Discriminative Low-Rank Gabor Filtering for Spectral–Spatial Hyperspectral Image Classification

Lin He, *Member, IEEE*, Jun Li, *Member, IEEE*, Antonio Plaza, *Fellow, IEEE*, and Yuanqing Li, *Senior Member, IEEE*

Abstract—Spectral–spatial classification of remotely sensed hyperspectral images has attracted a lot of attention in recent years. Although Gabor filtering has been used for feature extraction from hyperspectral images, its capacity to extract relevant information from both the spectral and the spatial domains of the image has not been fully explored yet. In this paper, we present a new discriminative low-rank Gabor filtering (DLRGF) method for spectral–spatial hyperspectral image classification. A main innovation of the proposed approach is that our implementation is accomplished by decomposing the standard 3-D spectral–spatial Gabor filter into eight subfilters, which correspond to different combinations of low-pass and bandpass single-rank filters. Then, we show that only one of the subfilters (i.e., the one that performs low-pass spatial filtering and bandpass spectral filtering) is actually appropriate to extract suitable features based on the characteristics of hyperspectral images. This allows us to perform spectral–spatial classification in a highly discriminative and computationally efficient way, by significantly decreasing the computational complexity (from cubic to linear order) compared with the 3-D spectral–spatial Gabor filter. In order to theoretically prove the discriminative ability of the selected subfilter, we derive an overall classification risk bound to evaluate the discriminating abilities of the features provided by the different subfilters. Our experimental results, conducted using different hyperspectral images, indicate that the proposed DLRGF method exhibits significant improvements in terms of classification accuracy and computational performance when compared with the 3-D spectral–spatial Gabor filter and other state-of-the-art spectral–spatial classification methods.

Index Terms—Discriminative spectral–spatial filtering, fast computation, Gabor filtering, hyperspectral imaging, low-rank filtering, spectral–spatial classification.

I. INTRODUCTION

HYPERSPECTRAL imaging collects hundreds of bands, at different wavelength channels, for the same area on the surface of the earth [1]–[3]. Classification is one of the most important techniques for remotely sensed hyperspectral image exploitation [4]–[6], [8]–[10]. The goal of classification is to assign the pixels (vectors) of a hyperspectral image into a set of unique and mutually exclusive class labels. Techniques for hyperspectral image classification can be roughly divided into two groups, namely, pixel-wise and spectral–spatial approaches [5]–[7]. The former group exclusively utilizes the spectral signatures of individual pixels to perform classification, whereas the latter exploits both the spectral and the spatial–contextual information [8].

Spectral–spatial classification techniques have become very popular due to their potential to better exploit available sources of information in hyperspectral images in order to achieve higher discriminability. They can be further divided into three categories according to the way spatial information is integrated into the (traditionally spectral-based) classification process: integrated spectral–spatial techniques, preprocessing-based techniques, and postprocessing-based techniques [8], [11].

1) Integrated techniques combine both the information derived from spectral signatures and spatial–contextual information into the classification decision directly. In this case, the exploitation of spectral and spatial information is not explicitly separated in the classification process. Among such a group of approaches, Bruzzone and Persello [12] and Li *et al.* [13] used spatial–contextual information to modify the objective and constraints of the support vector machine (SVM) pixel-wise classifier. Chen *et al.* [14], Li *et al.* [15], and Fang *et al.* [16] constructed a training sample-based collaborative dictionary and subsequently introduced contextual regression signals (pixels) to build a joint sparse/collaborative representation for classification purposes. Veganzones *et al.* [17] developed a spectral–spatial region model in which a region merging criterion and a set of spectral unmixing-based

pruning strategies were used to construct a binary partition tree able to segment hyperspectral images. Zhong and Wang [18] employed a Gibbs distribution to directly represent the joint *a posteriori* probability of labels conditioned on the whole hyperspectral image, so as to build a discriminative conditional random field for hyperspectral image classification.

2) In postprocessing-based techniques, a pixel-wise preliminary belief assignment of class belonging (typically conducted using spectral information alone) is improved by means of a spatial postprocessing regularizer. Kang *et al.* [19] utilized SVM to obtain preliminary class-belonging probability, and then used guiding images that reflect the homogeneity disruption to guide an edge-preserving classification. Li *et al.* [20], [21] used the multinomial logistic regression (MLR) to obtain the pixel-wise *a posteriori* classification output; then, a Markov random field (MRF) regularizer connected to the *a priori* probability of the labels was serialized, which was solved by the α-expansion or the loopy belief propagation strategies, to yield the final spectral–spatial *a posteriori* classification. Similar developments based on postprocessing with the MRF were presented in [22]–[25]; in the former two works, a probabilistic SVM and a Gaussian mixture classifier are alternatively used to model the spectral information, respectively, while in the latter two works, a homogeneity index and a total variation regularization are introduced to adaptively adjust the weight of spatial contribution, respectively. Specifically, Zhang *et al.* [26] applied the thread of the MRF-based postprocessing on the superpixels, each of which is a group of neighbored pixels.

3) In preprocessing-based techniques, the classification process is typically divided into two steps: 1) spectral–spatial feature extraction, which aims at extracting representative spectral–spatial features contained in the original hyperspectral image and 2) feature-wise classification, performed using standard classifiers such as SVM and MLR. Obviously, the feature extraction step is the key to the performance of these techniques [27]. Among many developments in this direction [8], [27], Benediktsson *et al.* [28], Plaza *et al.* [29], [30] and Fauvel [8], [31] *et al.* employed mathematical morphology operators with structural elements of various sizes to obtain extended morphological profiles that are used as features for classification purposes; lately, Mura *et al.* [32], Ghamisi *et al.* [33], Xia *et al.* [34], and Falco *et al.* [35] used more sophisticated morphological operations to develop morphological attribute profiles. He *et al.* [11] constructed a spatial translation-invariant spline wavelet transform to extract spectral–spatial features and then utilized primal–dual barrier-based linear programming to carry out a sparse representation-based classification. Tsai and Lai [36] utilized a modified 3-D gray-level coocurrence matrix to obtain spectral–spatial texture features. He *et al.* [37] designed spectral–spatial features using empirical mode decomposition and the

morphological wavelet transform, and then fed these features to a sparse multitask learning approach.

In addition to the aforementioned spectral–spatial techniques at a preprocessing level, the use of Gabor filters has been recently become widespread for hyperspectral image classification purposes [39], [40]. These filters, which have been successfully applied in many types of classification tasks, including biometric recognition [41]–[43], texture segmentation [44], [45], color image segmentation [46], signature recognition [47], and motion tracking [48], can yield optimal joint time–frequency resolution [38], [41], which makes them suitable for feature extraction purposes. In the context of hyperspectral imaging, Bau *et al.* [49] used the real part of 3-D Gabor filters to obtain the dominant energy features of regions in the hyperspectral image, suggesting the effectiveness and superiority of 3-D Gabor filtering in extracting spectral–spatial features. Shen and Jia [50] exploited both the real and imaginary parts of 3-D Gabor filters to extract suitable magnitude features. Jia *et al.* [51] combined such features with a least squares (LS) regression-based classifier to avoid selection of a massive number of Gabor features. Despite the success of these techniques in applying Gabor filtering to hyperspectral image processing, there are still some remaining challenging issues. First, directly convolving high-dimensional Gabor filters with a hyperspectral image is nearly prohibitive from a computational standpoint. Second, existing methods adopt a standard 3-D spectral–spatial Gabor filtering strategy without considering the special characteristics of hyperspectral images, in which spatial smoothing and spectral information preservation are the most critical aspects to extract suitable features for classification purposes [8]. In this regard, the standard 3-D spectral–spatial Gabor filtering strategy may generate additional features that not only increase computational complexity significantly but may also increase the classification risk. Based on the aforementioned observations, a straightforward question arises: is there a way to design a more discriminative and efficient technique based on Gabor filtering for spectral–spatial classification of hyperspectral images?

In order to address this research question, we develop a new discriminative low-rank Gabor filtering (DLRGF) method for spectral–spatial feature extraction prior to classification of hyperspectral data. A main innovation of the proposed approach is that our implementation is accomplished by decomposing the standard 3-D spectral–spatial Gabor filter into eight subfilters, which correspond to different combinations of low-pass and bandpass single-rank filters. To eliminate the interference of low-discriminating filtering components, only the subfilter connected to the low-pass spatial filtering and the bandpass spectral filtering is employed to extract spectral–spatial features. This is related to the fact that smoothness along spatial dimension and differential along the spectral dimension are highly beneficial to class separability, which is a special property of hyperspectral images not to be found in other kinds of images in which Gabor filtering has been used in the past [11], [52]–[54]. As compared with classic (3-D) Gabor filtering, our method offers two significant

advantages: 1) the possibility to obtain highly discriminative features, without the interference of low-discriminating filtering components and 2) a significant increase in computational efficiency since only the features that are suitable for discrimination purposes are retained for the subsequent classification. In fact, our interpretation of Gabor filtering (which is specifically tailored to the characteristics of hyperspectral data) represents a highly novel contribution with regard to previous developments, in which Gabor filtering was applied in a blind way without fully considering the properties of hyperspectral images. Another contribution of our proposed approach is that, resulting from the derivation of the proposed DLRGF, we also present a fast version of the traditional 3-D Gabor filtering approach that directly utilizes the low-rank decomposition to significantly speed up the calculation. A final contribution of our work is the derivation of an overall classification risk bound that is capable of measuring multiclass separability. This metric, constructed based on the Bayes smoothing on the upper bound of the class pair conditioned risk, is then used to demonstrate that the selected low-pass spatial and bandpass spectral subfilter achieves the highest class separability among all possible rank-1 spectral–spatial Gabor filters, justifying the validity of our fast discriminative Gabor filtering method.

The remainder of this paper is organized as follows. Section II describes the related work. The proposed spectral–spatial classification method is presented in detail in Section III. Section IV demonstrates the effectiveness of our newly developed technique using different hyperspectral images. Finally, Section V concludes with some remarks and hints at plausible future research lines.

II. RELATED WORK

Let us denote by (x, y, b) the spatial–spectral domain of a hyperspectral image, where components (x, y) denote the spatial domain and component b denotes the spectral domain. A standard spatial–spectral Gabor filter in this context can be mathematically formulated using a 3-D harmonic modulated with a 3-D Gaussian envelope as follows:

$$\mathbf{G}_{\boldsymbol{\omega},\theta,\varphi}(x, y, b) = \frac{1}{(2\pi)^{\frac{3}{2}}\sigma_x\sigma_y\sigma_b}$$
$$\times \exp\left\{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} + \frac{b^2}{\sigma_b^2}\right)\right\}$$
$$\times \exp\{j(x\omega_x + y\omega_y + b\omega_b)\} \quad (1)$$

where $\omega_x = |\boldsymbol{\omega}|\sin\varphi\cos\theta$, $\omega_y = |\boldsymbol{\omega}|\sin\varphi\sin\theta$, and $\omega_b = |\boldsymbol{\omega}|\cos\varphi$ are the projections of the given angular frequency $\boldsymbol{\omega}$ onto the x, y, and b directions, respectively; φ is the angle between the frequency $\boldsymbol{\omega}$ and the b-direction; θ is the angle between the projection of $\boldsymbol{\omega}$ onto the (x, y) plane and the x-direction; $|\boldsymbol{\omega}| = (\omega_x^2 + \omega_y^2 + \omega_b^2)^{\frac{1}{2}}$ is the magnitude of $\boldsymbol{\omega}$; and j stands for the imaginary unit. Let $\mathbf{h}(x, y, b)$ be a hyperspectral image; let $h(m, n, l)$ represent the value of a pixel of the hyperspectral image at the lth band and at the spatial location (m, n). Then, the standard 3-D

spatial–spectral Gabor filter (hereinafter, 3-DGF for convenience) can be defined by means of the following projection operator g:

$$g\mathbf{h}_{\boldsymbol{\omega},\theta,\varphi}(x, y, b)$$
$$= \langle\mathbf{h}, \mathbf{G}_{\boldsymbol{\omega},\theta,\varphi}(m - x, n - y, l - b)\rangle_{m,n,l}$$
$$= \sum_m \sum_n \sum_l h(m, n, l)\mathbf{G}_{\boldsymbol{\omega},\theta,\varphi}^c(m - x, n - y, l - b)$$
$$= \mathbf{h}(x, y, b) * \mathbf{G}_{\boldsymbol{\omega},\theta,\varphi}^c(-x, -y, -b) \quad (2)$$

where $\langle\cdot, \cdot\rangle$ denotes the inner product of two 3-D signals, $*$ denotes the convolution operation, and the exponent c denotes the conjugation operation.

Having the aforementioned notation in mind, Bau *et al.* [49] used only the real part of the 3-DGF shown in (2) for spectral–spatial representation. After the filtering process, the energy feature is calculated as follows:

$$E(x, y) = \sum_b \sum_m \sum_n \{g\mathbf{h}_{\boldsymbol{\omega}',\theta'}(m, n, b)\}^2. \quad (3)$$

It should be noted that, although this approach (referred to hereinafter as real-part Gabor energy feature) is initially designed for region-based characterization, it can be used to classify hyperspectral image pixels individually.

On the other hand, Shen and Jia [50] utilized the 3-DGF as defined in (2) to perform filtering on the original hyperspectral image cube. Then, the magnitudes calculated from the filtering result were used as features for classification purposes.

Finally, Jia *et al.* [51] also utilized the aforementioned Gabor magnitude features for hyperspectral image classification, where the high-dimensional features obtained under different parameter settings were directly fed into an LS-based collaborative representation classifier to avoid the feature selection step used in many conventional classifiers.

III. DISCRIMINATIVE LOW-RANK GABOR FILTERING

In this section, we describe the proposed DLRGF approach. Using Euler's relation, the 3-DGF in (2) can be rewritten as

$$g\mathbf{h}_{\boldsymbol{\omega},\theta,\varphi}(x, y, b) = \mathbf{h}(x, y, b) * \mathfrak{Re}\{\mathbf{G}_{\boldsymbol{\omega},\theta,\varphi}(-x, -y, -b)\}$$
$$- j\mathbf{h}(x, y, b) * \mathfrak{Im}\{\mathbf{G}_{\boldsymbol{\omega},\theta,\varphi}(-x, -y, -b)\}$$
$$(4)$$

where $\mathfrak{Re}\{\cdot\}$ and $\mathfrak{Im}\{\cdot\}$ represent the real and imaginary parts of a 3-D signal, respectively. As shown in (4), the 3-DGF involves two 3-D subfilters, related to the real and imaginary components. To further understand the mechanism of the 3-DGF in (4), in the following, we will decompose the filter into subfilters and obtain a fast low-rank spectral–spatial Gabor filter (LRGF). Furthermore, as more filtering components are obtained, we will be able to analyze 3-DGF structure in more detail. By taking the special characteristics of hyperspectral images into consideration, we will finally choose only the subfilter yielding the highest feature discriminability to build our final DLRGF.

A. Low-Rank Spectral–Spatial Gabor Filtering

Through a detailed inspection of (2) and (4), we can observe that the exponential harmonic $\exp\{j(x\omega_x + y\omega_y + b\omega_b)\}$ [or the corresponding sine and cosine harmonics $\sin\{j(x\omega_x + y\omega_y + b\omega_b)\}$ and $\cos\{j(x\omega_x + y\omega_y + b\omega_b)\}$] used by the 3-DGF are formed with the coupling of x, y, and b factors, whereas the Gaussian envelope $(1/((2\pi)^{(3/2)}\sigma_x\sigma_y\sigma_b^2))$ $\exp\{-(1/2)((x^2/\sigma_x^2) + (y^2/\sigma_y^2) + (b^2/\sigma_b^2))\}$ is easier to separate along the x, y, and b directions. For simplicity, we let $k = (1/((2\pi)^{(3/2)}\sigma_x\sigma_y\sigma_b))\exp\{-(1/2)((x^2/\sigma_x^2) + (y^2/\sigma_y^2) + (b^2/\sigma_b^2))\}$ represent the Gaussian envelope. Then, we have

$$\Re\{\Omega_{\boldsymbol{\omega},\theta,\varphi}(-x, -y, -b)\} = k\cos(-x\omega_x - y\omega_y - b\omega_b)$$
$$\Im\{\Omega_{\boldsymbol{\omega},\theta,\varphi}(-x, -y, -b)\} = k\sin(-x\omega_x - y\omega_y - b\omega_b).$$
$$(5)$$

The real and imaginary parts of the spectral–spatial Gabor filter above can be decomposed as

$$\Re\{\Omega_{\boldsymbol{\omega},\theta,\varphi}(-x, -y, -b)\}$$
$$= k\cos(-x\omega_x)\cos(-y\omega_y - b\omega_b) - k\sin(-x\omega_x)$$
$$\times \sin(-y\omega_y - b\omega_b)$$
$$= k\cos(x\omega_x)\cos(y\omega_y)\cos(b\omega_b) - k\cos(x\omega_x)$$
$$\times \sin(y\omega_y)\sin(b\omega_b) - k\sin(x\omega_x)\sin(y\omega_y)\cos(b\omega_b)$$
$$- k\sin(x\omega_x)\cos(y\omega_y)\sin(b\omega_b)$$
$$= g_{\cos}^{(x)} \cdot g_{\cos}^{(y)} \cdot g_{\cos}^{(b)} - g_{\cos}^{(x)} \cdot g_{\sin}^{(y)} \cdot g_{\sin}^{(b)}$$
$$- g_{\sin}^{(x)} \cdot g_{\sin}^{(y)} \cdot g_{\cos}^{(b)} - g_{\sin}^{(x)} \cdot g_{\cos}^{(y)} \cdot g_{\sin}^{(b)} \qquad (6)$$

and

$$\Im\{\Omega_{\boldsymbol{\omega},\theta,\varphi}(-x, -y, -b)\}$$
$$= k\sin(-x\omega_x)\cos(-y\omega_y - b\omega_b) + k\cos(-x\omega_x)$$
$$\times \sin(-y\omega_y - b\omega_b)$$
$$= -k\sin(x\omega_x)\cos(y\omega_y)\cos(b\omega_b) + k\sin(x\omega_x)$$
$$\times \sin(y\omega_y)\sin(b\omega_b) - k\cos(x\omega_x)\sin(y\omega_y)\cos(b\omega_b)$$
$$- k\cos(x\omega_x)\cos(y\omega_y)\sin(b\omega_b)$$
$$= -g_{\sin}^{(x)} \cdot g_{\cos}^{(y)} \cdot g_{\cos}^{(b)} + g_{\sin}^{(x)} \cdot g_{\sin}^{(y)} \cdot g_{\sin}^{(b)}$$
$$- g_{\cos}^{(x)} \cdot g_{\sin}^{(y)} \cdot g_{\cos}^{(b)} - g_{\cos}^{(x)} \cdot g_{\cos}^{(y)} \cdot g_{\sin}^{(b)} \qquad (7)$$

where

$$g_{\cos}^{(x)} = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_x}\exp\left(-\frac{x^2}{2\sigma_x^2}\right)\cos(x\omega_x)$$

$$g_{\cos}^{(y)} = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_y}\exp\left(-\frac{y^2}{2\sigma_y^2}\right)\cos(y\omega_y)$$

$$g_{\cos}^{(b)} = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_b}\exp\left(-\frac{b^2}{2\sigma_b^2}\right)\cos(b\omega_b)$$

$$g_{\sin}^{(x)} = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_x}\exp\left(-\frac{x^2}{2\sigma_x^2}\right)\sin(x\omega_x)$$

$$g_{\sin}^{(y)} = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_y}\exp\left(-\frac{x^2}{2\sigma_y^2}\right)\sin(y\omega_y)$$

and

$$g_{\sin}^{(b)} = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_b}\exp\left(-\frac{b^2}{2\sigma_b^2}\right)\sin(b\omega_b).$$

Algorithm 1 Generation of Rank-1 Subfilters for Spectral–Spatial Gabor Filtering

1) Compute the rank-1 Gaussian envelope:
   $en^{(i)} = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_i}\exp(-\frac{i^2}{2\sigma_i^2})|_{i=x,y,b}.$
2) Compute the rank-1 cosine harmonic:
   $har_{\cos}^{(i)} = \cos(i\omega_i)|_{i=x,y,b}$ and the sinusoidal harmonic:
   $har_{\sin}^{(i)} = \sin(i\omega_i)|_{i=x,y,b}.$
3) Compute the rank-1 subfilters
   $g_{\cos}^{(i)} = en^{(i)} \cdot har_{\cos}^{(i)}|_{i=x,y,b}$ and
   $g_{\sin}^{(i)} = en^{(i)} \cdot har_{\sin}^{(i)}|_{i=x,y,b}.$

Equations (6) and (7) suggest that a 3-DGF is actually formed by the superposition of eight subfilters, rather by merely two real and imaginary components. Furthermore, each of these subfilters is constituted by the separable tensor product of three rank-1 filters along the x, y, and b-directions, respectively. Therefore, (6) and (7) reveal the structure of the spectral–spatial Gabor filter in a much more elaborate manner than (1) and (4), thus offering us the possibility of conducting a deeper exploration of the mechanism of the filter. The generation of the aforementioned rank-1 filters in (6) and (7) is described in Algorithm 1.

For convenience, the four terms on the right-hand side of the real-part representation in (6), i.e., the four subfilters resulting from the real part of the 3-DGF, are hereinafter denoted by $\text{Re}1(x, y, b)$, $\text{Re}2(x, y, b)$, $\text{Re}3(x, y, b)$, and $\text{Re}4(x, y, b)$, respectively. Similarly, the four terms in the imaginary part representation in (7), i.e., the other four subfilters from the imaginary part of the 3-DGF, are hereinafter denoted by $\text{Im}1(x, y, b)$, $\text{Im}2(x, y, b)$, $\text{Im}3(x, y, b)$, and $\text{Im}4(x, y, b)$, respectively. According to (4), (6), and (7), a spatial–spectral Gabor filter for a hyperspectral image is associated with the two convolutions $\mathbf{h}(x, y, b) * \Re\{\mathbf{G}_{\boldsymbol{\omega},\theta,\varphi}(-x, -y, -b)\}$ and $\mathbf{h}(x, y, b) * \Im\{\mathbf{G}_{\boldsymbol{\omega},\theta,\varphi}(-x, -y, -b)\}$, and is thus associated with the eight convolutions between $\mathbf{h}(x, y, b)$ and $\text{Re}1(x, y, b)$, $\text{Re}2(x, y, b)$, $\text{Re}3(x, y, b)$, $\text{Re}4(x, y, b)$, $\text{Im}1(x, y, b)$, $\text{Im}2(x, y, b)$, $\text{Im}3(x, y, b)$, and $\text{Im}4(x, y, b)$. Taking $\text{Re}1(x, y, b)$ as an example, its convolution with a hyperspectral image $\mathbf{h}(x, y, b)$ can be denoted as follows:

$$\mathbf{h}(x, y, b) * Re1(x, y, b)$$
$$= \sum_m \sum_n \sum_l h(m, n, l) \cdot \text{Re}1_{\boldsymbol{\omega},\theta,\varphi}(x - m, y - n, b - l)$$
$$= \sum_l \sum_n \sum_m h(m, n, l) \cdot g_{\cos}^{(x)}(x - m) \cdot g_{\cos}^{(y)}(y - n)$$
$$\cdot g_{\cos}^{(b)}(b - l)$$
$$= \mathbf{h}(x, y, b) * g_{\cos}^{(x)}(x) * g_{\cos}^{(y)}(y) * g_{\cos}^{(b)}(b). \qquad (8)$$

Thus, the spectral–spatial Gabor filtering operation on $\mathbf{h}(x, y, b)$ can be reformulated as follows:

$$g\mathbf{h} = \mathbf{h} * g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\cos}^{(b)} - \mathbf{h} * g_{\cos}^{(x)} * g_{\sin}^{(y)} * g_{\sin}^{(b)}$$
$$- \mathbf{h} * g_{\sin}^{(x)} * g_{\sin}^{(y)} * g_{\cos}^{(b)} - \mathbf{h} * g_{\sin}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$$
$$- j\{\mathbf{h} * g_{\sin}^{(x)} * g_{\cos}^{(y)} * g_{\cos}^{(b)} - \mathbf{h} * g_{\sin}^{(x)} * g_{\sin}^{(y)} * g_{\sin}^{(b)}$$
$$+ \mathbf{h} * g_{\cos}^{(x)} * g_{\sin}^{(y)} * g_{\cos}^{(b)} + \mathbf{h} * g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}\} \qquad (9)$$

Algorithm 2 LRGF
1) Input: a hyperspectral image $\mathbf{h}(x, y, b)$.
2) Compute rank-1 subfilters $g_{\cos}^{(i)}\|_{i=x,y,b}$ and $g_{\sin}^{(i)}\|_{i=x,y,b}$ using Algorithm 1.
3) Compute the eight subfilters: $g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\cos}^{(b)}$, $g_{\cos}^{(x)} * g_{\sin}^{(y)} * g_{\sin}^{(b)}$, $g_{\sin}^{(x)} * g_{\sin}^{(y)} * g_{\cos}^{(b)}$, $g_{\sin}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$, $g_{\sin}^{(x)} * g_{\cos}^{(y)} * g_{\cos}^{(b)}$, $g_{\sin}^{(x)} * g_{\sin}^{(y)} * g_{\sin}^{(b)}$, $g_{\cos}^{(x)} * g_{\sin}^{(y)} * g_{\cos}^{(b)}$, and $g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$.
4) Perform Gabor filtering using (9).

where $\mathbf{h}$ denotes $\mathbf{h}(x, y, b)$ and $g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\cos}^{(b)}$ denotes $g_{\cos}^{(x)}(x) * g_{\cos}^{(y)}(y) * g_{\cos}^{(b)}(b)$.

If we compare (4) with (9), we can conclude that the direct calculation of 3-DGF involves two 3-D convolutions, whereas the proposed low-rank filtering in (9) involves serial 1-D convolutions along the x, y, and b directions. If we assume that the sizes of a spectral–spatial Gabor filter along the x, y, and b directions are always L, in this case, the computational complexity of the proposed approach in (9) is linear (with the order of L), whereas the computational complexity of 3-DGF in (4) is cubic. Therefore, our LRGF method (as shown in Algorithm 2) has much lower computation complexity than the 3-DGF and can be in fact considered as a fast implementation of 3-DGF.

B. Discriminative Low-Rank Gabor Filtering

Based on the previously introduced efficient LRGF method, in this section, we develop our new DLRGF approach designed to obtain highly discriminative features and for spectral–spatial hyperspectral image classification with low computational complexity. This section is organized as follows. First, we study the characteristics of the subfilters that integrate the LRGF. Then, we present the DLRGF formulation that is based on an intelligent selection of the most discriminative subfilter.

1) Characteristics of the Rank-1 Filters in LRGF: As shown in (9), by utilizing a low-rank convolution mechanism, the LRGF is able to significantly reduce the computational complexity of 3-DGF from cubic to linear order. Following the works in [49]–[51], we can directly use the resulting features from LRGF for hyperspectral image classification, with the advantage that the proposed LRGF has considerably lower computational complexity than the 3-DGF. However, (9) reveals that the LRGF is more than a fast algorithm. According to such a mathematical representation, instead of simply the real and imaginary parts, the LRGF involves a total of eight components after the spectral–spatial filtering, each of which is produced by a subfilter that is associated with three single rank-1 filters. A natural question arising at this point is the role of each subfilter in the subsequent spectral–spatial classification.

As observed in (9), the rank-1 filters that are used by the LRGF are 1-D Gaussian-modulated sine or cosine harmonics along different directions. Without loss of generality, let us consider the 1-D Gaussian-modulated cosine and sine
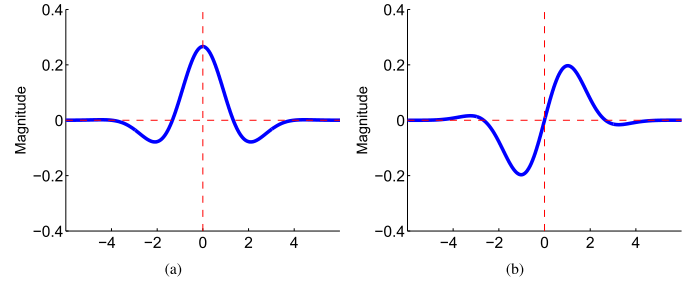

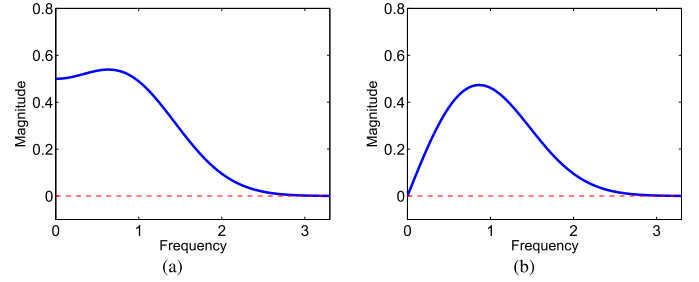Fig. 1. Gaussian enveloped (a) cosine and (b) sinusoidal harmonics.

Fig. 2. Frequency magnitude of Gaussian enveloped (a) cosine and (b) sinusoidal harmonics.

harmonics (as shown in Fig. 1)

$$g_{\cos}^{(i)} = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma} \exp\left(-\frac{i^2}{2\sigma^2}\right) \cos(\omega i)\|_{i=x,y,b}$$

and

$$g_{\sin}^{(i)} = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma} \exp\left(-\frac{i^2}{2\sigma^2}\right) \sin(\omega t)\|_{i=x,y,b}.$$

For a given frequency ω_0, their mathematical representations in the frequency domain are

$$
\begin{aligned}
\widehat{g}_{\cos}^{\omega_0}(\omega) \\
&= \frac{1}{(2\pi)^{\frac{1}{2}}\sigma} \int_{-\infty}^{+\infty} \exp\left(-\frac{i^2}{2\sigma^2}\right) \cos(\omega_0 i) \exp(-j\omega i) di \\
&= \frac{1}{2(2\pi)^{\frac{1}{2}}\sigma} \int_{-\infty}^{+\infty} \exp\left\{-\frac{i^2}{2\sigma^2} - j(\omega - \omega_0)i\right\} di \\
&\quad + \frac{1}{2(2\pi)^{\frac{1}{2}}\sigma} \int_{-\infty}^{+\infty} \exp\left\{-\frac{i^2}{2\sigma^2} - j(\omega + \omega_0)i\right\} di
\end{aligned}
$$
(10)

and

$$
\begin{aligned}
\widehat{g}_{\sin}^{\omega_0}(\omega) \\
&= \frac{1}{(2\pi)^{\frac{1}{2}}\sigma} \int_{-\infty}^{+\infty} \exp\left(-\frac{i^2}{2\sigma^2}\right) \sin(\omega_0 i) \exp(-j\omega i) di \\
&= -j\frac{1}{2(2\pi)^{\frac{1}{2}}\sigma} \int_{-\infty}^{+\infty} \exp\left\{-\frac{i^2}{2\sigma^2} - j(\omega - \omega_0)i\right\} di \\
&\quad -j\frac{1}{2(2\pi)^{\frac{1}{2}}\sigma} \int_{-\infty}^{+\infty} \exp\left\{-\frac{i^2}{2\sigma^2} - j(\omega + \omega_0)i\right\} di
\end{aligned}
$$
(11)

TABLE I

OVERALL CLASSIFICATION RISK BOUNDS (10^{-4}). LABELS 1–8 STAND FOR THE DIFFERENT SUBFILTERS: $g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\cos}^{(b)}$ (1), $g_{\cos}^{(x)} * g_{\sin}^{(y)} * g_{\sin}^{(b)}$ (2), $g_{\sin}^{(x)} * g_{\sin}^{(y)} * g_{\cos}^{(b)}$ (3), $g_{\sin}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$ (4), $g_{\sin}^{(x)} * g_{\cos}^{(y)} * g_{\cos}^{(b)}$ (5), $g_{\sin}^{(x)} * g_{\sin}^{(y)} * g_{\sin}^{(b)}$ (6), $g_{\cos}^{(x)} * g_{\sin}^{(y)} * g_{\cos}^{(b)}$ (7), AND $g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$ (8). THE 3-DGF (9) IS ALSO REPRESENTED

Hyperspectral image	1	2	3	4	5	6	7	8	9
Zaoyuan	6.11	861.14	1241.92	878.23	733.41	1282.31	349.56	**1.46**	4.97
Kennedy Space Center	2.29	281.54	705.12	316.26	517.40	311.33	526.30	**0.39**	1.60
Indian Pines	104.39	303.64	960.77	461.57	520.46	1161.77	355.70	**33.16**	53.50
Salinas	51.31	109.99	201.94	101.86	148.01	137.11	154.55	**28.79**	40.49
Pavia University	14.08	413.59	240.49	361.50	293.77	483.30	370.40	**5.28**	13.80

respectively, where

$$\int_{-\infty}^{+\infty} \exp\left\{-\frac{i^2}{2\sigma^2} - j(\omega - \omega_0)i\right\} di$$
$$= \int_{-\infty}^{+\infty} \exp\left\{-\frac{\{i + j\sigma^2(\omega - \omega_0)\}^2}{2\sigma^2} - \frac{\sigma^2(\omega - \omega_0)^2}{2}\right\} di$$
$$= \sqrt{2\pi}\sigma \exp\left\{-\frac{\sigma^2(\omega - \omega_0)^2}{2}\right\} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left\{-\frac{i^2}{2}\right\} di$$
$$= \sqrt{2\pi}\sigma \exp\left\{-\frac{\sigma^2(\omega - \omega_0)^2}{2}\right\} \tag{12}$$

and

$$\int_{-\infty}^{+\infty} \exp\left\{-\frac{i^2}{2\sigma^2} - j(\omega + \omega_0)i\right\} di$$
$$= \sqrt{2\pi}\sigma \exp\left\{-\frac{\sigma^2(\omega + \omega_0)^2}{2}\right\}. \tag{13}$$

Substituting both (12) and (13) into (10) and (11), respectively, we obtain the following frequency magnitudes:

$$\left|\hat{g}_{\cos}^{\omega_0}(\omega)\right| = \frac{1}{2}\left|\exp\left\{-\frac{\sigma^2(\omega - \omega_0)^2}{2}\right\}\right.$$
$$\left. + \exp\left\{-\frac{\sigma^2(\omega + \omega_0)^2}{2}\right\}\right| \tag{14}$$

and

$$\left|\hat{g}_{\sin}^{\omega_0}(\omega)\right| = \frac{1}{2}\left|\exp\left\{-\frac{\sigma^2(\omega + \omega_0)^2}{2}\right\}\right.$$
$$\left. - \exp\left\{-\frac{\sigma^2(\omega - \omega_0)^2}{2}\right\}\right|. \tag{15}$$

If the frequency ω in (14) and (15) is zero, then $|\hat{g}_{\cos}^{\omega_0}(\omega)|$ is nonzero and $|\hat{g}_{\sin}^{\omega_0}(\omega)|$ is zero, which implies that $\hat{g}_{\cos}^{\omega_0}(\omega)$ is low-frequency pass and $\hat{g}_{\sin}^{\omega_0}(\omega)$ is low-frequency resistant. Fig. 2 shows the appearance of the frequency-domain representations $|\hat{g}_{\cos}^{\omega_0}(\omega)|$ and $|\hat{g}_{\sin}^{\omega_0}(\omega)|$; clearly, $g_{\cos}^{(i)}$ is a low-pass filter, while $g_{\sin}^{(i)}$ is a bandpass filter.

2) Derivation of the DLRGF: Since $\hat{g}_{\cos}^{\omega_0}(\omega)$ is a low-pass filter and $\hat{g}_{\sin}^{\omega_0}(\omega)$ is a bandpass filter, eight subfilters $g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\cos}^{(b)}$, $g_{\cos}^{(x)} * g_{\sin}^{(y)} * g_{\sin}^{(b)}$, $g_{\sin}^{(x)} * g_{\sin}^{(y)} * g_{\cos}^{(b)}$, $g_{\sin}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$, $g_{\sin}^{(x)} * g_{\cos}^{(y)} * g_{\cos}^{(b)}$, $g_{\sin}^{(x)} * g_{\sin}^{(y)} * g_{\sin}^{(b)}$, $g_{\cos}^{(x)} * g_{\sin}^{(y)} * g_{\cos}^{(b)}$, and $g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$, associated with the eight components of the spectral–spatial Gabor filter as shown in (9), are built by combining various rank-1 low-pass and bandpass filters along different spectral and spatial directions. Previous research has demonstrated that, in order to enhance the discriminability in hyperspectral images, smoothness in the spatial domain [11]

and the differential between various bands in the spectral domain [53], [54] are preferred. This is because, in hyperspectral images, the discriminating information tends to appear on low frequencies in the spatial dimension and away from low frequencies in the spectral dimension. Inspired by this observation, our introspection is that (among the eight considered subfilters) the filter $g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$—which corresponds to a low-pass filter in the spatial dimension and a bandpass filter in the spectral dimension—is more discriminative than the other subfilters. To quantitatively justify this introspection regarding the potential superiority of the filtering $\mathbf{h} * g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$, we introduce an overall classification risk bound to measure the separability of multiple classes, using the Bhattacharyya bound defined in [55]. The classification risk can be represented as

$$R(err) = \int_{\mathbb{R}^B} P(\text{err}, \mathbf{h}_p) d\mathbf{h}_p$$
$$= \int_{\mathbb{R}^B} \sum_j \sum_i P(\text{err}, \mathbf{h}_p | (c_i, c_j)) P(c_i, c_j) d\mathbf{h}_p$$
$$= \sum_j \sum_i P(c_i, c_j) \int_{\mathbb{R}^B} P(\text{err}, \mathbf{h}_p | (c_i, c_j)) d\mathbf{h}_p$$
$$\leq \sum_j \sum_i P(c_i, c_j) P(c_i | (c_i, c_j))^{\frac{1}{2}} P(c_j | (c_i, c_j))^{\frac{1}{2}}$$
$$\times \exp\{-BD(c_i, c_j)\}$$
$$= 2 \sum_{j, j < i} \sum_i \left(\frac{N_i N_j}{\sum_j \sum_i (N_i N_j)}\right) \left(\frac{N_i^{\frac{1}{2}} N_j^{\frac{1}{2}}}{N_i + N_j}\right)$$
$$\times \exp\{-BD(c_i, c_j)\} \tag{16}$$

where B is the number of spectral bands, and $\mathbf{h}_p$ denotes a pixel of the hyperspectral image; $P(c_i, c_j)$ is a probability of occurrence of the classification between classes i and j, and $P(c_i | (c_i, c_j))$ stands for the probability of class i under such a classification task; $P(c_i | (c_i, c_j))^{(1/2)} P(c_j | (c_i, c_j))^{(1/2)} \exp\{BD(c_i, c_j)\}$ represents the Bhattacharyya bound given by classes i and j; $BD(c_i, c_j) = (1/8)(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T ((\Sigma_i + \Sigma_j)/2)^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) + (1/2)\ln(|(\Sigma_i + \Sigma_j)2)|)/(|\Sigma_i||\Sigma_j|)^{(1/2)}$ denotes the Bhattacharyya distance between classes i and j [55], $(\boldsymbol{\mu}_i, \Sigma_i)$ and $(\boldsymbol{\mu}_j, \Sigma_j)$ are the mean and covariance of classes i and j, respectively, and N_i is the number of pixels of the ith class. Note that, due to the classification context, we have $P(c_i, c_j) = \begin{cases} 0, & i = j \\ P(c_j, c_i), & \text{otherwise} \end{cases}$. The quantity in the last
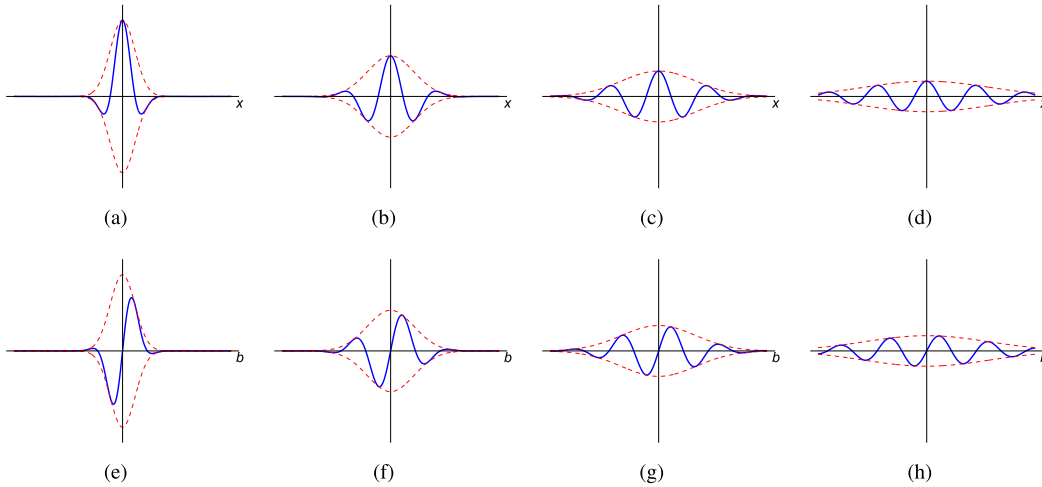
Fig. 3. Impact of the envelope variance on Gaussian enveloped harmonics when the harmonic frequency keeps unchanged. (a)–(d) Cosine harmonics [the envelope variance increases from (a) to (d)]. (e)–(h) Sinusoidal harmonics [the envelope variance increases from (e) to (h)].

line of (16)

$$ORB = 2 \sum_{j, j < i} \sum_{i} \left(\frac{N_i N_j}{\sum_{j} \sum_{i} (N_i N_j)} \right) \left(\frac{N_i^{\frac{1}{2}} N_j^{\frac{1}{2}}}{N_i + N_j} \right) \times \exp\{-BD(c_i, c_j)\} \quad (17)$$

is an upper bound of the overall classification risk $R(err)$. We can utilize this metric to evaluate the separability of multiple classes.

Table I illustrates the overall classification risk bounds calculated for different subfilters with five real hyperspectral data sets collected by different instruments over the regions of Zaoyuan (China), Kennedy Space Center, Indian Pines and Salinas Valley (United States), and the University of Pavia (Italy). Three of these images (Zaoyuan, Kennedy Space Center, and Indian Pines) will also be used in the detailed experiments reported in Section IV. In Table I, the subfilter $g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$ clearly presents the lowest classification risk bound on all considered hyperspectral images (as indicated by the bolded numbers), which means that the risk of classification error achieved by this filter is the lowest. Accordingly, such a filter is expected to yield higher class separability than all other subfilters. It should be especially be noted that the subfilter $\mathbf{h} * g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)}$ achieves better class separability than the 3-DGF. This is due to the fact that, by preserving only the low-pass filtering in the spatial domain and the bandpass filtering in the spectral domain, the other seven (low discriminative) filtering components that will interfere with classification are eliminated. Therefore, with the filtering

$$g'\mathbf{h} = \mathbf{h} * g_{\cos}^{(x)} * g_{\cos}^{(y)} * g_{\sin}^{(b)} \quad (18)$$

a fast DLRGF can be simply obtained as indicated in Algorithm 3.

Equation (18) suggests that our DLRGF is a filtering in a 3-D spatial–spectral space that is realized by convolving a hyperspectral image serially with three 1-D local cosine/sinusoidal harmonics (Gaussian enveloped harmonics), instead of a 3-D filter. Fig. 3 shows the local

---

Algorithm 3 DLRGF

1) Input: a hyperspectral image $\mathbf{h}(x, y, b)$.
2) Compute low-rank filters: $g_{\cos}^{(x)}$, $g_{\cos}^{(y)}$ and $g_{\sin}^{(b)}$ using Algorithm 1.
3) Perform Gabor filtering with (18).

---

harmonics with the varying variance of the Gaussian envelope, while the frequency magnitude is fixed. Fig. 3(a)–(d) shows the cosine harmonics along the x-dimension used in DLRGF (what are used along the y-dimension are the similar cosine harmonics), whereas Fig. 3(e)–(h) shows the sinusoidal harmonics applied along the spectral dimension (i.e., the b-dimension). As observed in these figures, with the increase in the envelope variance, the cover of a harmonic grows along spatial/spectral dimension, and thus the spatial/spectral locality of the harmonic is reduced. Fig. 4 shows the local harmonics with the varying frequency magnitude, while the variance of the Gaussian envelope is fixed. As observed, with the growing of the frequency magnitude, a harmonic tends to get to greater oscillation and hence tends to capture the signal with faster change.

Our proposed DLRGF method has two key advantages: 1) low computational complexity and 2) high discriminability. Given a spectral–spatial Gabor kernel of size L, the DLRGF (as shown in Algorithm 3) involves only 1-D convolution with computation complexity on the linear order of L, while the 3-DGF [as shown in (2)] involves convolutions of two 3-D tensors with computation complexity on the cubic order of L. In addition, both the DLRGF and LRGF (i.e., the fast implementation of 3-DGF developed in Algorithm 2) exhibit linear computational complexity. However, the DLRGF uses only one of eight subfilters adopted in LRGF; hence, it is expected to run much faster in practice. By preserving only the most discriminative components and consequently ignoring the interference of the low discriminative components, our DLRGF is expected to extract more discriminative spectral–spatial features than the 3-DGF (and the LRGF, which can be seen as a fast
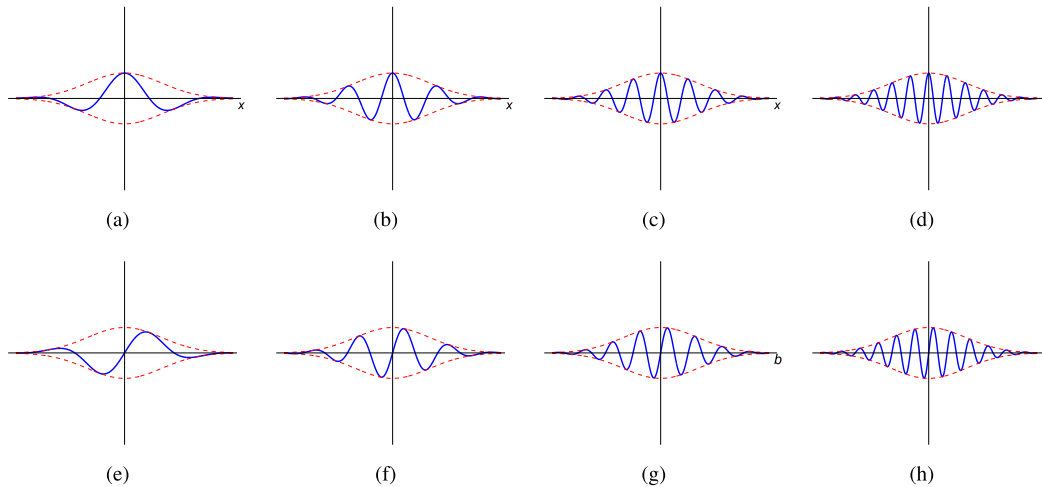
Fig. 4.   Impact of the harmonic frequency on Gaussian enveloped harmonics when the envelope variance keeps unchanged. (a)–(d) Cosine harmonics [the harmonic frequency increases from (a) to (d)]. (e)–(h) Sinusoidal harmonics [the harmonic frequency increases from (e) to (h)].

implementation of 3-DGF). In the following section, we will analyze the implications on the obtained classification accuracy.

IV. EXPERIMENTAL RESULTS

This section evaluates the proposed DLRGF using three real hyperspectral image data sets. For comparison, three spectral–spatial Gabor filtering methods including our proposed DLRGF, the classic 3-DGF, and the real-part standard spectral–spatial Gabor filtering (ReGF) that uses only the real part of 3-DGF are considered. For each of them, the frequency length $|\omega|$ is set to be $(\pi/16)$, $(\pi/8)$, $(\pi/4)$, and $(\pi/2)$; the angle between the frequency and the spectral dimension φ is set to 0, $(\pi/4)$, $(\pi/2)$, and $(3\pi/4)$; and the angle between the projection of the frequency onto the (x, y) plane and the x-direction (θ) is also set to values 0, $(\pi/4)$, $(\pi/2)$, and $(3\pi/4)$. Therefore, there are in total 52 Gabor filters considered for each filtering operation (note that if φ is set to be 0, various θs yield the same filter), and the magnitude of the filtering result is used as feature. Hence, our experiments comprise 52 features per band and $52 \times B$ features for a hyperspectral image if it is composed of B bands. The experiments conducted with each of the hyperspectral images are structured as follows.

1) We compute the running times of various Gabor filtering operations. Here, we consider our proposed DLRGF (see Algorithm 3), the classic 3-DGF [see (4)], and the ReGF that uses only the real part of the 3-DGF. The computational complexity of DLRGF is linear, while the complexities of 3-DGF and ReGF are cubic. In addition, the developed fast implementation of 3-DGF (called LRGF; see Algorithm 2) is also considered. This algorithm has linear computational complexity as well.

2) Then, we evaluate the classification accuracy of the considered approaches. In our experiments, we first apply Gabor filtering to obtain spectral–spatial features; then, the LS-based collaborative representation classifier and the SVM classifier are used to perform the subsequent classification. The LS collaborative representation classifier is directly applied to the original

TABLE II
GROUND-TRUTH CLASSES IN THE OMIS ZAOYUAN
HYPERSPECTRAL IMAGE

Number	Class name	Number of samples
1	Vegetable	2625
2	Grape	1302
3	Dry vegetable	3442
4	Pear	10243
5	Corn	1425
6	Terrace/Grass	1484
7	Bush-Lespedeza	1808
8	Peach	1495
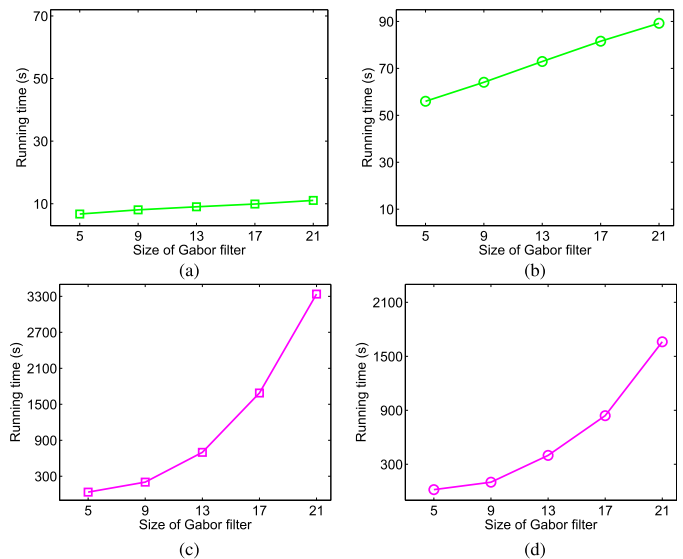
Fig. 5.   Computational time (in seconds) as a function of the size of the Gabor filters for the OMIS Zaoyuan image. (a) DLRGF. (b) LRGF. (c) 3-DGF. (d) ReGF.

features resulting from Gabor filtering, while the SVM is applied to the principal components of the Gabor features. The kernel function employed by the SVM is a radial basis function (RBF) kernel. Three Gabor-based methods: our proposed DLRGF, the classic 3-DGF, and the ReGF are considered. Since we have two different classifiers, our comparison among Gabor-related classifications comprises six approaches in total, namely,

TABLE III

OVERALL ACCURACY (%) AND KAPPA COEFFICIENT (IN THE PARENTHESES) FOR THE OMIS ZAOYUAN IMAGE AS A FUNCTION OF THE TRAINING SAMPLE SIZE (% OF LABELED PIXELS). THE BEST RESULTS ARE HIGHLIGHTED IN BOLD TYPEFACE

Sample size	0.1	0.2	0.5	0.8	1
DLRGF-LS	**77.51 (0.71)**	**86.51 (0.82)**	**92.75 (0.90)**	**93.15 (0.91)**	**94.71 (0.93)**
3DGF-LS	70.78 (0.62)	73.76 (0.66)	84.38 (0.80)	87.49 (0.84)	88.83 (0.85)
ReGF-LS	67.22 (0.59)	71.78 (0.64)	80.64 (0.75)	80.25 (0.74)	83.16 (0.78)
DLRGF-SVM	74.93 (0.67)	84.56 (0.80)	91.80 (0.89)	92.93 (0.91)	93.47 (0.91)
3DGF-SVM	71.72 (0.62)	83.06 (0.77)	89.27 (0.86)	91.14 (0.88)	91.81 (0.89)
ReGF-SVM	72.21 (0.61)	81.51 (0.75)	88.51 (0.85)	91.17 (0.88)	91.30 (0.89)
3DW-DB1	66.18 (0.54)	75.88 (0.67)	85.92 (0.81)	87.78 (0.84)	88.45 (0.85)
CKSVM	73.90 (0.64)	82.58 (0.78)	89.34 (0.86)	91.63 (0.89)	91.73 (0.89)
SVM-MRF	67.85 (0.53)	83.04 (0.78)	90.48 (0.88)	91.66 (0.89)	90.58 (0.87)
SVM	64.07 (0.50)	70.49 (0.62)	79.83 (0.74)	82.99 (0.78)	84.72 (0.80)

TABLE IV

INDIVIDUAL CLASSIFICATION ACCURACIES (%) FOR THE OMIS ZAOYUAN IMAGE, USING ONLY 0.5% OF THE LABELED PIXELS PER CLASS FOR TRAINING AND THE REMAINING LABELED PIXELS FOR TESTING

| Class	DLRGF-LS	3DGF-LS	ReGF-LS	DLRGF-SVM	3DGF-SVM	ReGF-SVM	3DW-DB1	CKSVM	SVM-MRF	SVM
1	90.62	74.99	71.51	94.22	92.95	93.30	74.42	83.03	98.09	74.11
2	90.04	86.02	88.34	91.27	88.11	85.10	93.51	86.10	92.66	87.57
3	95.94	94.33	94.51	93.84	93.31	93.34	89.92	95.56	88.99	79.88
4	97.14	91.17	85.31	97.50	95.84	96.16	91.52	89.71	87.77	81.91
5	77.76	64.64	50.95	70.92	83.27	79.53	76.15	87.44	94.50	60.62
6	88.55	83.67	79.95	89.97	75.95	76.63	84.96	88.48	89.63	78.05
7	92.77	67.96	70.58	91.66	90.55	80.09	88.60	99.22	99.22	98.83
8	79.72	69.27	67.12	66.17	46.83	49.93	58.89	77.02	83.56	65.84
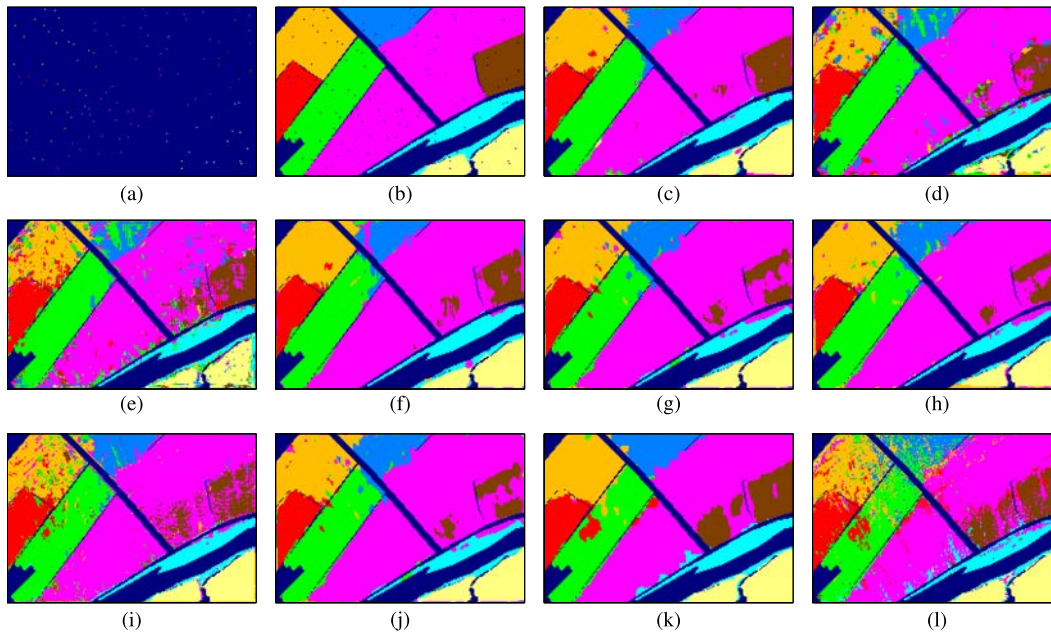
Fig. 6. Classification results for the OMIS Zaoyuan image using only 0.5% of the labeled pixels available per class for training. (a) Training samples. (b) Test samples. (c) DLRGF-LS. (d) 3-DGF-LS. (e) ReGF-LS. (f) DLRGF-SVM. (g) 3-DGF-SVM. (h) ReGF-SVM. (i) 3-DW-DB1. (j) CKSVM. (k) SVM-MRF. (l) SVM.

DLRGF-LS, 3-DGF-LS, ReGF-LS, DLRGF-SVM, 3-DGF-SVM, and ReGF-SVM. In our experiments, we use the fast implementations of 3-DGF and ReGF to perform classification because those implementations yield exactly the same results as the 3-DGF and ReGF while being more computationally efficient. In this case, we still use the names "3-DGF" and "ReGF" when we report the classification results achieved by the fast methods. Besides such six Gabor feature-based methods, the state-of-the-art SVM pixel-wise classifier [59], [61] and two representative spectral–spatial classifiers, MRF based on SVM preprocessing (SVM-MRF) [22], and the composite kernel-based SVM (CKSVM) [11], [58] are also included in the comparison. The RBF kernel is also used for the implementations of these three classifiers. In addition, 3-D wavelet feature-based classification is also implemented, where the 3-D redundant wavelet coefficients [56], [57] are extracted with Daubechies 1 filter and then RBF kernel SVM is run on the principal components of those wavelet coefficients (this classification is called 3-DW-DB1 hereafter). Several metrics are employed to quantitatively evaluate the

TABLE V

OVERALL ACCURACY (%) AND KAPPA COEFFICIENT (IN THE PARENTHESES) FOR THE AVIRIS KENNEDY SPACE CENTER IMAGE AS A FUNCTION OF THE TRAINING SAMPLE SIZE (% OF LABELED PIXELS). THE BEST RESULTS ARE REPORTED IN BOLD TYPEFACE

Sample size	1	2	3	4	5
DLRGF-LS	**89.40 (0.88)**	**96.84 (0.96)**	**97.31 (0.97)**	**97.32 (0.97)**	**98.73 (0.99)**
3DGF-LS	82.10 (0.80)	88.49 (0.87)	93.46 (0.92)	95.25 (0.94)	96.84 (0.96)
ReGF-LS	78.80 (0.76)	87.29 (0.86)	91.62 (0.91)	93.57 (0.93)	95.60 (0.95)
DLRGF-SVM	85.05 (0.83)	94.68 (0.94)	96.65 (0.96)	97.78 (0.98)	98.40 (0.98)
3DGF-SVM	82.72 (0.81)	91.70 (0.91)	93.82 (0.93)	94.83 (0.94)	95.25 (0.95)
ReGF-SVM	82.80 (0.81)	92.00 (0.91)	93.98 (0.93)	94.87 (0.95)	95.91 (0.95)
3DW-DB1	77.32 (0.75)	84.50 (0.83)	89.03 (0.88)	90.25 (0.89)	91.34 (0.90)
CKSVM	85.38 (0.84)	93.13 (0.92)	94.26 (0.94)	95.13 (0.95)	95.85 (0.95)
SVM-MRF	80.82 (0.79)	86.31 (0.85)	88.43 (0.87)	91.73 (0.91)	93.10 (0.92)
SVM	76.00 (0.73)	82.50 (0.80)	84.75 (0.83)	88.67 (0.87)	89.01 (0.88)

TABLE VI

INDIVIDUAL CLASSIFICATION ACCURACIES (%) FOR THE AVIRIS KENNEDY SPACE CENTER IMAGE, USING ONLY 2% OF THE LABELED PIXELS PER CLASS FOR TRAINING AND THE REMAINING LABELED PIXELS FOR TESTING

| Class	DLRGF-LS	3DGF-LS	ReGF-LS	DLRGF-SVM	3DGF-SVM	ReGF-SVM	3DW-DB1	CKSVM	SVM-MRF	SVM
1	98.21	100.00	80.80	96.88	94.64	95.98	48.21	84.38	55.80	40.63
2	65.85	59.35	42.28	70.33	51.63	43.90	62.20	67.89	64.63	53.66
3	82.17	80.89	79.62	80.89	71.97	77.71	69.43	75.16	81.53	76.43
4	99.20	78.40	83.20	99.20	79.60	80.40	55.20	93.60	23.20	62.40
5	100.00	95.57	94.77	96.24	97.58	98.39	86.58	99.19	97.72	93.83
6	100.00	90.13	95.19	93.42	95.95	97.22	85.57	96.96	59.49	65.32
7	97.56	68.90	77.64	98.58	85.37	94.72	83.54	92.68	100	89.84
8	99.76	80.09	73.46	97.16	98.10	84.36	79.86	96.21	99.76	80.09
9	100.00	100.00	99.80	100.00	99.61	100.00	98.43	97.64	100	89.98
10	100.00	100.00	100.00	100.00	100.00	100.00	99.89	100.00	100	99.45
11	89.08	68.49	65.55	79.41	69.33	79.41	76.89	70.17	75.21	67.22
12	99.02	95.37	93.66	90.98	98.05	96.83	97,32	97.56	86.83	88.05
13	100.00	100.00	100.00	100.00	100.00	100.00	77.45	82.35	100	86.27
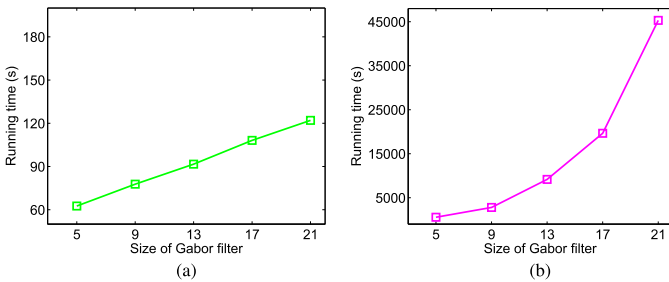
Fig. 7. Computational time (in seconds) as a function of the size of the Gabor filters for the AVIRIS Kennedy Space Center image. (a) DLRGF. (b) 3-DGF.

obtained classification results, including the overall accuracy [11], [59], the kappa coefficient [11], [14], and the individual class accuracy [11], [59]. In our experiments, we use a limited number of randomly selected labeled pixels as training samples, whereas the remaining pixels are used for testing.

Concerning the free parameters involved in the considered methods, we use the following settings.

1) For the DLRGF-LS, 3-DGF-LS, ReGF-LS, DLRGF-SVM, 3-DGF-SVM, and ReGF-SVM, the scale factors involved in the Gabor filtering are set by fivefold cross validation over the span [0.5, 5].

2) For the DLRGF-SVM, 3-DGF-SVM, ReGF-SVM, and 3-DW-DB1, the numbers of principal components fed into the SVM classifier are identical to allow a fair comparison, and are all assigned under 3-DGF by fivefold cross validation in the range [5, 100], with an interval of 5.

3) The regularization and kernel parameters of the SVM used in the DLRGF-SVM, 3-DGF-SVM, and ReGF-SVM (and those of the CKSVM and the 3-DW-DB1) are all set by fivefold cross validation of exponentially increasing sequences: $\{2^{-3}, 2^{-1}, \ldots, 2^{15}\}$ and $\{2^{-8}, 2^{-6}, \ldots, 2^3\}$, respectively. In addition, the weight parameter in the CKSVM and the scale to decomposed to is also set by fivefold cross validation in the range [0, 1].

All our experiments have been conducted using MATLAB R2013a in a desktop PC equipped with an Intel Core i7 CPU (at 3.6 GHz) and 32 GB of RAM.

A. Experiments With the OMIS Zaoyuan Hyperspectral Image

The first hyperspectral image used in experiments was collected by the operational modular imaging spectrometer (OMIS) over the Zaoyuan region, China, in 2001. The flight altitude was approximately 1200 m. The data set comprises 137×202 pixels and 80 bands covering the region: 0.4–1.7 μm, where the first 64 bands cover the region: 0.4–1.1 μm and the last 16 bands cover the region: 1.06–1.7 μm. The eight land-cover classes to be classified in the scene are illustrated in Table II.

Fig. 5 shows the computational time as a function of the size of the Gabor filters for the experiments conducted with the Zaoyuan image, which is the average over ten runs. The considered filtering methods include our DLRGF, the 3-DGF and its fast implementation (LRGF), and the ReGF. Fig. 5 clearly shows that DLRGF outperforms 3-DGF in terms of computational cost. For a relatively large size of the Gabor Filter ($L = 21$), DLRGF requires only slightly more than 10 s
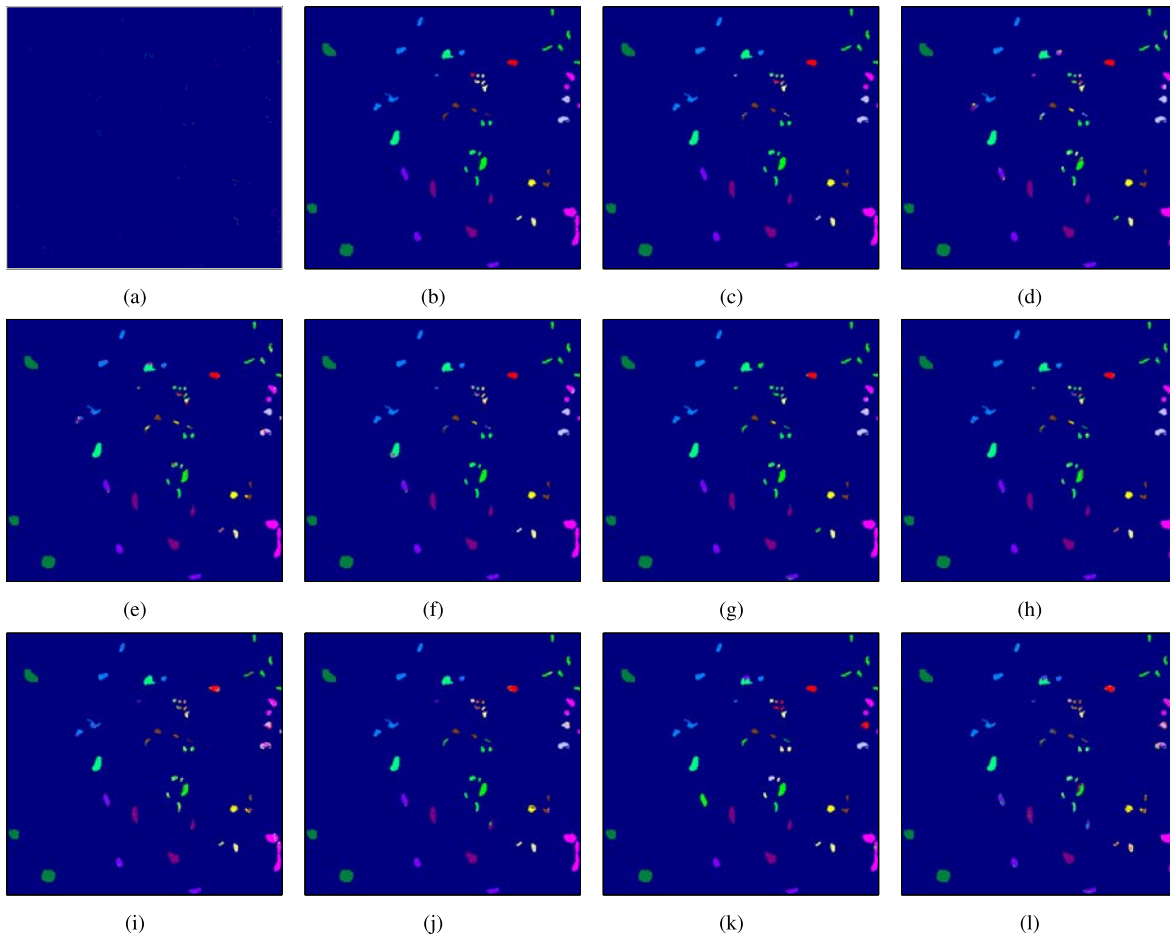
Fig. 8. Classification results for the AVIRIS Kennedy Space Center image using only 2% of the labeled pixels available per class for training. (a) Training samples. (b) Test samples. (c) DLRGF-LS. (d) 3-DGF-LS. (e) ReGF-LS. (f) DLRGF-SVM. (g) 3-DGF-SVM. (h) ReGF-SVM. (i) 3-DW-DB1. (j) CKSVM. (k) SVM-MRF. (l) SVM.
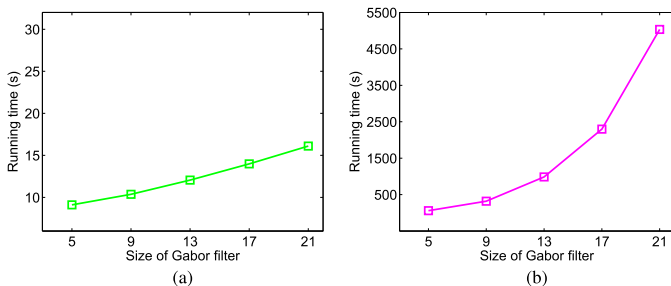

Fig. 9. Computational time (in seconds) as a function of the size of the Gabor filters for the AVIRIS Indian Pines image. (a) DLRGF. (b) 3-DGF.

in the considered computing environment, whereas the 3-DGF requires more than 3300 s (i.e., almost 1 h). Furthermore, it can also be observed that the computational time of DLRGF increases as a linear function of the size of the Gabor filter, whereas the computational time of the 3-DGF is of a high polynomial order. Fig. 5 also shows that the ReGF requires approximately half the time of the 3-DGF to complete the computation. This is expected, because the ReGF calculates only the real part of the 3-DGF. In addition, the computation time of LRGF is approximately eight times that of DLRGF. This is because the LRGF has to cope with eight subfilters, whereas DLRGF performs only one of the subfilters. These observations are consistent with our analyses of computation complexity in Section III and verify the superiority of our DLRGF in terms of computational efficiency.

Finally, we report the classification experiments conducted with this hyperspectral image. We randomly choose a very small percentage of labeled pixels in each class as training samples and used the rest of the pixels as test samples. Table III shows the overall classification accuracies and kappa statistic for different classifiers using different percentages of training samples selected per class, from 0.1% to 1%. As shown in Table III, for the same classifier, our DLRGF always results in the highest classification accuracies. On the other hand, Table IV shows the individual classification accuracies obtained by the different classifiers by considering 0.5% of the labeled samples for training and the rest of the labeled pixels for testing. From Table IV, we can conclude that DLRGF obtains better individual classification accuracies for all classes. For illustrative purposes, Fig. 6 shows the classification maps obtained by different classifiers [with Fig. 6(a) showing the training samples and Fig. 6(b) showing the test samples]. Visual interpretation of these maps reveals that the proposed DLRGF provides highly competitive results compared with the other tested methods.

B. Experiments With the AVIRIS Kennedy Space Center Image

The data set used in this experiment was acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) over the Kennedy Space Center in Florida. The image orig-
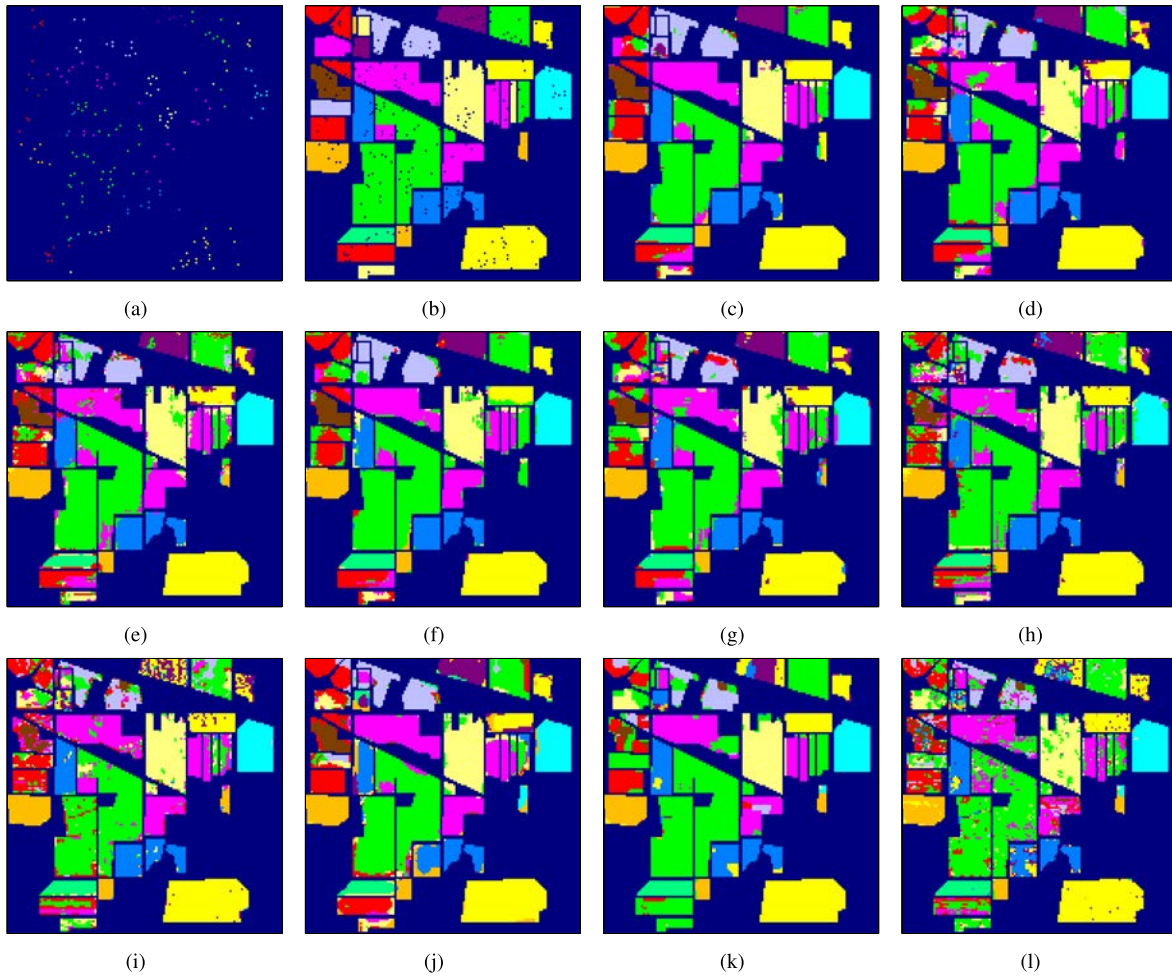
Fig. 10. Classification results for the AVIRIS Indian Pines image using only 2% of the labeled pixels available per class for training. (a) Training samples. (b) Test samples. (c) DLRGF-LS. (d) 3-DGF-LS. (e) ReGF-LS. (f) DLRGF-SVM. (g) 3-DGF-SVM. (h) ReGF-SVM. (i) 3-DW-DB1. (j) CKSVM. (k) SVM-MRF. (l) SVM.

TABLE VII

OVERALL ACCURACY (%) AND KAPPA COEFFICIENT (IN THE PARENTHESES) FOR THE AVIRIS INDIAN PINES IMAGE AS A FUNCTION OF THE TRAINING SAMPLE SIZE (% OF LABELED PIXELS). THE BEST RESULTS ARE HIGHLIGHTED IN BOLD TYPEFACE

Sample size	1	2	3	4	5
DLRGF-LS	**83.59 (0.81)**	**89.54 (0.88)**	**93.29 (0.92)**	**94.90 (0.94)**	**96.16 (0.96)**
3DGF-LS	77.12 (0.74)	84.04 (0.82)	90.70 (0.89)	93.17 (0.92)	94.92 (0.94)
ReGF-LS	73.42 (0.69)	83.27 (0.81)	90.11 (0.89)	91.83 (0.91)	93.75 (0.93)
DLRGF-SVM	77.96 (0.74)	86.69 (0.85)	92.07 (0.91)	94.28 (0.93)	96.29 (0.96)
3DGF-SVM	72.09 (0.67)	82.84 (0.80)	87.34 (0.85)	90.90 (0.90)	92.41 (0.91)
ReGF-SVM	71.36 (0.67)	81.35 (0.78)	84.60 (0.82)	88.11 (0.86)	90.04 (0.89)
3DW-DB1	74.39 (0.71)	78.94 (0.76)	83.14 (0.81)	85.17 (0.85)	87.08 (0.87)
CKSVM	74.22 (0.70)	83.07 (0.81)	88.49 (0.87)	90.89 (0.91)	92.67 (0.92)
SVM-MRF	74.26 (0.70)	79.56 (0.76)	83.04 (0.80)	86.00 (0.84)	87.42 (0.85)
SVM	70.92 (0.66)	73.81 (0.70)	76.58 (0.73)	78.94 (0.76)	80.42 (0.77)

inally contained 224 bands. After removing the noisy bands due to water absorption, 176 bands remained. Thirteen labeled classes, namely, Oak/broadleaf hammock, Cabbage palm/oak hammock, Slash pine, Cabbage palm hammock, Scrub, Cattail marsh, Mud flats, Graminoid marsh, Spartina marsh, Water, Willow swamp, Salt marsh, and Hardwood swamp are considered in an area comprising 428×450 pixels (additional details about the image and the classes are provided in [60]).

Fig. 7 shows the computational time our DLRGF and the conventional 3-DGF for the Kennedy Space Center data set. Similar to the experimental results reported for the Zaoyuan

image, the running time of our DLRGF is considerably less than that of the 3-DGF. For instance, for a Gabor filter of size $L = 21$, the DLRGF requires only approximately 2 min, whereas 3-DGF requires more than 450 000 s (i.e., 12.5 h). Furthermore, the cost of the DLRGF increases as an approximately linear function of the Gabor filter size, whereas the cost of the 3-DGF follows a high-order polynomial function. Due to space considerations and a similar trend exhibited by the ReGF and LRGF compared with the 3-DGF and DLRGF (as discussed in the experiment on OMIS Zaoyuan image), the running times of ReGF and LRGF are not presented here. This

TABLE VIII
INDIVIDUAL CLASSIFICATION ACCURACIES (%) FOR THE AVIRIS INDIAN PINES IMAGE, USING ONLY 2% OF THE
LABELED PIXELS PER CLASS FOR TRAINING AND THE REMAINING LABELED PIXELS FOR TESTING

Class	DLRGF-LS	3DGF-LS	ReGF-LS	DLRGF-SVM	3DGF-SVM	ReGF-SVM	3DW-DB1	CKSVM	SVM-MRF	SVM
1	87.33	79.36	80.07	81.57	80.71	80.00	81.07	80.93	67.19	62.42
2	80.05	70.75	78.58	78.34	61.32	57.65	62.18	87.03	47.86	54.58
3	96.51	91.27	76.42	88.21	79.04	64.63	48.03	79.48	25.33	34.06
4	83.37	80.90	86.04	82.14	84.19	83.37	93.43	85.63	89.12	83.78
5	96.45	90.44	91.80	91.39	92.62	92.08	95.63	75.96	90.03	81.28
6	100.00	100.00	99.58	97.91	91.65	88.31	99.37	95.41	99.58	98.96
7	81.54	77.74	75.32	74.16	71.84	75.63	79.43	75.95	71.73	67.09
8	93.63	91.27	89.00	96.24	91.73	88.88	80.15	85.28	99.42	85.15
9	70.88	65.72	59.57	68.72	58.74	57.24	45.59	73.54	51.41	48.92
10	94.20	78.74	71.98	97.58	94.20	78.74	86.96	83.09	99.03	97.58
11	99.36	91.72	90.38	94.72	92.98	95.03	92.82	89.35	99.05	94.48
12	83.33	72.84	71.77	71.24	76.34	75.54	41.66	77.15	29.57	23.39

experiment clearly suggests the superiority of our DLRGF in terms of computational efficiency, as it was the case in the experiment with the OMIS Zaoyuan image.

The classification results for the AVIRIS Kennedy Space Center data are summarized in Tables V and VI, and the classification maps are shown in Fig. 8. In this case, we considered sample sizes comprising 1%–5% of the available labeled pixels per class for training, and reported the individual classification accuracies for the case in which 2% of the available labeled pixels per class were used for training. Similar to the results reported for the OMIS Zaoyuan image, the proposed DLRGF approach achieved the best results in all cases.

C. Experiments With the AVIRIS Indian Pines Image

The Indian Pines data set was acquired by the AVIRIS sensor over northwestern Indiana. The data set include 220 spectral bands and 145×145 pixels. There are a total of 16 land-cover classes in the scene; 40 water absorption and noisy bands were removed, and the remaining 180 bands (6–100th, 112–147th, and 167–215th) were used for classification purposes. Since some classes have very few samples, 12 land-cover classes were considered in our experiments: Corn-notill, Corn-min, Corn, Grass/Pasture, Grass/Trees, Hay-windrowed, Soybeans-notill, Soybeans-min, Soybean-clean, Wheat, Woods, and Bldg-Grass-Tree-Drives (details of these classes can be found in [8] and [14]). The results obtained with the AVIRIS Indian Pines are shown in Figs. 9 and 10 and in Tables VII and VIII. As with the previous experiments, the proposed DLRGF outperforms the other tested methods in terms of both computation time and classification accuracy.

V. CONCLUSION

In this paper, we have developed a fast DLRGF approach for spectral–spatial classification of remotely sensed hyperspectral images. The proposed approach represents an efficient strategy for selecting spatial–spectral features and offers significant novelty in the way Gabor filtering is adapted to the special properties of hyperspectral images, which require spatial smoothness and spectral differential information to enhance class separability. The main properties of the DLRGF are: 1) its high computational efficiency and 2) its capacity to provide spectral–spatial features with higher discriminability. Specifically, in the proposed DLRGF, the standard 3-DGF is first derived as the superposition of eight rank-1 subfilters, so

that computational complexity reduces from cubic to linear order. Furthermore, we choose only the subfilter associated with both the spatial low-pass and spectral bandpass characteristics to perform Gabor filtering, thus increasing computational efficiency and reducing the interference from subfilters that are not highly discriminative in the context of hyperspectral images. To quantitatively analyze the feature discriminability under various filtering conditions, we designed an overall classification risk bound that confirmed that the features related to the spatial low-pass and the spectral bandpass Gabor processing result in higher discrimination ability for classification purposes than by the other components. Another important advantage of using only the most discriminative component is that the DLRGF can avoid the interference from the low discriminating components, while other strategies produce lower classification results due in part to this interference. Our experimental results, conducted with several hyperspectral images, confirm that our newly developed DLRGF can yield high computation efficiency and classification results that are comparable or better than those obtained by other state-of-the-art classifiers. It is worth mentioning that the low-pass spatial filtering may degrade the signals at class boundaries if the spatial filter is fixed for all pixels, which implies the homogeneous spatial dependency across the whole hyperspectral image that is in fact disrupted along the class boundaries. In the future, we will pursue incorporating an additional boundary-sensitive factor into our Gabor filtering.

REFERENCES

[1] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, no. 4704, pp. 1147–1153, Jun. 1985.

[2] A. F. H. Goetz, "Three decades of hyperspectral remote sensing of the earth: A personal view," *Remote Sens. Environ.*, vol. 113, no. 1, pp. 5–16, Sep. 2009.

[3] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 12–16, Jan. 2002.

[4] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 17–28, Jan. 2002.

[5] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson, "Advances in hyperspectral image classification: Earth monitoring with statistical learning methods," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 45–54, Jan. 2014.

[6] A. Plaza *et. al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, no. 1, pp. S110–S122, Sep. 2009.

[7] G. Cheng, J. Han, L. Guo, Z. Liu, S. Bu, and J. Ren, "Effective and efficient midlevel visual elements-oriented land-use classification using VHR remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4238–4249, Aug. 2015.

[8] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral–spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.

[9] D. Lunga, S. Prasad, M. M. Crawford, and O. Ersoy, "Manifold-learning-based feature extraction for classification of hyperspectral data: A review of advances in manifold learning," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 55–66, Jan. 2014.

[10] Y. Qian, F. Yao, and S. Jia, "Band selection for hyperspectral imagery using affinity propagation," *IET Comput. Vis.*, vol. 3, no. 4, pp. 213–222, Dec. 2009.

[11] L. He, Y. Li, X. Li, and W. Wu, "Spectral–spatial classification of hyperspectral images via spatial translation-invariant wavelet-based sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2696–2712, May 2015.

[12] L. Bruzzone and C. Persello, "A novel context-sensitive semisupervised SVM classifier robust to mislabeled training samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2142–2154, Jul. 2009.

[13] C.-H. Li, B.-C. Kuo, C.-T. Lin, and C.-S. Huang, "A spatial–contextual support vector machine for remotely sensed image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 784–799, Mar. 2012.

[14] 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.

[15] J. Li, H. Zhang, Y. Huang, and L. Zhang, "Hyperspectral image classification by nonlocal joint collaborative representation with a locally adaptive dictionary," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3707–3719, Jun. 2014.

[16] L. Fang, S. Li, X. Kang, and J. A. Benediktsson, "Spectral–spatial hyperspectral image classification via multiscale adaptive sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 12, pp. 7738–7749, Dec. 2014.

[17] M. A. Veganzones, G. Tochon, M. Dalla Mura, A. J. Plaza, and J. Chanussot, "Hyperspectral image segmentation using a new spectral unmixing-based binary partition tree representation," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3574–3589, Aug. 2014.

[18] P. Zhong and R. Wang, "Learning conditional random fields for classification of hyperspectral images," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1890–1907, Jul. 2010.

[19] X. Kang, S. Li, and J. A. Benediktsson, "Spectral–spatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, May 2014.

[20] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral–spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 809–823, Mar. 2012.

[21] J. Li, J. 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.

[22] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM- and MRF-based method for accurate classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 736–740, Oct. 2010.

[23] W. Li, S. Prasad, and J. E. Fowler, "Hyperspectral image classification using Gaussian mixture models and Markov random fields," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 153–157, Jan. 2014.

[24] B. Zhang, S. Li, X. Jia, L. Gao, and M. Peng, "Adaptive Markov random field approach for classification of hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 973–977, Sep. 2011.

[25] L. Sun, Z. Wu, J. Liu, L. Xiao, and Z. Wei, "Supervised spectral–spatial hyperspectral image classification with weighted Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1490–1503, Mar. 2015.

[26] G. Zhang, X. Jia, and J. Hu, "Superpixel-based graphical model for remote sensing image mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 5861–5871, Nov. 2015.

[27] 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.

[28] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.

[29] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles," *Pattern Recognit.*, vol. 37, no. 6, pp. 1097–1116, Jun. 2004.

[30] A. Plaza, P. Martinez, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 466–479, Mar. 2005.

[31] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, Nov. 2008.

[32] M. Dalla Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 542–546, May 2011.

[33] P. Ghamisi, J. A. Benediktsson, and J. R. Sveinsson, "Automatic spectral–spatial classification framework based on attribute profiles and supervised feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5771–5782, Sep. 2014.

[34] J. Xia, M. Dalla Mura, J. Chanussot, P. Du, and X. He, "Random subspace ensembles for hyperspectral image classification with extended morphological attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 4768–4786, Sep. 2015.

[35] N. Falco, J. A. Benediktsson, and L. Bruzzone, "Spectral and spatial classification of hyperspectral images based on ICA and reduced morphological attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 6223–6240, Nov. 2015.

[36] F. Tsai and J. S. Lai, "Feature extraction of hyperspectral image cubes using three-dimensional gray-level cooccurrence," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3504–3513, A Jun. 2013.

[37] Z. He, Q. Wang, Y. Shen, and M. Sun, "Kernel sparse multitask learning for hyperspectral image classification with empirical mode decomposition and morphological wavelet-based features," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5150–5163, Aug. 2014.

[38] D. Gabor, "Theories of Communication," *J. Inst. Elect. Eng.*, vol. 93, pp. 429–457, 1946.

[39] L. Shen and L. Bai, "3-D Gabor wavelets for evaluating SPM normalization algorithm," *Med. Image Anal.*, vol. 12, no. 3, pp. 375–383, Jun. 2008.

[40] Z. Zhu, S. Jia, S. He, Y. Sun, Z. Ji, and L. Shen, "Three-dimensional Gabor feature extraction for hyperspectral imagery classification using a memetic framework," *Inf. Sci.*, vol. 298, pp. 274–287, Mar. 2015.

[41] J.-K. Kamarainen, V. Kyrki, and H. Kalviainen, "Invariance properties of Gabor filter-based features-overview and applications," *IEEE Trans. Image Process.*, vol. 15, no. 5, pp. 1088–1099, May 2006.

[42] W.-S. Chen, J. Huang, J. Zou, and B. Fang, "Wavelet-face based subspace LDA method to solve small sample size problem in face recognition," *Int. J. Wavelets Multiresolution Inf. Process.*, vol. 7, no. 2, pp. 199–214, 2009.

[43] W.-S. Chen, W. Wang, J.-W. Yang, and Y. Y. Tang, "Supervised regularization locality-preserving projection method for face recognition," *Int. J. Wavelets Multiresolution Inf. Process.*, vol. 10, no. 6, p. 1250053, 2012.

[44] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 1, pp. 55–73, Jan. 1990.

[45] F. Bianconi and A. Fernández, "Evaluation of the effects of Gabor filter parameters on texture classification," *Pattern Recognit.*, vol. 40, no. 12, pp. 3325–3335, Dec. 2007.

[46] J. F. Khan, R. R. Adhami, and S. M. A. Bhuiyan, "A customized Gabor filter for unsupervised color image segmentation," *Image Vis. Comput.*, vol. 27, no. 4, pp. 489–501, Mar. 2009.

[47] J. Yi, C. Lee, and J. Kim, "Online signature verification using temporal shift estimated by the phase of Gabor filter," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 776–783, Feb. 2005.

[48] T. Chen, X. Wang, S. Chung, D. Metaxas, and L. Axel, "Automated 3-D motion tracking using Gabor filter bank, robust point matching, and deformable models," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 1–11, Jan. 2010.

[49] T. C. Bau, S. Sarkar, and G. Healey, "Hyperspectral region classification using a three-dimensional Gabor filterbank," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3457–3464, Sep. 2010.

[50] L. Shen and S. Jia, "Three-dimensional Gabor wavelets for pixel-based hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 5039–5046, Dec. 2011.

[51] S. Jia, L. Shen, and Q. Li, "Gabor feature-based collaborative representation for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 1118–1129, Feb. 2015.

[52] P.-F. Hsieh and D. Landgrebe, "Lowpass filter for increasing class separability," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 1998, pp. 2691–2693.

[53] F. Tsai and W. D. Philpot, "A derivative-aided hyperspectral image analysis system for land-cover classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 2, pp. 416–425, Feb. 2002.

[54] A. Ertürk, M. K. Güllü, and S. Ertürk, "Hyperspectral image classification using empirical mode decomposition with spectral gradient enhancement," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2787–2798, May 2013.

[55] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. San Diego, CA, USA: Academic, 1990, ch. 3.

[56] X. Zhou, S. Prasad, and M. M. Crawford, "Wavelet-domain multiview active learning for spatial–spectral hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4047–4059, Sep. 2016, doi: 10.1109/JSTARS.2016.2552998.

[57] Y. Qian, M. Ye, and J. Zhou, "Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2276–2291, Aug. 2012.

[58] G. Camps-Valls, L. Gomez-Chova, J. Muñoz-Marí, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.

[59] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[60] J. Ham, Y. Chen, M. M. Crawford, and J. Ghosh, "Investigation of the random forest framework for classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 492–501, Mar. 2005.

[61] G. Camps-Valls *et al.*, "Robust support vector method for hyperspectral data classification and knowledge discovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1530–1542, Jul. 2004.

Lin He (S'05–M'12) received the B.S. degree from the Xi'an Institute of Technology, Xi'an, China, in 1995, the M.S. degree from Chongqing University, Chongqing, China, in 2003, and the Ph.D. degree from Northwestern Polytechnical University, Xi'an, in 2007.

Since 2007, he has been with the College of Automation Science and Engineering, South China University of Technology, Guangzhou, China, where he is currently an Associate Professor. His current research interests include statistical pattern recognition, hyperspectral image processing, and biomedical signal processing.

Jun Li (M'12) received the B.S. degree in geographic information systems from Hunan Normal University, Changsha, China, in 2004, the M.E. degree in remote sensing from Peking University, Beijing, China, in 2007, and the Ph.D. degree in electrical engineering from the Instituto de Telecomunicações, Instituto Superior Técnico (IST), Universidade Técnica de Lisboa, Lisbon, Portugal, in 2011.

From 2007 to 2011, she was a Marie Curie Research Fellow with the Departamento de Engenharia Electrotcnica e de Computadores and the Instituto de Telecomunicações, IST, Universidade Técnica de Lisboa, in the framework of the European Doctorate for Signal Processing. She has also been actively involved in the Hyperspectral Imaging Network, a Marie Curie Research Training Network involving 15 partners in 12 countries and intended to foster research, training, and cooperation on hyperspectral imaging at the European level. Since 2011, she has been a Post-Doctoral Researcher with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, Cáceres, Spain. Currently, she is a Professor with Sun Yat-Sen University, Guangzhou, China. Her current research interests include hyperspectral image classification and segmentation, spectral unmixing, signal processing, and remote sensing.

Dr. Li is an Associate Editor of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and has been a Reviewer for several journals.

Antonio Plaza (M'05–SM'07–F'15) was born in Cáceres, Spain, in 1975.

He is an Associate Professor (with accreditation for Full Professor) with the Department of Technology of Computers and Communications, University of Extremadura, Cáceres, where he is the Head of the Hyperspectral Computing Laboratory, one of the most productive research groups involved in remotely sensed hyperspectral data processing worldwide. He has been the Advisor of 12 Ph.D. dissertations and more than 30 M.Sc. dissertations. He was the Coordinator of the Hyperspectral Imaging Network, a European project with a total funding of € 2.8 million. He has authored more than 500 publications, including 152 journal papers (more than 100 in IEEE journals), 22 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) and has guest edited nine special issues on hyperspectral remote sensing for different journals. His current research interests include hyperspectral data processing and parallel computing of remote sensing data.

Dr. Plaza was recognized as the Best Reviewer for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS in 2009 and for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2010, a journal for which he was an Associate Editor from 2007 to 2012. He was a recipient of the 2013 Best Paper Award of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS), and of the most highly cited paper in 2005–2010 in the *Journal of Parallel and Distributed Computing*. He received the Best Paper Awards at the IEEE International Conference on Space Technology and the IEEE Symposium on Signal Processing and Information Technology. He was a recipient of the Best Ph.D. Dissertation Award at the University of Extremadura; a recognition was also received by six of his Ph.D. students. He was the Director of the Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) from 2011 to 2012, and has been the President of the Spanish Chapter of the IEEE GRSS since 2012. He has served as a Proposal Evaluator for the European Commission, the National Science Foundation, the European Space Agency, the Belgium Science Policy, the Israel Science Foundation, and the Spanish Ministry of Science and Innovation. He has reviewed more than 500 manuscripts for over 50 different journals. He is also an Associate Editor of the IEEE ACCESS, and was an Editorial Board Member of the IEEE GEOSCIENCE AND REMOTE SENSING NEWSLETTER from 2011 to 2012, and the IEEE GEOSCIENCE AND REMOTE SENSING MAGAZINE in 2013. He was also a member of the Steering Committee of the IEEE JSTARS. He is currently the Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING journal.

Yuanqing Li (M'05–SM'14) received the B.S. degree in applied mathematics from Wuhan University, Wuhan, China, in 1988, and the M.S. degree in applied mathematics and the Ph.D. degree in control theory and applications from South China Normal University, Guangzhou, China, 1994 and 1997, respectively.

Since 1997, he has been with the South China University of Technology, Guangzhou, where he became a Full Professor in 2004. From 2002 to 2004, he was a Researcher at the Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Wakō, Japan. From 2004 to 2008, he was a Research Scientist at the Laboratory for Neural Signal Processing, Institute for Inforcomm Research, Singapore. He has authored more than 70 scientific papers in journals and conference proceedings. He is currently a Yangtse River Scholar Chair Professor with the Ministry of Education, China. His current research interests include blind signal processing, sparse representation, machine learning, BCI, and brain data analysis.