A Framework of Mixed Sparse Representations for Remote Sensing Images

Feng Li, *Senior Member, IEEE*, Lei Xin, Yi Guo, Junbin Gao, and Xiuping Jia, *Senior Member, IEEE*

***Abstract*— In this paper, a new framework of mixed sparse representations (MSRs) is proposed for solving ill-conditioned problems with remote sensing images. In general, it is very difficult to find a common sparse representation for remote sensing images because of complicated ground features. Here we regard a remote sensing image as a combination of subimage of smooth, edges, and point-like components, respectively. Since each domain transformation method is capable of representing only a particular kind of ground object or texture, a group of domain transformations are used to sparsely represent each subimage. To demonstrate the effect of the framework of MSR for remote sensing images, MSR is regarded as a prior for maximum *a posteriori* when solving ill-conditioned problems such as classification and super resolution (SR), respectively. The experimental results show that not only the new framework of MSR can improve classification accuracy but also it can construct a much better high-resolution image than other common SR methods. The proposed framework MSR is a competitive candidate for solving other remote sensing images-related ill-conditioned problems.**

***Index Terms*— Classification, compressive sensing (CS), mixed sparse representations (MSRs), super-resolution (SR).**

I. Introduction

A. Sparsity for Remote Sensing Images

SPARSITY describes a signal with few nonzero elements or few nonzero coefficients in a particular transform domain or basis. Sparsity driven techniques have become very popular with the development of L1-norm-based optimization methods [1] and the emergence of the sampling theory of compressive sensing (CS) [2]–[4]. CS works well for sparse signals or sparse representations with respect to a particular transform domain or basis. So far, it has been applied in many research areas [5]–[11]. Sparsity is one of the important ingredients of CS and is a very useful property for image processing such as compression [12], denoising [13], deblurring [14], and unmixing [15]. In [14] and [16], sparsity in astronomical images is studied. However, modeling the sparsity in the remote sensing images [17]–[19] is far more complicated than point sources or extended sources in astronomical images.

In general, either a dictionary learning method or a domain transformation method has been used to explore the sparsity in image processing community. However, dictionary learning methods highly depend on the training data set, and it is questionable when a set of well-trained dictionaries is extended to some previously unseen data. As far as domain transformation methods are concerned, a particular transformation is good at providing sparse representation for a certain type of input signal. For example, wavelet transform [20], [21] is a suitable tool of sparse representing smooth content in images; curvelet transform [22] can well represent global line structures of any directions; discrete cosine transform is excellent in representing either smooth or periodic components in images.

Wavelet transform has been well known for its multiscale/multiresolution analysis ability in signal processing regime. This has brought significant progresses in many research areas such as image compression [12], [23], image denoising [24], [25], and image restoration [17], [26]–[28]. Though wavelets are certainly suitable for dealing with objects of smooth regions, they are ill suited for detecting, organizing, or providing a compact representation of intermediate dimensional structures.

Curvelet transform [22], which was created to overcome the disadvantages of multiscale/multiresolution representations such as wavelets, is a multiscale pyramid with many directions and positions at each length scale and needle-shaped elements at fine scales. Based on [22], we know that curvelets can always offer a much more optimal performance than wavelets when representing edges of objects.

Generally speaking, each domain transformation method is capable of representing only a particular kind of ground object or texture. In this paper, we propose a general framework of sparse representations for remote sensing images, in which a remote sensing image is regarded as a combination of subimages of smooth, edges, and point-like components, respectively, and a group of domain transformations are used to sparsely represent each subimage. To verify the effectiveness of the framework of mixed sparse representations (MSRs) for remote sensing images, MSR is regarded as a prior for maximum *a posteriori* (MAP) when solving ill-conditioned problems such as classification and super resolution (SR), respectively.

Manuscript received May 24, 2016; revised August 30, 2016; accepted October 15, 2016. Date of publication November 16, 2016; date of current version December 29, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 41371415 and in part by Australian Research Council under Grant DP130100364. *(Corresponding author: Feng Li.)*

F. Li and L. Xin are with the Qian Xuesen Laboratory of Space Technology, Beijing 100094, China (e-mail: lifeng@qxslab.cn; xinlei@qxslab.cn).

Y. Guo is with the School of Computing, Engineering and Mathematics, Western Sydney University, Parramatta, NSW 2150, Australia (e-mail: y.guo@westernsydney.edu.au).

J. Gao is with the Discipline of Business Analytics, The University of Business School, The University of Sydney, NSW 2006, Australia (e-mail: junbin.gao@sydney.edu.au).

X. Jia is with the School of Engineering and Information Technology, University of New South Wales at Canberra, ACT 2600, Australia (e-mail: X.JIA@adfa.edu.au).

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.2016.2621123

B. Classification

Hyperspectral sensors used in remote sensing usually capture several images from frequency bands/channels in the visual and nonvisual range with 5–20-nm wavelength between bands/channels. Hyperspectral images-based classification [29], [30] is a very important branch of remote sensing community. Based on the spectral measurements at a position, the object's class may be recognized; therefore, the main task of classification is to assign a label to each pixel in order to generate a land cover map [31]–[33].

Classification methods are typically divided into supervised and unsupervised classification. Here, we focus on the supervised classification only for better evaluation purpose. The framework proposed in this paper can be easily extended to unsupervised classification as well. There are several ways to improve classification accuracy such as developing more effective classification methods. Classification methods have been investigated for many years, and dozens of new methods are proposed each year [34]–[39]. One of the new trends in this research area is to incorporate both spectral and spatial constraints in the model [36], [40], [41]. Rather than developing new classification methods to improve the labeling performance, we will test the proposed framework to add spatial constraints to the input data by augmenting the original hyperspectral data cube.

C. Super-Resolution

SR has many different definitions in different research communities. In this paper, SR means a technique of restoring a high spatial resolution image from a series of lower resolution images of the same scene captured at different times in a short period. Optical resolution in itself is a big topic, which means the ability of resolving details of the object that being imaged. Roughly speaking, it involves two critical concepts: lens resolution and sensor resolution. First, lens resolution that is effected and limited by the Rayleigh criterion, i.e., the angular resolution $\theta = 1.22(\lambda_{\text{light}}/D_{\text{aperture}})$, where λ_{light} denotes the incoming light and D_{aperture} is the diameter of the lens aperture. An Airy disk is formed, of which the diameter d_{Airy} is

$$d_{\text{Airy}} = 2.44\frac{f_{\text{length}}\lambda_{\text{light}}}{D_{\text{aperture}}} \tag{1}$$

where f_{length} is the focal length of the lens.

Sensor resolution indicates the ability to accurately capture all of the details from the incoming light, and it depends on the sampling interval. Based on the Nyquist sampling theorem, the sampling interval has to be twice of the highest specimen spatial frequency in order to accurately preserve all details in a digital image. In other words, two detectors have to be placed in the Airy disk [42]. When and only when lens resolution matches sensor resolution, the best optical resolution can be achieved. Strictly speaking, given a fixed lens resolution, doubling the density of the sensor array with the same area of the array size does not improve twice of the spatial resolution of the imaging system, but it can help to improve the accuracy of measuring the features from the captured images. Thus, we should bear in mind that the SR technology discussed in this paper mainly works in the same way.

Since the first SR technology was proposed in 1984 [43], many SR methods have emerged and can be roughly subdivided into the following main categories: the frequency domain approaches, spatial domain methods, and learning-based methods. The frequency domain approaches are early methods used for SR from 1984 [43]. Assuming that the images suffer global translational motion [44], the frequency domain-based methods combine all high-frequency information in the Fourier domain and restore a high-resolution image. The main advantages of the frequency domain approaches include theoretical simplicity and straightforward implementation, which can be paralleled.

The spatial domain methods include the projection onto convex sets (POCS) approach, iterative back projection (IBP) approach, and the probabilistic reconstruction approach such as MAP. According to the method of POCS, incorporating prior knowledge into the formulation can be interpreted as a restriction on global optimality on a convex set. Simplicity is the major advantage of POCS [45]; however, nonuniqueness of the solution, slow convergence, and high computational cost are the main disadvantages.

IBP [46] is similar to the back projection used in tomography. The algorithm emulates the imaging process, back projects the error between the emulated and actual low-resolution (LR) images to the SR image, and iteratively updates it. The advantage of IBP is that it is intuitive and easy to understand; a shortcoming is that the method cannot deal with noise contaminated LR images, because the noise will be regarded as an error to be projected back to the SR image.

MAP is one of the most popular probabilistic methods. This stochastic regularization technique requires a probability density function of the data known as a prior image model that represents the distribution of the image pixel values. Li *et al.* [17], [47] proposed a MAP method based on the universal hidden Markov tree model MAP-uHUM. The method regards a universal hidden Markov tree model, which can be calculated by training in the wavelet domain, as a prior model. The results on real data show that the method works well. The main advantages of MAP-based methods [48], [49] are the direct inclusion of *a priori* constraints, such as the Huber function [50], for the ill-conditioned problem. However, the disadvantage [48], [49] is that the blur of the LR images is assumed to be due to simple averaging, which is not suitable for real applications.

A fast and robust SR algorithm was proposed in [51], in which a cost function of the difference between the model of the observations and the actual observations was optimized, plus a regularization term: bilateral total variation (BTV). Although it is robust against minor incorrect estimation of warping model, blur kernel, and noise variance, it cannot deal with remote sensing images properly, because they assume global translational shifts only for the warping model and severe misalignment largely decreases the reconstruction quality. Recently, Bayesian image SR methods were proposed [52], [53]. By marginalizing over registration parameters, this approach allows the calculation of the registration

parameters concurrently while obtaining the super-resolved images.

In general, the learning-based SR methods work in the following way [54]–[56]: learn a degeneration model of high-resolution images from LR images first, and then reconstruct the original high-resolution images based on the degeneration model. Miura *et al.* [55] proposed a learning-based BTV SR method, and the method uses principle components analysis to remove noise. The results indicate that the method can provide good performance turning National Television Standards Committee TV signal to HDTV. The main shortcoming of this kind of method is slowness. The step of searching the matched high-resolution block from a huge data set consumes a lot of time.

Recently, Candès and Fernandez-Granda [57] wrote a paper on SR. A theorem was proposed that if we assume nonzero components of high-resolution signal being separated by at least $2/f_c$ distance of sampling grid, where f_c is the cut-off frequency, perfect SR results can be achieved via L1-norm minimization. Their efforts shined a light on the behind theorem of SR from a mathematical point of view. However, there is limited shining stuff we can borrow from the paper. First, remote sensing images are not sparse in the spatial domain and we cannot guarantee that the high-resolution signal is separated by at least $2/f_c$ distance of sampling grid. Second, they did not consider the circumstance of which there are several LR sampled signals without a uniform sampling grid.

In summary, SR methods suffer some problems in real applications for the remote sensing community. One of the most challenging problems is that there is rarely general sparse representation for solving the typical ill-conditioned problem of SR.

II. Methodology

A. MSR for Classification

It is very difficult to find a general transform that can explore sparsity for the images with complex features, especially, for remote sensing images where heterogeneous contents are present all the time. We assume that a remote sensing image $\mathbf{z}$ consists of three types of subimages: the point components $\mathbf{z}_s$ in the image domain itself, the smooth components $\mathbf{z}_w$ showing sparsity in the wavelet domain, and curve/line components $\mathbf{z}_c$ showing sparsity in the curvelet domain. Therefore, we have

$$\mathbf{z} = \mathbf{z}_s + \mathbf{z}_w + \mathbf{z}_c \tag{2}$$

where $\mathbf{z}_s$, $\mathbf{z}_w$ and $\mathbf{z}_c$ are of the same size as $\mathbf{z}$. In some circumstances, sparsity is regarded as a prior for signal in MAP, so our intention is to apply the MSRs as a prior for solving a remote sensing image-related MAP problem. From (2), we can see that there are many potential combinations for those subimages, and therefore, it is a typical ill-conditioned problem.

Because noise contamination is inevitable for all imaging facilities including hyperspectral imagers, we proposed a subimage separation model for classification

$$\begin{aligned}\min\ &\{\lambda_1\|W(\mathbf{z}_w)\|_1 + \lambda_2\|C(\mathbf{z}_c)\|_1 + \lambda_3\|\mathbf{z}_s\|_1\}\\ &\text{s.t. } \mathbf{z} = \mathbf{z}_s + \mathbf{z}_w + \mathbf{z}_c + \mathbf{e}, \quad \|\mathbf{e}\|_\infty \le \delta\end{aligned} \tag{3}$$

where W denotes the wavelet transform matrix, C is the curvelet transform matrix, λ_1, λ_2 and λ_3 denote different weights for the contribution from different transformations, $\mathbf{e}$ is the noise following some distributions, say Gaussian distribution, and δ is the variance of the noise. If we select a large λ_1, which means we put more weight on the smooth components and more smooth-like structures will be shown in the resolved subimage $\mathbf{z}_s$. In other words, λ_1–λ_3 mainly work as balancing parameters among those different structures. If we set both λ_1 and λ_2 as 0, then (3) will degenerate to a basis pursuit denoising problem [4]. In general, those balancing parameters can be set based on experience and the understanding of the input images. The above separation method can be run for each channel of the hyperspectral image (which is stored as a data cube) by solving (3) and will create a three times larger data cube.

In the above equations, we use L1-norm minimization, which is the foundation of CS [4] and it has the property of producing many coefficients with zero values and very few nonzero coefficients, i.e., it stimulates the sparsity. We can apply the augmented Lagrange multiplier (ALM) method to solve (3), and the convergence analysis can be derived as follows. By introducing variables $\mathbf{u} = W(\mathbf{z}_w)$ and $\mathbf{v} = C(\mathbf{z}_c)$ in (3), we have

$$\begin{aligned}&\min\ \{\lambda_1\|W(\mathbf{z}_w)\|_1 + \lambda_2\|C(\mathbf{z}_c)\|_1 + \lambda_3\|\mathbf{z}_s\|_1\}\\ &\text{s.t. } \mathbf{z} = \mathbf{z}_s + \mathbf{z}_w + \mathbf{z}_c + \mathbf{e}; \quad \|\mathbf{e}\|_\infty \le \delta\\ &\qquad \mathbf{u} = W(\mathbf{z}_w); \quad \mathbf{v} = C(\mathbf{z}_c).\end{aligned} \tag{4}$$

Equation 4 can now be easily converted into an unconstrained problem under the ALM framework [58], [59] as (ignoring $\|\mathbf{e}\|_\infty \le \delta$ at the moment)

$$\begin{aligned}\min\ \Big\{&\lambda_1\|\mathbf{u}\|_1 + \lambda_2\|\mathbf{v}\|_1 + \lambda_3\|\mathbf{z}_s\|_1 + \langle\mathbf{y}_1, \mathbf{u} - W(\mathbf{z}_w)\rangle\\ &+\frac{\eta_w\mu}{2}\|\mathbf{u} - W(\mathbf{z}_w)\|_2^2 + \langle\mathbf{y}_2, \mathbf{v} - C(\mathbf{z}_c)\rangle\\ &+\frac{\eta_c\mu}{2}\|\mathbf{v} - C(\mathbf{z}_c)\|_2^2 + \langle\mathbf{y}_3, \mathbf{z} - (\mathbf{z}_s + \mathbf{z}_w + \mathbf{z}_c + \mathbf{e})\rangle\\ &+\frac{\eta_e\mu}{2}\|\mathbf{z} - (\mathbf{z}_s + \mathbf{z}_w + \mathbf{z}_c + \mathbf{e})\|_2^2\Big\}\end{aligned} \tag{5}$$

where $\mathbf{y}_1$, $\mathbf{y}_2$, and $\mathbf{y}_3$ are the Lagrange multipliers in vector format; $\langle .,. \rangle$ denotes the inner product of two vectors; and η_w, η_c, and η_e are the penalty parameters. Now, we can propose the following algorithm for solving (5).

1) Fixing $\mathbf{u}$, $\mathbf{v}$, $\mathbf{z}_w$, $\mathbf{z}_c$, and $\mathbf{e}$, solve for $\mathbf{z}_s$ by

$$\min_{\mathbf{z}_s} \lambda_3\|\mathbf{z}_s\|_1 + \frac{\eta_e\mu}{2}\left\|\mathbf{z}_s - (\mathbf{z} - \mathbf{z}_w - \mathbf{z}_c - \mathbf{e}) - \frac{1}{\eta_e\mu}\mathbf{y}_3\right\|_2^2$$

which has a closed-form solution.

2) Fixing $\mathbf{u}$, $\mathbf{v}$, $\mathbf{z}_s$, $\mathbf{z}_c$, and $\mathbf{e}$, solve for $\mathbf{z}_w$ by

$$\begin{aligned}\min_{\mathbf{z}_w}\ &-\langle\mathbf{y}_1, W(\mathbf{z}_w)\rangle + \frac{\eta_w\mu}{2}\|\mathbf{u} - W(\mathbf{z}_w)\|_2^2 - \langle\mathbf{y}_3, \mathbf{z}_w\rangle\\ &+\frac{\eta_e\mu}{2}\|\mathbf{z}_w - (\mathbf{z} - \mathbf{z}_s - \mathbf{z}_c - \mathbf{e})\|_2^2\end{aligned}$$

which is a least squares problem with a closed-form solution.

3) Fixing $\mathbf{u}$, $\mathbf{v}$, $\mathbf{z}_s$, $\mathbf{z}_w$, and $\mathbf{e}$, solve for $\mathbf{z}_c$ by

$$\min_{\mathbf{z}_c} \; -\langle \mathbf{y}_2, C(\mathbf{z}_c)\rangle + \frac{\eta_c\mu}{2}\|\mathbf{v} - C(\mathbf{z}_c)\|_2^2 - \langle \mathbf{y}_3, \mathbf{z}_c\rangle + \frac{\eta_e\mu}{2}\|\mathbf{z}_c - (\mathbf{z} - \mathbf{z}_s - \mathbf{z}_w - \mathbf{e})\|_2^2$$

which is a least square problem with a closed-form solution.

4) Fixing $\mathbf{u}$, $\mathbf{v}$, $\mathbf{z}_w$, $\mathbf{z}_c$, and $\mathbf{z}_s$, solve for $\mathbf{e}$ by

$$\min_{\hat{\mathbf{e}}} -\langle \mathbf{y}_3, \hat{\mathbf{e}}\rangle + \frac{\eta_e\mu}{2}\|\hat{\mathbf{e}} - (\mathbf{z} - \mathbf{z}_s - \mathbf{z}_w - \mathbf{x}_c)\|_2^2$$

which has a closed-form solution. To obtain $\mathbf{e}$, we project $\hat{\mathbf{e}}$ to the convex set $\|\mathbf{e}\|_\infty \le \delta$, which is a straightforward thresholding on entries in $\hat{\mathbf{e}}$.

5) Fixing $\mathbf{v}$, $\mathbf{z}_w$, $\mathbf{z}_c$, $\mathbf{z}_s$, and $\mathbf{e}$, solve for $\mathbf{u}$ by

$$\min_{\mathbf{u}} \lambda_1\|\mathbf{u}\|_1 + \frac{\eta_w\mu}{2}\left\|\mathbf{u} - W(\mathbf{z}_w) + \frac{1}{\eta_w\mu}\mathbf{y}_1\right\|_2^2$$

which has a closed-form solution.

6) Fixing $\mathbf{u}$, $\mathbf{z}_w$, $\mathbf{z}_c$, $\mathbf{z}_s$, and $\mathbf{e}$, solve for $\mathbf{v}$ by

$$\min_{\mathbf{v}} \lambda_2\|\mathbf{v}\|_1 + \frac{\eta_c\mu}{2}\left\|\mathbf{v} - C(\mathbf{z}_c) + \frac{1}{\eta_c\mu}\mathbf{y}_2\right\|_2^2$$

which has a closed-form solution.

7) Updating $\mathbf{y}_1$, $\mathbf{y}_2$, and $\mathbf{y}_3$ by

$$\begin{aligned}\mathbf{y}_1 &\leftarrow \mathbf{y}_1 + \mu(\mathbf{u} - W(\mathbf{z}_w))\\ \mathbf{y}_2 &\leftarrow \mathbf{y}_2 + \mu(\mathbf{v} - C(\mathbf{z}_c))\\ \mathbf{y}_3 &\leftarrow \mathbf{y}_2 + \mu(\mathbf{z} - \mathbf{z}_s - \mathbf{z}_w - \mathbf{z}_c - \mathbf{e}).\end{aligned}$$

8) Some update for μ as an adaptive way.

The above algorithm can be guaranteed to be convergent (see [58], [59]). In fact, the idea behind the above algorithm is very simple. In each iteration, by projecting the residual to each domain, the algorithm keeps large magnitude components and sets small magnitude coefficients to zero, and this is in accordance with L1-norm minimization. As the iterations proceed, the residual is getting smaller and smaller. The different types of structured components can be reconstructed in the wavelet domain, the curvelet domain, and the spatial domain, respectively, in each iteration. When either a certain number of iterations are reached or the algorithm converges to a minimum, the algorithm terminates. As a result, the point-like component $\mathbf{z}_s$, the smooth component $\mathbf{z}_w$, and the edge component $\mathbf{z}_c$ can be eventually separated. Moreover, we can also regard the above algorithm as a denoising method and the bonus of the separation is the cleaned image that is the summation of these subimages.

B. MSR for SR

To reconstruct a high-resolution image, we need a good understanding of the whole process of capturing LR images. As we know that onboard panchromatic cameras can capture only the reflected light from the surface of the earth, the reflected light will suffer from the air turbulence and diffusion from the optical lens system. Therefore, the procedure for obtaining observed LR images on board can be modeled by

$$\mathbf{g}_i = \mathrm{DH}_i M_i \mathbf{z} + \mathbf{n}_i, \quad i = 1, \ldots, K \tag{6}$$

where $\mathbf{z}$ (the original SR image) can be regarded as an ideal set of pixels obtained by sampling a continuous scene at high resolution, K is the number of LR images, $\mathbf{g}_i$ is the ith observed LR image, M_i is a warping matrix representing the step of warping the original HR image to the grid of the observed image, H_i is a blur matrix representing all the inevitable blurriness caused by imperfected optical systems within onboard cameras, D is a subsampling matrix representing the degenerate step of the original HR image to an observed LR image, and $\mathbf{n}_i$ represents the inevitable noise contamination with an additive noise of Gaussian distribution with zero mean and variance δ_i.

Choosing the jth LR image as a reference image, then $\mathbf{g}_j = \mathrm{DH}_j\mathbf{z}$ or $\mathbf{g}_j = \mathrm{DH}_j M_j\mathbf{z}$, where M_j is an identity matrix in this case. In general, the mapping matrix M_i must be estimated initially from the LR images between $\mathbf{g}_i$ and the reference LR image $\mathbf{g}_j$, and the blurring matrix H_i can be approximated by a Gaussian blur kernel whose size is estimated from sharp points or edges in the images. Let $N \times N$ be the number of pixels in the original SR image and d be the decimating rate (subsampling rate). $\mathbf{g}_i$, $\mathbf{z}$, and $\mathbf{n}_i$ are all symbolized as lexicographically ordered vectors of lengths N^2/d^2, N^2, and N^2/d^2, respectively. Thus, H_i and M_i are $N^2 \times N^2$ matrices. The size of D is $N^2/d^2 \times N^2$.

From (6), we can see that the imaging procedure is not invertible, and the SR technique represents a typical ill-conditioned problem. This implies that there is no unique solution. A well-posed problem is preferred, and restoring the original SR image $\mathbf{z}$ is desired.

Here, we propose the following solution for the SR problem with the framework of MSR:

$$\min \{\lambda_1\|W\mathbf{z}_w\|_1 + \lambda_2\|C(\mathbf{z}_c)\|_1 + \lambda_3\|(\mathbf{z}_s)\|_1\} \quad \text{s.t. } \|\mathbf{g}_i - \mathrm{DH}_i M_i\mathbf{z}\|_2 \le \delta_i, \quad i = 1, \ldots, K \tag{7}$$

where W and C are the same as defined in the above section and δ_i denotes the variance of the Gaussian noise in the ith image. In (7), λ_1–λ_3 denote different weights for the contribution from different transformations. Since the proposed SR method is an MSR-based SR for remote sensing images, the method will be shorted as MSR-SR for simplicity.

Here, we use the iterative shrinkage-thresholding algorithm [60] for solving the SR problem as follows.

1) Initialization:
   a) Choose parameters: the soft threshold τ (this can be set by 1 for most normalized circumstances) and the stopping threshold δ (this can be set by noise level).
   b) The total number of iterations $l = \mathrm{floor}(\tau/\delta)$.
   c) Set the weighting parameters $\lambda_1 \cdots \lambda_3$.
   d) $\mathbf{z}_s = 0$; $\mathbf{z}_w = 0$; $\mathbf{z}_c = 0$; $\mathbf{z} = \mathbf{z}_s + \mathbf{z}_w + \mathbf{z}_c$.
2) Within l iterations:
   a) Reconstructing the sparse components $\mathbf{z}_s$ for each LR image g_i.
      i) Calculate the residual

$$\mathbf{r} = \sum_{i=1}^{K}\{\mathrm{DH}_i M_i\mathbf{z} - \mathbf{g}_i\}.$$

ii) Calculate the gradient $\mathbf{d} = \sum_{i=1}^{K} \{M_i^* H_i^* D^* \mathbf{r}\}$.
iii) Update $\mathbf{z}_s = \mathbf{z}_s - \mathbf{d}$.
iv) Soft threshold ($\mathbf{z}_s$), i.e., set any absolute values below (τ/λ_3) to zero and update $\mathbf{z}_s$.

b) Reconstructing the smooth components $\mathbf{z}_w$ for each LR image $\mathbf{g}_i$.
i) Calculate the residual

$$\mathbf{r} = \sum_{i=1}^{K} \{\mathrm{DH}_i M_i \mathbf{z} - \mathbf{g}_i\}.$$

ii) Calculate the gradient $\mathbf{d} = \sum_{i=1}^{K} \{M_i^* H_i^* D^* \mathbf{r}\}$.
iii) Update $\mathbf{z}_w = \mathbf{z}_w - \mathbf{d}$.
iv) Calculate the DWT coefficients for $\mathbf{z}_w$.
v) Soft threshold the wavelet coefficients of $W\mathbf{z}_w$, i.e., set any values below $\frac{\tau}{\lambda_1}$ to zero and update $W\mathbf{z}_w$.
vi) Calculate the inverse wavelet transform for the wavelet coefficients, and then update $\mathbf{z}_w$.

c) Reconstructing the edge structures $\mathbf{z}_c$ for each LR image $\mathbf{g}_i$.
i) Calculate the residual

$$\mathbf{r} = \sum_{i=1}^{K} \{\mathrm{DH}_i M_i \mathbf{z} - \mathbf{g}_i\}.$$

ii) Calculate the gradient $\mathbf{d} = \sum_{i=1}^{K} \{M_i^* H_i^* D^* \mathbf{r}\}$.
iii) Update $\mathbf{z}_c = \mathbf{z}_c - \mathbf{d}$.
iv) Calculate the curvelet coefficients for $\mathbf{z}_c$.
v) Soft threshold the wavelet coefficients of $C\mathbf{z}_c$, i.e., set any values below (τ/λ_2) to zero and update $C\mathbf{z}_c$.
vi) Calculate the inverse curvelet transform for the curvelet coefficients, and then update $\mathbf{z}_c$.

d) $\tau = \tau - \delta$.

3) Reconstructed SR image $\mathbf{z} = \mathbf{z}_s + \mathbf{z}_w + \mathbf{z}_c$.

In the above pseudocode, $*$ denotes the back projection step of (6) when calculating the gradient. For example, M_i^* denotes the warping matrix representing the warp of the observed LR image $\mathbf{g}_i$, and H_i^* is the conjugate transpose of the blur matrix H_i. Note that D_i^* represents an interpolating step that can be implemented with the traditional bilinear interpolation method. The selection of those weighting parameters is a bit tricky, the more weight put on a particular type of structure, for example, we say point components, then the threshold (τ/λ_1) of the shrinking operator would be very small. That means few structures or energy will be left for the following two steps of resolving smooth components and the edge-like components. On the other hand, if we set a small λ_1, then the threshold (τ/λ_1) of the shrinking operator would be very big, which will lead to squeezing most of structures or energy from $\mathbf{z}_s$ to those smooth components and edge-like components.

As for the MSR-based classification, the residual will be projected to each representative domain, respectively, and is getting smaller and smaller in each iteration when solving the SR problem. As iterations progress, the different types of structured components can be reconstructed in the spatial domain, the wavelet domain, or the curvelet domain. Either a certain number of iterations are reached or the algorithm converges to a minimum and the algorithm terminates. The super-resolved image $\mathbf{z}$ can be reconstructed simply by adding $\mathbf{z}_s$, $\mathbf{z}_w$ and $\mathbf{z}_c$.

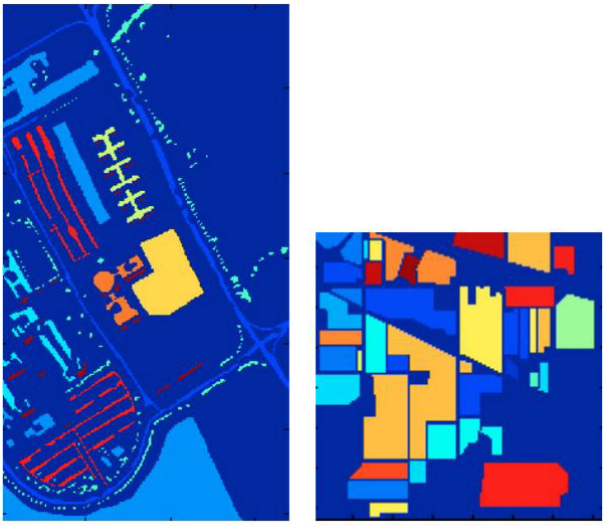
III. Experimental Results

To test the proposed framework of MSR, it is applied to solve the ill-conditioned problem: classification and SR introduced above, respectively.

A. MSR for Classification

In general, point-like components in hyperspectral images do not help in terms of adding spatial constraints in classification, and therefore, we ignore the resolved point-like components and utilize the smooth components and the edge components only. We simply augment the original hyperpspectral image by attaching the separated smooth subimage and the curvelet resolved subimage as a spatial constraint to the original data cube and use this expended data cube in traditional classification methods without any changes in terms of input parameters. In this paper, three classification methods are used to verify the improvement of using the expanded hyperspectral data cube. One is the widely used spectral classification method support vector machine (SVM) in remote sensing community. The other spectral-based classification method used in this paper is logistic regression via variable splitting and augmented Lagrangian (LORSAL) algorithm [61], which mainly implements a multinomial logistic regression classifier. The spectral–spatial classification method for hyperspectral image data proposed in [37] is abbreviated as MAP marginal (MPM)-loopy belief propagation (LBP). This method serves as an engine in the context of active learning to exploit both spatial and spectral information simultaneously. Li *et al.* [37] used the LBP algorithm to estimate the MPM solution. The reason we use those two methods is that LORSAL represents a typical classification method that considers *no* spatial constraint, while MPM-LBP is a well-performed spatial-spectral-based classification method. Our intention is to test whether the data augmentation by MSR can improve the classification accuracy for both spectral-*only*-based classification methods and spectral–spatial-based classification methods. Again, we are not intent to propose a new classification method. On the contrary, our contribution is the MSR framework that augments the original data cube and helps improve the classification accuracy for most of any classification algorithms at hand.

As far as the testing data are concerned, we use two popular data cubes to test the above algorithm. One is the hyperspectral data set that was collected by the ROSIS optical sensor over the urban area of the University of Pavia, Italy. The flight was operated by the Deutschen Zentrum for Luftund Raumfahrt, and managed and sponsored by the European Union. The image size is 610×340 pixels, and its spatial resolution is 1.3 m/pixel. The number of data channels captured by the sensor is 115 with wavelength range from 0.43 to 0.86 μm. The ground truth of the data set can be seen in Fig. 1(a). As can



(a) PaviaUni (b) Indian Pines

Fig. 1. Ground truth of PaviaUni data and Indian Pines' AVIRIS data.

be seen in Fig. 1(a), there are nine mutually exclusive ground truth classes in the region. After removing the severely noise contaminated frames, there are 103 frames left in the PaviaUni data cube. We run the algorithm 103 times channel by channel and separate the whole data cube. Therefore, the original data cube is expanded into three times larger $610 \times 340 \times 412$ (in which 1–103 frames are the original data cube, 104–206 frames are the smooth components, 207–309 frames are the edge components, and 310–412 frames are the point components in the new expanded data cube). The original image of channel 100 can be found in Fig. 2(a). After the above segmentation step, the smooth components (channel 203 in the new expanded data cube), edge components (channel 306 in the new expanded data cube), and point components (channel 409 in the new expanded data cube) within channel 100 are separated and shown in Fig. 2(b)–(d), respectively.

The other data cube used in this paper was captured by the AVIRIS sensor over the Indian Pines region in Northwestern Indiana in 1992. This scene is of 145×145 pixels whose spatial resolution is 20 m. The number of data channels captured by the sensor is 220 with wavelength range from 0.4 to 2.5 μm. The ground truth of the data set can be seen in Fig. 1(b). As can be seen in Fig. 1(b), there are 16 mutually exclusive ground truth classes in the region. After removing the severely noise contaminated frames, there are 191 frames left in the AVIRIS data cube in our test. Again, we run the algorithm 191 times channel by channel and separate the whole data cube into three kinds of components. Therefore, the original data cube is expanded into three times larger $145 \times 145 \times 764$ (in which 1–191 frames are the original data cube, 192–382 frames are the smooth components, 383–573 frames are the edge components, and 574–764 frames are the point components in the new expanded AVIRIS data cube). The original image of channel 9 of Indian Pines' AVIRIS data can be found in Fig. 3(a) as an example. After the above segmentation step, the smooth components (channel 200 in the new expanded data cube), edge components (channel 391 in the new expanded data cube), and point components (channel 582 in the new expanded data cube) within channel 9 are separated and shown in Fig. 3(b)–(d), respectively.

TABLE I
CLASSIFICATION RESULTS OF SVM

	PaviaUni data cube		**AVIRIS data cube**	
Frames	Original 103	Expanded 309	Original 191	Expanded 573
Overall Accuracy (OA)	0.9309	**0.9511**	0.8568	**0.9256**
Average Accuracy (AA)	0.9032	**0.9281**	0.8492	**0.8993**
Kappa (k)	0.9079	**0.9350**	0.8358	**0.9151**

TABLE II
CLASSIFICATION RESULTS OF LORSAL

	PaviaUni data cube		**AVIRIS data cube**	
Frames	Original 103	Expanded 309	Original 191	Expanded 573
Overall Accuracy (OA)	0.9313	**0.9569**	0.8726	**0.9476**
Average Accuracy (AA)	0.9048	**0.9344**	0.8079	**0.9225**
Kappa (k)	0.9088	**0.9427**	0.8544	**0.9403**

TABLE III
CLASSIFICATION RESULTS OF MPM-LBP

	PaviaUni data cube		**AVIRIS data cube**	
Frames	Original 103	Expanded 309	Original 191	Expanded 573
Overall Accuracy (OA)	0.9781	**0.9823**	0.9555	**0.9709**
Average Accuracy (AA)	0.9598	**0.9683**	0.9032	**0.9452**
Kappa (k)	0.9709	**0.9765**	0.9493	**0.9669**

In the experiments, we use fixed parameter settings to concentrate more on the proposed framework and less on its parameter tuning for SVM, LORSA, and MPM-LBP. As for many other classification methods, we use those standard evaluation formulas overall accuracy (OA), average accuracy (AA), kappa k statistic to carry out the comparison. In the experiments of PaviaUni data, we randomly select 0.05% of each class in the ground truth. That is to say that there are 2138 training samples in the training set. Then the whole data set is used to evaluate the performance of the classification. Similarly, in the experiments of Indian Pines' AVIRIS data set, we randomly selected 1024 training samples, i.e., 10% of each class in the ground truth. The whole augmented Indian Pines' data set is used to evaluate the performance of the classification. As can be seen in Tables I–III, without changing anything but using the augmented data, the classification accuracy can be improved around 5%, especially, in terms of OA and kappa coefficient. As we can see from Tables I–III,

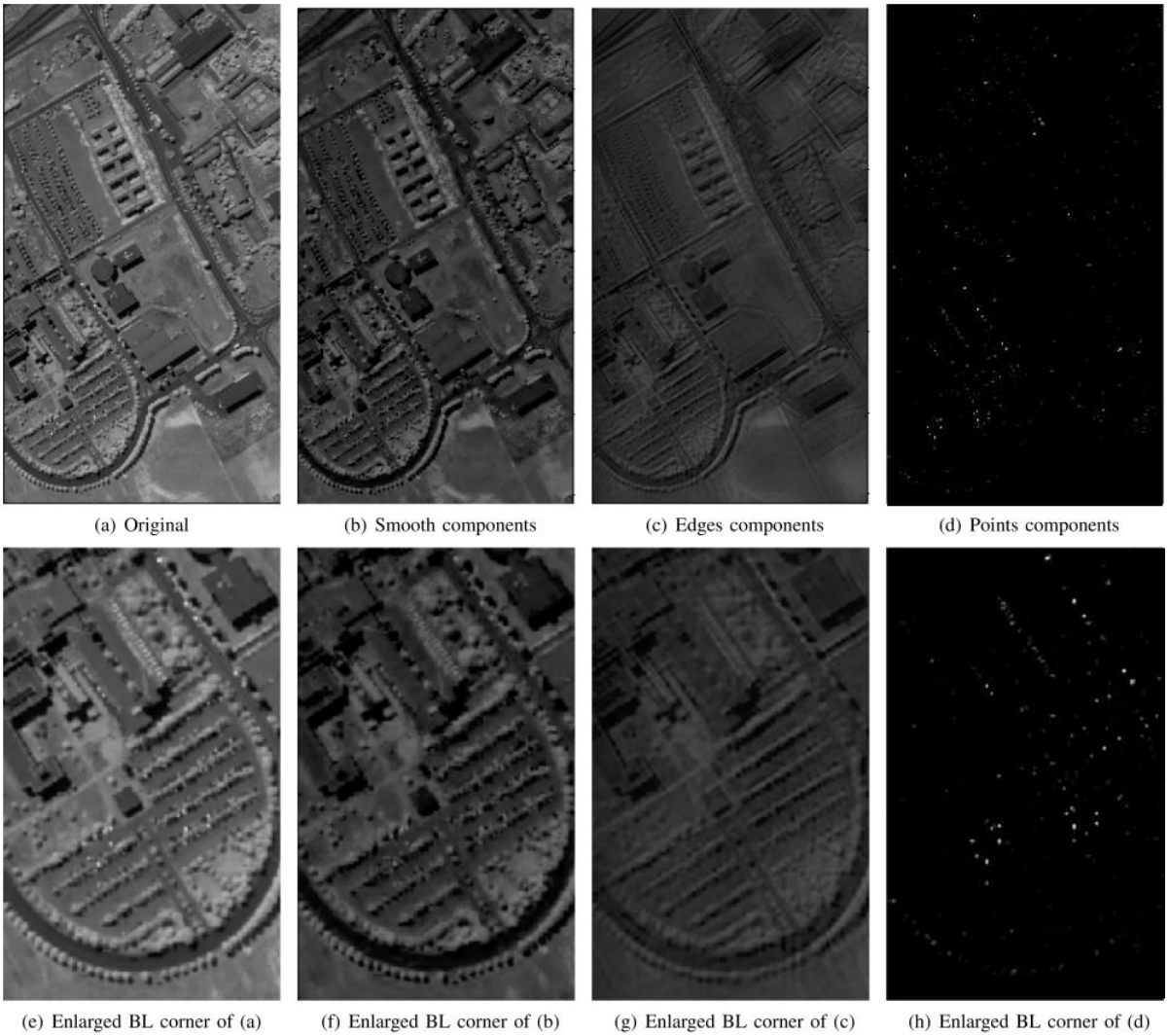


Fig. 2. Segmentation results of Pavia University data. (a) Original channel 100 image. (b) Smooth components within channel 100. (c) Edge components within channel 100. (d) Point components within channel 100. Based on (2), we know that those subfigures (a) $\approx$ (b) + (c) + (d). To help readers to have a better understanding, (e)–(h) enlarged bottom-left corner of those subimages are shown.

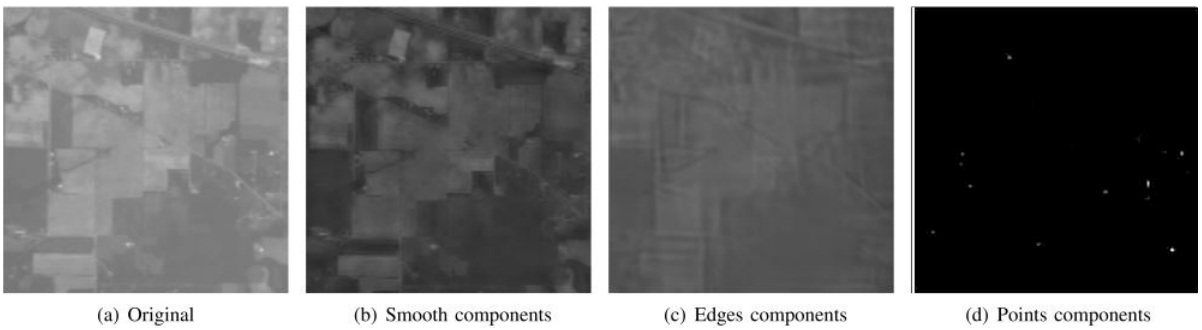
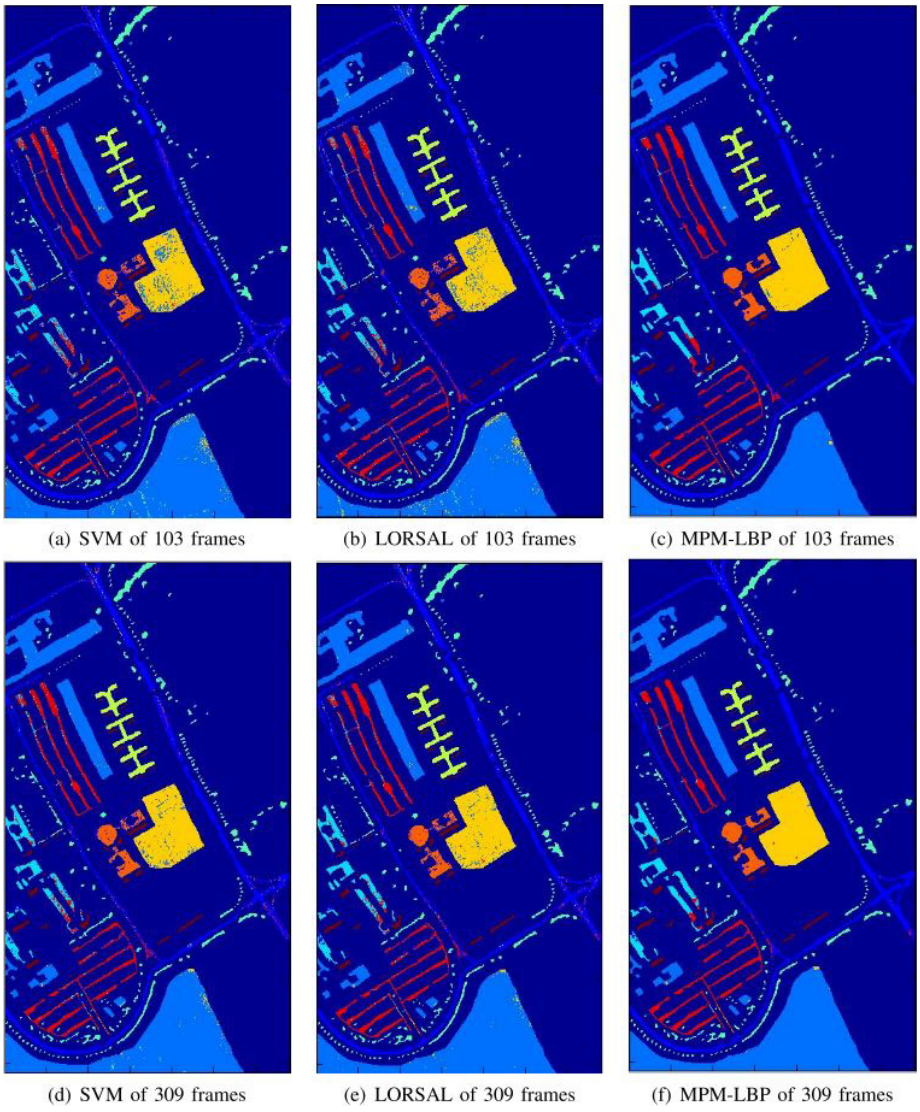


Fig. 3. Segmentation results of Indian Pines' AVIRIS data. (a) Original channel 9 image. (b) Smooth components within channel 9. (c) Edge components within channel 9. (d) Point components within channel 9. Based on (2), we know that those subfigures (a) $\approx$ (b) + (c) + (d).

the framework works much more effectively for spectral-only-based classification methods than those spatial–spectral joined classification methods. This is because that those spatial–spectral-based classification methods have already utilized the spatial correlation to some extent. Moreover, we may note that the performance of the framework works more efficient for Indian Pines' AVIRIS data than PaviaUni data. This is mainly because that the number of training samples for PaviaUni data is quite big, and the classification results from the original data set are already good enough, so the improvement is limited. To help readers have a direct understanding of the performance of the proposed frame work, we list the experimental results of PaviaUni data and Indian Pines' AVIRIS data in Figs. 4 and 5, respectively. The first rows of Figs. 4 and 5 from left to



(a) SVM of 103 frames (b) LORSAL of 103 frames (c) MPM-LBP of 103 frames

(d) SVM of 309 frames (e) LORSAL of 309 frames (f) MPM-LBP of 309 frames

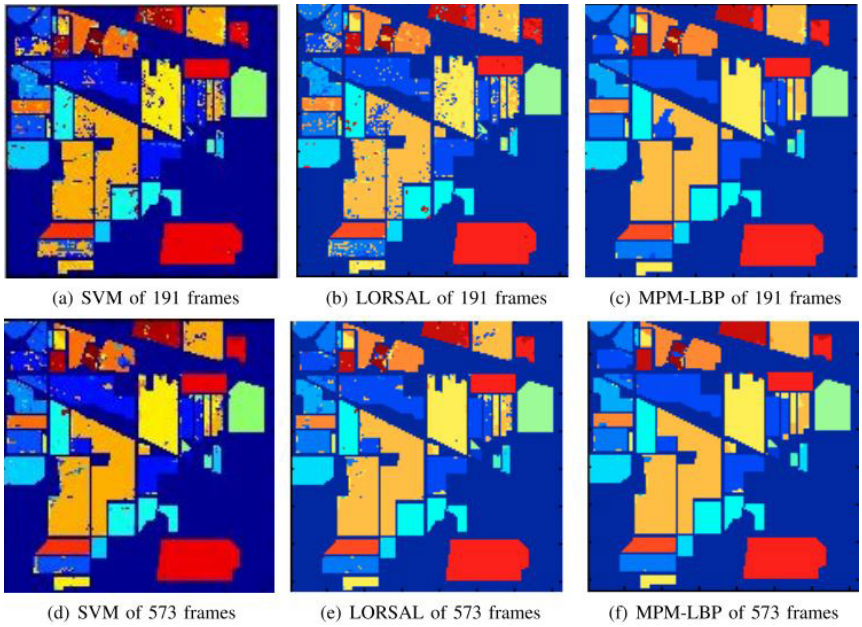
Fig. 4. Classification results for PaviaUni data. The first row from left to right shows the classification results of SVM, LORSAL, and MPM-LBP with the original PaviaUni data. The bottom row from left to right shows the improved classification results with the augmented data.

right show the classification results of SVM, LORSAL, and MPM-LBP with the original data set. The bottom rows from left to right show the classification results with the augmented data. By comparing those classification results of the original data and the augmented data, we can see that improved classification results can be achieved.

B. MSR for Super-Resolution

In this section, we test and compare our method using images of size 256×256 pixels from the panchromatic channel of Landsat7 (spatial resolution of 15 m) captured on four different dates (September 13, September 29, December 2, and December 18, 2000) within the Canberra region. The means of those four aligned and bilinear interpolated images have been shown in Fig. 6(a), in which the image of September 29 is chosen as the reference. In general, data captured with minimum time gap are good for SR; however, the repeating cycle of Landsat7 is 16 days and clouds appeared between October and November in 2000, so the selection of those four images is a compromise under these circumstances. Therefore, changes of objects on the ground over time will be ignored. The illumination of the four data groups varies over a large range because of the different atmospheric conditions. Thus, we rescaled the four images to [0, 1], and performed histogram matching among them to reduce the effects due to different atmospheric and illumination conditions. In this test, we set the SR factor as 2, which means that the spatial resolution of the super-resolved image will achieve a spatial resolution of 15 m.

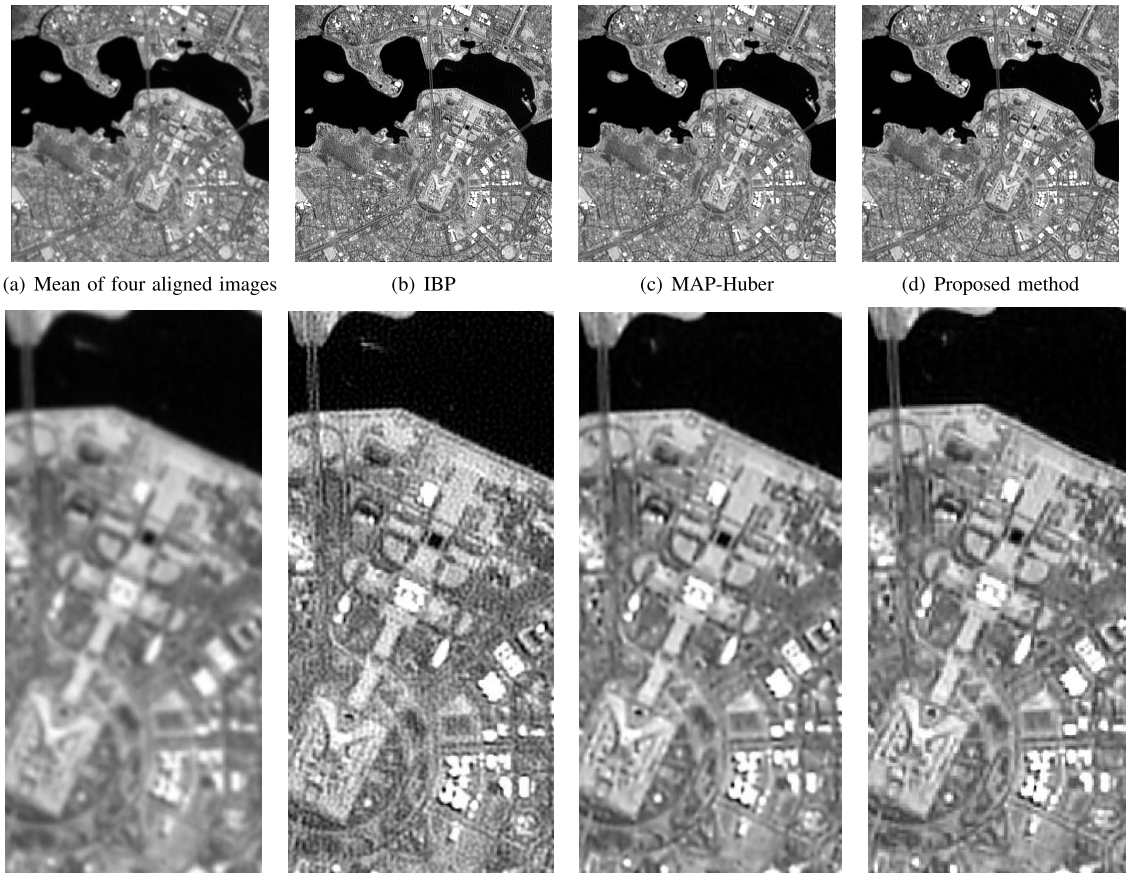
From (7), we can see that matrices M_i, H_i, and D need to be worked out beforehand. In practice, there is no need to construct real matrices for M_i, H_i, and D because of memory and computational limitation. Matrices D and D^* can be replaced by a downsampling and a bilinear interpolation operator ($d = 2$). In this test, matrices H_i and H_i^* can be approximated by a Gaussian blur kernel of size 5, which is roughly estimated from sharp edges in those LR images. Similarly, the elastic image registration method [62] is used to calculate the warps between LR images rather than to construct



(a) SVM of 191 frames (b) LORSAL of 191 frames (c) MPM-LBP of 191 frames

(d) SVM of 573 frames (e) LORSAL of 573 frames (f) MPM-LBP of 573 frames

Fig. 5. Classification results for Indian Pines' AVIRIS data. The first row from left to right shows the classification results of SVM, LORSAL, and MPM-LBP with the original Indian Pines' AVIRIS data. The bottom row from left to right shows the improved classification results with the augmented data.

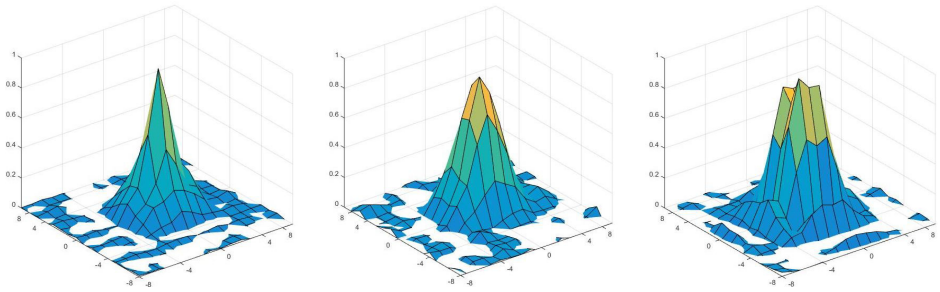


(a) Mean of four aligned images (b) IBP (c) MAP-Huber (d) Proposed method

(e) Local of the mean image (f) Local of IBP (g) Local of MAP-Huber (h) Local of our method

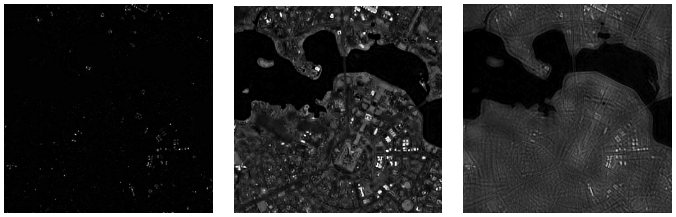
Fig. 6. Experimental results on real Landsat7 data. (a) Mean of those four aligned Landsat7 images. (b) IBP. (c) MAP-Huber. (d) Our method. (e) Local area of the aligned mean image in (a). (f) Local area of the IBP reconstructed image in (b). (g) Local area of MAP-Huber reconstructed image in (c). (h) Local area of the IBP reconstructed image in (d).

real warping matrices M_i and M_i^*. Common rigid registration methods may fail to handle irregular local warps and transformation caused by platform vibration and air turbulence in LR remote sensing images. Therefore, we use an elastic registration method to model the warps and to improve the reconstruction image quality.



(a) (b) (c)

Fig. 7. PSFs and the variance of those equivalent Gaussian σ^2 kernel for each SR method. (a) IBP, $\sigma^2 = 2.16$. (b) Huber, $\sigma^2 = 2.22$. (c) Proposed method, $\sigma^2 = 2.72$.



(a) (b) (c)

Fig. 8. Byproducts of MSR-SR in Fig. 6(d). (a) Sparse components $\mathbf{z}_s$ in the reconstructed image. (b) Smooth components $\mathbf{z}_w$. (c) Edge components. Therefore, the reconstructed image $\mathbf{z} = \mathbf{z}_s + \mathbf{z}_w + \mathbf{z}_c$.

In order to evaluate the performance of the proposed method, we compare it with other popular methods: IBP [46] reconstruction and MAP-Huber [48]. The original sizes of the LR images are 256×256 pixels. To carry out a better comparison, we also display the central part of the reconstructed images in an enlarged scale. The experimental results are shown in Fig. 6. The mean of four aligned images is in Fig. 6(a). We can see that it is very blurred and most of details are lost. Although all SR reconstructed images from Fig. 6(b)–(d) are much sharper and clearer than the mean of four aligned images, the proposed method provides the best reconstruction under a visual comparison. For example, as shown in the right corner of Fig. 6(e)–(h), there is a three-way junction or "T" junction. We cannot see it clearly in Fig. 6(e)–(g), but the three-way junction can be easily recognized in Fig. 6(h). Moreover, the bridge in the top left of the enlarged version in Fig. 6(h) provides a much more decent appearance than other methods.

In general, it is difficult to carry out a quantitative comparison for those super-resolved images by different methods with real data. Here, point spread functions (PSFs) are computed by assuming that Fig. 6(a) is a blurred version of those super-resolved images [Fig. 6(b)–(d)]. In this case, a wider PSF indicates better resolution, since the blurred image is the same for all super-resolved images. The PSFs are computed using a Wiener filter, and the results are shown in Fig. 7, in which the PSF of our method shows the largest half peak width (best resolution). To quantify this further, circularly symmetric Gaussians are fitted to the PSFs based on equivalent volume, and comparison of the resulting variance of the equivalent Gaussian σ^2 also indicates that our method achieves the best resolution $\sigma^2 = 2.72$.

Interestingly, as a bonus from MSR-SR, full version of the sparse components $\mathbf{z}_s$, the smooth components $\mathbf{z}_w$, and the edge components of $\mathbf{z}_c$ can be reconstructed separately and are shown in Fig. 8(a)–(c), respectively. Those resolved subimages by MSR-SR may help multispectral/hyperspectral-based classification methods by easily combining spatial constraint in a certain way.

IV. Discussion

Ideally, what we wish is that any components in an image can show sparse representation only in a particular domain. In other words, a component can show up sparsity only in a particular domain. As shown in Fig. 8(b) and (c), however, there are overlapped regions between the wavelet represented components and the curvelet represented components. This is mainly because those two transformations (wavelet and curvelet) are not orthogonal. So here is an open problem, "incoherence"-like property, which is widely used in the CS community denoting the two completely uncorrelated transform domains [2], is preferred when selecting the combination of transformations and the best combination needs further investigation.

The initialization of those methods is not sensitive, and the initial Z_s, Z_s, and Z_w can be empty images. The proposed method will converge to a separation because it is a convex optimization problem. As far as the computational efficiency of the framework is concerned, MSR-SR is much slower than the classification case. This is mainly because there are extra procedures: such as interpolation, warping, and blurring in each iteration. Moreover, the elastic registration is built in with MSR-SR for solving irregular warping among those LR images, which definitely decreases the computational

efficiency. In our test, for example, reconstructing a 512×512 super-resolved image from four 256×256 images using MSR-SR takes around 5 min. The computer is a 3.20-GHz Windows machine with 8-GB RAM. All routines were created under MATLAB.

V. Conclusion

It is very difficult to find a common sparse representation for remote sensing images because of complicated ground features. In this paper, we proposed a generalized framework of MSRs for remote sensing images. By assuming that a remote sensing image consists of subimages of smooth, edges, and point-like components, a group of domain transformations is used to sparsely represent each subimage, since each domain transformation method is capable of representing only a particular kind of ground object or texture. The performance of the framework was tested by solving two typical ill-conditioned problems: classification and SR. Our experimental results show that not only the new framework of MSR can improve classification accuracy but also can construct a much better high-resolution image than other common SR methods. The proposed framework MSR is a competitive candidate prior for solving other remote sensing images-related MAP problems.

Acknowledgment

The authors would like to thank RICE university for the CurveLab Toolbox [22] and the wavelet toolbox [63].

References

[1] E. J. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inv. Prob.*, vol. 23, no. 3, p. 969, 2007.

[2] E. J. Candès and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Jan. 2008.

[3] E. J. Candès and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[4] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[5] M. Wakin, J. Laska, M. Duarte, and D. Baron, "An architecture for compressive imaging," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Atlanta, GA, USA, Oct. 2006, pp. 1273–1276.

[6] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.

[7] J. Zhang, T. Xiong, T. Tran, S. Chin, and R. Etienne-Cummings, "Compact all-CMOS spatiotemporal compressive sensing video camera with pixel-wise coded exposure," *Opt. Exp.*, vol. 24, no. 8, pp. 9013–9024, Apr. 2016.

[8] M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub—Nyquist rates," *IET Circuits, Devices Syst.*, vol. 5, no. 1, Jan. 2011.

[9] X. Gao, F. Jiang, S. Liu, W. Che, X. Fan, and D. Zhao, "Hierarchical frame based spatial–temporal recovery for video compressive sensing coding," *Neurocomputing*, vol. 174, pp. 404–412, Jan. 2016.

[10] N. Mohan, I. Stojanovic, W. C. Karl, B. E. A. Saleh, and M. C. Teich, "Compressed sensing in optical coherence tomography," *Proc. SPIE*, vol. 7570, p. 7570, Feb. 2010.

[11] J. Hunt *et al.*, "Metamaterial apertures for computational imaging," *Science*, vol. 339, no. 6117, pp. 310–313, 2012.

[12] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 243–250, Jun. 1996.

[13] J. L. Starck, E. J. Candès, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Image Process.*, vol. 11, no. 6, pp. 670–684, Jun. 2002.

[14] F. Li, T. J. Cornwell, and F. de Hoog, "The application of compressive sampling to radio astronomy—I. Deconvolution," *Astron. Astrophys.*, vol. 528, Apr. 2011, Art. no. A31.

[15] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[16] R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "Sparsity averaging reweighted analysis (SARA): A novel algorithm for radio-interferometric imaging," *Mon. Notices R. Astron. Soc.*, vol. 426, no. 2, pp. 1223–1234, 2012.

[17] F. Li, X. Jia, D. Fraser, and A. Lambert, "Super resolution for remote sensing images based on a universal hidden Markov tree model," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1270–1278, Mar. 2010.

[18] C. Huo, R. Zhang, and D. Yin, "An adaptive sparse representation for remote sensing image based on combination of wavelet and adaptive directional filter," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, Mar. 2010, pp. 990–993.

[19] Z. Zhihui, W. Bo, and S. Kang, "Single remote sensing image super-resolution and denoising via sparse representation," in *Proc. Int. Workshop Multi-Platform/Multi-Sensor Remote Sens. Mapping*, Jan. 2011, pp. 1–5.

[20] C. K. Chui, *An Introduction to Wavelets*. San Diego, CA, USA: Academic, 1992.

[21] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA, USA: SIAM, 1992.

[22] E. J. Candès, L. Demanet, D. Donoho, and X. Ying, "Fast discrete curvelet transforms," *Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861–899, Sep. 2006.

[23] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.

[24] P. Chen and D. Suter, "Shift-invariant wavelet denoising using inter-scale dependency," in *Proc. Int. Conf. Image Process. (ICIP)*, vol. 2. Oct. 2004, pp. 1005–1008.

[25] I. Hussain and H. Yin, "A novel wavelet thresholding method for adaptive image denoising," in *Proc. Int. Symp. Commun., Control Signal Process.*, Mar. 2008, pp. 1252–1256.

[26] M. Belge and E. Miller, "Wavelet domain image restoration using edge preserving prior models," in *Proc. IEEE Int. Conf. Image Process.*, Mar. 1998, pp. 103–107.

[27] M. Belge, M. E. Kilmer, and E. L. Miller, "Wavelet domain image restoration with adaptive edge-preserving regularization," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 597–608, Apr. 2000.

[28] A. Temizel and T. Vlachos, "Wavelet domain image resolution enhancement using cycle-spinning," *Electron. Lett.*, vol. 41, no. 3, pp. 119–121, Feb. 2005.

[29] X. Jia and J. A. Richards, "Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 538–542, Jan. 1999.

[30] X. Jia, B.-C. Kuo, and M. M. Crawford, "Feature mining for hyperspectral image classification," *Proc. IEEE*, vol. 101, no. 3, pp. 676–697, Mar. 2013.

[31] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis*. Berlin Germany: springer-Verlag, 1996.

[32] A. J. Brown, "Spectral curve fitting for automatic hyperspectral data analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1601–1608, Jun. 2006.

[33] A. J. Brown, B. Sutter, and S. Dunagan, "The MARTE VNIR imaging spectrometer experiment: Design and analysis," *Astrobiology*, vol. 8, no. 5, pp. 1001–1011, 2008.

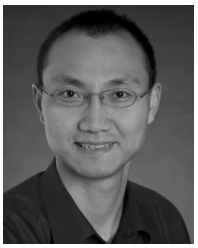
[34] X. Jia and J. A. Richards, "Cluster-space representation for hyperspectral data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 3, pp. 593–598, Mar. 2002.

[35] J. A. Gualtieri and R. F. Cromp, "Support vector machines for hyperspectral remote sensing classification," *Proc. SPIE*, vol. 3584, p. 221, Oct. 1998.

[36] C.-H. Li, H.-S. Chu, B.-C. Kuo, and C.-T. Lin, "Hyperspectral image classification using spectral and spatial information based linear discriminant analysis," in *Proc. IEEE Int. Symp. Geosci. Remote Sens.*, Jul. 2011, pp. 1716–1719.

[37] J. Li, M. Bioucas-Dias, and A. Plaza, "Spectral–spatial classification of hyperspectral data using loopy belief propagation and active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 844–856, Feb. 2013.

[38] J. Yin, Y. Wang, and J. Hu, "A new dimensionality reduction algorithm for hyperspectral image using evolutionary strategy," *IEEE Trans. Ind. Informat.*, vol. 8, no. 4, pp. 935–943, Nov. 2012.
[39] J. Yin, C. Gao, and X. Jia, "Using hurst and Lyapunov exponent for hyperspectral image feature extraction," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 705–709, Jul. 2012.
[40] K. Kavitha, S. Arivazhagan, and N. Kayalvizhi, "Wavelet based spatial—Spectral hyperspectral image classification technique using support vector machines," in *Proc. Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, Jul. 2010, pp. 1–6.
[41] C. Chen, W. Li, E. W. Tramel, M. Cui, S. Prasad, and J. E. Fowler, "Spectral–spatial preprocessing using multihypothesis prediction for noise-robust hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1047–1059, Apr. 2014.
[42] L. Don, "A basis for estimating digital camera parameters," *Photogram. Eng. Remote Sens.*, vol. 70, no. 3, pp. 297–300, 2004.
[43] R. Y. Tsai and T. S. Huang, "Multi-frame image restoration and registration," *Adv. Comput. Vis. Image Process.*, vol. 1, no. 2, pp. 317–339, 1984.
[44] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.
[45] S. Borman and R. L. Stevenson, "Super-resolution from image sequences—A review," in *Proc. Midwest Symp. Circuits Syst.*, vol. 5. 1998, pp. 1–5.
[46] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP, Graph. Models Image Process.*, vol. 53, no. 3, pp. 231–239, May 1991.
[47] F. Li, X. Jia, and D. Fraser, "Universal HMT based super resolution for remote sensing images," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 333–336.
[48] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. Image Process.*, vol. 5, no. 6, pp. 996–1011, Jun. 1996.
[49] H. Shen, M. K. Ng, P. Li, and L. Zhang, "Super-resolution reconstruction algorithm to MODIS remote sensing images," *Comput. J.*, vol. 52, no. 1, pp. 90–100, 2009.
[50] P. J. Huber, *Robust Statistics*. Hoboken, NJ, USA: Wiley, 2004.
[51] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.
[52] M. E. Tipping and C. M. Bishop, "Bayesian image super resolution," U.S. Patent 7 106 914, Sep. 12, 2006.
[53] L. C. Pickup, D. P. Capel, S. J. Roberts, and A. Zisserman, "Bayesian image super-resolution, continued," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 19. 2006, pp. 1089–1096.
[54] J. Zhang, D. Zhao, and W. Gao, "Group-based sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3336–3351, Aug. 2014.
[55] S. Miura, Y. Kawamoto, S. Suzuki, T. Goto, S. Hirano, and M. Sakurai, "Image quality improvement for learning-based super-resolution with PCA," in *Proc. IEEE 1st Global Conf. Consum. Electron. (GCCE)*, Oct. 2012, pp. 572–573.
[56] C. M. Bishop, A. Blake, and B. Marthi, "Super-resolution enhancement of video," in *Proc. Artif. Intell. Statist.*, vol. 2. Key West, FL, USA, 2003.
[57] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure Appl. Math.*, vol. 67, no. 6, pp. 906–956, Jun. 2014.
[58] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Advances in Neural Information Processing Systems*. 2011, pp. 612–620.
[59] S. Tierney, J. Gao, and Y. Guo, "Subspace clustering for sequential data," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Ohio, USA, Jun. 2014, pp. 1019–1026.
[60] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
[61] J. Bioucas-Dias and M. Figueiredo, "Logistic regression via variable splitting and augmented Lagrangian tools," Inst. Superior Técnico, Lisbon, Portugal, Tech. Rep., 2009.
[62] F. Li, C. Li, L Tang, Y.Guo, "Elastic registration for airborne multispectral line scanners," *J. Appl. Remote Sens.*, vol. 8, no. 1, p. 083614, 2014.
[63] *Rice Wavelet Toolbox*. accessed on Nov. 11, 2016. [Online]. Available: http://dsp.rice.edu/software/rice-wavelet toolbox



Feng Li (M'07–SM'10) received the B.S.E.E. degree from Lanzhou Railway University, Lanzhou, China, in 1999, the M.Eng. degree from the China Academy of Space Technology, Beijing, China, in 2002, and the Ph.D. degree in electrical engineering from the University of New South Wales, Kensington, NSW, Australia, in 2009.

In the following several years, he has been invloved in astronomical image processing with the Commonwealth Science and Industrial Research Organisation, Sydney, Australia, and in remote sensing image processing with the Chinese Academy of Science. He is currently a Principal Investigator with the Qian Xuesen Laboratory of Space Technology, China. His current research interests include image registration, super resolution, and compressive sensing.

Lei Xin received the B.Sc. and M.Eng. degrees in instrumentation science and technology from Beihang University, Beijing, China, in 2010 and 2013, respectively.

She is currently a Research Associate with the Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology, Beijing, China. Her research interests include compressed sensing, super-resolution reconstruction, image processing, and particle sizing.

Yi Guo received the B.Eng. degree (Hons.) in instrumentation from the North China University of Technology, Beijing, China, in 1998, the M.Eng. degree in automatic control from Central South University, Changsha, China, in 2002, and the Ph.D. degree in computer science from the University of New England, Armidale, NSW, Australia, in 2008, focusing on dimensionality reduction for nonvectorial data representation.

From 2008 to 2016, he was with the Commonwealth Science and Industrial Research Organisation, Sydney, Australia, as a Computational Statistician on various projects in spectroscopy, remote sensing, and materials sciences. He joined the Centre for Research in Mathematics, School of Computing, Engineering and Mathematics, Western Sydney University, Werrington, NSW. His current research interests include machine learning, computational statistics, and big data.

Junbin Gao received the B.Sc. degree in computational mathematics from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 1982, and the Ph.D. degree from the Dalian University of Technology, Dalian, China, in 1991.

From 1982 to 2001, he was an Associate Lecturer, a Lecturer, an Associate Professor, and a Professor with the Department of Mathematics, HUST. He was a Professor in computer science with the School of Computing and Mathematics, Charles Sturt University, Bathurst, NSW. He was a Senior Lecturer and a Lecturer in computer science with the University of New England, Armidale, NSW, from 2001 to 2005. He is currently a Professor of Big Data Analytics with the University of Sydney Business School, University of Sydney, Camperdown, NSW, Australia. His research interests include machine learning, data analytics, Bayesian learning and inference, and image analysis.

Xiuping Jia (M'93–SM'03) received the B.Eng. degree from the 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, Kensington, NSW, Australia, in 1996.

Since 1988, she has been with the School of Engineering and Information Technology, the University of New South Wales at Canberra, Australia, where she is currently a Senior Lecturer. She has authored or coauthored over 170 referred papers, including 68 journal papers. She has coauthored the remote sensing textbook *Remote Sensing Digital Image Analysis* (Springer-Verlag, 3rd (1999) and 4th eds. (2006)) Her current research interests include remote sensing, image processing, and spatial data analysis.

Dr. Jia is a Subject Editor of the *Journal of Soils and Sediments* and an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.