Cloud Removal Based on Sparse Representation via Multitemporal Dictionary Learning

Meng Xu, *Student Member, IEEE*, Xiuping Jia, *Senior Member, IEEE*, Mark Pickering, *Member, IEEE*, and Antonio J. Plaza, *Fellow, IEEE*

***Abstract*—Cloud covers, which generally appear in optical remote sensing images, limit the use of collected images in many applications. It is known that removing these cloud effects is a necessary preprocessing step in remote sensing image analysis. In general, auxiliary images need to be used as the reference images to determine the true ground cover underneath cloud-contaminated areas. In this paper, a new cloud removal approach, which is called multitemporal dictionary learning (MDL), is proposed. Dictionaries of the cloudy areas (target data) and the cloud-free areas (reference data) are learned separately in the spectral domain. The removal process is conducted by combining coefficients from the reference image and the dictionary learned from the target image. This method could well recover the data contaminated by thin and thick clouds or cloud shadows. Our experimental results show that the MDL method is effective in removing clouds from both quantitative and qualitative viewpoints.**

***Index Terms*—Cloud removal, dictionary learning, image reconstruction, multitemporal, sparse representation.**

I. Introduction

With the development of remote sensing technology, satellite images have become very useful in a variety of applications, including the following: Earth observation, climate change, and environmental monitoring. However, optical remote sensing images are often contaminated by clouds and cloud shadows since optical sensors acquire data at the visible and near-infrared wavelengths [1]. Cloud cover is considered to be a severe problem in optical images because it also leads to cloud shadow emerging. Both clouds and cloud shadows will degrade the utilization of image data and limit the use of these optical remote sensing images in further applications. For this reason, the process of removing clouds is necessary for improving the usefulness of optical remote sensing images.

Cloud effects vary due to their different compositions and heights. Opaque clouds obscure all the reflectance from the Earth's surface, allowing no ground cover signal to be collected by sensors. In this paper, we particularly focus on the case of thin clouds that do not entirely cover the signal corresponding to the underlying objects, as opposed to opaque clouds which entirely dominate the pixel spectral signature. The task of removing clouds is treated as an image restoration problem. A number of cloud removal methods have been developed to address this problem. The related approaches can be classified into two categories: individual-based and multitemporal-based methods.

In individual-based methods, cloud removal is implemented by making use of other bands from the individual image to model the cloud effects. An automatic thin cloud removal method utilizing the cirrus band as the auxiliary multispectral data was proposed in [2]. The relationships between the cirrus band in Landsat 8 Operational Land Imager (OLI) and other bands were used to model cloud effects. In [3], the haze optimized transformation method was proposed based on the high correlation between visible bands under clear atmospheric conditions and corrected for cloud effects by recovering how the linear relationship deviated as a result of clouds. The author in [4] used the near-infrared bands to estimate the spatial distribution of haze in visible bands by building a linear model over deep water regions. However, overcorrection will appear when clouds are not relatively thin. Generally, these individual-based cloud removal methods have difficulty mitigating the effects of thick clouds.

In contrast, multitemporal-based methods are capable of dealing with both thin and thick clouds. Since satellites revisit the same geographical location periodically, multitemporal cloud-free images can be acquired for the same location at different dates. A cloud removal method based on clone information was developed in [5]. Data from multitemporal images with no clouds were cloned to the cloudy patches based on the Poisson equation and a global optimization process. A cloud removal method by applying image fusion and a multiscale wavelet-based approach was proposed in [6], using target cloudy and multitemporal clear images to generate cloud-free images. Some methods designed for filling the gaps in Landsat ETM+Scan Line Corrector (SLC)-off imagery based on this idea have been also applied to the cloud removal task [7], [8]. For instance, a modified neighborhood similar pixel interpolator (MNSPI) cloud removal approach was proposed in [9] to predict the value of pixels blocked by thick clouds from its neighboring similar pixels with the help of auxiliary no-cloud images.

Cloud shadows are another problem which is even more difficult to cope with. Cloud shadows occur when the cloud

Manuscript received July 6, 2015; revised October 27, 2015; accepted December 10, 2015. Date of publication January 14, 2016; date of current version March 25, 2016.

M. Xu, X. Jia, and M. Pickering are with the School of Engineering and Information Technology, University of New South Wales, Canberra, A.C.T. 2600, Australia (e-mail: Meng.Xu@student.adfa.edu.au; X.JIA@adfa.edu.au; M.Pickering@adfa.edu.au).

A. J. Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, 10003 Cáceres, Spain (e-mail: aplaza@unex.es).

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/TGRS.2015.2509860

occludes the sunlight and prevents it from reaching the land surface [10]. Therefore, the areas contaminated by cloud shadows have normally lower reflectance than other regions. Lu [11] developed a cloud/shadow detection and substitution method based on maximum and minimum filters. This method needs thresholds determined by trial and error to extract clouds or shadows. Hence, the method exhibits limited accuracy when shadows cover high-reflectance surfaces or for thin clouds, which allow image regions to remain bright in these cases. In [12], a masking method was developed to detect cloud and cloud shadow in Landsat imagery over water and land, separately. Cloud shadow effects were detected by using a near-infrared band to set up a shadow layer.

Dictionary learning techniques are another option for cloudy data correction. Dictionary learning is a key component in sparse representation theory which has received a growing interest for the decomposition of signals into a subset of linear projections from an overcomplete dictionary. Dictionary learning methods aim at searching a data set to best represent the signals by only using a small subset of the dictionary (i.e., a few atoms). Sparse representation has been employed in many applications, such as image denoising [13], face recognition [14], feature extraction and classification [15], and image superresolution [16]. With regard to cloud removal, a method was proposed in [17] for the reconstruction of cloudy areas based on compressive sensing theory. The basis pursuit and orthogonal matching pursuit approaches were adopted for addressing the sparse representation. Genetic algorithms have been also exploited to obtain the best solution of an ℓ_0-norm optimization problem. Most recently, a method has been developed in [18] using two multitemporal dictionary learning methods based on the expanded K-SVD (K-means clustering process) and Bayesian algorithms to recover quantitative remote sensing products contaminated by thick clouds and shadows. This method sorts a set of time series data according to the temporal correlations for K-SVD and adjusts the weights in a Bayesian scheme. So far, cloud removal methods based on dictionary learning are very limited and have been only employed in the spatial domain.

In this paper, a cloud removal method is developed by combining multitemporal and dictionary learning methods. Specifically, we propose a new data restoration approach, which is called multitemporal dictionary learning (MDL), based on sparse representation to remove cloud and cloud shadow effects. Dictionaries of the cloudy areas (target data) and the cloud-free areas (reference data) are learned separately in the spectral domain, where each atom is associated with a ground cover component under the two aforementioned conditions. Meanwhile, cloud detection is not a required preprocessing step in the proposed method. The cloud information is reflected in the coefficients of the cloudy target image. The removal process is conducted by combining the coefficients from the reference image and the dictionary from the target image.

The remainder of this paper is organized as follows. Section II describes the proposed MDL method for reconstructing cloudy images. Section III gives an assessment of our experimental results. Finally, discussions and conclusions are presented in Section IV.

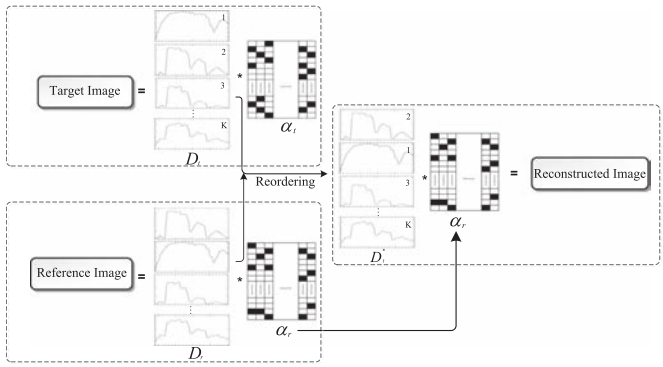

Fig. 1. Flowchart of the proposed cloud removal method.

II. Methodology

The proposed method is shown in Fig. 1. The cloudy image to be corrected is referred to as the target image. A clear image of the same location is used as a reference image. The method has three parts: 1) dictionary learning from target and reference images, respectively; 2) reordering dictionaries; and 3) reconstructing the cloud/shadow-free image. These three parts are presented in the following.

A. Training Dictionaries From Target and Reference Images

The goal of sparse representation is to define a given data vector $\boldsymbol{x} \in \mathbb{R}^n$ as a weighted linear combination of a small number of basis vectors. This representation is usually expressed as

$$\boldsymbol{x} = \boldsymbol{D}\boldsymbol{\alpha} \tag{1}$$

where $\boldsymbol{D} \in \mathbb{R}^{n\times k}$ is called the dictionary matrix, and typically, $k > n$. The columns of $\boldsymbol{D}$ are the basis vectors $\{\boldsymbol{d}_i\}_{i=1}^{k}$, which are often referred to as the *atoms* of $\boldsymbol{D}$. These atoms represent the dominant patterns in the input data $\boldsymbol{x}$. The elements of $\boldsymbol{\alpha} \in \mathbb{R}^k$ are sparse *coefficients*, which satisfy the constraint that the number of nonzero elements in $\boldsymbol{\alpha}$ is much smaller than k.

If $\boldsymbol{D}$ is given, the process of finding $\boldsymbol{\alpha}$ is called *sparse coding*. On the contrary, if $\boldsymbol{\alpha}$ is known, we use *dictionary learning* to find $\boldsymbol{D}$. *Compressed sensing* is a special case of sparse coding where the dictionary matrix takes a special form. Suppose $\boldsymbol{y} \in \mathbb{R}^k$ and $\Psi \in \mathbb{R}^{k\times k}$ represents an orthogonal basis. If $\boldsymbol{y}$ can be represented sparsely in the domain defined by the matrix Ψ, then $\boldsymbol{y}$ can be expressed as

$$\boldsymbol{y} = \Psi\boldsymbol{\alpha}. \tag{2}$$

The subsampled data vector $\boldsymbol{x} \in \mathbb{R}^n$ can be then defined as

$$\boldsymbol{x} = \Phi\boldsymbol{y} = \Phi\Psi\boldsymbol{\alpha} \tag{3}$$

where the matrix $\Phi \in \mathbb{R}^{n\times k}$ typically has the form of an identity matrix with random rows removed. Sparse coding can be then used to find $\boldsymbol{\alpha}$ given $\boldsymbol{x}$, Φ, and Ψ; and $\boldsymbol{y}$ can be found using (2).

In most studies, sparse representation theory has been developed for denoising and feature extraction. In these scenarios,

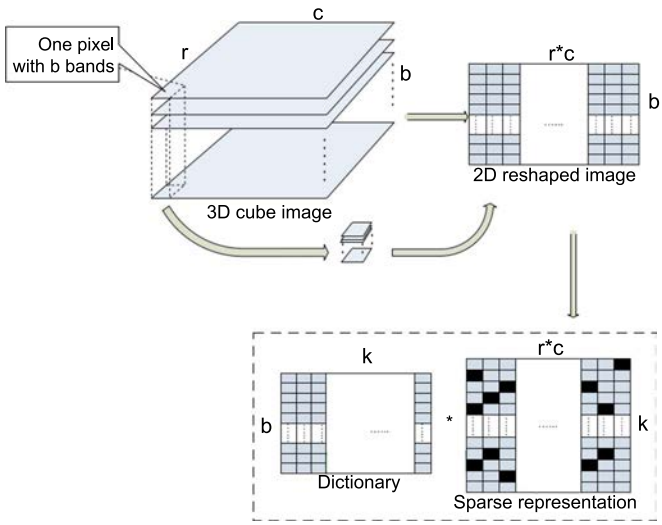

Fig. 2. Illustration of dictionary learning in the spectral domain.

sparse representations of the data in the image domain have been exploited, with a sliding window of various sizes. Given that cloud and cloud shadow effects do not follow a distinct pattern within a spatial window, the cloud contaminations are different for different pixels. On the other hand, the contamination on each spectral band of a given pixel is highly related, i.e., the true reflected spectrum from each pixel is modified when there is a cloud/shadow cover. Therefore, we propose to exploit sparse representations of the data in the spectral domain. The window size is $b \times 1$, where b is the number of spectral bands. A block diagram showing how the data are expressed, in our method, by a sparse representation in the spectral domain, is shown in Fig. 2.

Let $\mathbf{X} = [\boldsymbol{x_1}, \ldots, \boldsymbol{x_n}] \in \mathbb{R}^{b \times n}$ represent a spectral image signal with b bands and $n = r \times c$ pixels. Then, let $\boldsymbol{x}_i$ satisfy

$$\min_{\boldsymbol{D} \in \mathcal{C}, \alpha \in \mathbb{R}^{b \times n}} \sum_{i=1}^{n} \frac{1}{2} \|\boldsymbol{x}_i - \boldsymbol{D}\boldsymbol{\alpha_i}\|_2^2 \text{ s.t. } \boldsymbol{D} \geq 0, \quad \forall \boldsymbol{\alpha}_i \geq 0 \tag{4}$$

where $\boldsymbol{D} \in \mathbb{R}^{b \times k}$ is the dictionary, and each column $\boldsymbol{d}_i$ is a basis vector. Nonnegative constraints are enforced in the decomposition process to make the values of $\boldsymbol{\alpha}$ meaningful. $\mathcal{C}$ is the convex set of matrices with the following constraint:

$$\mathcal{C} \triangleq \left\{ \boldsymbol{D} \in \mathbb{R}^{b \times k} \text{ s.t. } \quad \forall j = 1, \ldots, k, \|\boldsymbol{d_j}\|_2 \leq 1 \text{ and } \boldsymbol{d}_j \geq 0 \right\}. \tag{5}$$

This is a joint optimization problem with respect to the dictionary $\boldsymbol{D}$ and the coefficients $\boldsymbol{A} = [\boldsymbol{\alpha_1}, \ldots, \boldsymbol{\alpha_n}] \in \mathbb{R}^{\boldsymbol{k} \times \boldsymbol{n}}$. This joint minimization problem is not convex. However, a convex minimization problem can be formulated with respect to each of $\boldsymbol{D}$ and $\boldsymbol{A}$ if the other is fixed. The online dictionary learning method proposed in [19] takes this approach, and we propose to use this method to solve the joint optimization problem described earlier.

The online dictionary learning method can be summarized as follows. The initial dictionary $\boldsymbol{D}_0$ is provided with random elements from a training set. Sparse coding steps are used to compute the coefficients $\boldsymbol{A}$ using the Least Angle Regression (LARS) method [20]. Then, the updated dictionary $\boldsymbol{D}_t$ is found by minimizing

$$\boldsymbol{D}_t \triangleq \arg\min_{\boldsymbol{D} \in \mathcal{C}} \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} \|\boldsymbol{x}_i - \boldsymbol{D}\boldsymbol{\alpha_i}\|_2^2 \tag{6}$$

using a block-coordinate descent approach with respect to the jth column $\boldsymbol{d}_j$ while keeping the other columns fixed with the constraint $\|\boldsymbol{d_j}\|_2 \leq 1$. This process is then repeated using the updated dictionary until a convergence criterion is satisfied (for a more detailed explanation of the online dictionary learning method, please refer to [19]).

This spectral domain sparse representation approach is applied to the problem of cloud and cloud shadow removal in the following manner. Let $\boldsymbol{X}_t \in \mathbb{R}^{b \times n}$ be the target image contaminated by clouds and $\boldsymbol{X}_r \in \mathbb{R}^{b \times n}$ be the reference image of the same geographical region from a clear day. Let $n = r \times c$ be the total number of pixels of an image of r rows by c columns, and b is the number of spectral bands. The columns of $\boldsymbol{X}_t$ and $\boldsymbol{X}_r$ correspond to the spectral vectors of each pixel. k is the number of columns in the dictionary. Let $\boldsymbol{D_t}$ and $\boldsymbol{D_r}$ denote the dictionaries of the target and reference images, respectively. Then, the sparse representation of the target and reference images can be expressed as

$$\boldsymbol{X}_t = \boldsymbol{D}_t \boldsymbol{A}_t + \boldsymbol{\varepsilon}_t \tag{7}$$

$$\boldsymbol{X}_r = \boldsymbol{D}_r \boldsymbol{A}_r + \boldsymbol{\varepsilon}_r. \tag{8}$$

Dictionaries $\boldsymbol{D}_t$ and $\boldsymbol{D}_r$ are learned separately using the online dictionary learning method introduced earlier.

In general, dictionaries are an overcomplete matrix and can reflect the basic patterns contained in each window. In our case, the dictionaries represent the spectra of fundamental components which all the pixels contain, and the sparse coefficients are the weights of the associated components. In other words, each measured pixel spectrum is a weighted sum of the spectra of a few selected fundamental components (atoms).

B. Dictionary Reordering

When there is a cloud cover, we can expect that the fundamental components that a pixel contains remain unchanged. Therefore, the same number of atoms for the two dictionaries is selected. However, the two dictionaries are not identical due to two reasons. First, the order of the atoms may not be the same since they are generated separately during each dictionary training. Second, the atoms corresponding to the same spectral components (subclasses) change with imaging conditions. Nevertheless, the corresponding atoms from the cloudy/shadowy image should have the same pattern as those from the clear image and are highly correlated. It is important that the two dictionaries are in the same order for performing the next step of the proposed method. Hence, reordering is performed using the correlation matrix (CM) between the atoms in the two dictionaries. The CM is determined by calculating the correlation coefficient (CC) between each atom in $\boldsymbol{D_t}$ and $\boldsymbol{D_r}$. Every element in the CM is the CC of the two dictionaries. As a result, the CM is a $k \times k$ matrix in which the nth column

represents the CC of the nth atom in $\boldsymbol{D_t}$ with all the atoms in $\boldsymbol{D_r}$. The reordered matrix is generated by searching the highest CC value of each column in the CM. Each column is reordered by moving the highest CC value to the nth row in the nth column. Then, each column in $\boldsymbol{D_t}$ will change according to the movement in the CM. These steps are used to make sure that the two dictionaries have the same order. The new reordered dictionary is denoted by $\boldsymbol{D_t^*}$.

It has been observed that both dictionaries are composed of the spectral vectors of various fundamental components of land cover materials. There is no cloud component in the dictionary which is learned from the cloudy image (target image), as illustrated in the experimental section. This result indicates that the dictionary is not affected by the cloud cover. This is an important property and leads to the proposed method for cloudy image restoration.

C. Cloudy Image Restoration

Based on the concept of sparse representation, we know that $\boldsymbol{A}_r$ determines which atoms are associated with a given pixel and what the corresponding weights are. This association will not be changed due to cloud cover. However, the magnitudes of the atoms may be different and are affected by the different imaging conditions, such as the changes in atmospheric parameters. Therefore, to recover the signals for the target image, we propose to use the (reordered) target image's dictionary $\boldsymbol{D_t^*}$ and the reference image's coefficients $\boldsymbol{A}_r$ to reconstruct the clear image. Specifically, the clear image is obtained as follows:

$$\boldsymbol{X}_t^c = \boldsymbol{D}_t^* \boldsymbol{A}_r \tag{9}$$

where $\boldsymbol{X}_t^c$ is the reconstructed image after removing cloud and cloud shadows.

Our newly proposed MDL method can be applied to various types of remote sensing data, including multispectral and hyperspectral images. The only parameter that needs to be adjusted is the size of the dictionary k. Differing from conventional cloud/shadow removal approaches, the proposed method can be conducted without screening clouds or cloud shadows as a preprocessing step. Moreover, it is not affected by the size of the cloud/shadow cover area and the complexity of the background.

III. Experiments and Results

A. Experiments on a Simulated Image

1) Generation of a Simulated Cloudy Image: In order to quantify the performance of the MDL approach, simulation experiments are described here. In this experiment, cloud cover is simulated by utilizing the cirrus band (band 9) from Landsat OLI. The cirrus band at 1.375 μm has very strong water vapor absorption. Therefore, the background underneath clouds will become dark in this particular band. The brighter the cloudy pixels, the thicker the clouds are in this band. The boundary of clouds is displayed clearly in the single cirrus cloud image in Fig. 3. In order to make the simulated cloud cover closer to the true case, we also apply a cloud signature to this image. This way, a simulated cloudy image was generated, as shown

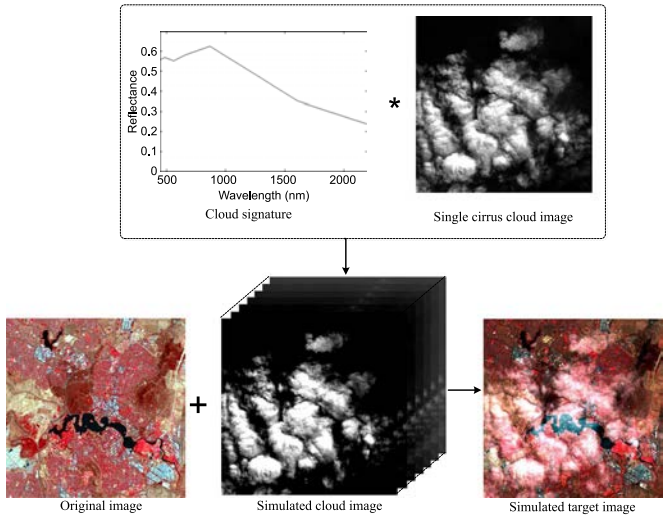

Fig. 3. Flowchart of the procedure adopted to generate a simulated cloudy image.

in Fig. 3. The cloudy target image is simulated by overlaying the original clear image acquired on January 15, 2014 with the simulated cloud data, as shown in Fig. 3. The simulated target image is contaminated by clouds over different areas and with varying thickness.

2) Cloudy Image Dictionaries: With regard to the dictionary learning process, the online dictionary learning algorithm was exploited for sparse representation because of its good performance on large data sets (e.g., pixelwise spectral image processing) and its increased speed from using the SPArse Modelling Software toolbox [19], [21]. A nonnegative matrix factorization [22] is also included to ensure that the elements in $\boldsymbol{D}$ and $\boldsymbol{\alpha}$ are positive. Due to the dictionary learning method being applied in the spectral domain, the patch size in the training process is fixed, i.e., 7×1 for OLI data.

In order to test the dictionary learning method over different types of ground cover, two regions were selected, i.e., cloudy water and vegetation areas. In this test, the two dictionaries are learned when $k = 5$ and shown in Fig. 4(a) and (b), respectively. It is clear that no cloud element appears in the two plots of water and vegetation dictionaries, although the cloudy image is simulated by cloud signature in Fig. 3. The online dictionary learning was also conducted on the whole target cloudy image. Fig. 4(c) shows the plot of dictionary atoms extracted from the whole simulated target image when $k = 20$. From all the three plots, we can see that dictionary learning can extract the fundamental materials from the image contaminated with clouds and a cloud signature is not selected in this process.

3) Comparisons of Target and Reference Image Dictionaries: A reference image was collected on December 30, 2013 in the same region as the original image. This image scene covers the Canberra region in Australia and consists of heterogeneous urban, mountain, and lake areas. The reference dictionary is also learned when $k = 20$ in the target dictionary. The two extracted dictionaries are shown in Fig. 5 as $\boldsymbol{D_t}$ and $\boldsymbol{D_r}$. Although the components from the two images exhibit a similar pattern, it is still hard to ensure that they are of the same order. Therefore, dictionary reordering is implemented based on the

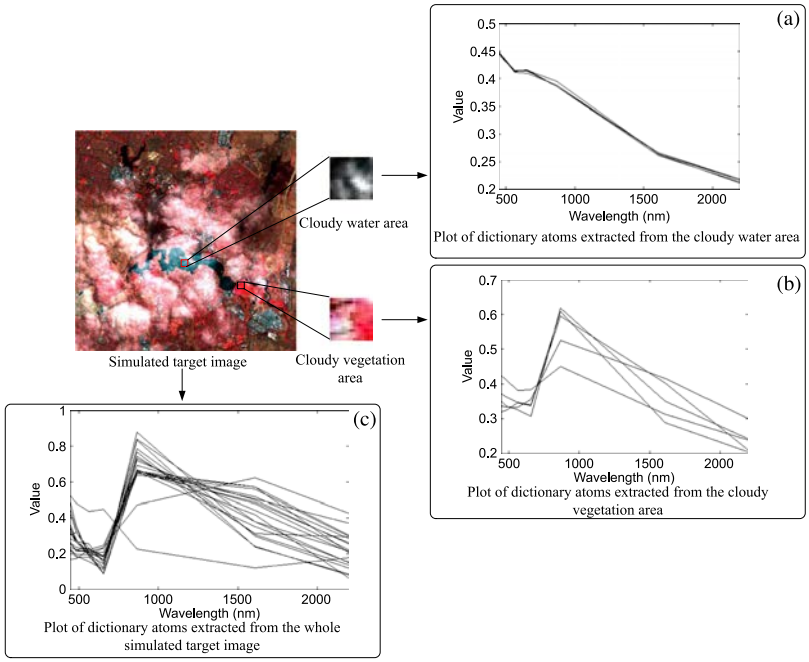

Fig. 4. Illustration of dictionary learning on the simulated target image. The original image was acquired on January 15, 2014.

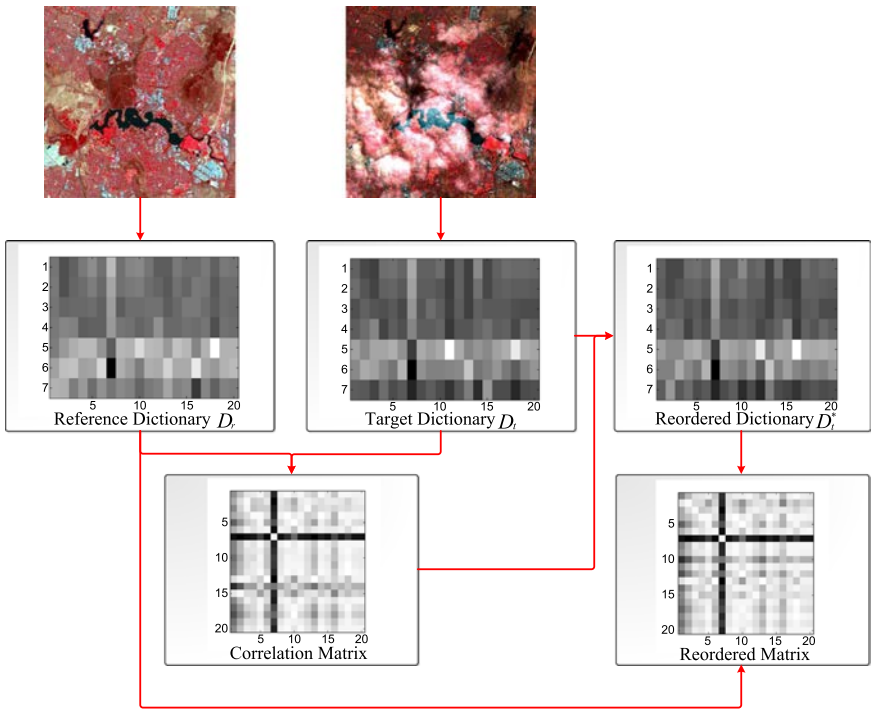

Fig. 5. Illustration of the reordering procedure on the target dictionary according to the dictionary of the reference image.

proposed method. The CM is first calculated to present the correlation coefficients between each column in $\boldsymbol{D_t}$ and $\boldsymbol{D_r}$. The purpose of reordering is to make the atom in $\boldsymbol{D_t}$ have the highest correlation with the corresponding atom in $\boldsymbol{D_r}$. The reordered dictionary $\boldsymbol{D_t^*}$ is adjusted this way. The CM after reordering is shown in Fig. 5 as well. There is little difference between the two matrices, and it should be noted that the correlation coefficients of each of the corresponding elements are reasonably high before reordering. The mean value of the correlation coefficients for each corresponding element in the target and reference dictionaries increased from 0.87 to 0.92 after reordering. Therefore, we can conclude that the reordering procedure is a necessary process to make sure that the sparse coefficients represent the associated components as accurately as possible.

4) Cloud Removal Results: To draw comparisons, the MNSPI method [9] was also applied to the simulated image in our experiments. Mean absolute percentage error (MAPE),

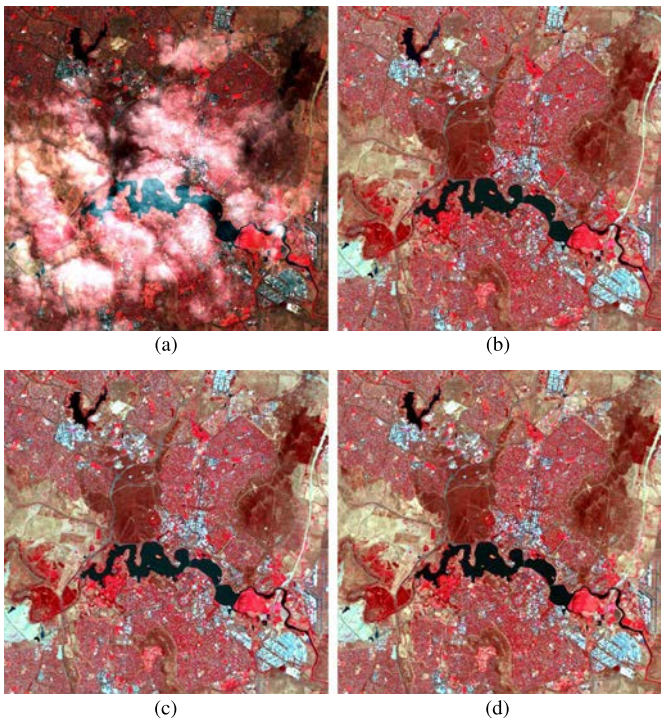

Fig. 6. Landsat 8 OLI images used in the simulation experiments, with false color composite $R =$ band 5, $G =$ band 4, $B =$ band 3. (a) Simulated target cloudy image. (b) Reference image acquired on December 30, 2013. (c) Recovered image using the MNSPI method. (d) Recovered image using the proposed MDL method.

TABLE I
QUANTITATIVE RESULTS OF SIMULATION EXPERIMENTS USING THE PROPOSED MDL AND MNSPI METHODS

Band	Method	MAPE	PSNR	CC
Band 3	Simulation	1.15	10.74	0.27
	MNSPI	0.86	11.76	0.78
	MDL	**0.18**	**23.76**	**0.84**
Band 4	Simulation	1.17	10.51	0.38
	MNSPI	0.88	11.28	0.78
	MDL	**0.2**	**23.6**	**0.84**
Band 5	Simulation	0.5	14.42	0.6
	MNSPI	0.38	15.05	0.84
	MDL	**0.2**	**18.51**	**0.88**

peak signal-to-noise ratio (PSNR), and CC were used to assess the proposed method quantitatively. The recovered images using MNSPI and the proposed MDL method are shown in Fig. 6(c) and (d), respectively. The MAPE, PSNR, and CC indexes of the simulated cloudy image and the images recovered by MNSPI and MDL versus the original clear image are shown in Table I. The MAPE is defined as $(1/N)\sum_{i=1}^{n}|(x_o(i) - x_c(i))/x_o(i)|$, where N is the total number of pixels contaminated by clouds, and $x_o(i)$ and $x_c(i)$ are the original and recovered image values, respectively. Three bands are evaluated in Table I. The recovered image using the MDL method has the lowest MAPE value, which means that it is closest to the original true data. The simulated cloudy image has a very low PSNR value (10.74 dB, 10.51 dB, 14.42 dB), indicating that the simulated clouds strongly affect the quality of the image. The CCs for MDL improve more than for the MNSPI method in all three bands. The MNSPI method is effective in recovering the edges of regions, but is sensitive to the size of the clouds. The large patches of simulated clouds decrease the accuracy of recovery by MNSPI. From the simulation experiments, we can conclude that the proposed MDL method provides more accurate restoration results.

B. Experiments on Real Images

1) Cloud Removal for a Hyperion Image: In order to better substantiate our restoration results, we also performed experiments on real hyperspectral data. One data set consisting of EO-1 Hyperion images was tested. Two Hyperion images acquired on March 22, 2003 and March 6, 2003 were selected as target and reference images, respectively. The target image is contaminated by clouds and cloud shadows over large areas of the image. Fig. 7 illustrates the details of the recovery process using the proposed MDL method on EO-1 data. The first column shows the target, reference, and recovered Hyperion images with composite colors of R: band 49 (0.844 μm), G: band 45 (0.803 μm), and B: band 10 (0.447 μm). The dictionaries and coefficients extracted from the target and reference images, respectively, are shown in the second column in Fig. 7. The number of elements in the dictionary is 20. The recovered image was generated by combining the reordered target image dictionary $\boldsymbol{D}_t^*$ and the reference image coefficients $\boldsymbol{A}_r$. A cloudy pixel, which is labeled by the red X, was selected as an illustrative example. The spectral profiles of this pixel in the target, reference, and recovered images are shown in the third column. Each legend shows the coefficients corresponding to each extracted atom (spectra) in the dictionary. The marked pixel is generated by the sum of these values. Spectral profiles of the labeled pixel show that the MDL method can restore the signatures of the ground covered by clouds. The difference between the reference and recovered spectra is regarded as the difference in atmospheric conditions for the two dates. The cloud shadow regions are also recovered in the results.

2) Cloud Removal for an OLI Image: Here, one real data set from Landsat 8 OLI data is investigated using the proposed MDL method. The experimental OLI images were downloaded from the National Aeronautics and Space Administration website (http://earthexplorer.usgs.gov/). The OLI is a 30-m spatial resolution optical sensor with a 16-day revisit time. Therefore, it is possible to acquire time series data in the same geographical area. The experimental target image was acquired on December 4, 2013, and the reference image was acquired on September 18, 2013. These images contain 500 $\times$ 500 pixels with seven spectral bands.

In Fig. 8, the target image is contaminated by relatively thin clouds over almost the entire image. The recovered image, as shown in Fig. 8(c), is visually clear and close to the reference image acquired on September 18. The number of dictionary atoms is 30 in this experiment. The homogeneous and heterogeneous areas are both corrected. In the top right corner of the scene, the urban area is restored well. Fig. 8(e)–(f) shows zoomed images revealing clear edges and high quality. The results indicate that the proposed MDL method can recover data affected by different types of clouds and is not sensitive to the type of ground covers.

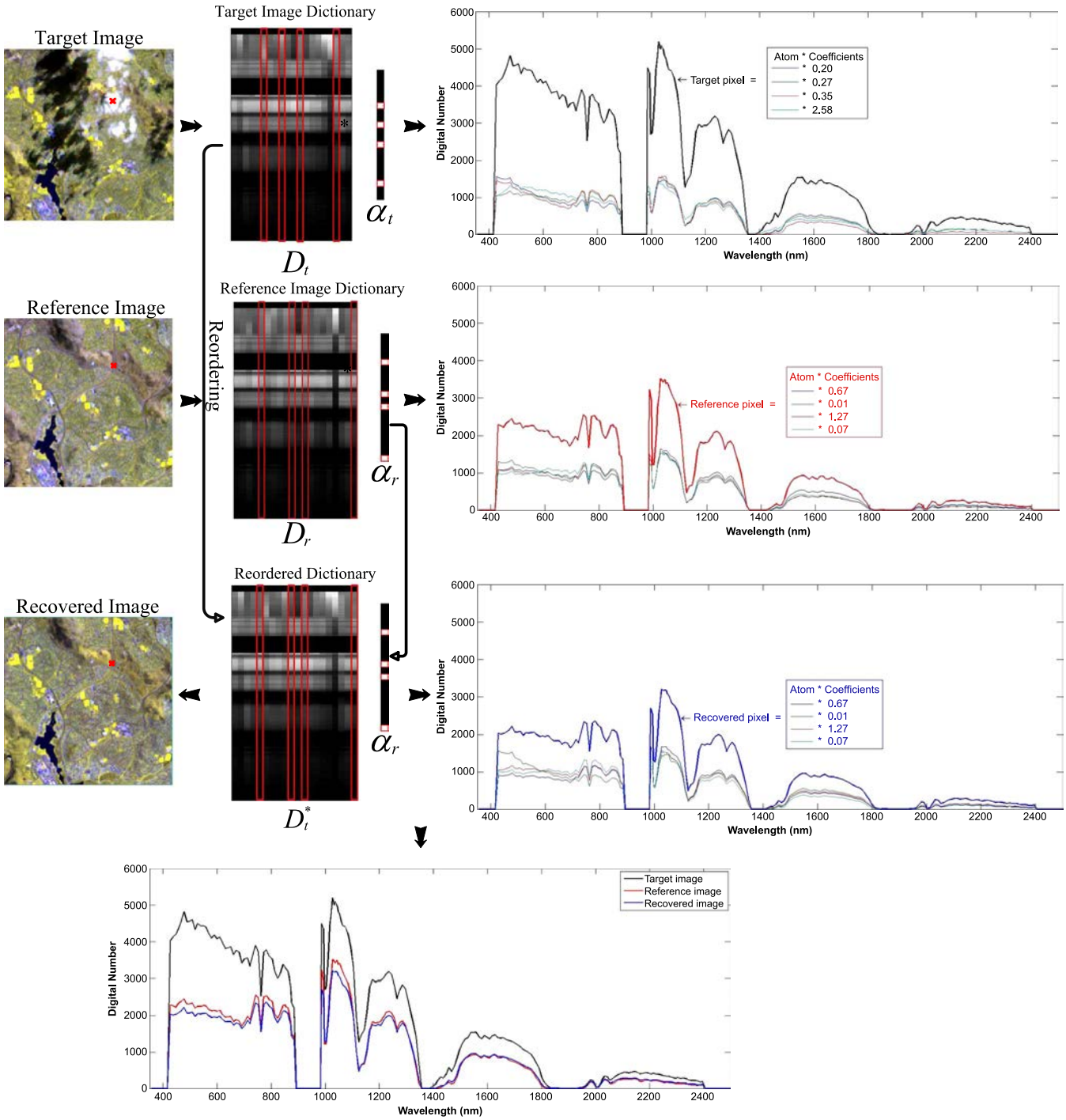

Fig. 7. Details of cloud removal results for EO-1 Hyperion data. The target and reference images were acquired on March 22, 2003 and March 6, 2003, respectively.

IV. Discussion and Conclusions

In this paper, a cloud removal method, which is called MDL method, has been proposed. A reference image is required as auxiliary data for determining the fundamental components in each pixel contaminated by clouds. The dictionary learned from the target image is reordered to match the atoms in the reference image. The recovery process is conducted by combining the reordered target dictionary and the sparse coefficients from the reference image. Simulated and real data were both quantitatively and qualitatively investigated in our experiments. The recovery results show better performance of the proposed MDL method when compared with the MNSPI method. Our proposed method can handle thin and thick clouds, as well as cloud shadows. The size of the contaminated areas is not a limitation for the proposed method. It should be noted that the only parameter that needs to be set manually in the recovery procedure is the number of atoms in the dictionary. In the experiments, various values for this parameter were tested to

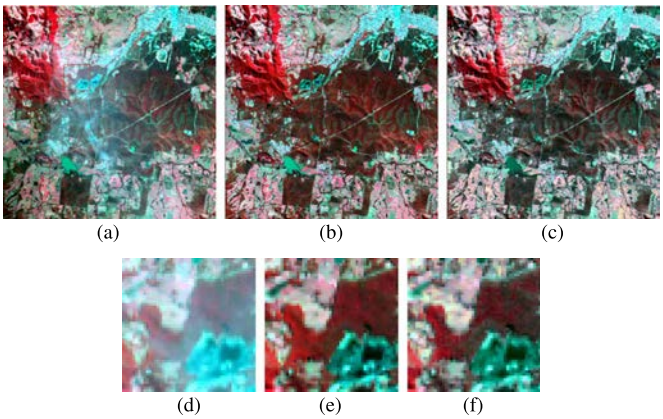

Fig. 8. Cloud removal results for the Landsat 8 OLI images in real data experiments. (a) Target image acquired on December 4, 2013. (b) Reference image acquired on September 18, 2013. (c) Recovered image using the proposed MDL method. (d)–(f) Zoomed images of the square area marked in (a)–(c).

obtain the best restoration results. This testing process indicated that the optimal value of this parameter varies for different image scenes. In the future, the reasons for this variation will be investigated. Further applications of the MDL method will be also explored.

ACKNOWLEDGMENT

The authors would like to thank the Editors and the anonymous reviewers for their outstanding comments and suggestions, which greatly helped them to improve the technical quality and presentation of this paper.

REFERENCES

[1] J. Ju and D. P. Roy, "The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 1196–1211, Mar. 2008.
[2] M. Xu, X. Jia, and M. Pickering, "Automatic cloud removal for Landsat 8 OLI images using cirrus band," in *Proc. IEEE IGARSS*, 2014, pp. 2511–2514.
[3] Y. Zhang, B. Guindon, and J. Cihlar, "An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images," *Remote Sens. Environ.*, vol. 82, no. 2/3, pp. 173–187, Oct. 2002.
[4] C. Ji, "Haze reduction from the visible bands of Landsat TM and ETM+ images over a shallow water reef environment," *Remote Sens. Environ.*, vol. 112, no. 4, pp. 1773–1783, Apr. 2008.
[5] C.-H. Lin, P.-H. Tsai, K.-H. Lai, and J.-Y. Chen, "Cloud removal from multitemporal satellite images using information cloning," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 232–241, Jan. 2013.
[6] D.-C. Tseng, H.-T. Tseng, and C.-L. Chien, "Automatic cloud removal from multi-temporal SPOT images," *Appl. Math. Comput.*, vol. 205, no. 2, pp. 584–600, Nov. 2008.
[7] J. Chen, X. Zhu, J. E. Vogelmann, F. Gao, and S. Jin, "A simple and effective method for filling gaps in Landsat ETM+ SLC-off images," *Remote Sens. Environ.*, vol. 115, no. 4, pp. 1053–1064, Apr. 2011.
[8] X. Zhu, D. Liu, and J. Chen, "A new geostatistical approach for filling gaps in Landsat ETM+ SLC-off images," *Remote Sens. Environ.*, vol. 124, pp. 49–60, Sep. 2012.
[9] X. Zhu, F. Gao, D. Liu, and J. Chen, "A modified neighborhood similar pixel interpolator approach for removing thick clouds in Landsat images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 3, pp. 521–525, May 2012.
[10] A. Shahtahmassebi, N. Yang, K. Wang, N. Moore, and Z. Shen, "Review of shadow detection and de-shadowing methods in remote sensing," *Chin. Geogr. Sci.*, vol. 23, no. 4, pp. 403–420, Aug. 2013.
[11] D. Lu, "Detection and substitution of clouds/hazes and their cast shadows on ikonos images," *Int. J. Remote Sens.*, vol. 28, no. 18, pp. 4027–4035, Sep. 2007.
[12] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83–94, Mar. 2012.
[13] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.
[14] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
[15] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3973–3985, Oct. 2011.
[16] 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.
[17] L. Lorenzi, F. Melgani, and G. Mercier, "Missing-area reconstruction in multispectral images under a compressive sensing perspective," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 3998–4008, Jul. 2013.
[18] X. Li *et al.*, "Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 7086–7098, Nov. 2014.
[19] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *J. Mach. Learn. Res.*, vol. 11, pp. 19–60, 2010.
[20] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Stat.*, vol. 32, no. 2, pp. 407–499, Apr. 2004.
[21] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, pp. 689–696.
[22] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 556–562.

Meng Xu (S'13) received the B.S. and M.S. degrees in electrical engineering from Ocean University of China, Qingdao, China, in 2011 and 2013, respectively. She is currently working toward the Ph.D. degree in the School of Engineering and Information Technology, University of New South Wales, Canberra, Australia.

Her research interests include cloud removal and remote sensing image processing.

Xiuping Jia (M'93–SM'03) received the B.Eng. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 1982 and the Ph.D. degree in electrical engineering from the University of New South Wales, Canberra, Australia, in 1996.

Since 1988, she has been with the School of Information Technology and Electrical Engineering, University of New South Wales, where she is currently a Senior Lecturer. She is also a Guest Professor with Harbin Engineering University, Harbin, China, and an Adjunct Researcher with the National Engineering Research Center for Information Technology in Agriculture, Beijing. She has coauthored the remote sensing textbook *Remote Sensing Digital Image Analysis* [Springer-Verlag, 3rd (1999) and 4th eds. (2006)]. Her research interests include remote sensing, imaging spectrometry, and spatial data analysis.

Dr. Jia is an Editor of the *Annals of GIS* and an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Mark Pickering (S'92–M'95) was born in Biloela, Australia, in 1966. He received the B.Eng. degree from the Capricornia Institute of Advanced Education, Rockhampton, Australia, in 1988, and the M.Eng. and Ph.D. degrees from the University of New South Wales, Canberra, Australia, in 1991 and 1995, respectively, all in electrical engineering.

He was a Lecturer from 1996 to 1999 and a Senior Lecturer from 2000 to 2009 with the School of Electrical Engineering and Information Technology, University of New South Wales, where he is currently an Associate Professor. His research interests include video and audio coding, medical imaging, data compression, information security, data networks, and error-resilient data transmission.

Antonio J. Plaza (M'05–SM'07–F'15) received the Bachelor's, M.Sc., and Ph.D. degrees from the University of Extremadura, Extremadura, Spain, in 1997, 1999, and 2002, respectively, all in computer engineering.

He is currently an Associate Professor (with accreditation for Full Professor) with the Department of Technology of Computers and Communications, University of Extremadura, where he is the Head of the Hyperspectral Computing Laboratory. From 2007 to 2011, he was the Coordinator of the Hyperspectral Imaging Network, a European project with a total funding of 2.8 million euros. He has authored over 400 publications, including 126 JCR journal papers (78 in IEEE journals), 20 book chapters, and over 240 peer-reviewed conference proceeding papers (94 in IEEE conferences). He has edited the book *High-Performance Computing in Remote Sensing* (CRC Press/Taylor and Francis) (the first book on this topic in the published literature) and guest edited eight special issues on hyperspectral remote sensing for different journals.

Dr. Plaza served as the Chair for the 2011 IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing and the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) from 2011 to 2012. He has been serving as the President of the Spanish Chapter of IEEE GRSS since November 2012. He served as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING from 2007 to 2012. He was a member of the Editorial Board of the IEEE Geoscience and Remote Sensing Newsletter from 2011 to 2012 and a member of the steering committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING in 2012. He is a Guest Editor of seven special issues on JCR journals (three in IEEE journals). He is also an Associate Editor of IEEE ACCESS and the IEEE GEOSCIENCE AND REMOTE SENSING MAGAZINE. He has been serving as the Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING since January 2013. He was recipient of the Best Reviewers of the IEEE Geoscience and Remote Sensing Letters in 2009 and of the Best Reviewers of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2010.