Image Anal Stereol 2023;42:93-99 doi: 105566/ias.2928
Original Research Paper

NOISE ROBUST HYPERSPECTRAL IMAGE CLASSIFICATION WITH MNF-BASED EDGE PRESERVING FEATURES

GUANG YI CHEN✉,[1], ADAM KRZYZAK[1] AND SHEN-EN QIAN[2]

[1]Department of Computer Science and Software Engineering, Concordia University, Montreal, QC, Canada H3G 1M8; [2] Space Science and Technology, Canadian Space Agency, St-Hubert, QC, Canada J3Y 8Y9.
e-mail: guang_c@cse.concordia.ca, krzyzak@cse.concordia.ca, shen-en.qian@canada.ca.
(Received May 18, 2023; accepted June 16, 2023)

ABSTRACT

Hyperspectral image (HSI) classification is an important topic in remote sensing. In this paper, we improve the principal component analysis (PCA)-based edge preserving features (EPFs) for HSI classification. We select to use minimum noise fraction (MNF) instead of PCA to reduce the dimensionality of the hyperspectral data cube to be classified. We keep all the rest steps from the PCA-based EPFs for HSI classification. Since MNF can preserve fine features of a HSI data cube better than PCA, our new method can outperform PCA-EPFs for HSI classification significantly. Experimental results show that our new method performs better than the PCA-based EPFs under such noisy environment as Gaussian white noise and shot noise. In addition, our MNF+EPFs outperform the PCA+EPFs even when no noise is added to the HSI data cubes for most testing cases, which is very desirable in remote sensing.

Keywords: Edge preserving features (EPFs), hyperspectral image (HSI) classification, minimum noise fraction (MNF), principal component analysis (PCA), support vector machine (SVM).

INTRODUCTION

Hyperspectral imaging (HSI) generates vast amount of spectral and spatial information, which allows for a better characterization and exploitation of the Earth surface by combining the rich spectral and spatial information. Nevertheless, HSI introduces major challenges for supervised classification methods due to the high dimensionality of the data and the limited availability of training samples. These problems along with the high intraclass variability (and interclass similarity) may compromise the effectiveness of HSI classifiers.

We briefly review several published methods for HSI classification in recent years here. Liu *et al.* (2020) developed a novel approach for dimensionality reduction of hyperspectral images based on improved spatial-spectral weight manifold embedding and classifications. Zhou *et al.* (2019) analyzed hyperspectral image classification by using spectral-spatial long- and short-term memory (LSTM). Chen *et al.* (2013) studied hyperspectral image classification by means of kernel sparse representation. Melgani and Bruzzone (2004) investigated the classification of hyperspectral remote sensing images with support vector machines. Fauvel *et al.* (2012) worked on a spatial-spectral kernel-based approach for the classification of remote-sensing images. Camps-Valls and Bruzzone (2005) investigated kernel-based methods for HSI classification. Li *et al.* (2013) presented a generalized composite kernel framework for HSI classification. Chen *et al.* (2011) analyzed HSI classification by using dictionary-based sparse representation. Li *et al.* (2013) studied on spectral-spatial classification of hyperspectral data using loopy belief propagation and active learning. Kang *et al.* (2014) worked on spectral-spatial HSI classification with edge-preserving filtering. Cheng *et al.* (2016) analyzed semi-supervised HSI classification via discriminant analysis and robust regression. Chen *et al.* (2014) investigated deep learning-based classification of hyperspectral data with good results. Kang *et al.* (2017) proposed a novel principal component analysis (PCA)-based edge preserving features (PCA-EPFs) method for HSI classification. First, the standard EPFs are constructed by applying edge-preserving filters with different parameter settings to the considered image, and the resulting EPFs are stacked together. Then, the spectral dimension of the stacked EPFs is reduced by the PCA, which not only can represent the EPFs in the mean square sense but also highlight the separability of pixels in the EPFs. Finally, the resulting PCA-EPFs are classified by a support vector machine (SVM) classifier.

In this paper, we improve the PCA-based edge preserving features (EPF) method for HSI classification. We select minimum noise fraction (MNF) instead of

PCA to reduce the dimensionality of the input data cube. We then choose the MNF to represent the stacked EPFs and keep all the rest steps as same as the PCA-EPFs for HSI classification. Because MNF retains more fine features of a HSI data cube than PCA, our new method can perform better than PCA-EPFs for HSI classification. Experiments demonstrate that our new method performs better than PCA-EPFs for both Gaussian white noise (GWN) and shot noise significantly. More importantly, our MNF+EPFs outperform the PCA+EPFs even when no noise is added to the HSI data cubes for most testing cases, which is very desirable in remote sensing.

The organization of this paper is as follows. Section II proposes a novel method for HSI classification by using MNF-based EPFs. Section III performs experiments to demonstrate the effectiveness of our proposed method. Finally, Section IV draws the conclusions of the paper and proposes future research topics.

THE PROPOSED METHOD

In this section, we briefly review the MNF, the PCA, the SVMs, and the edge-preserving features (EPF). We provide their definitions and their basic properties here first, then introduce our method proposed in this paper.

The MNF (Green *et al.* 1988) transform is a linear transform that is made from two principal components (PCs) analysis rotations. The first rotation finds the PCs of the noise covariance matrix to decorrelate and rescale the noise in the data (a process known as noise whitening), resulting in transformed data in which the noise has unit variance and no band-to-band correlations. The second rotation picks the PCs derived from the original image data after they have been noise-whitened by the first rotation and rescaled by the noise standard deviation. Because further spectral processing will take place, the inherent dimensionality of the data is determined by examining the final eigenvalues and the associated images. We can divide the data space into two parts: one part associated with large eigenvalues and coherent eigenimages, and a complementary part with near-unity eigenvalues and noise-dominated images. Using only the coherent portions separates the noise from the data, thus improving spectral processing results.

PCA (Jolliffe 2002) is a popular technique to reduce the dimensionality of a data set, which can increase interpretability and reduce information loss. The PCA builds new uncorrelated variables that can successively maximize variance. Nevertheless, the PCA requires to solve an eigenvalue/eigenvector problem. The PCA is adaptive since variants of the technique have been developed, which are suited for various data types and structures. The PCA projects every data point onto only the first a few principal components to obtain lower-dimensional data while preserving as much of the data's variation as possible. The first PCs can equivalently be defined as a direction that maximizes the variance of the projected data. The rest PCs can be taken as a direction orthogonal to the first a few PCs that maximize the variance of the projected data.

SVMs (Cortes and Vapnik 1995) are supervised learning method for classification, regression, and outlier's detection. They are the most robust prediction methods that are based on the statistical learning framework. For a given set of training samples, the SVM constructs a model which assigns new samples to one category or the other, making it a robust binary linear classifier. The SVM is a representation of the samples as points in space, mapped so that the samples of the separate categories are divided by a clear gap that is as wide as possible. New samples are then mapped into that same space and predicted to belong to a category based on the side of the gap on which they fall.

The EPF (Kang *et al.* 2014) is a popular method that can smooth away textures and noise and retain significant parts of the image content, typically edges, lines, or other details which are important for the interpretation of the HSI images. Due to this advantage, the EPF receives reasonable attention from the image processing and computer vision research communities for the last two decades. Recently, more powerful EPFs have been successfully applied to hyperspectral remote sensing applications.

Kang *et al*. (2017) proposed a PCA-EPFs classification, which consist of the following steps :

1. The standard multi-parametric EPFs are constructed and stacked together by performing edge preserving features on the dimension-reduced (DR) hyperspectral data.
2. The PCA is used to represent the stacked EPFs in the mean square sense and highlight the spectral differences.
3. The SVM is used for the classification of the PCA-EPFs.

Experimental results reported in Kang *et al.* (2017) show that, by comparing with the standard EPFs and other recently published HSI classification techniques, their proposed PCA-EPFs can obtain much higher classification accuracies on three real hyperspectral data sets, especially when the number of training samples is limited.

In this paper, we improve the PCA-EPFs method for HSI classification from two aspects. First, we use MNF to reduce the dimensionality of the input HSI data cube instead of PCA. Since MNF can preserve more fine features of a HSI data cube than PCA, our method can outperform PCA-EPFs for HSI classification significantly. Second, we choose the MNF to represent the stacked EPFs in the mean square sense and highlight the spectral differences. Experiments show that our new method is better than PCA-EPFs under the noisy environment for all three testing HSI data cubes. Furthermore, our MNF+EPFs outperform the PCA+EPFs even when no noise is added to the HSI data cubes for most testing cases, which is very desirable in remote sensing.

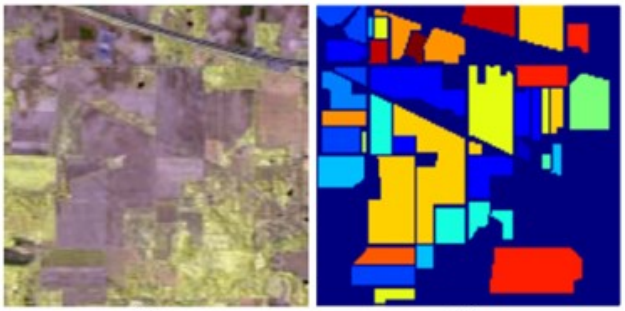
Our proposed MNF-EPFs can be summarized as the following steps:

1. By performing edge preserving features on the DR hyperspectral data, the standard multiparametric EPFs are constructed and stacked together.
2. We choose the MNF to represent the stacked EPFs in the mean square sense and highlight the spectral differences.
3. We select the SVM for the classification of the MNF-EPFs.

EXPERIMENTAL RESULTS

We test three HSI data cubes for HSI classification, which can be described as follows.

a. Indian Pines. This data cube was acquired by the airborne visible/infrared imaging spectrometer (AVIRIS) sensor over the Indian Pine test site in northwestern Indiana, USA, on June 12, 1992. This data cube has 145×145 pixels with 200 spectral bands. Fig. 1 depicts the false-colour composite image and the ground-truth map. Table 1 tabulates ground truth classes and the pixel number for every class in this data cube.



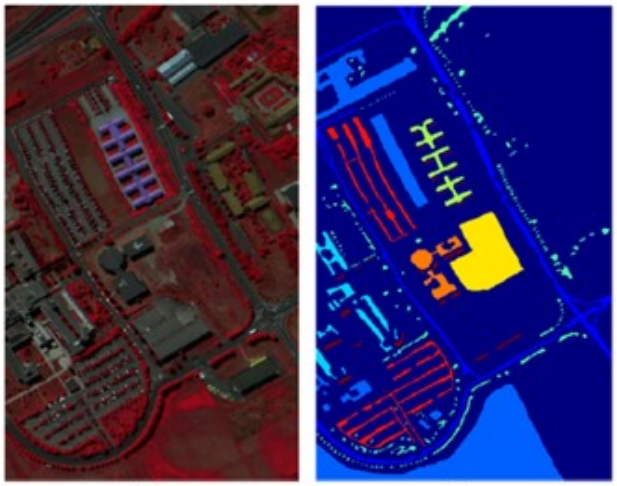
(a) (b)

Fig. 1. *Indian Pines data cube. (a) False-color composite image (b) Ground-truth map with 16 classes.*

Table 1. *Ground truth classes and the total pixel number for each class in Indian Pines data cube.*

No	Class Names	Total Samples
C1	Alfalfa	46
C2	Corn notill	1428
C3	Corn mintill	830
C4	Corn	237
C5	Grass pasture	483
C6	Grass trees	730
C7	Grass pasture mowed	28
C8	Hay windrowed	478
C9	Oats	20
C10	Soybean notill	972
C11	Soybean mintill	2455
C12	Soybean clean	593
C13	Wheat	205
C14	Woods	1265
C15	Buildings Grass Trees Drives	386
C16	Stone Steel Towers	93

b. Pavia University. This data cube was acquired by the ROSIS sensor during a flight campaign over Pavia, northern Italy, on July 8, 2002. This data cube has 610×340 pixels with 115 spectral bands. Fig. 2 depicts a 3-band false-colour composite image and the ground truth map. There are 9 classes of land covers, which is tabulated in Table 2.



(a) (b)

Fig. 2. *Pavia University data cube. (a) False-color composite image (b) Ground- truth map with 9 classes.*

Table 2. *Ground truth classes and the total pixel number for each class in Pavia University data cube.*

No	Class Names	Total Samples
C1	Asphalt	6631
C2	Meadows	18649
C3	Gravel	2099
C4	Trees	3064
C5	Painted metal sheets	1345
C6	Bare Soil	5029
C7	Bitumen	1330
C8	Self-Blocking Bricks	3682
C9	Shadows	947

c. Salinas. This data cube was collected by the AVIRIS sensor over Salinas Valley, California. This data cube has 512×217 pixels with 204 spectral bands. Fig. 3 demonstrates a false-colour composite image and the ground truth map. Table 3 shows ground truth classes and the total pixel numbers for all classes.

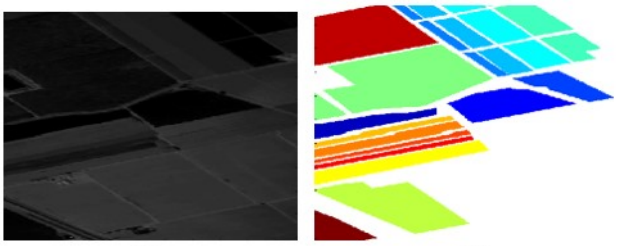


Fig. 3. *Salinas data cube. (a) False-color composite image (b) Ground-truth map with 16 classes.*

Table 3. *Ground truth classes and the total pixel number for each class in Pavia University data cube.*

No	Class Names	Total Samples
C1	Broccoli green weeds 1	2009
C2	Broccoli green weeds 2	3726
C3	Fallow	1976
C4	Fallow rough plow	1394
C5	Fallow smooth	2678
C6	Stubble	3959
C7	Celery	3579
C8	Grapes untrained	11271
C9	Soil vinyard develop	6203
C10	Corn senesced green weeds	3278
C11	Lettuce romaine 4wk	1068
C12	Lettuce romaine 5wk	1927
C13	Lettuce romaine 6wk	916
C14	Lettuce romaine 7wk	1070
C15	Vinyard untrained	7268
C16	Vinyard vertical trellis	1807

We test our proposed method when Gaussian white noise (GWN) is added to the noise-free HSI data cubes with noise standard deviation σn=100, 200, 300, 400, and 500. A noisy HS data cube can be generated by adding GWN to a noise-free data cube.

$$B=A+\sigma_n Z, \quad (1)$$

where A is the noise-free data cube, B is the simulated noisy data cube, Z obeys normal distribution with zero mean and unit variance, and □n is the noise standard deviation. We also conduct experiments with shot noise added to the noise free data cube using the following Matlab command:

$$B(i) = \text{imnoise}(A(i),'\text{poisson}'), \quad (2)$$

where A(i) is the i=th noise-free spectral band image of the HSI data cube and B(i) is the noisy spectral band image. Shot noise does not have a noise level in the generation of the noisy data cube, which is totally different from the GWN.

The overall classification accuracy and standard deviation (STD) are shown in Tables 4-6 for the Indian Pines data cube, the Pavia University data cube, and the Salinas data cube, respectively. We randomly select 2%, 4%, 6%, 8%, and 10% of pixels as training data set and the rest of pixels as testing data set for each class. We run the two methods for five times and take the mean and STD in our experiments. The best classification results are highlighted in bold font. Our proposed method in this paper improves upon PCA-EPFs significantly in classification accuracies for all noise levels in our experiments. Our new method outperforms the PCA-EPFs even when no noise is added to the HSI data cubes for most testing cases, which is very desirable. Fig. 4 shows the original noise-free band (#50) from the Indian Pines data set and the noisy bands with different noise levels whereas Fig.5 displays the noise-free band and that with shot noise added.

Table 4. *Overall classification accuracy (%) and STD for the Indian Pines data set (Overall Accuracy ± STD). The best results are highlighted in bold font.*

Method	Training Sample Percentage	No Noise	GWN (σ_n) 100	200	300	400	500	Shot Noise
PCA	2	**97.71±1.56**	90.12±0.80	85.16±1.28	84.23±1.20	84.01±1.50	82.81±0.63	42.27±1.42
	4	**97.70±0.42**	93.96±0.91	93.16±0.45	91.06±0.74	91.40±1.16	89.73±0.54	53.42±0.41
	6	**98.39±0.34**	96.45±0.20	94.58±0.68	94.30±0.56	92.75±0.58	93.07±0.43	59.53±0.84
	8	98.62±0.40	97.43±0.70	95.98±0.50	94.88±0.46	95.49±0.39	94.07±0.44	64.45±0.51
	10	**99.02±0.16**	97.74±0.34	96.79±0.57	96.13±0.45	96.14±0.23	95.68±0.30	67.15±0.98
MNF	2	93.39±0.97	**91.16±1.11**	**90.62±1.26**	**90.48±1.55**	**89.79±1.97**	**88.32±1.24**	**72.25±2.79**
	4	97.01±0.82	**95.78±0.98**	**95.56±0.71**	**95.03±0.42**	**94.45±0.89**	**95.13±0.34**	**85.09±2.23**
	6	98.21±0.29	**97.50±0.16**	**97.27±0.41**	**96.68±0.38**	**96.34±0.55**	**96.70±0.51**	**91.46±1.05**
	8	**98.75±0.46**	**97.77±0.29**	**98.13±0.36**	**97.44±0.49**	**97.05±0.32**	**97.05±0.70**	**93.56±0.58**
	10	98.93±0.28	**98.55±0.22**	**98.38±0.18**	**98.10±0.29**	**97.85±0.18**	**98.16±0.18**	**94.55±0.50**

Table 5. *Overall classification accuracy (%) and STD for the Pavia University data set (Overall Accuracy ± STD). The best results are highlighted in bold font.*

Method	Training Sample Percentage	No Noise	GWN (σ_n) 100	200	300	400	500	Shot Noise
PCA	2	98.63±0.43	98.57±0.30	98.59±0.46	98.43±0.23	98.30±0.42	98.31±0.17	88.39±1.73
	4	99.41±0.09	99.48±0.13	99.38±0.18	99.38±0.12	99.21±0.17	99.00±0.16	92.32±1.02
	6	99.64±0.07	99.64±0.09	99.44±0.05	99.45±0.06	99.40±0.14	99.37±0.11	92.92±0.89
	8	99.68±0.09	99.73±0.08	99.70±0.08	99.62±0.06	99.57±0.07	99.56±0.09	94.61±0.54
	10	99.78±0.06	99.70±0.05	99.74±0.06	99.69±0.05	99.69±0.04	99.59±0.09	94.68±0.94
MNF	2	**98.86±0.29**	**98.75±0.27**	**98.83±0.42**	**98.82±0.24**	**98.74±0.13**	**99.08±0.26**	**90.90±0.45**
	4	**99.48±0.17**	**99.52±0.08**	**99.42±0.15**	**99.39±0.07**	**99.28±0.22**	**99.39±0.12**	**94.48±0.59**
	6	**99.65±0.06**	**99.69±0.09**	**99.62±0.05**	**99.61±0.10**	**99.64±0.10**	**99.59±0.04**	**95.61±0.24**
	8	**99.71±0.10**	**99.74±0.04**	**99.71±0.07**	**99.76±0.06**	**99.75±0.08**	**99.76±0.05**	**96.48±0.44**
	10	**99.82±0.05**	**99.79±0.08**	**99.78±0.03**	**99.82±0.02**	**99.78±0.05**	**99.77±0.11**	**97.26±0.17**

Table 6. *Overall classification accuracy (%) and STD for the Salinas data set (Overall Accuracy ± STD). The best results are highlighted in bold font.*

Method	Training Sample Percentage	No Noise	GWN (σ_n) 100	200	300	400	500	Shot Noise
PCA	2	**99.76±0.01**	**99.76±0.02**	99.52±0.14	99.20±0.45	99.20±0.09	98.89±0.24	90.90±0.51
	4	99.84±0.05	99.78±0.04	99.60±0.12	99.66±0.05	99.52±0.20	99.48±0.16	95.47±0.24
	6	99.85±0.04	99.84±0.04	99.83±0.03	99.72±0.04	99.71±0.07	99.64±0.08	96.49±0.45
	8	99.88±0.04	99.88±0.02	99.85±0.06	99.83±0.03	99.77±0.02	99.81±0.03	97.72±0.13
	10	99.92±0.04	99.86±0.02	99.88±0.04	99.85±0.04	99.79±0.03	99.81±0.06	98.15±0.26
MNF	2	99.75±0.05	99.63±0.09	**99.60±0.12**	**99.61±0.07**	**99.50±0.10**	**99.35±0.19**	**96.02±0.50**
	4	**99.84±0.03**	**99.79±0.02**	**99.81±0.05**	**99.77±0.11**	**99.78±0.06**	**99.72±0.07**	**97.99±0.31**
	6	**99.88±0.04**	**99.89±0.03**	**99.85±0.03**	**99.83±0.05**	**99.87±0.02**	**99.86±0.04**	**98.78±0.31**
	8	**99.92±0.03**	**99.90±0.03**	**99.91±0.03**	**99.89±0.04**	**99.92±0.03**	**99.82±0.05**	**99.05±0.11**
	10	**99.93±0.05**	**99.93±0.01**	**99.92±0.02**	**99.89±0.03**	**99.88±0.02**	**99.90±0.03**	**99.35±0.08**

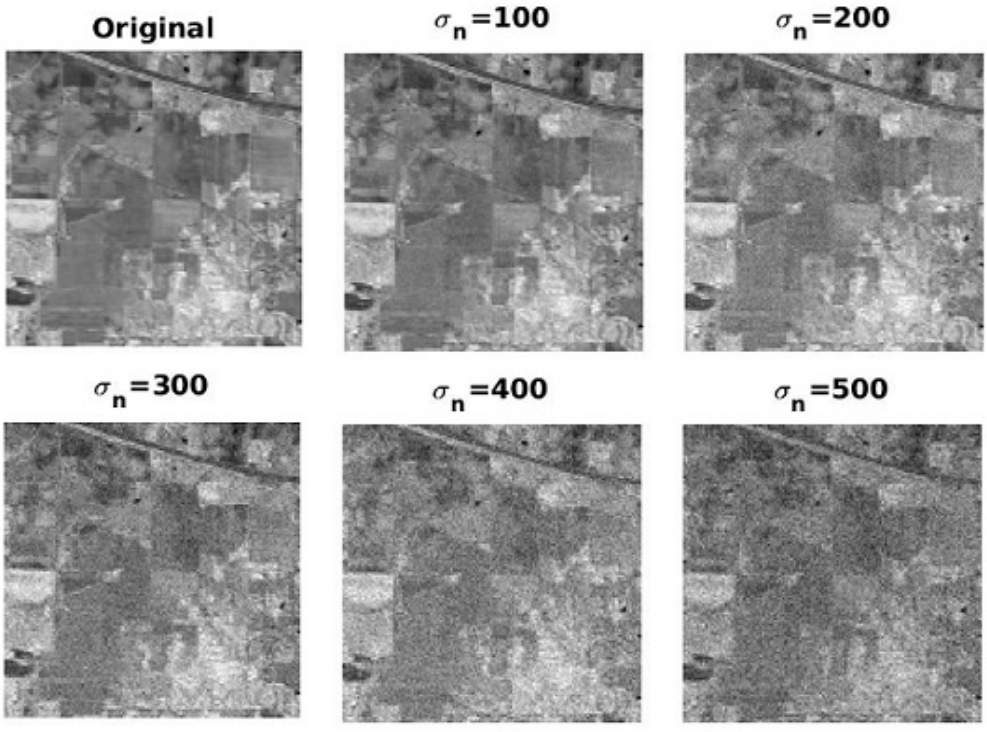


Fig. 4. *The noise-free #50 band in the Indian Pines data cube and the noisy spectral bands with GWN added.*

Fig. 5. *The noise-free #50 band in the Indian Pines data cube and the noisy band with shot noise added.*

Table 7 tabulates the execution time in seconds by using our unoptimized Matlab code with 2% of pixels as training samples and the rest of pixels as testing samples for the Indian pines data cube, the Pavia University data cube, and the Salinas data cube. Our experiments are done under the Linux operating system with Intel(R) Xeon(R) CPU E5-2697 v2 at 2.70GHz and 131 GB of random-access memory (RAM). Our new method is faster than PCA-EPFs for the Indian Pines data cube, but it is slower than PCA-EPFs for the Pavia University data cube and the Salinas data cube for HSI classification. Nevertheless, we can optimize our MNF-EPFs so that it can be faster than PCA-EPFs significantly.

Table 7. *The execution time in seconds by using our unoptimized Matlab code with 2% of pixels as training samples for the Indian pines data cube, the Pavia University data cube, and the Salinas data cube. The best results are highlighted in bold font.*

Method	Indian Pines	Pavia University	Salinas
PCA	8.49	**31.18**	**39.69**
MNF	**4.48**	33.53	41.41

CONCLUSIONS

Image classification is a phenomenal mechanism for analyzing diversified land cover in remotely sensed hyperspectral images. In remote sensing, HSI classification has been an established research topic, and hence, the inherent primary challenges are (a) curse of dimensionality and (b) insufficient samples pool during training. With the available high spectral resolution, subtle objects and materials can be extracted by hyperspectral imaging sensors with very narrow diagnostic spectral bands.

In this paper, we have presented a novel method for HSI classification by using MNF-based EPFs instead of PCA-based EPFs method. Our new method MNF-EPF is very robust to both GWN and shot noise for HSI classification. Experiments demonstrate that our new method compares favorably to PCA-EPFs for HSI classification especially under the noisy environment. In addition, our MNF+EPFs outperform the PCA+EPFs even when no noise is added to the HSI data cubes for most testing cases, which is very desirable in remote sensing.

Future research will be carried out in the following aspects. We may perform PCA-based denoising methods (Chen ang Qian 2011, Chen *et al.* 2014, Luo *et al.* 2016) for HSI image classification as a preprocessing step. We may investigate deep learning for HSI classification. Also, we may use low-rank matrix approximation for HSI classification. Furthermore, it is important to take advantages of both spatial and spectral information for HSI classification.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Xudong Kang and Dr. Shutao Li for providing their Matlab code of the PCA-EPFs method to us. Without this Matlab code, this work could not have been done.

REFERENCES

Camps-Valls G and Bruzzone L (2005). Kernel-based methods for hyperspectral image classification. IEEE T Geosci Remote, 43:1351-62.

Chen GY, Bui TD, Quach KG and Qian SE (2014). Denoising hyperspectral imagery using principal component analysis and block matching 4D filtering. Can J Remote Sens, 40:60-7.

Chen GY and Qian SE (2011). Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage. IEEE T Geosci Remote, 49:973-80.

Chen Y, Lin Z, Zhao X, Wang G and Gu Y (2014). Deep learning-based classification of hyperspectral data. IEEE J-Stars 7:2094-107.

Chen Y, Nasrabadi NM and Tran TD (2011). Hyperspectral image classification using dictionary-based sparse representation. IEEE T Geosci Remote 49:3973-85.

Chen Y, Nasrabadi NM and Tran TD (2013). Hyperspectral image classification via kernel sparse representation. IEEE T Geosci Remote 51:217-31.

Cheng G, Zhu, F, Xiang S, Wang Y and Pan X (2016). Semisupervised hyperspectral image classification via discriminant analysis and robust regression. IEEE J-Stars, 9:595-608.

Cortes C and Vapnik VN (1995). Support-vector networks. Machine Learning 20:273–97.

Green AA, Berman M, Switzer P and Craig MD (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE T Geosci Remote 26:65-74.

Jolliffe IT (2002). Principal Component Analysis, second edition (Springer).

Kang XD, Li S and Benediktsson JA (2014). Spectral-spatial hyperspectral image classification with edge-preserving filtering. IEEE T Geosci Remote 52:2666-77.

Kang X, Xiang X, Li S and Benediktsson JA (2017). PCA-Based Edge-Preserving Features for Hyperspectral Image Classification. IEEE T Geosci Remote 55:7140-51.

M. Fauvel, J. Chanussot and J. A. Benediktsson (2012). A spatial–spectral kernel-based approach for the classification of remote-sensing images. Pattern Recogn, vol. 45(1): 381-392.

Li J, Bioucas-Dias JM and Plaza A (2013). Spectral-spatial classification of hyperspectral data using loopy belief propagation and active learning. IEEE T Geosci Remote, 51:844-56.

Li J, Marpu PR, Plaza A, Bioucas-Dias JM and Benediktsson JA (2013). Generalized composite kernel framework for hyperspectral image classification. IEEE T Geosci Remote 51:4816-29.

Liu H, Xia K, Li T, Ma J and Owoola E (2020). Dimensionality reduction of hyperspectral images based on improved spatial–spectral weight manifold embedding. Sensors 20:4413.

Luo GC, Chen GY, Tian L, Qin K and Qian SE (2016). Minimum noise fraction versus principal component analysis as a preprocessing step for hyperspectral imagery denoising. Can J Remote Sens 42:106-16.

Melgani F and Bruzzone I (2004). Classification of hyperspectral remote sensing images with support vector machines. IEEE T Geosci Remote 42:1778-90.

Zhou F, Hang R, Liu Q and Yuan X (2019). Hyperspectral image classification using spectral-spatial LSTMs. Neurocomputing 328:39-47.