HYPERSPECTRAL IMAGE CLASSIFICATION USING SET-TO-SET DISTANCE

Junjun Jiang[1], Chen Chen[2], Xin Song[1], and Zhihua Cai[1]*

[1] School of Computer Science, China University of Geosciences, Wuhan 430074, China
[2] Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75080 USA

ABSTRACT

Hyperspectral image (HSI) classification has attracted much attention and extensive research efforts over the past decade. Due to few labeled samples versus high dimensional features, it is a challenging problem in practice. Recently, combining the pixel spectral information and the spatial (neighborhood) information has been verified to be effective for HSI classification. In this paper, we introduce a novel method for HSI classification using *set-to-set distance (SSD)*. Based on the assumption that neighbor pixels tend to belong to the same class with high probability, we model a test pixel and its neighbor pixels as a testing set (or a neighbor set) inspired by bilateral filtering. Meanwhile, the training pixels belong to the same class are modeled as a training set. Therefore, the classification is based on comparisons of sets distances. Experiments on a real HSI dataset show that our proposed method outperforms a number of existing state-of-the-art approaches.

Index Terms—Remote Sensing, Hyperspectral image classification, spatial information, set-to-set distance, set classification.

1. INTRODUCTION

Recent advances in Hyperspectral Remote Sensor (HRS) technology allow the simultaneous acquisition of hundreds of narrow contiguous spectral bands with a fine spectral resolution for each image pixel. This detailed spectral information about the materials in the scene increases the possibility of more accurately discriminating ground objects [1-6]. Hyperspectral image (HSI) classification plays an important role in remote sensing. It has been used in various thematic applications including ecological science, geological science, hydrological science, precision agriculture and military applications.

Hyperspectral images with high spatial resolution provide new opportunities of analyzing small spatial structures in images. In general, hyperspectral images usually have flat and large homogeneous regions that consist of the same type of material (i.e., same class). Compared to the classifiers solely using spectral information, classifying image pixels by taking into account the spatial (neighborhood) information has been verified to be very effective, especially for high spatial resolution images.

For example, morphological profile (MP) generated by the morphological operators (e.g., opening and closing), is applied to model structural information in [7]. Qian et al. [8] introduced the 3-D discrete wavelet transform to effectively capture the spatial information of HSI in different scales and orientations. In [9, 10], Markov random field (MRF) was employed as a post-processing step to enhance the results of support vector machine (SVM) [11] by incorporating the spatial-context information based on the results of SVM, namely SVM-MRF. In addition, researchers have also explored simultaneous spectral and spatial information within the designed classifier. Camps-Valls et al. designed a composite kernel to combine both the spectral and spatial information for SVM. The resulting classifier is referred to as SVM-CK [12]. As a powerful tool for statistical signal modeling, sparse representation (or sparse coding) has been successfully used in image processing [13-16] and computer vision [17, 18] applications, and recently has led to promising results in HSI classification [19]. In [19], simultaneous orthogonal matching pursuit (SOMP) and OMP with smoothing (OMP-S) were developed to incorporate the contextual information for the HSI classification task. Li et al. proposed a joint collaborative representation with Tikhonov regularization (JCR-TR) considering four different strategies of incorporating contextual information [20].

To simultaneously incorporate the spectral and spatial information for HSI classification, in this paper we propose a novel algorithm based on set-to-set distance (SSD). Experimental results on one real HSI dataset demonstrate the proposed method outperforms other state-of-the-art spectral-spatial based HSI classification techniques.

The structure of this paper is as follows: Section 2 presents the motivation and contributions of this paper. Section 3 provides a detailed description of the proposed method by firstly introducing the strategy of generating the neighbor set, secondly showing the set distance measurement approach, and lastly summarizing the overall SSD based HSI classification algorithm. The experimental results and their discussion are stated in Section 4. Finally, the conclusion appears in Section 5.

The research was supported by the National Natural Science Foundation of China under Grant 61501413.

*Corresponding author: junjun0595@163.com

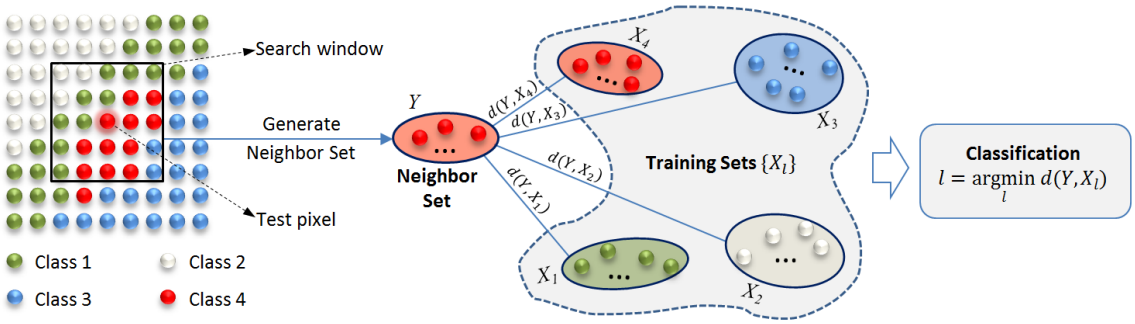


Fig.1. Flow diagram of the proposed SSD based HSI classification method.

2. MOTIVATION AND CONTRIBUTIONS

As far as utilizing space neighbor information for HSI classification, existing methods mostly use the coding-residual approach. They first encode the pixel (with/without the spatial information) on the training samples collaboratively or sparsely, and then use the residual errors of the representation to classify the test pixels. In this paper, we propose an SSD based HSI classification approach. Fig. 1 illustrates the diagram of using SSD for HSI classification. As shown in the flow diagram, for a test pixel to be classified, we first search its neighbor set (including itself) and utilize this neighbor set to represent it. Then, we use the SSD to measure the distance between the neighbor set of the test pixel and the training samples within each class, which can be seen as a set modeled by the affine hulls of the spectral feature vectors. Lastly, the class label of the test pixel is assigned according to the minimum set distance. In summary, the main contributions of this paper lie in the following three aspects:

- It casts the HSI classification as the problem of measuring the set distance.
- Inspired by the merits of bilateral filtering, we propose a neighbor set generation method that considers the pixel density similarity and the spatial distance simultaneously.
- Experimental results demonstrate that our proposed SSD based method is superior to the state-of-the-art spectral-spatial based HSI classification methods.

3. PROPOSED METHOD

Consider a dataset with n total training samples $X=\{x_i\}_{i=1}^{n}$ in R^d (d is the dimensionality) and class labels $w_i \in \{1,2,...,C\}$, where C is the number of classes. Let n_l be the number of available training samples for the l-th class, $\sum_{l=1}^{C} n_l = n$. Thus the training samples X can be partitioned into l different class-specific subset according to their class labels, $X_l=\{x_i \mid \forall\, i \text{ s.t. } w_i=l\}$. Given a test pixel y, the task of HSI classification is to assign a class label to y.

3.1 Neighbor set generation

In this subsection, we give the details of how to generate the neighbor set of a given test pixel y. Bilateral filter is an edge-preserving and noise-reducing smoothing filter for images. The intensity value of each pixel in an image is replaced by a weighted average of intensity values from nearby pixels. The weights depend not only on the Euclidean distance of pixels, but also on the pixel intensity. In this paper, we borrow the idea of bilateral filtering to form a neighbor set for each test pixel. More specifically, the selected pixels in a neighbor set should be *near* and *similar* to the test pixel. To this end, we first set a spatial window centered at the test pixel. We then choose the pixels that are similar to the test pixel as the neighbor set. Thus, the neighbor set Y can be mathematically defined as follows:

$$Y=\{y_i \mid y_i \in \Omega(y)\} \cap \{y_i \mid dist(y-y_i) < c\frac{1}{L^2}\sum_i dist(y-y_i)\}, \quad (1)$$

where $\Omega(y)$ is the pixel set within the spatial window of size $L\times L$, $\frac{1}{L^2}\sum_i dist(y-y_i)$ is the mean distance between the test pixel y and each pixel in the window, and c is a parameter that controls the number of pixels in the neighbor set. In Eq. (1), $y_i \in \Omega(y)$ can ensure that the selected pixels in neighbor set is *near* (or *close*) to the test pixel, and $dist(y-y_i) < c\frac{1}{L^2}\sum_i dist(y-y_i)$ can ensure the selected pixels in neighbor set are *similar* to the test pixel.

3.2 Set distance measurement

We cast the HSI classification as the problem of measuring the set distance between the neighbor set Y and the training subsets X_l (each training set is composed of the training samples with the same class label). Mathematically, given two pixel sets $Y=\{y_i\}_{i=1}^{t}$ and $X_l=\{x_l\}_{l=1}^{n_l}$, our goal is to calculate the set to set distance between them, i.e., $d(Y,X_l)$.

3.2.1 Pixel set model

A pixel set is usually represented by a hull, i.e., a subspace spanned by all the available pixels in the set. The hull of Y and X_l are defined as $H(Y)=\{Y\alpha\}$ and $H(X_l)=\{X_l\beta\}$ respectively, where $\alpha=[\alpha_1,\alpha_2,...,\alpha_t]^{\mathrm{T}}$ and $\beta=[\beta_1,\beta_2,...,\beta_{n_l}]^{\mathrm{T}}$. Usually, $\sum_i\alpha_i=1$ and $\sum_l\beta_l=1$ are required:

$$H(Y)=\{\sum_{i=1}^{t}y_i\alpha_i \mid \sum_i\alpha_i=1\} \quad (2)$$

$$H(X_l)=\{\sum_{l=1}^{n_l}x_l\beta_l \mid \sum_l\beta_l=1\}. \quad (3)$$

3.2.2 Set-to-set distance (SSD)

Given two pixel sets Y and X_l, the SSD between them is defined as

$$d(Y,X_l)=\|Y\hat{\alpha}-X_l\hat{\beta}\|_2^2 \quad \text{s.t.} \quad \sum_i\hat{\alpha}_i=1 \text{ and } \sum_l\hat{\beta}_l=1, \quad (4)$$

where $\hat{\alpha}$ and $\hat{\beta}$ can be solved by:

$$(\hat{\alpha},\hat{\beta})=\arg\min_{\alpha,\beta}\|H(Y)-H(X_l)\|_2^2$$
$$\text{s.t.} \quad \sum_i\alpha_i=1 \text{ and } \sum_l\beta_l=1. \quad (5)$$

Based on the constraints that $\sum_i\alpha_i=1$ and $\sum_l\beta_l=1$, (5) can be converted to the following unconstrained optimization problem:

$$\begin{aligned}(\hat{\alpha},\hat{\beta})&=\arg\min_{\alpha,\beta}\|H(Y)-H(X_l)\|_2^2=\|Y\alpha-X_l\beta\|_2^2\\&=\|\{[\hat{y}_1,\hat{y}_2,..,\hat{y}_{t-1}][\alpha_1,\alpha_2,...,\alpha_{t-1}]^{\mathrm{T}}+y_t\}\\&\quad-\{[\hat{x}_1,\hat{x}_2,..,\hat{x}_{n_l-1}][\beta_1,\beta_2,...,\beta_{n_l-1}]^{\mathrm{T}}+x_{n_l}\}\|_2^2\\&=Z\gamma+y_t-x_{n_l},\end{aligned} \quad (6)$$

where $\hat{y}_i=y_i-y_t$, $i=1,2,...,t-1$, $\hat{x}_l=x_l-x_{n_l}$, $l=1,2,...,n_l-1$, $Z=[\hat{y}_1,..,\hat{y}_{t-1},\hat{x}_1,..,\hat{x}_{n_l-1}]$, and $\gamma=[\alpha_1,...,\alpha_{t-1},\beta_1,...,\beta_{n_l-1}]^{\mathrm{T}}$. Therefore, the solution to (5) is:

$$\hat{\alpha}=[\hat{\gamma}_1,\hat{\gamma}_2,...,\hat{\gamma}_{t-1},1-\textstyle\sum_{i=1}^{t-1}\hat{\gamma}_i]^{\mathrm{T}}$$
$$\text{and } \hat{\beta}=[\hat{\gamma}_t,\hat{\gamma}_{t+1},...,\hat{\gamma}_{t+n_l-2},1-\textstyle\sum_{l=t}^{t+n_l-2}\hat{\gamma}_l]^{\mathrm{T}},$$

where $\hat{\gamma}=(Z^{\mathrm{T}}Z)^{-1}Z^{\mathrm{T}}(x_{n_l}-y_t)$. Therefore, the distance between two pixel sets Y and X_l is calculated as

$$d(Y,X_l)=Z(Z^{\mathrm{T}}Z)^{-1}Z^{\mathrm{T}}(x_{n_l}-y_t)+y_t-x_{n_l}. \quad (7)$$

3.3 Summary of the proposed algorithm

To better understand the proposed SSD based HSI classification algorithm, we summarize each step of the algorithm as follows:

a) Search the neighbors of a test pixel to get the corresponding neighbor set Y using the strategy presented in Section 3.1.

b) For each class of the training samples, construct a matrix X_l.

c) Compute the matrix Z through $Z=[\hat{y}_1,\hat{y}_2,..,\hat{y}_{t-1},\hat{x}_1,\hat{x}_2,..,\hat{x}_{n_l-1}]$.

d) Find the estimation of $\hat{\gamma}$ for (6) according to $\hat{\gamma}=(Z^{\mathrm{T}}Z)^{-1}Z^{\mathrm{T}}(x_{n_l}-y_t)$.

e) Calculate the distance between two pixel sets Y and X_l via (7).

f) Repeat c) to e) and assign the test pixel to the class that has the smallest SSD.

3.4 Connections to existing coding-residual based methods

Let's revisit the objective functions of the coding-residual (CR) based methods and the proposed SSD method:

(**CR**) $\quad \hat{\beta}=\arg\min_{\beta}\|Y-X_l\beta\|_2^2+\pi\Omega(\beta)$

(**SSD**) $\quad (\hat{\alpha},\hat{\beta})=\arg\min_{\alpha,\beta}\|Y\alpha-X_l\beta\|_2^2$

$\quad\quad$ s.t. $\quad \sum_i\alpha_i=1$ and $\sum_l\beta_l=1$.

In CR, Y can be a test pixel only or a test pixel and its nearest neighbors, and $\Omega(\beta)$ denotes some priors about the solution of β, which can be modeled by collaborative, sparse or local regularization. From these two objective functions, we learn that the core difference between CR and SSD lies in their ways of representation: CR method is based on single representation while SSD method is dual. In other words, CR method fixes the test samples, while SSD method gives different freedom to each test sample by an additional vector α. Therefore, SSD method has more powerful ability in utilizing the spatial information (i.e., test samples).

4. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, different classification strategies using spectral and spatial information are compared using a real HSI dataset, Pavia University (PaviaU) dataset. PaviaU dataset has a spatial dimension of 610×340 pixels. Twelve channels (or spectral bands) were removed due to low signal to noise ratio (SNR). The remaining 103 spectral channels were left for further analysis. As shown in Fig. 2, there are nine ground-truth classes: tree, asphalt, bitumen, gravel, metal sheet, shadows, bricks, meadow, and soil. There are 42776 labeled pixels for the PaviaU dataset. Detailed information of the number of training and testing samples used for this dataset is summarized in Table 1.

Parameter tuning. We investigate the parameters of neighbor set generation in the proposed method, i.e., the window size and the parameter c which controls the number of neighbors in the neighbor set. In Fig. 1, we report the performance demonstrating the sensitivity of the proposed method over a wide range of these two parameters, e.g., the window size varies from 1 to 15 and the parameter c varies from 0.1 to 100. As for the window size, when it is set to 7 ×7, the proposed method achieves the best performance. This is mainly because that small window size may not able

Table 1. Classification accuracy (%) for the University of Pavia dataset.

Class	# samples		Classification algorithms								
	Train	Test	SVM	SVM-CK	SVM-MRF	OMP	SOMP	OMP-S	NRS	JCR-TR	SSD
Asphalt	60	6631	80.29	84.68	95.39	57.58	50.08	55.63	82.05	93.30	**97.35**
Bare soil	60	18649	84.32	96.85	89.89	73.04	79.87	78.89	80.05	97.36	**99.40**
Bitumen	60	2099	82.84	89.52	90.95	72.70	85.66	86.61	79.80	**93.66**	91.66
Bricks	60	3064	92.26	96.64	96.74	91.84	92.23	94.71	95.76	**97.52**	95.04
Gravel	60	1345	99.11	99.70	99.41	99.33	100	100	99.55	**100**	99.93
Meadows	60	5029	89.12	93.64	99.54	64.86	70.87	72.94	92.27	99.20	**99.98**
Metal sheets	60	1330	92.01	97.44	96.32	89.17	97.59	93.08	94.89	98.50	**99.92**
Shadows	60	3682	79.71	88.10	**95.93**	62.19	67.93	63.96	87.06	94.38	93.64
Trees	60	947	99.79	99.79	**99.79**	97.04	90.50	98.84	99.47	96.62	95.78
Overall Accuracy (%)			84.06	93.63	93.66	71.94	75.76	76.36	83.94	96.62	97.92

to capture important spatial information, while large window size may introduce some irrelative information, e.g., pixels from other classes. The parameter c plays an important role in selecting similar pixels with the same class while rejecting dissimilar pixels belong to different classes. When $c=1.1$, the proposed method achieves the highest classification accuracy. Therefore, in all our experiments, we set the window size to 7×7 and parameter c to 1.1.

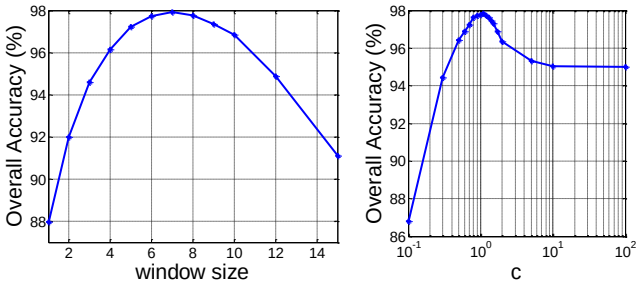
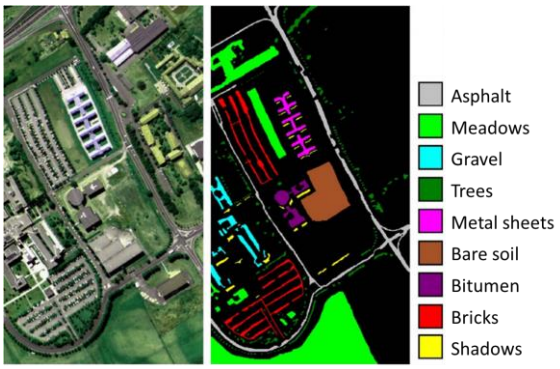


Fig. 1. The overall classification accuracy (%) versus the window size (left) and the parameter c (right).

Comparisons to state-of-the-art results. Eight state-of-the-art methods for HSI classification are used to make holistic comparisons with our method. These methods include three SVM based approaches (SVM [11], SVM-CK [12] and SVM-MRF [10]), three sparse representation based approaches [19] (OMP, SOMP and OMP-S), and two collaborative representation based methods (JCR-TR [20] and NSR [21]). To facilitate a fair comparison, we follow the same experimental setting in [20]. Specifically, the numbers of test and train samples are shown in Table 1. In this table, we also report the quantitative evaluation results in terms of the classification accuracies (class-specific and overall accuracy) of different methods on the PaviaU dataset. As evident from this table, our proposed SSD method achieves the best performance in terms of overall accuracy (i.e., 97.92%). The improvement of SSD over the second best method, i.e., Li et al.'s JCR [20], is 1.30%. In the competition of all 9 classes, our proposed method obtains the best classification accuracy in four classes ("Asphalt", "Bare soil", "Meadows", and "Metal sheets"), and show strong competitiveness in other classes. This improvement can be explained as follows: by modeling the test and training sets as affine hulls, our proposed method has much more flexible in using the spectral and spatial information in HSI classification. In other words, SSD can gives different freedom to each test sample, while the traditional methods treat all test samples equally.



(a) (b)

Fig. 2. Pavia University image. (a) is the false color image using three spectral bands, and (b) is the ground truth map.

5. CONCLUSIONS

In this paper, we developed a novel method for Hyperspectral image (HSI) classification based on set-to-set distance (SSD). It is the first time that we introduce the set distance for HSI classification. Our proposed method characterizes each training set (where the samples all have the same class label) from the training sample pixels and the test set (includes the test pixel and its neighbor pixels selected by bilateral filtering criterion) in terms of the affine hulls of the spectral feature vectors. Classification is performed by finding the training set that is closest to the test set in the sense of minimum distance between convex sets. Experimental results on a real hyperspectral image verified that the proposed algorithm is superior to the state-of-the-art spectral-spatial based HSI classification methods.

6. REFERENCES

[1] Fauvel, M., Tarabalka, Y., Benediktsson, J. A., Chanussot, J., and Tilton, J. C., "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652-675, March 2013.

[2] W. Li, C. Chen, H. Su, and Q. Du, "Local Binary Patterns for Spatial-Spectral Classification of Hyperspectral Imagery,"IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 7, pp. 3681-3693, July 2015.

[3] C. Chen, W. Li, E. W. Tramel, M. Cui, S. Prasad, and J. E. Fowler, "Spectral-Spatial Preprocessing Using Multihypothesis Prediction for Noise-Robust Hyperspectral Image Classification," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 4, pp. 1047-1059, April 2014.

[4] C. Chen, J. Jiang, B. Zhang, W. Yang, and J. Guo, "Hyperspectral Image Classification Using Gradient Local Auto-Correlations," the 3rd IAPR Asian Conference on Pattern Recognition, November 3-6, 2015, Kuala Lumpur, Malaysia.

[5] L. Ma, M. M. C rawford, and J. W. Tian, "Local manifold learning based k-nearest-neighbor for hyperspectral image classification," IEEE Trans. Geosci. Remote Sens., vol. 48, no. 11, pp. 4099–4109, Nov. 2010.

[6] L. Ma, M. M. C rawford, X. Yang, and Y. Guo, "Local manifold learning based g raph construction f or semisupervised hyperspectral image classifi-cation,"IEEE Trans. Geosci. Remote Sens., vol. 53, no. 5, pp. 2832–2844, May 2015.

[7] M. Pesaresi, and J. Benediktsson, "A new approach for the morphological segmentation of high resolution satellite imagery," IEEE Transactions on Geoscience and Remote Sensing, vol. 39, no. 2, pp. 309-320, Feb. 2001.

[8] Y. Qian, M. Ye, and J. Zhou, "Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features," IEEE Transactions on Geoscience and Remote Sensing, vol. 51, no. 4, pp. 2276-2291, Apr. 2013.

[9] G. Moser and S. B. Serpico, "Combining support vector machines and markov random fields in an integrated framework for spatial image classification," IEEE Trans. Geosci. Remote Sens., vol. 51, no. 5, pp. 2734–2754, May 2013.

[10] W.Li, S.Prasad,andJ. E.Fowler, "Hyperspectral image classification using Gaussian mixture model and Markov random field," IEEE Geosci. Remote Sens. Lett., vol. 11, no. 1, pp. 153–157, Jan. 2014.

[11] 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.

[12] 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.

[13] J. Jiang, X. Ma, R. Hu, Z. Cai. "Sparse Support Regression for Image Super-Resolution," IEEE Photonics Journal, vol. 7, no.5, pp. 1-11, Oct. 2015.

[14] J. Jiang, R. Hu, Z. Han, Z. Wang, J. Chen. "Two-step super-resolution approach for surveillance face image through radial basis function-partial least squares regression and locality-induced sparse representation," Journal of Electronic Imaging, vol. 22, no. 4, Oct 2013.

[15] J. Jiang, R. Hu, Z. Han, Z. Wang. "Low-Resolution and Low-Quality Face Super-resolution in Monitoring Scene via Support-Driven Sparse Coding," Journal of Signal Processing Systems, vol. 75, no. 3, pp. 245-256, Jun 2014.

[16] Z. Zhu, F. Guo, H. Yu, and C. Chen, "Fast Single Image Super-Resolution via Self-Example Learning and Sparse Representation," IEEE Transactions on Multimedia, vol. 16, no. 8, pp. 2178-2190, December 2014.

[17] J. Ma, J. Zhao, and A. L. Yuille. "Non-Rigid Point Set Registration by Preserving Global and Local Structures," IEEE Transactions on Image Processing, vol. 25, no. 1, pp. 53-64, 2016.

[18] J. Ma, J. Zhao, J. Tian, A. Yuille, and Z. Tu, "Robust Point Matching via Vector Field Consensus," IEEE Transactions on Image Processing, vol. 23, no. 4, pp. 1706-1721, 2014.

[19] 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.

[20] W. Li and Q. Du, "Joint within-class collaborative representation for hyperspectral image classification," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 7, no. 6, pp. 2200–2208, Jun. 2014.

[21] W. Li, E. W. Tramel, S. Prasad, and J. E. Fowler, "Nearest regularized subspace for hyperspectral classi fication, " IEEE Trans. Geosci. Remote Sens., vol. 52, no. 1, pp. 477 – 489, Jan. 2014.