

Original Research Paper

The Use of Band Selection in Weighted Linear Prediction for Hyperspectral Image Classification

¹Ibtissam Banit'ouagua, ²Mounir Ait Kerroum and ³Ahmed Hammouch

¹UFR LRIT, Faculty of Sciences, Mohamed V University in Rabat, B.P. 1014, Morocco

²LaRIT Laboratory, Team Network, Telecommunications and Artificial Intelligence, Faculty of Science, University Ibn Tofail Kenitra, Morocco

³GTI-LGE, ENSET, B.P. 6207, Rabat, Morocco

Article history

Received: 13-04-2018

Revised: 28-08-2018

Accepted: 3-09-2018

Corresponding Author:

Ibtissam Banit'ouagua
UFR LRIT, Faculty of
Sciences, Mohamed V
University in Rabat, B.P. 1014,
Morocco
Email: ibtissamebani@gmail.com

Abstract: The advancement of imaging procedures has made hyperspectral sensors fit for acquiring spectral data in many restricted and bordering bands, which brings about a high relationship between's neighboring bands and high information excess. It is important to decrease these bands previously advance analysis utilizing land cover classification and target location. Going for the classification undertaking, this paper displays another weighted technique for band selection, in view of band comparability estimation through Weighted Linear Prediction-based Band Selection (WLPBS). This method removes bands according to the correlation using a weighted linear prediction criterion. This makes it less demanding to accelerate the learning procedure and to enhance overall classification accuracy. Experimental results using Support a Vectors Machine (SVM) classifier on Airborne Visible Infrared Imaging Spectrometer (AVIRIS) datasets showed the effectiveness of our WLPBS method to improves the classification accuracy and to select informative and distinctive bands compared with the widely used Linear Prediction-based Band Selection (LPBS) technique and to those in the state-of-the-art. The maximum result obtained for classification of the selected band from Indian Pine and Salinas's datasets successively was 91.07% and 95.30% using WLPBS.

Keywords: Band Selection, Data Whitening, Weighted Linear Prediction, Similarity Measurement, Hyper-Spectral Image, SVM Classifier

Introduction

Hyperspectral imaging is a quickly developing territory in remote sensing as announced in (Chang, 2003). It extends and enhances the capacity of multispectral image investigation. It exploits several coterminous spectral channels to reveal materials that typically can't be settled by multispectral sensors. Therefore, with such high spectral resolution, numerous unobtrusive articles and materials can presently be revealed and extricated by hyperspectral imaging sensors with exceptionally thin analytic otherworldly bands for detection, discrimination, classification, identification, recognition, and quantification. A considerable lot of its applications are yet to be investigated. It has been the presence of mind to consider hyperspectral imaging as a characteristic expansion of multispectral imaging with band development (Chang, 2007). We about dependably

will confront a mixed pixel problem where the estimation speaks to a reaction from a composite of different materials. For instance, a field of grass contains cutting edges ordinarily of various species mixed together with different weeds, in addition to commitments from the hidden soil that can be comprised of different natural mixes and soil composes (Chang, 2007).

We perceive, as detailed in (Chang, 2003), that one of the awesome difficulties for hyperspectral imaging is subpixel recognition, which isn't dealt with in standard spatial-based image processing. After a target is recognized, the following stage is to characterize identified focuses as indicated by their spatial or otherworldly properties. The main contribution of the present work is a new approach called WLPBS for band selection to select informative bands for the classification task.

Hyperspectral band selection, which happens in a diversity of algorithms for information preprocessing, has gotten extensive consideration, with numerous researchers consistently advancing strategies and thoughts of their own. As to these techniques, they can be generally assembled into two sorts, namely supervised and unsupervised. The supervised methods need some earlier data, for example, training samples and target signatures. This is seen for example in (Kuo *et al.*, 2014; Yang *et al.*, 2011). In general, these techniques initially characterize a few criteria, for example, the Jeffreys-Matusita (JM) distance (Yang *et al.*, 2011), class divergence and signature angle (Keshava, 2004). At that point, the subset of bands that optimizes the criteria for training samples is selected. However, because the prior information is habitually not available in practice, supervised techniques are not appropriate for hyperspectral band selection (Martinez-Uso *et al.*, 2007). Hence, this paper centers unsupervised band selection methods.

The unsupervised band selection systems are usually built in data assessment techniques and attempt to choose the subset that has the most extreme information (Chavez *et al.*, 1982; Chang *et al.*, 1999), maximum information divergence (Chang and Wang, 2006), or other information measurements (Jia *et al.*, 2012; Mojaradi *et al.*, 2008). However, these prioritization-based strategies do not take band correlations into consideration. Thus, the bands acquired by these strategies are highly correlated in general. Others consider band correlation, for instance with a new similarity measure. The criterion of Linear Prediction Error (LPE) (Du and Yang, 2008) decreases the dimensionality by choosing a subset of spectral bands with distinctive and informative bands. However, a large portion of these supervised or unsupervised techniques considers band information into account. Though, few works focus on the quality of these bands. For the most part, the high level of outliers in bands can bring about extensive differences. Therefore, the outlying observations (pixels) in bands will have higher selection weights if the quality of the band is ignored, which will most likely prompt absurd outcomes. To overcome these limitations, the aim of this research was to give smaller weights to possible outliers observations using a weighted least-square function.

In this study, an enhancement to reduce the influence of outliers in the algorithm of selection and minimize the correlation between the bands selected is proposed using the Weighted Linear Prediction (WLP) criterion. The proposed feature selection method is called Weighted Linear Prediction-based Band Selection (WLPBS).

To guarantee that they chose bands are particular as well as instructive, data preprocessing, including terrible band pre-removal and data whitening, is required.

Experimental results demonstrate that our method WLPBS reduces the influence of outliers and minimizes

the correlation between selected bands. The results with AVIRIS databases demonstrate that the proposed similarity-based band selection method has an obvious performance improvement in terms of information conservation and class separability compared with the widely used technique shown in (Du and Yang, 2008) and to those in the state-of-the-art MIFS (Banit'ouagua *et al.*, 2016) and MI-est (Guo *et al.*, 2006) methods, based on the selection of complementary bands via MI evaluation.

The rest of this paper is organized as follows. Section 2 reviews some relevant works on the topic of unsupervised hyperspectral band selection. In section 3 we present the proposed approach. The experimental results are presented in section 4. Finally, we give some closing comments in section 5.

Related Works

In the important writing, hyperspectral bands are frequently described by susceptible spectral information. Specialists have examined different unsupervised criteria to recognize these bands, for example, the spectral distance, band correlation, orthogonally, Linear Prediction (LP), Mutual Information (MI) and Kullback-Leibler Divergence (KLD). Among these criteria, the spectral distance is conceivably the most fundamental and well known.

Some unsupervised techniques take the linear characteristics of bands into thought and look for the particular bands in the vector space. The latter procedure can be separated into the accompanying two categories: One removing distinctive pixels based on similarity measurement and the other utilizing the geometry concept, for example, a simplex strategy. The endmember extraction algorithms employing Unsupervised Fully Compelled Least-Squares Linear Unmixing (UFCLSLU) in (Heinz and Chang, 2001) and Orthogonal Subspace Projection (OSP) in (Ren and Chang, 2003) have a place with the primary category, though the well-known pixel purity index (Boardman *et al.*, 1995) and NFINDR calculations (Winter, 1999) (Wang *et al.*, 2007) have a place in the second category. The linear prediction error is likewise used to decrease the bands by estimating the similarity between bands (Du and Yang, 2008). Sun *et al.* (2014), an Autocorrelation network based Band Selection (ACMBS) strategy, which utilizes the base Linear Prediction (LP) error as the selection criterion and searches the suboptimal subset by Sequential Backward Selection (SBS) acquired promising results. The authors of (Han *et al.*, 2017) propose an Improved Similarity Measurement method in light of LP (ISMLP) for hyperspectral ocean ice identification.

The clustering methods are frequently used to select a representative and different bands (Martinez-Uso *et al.*, 2007; Cao *et al.*, 2016). Positioning based techniques will rank bands in light of some ranking scores (Sun *et al.*,

2015). Mutuelle Information (MI) is also favored by numerous researchers for their nonlinear attributes (Banit'ouagua *et al.*, 2016; Guo *et al.*, 2006).

In view of the defined criteria, two major kinds of optimization strategies, i.e., an incremental (Whitney, 1971) search and an integrated search (Marill and Green, 1963), are utilized in the literature to search for the particular bands. An incremental pursuit is a 'down-top' strategy which method which starts with a single band followed by incrementally adding new bands. If the combination of a certain band and the preselected bands meet the optimal criterion, this band is recognized as the new band to be included. The procedure is repeated until the point that the quantity of bands is sufficiently extensive. An integrated search is the complement of an incremental search. It is a 'down-top' search method and treats the complete set of bands as the candidates to be selected at first. Then, the bands in the candidate set are detached one by one until reaching the wanted number of remaining selected bands.

Without a reasonable stopping criterion, the band selection process may run pointlessly long or perhaps forever depending on search strategy. Optimization strategies and evaluation functions can impact the decision of a stopping criterion. Stopping criteria based on a generation procedure include: (i) When a predefined number of features is chosen and (ii) when a predefined number of emphases is come to. Stopping criteria based on an evaluation function can include: (i) When the expansion (or erasure) of any element does not produce a better subset; and (ii) when an ideal subset as indicated by some evaluation function is acquired. The loop proceeds until the stopping criterion is satisfied (Dash and Liu, 1997).

As of late, numerous pertinent works have turned to the idea of Virtual Dimensionality (VD) (Chang and Liu, 2014) or subspace identification strategies (Ghamary Asl *et al.*, 2014) to estimate the number of spectrally distinct signatures that describe the information.

Similarity-Based Band Selection

Data Preprocessing

The experimental results in (Du and Yang, 2008) recommend that using data whitening (after bad band deletion) for selected bands may offer a somewhat preferable execution over utilizing original bands. This is because the noise component in different bands is varied and if the noise component is larger, a band may appear to be different from others, although it may not be informatively distinct. In this way, in this study, we apply data whitening to the original bands (after bad band removal), which can be easily achieved by the eigendecomposition of the data

covariance matrix. The readers are referred to (Du *et al.*, 2003) for a demonstration that the net effect of noise whitening and data whitening is similar.

Linear Prediction-Based Band Selection Method

To choose the most unique or different bands, a likeness metric should be assigned. The authors of (Du and Yang, 2008) utilized an approach where band similarity can be assessed mutually rather than pairwise. It starts with the best band-band combination and afterward the two-band combination is in this way expanded to three, four, et cetera, until the desired numbers of bands are selected. Linear Prediction (LP), a new criterion for similarity comparison, can together assess the similarity between a solitary band and various bands.

There are two parameters for LP-based Band Selection (BS): Initial bands and number of bands to be selected. The last must be pre-decided. The accompanying algorithm can be utilized to locate the first two best groups (Du and Yang, 2008).

1. Arbitrarily select a band B_1 from the original data set and undertaking the various $N-1$ bands to its orthogonal subspace $\langle B_1 \rangle^\perp$. Discovery the band B_2 with the most extreme projection in $\langle B_2 \rangle^\perp$, which is viewed as the most, not at all like B_1
2. For the band B_2 , project the various $N-1$ bands to its orthogonal subspace $\langle B_2 \rangle^\perp$. Find the band B_3 with the most extreme projection in $\langle B_2 \rangle^\perp$
3. In the event that $B_3 = B_1$, the algorithm is terminated because B_1 and B_3 are affirmed to the combine with the most noteworthy dissimilarity. At that point either B_1 or B_3 can be utilized as the underlying band. If $B_3 \neq B_1$, go to the following stage
4. For B_i , locate the most dissimilar band until $B_{i+1} = B_{i-1}$ at that point, either B_{i-1} or B_i can be utilized as the initial band (or the two bands are utilized as the initial band pair)

LP-based BS exploits the advantage of a simple algorithm concept and has shown high productivity in selecting appropriate bands. In view of the hypothesis of this strategy, the high linear independence of bands shows their appropriateness for selection as best bands. Though, this criterion does not generally fulfill the classification purpose it may fail when the candidate band to be selected has many outliers, as these outlying observations (pixels) in the band will have higher selection weights. This leads to an unreasonable result. To resolve this problem, we proposed WLPBS, which has both the ability to exclude the correlated bands and minimize the weights of outlying observations (pixels) in the band.

Proposed Algorithm: Weighted Linear Prediction-Based Band Selection Method

The basic steps of the WLP-based BS algorithm can be described as follows:

1. As in Linear Prediction-based Band Selection (LPBS), initialize the algorithm by choosing a pair of bands B_1 and B_2 ; the selected subset is denoted as $S = \{B_1, B_2\}$
2. Find a third band B_3 that is the most dissimilar to all the bands in the current S by using a certain criterion (discussed below) and then the selected band subset is updated as $S = S \cup B_3$
3. Continue step 2 until the stopping condition is satisfied

The criterion used in the WLPBS algorithm is described as takes after. Expect that the present band subset S incorporates B_1 and B_2 . To find a third band that is the most different to B_1 and B_2 , let each band B be assessed as:

$$a_0 + a_1 B_1 + a_2 B_2 = B' \tag{1}$$

where, B' is the weighted linear prediction of band B by B_1 and B_2 and a_0, a_1 and a_2 are the parameters that can minimize the linear prediction error: $e = \|B-B'\|$. Using a weighted least squares solution, $a = (a_0, a_1, a_2)^T$ can be estimated as:

$$a = (X^T W X)^{-1} X^T W y \tag{2}$$

Where X is an $N \times 3$ matrix whose first column is one, the second column includes all the pixels in B_1 and third column includes all the pixels in B_2 . y is $N \times 1$ vector with all the pixels in B . W is a matrix that is used to weight the outlier observations (pixels) in bands. WLS uses a kernel to weight nearby observations more heavily than other observations. The kernel assigns a weight given by:

$$W(i,i) = \exp\left(-\left(\frac{\|x_i - x\|}{k}\right)^2\right) \tag{3}$$

Where k is the number of already selected band, x_i and x are respectively candidate and estimate bands. This builds the weight matrix W , which has only diagonal elements.

The band that yields the greatest error e_{\min} (utilizing the ideal parameters in a) is considered as the most dissimilar band to B_1 and B_2 and will be chosen as B_3 then $S = \{B_1, B_2, B_3\}$. The schematically depiction of WLPBS approach is as appeared in Fig. 1.

The stopping condition: The research strategy in WLPBS stops when the linear prediction error cannot be further optimized. The performances of previously mentioned algorithms are assessed on two real AVIRIS datasets (Indian Pine and Salinas) as far as dimensionality reduction and classification accuracy utilizing SVM classifier. Every database has been separated into two parts: the training-base and the test one. Half of the pixels from each class were arbitrarily decided for training, with the staying 50% for the test set on which performance was evaluated.

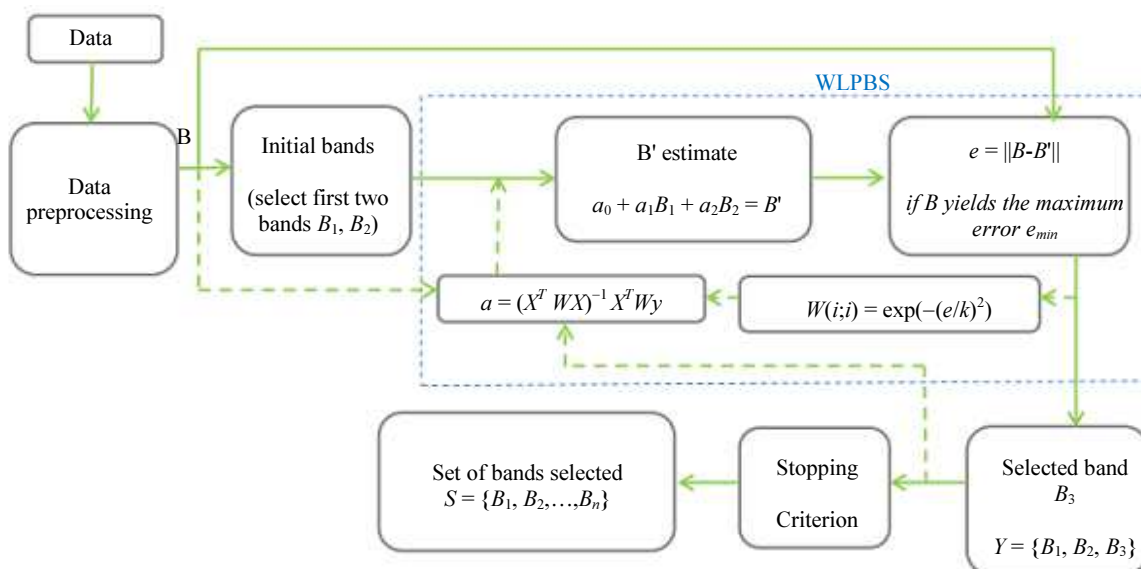


Fig. 1: Schematical description of weighted linear prediction-based band selection approach

The results presented here are comparable with the results of the original unsupervised approach LPBS (Du and Yang, 2008) and to those in the state-of-the-art MIFS (Banit'ouagua *et al.*, 2016) and MI-est (Guo *et al.*, 2006) methods, based on the selection of complementary bands via MI evaluation. We point out that the performances of MIFS outperform (Banit'ouagua *et al.*, 2016) any MI-based method using the associated reference map, an entropy-based method, and a correlation-based method. This is the reason why we chose this method to advocate the idea of using WLPBS algorithm to perform bands selection in hyperspectral images, by comparing its results.

The evaluation of our band reduction approach is performed by supervised classification utilizing Support Vector Machine (SVMs). One-Vs-One multiclass SVMs were trained attributable to the accessibility of ground truth and tested on the reduced datasets. For this, we utilize the Library for Support Vector Machines (LIBSVM) package with the Gaussian Radial Basis Function (RBF) as in past work (Banit'ouagua *et al.*, 2016). To improve the penalty parameter c and the aperture γ a cross-validation strategy is performed. The classification accuracy is processed for band reduction techniques.

It should be noted that the whitened data are used only for band selection processes, while the classifications are carried out on original data.

Experiment

Indian Pine Experiments

Data Description

Indian Pines scene: This image is a public hyperspectral dataset, which was acquired over a test site called Indian Pine in northwestern Indiana, U.S.A. The AVIRIS sensor nominally collects 224 spectral reflectance bands of data in the [0.4-2.5]- μm Visible and Near-Infrared (VNIR) range.

Among them, four contain only zeros and are discarded. Therefore, 220 bands from the 92AV3C dataset are used for the experiments. Each of the 220 band images is of 145×145 pixels. Around 49% of pixels are grouped into 16 different classes (Table 1). As to the remaining pixels, it is difficult to group them into any of the existing classes and they are identified as the background.

In our experiments, as specified in (Du and Yang, 2008) we removed the noisiest bands from this dataset and kept 196 spectral bands as the initial data for every method. More precisely, we removed bands [1.3], [104.108], [150.163] and [218.220].

Owing to a large number of pixels in the remote sensing image, the sizes of the matrices in the data processor are large, which decreases proficiency. Be that as it may, utilization of just a moderately little subset of pixels in the band determination process does not change the outcomes as a rule (Du and Yang, 2008) (Han *et al.*, 2017). This is on the grounds that a high spatial correlation exists between bands of hyperspectral data. Here, for band decrease, we utilized just 10% of N pixels scattered all through the labeled pixels to decide the bands to be selected.

Results and Discussion

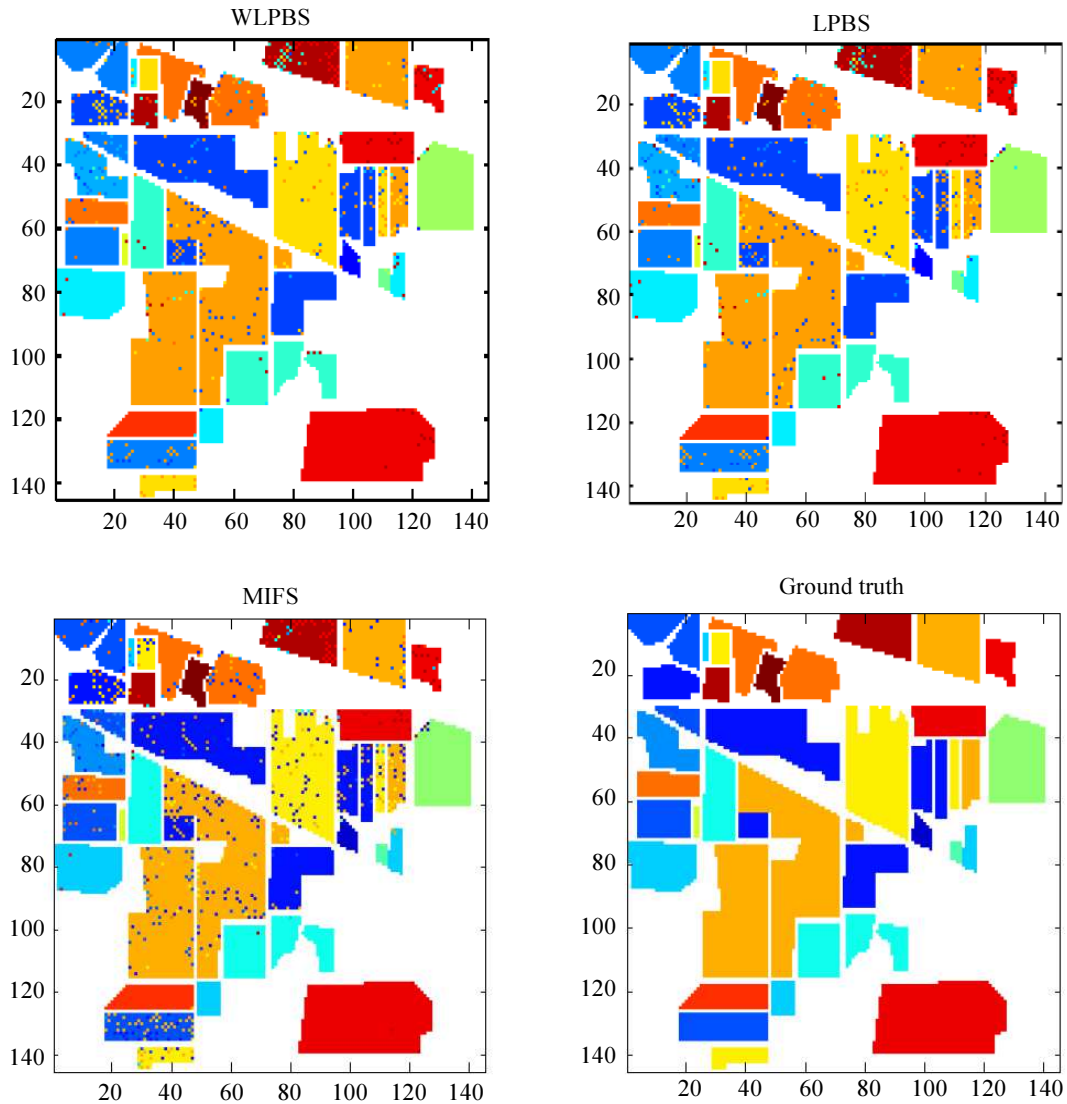
Table 2 shows the overall classification accuracy using an SVM classifier from the Indian Pine data. As we can see, the proposed WLP-based band selection algorithm significantly outperforms the LPBS, MIFS (Banit'ouagua *et al.*, 2016) and MI-est (Guo *et al.*, 2006) methods. The reason is that LPBS is influenced by outliers while WLPBS minimizes their weight. Accordingly, the bands selected by WLPBS are the best for classification. Compared to the MIFS algorithm (Banit'ouagua *et al.*, 2016) WLPBS fast in terms of execution time and does not require many parameters because it's used just 10% of N pixels scattered throughout the labeled pixels to determine the bands to be selected.

Table 1: Number of training and testing pixels in each class for the Indian Pine and Salinas datasets

| Indian Pine | | | | Salinas | | |
|-------------|------------------------|------------------------|-----------------------|---------------------------|------------------------|-----------------------|
| No | Class | Pixels in training set | Pixels in testing set | Class | Pixels in training set | Pixels in testing set |
| 1 | Alfalfa | 29 | 25 | Broccoli-green-weeds-1 | 1005 | 1004 |
| 2 | Corn-notill | 719 | 715 | Broccoli-green-weeds-2 | 1863 | 1863 |
| 3 | Corn-min | 419 | 415 | Fallow | 988 | 988 |
| 4 | Corn | 117 | 117 | Fallow-rough-plow | 697 | 697 |
| 5 | Grass/Pasture | 249 | 248 | Fallow-smooth | 1339 | 1339 |
| 6 | Grass/Trees | 374 | 373 | Stubble | 1980 | 1979 |
| 7 | Grass/pasture-mowed | 13 | 13 | Celery | 1790 | 1789 |
| 8 | Hay-windrowed | 243 | 246 | Grapes-untrained | 5636 | 5635 |
| 9 | Oats | 10 | 10 | Soil-vinyard-develop | 3102 | 3101 |
| 10 | Soybeans-notill | 483 | 485 | Corn-senesced-green-weeds | 1639 | 1639 |
| 11 | Soybeans-min | 1234 | 1234 | Lettuce-romaine-4wk | 534 | 534 |
| 12 | Soybean-clean | 304 | 310 | Lettuce-romaine-5wk | 964 | 963 |
| 13 | Wheat | 108 | 104 | Lettuce-romaine-6wk | 458 | 458 |
| 14 | Woods | 644 | 650 | Lettuce-romaine-7wk | 535 | 535 |
| 15 | Bldg-Grass-Tree-Drives | 190 | 190 | Vinyard-untrained | 3634 | 3634 |
| 16 | Stone-steel towers | 46 | 49 | Vinyard-vertical-trellis | 904 | 903 |

Figure 4 presents the linear prediction error for 25 bands selected using WLPBS and LPBS. We can observe from this figure that WLPBS provides the lowest prediction error (the optimal parameters in a) as compared to LPBS for the

Indian Pine data. With WLPBS we have the best-estimated band, used to select the most dissimilar band to B_1 and B_2 in order to maximize e_{\min} . LPBS does not exceed WLPBS in the quality of the band selected.



| | | | | |
|---------------|----------------|------------------------------|---------------------|------|
| Unlabeled | Alfalfa | Corn-notill | Corn-mintill | Corn |
| Grass-pasture | Grass-tress | Grass-pasture-moved | Hay-windrowed | |
| Oats | Soybean-notill | Soybean-mintill | Soybean-clean | |
| Wheat | Woods | Buildings-grass-trees-drives | Stones-steel-towers | |

Fig. 2: The ground truth map of AVIRIS 92AV3C and a thematic map of 25 selected bands using the WLPBS, LPBS and MIFS algorithms

The experimental results in Fig. 4 propose that utilizing WLPBS for select the bands may offer a marginally preferred performance over utilizing LPBS. This is on account of the outliers observations in various bands are changed, if the outliers observation is bigger, a band may appear to be different from others, despite the fact that it may not be informatively distinct.

Without minimizing the weight of outliers, the bands selected by LPBS are almost similar, which justifies the low variation of LPE for LPBS method.

Figure 2 presents classification maps when using 25 bands selected by WLPBS and LPBS. We can see that 25 bands are adequate to distinguish materials contained in the district (Table ref tab 4). It additionally appears in this figure that the thematic map delivered utilizing the 25 bands chosen by WLPBS algorithm is the most like the Ground Truth (GT).

Salinas Experiments

Data Description

Salinas scene was gathered utilizing the 224 bands AVIRIS sensor over Salinas Valley, California and is described by high spatial resolution (3.7 m pixels).

The region covered includes 512 lines \times 217 samples. In this study, the utilized Salinas scene data have well-labeled ground references containing 16 classes (Table 1). This datum comprises 205 spectral bands, after noisy band elimination ([108-112], [154-167] and 224). In our experiments with the AVIRIS Salinas data set in Table 3, the size of the smaller classes is bigger when compared to those in the AVIRIS Indian Pines data set. We decided to use only 1% of the available ground-truth pixels in Table 1 to compute the relevant bands to be selected.

Results and Discussion

The supervised classification results got from Salinas band sets are shown in Table 3. This can be ascribed to the fact that LPBS method is based on band correlation and has little consideration of data quality. In this manner, classification results in light of the bands gotten by this strategy are less accurate using whitened bands. The most elevated accuracy is accomplished when utilizing WLPBS, that considers band correlations and outliers.

The results are given in Fig. 5 showing that the WLPBS algorithm achieves better performance in LPE reduction (best B' with optimal parameters in a) than LPBS in the Salinas dataset. WLPBS selects the most dissimilar set of bands.

Figure 3 presents classification maps for 25 selected bands using WLPBS and LPBS. As in the Indian Pine scene, the thematic map produced using 25 bands selected from the Salinas dataset with the WLPBS algorithm is most similar to the GT map. Table 4 presents the classification accuracy of each class, with respect to several methods.

The Selected Number of Bands (NB)

The VD utilized by (Du and Yang, 2008) might be a sensible marker of the fitting number of bands to be selected. However, this criterion may work effectively when noise is independently and identically distributed. This is, unfortunately, not true when is applied to real image data. For this reason, we here use LPE as a criterion to stop the selection process, we stop the selection if any band from the candidate ones can't optimize more the linear prediction error.

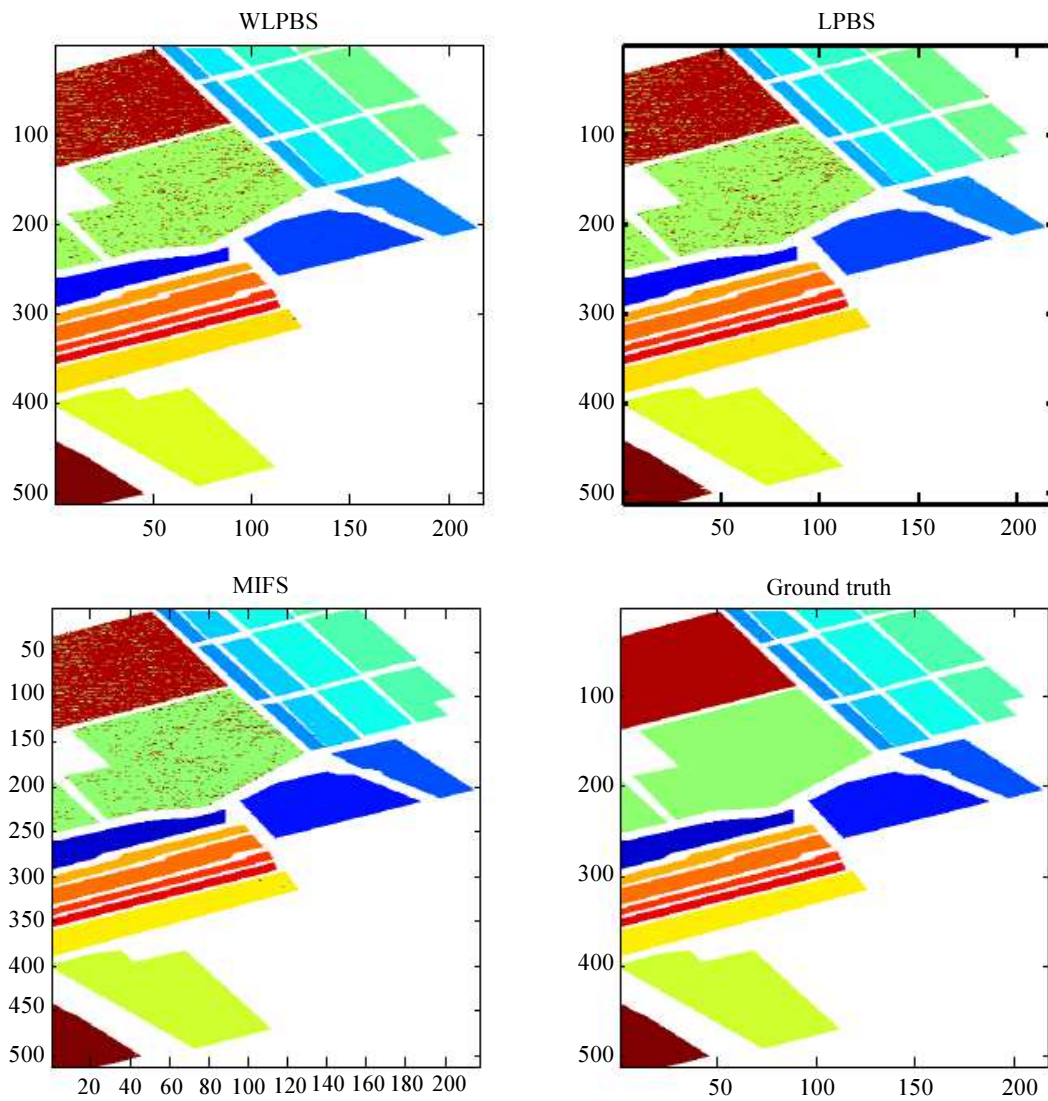
Table 2: A comparison of classification accuracy for Indian pine dataset

| Number of Bands | Indian Pine data | | | |
|-----------------|------------------|-------|-------|---------|
| | WLPBS | LPBS | MIFS | MI-est* |
| 20 | 88.00 | 86.82 | 85.19 | 86.57 |
| 25 | 88.79 | 87.52 | 88.02 | 87.18 |
| 30 | 90.20 | 87.70 | 88.35 | 88.23 |
| 35 | 90.84 | 88.02 | 89.06 | 89.00 |
| 40 | 91.07 | 89.01 | 88.75 | 89.38 |

Notes: The outcomes held of MI-est [marked by (*)] is the result of Guo *et al.* (2006).

Table 3: A comparison of classification accuracy for Salinas dataset

| Number of Bands | Salinas data | | |
|-----------------|--------------|-------|-------|
| | WLPBS | LPBS | MIFS |
| 20 | 94.46 | 93.85 | 93.99 |
| 25 | 94.93 | 94.50 | 94.61 |
| 30 | 95.11 | 94.65 | 94.39 |
| 35 | 95.24 | 94.88 | 94.80 |
| 40 | 95.30 | 94.98 | 95.07 |



| | | | | |
|----------------------|--------------------------|----------------------|-------------------------|-------------------|
| Uabeled | Brocoli-greenweeds-1 | Brocoli-greenweeds-2 | Fallow | Fallow-rough-plow |
| Fallow-smooth | Stubble | Celery | Grapes-untrained | |
| Soil-vinyard-develop | Corn-senesced-greenweeds | Lettuce-romaine-4wk | Lettuce-romaine-5wk | |
| Lettuce-romaine-6wk | Lettuce-romaine-7wk | Vinyard-untrained | Vinyard-verticaltrellis | |

Fig. 3: The ground truth map of Salinas dataset and thematic map of 25 selected bands selected with the WLPBS, LPBS and MIFS algorithms

For the Indiana Pine scene, $NB = 25$ when the $LPE = 10^{-3}$ for the WLPBS method and $NB = 25$ when $LPE = 10^{-1}$ for the LPBS algorithm (Fig. 4). When 25 bands were selected, the classification accuracy from the WLPBS was 88.79% and for LPBS it was

87.50% (Table 2). If the LPE decreases, the number of bands selected and the classification values are increased. Thus, the LPE might be a sensible pointer of the fitting number of bands to be selected in the real data.

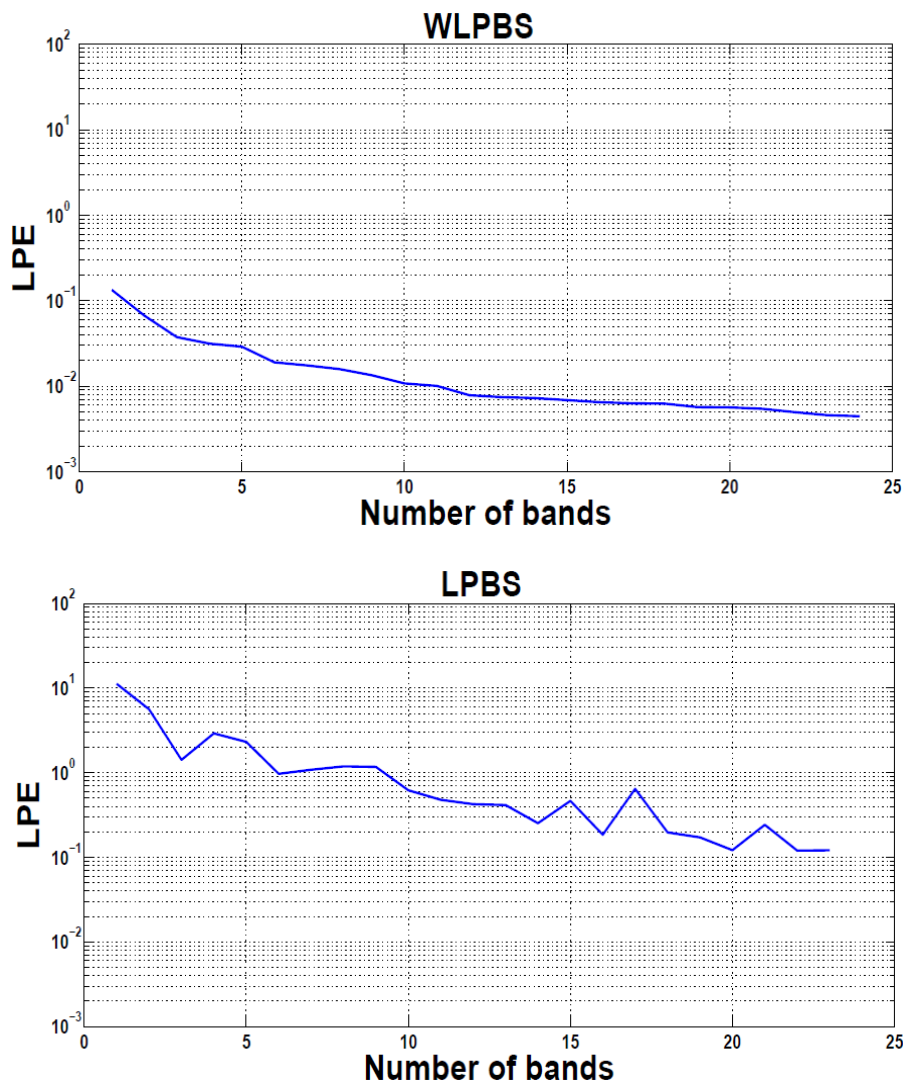
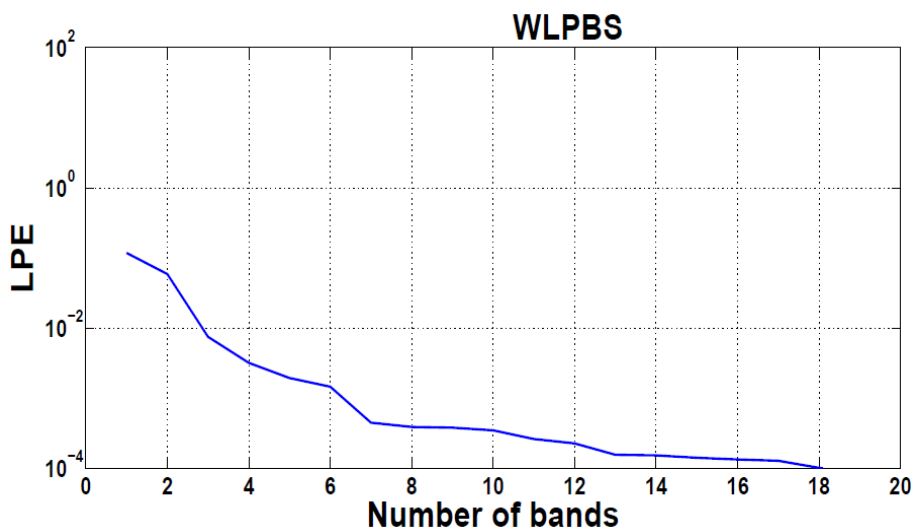


Fig. 4: The linear prediction error (LPE) for 25 bands selected by LPBS and WLPBS for Indian Pine dataset



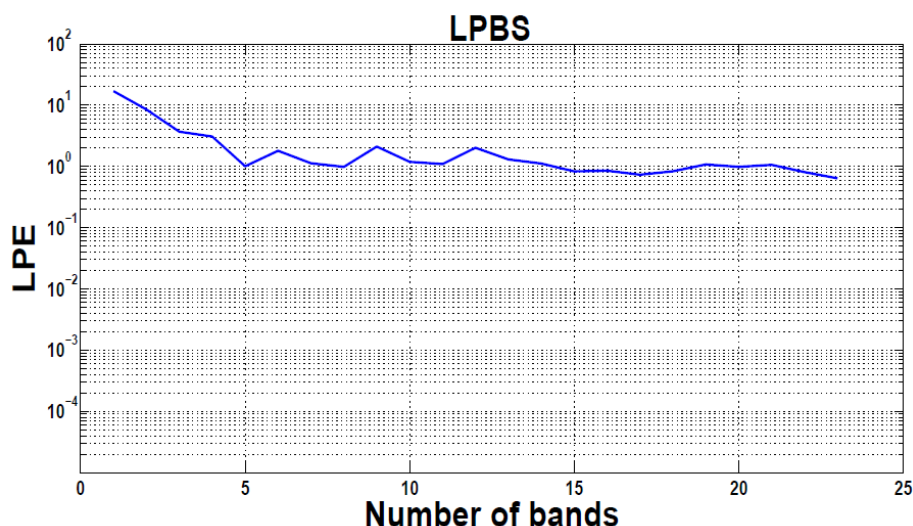


Fig. 5: Linear prediction error (LPE) for 25 bands selected by LPBS and WLPBS for Salinas dataset

Table 4: Classification result using the 25 selected bands with the WLPBS, LPBS and MIFS algorithms

| Class No | Accuracy % | | | | | |
|----------|-------------|-------|--------|---------|--------|--------|
| | Indian Pine | | | Salinas | | |
| | WLPBS | LPBS | MIFS | WLPBS | LPBS | MIFS |
| 1 | 92.00 | 84.00 | 100.00 | 100.00 | 100.00 | 99.60 |
| 2 | 85.03 | 83.92 | 85.59 | 100.00 | 100.00 | 100.00 |
| 3 | 80.96 | 81.45 | 78.07 | 99.39 | 100.00 | 99.60 |
| 4 | 82.05 | 79.49 | 83.76 | 99.57 | 99.43 | 99.43 |
| 5 | 97.98 | 97.98 | 96.77 | 99.33 | 99.48 | 99.18 |
| 6 | 97.59 | 97.05 | 98.93 | 99.85 | 99.80 | 99.90 |
| 7 | 92.31 | 92.31 | 92.31 | 100.00 | 100.00 | 99.83 |
| 8 | 97.97 | 97.97 | 97.97 | 87.38 | 86.64 | 87.97 |
| 9 | 100.00 | 90.00 | 100.00 | 99.71 | 99.74 | 99.58 |
| 10 | 89.90 | 86.39 | 80.00 | 99.15 | 99.21 | 98.54 |
| 11 | 87.60 | 84.36 | 86.22 | 99.63 | 99.63 | 99.63 |
| 12 | 86.45 | 82.26 | 84.52 | 99.90 | 99.79 | 100.00 |
| 13 | 99.04 | 99.04 | 99.04 | 99.78 | 99.78 | 99.56 |
| 14 | 94.62 | 95.38 | 96.00 | 99.07 | 98.50 | 98.69 |
| 15 | 62.11 | 70.00 | 75.79 | 83.30 | 81.12 | 80.68 |
| 16 | 100.00 | 95.92 | 100.00 | 99.78 | 99.89 | 99.67 |

Conclusion

Band selection in hyperspectral imaging is imperative, regardless of the learning algorithm which is utilized to train the hyperspectral dataset. Because of the presence of irrelevant and redundant bands, by choosing just the relevant bands of the data, higher predictive accuracy can be obtained. In this study, the problem of band selection (selection of just the important band in the classification procedure) is achieved.

Among the numerous arrangements of this problem, algorithm-based band similarity measures are ideal for they think about band correlations while the bands gotten by the others are highly correlated in general.

Existing LPBS, one of these algorithms, excludes correlated bands effectively, but it may fail when the candidate band to be selected has many outliers.

To resolve this problem, an unsupervised band selection method called WLPBS has been proposed. WLPBS has the ability to both exclude the correlated bands and minimize the weights of outlying observations (pixels) in the band.

The investigations demonstrate that WLPBS can remove the bands with low qualities and accordingly get high classification accuracy. Nonetheless, the major problem of the WLP-based band selection is that computational cost is high all the pixels are used. Later on, we will build up The N-FINDR algorithm

for pixel selection and then the selected pixels are used for band selection.

Acknowledgment

We would like to dedicate this work to Driss Aboutajdine, who died in March 2017.

Author's Contributions

Ibtissam Banit'ouagua: Development of the main idea, experimental tests execution, and text writing.

Mounir Ait Kerroum: Participated in all experiments, coordinated the data analysis and contributed to the writing of the manuscript.

Ahmed Hammouch: Designed the research plan and organized the study.

Ethics

Authors confirm that this manuscript has not been published elsewhere and that no ethical issues are involved.

References

- Banit'ouagua, I., M. Ait Kerroum, A. Hammouch and D. Aboutajdine, 2016. Band selection by mutual information for hyper-spectral image classification. *Int. J. Adv. Intell. Paradigms*, 8: 98-118. DOI: 10.1504/IJAIP.2016.074791
- Boardman, J.W., F.A. Kruse and R.O. Green, 1995. Mapping target signatures via partial unmixing of AVIRIS data. *Proceedings of the Summaries JPL Airborne Earth Science Workshop, (ESW' 95)*.
- Cao, X., J. Han, S. Yang, D. Tao and L. Jiao, 2016. Band selection and evaluation with spatial information. *Int. J. Remote Sens.*, 37: 4501-4520. DOI: 10.1080/01431161.2016.1214301
- Chavez, P.S., G.L. Berlin and B. Sowers, 1982. Statistical method for selecting Landsat MSS ratios. *J. Applied Photograph. Eng.*, 1: 23-30.
- Chang, C.I. and K.H. Liu, 2014. Progressive band selection of spectral unmixing for hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.*, 52: 2002-2017. DOI: 10.1109/TGRS.2013.2257604
- Chang, C.I. and S. Wang, 2006. Constrained band selection for hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.*, 44: 1575-1585. DOI: 10.1109/TGRS.2006.864389
- Chang, C.I., 2003. *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. 1st Edn., Springer Science and Business Media, New York, ISBN-10: 0306474832, pp: 370.
- Chang, C.I., 2007. *Hyperspectral Data Exploitation: Theory and Applications*. 1st Edn., John Wiley and Sons, Hoboken, ISBN-10: 047012461X, pp: 456.
- Chang, C.I., Q. Du, T.L. Sun and T.L. Sun, 1999. A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.*, 37: 2631-2641. DOI: 10.1109/36.803411
- Dash, M. and H. Liu, 1997. Feature selection for classification. *Intell. Data Anal.*, 1: 131-156. DOI: 10.1016/S1088-467X(97)00008-5
- Du, Q. and H. Yang, 2008. Similarity-based unsupervised band selection for hyperspectral image analysis. *IEEE Geosci. Remote Sens. Lett.*, 5: 564-568. DOI: 10.1109/LGRS.2008.2000619
- Du, Q., H. Ren and C.I. Chang, 2003. A comparative study for orthogonal subspace projection and constrained energy minimization. *IEEE Trans. Geosci. Remote Sens.*, 41: 1525-1529. DOI: 10.1109/TGRS.2003.813704
- Ghamary Asl, M., M.R. Mobasheri and B. Mojaradi, 2014. Unsupervised feature selection using geometrical measures in prototype space for hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.*, 52: 3774-3787. DOI: 10.1109/TGRS.2013.2275831
- Guo, B., S.R. Gunn, R.I. Damper and J.D.B. Nelson, 2006. Band selection for hyperspectral image classification using mutual information. *IEEE Geosci. Remote Sens. Lett.*, 3: 522-526. DOI: 10.1109/LGRS.2006.878240
- Han, Y., J. Li, Y. Zhang, Z. Hong and J. Wang, 2017. Sea ice detection based on an improved similarity measurement method using hyperspectral data. *Sensors*, 17: 1124-1124. DOI: 10.3390/s17051124
- Heinz, D.C. and C.I. Chang, 2001. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.*, 39: 529-545. DOI: 10.1109/36.911111
- Jia, S., Z. Ji, Y. Qian and L. Shen, 2012. Unsupervised band selection for hyperspectral imagery classification without manual band removal. *IEEE J. Sel. Top. Applied Earth Observ. Remote Sens.*, 5: 531-543. DOI: 10.1109/JSTARS.2012.2187434
- Keshava, N., 2004. Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. *IEEE Trans. Geosci. Remote Sens.*, 42: 1552-1565. DOI: 10.1109/TGRS.2004.830549
- Kuo, B.C., H.H. Ho, C.H. Li, C.C. Hung and J.S. Taur, 2014. A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification. *IEEE J. Sel. Top. Applied Earth Observ. Remote Sens.*, 7: 317-326. DOI: 10.1109/JSTARS.2013.2262926

- Marill, T. and D.M. Green, 1963. On the effectiveness of receptors in recognition systems. *IEEE Trans. Inform. Theory*, 9: 11-17.
DOI: 10.1109/TIT.1963.1057810
- Martinez-Uso, A., F. Pla, J.M. Sotoca and P. García-Sevilla, 2007. Clustering-based hyperspectral band selection using information measures. *IEEE Trans. Geosci. Remote Sens.*, 45: 4158-4171.
DOI: 10.1109/TGRS.2007.904951
- Mojaradi, B., H. Emami, M. Varshosaz and S. Varshosaz, 2008. A novel band selection method for hyperspectral data analysis. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 37: 447-454.
DOI: 10.1016/j.engappai.2013.07.010
- Ren, H. and C.I Chang, 2003. Automatic spectral target recognition in hyperspectral imagery. *IEEE Trans. Aerosp. Electron. Syst.*, 39: 1232-1249.
DOI: 10.1109/TAES.2003.1261124
- Sun, K., X. Geng and J. Luyan, 2015. Exemplar component analysis: A fast band selection method for hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett.*, 12: 998-1002.
DOI: 10.1109/LGRS.2014.2372071
- Sun, K., X. Geng and L. Ji, 2014. An efficient unsupervised band selection method based on an autocorrelation matrix for a hyperspectral image. *Int. J. Remote Sens.*, 35: 7458-7476.
DOI: 10.1080/01431161.2014.968686
- Wang, L., X. Jia and Y. Zhang, 2007. A novel geometry-based feature-selection technique for hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett.*, 4: 171-175.
DOI: 10.1109/LGRS.2006.887142
- Whitney, A.W., 1971. A direct method of non parametric measurement selection. *IEEE Trans. Comput. C*, 20: 1100-1103.
DOI: 10.1109/T-C.1971.223410
- Winter, M.E., 1999. NFINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data. *Proc. SPIE*, 3753: 266-275.
DOI: 10.1117/12.366289
- Yang, H., Q. Du, H. Su and Y. Sheng, 2011. An efficient method for supervised hyperspectral band selection. *IEEE Geosci. Remote Sens. Lett.*, 8: 138-142.
DOI: 10.1109/LGRS.2010.2053516

Abbreviations

The following abbreviations are used in this manuscript:

- WLPBS: Weighted Linear Prediction based band selection
LPBS: Linear Prediction based band selection
WLP: Weighted Linear Prediction
LP: Linear Prediction
LPE: Linear Prediction Error
VD: Vertical Dimensionality
GT: Ground Truth
SVM: Support a Vector Machine
JM: Jeffreys-Matusita
BS: Band Selection
RBF: Radial Basis Function
NB: Number of Bands
MI: Mutual Information
MIFS: Mutual Information Feature Selection