventing spectral distortion. For the Q4 index, the proposed method is better than the other methods, followed by the AWLP method and the IFLD method. On the whole, the proposed method performs better than the other fusion methods with this IKONOS data set.

B. Image Fusion Results With the WorldView-2 Data

In this section, we describe and analyze the experimental results with the WorldView-2 satellite data, which provides one Pan image and eight MS bands. Unlike the IKONOS data, the spectral range of the Pan image does not cover the whole interval of all the MS bands. It can be observed from Fig. 3 that the spectral range of the Pan image is mainly concentrated in the visible light region, which leads to the low correlations between the Pan image and those bands with a spectral range beyond the visible light region. It is demonstrated in the following paragraphs that this characteristic of the WorldView-2 data leads to results that are sensitive to the different pan-sharpening methods. The size of the LR MS image in the simulated WorldView-2 data set experiments was also 151×151 , with the corresponding Pan image sized 604×604 .

With the aim of reaching the best evaluation values, parameter ε in the proposed OCDL method and the SparseFI method was both set to 100, respectively. The other parameter settings were set to the same as those for the IKONOS data sets. It should be noted that we omit the CSIF method and the IFLD method with the WorldView-2 data, in which the fitting coefficients in the

linear fitting model between the Pan image and the MS bands are unavailable.

Fig. 6 shows scenes with the 5-3-2 bands as a red-green-blue composite for the HR MS fusion results of the different fusion methods with the WorldView-2 data set. Unlike the IKONOS data, whose Pan image covers the whole interval of all the MS bands, it can be observed that the GS method introduces more spectral distortions in this experiment, since the setting that averages all the MS bands to the LR Pan image will change the low-frequency components of the original MS image for the WorldView-2 data set. In addition, it can be seen in Fig. 6 that the AWLP method and the proposed OCDL method have more advantages than the other methods in maintaining the spectral information of the original MS image with these three specific MS bands, as does the SparseFI method. The running times of all the methods with the WorldView-2 data set are provided in Table VII as well. All the methods cost more time than the IKONOS datasets because of the larger number of spectral channels, especially the sparsity regularization based fusion methods.

The visual interpretation can only show the quality of three bands of the fusion results. We evaluate the fusion results for all the MS bands via a comprehensive quantitative evaluation, as shown in Table VIII. Although all the methods show excellent preservation of the CC values of the original MS bands, the proposed OCDL method maintains the highest correlation between the fusion result and the original MS image. In terms of the SAM index, AWLP and the proposed OCDL method present results are remarkably better than the other methods, from the aspect of spectral information protection. Moreover, the best performance is also provided by the proposed OCDL method with the ERGAS index. It is worth mentioning here that almost all the methods are affected by the characteristics of the WorldView-2 data, in that the Pan image has low correlations with the last two NIR image bands. The proposed OCDL method shows significant advantages over the SparseFI method in the last two NIR MS bands; that is to say that the proposed method can ensure relatively stable fusion results due to the integration of the MS image information in the dictionary construction stage. Overall, the proposed method is suitable for the pan-sharpening of WorldView-2 data and can achieve a good fusion result.

TABLE VIII
QUANTITATIVE ASSESSMENT RESULTS OF THE SIMULATED EXPERIMENT SHOWN IN FIG. 6

	EXP	GS	AIHS	AWLP	CBD	SparseFI	OCDL	
CC	C	0.9245	0.9707	0.9679	0.9822	0.9774	<u>0.9833</u>	0.9843
	B	0.9222	0.9743	0.9767	0.9835	0.9822	<u>0.9860</u>	0.9868
	G	0.9166	0.9779	0.9691	0.9830	0.9735	<u>0.9875</u>	0.9877
	Y	0.9157	0.9780	0.9727	0.9828	0.9765	<u>0.9867</u>	0.9874
	R	0.9225	0.9753	0.9660	0.9835	0.9819	<u>0.9864</u>	0.9869
	R-E	0.9050	0.9720	0.9667	0.9763	0.9735	<u>0.9797</u>	0.9828
	NIR-1	0.9155	0.9344	0.9529	0.9560	<u>0.9575</u>	0.9566	0.9616
	NIR-2	0.9137	0.9287	0.9534	0.9526	<u>0.9540</u>	0.9526	0.9556
	Avg.	0.9170	0.9639	0.9657	0.9750	0.9720	<u>0.9774</u>	0.9791
RMSE	C	50.0410	31.9977	34.0902	26.4789	30.7957	<u>23.7793</u>	23.0557
	B	68.6409	40.8283	41.7271	34.9388	33.8731	<u>29.5146</u>	28.6494
	G	110.9691	60.0991	74.0010	57.3425	63.3577	<u>43.5700</u>	43.1575
	Y	99.1001	53.3447	62.8429	49.9781	53.0175	<u>39.8717</u>	38.8807
	R	56.8345	33.2639	38.7123	29.1731	30.0582	24.1227	23.6743
	R-E	88.5645	54.0217	54.6167	50.6353	47.3831	<u>41.6019</u>	38.3218
	NIR-1	70.9260	65.0293	53.5045	54.0519	<u>50.9161</u>	51.0788	48.3438
	NIR-2	96.4733	90.7999	71.3332	75.1433	<u>70.7912</u>	71.7942	70.1613
	Avg.	80.1937	53.6731	53.8535	47.2177	47.5241	<u>40.6666</u>	39.2806
SAM	3.9736	5.0861	4.1707	3.5146	4.0528	3.7890	<u>3.7272</u>	
ERGAS	5.3505	3.7039	3.6289	3.2200	3.1900	<u>2.8301</u>	2.7289	

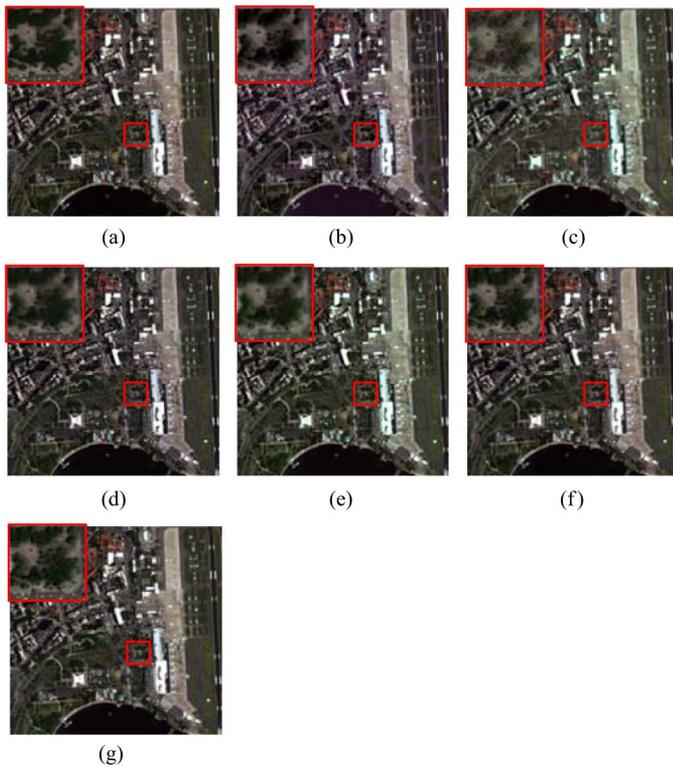


Fig. 6. Simulated image experiments: (a) original HR MS image (RGB, 604×604 , 4 m); (b) GS; (c) AIHS; (d) AWLP; (e) CBD; (f) SparseFI; and (g) OC DL.

C. Correlation Analysis

In this sub-section, we select the first IKONOS data set and the WorldView-2 data set to further analyze the effectiveness of the proposed method. In Figs. 7 and 8, the horizontal axis indicates the band name, and the vertical axis shows the corresponding CC between each LR/HR MS band and the LR/HR image for dictionary construction, respectively. In fact, the image utilized for the dictionary refers to the intermediate image in the proposed method and the Pan image in the SparseFI method. It is worth

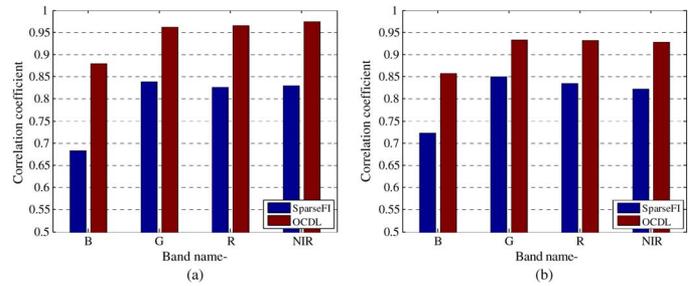


Fig. 7. Correlations between the MS bands and the image used for the dictionary construction, with the IKONOS data. (a) Low resolution. (b) High resolution.

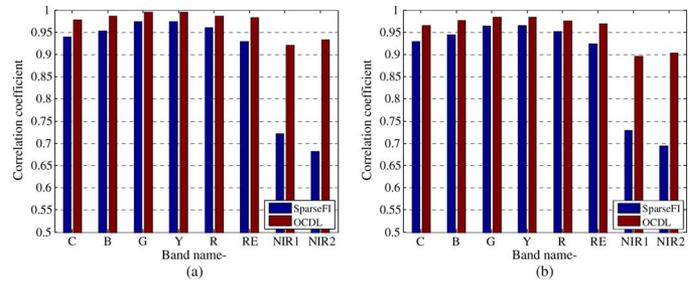


Fig. 8. Correlations between the MS bands and the image used for the dictionary construction, with the WorldView-2 data. (a) Low resolution. (b) High resolution.

mentioning that the original HR MS image is available for calculating the correlation between the original HR MS image and the HR image used for HR dictionary construction in the simulated image experiments.

From Fig. 7 (a), it can be observed that the LR intermediate image has a higher correlation with the LR MS image for each band than the down-sampled Pan image, which corresponds to the proposed method and the SparseFI method, respectively. Therefore, the atoms of the dictionary generated from the intermediate image are more relevant to the MS image patches than those generated from the Pan image, which leads to a better SR result. Likewise, the HR intermediate image has a higher

correlation with the original HR MS image for each band than the Pan image, as expected. Based on the fusion results in Section IV-A, we conclude that each fused MS band benefits from the novel coupled dictionary learning approach.

Fig. 8 depicts the correlations between each LR/HR WorldView-2 image band and the LR/HR image used for the LR/HR dictionary construction. For the WorldView-2 data, the LR intermediate image has a higher correlation with the LR MS image for each band than the down-sampled Pan image, which is consistent with the IKONOS data. Moreover, it is also noticeable that the Pan image is highly correlated with the first six MS bands, whereas it has relatively low correlations with the last two NIR bands. The proposed method, however, can maintain a high correlation between the MS image and the intermediate image used for dictionary construction, even in the circumstance of the Pan image being poorly correlated with the MS image. This is the reason why the last two NIR fusion bands improve more than the other six bands, as shown in the experiment. In conclusion, the proposed OCDL method shows more advantages than the other methods in sharpening WorldView-2 data, especially for the last two NIR image bands.

V. CONCLUSION

In this paper, an OCDL approach is proposed for the pansharpening of a LR MS image. The online coupled dictionaries are iteratively updated with the fused HR MS image, in which the atoms of the constructed dictionary are more relevant to the MS image patches, and lead to a better fusion result. In the proposed coupled dictionary learning stage, the available Pan image and the LR MS image are fully utilized to decrease the spectral distortion and to enhance the spatial information. The proposed method was compared with other state-of-the-art fusion methods, using IKONOS image data and WorldView-2 data. Overall, the proposed OCDL method showed satisfactory fusion results, even if the Pan image had a low correlation with some of the MS bands. The experimental results confirm the effectiveness of the proposed method, from both the spatial and spectral perspectives.

ACKNOWLEDGMENT

The authors would like to thank the editor and the anonymous reviewers, whose comments greatly helped in improving the presentation and quality of this paper. The authors would also like to thank DigitalGlobe for providing the WorldView-2 scenes from Rio de Janeiro for the IEEE-IGARSS 2011 Data Fusion Contest, and Prof. S. Li and Dr. H. Yin for providing the code of the IFLD method.

REFERENCES

- [1] L. Wald, "Some terms of reference in data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1190–1193, May 1999.
- [2] A. Filippidis, L. C. Jain, and N. Martin, "Multisensor data fusion for surface land-mine detection," *IEEE Trans. Syst. Man Cybern. C, Appl. Rev.*, vol. 30, no. 1, pp. 145–150, Feb. 2000.
- [3] J. Li, H. Zhang, and L. Zhang, "Supervised segmentation of very high resolution images by the use of extended morphological attribute profiles and a sparse transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 8, pp. 1409–1413, Aug. 2014.
- [4] J. Li, H. Zhang, Y. Huang, and L. Zhang, "Hyperspectral image classification by nonlocal joint collaborative representation with a locally adaptive dictionary," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3707–3719, Aug. 2013, doi: 10.1109/TGRS.2013.2274875.
- [5] A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE pansharpening of very high resolution multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 228–236, Jan. 2008.
- [6] B. Aiuzzi, S. Baronti, F. Lotti, and M. Selva, "A comparison between global and context-adaptive pansharpening of multispectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 302–306, Apr. 2009.
- [7] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE Geosci. Remote Sens.*, vol. 49, no. 1, pp. 295–309, Jan. 2011.
- [8] L. Wald, *Data Fusion: Definitions and Architectures: Fusion of Images of Different Spatial Resolutions*. Paris, France: Les Presses de l'École des Mines, 2002, p. 197.
- [9] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, "Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1301–1312, May 2008.
- [10] T.-M. Tu, P. S. Huang, C.-L. Hung, and C.-P. Chang, "A fast intensity hue-saturation fusion technique with spectral adjustment for IKONOS imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 309–312, Oct. 2004.
- [11] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An adaptive IHS pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 746–750, Oct. 2010.
- [12] P. S. Chavez, S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data Landsat TM and SPOT panchromatic," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 295–303, Mar. 1991.
- [13] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," Eastman Kodak Company, U.S. Patent 6011875, Jan. 4, 2000.
- [14] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly correlated images. I. Decorrelation and HSI contrast stretches," *Remote Sens. Environ.*, vol. 20, no. 3, pp. 209–235, Dec. 1986.
- [15] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rougé, "A variational model for P + XS image fusion," *Int. J. Comput. Vis.*, vol. 69, no. 1, pp. 43–58, Aug. 2006.
- [16] T. Ranchin, B. Aiuzzi, L. Alparone, S. Baronti, and L. Wald, "Image fusion—the ARSIS concept and some successful implementation schemes," *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 1/2, pp. 4–18, Jun. 2003.
- [17] X. Otazu, M. Gonzalez-Audicana, O. Fors, and J. Nunez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, Oct. 2005.
- [18] B. Aiuzzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, Oct. 2002.
- [19] S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 738–746, Feb. 2011.
- [20] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [21] C. Jiang, H. Zhang, H. Shen, and L. Zhang, "A practical compressed sensing-based pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 629–633, Jul. 2012.
- [22] S. Li, H. Yin, and L. Fang, "Remote sensing image fusion via sparse representations over learned dictionaries," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4779–4789, Sep. 2013.
- [23] X. X. Zhu and R. Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2827–2836, May 2013.
- [24] C. Jiang, H. Zhang, H. Shen, and L. Zhang, "Two-step sparse coding for the pan-sharpening of remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, doi: 10.1109/JSTARS.2013.2283236, to be published.
- [25] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer-Verlag, 2010.
- [26] H. Zhang, J. Li, Y. Huang, and L. Zhang, "A nonlocal weighted joint sparse representation classification method for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, doi: 10.1109/JSTARS.2013.2264720, to be published.

- [27] M. Elad, M. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proc. IEEE*, vol. 98, no. 6, pp. 972–982, Jun. 2010.
- [28] D. L. Donoho, "For most large underdetermined systems of linear equations, the minimal-norm near-solution approximates the sparsest near-solution," *Commun. Pure Appl. Math.*, vol. 59, no. 6, pp. 797–829, Jun. 2006.
- [29] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.
- [30] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc. B.*, vol. 58, no. 1, pp. 267–288, Mar. 1996.
- [31] S. Wang, L. Zhang, Y. Liang, and Q. Pan, "Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch image synthesis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 2216–2223.
- [32] J. Yang, Z. Wang, Z. Lin, S. Cohen, and T. Huang, "Coupled dictionary training for image super-resolution," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3467–3478, Aug. 2012.
- [33] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691–699, Jun. 1997.
- [34] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRSS data fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.
- [35] H. Zhang, H. Shen, and L. Zhang, "A super-resolution reconstruction algorithm for hyperspectral images," *Signal Process.*, vol. 92, no. 9, pp. 2082–2096, Sep. 2012.
- [36] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.



Min Guo received the B.S. degree from China University of Petroleum (East China), Qingdao, China, in 2012. She is currently working toward the Ph.D. degree with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China.

Her research interests include image fusion, sparse representation, pattern recognition, and computer vision in remote sensing images.



Hongyan Zhang (M'13) received the B.S. degree in geographic information system and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2005 and 2010, respectively.

Since 2010, he has been a Lecturer with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. His current research interests include image reconstruction, sparse representation, and low-rank methods in sensing image images.

Dr. Zhang is a Reviewer of several international academic journals, including IEEE TRANSACTION ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTION ON IMAGE PROCESSING, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and so on.



Jiayi Li (S'13) received the B.S. degree from Central South University, Changsha, China, in 2011. She is currently working toward the Ph.D. degree with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China.

Her research interests include hyperspectral imagery, sparse representation, pattern recognition, and computer vision in remote sensing images.



Liangpei Zhang (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, Changsha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is currently as the Head of the Remote Sensing Division, State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing,

Wuhan University. He is also a "Chang-Jiang Scholar" Chair Professor appointed by the Ministry of Education, China. He is now a Principal Scientist for the China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the remote sensing program in China. He has more than 260 research papers and is the holder of 5 patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang regularly serves as a Co-Chair of the series Society of Photo-Optical Instrumentation (SPIE) Engineers Conferences on Multispectral Image Processing and Pattern Recognition (MIPPR), Conference on Asia Remote Sensing, and many other conferences. He edits several conference proceedings, issues, and the Geoinformatics Symposiums. He also serves as an Associate Editor of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, *International Journal of Ambient Computing and Intelligence* (IJACI), *International Journal of Image and Graphics*, *International Journal of Digital Multimedia Broadcasting*, *Journal of Geo-spatial Information Science*, and the *Journal of Remote Sensing*. He is a Fellow of the Institution of Engineering and Technology (IET), Executive Member (Board of Governor) of the China National Committee of International Geosphere-Biosphere Programme, and Executive Member for the China Society of Image and Graphics.



Huanfeng Shen (M'11–SM'13) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

In July 2007, he joined the School of Resource and Environmental Science, Wuhan University, where he is currently a full professor. His research interests include image processing (for quality improvement), remote sensing application, and data fusion and assimilation. He has published more than 60 research papers.

Dr. Shen has been supported by several talent programs, including the New Century Excellent Talents by the Ministry of Education of China (2011) and the Hubei Science Fund for Distinguished Young Scholars (2011).