A greedy modular eigenspace-based band selection approach for hyperspectral imagery

Yang-Lang Chang\textsuperscript{a} and Hsuan Ren\textsuperscript{b}

\textsuperscript{a}Department of Information Management, National Taipei College of Business, Taipei, Taiwan
\textsuperscript{b}Center for Space and Remote Sensing Research, National Central University, Chung-Li, Taiwan

ABSTRACT

The greedy modular eigenspaces (GME) has shown effective in hyperspectral feature extraction. The GME was developed by grouping highly correlated hyperspectral bands into a smaller subset of band modular regardless of the original order in terms of wavelengths. It utilizes the inherent separability of different classes in hyperspectral images to reduce dimensionality and further to generate a unique GME feature. This paper takes advantage of the GME to develop a GME-based band selection (GMEBS) for hyperspectral imagery. It selects a subset of non-correlated hyperspectral bands for hyperspectral images using the unique ability of the GME in class separability. The proposed GMEBS algorithm provides a fast procedure to select the most significant features and speeds up the distance decomposition compared to GME features. It also avoids the bias problems of transforming the information into linear combinations of bands as does the traditional principal components analysis (PCA). The proposed GMEBS approach selects each band by a simple logical operation, call GME feature scale uniformity transformation (GME/FSUT), to include different classes into the most common feature modular subset of bands. Interestingly, experimental results show that this simple GMEBS approach is very effective and can be used as an alternative to other band selection algorithms.

Keywords: Hyperspectral, Band selection, Greedy modular eigenspace-based band selection (GMEBS), GME feature scale uniformity transformation (GME/FSUT), principal component analysis (PCA)

1. INTRODUCTION

High-dimensional spectral images obtained from multispectral, hyperspectral or even ultraspectral bands generally provide enormous spectral information for data analysis. It covers an abundance of applications from satellite imaging, monitoring systems to medical imaging and industrial product inspection. In this paper, a novel technique is proposed for hyperspectral band selection of earth remote sensing. It utilizes the inherent separability of different classes in hyperspectral images to reduce dimensionality and further to generate a unique greedy modular eigenspace (GME)\textsuperscript{1} feature. The method makes use of the GME to develop a GME-based band selection (GMEBS) for hyperspectral imagery. It’s developed for land cover classification based on band selection of the same scene collected from hyperspectral remote sensing images. It presents a framework for band selection of hyperspectral remote sensing images, which consists of two algorithms, referred to as the GME and feature scale uniformity transformation (GME/FSUT).\textsuperscript{2,3} The GME method is designed to extract features by a simple and efficient GME feature module, while the FSUT is performed to fuse most correlated features from different data sources.

One of most common issues in hyperspectral classification is to improve class separability without incurring the curse of dimensionality.\textsuperscript{4} The term has been given a great deal of attention by researchers in the statistic, database, and data mining communities to describe the difficulties associated with the feasibility of distribution estimation in high-dimensional datasets. Numerous techniques have been developed for feature extraction to reduce dimensionality without loss of class separability. The most widely used approach is the principal components analysis (PCA) which reorganizes the data coordinates in accordance with data variances so that features
are extracted based on the magnitudes of their corresponding eigenvalues. Another well-known approach is orthogonal subspace projection. It projects all undesired pixels into a space orthogonal to the space generated by the desired pixels to achieve high-dimensionality classifications. Most of them focus on the estimation of statistics at full dimensionality to extract classification features. For example, conventional PCA assumes the covariances of different classes are the same and the potential differences between class covariances are not explored. In contrast, our proposed GMEBS method overcomes the dependency on global statistics, while preserving the inherent separability of the different classes. Most classifiers seek only one set of features that discriminates all classes simultaneously. This not only requires a large number of features, but also increases the complexity of the potential decision boundary. This paper will show that the proposed GMEBS method solves this problem and speed up the feature extraction processes significantly.

In the past, various techniques for hyperspectral image classifications have been proposed. The most widely used approach is the PCA which reorganizes the data coordinates in accordance with data variances so that features are extracted based on the magnitudes of their corresponding eigenvalues. The GME is a spectral-based technique that explores the correlation among bands. Reordering the bands regardless of the original order in terms of wavelengths in high-dimensional datasets is an important characteristic of GME. It performs a greedy iteration search algorithm which reorders the correlation coefficients in the data correlation matrix row by row and column by column to group highly correlated bands as GME feature eigenspaces that can be further used for feature extractions and band selections. Each ground cover type or material class has a distinct set of GME-generated feature eigenspaces. The FSUT makes use of the GME feature extraction method which tends to equalize all the bands in a subgroup with highly correlated variances to avoid a potential bias problem that may occur in conventional PCA. This fast band selection algorithm, called GMEBS, takes advantage of the special characteristics of GME to construct different classes into the most common feature subspaces.

We propose a fast band selection algorithm to unify GME sets of different classes. It takes advantage of the special characteristics of GME to concentrate GME sets of different classes into the most common feature subspaces. A distance measure based on GMEBS was then applied to decompose the similarity for land cover classification purposes. The goal is to develop a band selection technique that combines different datasets in such a way that new types of data products can be produced. In our previous work, the GME was developed by grouping highly correlated bands into a small set of bands. After finding a GME set, a FSUT is next performed to unify the feature scales of these GME. The GME/FSUT carries out a band selection for the GME of different classes. To demonstrate the advantages of the proposed method, we compared several different configurations, which were categorized by their use of different options and distance measures. This paper presents a novel band selection algorithm. The approach makes use of the statistical properties of the abundant feature characteristics of MASTER hyperspatial datasets while taking advantage of the GME feature extraction method. The performance of the propose method is evaluated by MODIS/ASTER airborne simulator (MASTER) images for land cover classification during the PacRim II campaign. Experimental results demonstrate that the proposed GMEBS approach is an effective method for feature extractions. The rest of this paper is organized as follows. In Section 2, the proposed GMEBS, GME/FSUT, classifier is described in detail. In Section 3, a set of experiments
is conducted to demonstrate the feasibility and utility of the proposed approach. Finally, in Section 4, several conclusions are presented.

2. METHODOLOGY

Referring to Fig. 1, there are four stages to implementing our proposed GMEBS classification scheme. 1) A GME/FSUT algorithm is applied to construct an identical GME set for the purposes of the feature extraction and the dimension reduction. 2) The second stage is a band selection to perform the feature selection based on the GMEBS algorithm described on the first stage. 3) The third stage is a threshold decomposition which normalizes the scales of different feature bands selected by the GMEBS algorithm. Then, a distance decomposition, also known as a similarity measure, is performed for the Euclidean minimum distance classification. 4) Finally, a classification process is performed for the classification.

2.1. Greedy Modular Eigenspaces and Feature Scale Uniformity Transformation

In our previous work, we proposed a GME set $\Phi_k$ which is composed of a group of modular eigenspaces, i.e. $\Phi^k = (\Phi_1^k, \ldots, \Phi_l^k, \ldots, \Phi_n^k)$, for the class $\omega_k$. Each modular eigenspace $\Phi^k_l$ includes a set of highly correlated bands regardless of the original order of wavelengths. It utilizes the inherent separability of the different classes in high-dimensional data to reduce dimensionality and formulate a unique GME feature. In Fig. 2, a visual correlation matrix pseudo-color map (CMPM) which was proposed by Lee and Landgrebe to emphasize the importance of second-order statistics in hyperspectral data is used to illustrate the magnitude of correlation matrices in our proposed GME method. We define a correlation submatrix $c_{\Phi^k_l}(m_l \times m_l)$ which belongs to the $l^{th}$ modular eigenspace $\Phi^k_l$ of a GME set $\Phi^k = (\Phi_1^k, \ldots, \Phi_l^k, \ldots, \Phi_n^k)$ for a land cover class $\omega_k$ in this dataset, where $m_l$ and $n_k$ represent, respectively, the number of bands (feature spaces) in modular eigenspaces $\Phi^k_l$, and the total number of modular eigenspaces of a GME set $\Phi^k$, i.e. $l \in \{1, \ldots, n_k\}$ as shown in Fig. 2.

The original correlation matrix $c_{X_k}(m_t \times m_t)$, where $m_t$ is the total number of original bands (i.e. $m_t = \sum_{l=1}^{n_k} m_l$), is decomposed into $n_k$ correlation submatrices $c_{\Phi^k_1}(m_1 \times m_1), \ldots, c_{\Phi^k_l}(m_l \times m_l), \ldots, c_{\Phi^k_{n_k}}(m_{n_k} \times m_{n_k})$ to build a GME set $\Phi^k$ for the class $\omega_k$. There are $m_t!$ (the factorial of $m_t$) possible combinations to construct a candidate GME set. Only one of them can be chosen as the GME set. It is computationally expensive to make an exhaustive search to construct a GME set if $m_t$ is a large number. In our previous work, we proposed a fast greedy band reordering algorithm, called greedy modular eigenspace transformation (GMET), based on the assumption that highly correlated bands often appear adjacent to each other in high-dimensional data. Reordering the bands regardless of the original order in terms of wavelengths in high-dimensional datasets is an important characteristic of GME. It performs a greedy iteration search algorithm which reorders the correlation coefficients in the data correlation matrix row by row and column by column to group highly correlated bands as GME feature eigenspaces that can be further used for feature extraction. Each land cover type or material class has a distinct set of GME-generated feature eigenspaces. A GME set, $\Phi^k = (\Phi_1^k, \ldots, \Phi_l^k, \ldots, \Phi_n^k)$, is composed for

![Image](image-url)
land cover class $\omega_k$. Fig. 2 illustrates the original correlation matrix map and the reordered one after a GMET. Each land cover type or material class has its uniquely ordered GME set. For instance, in our experiment, six land cover types were transformed into six different GME sets. Three of these six classes are shown in Fig. 3.

In this visualization scheme, we can build a GME efficiently and bypass the redundant procedures of rearranging the band order from the original high-dimensional datasets. Moreover, the GMET algorithm can tremendously reduce the eigen-decomposition computation compared to conventional PCA feature extraction. The computational complexity for conventional PCA is of the order of $O(m_t \times m_t)$ and it is $O(\sum_{k=1}^{n_k} m_t^2)$ for GME. The GME preserves the original information of a correlation matrix. After defining GME sets $\Phi^k$, a fast and effective GME/FSUT is performed to unify the feature scales of these GME sets to an identical GME set $\Phi_I$. We use intersection (AND) operations$^2$ applied to the band numbers inside each GME module $\omega_k$ to unify the feature scales of different classes produced by GMET and construct an identical intersection GME (IGME) $\Phi_I$ set for all classes $\omega_k$, $k \in \{1, \ldots, N\}$. A concept block diagram of GMEBS is shown in Fig. 4.
different class has the same IGME set $\Phi_I$ after the GMEBS.

The GME/FSUT performing a searching iteration to build an identical IGME set $\Phi_I$ is initially carried out on a newly formed IGME feature module $\Phi_{I_l}$, where $l \in \{1, \ldots, n\}$ and $\Phi_{I_l} \in \Phi_I$, in which the first band $b_i$, where $i \in \{1, \ldots, n\}$ and $i = 1$, of the largest GME module $\Phi^1_k$ is chosen to form an IGME set $\Phi_I$. Each band $b_i$ is assigned an attribute during a GME/FSUT. If the attribute of $b_i$ is set as available, it means this $b_i$ has not been yet assigned to any identical IGME set $\Phi_I$. If a $b_i$ is assigned to a $\Phi_I$, the attribute of this $b_i$ is set to used. All attributes of the original $b_i$, $i \in \{1, \ldots, n\}$, are first set as available. The proposed GMEBS algorithm is as follows:

**Step 1.** Initialization: a new IGME feature module $\Phi_{I_l}$, where $\Phi_{I_l} \in \Phi_I$, is initialized by a new band $b_i$ inside a GME module $\Phi^k_1$, where $b_i$ is defined as the first available band and $\Phi^k_1$ as the largest GME module $\Phi^k_1$ of the class $\omega_k$. This new band $b_i$ is assigned to the newly created IGME feature module $\Phi_{I_l}$ and is then set as the current $b_i$, i.e. the only one activated at the current time. Then, go to step 2. Note that this GME/FSUT algorithm is terminated if the last band $b_i$ is already set to used and the final IGME feature module $\Phi_{I_l}$, $\Phi_{I_l} \in \Phi_I$, has been obtained.

**Step 2.** If the current band $b_i$ and all its related bands have all been set to used, then a IGME feature module $\Phi_{I_l}$ is constructed with all used bands, these used bands are removed from the band list, and the algorithm goes to step 1 for another round to find a new IGME feature module $\Phi_{I_l}$. Otherwise, it goes to step 3.

**Step 3.** As illustrated in Fig. 5, the GME/FSUT links all the same bands $b_i$ located in different classes $\omega_k$, $k \in \{1, \ldots, N\}$, together with arrow lines in an intersection operation behavior. Then this band $b_i$ is set as used. Meanwhile, acting in an (AND) manner, GME/FSUT also greedily extends the band range to all of the other related bands (inside the black bold GME modular boxes) inside the current activated GME module $\Phi^F_k$ associated with the used bands for all classes $\omega_k$, $k \in \{1, \ldots, N\}$. Go to step 2.

Eventually, an identical IGME set $\Phi_I$ is composed. For convenience, we sort these IGME feature modules $\Phi_{I_l}$, where $l \in \{1, \ldots, n_I\}$, according to the number of their feature bands, i.e. the number of feature spaces in descending order. An example of the proposed GMEBS method is illustrated in Fig. 5. Each IGME feature
2.2. Band Selection

In this paper, we present a novel band selection GMEBS method which utilizes the inherent separability of different classes in hyperspectral images to reduce dimensionality and further to generate a unique IGME feature. It takes advantage of the GME/FSUT-generated bands to develop a GMEBS method for hyperspectral imagery. It selects a subset of non-correlated hyperspectral bands for hyperspectral images using the unique ability of the IGME in class separability. Here, we use the intersection property of the IGME to select one of the most similar (correlated) bands in each IGME module $\Phi_l$, $l \in \{1, \ldots, n_I\}$, for all class $\omega_k$, $k \in \{1, \ldots, N\}$, arbitrarily to compose an identical IGME set $\Phi_I$. Let us assume an identical IGME set $\Phi_I$ has $n_I$ feature bands $b_I = (b_1, \ldots, b_{n_I})$, where $b_1 \in \Phi_{I_1}, \ldots, b_l \in \Phi_{I_l}, \ldots, b_{n_I} \in \Phi_{I_{n_I}}$, as shown in Fig. 4. This GMEBS algorithm provides a quick bands selection procedure of the most significant features and an instant distance measure of the hyperspectral test samples compared to the conventional feature extraction methods.

2.3. Threshold and Distance Decompositions

The IGME can distinguish different classes well by the highly correlated feature characters of GME. It can make use of the Euclidean distance (ED) to extract the most significant bands selected from the IGME as a similarity measure from the proposed GME/FSUT-generated bands. Before finding the ED, a threshold decomposition is
needed to normalize the scales of different feature bands selected by GMEBS. The threshold decompositions are performed to create normalized band scale values $b_{I_{Nr}} = (b_{I_{1 Nr}}, \ldots, b_{I_{Nk}}, \ldots, b_{I_{Nk}})$. The $b_{I_{Nr}}$ is normalized to the range (0, 1) by the nonlinear sigmoid function:

$$\xi = \frac{b_{I_{Nr}} - \mu}{\sigma},$$

and

$$b_{I_{Nr}}(\xi) = \frac{1}{1 + \exp(-\xi t)},$$

where $\mu$, $\sigma$, and $t$ denote the mean and the standard deviation of the normalized band scale values $b_{I_{Nr}}$ and a threshold value respectively. This new normalized band scale values $b_{I_{Nr}}$ are converted into binary values. In Fig. 6, the threshold decomposition function $T(\cdot)$ transforms the $b_{I_{Nr}}$ into the binary values for all classes $\omega_k$.

After the band scale value normalization, a normalized ED $e_{I_{Nr}} = (e^{1}_{I_{Nr}}, \ldots, e^{k}_{I_{Nr}}, \ldots, e^{N}_{I_{Nr}})$ for all classes $\omega_k$, $k \in \{1, \ldots, N\}$, is then decomposed for the purpose of distance measure. Note that only one band for each IGME feature module $\Phi_l$, $l \in \{1, \ldots, n_l\}$, is selected to decompose the ED normalized distance, as shown in Fig. 5. In Fig. 6, a normalized ED $e_{I_{Nr}}$ function $E(\cdot)$ of an identical IGME set with $n_l$ feature bands is defined as:

$$e_{I_{Nr}}(x) = \sqrt{\sum_{i=1}^{n_l} \tilde{x}_i^2},$$

where $\tilde{x} = x - \bar{x}$ is the mean-normalized vector of sample $x$. In Fig. 6, the sample $x$ is equal to $b_{I_{Nr}} = (b_{I_{1 Nr}}, \ldots, b_{I_{Nk}}, \ldots, b_{I_{Nk}})$ for the training samples. Here, $e_{I_{Nr}}(x)$ represents the distance between the query test samples $X$ and the mean vector of training samples based on the IGME feature module $\Phi_l$. This distance decomposition is applied to all classes $\omega_k$ to generate an identical normalized ED $e_{I_{Nr}}$.

### 2.4. Minimum Distance Classification

After finding the normalized ED $e_{I_{Nr}}$ from the previous stage, a minimum distance classification is next performed. By comparing the normalized ED $e_{I_{Nr}}$ between training samples and test samples, it is easy to identify
what the best class the test samples are belonged to. In Fig. 6, the training samples $x$ are first applied to the GME/FSUT to construct the IGME feature module $\Phi_{I_l}$. The final determined classes are then induced by applying the test samples $X$ to the GME/FSUT stage for band selection and then to the ED normalized distance decomposition for the minimum distance classification.

3. EXPERIMENTAL RESULTS

A plantation area in Au-Ku on the east coast of Taiwan as shown in Fig. 7 was chosen for investigation. The image data was obtained by the MASTER as part of the PacRim II project. A ground survey was made of the selected six land cover types at the same time. The proposed GMEBS method was applied to 35 bands selected from the 50 contiguous bands (excluding the low signal-to-noise ratio mid-infrared channels) of MASTER. Six land cover classes, sugar cane A, sugar cane B, seawater, pond, bare soil and rice ($N = 6$) are used in the experiment. The criterion for calculating the classification accuracy of experiments was based on exhaustive test cases. One hundred and fifty labeled samples were randomly collected from ground survey datasets by iterating every fifth sample interval for each class. Thirty labeled samples were chosen as training samples, while the rest were used as test samples, i.e., the samples were partitioned into 30 (20%) training and 120 (80%) test samples ($M = 120$) for each test case, as shown in Fig. 6. Three correlation coefficient threshold values, $t_c = 0.75$, 0.80 and 0.85, were selected to carry out GMET. Finally, the accuracy was obtained by averaging all of the multiple combinations stated above.

We compared several different configurations. Three main groups are compared in Fig. 8. The first group is for GMEBS (the bolder lines in gray in Fig. 8). In this case, the GMEBS was applied to MASTER datasets to generate the IGME. One band was arbitrarily selected for each IGME feature module $\Phi_{I_l}^k$, $l \in \{1, \ldots, n_I\}$, to decompose the normalized ED and apply to minimum distance classifier. For the second group, the same datasets were used as for the first group, thus they was used to make a comparison between the conventional Euclidean distance and our proposed GMEBS methods. The Euclidean upper bound, Euclidean average bands and Euclidean lower bound represent respectively the best upper bound feature bands, the average bands (between upper bound and lower bound) and the worst lower bound feature bands obtained from original datasets. Euclidean PCA stands for the primary principal components of PCA from original datasets. Compared to the first group, only the Euclidean upper bound is better than proposed GMEBS method. Note that the Euclidean upper bound is chosen based on a prior knowledge of the MASTER instrument. It is hard to obtain the test dataset information without a priori knowledge of the instrument. It is obvious that our proposed GMEBS method is superior to the conventional Euclidean average band cases.

An interesting case is the three test group, in which there was a difference in classification accuracy between ED distance measures and the residual reconstruction error (RRE) distance decompositions, also known as a GME projection method. In spite of the fact that the feature module of GME has a better accuracy than
GMEBS feature modules when they were applied to different distance decompositions, the proposed GMEBS algorithm provides an efficient way to select the most significant feature band and to speed up the distance decomposition compared to GME features. Table 1 summarizes the evaluation of classification accuracy under different conditions to illustrate the validity of these unique properties of proposed GMEBS band selection method. These encouraging results showed that satisfactory classification accuracy could be achieved with only a few computational time and small training samples.

4. CONCLUSIONS

This paper presents a novel band selection GMEBS technique for hyperspectral image classification, which consists of two algorithms, referred to as the GME and the FSUT. The GME method makes use of the GMET developed by grouping highly correlated bands into a small set of bands regardless of the original order of wavelengths to extract the most significant feature bands from high-dimensional datasets, while the FSUT is performed to uniformity most correlated feature scales from different data sources. The proposed GMEBS algorithm is efficient with little computational complexity. It uses intersection (AND) operations applied to the band numbers inside each GME module to unify the feature scales of GME and construct an identical IGME feature module set. It can be implemented as a band selector to generate a particular feature band set for each material class.

The experimental results demonstrated that the feature bands selected from the MASTER datasets by the GMEBS algorithm contain discriminatory properties crucial to subsequent classification. Moreover, compared to conventional feature extraction techniques, the IGME feature modules have very good abilities to adapt to the minimum distance classifiers. They make use of the potential significant separability of GME to select a unique set of most important feature bands in high-dimensional datasets. The proposed GME/FSUT algorithm provides a fast way to find the most significant feature bands and to speed up the distance decomposition compared to GME features.
Table 1. Summary evaluation of classification accuracy for different band selection methods, distance measures, number of classes.

<table>
<thead>
<tr>
<th>Band selection methods (Distance measures)</th>
<th>Number of classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMEBS (ED)</td>
<td>3  4  5  6</td>
</tr>
<tr>
<td>Group one:</td>
<td></td>
</tr>
<tr>
<td>GMEBS</td>
<td>85.97% 83.15% 85.60% 86.52%</td>
</tr>
<tr>
<td>Group two:</td>
<td></td>
</tr>
<tr>
<td>Euclidean upper bound</td>
<td>89.06% 93.46% 94.19% 91.44%</td>
</tr>
<tr>
<td>Euclidean average bound</td>
<td>69.55% 63.34% 60.42% 56.79%</td>
</tr>
<tr>
<td>Euclidean lower bound</td>
<td>52.85% 46.19% 42.18% 40.63%</td>
</tr>
<tr>
<td>Euclidean PCA</td>
<td>59.00% 54.05% 51.89% 48.54%</td>
</tr>
<tr>
<td>Group three: (RRE)</td>
<td></td>
</tr>
<tr>
<td>GME</td>
<td>95.93% 97.03% 95.10% 92.99%</td>
</tr>
</tbody>
</table>

REFERENCES