Journal of Engineering and Applied Sciences 13 (7): 1606-1612, 2018
ISSN: 1816-949X

Ranking Based Classification in Hyperspectral Images

[1]Y. Aruna Suhasini Devi and [2]K.Manjunathachari
[1]CMR College of Engineering & Technology, Medchal, Hyderabad, India
[2]Professor & HOD ECE Department, GITAM UNIVERSITY, Hyderabad, India

Abstract: In recent years, the ranking based band selection has been very successful in remote sensing image classification. Hyperspectral imagery often contains hundreds of images, hence, dimensionality reduction should be applied to overcome the difficulty of Hughes phenomenon. In this study, two approaches for efficiency of band selection and robustness are applied. Classification is done using ranking based Fast Density Peek Clustering (FDPC) algorithm. For FDPC algorithm, first the ranking score of each band is computed by weighing the normalized local density and the intracluster distance rather than equally taking them into account. Secondly, an exponential-based learning rule is employed to adjust the cutoff threshold for a different number of selected bands where it is fixed in the FDPC. Finally, the selected features are processed by MPCA (Multilinear Principal Component Analysis) technique to reduce the data redundancy and increasing robustness. From the experimental analysis it is observed that the proposed ranking based classification is more than efficient and robust when compared with existing band selection techniques.

Key words: FDPC algorithm, MPCA (Multilinear Principal Component Analysis) technique, hyperspectral images, weighing, selected features, classification

INTRODUCTION

Hyperspectral images are produced by hyperspectral sensors at different spectral channels in a single collection. One of the purposes of hyperspectral images is to identify target materials present in a scene without physical contact. Hyperspectral Imaging (HSI) is a remote sensing technique that acquires spatial images in typically hundreds of contiguous bands of high spectral resolution covering the visible, near-infrared and shortwave infrared bands. This technique has been applied in several applications; for instance, face detection and identification of the ground surface as well as atmospheric composition, analysis of soil type, agriculture, mineral exploration and environmental monitoring (e.g., oil/gas leakage from pipelines or natural wells). This technology produces a signature for each pixel in the image in many highly correlated bands presenting considerable amounts of spectral redundancy. On the other hand, dimension reduction of multivariate images is one of the main subjects of interest for the hyperspectral community.

Target detection, image segmentation, pixel classification and spectra unfixing in HSI have the additional difficulty that pixels are located in a high dimension space increasing computational complexity and degrading accuracy (Bioucas-Dias *et al*., 2013; Hughes, 1968). However, compared with the large amount of spectral bands, it is difficult and laborious to obtain sufficient training samples in practice which often leads to ill-conditioned problems such as the Hughes phenomenon (i.e., for a limited number of training samples, the classification accuracy can be deteriorated in a higher dimensional feature space) (Hughes, 1968). Moreover, due to the high correlation among the neighboring bands which not only increases the computational complexity of the classifiers but also may have a negative impact on the classification accuracy? Dimensionality Reduction (DR) should be applied as a preprocessing step to discard the redundant information (MartInez-UsoMartinez-Uso *et al*., 2007; Khodr and Younes, 2011).

Feature extraction is one of the techniques in dimensionality reduction. It achieves higher classification accuracy and the obtained features are not related to the original wavelengths. Alternatively, band selection methods can preserve the relevant original information of the spectral bands which could significantly simplify the image acquisition process and save the data storage.

Band selection scheme is classified into two main parts, i.e., ranking based and clustering based methods. In ranking based band selection method it first quantifies the importance of each band according to a certain criterion such as non- Gaussianity (Chang and Wang, 2006; Chang, 2007), variance (Chang *et al*., 1999), mutual information (Guo *et al*., 2006), etc. Then a given number of top-ranked bands in the sorted sequence are selected to form the subset. Clearly, the key point in the ranking-based UBS methods is how well the criterion

Corresponding Author: Y. Aruna Suhasini Devi, CMR College of Engineering & Technology, Medchal, Hyderabad, India

characterizes the distinctiveness of spectral bands. Alternatively, the clustering-based UBS methods are performed in the similarity space defined on the bands. Through partitioning the bands into disjoint groups (clusters) such that bands in the same subset are more similar to each other than bands in different groups, the band close to the centroid of each cluster is picked out to form the chosen subset (Zhang *et al*., 2013a, b; Shi *et al*., 2013).

Zhang *et al*. (2013a, b) proposed a method for the Dimensionality Reduction (DR) of spectral-spatial features in Hyper Spectral Images (HSIs), from multilinear algebra of tensors. The proposed approach is a tensor extension of conventional supervised manifold-learning-based DR. In particular, we define a tensor organization scheme for representing a pixels spectral-spatial feature and develop Tensor Discriminative Locality Alignment (TDLA) for removing redundant information for subsequent classification. The optimal solution of TDLA is obtained by alternately optimizing each mode of the input tensors. The methods are tested on 3 public real HSI data sets collected by hyperspectral digital imagery collection experiment, reflective optics system imaging spectrometer and airborne visible/infrared imaging spectrometer. The classification results show significant improvements in classification accuracies while using a small number of features (Zhang *et al*., 2013a).

Shi *et al*. (2013) introduced an efficiency manifold learning based Spervised Graph Embedding (SGE) algorithm for Polarimetric Synthetic Aperture Radar (POLSAR) image classification. Linear dimensionality reduction technology named SGE is used to obtain a low-dimensional subspace which can preserve the discriminative information from training samples. Various POLSAR decomposition features are stacked into the input feature cube in the original high-dimensional feature space. The SGE is then implemented to project the input feature into the learned subspace for subsequent classification. The suggested method is validated by the full polarimetric airborne SAR system EMISAR in Foulum, Denmark. The experiments show that the SGE presents favorable classification accuracy and the valid components of the multi-feature cube are also distinguished (Shi *et al*., 2013).

Ranking based band selection: The ID method (Chang and Wang, 2006) uses a divergence criterion (Cover and Thomas, 1991) to assessthe discriminative potential of each band. Since, the class labels are not known in the BS, the non-Gaussianity of a band image is usually used instead of its real discriminative ability of classification. The Kurtosis of the probability distribution (Cardoso, 2003) of a band image and the Kullback-Leibler divergence (Cover and Thomas, 1991) to its associated Gaussian probability distribution are measured. After evaluating the probability distribution of each image and its associated Gaussian probability distribution, the whole band set can be sorted according to the resulting scores and those with higher values are selected. However, the main disadvantage of the ID method is that the information concerning the spectral correlation is ignored which causes the selected bands to still have a strong correlation, i.e., some bands can well represent the others and some bands can be removed without a significant loss of information.

Clustering based band selection: In other side, the hyperspectral BS problem can be conceptually formulated as a data clustering procedure which partitions the data set into groups of similar objects (clusters) without any class label information and has attracted increasing attention in the field of band selection. In clustering-based band selection approaches, each band is considered a data point and the bands are separated into several clusters on the basis of similarity matrix S. Through selecting a band in a cluster to represent all the bands in this cluster, the band decorrelation can be accomplished. If the similarity measure is the same (such as the Euclidean distance used here), the performanceof the clustering-based methods is better than that of the ranking-based methods (MartInez-UsoMartinez-Uso *et al*., 2007).

In this study, the two methods are combined together to carry out the band selection technique. Classification using a ranking based Fast Density Peek Clustering (FDPC) algorithm isproposed and MPCA (Multilinear Principal Component Analysis) technique is used to reduce the data redundancy and to increase robustness.

Ranking based clustering for hyperspectural images: Several popular ranking-based and clustering-based band selection methods were used as discussed in the previous studies. This study presentsclassification using a ranking based Fast Density Peek Clustering (FDPC) algorithm and MPCA (Multilinear Principal Component Analysis) technique to reduce the data redundancy and increase robustness.

MATERIALS AND METHODS

Fast Density Peek Clustering (FDPC) method: As described earlier, the FDPC is a rapid and effective method to find the cluster centers according to local density ρ and intracluster distance δ. Inspired by the idea of using density-peak-based ranking scores to carry out data clustering not only the stability of the chosen band subset can be guaranteed but also the difficulty of tuning parameters involved in the clustering process (such as AP) can be avoided, thus, the FDPC is a good candidate for hyperspectral band selection. However, FDPC method is not directly applied with the band selection task in hyperspectral imagery. From our approach, the appropriate number of selected bands is typically much larger than that of the clusters in the data. In FDPC clustering when the number of selected bands k is small, the representative bands are obliviously the cluster centers that are identified. As k increases, the FDPC may choose the bands with large ρ and relatively small δ, i.e., the points around the cluster centers which are highly correlated to the bands already selected (Fig. 1).

Weighted ranking score scheme: How to effectively integrate the two factors for hyperspectral band selection is still an unsolved issue. To combine local density ρ and distance δ from points of higher density is far more reliable than using either of the two factors alone. The weight value is important to promote the performance of the algorithm and the ranking score of each band shouldbe computed through weighting the two factors. Instead of taking δ and ρ into account equally, the proposed FDPC algorithm increases the weight of δ through a normalization and square product process. Before detailing the improvement, it is worth pointing out that a Gaussian kernel function is adopted to estimate local density ρ_i for each band image to decrease the negative impact of the statistical errors caused by the small number of bands, i.e.:

$$\rho_i = \sum_{j=1, j\neq i}^{L} \exp\left(-\left(\frac{D_{ij}}{d_c}\right)^2\right)$$

After δ has been computed by normalized scale of [0, 1]:

$$\delta = (\delta - \delta_{min})/(\delta_{max} - \delta_{min})$$

Where:
/ = Element wise division operator
ρ = The weight of δ has been significantly increased after the normalization

In order to compensate ρ for the loss of weight during the normalization procedure, ranking score γ_i for any band iis finally obtained by:

$$\gamma_i = \rho_i \times \delta_i^2$$

Exponential-based heuristic rule for cutoff threshold: The local density of ρ_i of any data point is computed based on cutoff threshold d_c:

$$dc = d_{ini}/\exp\left(\frac{k}{L}\right)$$

Where:
L = The spectral no of whole data
d_{ini} = The initial value of cutoff threshold

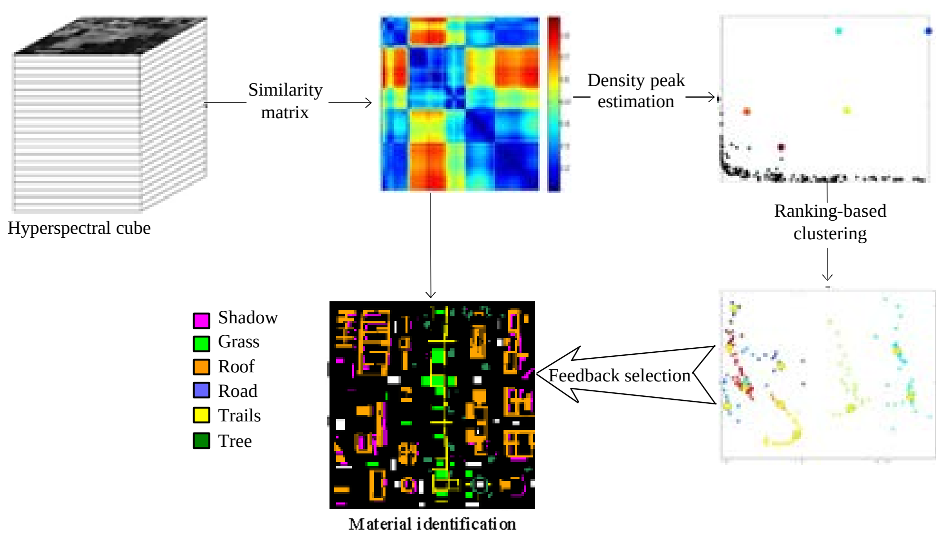

Fig. 1: Proposed FDPC algorithm for hyperspectral band selection

$$d_{ini} = 2\% \times L \times (L-1)$$

FDPC algorithm pseudocode:

BEGIN
$d_c = d_{ini}/\exp$ (k/L:% cutoff threshold
For i = 1; to L
do
$\rho i = \Sigma^{l}_{j=1, j \neq i} \exp(-(D_{ij}/d_c)^2)$ a Gaussian kernel function is used to estimate ρ
end for
[V, 1] = sort (ρ, decend)
$\delta_{I_1} = -1$
For i = 2 to L do
$\delta_{I_i} = \max(D_{ij})$
For j = 1 to i-1 do
If $D_{I_j I_i} < \delta_{I_i}$ then
$\delta_{I_i} = D_{I_j I_i}$
end if
end for
end for
$\delta_{I_1} = \max(\delta)$
$\rho = (\rho - \rho_{min})/(\rho_{max} - \rho_{min})$; ...normalizing two factors
$\delta = (\delta - \delta_{min})/(\delta_{max} - \delta_{min})$;
$\gamma = \rho . \times \delta^2$, …, square weight measure is applied on each element of δ
[V, 1] = sort (ρ, decend);
C = I(1:k);
For i = 1 to L do
$j = I_i$
$Aj = \min_{c \in c} D_{cj}$
end for
END

The pseudocode of the proposed E-FDPC band selection approach is outlined in Algorithm 1. It can be seen that the calculation time of the proposed E-FDPC approach can be divided into two parts. One part is related to the computation of the similarity matrix S that is used as the input of the subsequent clustering process, and the complexity is $0(2L^2MN)$ The other part is consumed for calculating the ranking scores of each band and the label assignment procedure, i.e., approximately $0(L^2)$ with a constant multiplier. Note that matrix S only needs to be computed once and the FDPC clustering process is not affected by the number of pixels in the scene, thus, the approach presented here is computationally efficient, even for the hyperspectral imagery with a large size.

Multilinear Principal Component Analysis (MPCA): MPCA is used to extract the most significant components for increasing robustness. The potential of the MPCA-based feature extraction has been used to high-order computer vision and pattern recognition applications such as face, handwriting and gait recognition. Tacking a set of N node integers as $\{A_m \in R^{L_1 \times L_2 \times .. \times L_N}\}, m = 1, 2, ..., M$ and the projection $U_d \in R^{L_d \times D_d}, d = 1, 2, ..., N$ matrices the low dimensional representation of input tensor $B = B \times_1 U^T_d \times, ..., \times_N U_N^T$ with $B \in R^{D_1 \times D_2 \times .. \times D_N}$ the procedure of MPCA is given.

Data centralization:

$$\breve{A}_m = A_m - \bar{A}, \bar{A} = \left(\frac{1}{M}\right) \Sigma^M_{m=1} A_m$$

At dth mode, unfold tensor A_m into matrix $mat_d A_m$ and edge composition:

$$\phi_d = \sum_{m=1}^{M} (mat_d A_m), (mat_d A_m)^T$$

Optimization of the projection matrices with the initial projection is formed by calculating:

$$\tilde{P}_m = \breve{A}_m \times_1 U_1^T \times, ..., \times_d U_d^T, ..., \times_N U_N^T$$

And the scatter matrix:

$$\varphi_0 = \sum_{m=1}^{m} \| \tilde{P}_m \|_2^p$$

In an iterative way, the projection:

$$\breve{P}_m = \breve{A}_m \times_1 U_d^T \times, ..., \times_d U_d^T \times, ..., \times_N U_N^T$$

If the difference of the $\varphi_K - \varphi_{k-1}$ is smaller than a predefined threshold. The low dimensional tensor projection matrix:

$$B^r = B_1 \times U_1^T \times, ..., \times_d U_d^T \times, ..., \times_N U_N^T$$

RESULTS AND DISCUSSION

Experimental evaluation: The proposed classification framework has been validated using hyperspectral image acquired from the instrument Hyperspectral Digital Imagery Collection Experiment (HYDICE) in Fig. 2.

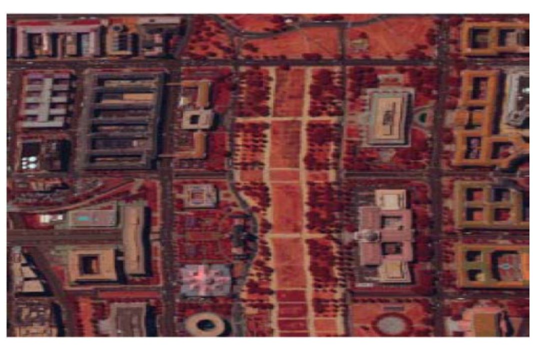

Fig. 2: Hyperspectral image-HYDICE Washington DC Mall

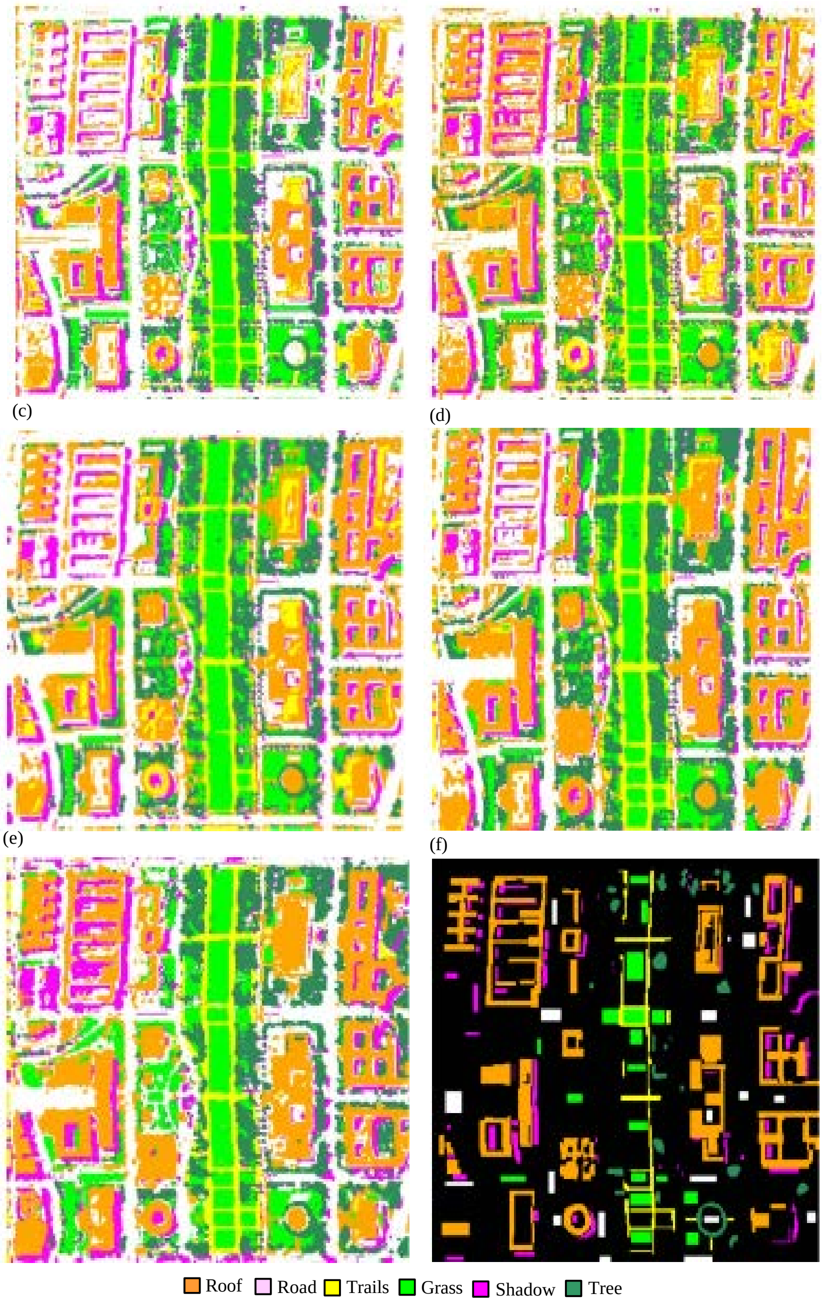

Fig. 3: Classification map obtained using different classification methods: a) K-cluster; b) AP method; c) ID method; d) DB scan; e) FDPC with MPCA and f) Ground truth reference

The HYDICE airborne hyperspectral data flight over the Washington DC Mall has been considered in experimental evaluation. These data consist of 210 spectral bands from 0.4-2.4 μm, among which 19 bands were discarded due to water absorption. The test image contains 280×307 pixels with a 5 m spatial resolution. The reference data consist of 17061 test samples as listed in Table 1, from which a limited number of 10 pixels per class are randomly selected as training samples. The corresponding classification maps are shown in Fig. 3.

To verify the effectiveness of the proposed ranking based clustering classification method by comparing the FDPC algorithm with four band selection techniques which include two ranking based techniques and two

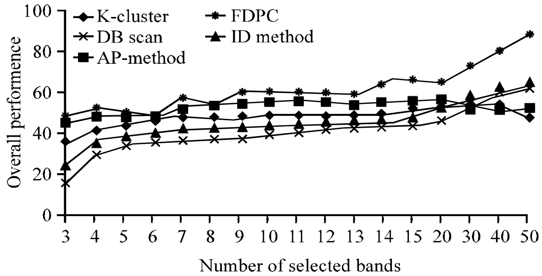

Fig. 4: Performance versus the selected bands of the different methods

Table 1: Sample dataset

Classes	Land cover type	No. of samples
C1	Roof	8806
C2	Road	1402
C3	Trails	1446
C4	Grass	1790
C5	Shadow	2423
C6	Tree	1194
Total		17061

clustering based techniques. The performance of the compared band selection methods is evaluated with different sample sizes, i.e., a fixed number of samples of the available labeled samples are selected from each class to form the training set. The remaining samples are then used as the test set for evaluation. Each experiment is repeated ten times with different training sets to reduce the influence of random effects and the average results are reported.

Figure 3 illustrates the classification results in which ten samples from each class are randomly picked out to form the training set. To investigate the impact of a different number of selected bands on the classification accuracy, five clustering based methods are applied (K-cluster, AP-method, ID method, DBscan and FDPC with MPCA) (Fig. 4). Compared with the four clustering-based methods, the FDPC with MPCA delivers the most stable and accurate results, proving the effectiveness of the proposed FDPC with MPCA approach for hyperspectral band selection.

Finally, the classification accuracy level as a function of the number of training samples in each class for the interval (3, ..., 15) is shown in Fig. 5. From experiments on the hyperspectral remote sensing data sets, some important results can be summarized. The spectral information of hyperspectral imagery contains large amounts of redundancy, thus, band selection is necessary and feasible. In general, the clustering-based methods usually achieve higher classification performance than the ranking based methods but the stability of the chosen band set and the relatively high complexity of the clustering-based methods are not as good as those of the

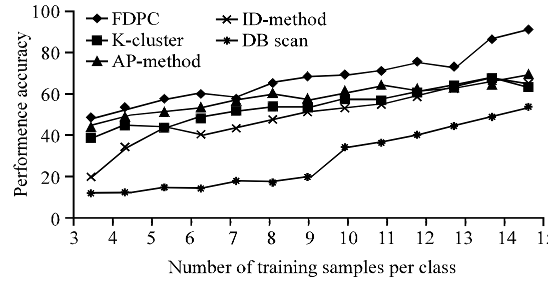

Fig. 5: Classification accuracy versus the different number of training samples

ranking-based methods. The proposed FDPC with MPCA clustering algorithm provides an artful band selection tool for combining the two kinds of methods. Compared with other methods, the FDPC with MPCA always produces good and stable classification performance with any classifier and in any hyperspectral data set.

CONCLUSION

In this study, we have presented a new ranking based clustering framework for dimension reduction in hyperspectural images. Particularly, we have develop a Fast Density Peek based Clustering (FDPC) algorithm and Multilinear Principle Component Analysis (MPCA) technique to reduce the data redundancy and increasing robustness. From our experimental analysis our proposed ranking based classification is more than efficient and robustness comparing with existing band selection techniques.

REFERENCES

Bioucas-Dias, J.M., A. Plaza, G. Camps-Valls, P. Scheunders and N. Nasrabadi *et al.*, 2013. Hyperspectral remote sensing data analysis and future challenges. IEEE. Geosci. Remote Sens. Mag., 1: 6-36.

Cardoso, J.F., 2003. Dependence, correlation and gaussianity in independent component analysis. J. Mach. Learn. Res., 4: 1177-1203.

Chang, C.I. and S. Wang, 2006. Constrained band selection for hyperspectral imagery. IEEE. Trans. Geosci. Remote Sens., 44: 1575-1585.

Chang, C.I., 2007. Hyperspectral Data Exploitation: Theory and Applications. John Wiley & Sons, Hoboken, NewJersey,USA.,ISBN:978-0-471-74697-3, Pages: 430.

Chang, C.I., Q. Du, T.L. Sun and M.L. Althouse, 1999. A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification. IEEE. Trans. Geosci. Remote Sens., 37: 2631-2641.

Cover, T.M. and J.A. Thomas, 1991. Elements of Information Theory. John Wiley and Sons, New York, USA.

Guo, B., S.R. Gunn, R.I. Damper and J.D. Nelson, 2006. Band selection for hyperspectral image classification using mutual information. IEEE. Geosci. Remote Sens. Lett., 3: 522-526.

Hughes, G., 1968. On the mean accuracy of statistical pattern recognizers. IEEE. Trans. Inf. Theor., 14: 55-63.

Khodr, J. and R. Younes, 2011. Dimensionality reduction on hyperspectral images: A comparative review based on artificial datas. Proceedings of the 4th International Congress on Image and Signal Processing (CISP'11) Vol. 4, October 15-17, 2011, IEEE, Shanghai, China, ISBN:978-1-4244-9304-3, pp: 1875-1883.

MartInez-UsoMartinez-Uso, A., F. Pla, J.M. Sotoca and P. Garcia-Sevilla, 2007. Clustering-based hyperspectral band selection using information measures. IEEE. Trans. Geosci. Remote Sens., 45: 4158-4171.

Shi, L., L. Zhang, J. Yang, L. Zhang and P. Li, 2013. Supervised graph embedding for polarimetric SAR image classification. IEEE. Geosci. Remote Sens. Lett., 10: 216-220.

Zhang, L., L. Zhang, D. Tao and X. Huang, 2013a. Tensor discriminative locality alignment for hyperspectral image spectral-spatial feature extraction. IEEE. Trans. Geosci. Remote Sens., 51: 242-256.

Zhang, X., Y. He, N. Zhou and Y. Zhengb, 2013. Semisupervised dimensionality reduction of hyperspectral images via. local scaling cut criterion. IEEE. Geosci. Remote Sens. Lett., 10: 1547-1551.