Novel hyperbolic clustering-based band hierarchy (HCBH) for effective unsupervised band selection of hyperspectral images.
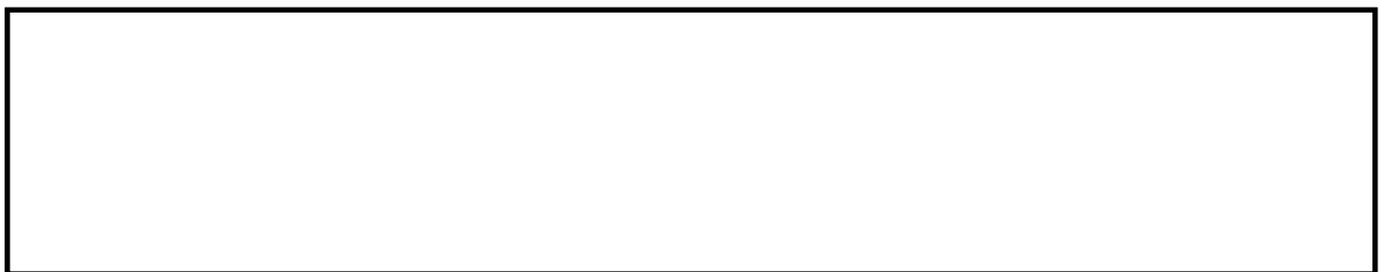
SUN, H., ZHANG, L., REN, J. and HUANG, H.

2022

Novel Hyperbolic Clustering-based Band Hierarchy (HCBH) for Effective Unsupervised Band Selection of Hyperspectral Images

He Sun[a], Lei Zhang[a,*], Jinchang Ren[b], Hua Huang[c]

[a]*School of Computer Science and Technology, Beijing Institute of Technology, China*
[b]*National Subsea Centre, Robert Gordon University, U.K.*
[c]*School of Artificial Intelligence, Beijing Normal University, China*

Abstract

For dimensionality reduction of HSI, many clustering-based unsupervised band selection (UBS) methods have been proposed due to their superiority of reducing the high redundancy between selected bands. However, most of these methods fail to reflect the data structure of HSI, leading to inconsistent results of band selection. To tackle this particular issue, we have proposed a novel hyperbolic clustering-based band hierarchy (HCBH) to fully represent the underlying spectral structure and obtain a more consistent band selection. With the proposed adaptive hyperbolic clustering, the performance can be effectively improved with the aid of geometrical information. By introducing a cluster-centre based ranking metric, the desired band subset can be naturally obtained during the clustering process. Experimental results on three popularly used datasets have validated the superior performance of the proposed approach, which outperforms a few state-of-the-art (SOTA) UBS approaches.

Keywords: Hyperspectral image; unsupervised band selection; hyperbolic space clustering; hierarchical clustering.

---

[*]Corresponding author
 Email address: `leizhang@bit.edu.cn` (Lei Zhang)

1. Introduction

Hyperspectral images (HSI) can provide rich spectral information with hundreds of contiguous bands, which is beneficial to a wide range of remote sensing applications, such as land cover analysis [1, 2, 3, 4]. However, numerous bands increase the dimension of HSI and result in challenging issues for HSI data processing. As there are unlikely sufficient training samples in real applications, the HSI processing may suffer from the Hughes phenomenon, which affect subsequent tasks, e.g. image classification/recognition and target detection. Additionally, the redundant bands within a narrow range may affect the effectiveness of data analysis. Furthermore, the high dimensionality of HSI may result a huge computational burden. Therefore, the dimensionality reduction becomes a major challenge in HSI.

The dimensionality reduction methods of HSI is generally classified into two categories, i.e. the feature extraction based and band selection based. Although feature extraction methods [5, 6] maintain the discriminative information of HSI, they rely heavily on some mathematical operations, such as the principal component analysis (PCA) and its variations folded-PCA [3], 1D/2D singular spectrum analysis (SSA) [7], etc. The band selection in HSI [8] can directly select a desired band subset from the original data whilst discarding the redundant bands. In this manner, the band selection methods can retain the sequential information of the original hypercube for ease of further physical interpretation, which is more preferred in the HSI data processing. The band selection methods are usually comprised of three groups, supervised, semi-supervised, and unsupervised based on if training samples are involved or not [9]. As it is not straightforward to acquire the label information, the unsupervised band selection (UBS) is more preferable [10].

Current UBS methods can be grouped into five main categories, i.e. ranking-based [11], searching-based [12, 13, 14, 15, 16], clustering-based [17, 18, 19, 20, 21], sparsity-based [22, 23], and embedding-based [24, 25]. Ranking-based methods weight each band and select the top-ranked bands, but the correlation

between selected bands are ignored. Searching-based methods normally define an objective function to obtain a band subset via an iterative optimization process, which usually results a huge computational burden. Clustering-based methods first group all bands into different clusters and then choose the most significant band from each group to form the band subset, which can effectively avoid the high redundancy among the chosen bands. Sparsity-based methods first assume that each band can be represented by several other bands, and the calculated sparse coefficients can reflect the importance of each band. However, the sparse coefficients are very sensitive to the employed optimization methods. The embedding-based UBS aim to find the desired band subset by investigating the intrinsic geometric structural information, but they often suffers from high computational burden due to their learning process. Among the above five categories of UBS methods, clustering-based have attracted more attention due to their superiority of lowering the redundancy among chosen bands.

Nevertheless, there are some challenging issues for the clustering-based UBS methods. Firstly, the geometrical structure of the HSI data is not fully investigated, which may affect clustering results. Besides, existing clustering-based UBS methods tend to apply a ranking strategy on the clustering result to obtain the desired band subset. Although this strategy has achieved a good performance, it increases the computational complexity.

As an effective tool for handling the high dimensional data, the hyperbolic distance has achieved great success due to its strong ability of reflecting the geometrical relationship within the hierarchical data [26, 27]. Compared to the Euclidean distance, the hyperbolic one grows in an exponential way, and the distance between nodes in a hierarchy also performs in an exponential manner with the increase of depth in the hierarchy [27]. In this manner, the hyperbolic distance is more suitable for a tree-like data. For HSI data, contained bands are contiguous according to their wavelength spectrum, and the HSI can be treated as a tree and each node represents a single band [21]. As the hyperbolic distance measures the similarity of two data points in consideration of both their distance and data representations, the geometrical structure of HSI can

be better revealed, and a more efficient ranking strategy can be designed after revealing the inherent structure of HSI.

Therefore, we have constructed a band hierarchy based on the hyperbolic clustering. From our knowledge, HCBH is the first UBS method in the hyperbolic space. We formulate the major contributions of this paper as follows:

1) A novel hyperbolic clustering-based band hierarchy is proposed for unsupervised hyperspectral band selection. With a tree-based hierarchy in the hyperbolic space, the proposed HCBH can effectively reveal the geometrical structure of the HSI dataset.

2) Compared to the widely used Euclidean metric, the introduced adaptive hyperbolic distance can effectively evaluate both geometrical and informative similarities between different bands, and the effectiveness of band selection can be improved even with less data samples.

3) Instead of applying a specific ranking metric, a hyperbolic distance-based cluster-centre ranking strategy is employed to weight the significance of each band and locate the centre of each cluster, which can directly output a better band selection result during the clustering process.

The reminder of this paper are organized as follows. Section II summarizes related clustering-based methods. In Section III, the detail of our proposed HCBH framework are given, including hyperbolic band hierarchy, adaptive hyperbolic distance, cluster-centre ranking strategy, and the merits of HCBH. Afterwards, section IV presents the experiments conducted on three HSI datasets. Finally, conclusions of this paper are give in the Section V.

2. Related Work

According to the defined strategy of our HCBH method, some related clustering-based UBS approaches will be introduced as follows.

Clustering-based UBS methods usually build a similarity matrix first based on predefined criteria, such as the Euclidean distance. Afterwards, HSI bands can be divided into different groups according to the similarity matrix, where

each cluster-centre is usually selected to form the desired band subset. As chosen bands are from different groups, there will be low correlation in the band subset.

In WaLuMi and WaLuDi [17], a hierarchical clustering structure has been proposed to perform the UBS. To simultaneously minimize the intra-cluster and maximize the inter-cluster variance, two information metrics, the mutual information and the Kullback-Leibler (K-L) divergence, are employed to reduce data redundancy between selected bands. With these two metrics, the Ward's linkage method is utilized to perform a hierarchical clustering and the total band set is divided into different groups. Afterwards, the band in each cluster with the highest averaged similarity to the rest is chosen. As noisy bands are not similar to their adjacent bands and can easily become single-band clusters, both WaLuMi and WaLuDi are sensitive to noisy bands. In addition, both methods have high computational costs due to their utilized information metrics.

Aiming to take advantages of both ranking-based and clustering-based UBS methods, an enhanced fast-peak-based clustering (E-FDPC) method [11] has been developed. Based on the hypothesis that a cluster-centre should be with a higher local density and a larger distance to other clusters, all bands are ranked based on the above metrics estimated from an Euclidean-based matrix, and the top-ranked bands are chosen eventually. Although the E-FDPC can be considered as an efficient clustering-based method, its chosen band subset still has a higher redundancy.

Recently, an optimal clustering framework (OCF) [18] has been proposed based on the dynamic programming. According to the top-ranked cut and the normalized cut mechanisms, OCF has attempted to optimize the clustering result with constraints for HSI. After segmenting the HSI bands into separated clusters, three ranking metrics, including the E-FDPC, maximum-variance PCA-based criterion, and information entropy (IE), have been applied for evaluating and choosing the most significant band within each cluster. The experiments have validated the superiority of combining a clustering framework with effective ranking strategies.

With a coarse-to-fine process, Wang *et. al.* has proposed a clustering-based

UBS method, an adaptive subspace partition strategy (ASPS) [19], to segment the whole HSI data into several sub-cubes by maximizing the ratio of inter-class to intra-class distance. In ASPS, a band is divided into several blocks, where the local mean and variance are utilized to represent the noisy level of each band. Afterwards, the band with minimum noise in each sub-cube is picked as part of the desired band subset. However, the noisy level estimation relies on the extracted blocks, which may heavily affect the performance of ASPS.

Similar to ASPS, a fast neighbourhood grouping band selection method (FNGBS) [20] has been further introduced in a coarse-to-fine mechanism. Based on the Euclidean distance, a similarity matrix is constructed to measure distances between bands and the coarse cluster centre. With the aid of contextual information, the clustering result is further refined so that continuous bands are grouped into a cluster based on their similarities. Following the clustering result, the band with a larger product of local density and information entropy is selected in each group.

In our former work[21], we have proposed an adaptive distance-based band hierarchy (ADBH) for extracting the hierarchical structure in spectral domain of HSI and generate a more consistent UBS result with varying numbers of selected bands. Firstly, the hierarchy clustering is performed to group similar bands according to a bottom-up manner. To solve the noisy sensitive problem of clustering-based UBS methods, an adaptive Euclidean-based distance is designed with the consideration of cluster density.

Although the above approaches have achieved a good performance for UBS, all these methods implement the clustering process with the aid of Euclidean-based metrics. In HSI, each band represents the reflectance of certain response with a range of wavelength, and similar bands usually share some physical characteristics, such as contiguous wavelengths. Hence, the inherent structure of a HSI may fit a tree hierarchy as each band represents a node and adjacent bands in HSI are closer. Recently, The hyperbolic distance-based methods have been successfully applied into several areas, including the 3D shape index [28] and the person re-identification [29], etc. It is found that a specific non-Euclidean

6

space, the hyperbolic space, is more suitable for a tree-based data structure [26], which motivates this work to combine the hyperbolic distance with band hierarchy for more effective UBS as detailed in the next section.
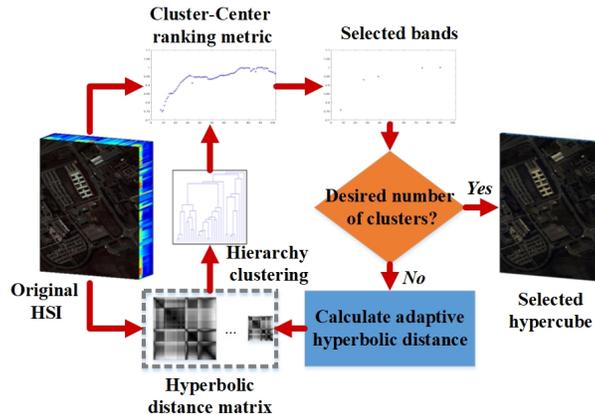
3. The Proposed Method

Figure 1: The overall architecture of our proposed HCBH.

In this section, we detail the proposed HCBH method, including our proposed tree-based hierarchy strategy, the description of the adaptive hyperbolic distance, a further developed cluster-centre strategy for choosing desired band subsets in the clustering process, and merits of our proposed HCBH. As shown in Fig. 1, a hierarchical clustering process is performed in a bottom-up manner after estimating the similarity matrix. With the cluster-centre ranking strategy, consistent results of band selection can be yielded. Relevant details are presented in the following subsections.

3.1. Hyperbolic Band Hierarchy

For clustering-based UBS methods, the whole band set is usually grouped into different clusters first, then the representative band in each cluster is selected. However, inconsistent results of band selection can be produced with different numbers of clusters. In our previous work ADBH [21], we have built

a tree-based band hierarchy for UBS in HSI. With a bottom-up clustering process, the ADBH can guarantee a consistent band selection result with various numbers of clusters, and the robust performance of ADBH on several datasets have shown the superiority of tree-based hierarchy for UBS.

Although ADBH has achieved certain success, it cannot fully reflect the hierarchical structure of HSI data within the used Euclidean space. In general, the Euclidean space is assumed to be flat and symmetric [26], which can not better fit the hierarchy structure [26]. On the contrary, the hyperbolic space, as a non-Euclidean space, has received more attention recently due to its strong ability to fit a tree-based hierarchy and provide a more flexible data representation [27]. For a HSI, its contained bands can fit in a tree-like structure, i.e. a hierarchy, based on their wavelength. With a Euclidean-based metric, the distance between two bands can be obtained, but it is hard to find out their locations within the tree-like structure. The location of each band on the tree can be easily obtained if their distances to the root of the tree is known. In a hyperbolic space, such as a Poincaré ball, the 'origin' of the ball can be assumed as the root and the distance between each point on the ball and the 'origin' can be estimated. This property is very useful to understand the hierarchical structure of HSI. By fitting a HSI in a hyperbolic space like the Poincaré ball, the distance between each band to the 'origin' of the ball can be estimated. In this way, each band can be naturally located in the tree-like hierarchy. Hence, we can build a band hierarchy in the hyperbolic space, the HCBH, for improving the band selection performance and revealing the hierarchical structure of HSI.

Let $X \in R^{W \times H \times L}$ be a 3D hypercube, where $W \times H$ is the spatial dimension and L is the spectral dimension. By projecting the hypercube into a 2-D matrix, a HSI can be reshaped as $X \in R^{N \times L} = (x_1, ..., x_l, ..., x_L)$ and normalized to $(0,1)$, where the number of pixels in the band equals to $N = W \times H$ and x_l represents the lth band. With the assumption that a HSI fit a tree hierarchy $T = (V, E, Z)$, $V = (1, 2, ..., L)$ denotes a set of nodes, where each node represents a spectral band, and the E denotes the linkage set between different nodes. Similar to [18, 20, 21], we assume that each band is more likely to be in the
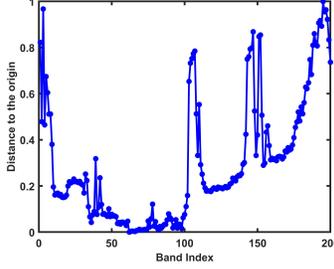
8

Figure 2: The introduced hierarchical distance Z of the Indian Pines dataset.

same cluster with its adjacent bands because of the contiguous characteristic of HSI in the spectral domain. For example, the lth node can be only clustered with the $(l-1)th$ and $(l+1)th$ node initially. Hence, the linkage set E is defined as $E = (e_1, ... e_l, ..., e_{L-1})$ that the 'edge' linked the lth node and the $(l+1)th$ node is denoted by e_l.

As a key factor of our proposed HCBH, a novel hierarchical distance $Z = (z_1, ..., z_l, ..., z_L)$ is introduced to illustrate the hierarchical information of each node. The basic idea is to embed our proposed T into the Poincaré ball model of the hyperbolic space [30], i.e. a Riemannian manifold, and the root of the tree is put at the 'origin' O of the Poincaré ball, i.e. the top of the tree-based hierarchy. All nodes will be placed around the 'origin', and the hierarchical information of the lth node will be represented by its distance z_l to the 'origin'. In the proposed HCBH, the 'origin' is set to be at the centre of the Poincaré ball and its value is 0. Although the 'origin' does not represent any physical meaning, it can help to locate the position of each band in the tree hierarchy. In the proposed band hierarchy, the band closer to the 'origin' will be in the higher level. Therefore, we assume that the band closer to the 'origin' is more important in its cluster. As shown in Fig. 2, the distance to the 'origin' of each band can be easily obtained, and spectral bands are placed around the 'origin' based on their hierarchical distance Z. With this proposed Poincaré ball model, the spectral band with a closer distance to the origin is more important than
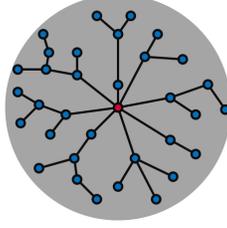
9

Figure 3: The obtained band hierarchy by the proposed method. The grey circle represents the Poincaré ball model, and the red circle denotes the 'origin' of the ball. Each spectral band is denoted by a blue circle and the relationship between different bands and the 'origin' are determined by the proposed hierarchical distance Z.

others. The definition of z_l is given by:

$$
\begin{aligned}
z_l &= acosh(1 + 2\frac{||x_l - O||_2^2}{(1 - ||x_l||_2^2)(1 - O) + \epsilon_1}) \\
&= acosh(1 + 2\frac{||x_l||_2^2}{(1 - ||x_l||_2^2) + \epsilon_1}) \\
&= acosh(\frac{1 + ||x_l||_2^2}{(1 - ||x_l||_2^2) + \epsilon_1}) \\
&= 2atanh(||x_l||_2)
\end{aligned}
\tag{1}
$$

where ϵ_1 is a small positive constant.

After that, the hierarchical clustering is implemented in a bottom-up manner as follows. Firstly, the initialization process is performed that each node, i.e. each spectral band, in the tree is treated as a single-band cluster. After that, a similarity matrix is built between different nodes. Based on the 'mutual nearest neighbouring' searching strategy [21], two neighbouring nodes are merged into one cluster if and only if both of them have the lightest edge compared to others. Following this criterion, the clustering process is implemented iteratively.

During the merging process, it is necessary to determine the relationship between two nodes, which can be used to indicate the latent structure of the HSI. In our previous work [21], this relationship is ignored as the position of each node or cluster cannot be located in the Euclidean space. Therefore, the ADBH treats each node equally. With the newly introduced Z, the position of each node in a tree hierarchy can be located by weighting the distance between each

node and the 'origin'. To this end, a 'parent-child' relationship between these two nodes can be established where the node closer to the 'origin' is assumed to be the 'parent'. For each new cluster, we utilize the mean spectral information to represent the whole cluster. This iterative process will continue until only one cluster is left, and the clustering results with the number of clusters from L to 1 can be obtained progressively. As a result, a band hierarchy as shown in Fig. 3 can be built by the proposed bottom-up clustering.

3.2. Adaptive Hyperbolic Distance

For hierarchical clustering, there are two crucial issues, i.e. the similarity metric and the effect of noisy bands. In hierarchical clustering, the similarity metric has to be calculated repeatedly when clusters are updated iteratively, a robust measurement can guarantee the clustering performance and avoid the huge computational burden simultaneously. Due to the nature of the bottom-up clustering, a noisy band is unlikely to be grouped for its large difference against its neighbours. If a noisy band becomes a single-band cluster after numerous iterations, it will be selected in the final band subset. Hence, an efficient and robust hyperbolic distance measurement is given as follows towards the above two issues.

In the Euclidean space, the most common way to determine the difference between two adjacent bands x_l and x_{l+1} is by:

$$d_{l,l+1} = ||x_l - x_{l+1}||^2 \tag{2}$$

In this paper, we introduce the fundamental hyperbolic distance h into the UBS work, where the distance between two adjacent bands is defined as:

$$h_{l,l+1} = acosh(1 + 2\frac{||x_l - x_{l+1}||_2^2}{(1 - ||x_l||_2^2)(1 - ||x_{l+1}||_2^2) + \epsilon_2}) \tag{3}$$

where ϵ_2 is a small positive constant.

Compared to the commonly used Euclidean distance, the hyperbolic distance has several advantages. As seen in (3), the hyperbolic distance includes the angle information with the *acosh* function. To this end, the hyperbolic distance can

11

derive more geometrical information between different bands, which is more favourable for a tree-based band hierarchy. Besides, the hyperbolic distance focuses more on HSI data representation itself by adding two data norms in the denominator part, where the Euclidean distance only measures the difference between two bands. As a result, the hyperbolic distance has a better capacity than the Euclidean one to estimate the similarities between two bands in an informative view. Moreover, the efficiency of hyperbolic distance is acceptable.

To avoid the single-band cluster caused by noisy bands, we have designed an adaptive mechanism on the hyperbolic distance by combining the cluster density and band representation. As a noisy band is easy to become a single-band cluster, the density m_l is introduced to illustrate that how many bands the lth cluster contains, initially $m_l = 1$. For an effective characteristic measurement, the $L2$ norm of each band is used to indicate the representation of spectral data. By combining the cluster density and norm of each band, the data representation μ_l of the lth band (cluster) is obtained as:

$$\mu_l = m_l * norm(x_l) \tag{4}$$

and the similarity between the lth band and neighbouring $(l + 1)$th can be estimated by an adaptive hyperbolic distance as:

$$\eta_{l,l+1} = h_{l,l+1} * \mu_l * \mu_{l+1} \tag{5}$$

In our proposed HCBH, the edge between the l and $(l+1)$ nodes is denoted as $\eta_{l,l+1}$. According to this adaptive mechanism, the cluster with less bands will be merged effortlessly than others, which can suppress noisy bands.

3.3. Cluster-Centre Ranking Strategy

In conventional clustering-based UBS approaches such as OCF [18] and ADBH [21], the most representative band is chosen by using some ranking metrics such as E-FDPC. Although this kind of strategy is found to be effective on some datasets, it has two major drawbacks. The first is the computational cost,

which may decrease the efficiency. Additionally, the ranking metric may have some crucial parameters and increase the complexity.

To tackle these drawbacks, we have proposed a simple yet effective ranking metric to select bands during the clustering process. As we have introduced a novel hierarchical distance $Z = [z_1, ..., z_l, ...z_L]$, and this corresponds to the importance of each band in our proposed band hierarchy. With the metric Z, the 'parent' of each cluster can be taken as the cluster-centre. To this end, the hierarchical distance Z is an effective metric to indicate the most vital band in each cluster. Therefore, we have applied Z as a cluster-centre ranking metric, and the specified ranking process is given as follows.

Let K be the desired number of selected bands, the clustering result can be represented as $R = r_1, ...r_k, ..., r_K$, where r_k indexes of kth cluster. For each cluster r_k, the band b_k with minimum distance z_k to the 'origin' is chosen, and the band subset $B = [b_1, ...b_k, ..., b_K]$ can be formed. With the above strategy, spectral bands closest to the 'origin' of the tree-based hierarchy in each cluster construct the band subset B during the clustering process, which improves the efficiency of the HCBH.

Algorithm 1 Hyperbolic Clustering-based Band Hierarchy (HCBH)

1: **Input:** The original HSI X, the desired number of bands K.

2: Compute the η between neighbouring bands using (3), (4) and (5);

3: If mutual neighbouring exist, then merge these cluster pairs sequentially. Otherwise, merge two clusters with the lightest edge.

4: Update clustering result after one merging operation.

5: Repeat the above operations until the number of clusters equals to K

6: Select the band with minimum distance to the origin in each cluster to form the subset B.

7: **Output:** The selected band subset B.

3.4. Merits of HCBH

With the aid of the hyperbolic distance, our HCBH aims to obtain the band selection result with any desired number of selected bands whilst reflecting the latent structure of the HSI in a tree-based band hierarchy. By fitting the HSI in a hyperbolic space with a bottom-up clustering process, the relationship between different bands can be better represented by their distances to the 'origin'. Moreover, an adaptive hyperbolic distance is introduced to characterize different bands by effectively considering both the geometrical and informative representation. With the adaptive mechanism, the single-band cluster caused by noisy bands can be avoided. Rather than selecting a most important band from each group after the clustering process, we have designed a ranking metric to locate the centre of each cluster, and the band selection result can be obtained directly. According to the pseudocode shown in Algorithm 1, the HCBH requires no input parameters and outputs a series of band selection results with any specific number of bands, which can thus reduce the computational complexity.

4. Experiments

The experimental section validates the effectiveness and robustness of our proposed HCBH method on several remote sensing datasets. The performance is evaluated by the quantitative results, such as classification results with the chosen band subset.

First, the datasets are briefly introduced. Then, the experimental settings are discussed, including compared methods, benchmarking classifiers, and classification metrics. The comparison between our proposed method and eight state-of-the-art UBS approaches are presented in the third subsection. In the fourth subsection, the difference between HCBH and ADBH are analysed, including some discussion about the hyperbolic and Euclidean space UBS. The effectiveness of the proposed cluster-centre ranking strategy is further discussed in the fifth subsection. To further verify the efficiency of our proposed HCBH method the computational time comparison are also given in the final part.

14

4.1. Datasets

In the experiments, we have utilized three public available datasets captured by three different sensors, which is summarized in Table 1. The pseudo color images of all datasets are shown in Fig. 4.
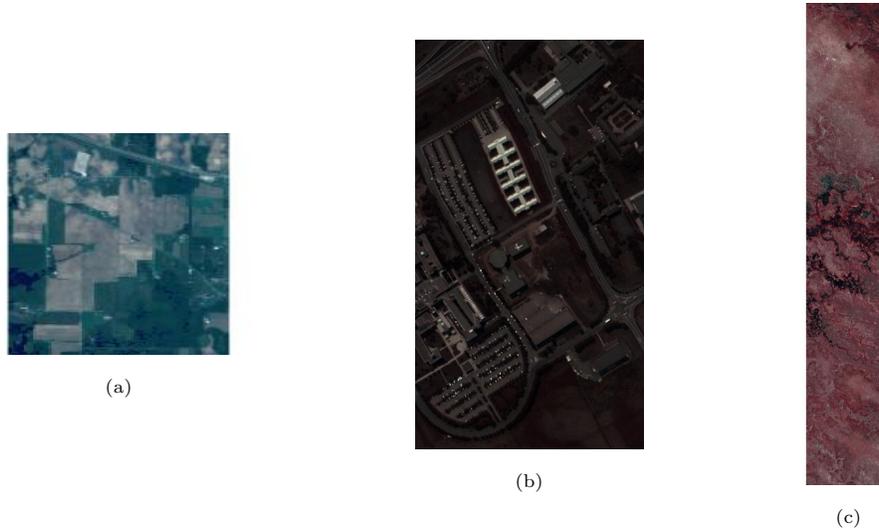

(a)

(b)

(c)

Figure 4: The pseudo color images of all datasets. (a) Indian Pines, (b) PaviaU, (c) Botswana.

Table 1: Summary of three utilized HSI datasets

	Indian Pines	PaviaU	Botswana
Number of bands	200	103	145
Spectral range(nm)	400-2500	430-860	400-2500
Spatial size	145×145	610×340	1476×256
Spatial resolution(m/pixel)	20	1.3	30
Number of classes	16	9	16
Sensors	AVIRIS	ROSIS	Hyperion

4.1.1. Indian Pines dataset

Collected by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) sensor, this dataset has 224 spectral bands with a spectral range from 400 to

15

$2500nm$ and a spatial size of 145×145 pixels. After removal of water absorption bands and severely noisy bands, only 200 bands in 16 semantic classes are utilized for testing.

Table 2: The number of training and testing samples of the Indian Pines dataset in the classification experiment.

Class	No.Train	No.Test	No.Samples	Class	No.Train	No.Test	No.Samples
Alfalfa	5	41	46	Corn-notill	143	1285	1428
Corn-mintill	83	747	830	Corn	24	213	237
Grass-pasture	49	434	483	Grass-trees	73	657	730
Grass-pas-turemowed	3	25	28	Hay-windrowed	48	430	478
Oats	2	18	20	Soybean-notill	98	874	972
Soybean- mintill	246	2209	2455	Soybean-clean	60	533	593
Wheat	21	184	205	Woods	127	1138	1265
Building-grass-trees	39	347	386	Stone-steel-towers	10	83	93
Total	1031	9218	10249				

4.1.2. Pavia University dataset (PaviaU)

With 103 spectral bands in a wavelength range of 430 to $860nm$, PaviaU was captured by the Reflective Optics System Imaging Spectrometer (ROSIS) sensor. After discarding some less useful pixels, a cropped dataset with a spatial size of 610×340 is used, in which 9 land-cover classes are labelled.

Table 3: The number of training and testing samples of the PaviaU dataset in the classification experiment.

Class	No.Train	No.Test	No.Samples	Class	No.Train	No.Test	No.Samples
Asphalt	664	5967	6631	Meadows	1865	16784	18649
Gravel	210	1889	2099	Trees	307	2757	3064
Painted metal sheets	135	1210	1345	Bare soil	503	4526	5029
Bitumen	133	1197	1330	Self-blocking bricks	369	3313	3682
Shadows	95	852	947	Total	4281	38495	42776

4.1.3. Botswana dataset

Captured by NASA EO-1 satellite Hyperion sensor and with 242 bands ranging from 400 to 2500nm in the spectral domain, this dataset has spatial size of 1476×256 pixels with 16 labelled semantic classes. Similar to the Indian Pines dataset, a corrected version with 145 bands is utilized.

Table 4: The number of training and testing samples of the Botswana dataset in the classification experiment.

Class	No.Train	No.Test	No.Samples	Class	No.Train	No.Test	No.Samples
Water	27	243	270	Hippo grass	10	91	101
SReeds	27	252	269	Firescar	26	233	259
Acacia woodlands	30	284	314	Acacia grasslands	30	275	305
Mixed mopane	27	241	268	FloodPlain grasses 1	25	226	251
FloodPlain grasses 2	22	193	215	Riparian	27	252	269
Island interior	20	183	203	Acacia shrublands	25	223	248
Short mopane	18	163	181	Expose soils	10	85	95
Total	4281	38495	42776				

4.2. Experimental Settings

To verify the efficacy of the proposed HCBH, three popular HSI classification methods, including the K-nearest neighbour (KNN) [31], the support vector machine (SVM) [32], and the edge preserving filter (EPF) [4] are utilized to evaluate the classification performance of chosen band subsets on the above three HSI datasets. Based on a 10-fold cross-validation, parameters of all the three classifiers are optimized. For each dataset, we have randomly chosen 10% of data samples as the training set for all classifiers and the rest are used for testing similar to other UBS literatures [18, 20, 19, 16], and the split of training and testing samples are shown in Table 2-4. Three commonly used metrics, the overall accuracy (OA), the average accuracy (AA), and the Kappa coefficient, are derived from the confusion matrix for quantitative evaluation of the experimental results.

For comparison, five SOTA clustering-based UBS approaches are compared

with the HCBH, including WaLuDi [17], OCF [18], ASPS [19], FNGBS [20], and ADBH [21]. Besides, two SOTA searching-based UBS methods and one SOTA ranking-based method, the dominant-set extraction UBS (DSEBS), the optimal neighborhood reconstruction (ONR), and E-FDPC [11], are compared to verify the effectiveness of the proposed HCBH:

1) WaLuDi [17]: A frequently cited clustering-based UBS method [18, 21, 20], which performs hierarchical clustering based on the construction of K-L divergence matrix. After that, the band in each cluster with the highest averaged similarity to others is chosen.

2) OCF [18]: With a dynamic programming-based strategy, the HSI data is divided into different clusters, followed by a ranking strategy to select the most representative band from the clustering results.

3) ASPS [19]: A novel clustering-based UBS approach, which groups the ordered HSI bands into different sub-cubes and focus more on the noisy effect to select a desired band subset.

4) FNGBS [20]: By fully mining the contextual information, a novel clustering-based method, namely FNGBS, divides the whole band set into different groups, and the product of the local density and the information entropy is utilized as the metric for ranking.

5) ADBH [21]: A tree-based band hierarchy is built to obtain consistent band selection results, and the noisy effect is suppressed by proposing an adaptive Euclidean-based distance. The ADBH is a SOTA method with a robust performance.

6) DSEBS [15]: As a searching-based method, the DSEBS has achieved a robust performance on several remote sensing datasets. It can output the band selection results by solving a greedy-searching problem, where both the band informativeness and independence have been considered.

7) ONR [16]: The ONR is a novel searching-based method with leading performance. It considers the UBS as a combinatorial optimization problem, and the chosen band subset are determined by evaluating their reconstruction ability of the original data.

8) E-FDPC [11]: the E-FDPC method takes both advantages of ranking-based and clustering-based UBS. By ranking each band based on jointly weighing two Euclidean-based metrics, it has a leading performance among ranking-based methods.

For the above methods, all the experiments are conducted on the original code provided by authors with default parameters. For our HCBH, it is a parameter-free method and only the desired number of bands k is required. For compared methods and HCBH, the classification experiments are repeated 10 times to reduce the randomness of chosen training samples for all classifiers, and the average results are reported, including the OA curves and quantitative results. We have compared the classification performance by using the original data for better indicating the effectiveness of the UBS methods, denoted as 'raw data' in this paper. All the methods are implemented on the MATLAB 2020b with their original codes using a 16-GB RAM Intel i7-10700K CPU.

4.3. Comparison Results

To demonstrate the effectiveness of our proposed HCBH, the experimental results are shown in two ways. First, the OA curves of three datasets are generated from 3 to 30 selected bands and shown in Figs. 5-7. Second, the quantitative results averaged on 3-30 selected bands are compared in Tables 5-7.
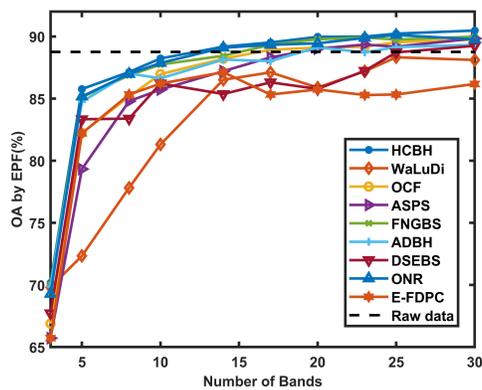
4.3.1. Indian Pines

As illustrated in the Fig. 5, our proposed method has a robust performance on the both KNN and EPF, and the best performance on the SVM. Although our OA curve on KNN is slightly lower than the ADBH when more than 20 bands are selected, the OA curve of HCBH is superior to all others. On the SVM classifier, our proposed HCBH has the best OA curve than all other methods. For the compared methods, most OA curves on the KNN become more stable with the increasing number of selected bands. In Table 5, the proposed HCBH has the best average quantitative results on all the three classifiers, which has fully
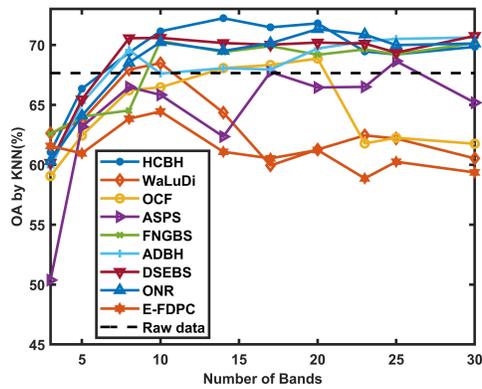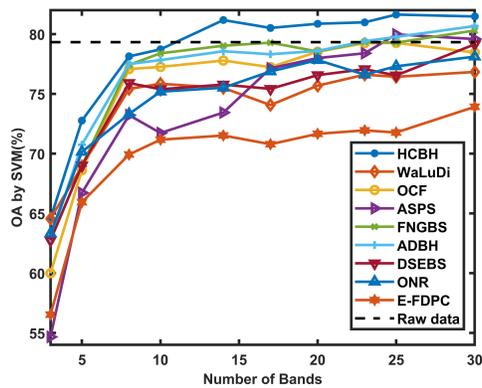
Figure 5: Overall accuracy curves for the Indian Pines dataset. (a) Results by KNN, (b) Results by SVM, (c) Results by EPF.
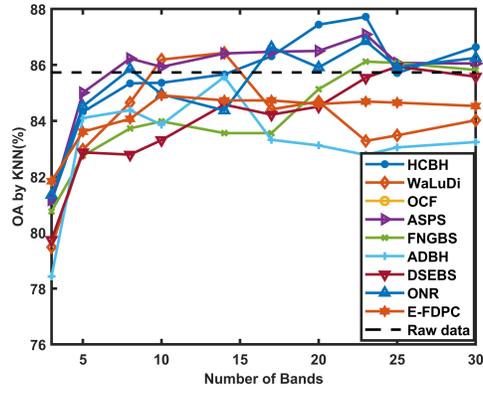
Table 5: Quantitative results on the Indian Pines dataset using different UBS methods. (Best results are labelled as bold except those from raw data).

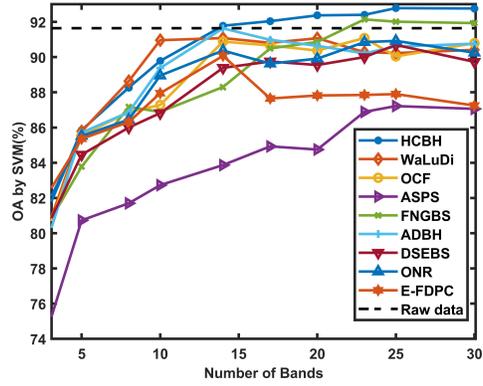| Classifier	WaLuDi[17]	OCF[18]	ASPS[19]	FNGBS[20]	ADBH[21]	DSEBS[15]	ONR[16]	E-FDPC[11]	HCBH	Raw data
OA by KNN(%)	63.35±2.89	64.52±3.45	64.27±5.24	67.9±2.95	67.94±2.74	68.74±3.39	68.51±3.52	61.21±1.7	**69.19±3.33**	67.65±0.72
AA by KNN(%)	51.48±3.01	55.03±3.96	52.77±5.66	56.87±4.02	57.55±2.3	54.99±2.68	55.34±4.48	47.01±2.87	**57.76±2.63**	54.22±1.31
Kappa by KNN	0.58±0.04	0.59±0.04	0.59±0.06	0.63±0.03	0.63±0.04	0.64±0.04	0.62±0.06	0.55±0.03	**0.67±0.04**	0.62±0.01
OA by SVM(%)	73.99±4.03	75.39±6.21	73.29±7.75	76.34±5.76	76.43±5.48	74.34±5.6	74.42±4.58	69.52±5	**78.09±5.45**	79.33±0.51
AA by SVM(%)	72.33±5.72	73.36±9.02	70.66±12.03	75.40±8.11	74.13±9.02	71.89±8.83	72.13±6.54	65.76±10.57	**76.67±6.02**	71.47±0.61
Kappa by SVM	0.70±0.05	0.72±0.07	0.70±0.1	0.72±0.08	0.73±0.08	0.70±0.07	0.71±0.05	0.65±0.06	**0.75±0.06**	0.75±0.01
OA by EPF(%)	82.45±6.86	85.60±6.99	84.85±7.42	86.70±6.07	86.07±5.07	84.33±6.15	86.69±6.31	83.45±6.36	**87.07±6.13**	88.76±0.92
AA by EPF(%)	80.18±5.91	83.17±6.26	83.02±7.12	84.89±6.11	84.32±6.07	83.32±5.83	**85.32±7.82**	81.79±5.33	85.29±5.69	86.92±5.03
Kappa by EPF(%)	0.81±0.08	0.84±0.06	0.84±0.08	0.86±0.07	0.86±0.06	0.85±0.09	**0.87±0.06**	0.83±0.05	**0.87±0.05**	0.88±0.04

demonstrated the effectiveness of HCBH. Although both ADBH and FNGBS have achieved a good performance, their OAs are about 2% lower than our HCBH method on both KNN and SVM. Both the DSEBS and ONR methods have achieved a good performance on the KNN but their OAs on the SVM are not robust enough. For the EPF classifier, all the compared methods have achieved a better performance and our proposed method has obtained the best OA. To sum up, HCBH has a significant superiority on the Indian Pines dataset.
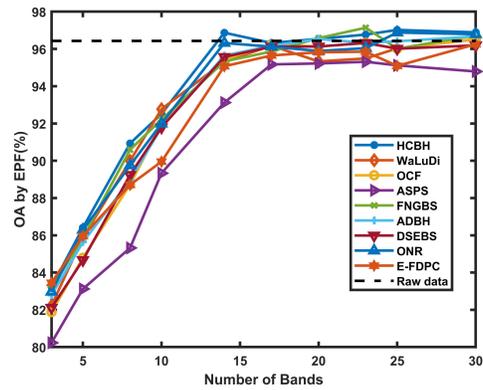
4.3.2. PaviaU

For the PaviaU dataset, Fig. 6 and Table 6 present the OA curves from all the three classifiers. As seen in Fig. 6 (a), the ASPS has the best performance

21

Figure 6: Overall accuracy curves for the PaviaU dataset. (a) Results by KNN, (b) Results by SVM, (c) Results by EPF.

Table 6: Quantitative results on the PaviaU dataset using different UBS methods (Best results are labelled as bold except those from raw data).

Classifier	WaLuDi[17]	OCF[18]	ASPS[19]	FNGBS[20]	ADBH[21]	DSEBS[15]	ONR[16]	E-FDPC[11]	HCBH	Raw data
OA by KNN(%)	83.96±1.94	83.19±1.72	**85.69±1.68**	84.15±1.69	83.18±1.86	81.92±2.3	85.26±1.61	84.24±0.92	85.56±0.02	85.73±1.15
AA by KNN(%)	79.86±2.27	79.12±2.03	82.45±1.92	80.62±1.83	78.76±2.86	76.38±3.2	82.18±4.01	80.68±1.56	**82.63±2.03**	82.02±0.98
Kappa by KNN	0.78±0.03	0.77±0.02	0.81±0.02	0.78±0.02	0.77±0.03	0.75±0.03	0.82±0.01	0.79±0.01	**0.83±0.02**	0.81±0.01
OA by SVM(%)	89±3.33	88.4±3.42	83.49±3.73	88.43±3.86	88.69±3.57	87.52±4.07	88.48±2.93	87.06±2.02	**89.97±3.72**	91.64±1.58
AA by SVM(%)	86±5.75	86.11±4.95	77.29±3.65	84.29±6.02	85.61±6.34	84.55±5.68	85.52±3.33	83.97±3.68	**87.14±4.99**	88.12±1.22
Kappa by SVM	0.85±0.05	0.85±0.05	0.78±0.03	0.84±0.05	**0.86±0.06**	0.83±0.05	0.84±0.03	0.83±0.03	**0.86±0.05**	0.89±0.00
OA by EPF(%)	92.57±4.97	92.41±5.43	90.67±5.79	92.95±4.96	92.69±5.17	92.42±5.34	92.90±5.00	92.18±4.76	**93.35±5.01**	96.43±1.52
AA by EPF(%)	89.13±3.55	88.91±3.19	87.11±4.22	89.32±2.95	89.11±3.44	88.71±3.18	88.92±3.13	88.56±4.23	**89.69±3.71**	93.79±1.98
Kappa by EPF(%)	0.88±0.02	0.88±0.03	0.87±0.01	**0.89±0.03**	0.88±0.02	0.88±0.03	0.88±0.02	0.87±0.03	0.88±0.03	0.94±0.01

when less than 20 bands are chosen, and our HCBH method has the leading performance when more than 20 bands are selected. Our method has achieved the most robust performance than the rest methods with respect to the SVM classifier, and the ASPS has the worst performance. For other compared methods, most approaches do not obtain a good result on both KNN and SVM when more than 20 bands are chosen. As shown in Fig. 6 (c), most methods have achieved a better performance, and our method has a more stable OA curve. In Table 6, our method has the best average quantitative results on both the SVM and EPF, and the second best on the KNN with a small gap behind the ASPS. ASPS has the best performance on the KNN but the worst on the SVM and

EPF, which indicates its poor robustness towards different classifiers. For the ONR, it has a good performance on the KNN and EPF but its classification results on the SVM are relatively poor. Therefore, our HCBH has the most robust performance for the PaviaU dataset, which indicates again its effectiveness.
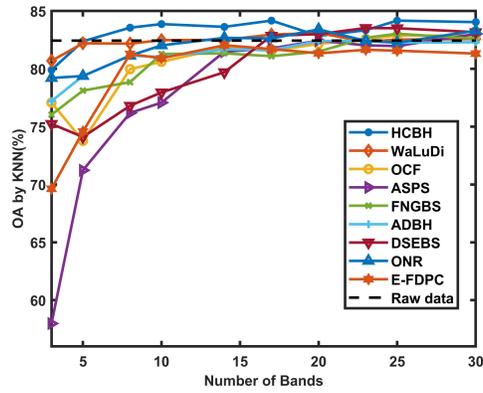
4.3.3. Botswana

As shown in Fig. 7, our HCBH method has the most stable OA curve against different numbers of bands on all the three classifiers. For all the compared methods, the WaLuDi has a relatively good performance with fewer chosen bands, and differences between other methods are not obvious except the ASPS. For the EPF classifier, most methods have achieved a better performance except the ASPS and the E-FDPC. As shown in Table 7, our HCBH has a consistent result of OA curves in Fig. 7, which has demonstrated its superiority with the best performance on all the three classifiers. Specifically, our proposed HCBH method has obtained a better average OA result on the KNN classifier than the raw data, which further validates the merit of HCBH.

According to comparison results, our proposed HCBH has achieved the best result on three public datasets. As shown in Figs. 5-7, HCBH has generated more stable OA curves than other compared methods, which illustrates the consistency of our band selection results. Furthermore, the great performance on all the three classifiers have validated the robustness of our proposed HCBH to different classifiers.

4.4. HCBH vs. ADBH

With similar tree-based band hierarchies, HCBH can be assumed as an enhanced version of ADBH. However, there are several differences between these two methods as compared below.
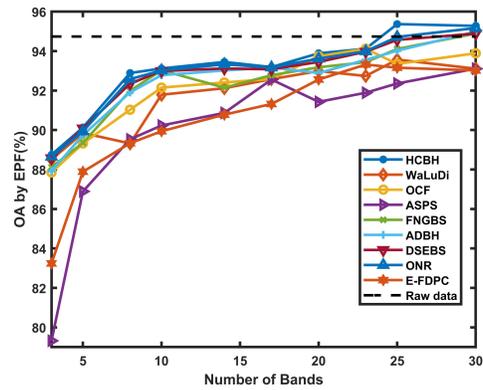
In principle, HCBH and ADBH fit the HSI data into a tree-based hierarchy in two spaces, i.e. the Hyperbolic and Euclidean space. For ADBH, two bands are grouped together based on their Euclidean distance, and it is not straightforward to understand the relationship between these two bands without other
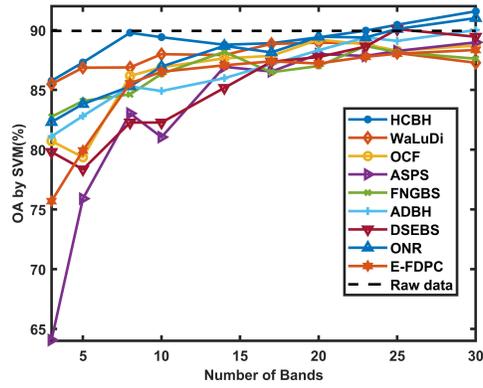
Figure 7: Overall accuracy curves for the Botswana dataset. (a) Results by KNN, (b) Results by SVM, (c) Results by EPF.

Table 7: Quantitative results on the Botswana dataset using different UBS methods (Best results are labelled as bold except those from raw data).

Classifier	WaLuDi[17]	OCF[18]	ASPS[19]	FNGBS[20]	ADBH[21]	DSEBS[15]	ONR[16]	E-FDPC[11]	HCBH	Raw data
OA by KNN(%)	82.36±0.66	80.53±2.95	77.52±7.83	80.65±2.27	81.15±1.66	79.99±3.71	81.85±1.49	79.6±4.13	**83.19±1.29**	82.44±0.63
AA by KNN(%)	79.72±0.75	77.81±3.06	74.95±7.77	78.24±2.16	78.52±1.69	77.39±3.76	78.83±2.45	76.98±4.13	**80.87±1.41**	80.11±0.72
Kappa by KNN	0.81±0.01	0.79±0.03	0.76±0.09	0.79±0.02	0.8±0.02	0.78±0.04	0.8±0.03	0.78±0.05	**0.82±0.02**	0.81±0.02
OA by SVM(%)	87.72±1.11	86.37±3.47	83.06±7.83	86.39±1.98	86.41±2.94	85.11±4.19	87.51±2.87	85.35±4.17	**89.13±1.63**	89.94±0.31
AA by SVM(%)	88.6±1.05	87.21±3.59	84.04±7.88	87.47±1.78	87.45±2.89	86.14±4.23	88.34±5.42	86.15±4.33	**90.05±1.48**	91.54±0.62
Kappa by SVM	0.87±0.01	0.85±0.04	0.82±0.08	0.85±0.02	0.85±0.03	0.84±0.05	0.86±0.02	0.84±0.04	**0.88±0.02**	0.89±0.01
OA by EPF(%)	91.66±1.78	92.05±2.08	89.82±4.11	92.30±2.08	92.40±2.09	92.71±1.97	92.82±2.04	90.46±3.12	**93.02±2.09**	94.74±0.61
AA by EPF(%)	93.23±1.96	93.78±1.79	91.36±3.93	93.95±2.35	93.97±2.02	94.31±2.09	94.15±1.89	92.01±3.64	**94.89±2.15**	95.92±0.47
Kappa by EPF(%)	0.92±0.02	0.92±0.01	0.88±0.01	0.93±0.02	0.92±0.01	0.92±0.01	**0.93±0.01**	0.90±0.02	0.92±0.02	0.93±0.01

metrics. Thus, the ADBH cannot weight bands within a cluster easily. Instead of regarding each node equally, the HCBH attempts to understand the relationship between bands in a cluster. The HSI data is firstly placed into a hyperbolic space, which is popular for its strong capacity of fitting tree-based data. Then, the hyperbolic distance is utilized in the clustering process to introduce more geometrical information. With the aid of the Poincaré ball, a hierarchical distance Z can be defined in (1), which relies on the hyperbolic distance between the 'origin' of Poincaré ball and each band. In this manner, the ranking between different bands can be obtained where bands closer to the 'origin' are assumed to be more critical, and the underlying structure can be better reflected. Besides,

the robust performance of ADBH depends on an effective ranking strategy on clustering results such as E-FDPC, which has lowered the efficiency of ADBH. On the contrary, the HCBH can output the band selection results directly by adopting the hierarchical distance Z to determine the cluster-centre as a selected band. From the classification results on three public datasets, it can be concluded that HCBH has shown a more effective performance than ADBH.

In the graph embedding field, an essential theory is that non-Euclidean space can embed the original data in a lower dimension to perform certain operations than the Euclidean space [33]. In this paper, we have designed a comparison experiment to verify the above theory in the UBS task. As introduced in Section III, the HSI data is firstly projected into a 2-D matrix with a size of $N \times L$. During our clustering process, each band is assumed to be a vector with a dimension of N, i.e. the number of samples in HSI.

In our experiments, we have randomly chosen part of samples to perform UBS task by HCBH and ADBH. By using less samples, the dimension of each band can be reduced, and the performance of HCBH and ADBH can be evaluated with a lower-dimensional data. After obtaining band selection results, we have utilized the SVM with the same setting in Section IV. B. As the samples are randomly chosen, we have repeated the band selection experiment five times and the averaged results are reported as follows.

In Fig. 8 (a), the HCBH with all data samples has the best OA curves than all others. With less than 15 selected bands, the ADBH with all data samples has the second-best OA curve. It is worth mentioning that even with 1000 and 5000 randomly chosen samples, i.e, 1000 and 5000 dimensions, HCBH has comparable results with ADBH that uses the whole dataset, which validates that band hierarchy in the hyperbolic space can generate a more flexible result than in the Euclidean space. As for ADBH with less samples, OA curves are not as robust as HCBH, especially when less than 20 bands are chosen.
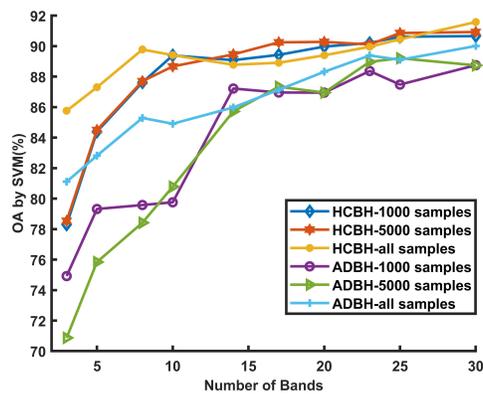
In Fig. 8 (b), the HCBH with all data samples has produced the most robust OA curve than all others on the PaviaU dataset. For ADBH, its performance is much poorer. With 5000 randomly chosen samples, the HCBH has achieved

27
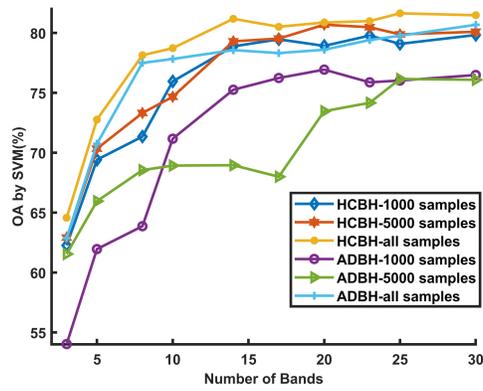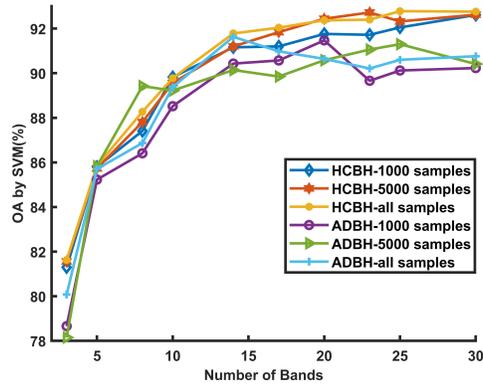
Figure 8: A comparison between HCBH and ADBH with different numbers of samples. (a) Indian Pines, (b) PaviaU, (c) Botswana.

a better OA curve than ADBH. In the circumstance of 1000 randomly picked pixels, the ADBH has obtained the worst result among all OA curves, where the HCBH has acquired a much better result than all ADBH-related curves. The robust performance of HCBH has demonstrated its effectiveness.

For the Botswana dataset, classification results with various numbers of chosen samples from both HCBH and ADBH are shown in Fig. 8 (c). Obviously, the HCBH with only 1000 samples has obtained a more robust OA curve than the ADBH methods with higher dimensions. With less samples, the ADBH cannot produce satisfactory results when less than 15 bands are selected. The superior performance with less samples have verified again the capacity of HCBH in handling the UBS task.

As verified in the results on all three datasets, HCBH can obtain more flexible results than ADBH even with less data samples. This phenomenon has validated the theory in [33], which shows a great potential of Hyperbolic space-based methods in UBS, or even other tasks in HSI.

4.5. The effectiveness of the proposed ranking strategy

Table 8: OA (%) by SVM with different ranking strategies (Best results are labelled as bold)

	HCBH+E-FDPC	HCBH+IE	The proposed one
Indian Pines	76.89	76.63	**78.09**
PaviaU	89.13	88.74	**89.97**
Botswana	87.89	88.62	**89.13**

To better investigate the effect of our proposed ranking strategy, we have compared two other popular strategies, i.e. the E-FDPC [11] and the information entropy (IE) [18]. By replacing the proposed strategy with E-FDPC and IE, we have compared the performance of all these three ranking metrics, and the experiment settings are the same as before and the OAs are generated by averaging the OAs from 3 to 30 bands with the SVM classifier. As shown in Table 8, the E-FDPC has achieved a better performance than the IE on the Indian Pines and the PaviaU datasets, and the IE has achieved a better performance

on the botswana dataset. For all the three datasets, our proposed method has achieved the best OA with the HCBH framework, which fully demonstrates the effectiveness of our proposed ranking strategy.

4.6. Computational Time Comparison

Table 9: The computational complexity between different UBS methods, including their running time (s) when 30 bands are selected and input parameters (Best results are labelled as bold).

	No.Param.	Indian Pines	PaviaU	Botswana
WaLuDi	4	30.27	78.72	112.45
OCF	0	0.17	0.23	0.57
ASPS	0	0.17	0.36	1.13
FNGBS	0	**0.05**	0.08	0.22
ADBH	0	0.24	0.92	0.82
DSEBS	4	0.21	0.89	2.55
ONR	1	0.08	0.33	0.57
E-FDPC	1	0.95	1.9	1.5
HCBH	0	0.06	**0.04**	**0.03**

As a significant issue in the UBS task, the computational complexity of our proposed HCBH is analysed. For a fair comparison, we have compared the computational time of each method with the same hardware and software platform. In Table 9, the number of key parameters and the processing time of each method with 30 chosen bands are depicted.

As seen in Table 9, most clustering-UBS methods, especially the recently proposed OCF, ASPS, FNGBS, and ADBH methods [18, 19, 20, 21], have no input parameters in their framework. Similar to these methods, our proposed HCBH method is also a parameter-free method. For the computational time, our HCBH is the most efficient one for both PaviaU and Botswana dataset and the second efficient one, slight below FNGBS method for the Indian Pines. For WaLuDi method, its computational burden is rather heavy on all three datasets. For DSEBS, ONR, and E-FDPC, all of them have one or more key parameters,

which may lower the robustness of these methods. Accordingly, HCBH is more efficient and flexible for robust and efficient clustering-based UBS of HSI.

5. Conclusion

In this paper, we have proposed a hyperbolic clustering-based band hierarchy for effective band selection in HSI. It is the first time the UBS task is performed in the hyperbolic space, which can construct a more flexible tree-based hierarchy to reflect the data structure within the HSI. By introducing the proposed adaptive hyperbolic distance, HCBH can effectively improve band selection results and reveal the geometrical information between different bands. The effectiveness of the proposed method has been further improved by generating band selection with a proposed cluster-centre ranking strategy during the clustering process. The experimental results on three public datasets have fully validated the effectiveness and robustness of our proposed HCBH framework.

Although the effectiveness of our HCBH method has been demonstrated by classification performance of selected bands in the experimental part, the HCBH has not been performed with other unsupervised tasks, such as anomaly detection [34, 35, 36] and object tracking [37]. For these applications, an effective UBS method can select more discriminative bands whilst reducing the computation cost. Hence, the utilization of HCBH in some other applications of HSI will be further explored in the future work.

References

[1] L. Yang, F. Zhang, P. S.-P. Wang, X. Li, Z. Meng, Multi-scale spatial-spectral fusion based on multi-input fusion calculation and coordinate attention for hyperspectral image classification, Pattern Recognit. 122 (2022) 108348.

[2] T. Qiao, Z. Yang, J. Ren, P. Yuen, H. Zhao, G. Sun, S. Marshall, J. A. Benediktsson, Joint bilateral filtering and spectral similarity-based sparse

representation: A generic framework for effective feature extraction and data classification in hyperspectral imaging, Pattern Recognit. 77 (2018) 316–328.

[3] J. Zabalza, J. Ren, M. Yang, Y. Zhang, J. Wang, S. Marshall, J. Han, Novel folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and sar in remote sensing, ISPRS J. Photogram. and Remote Sens. 93 (2014) 112–122.

[4] X. Kang, S. Li, J. A. Benediktsson, Spectral–spatial hyperspectral image classification with edge-preserving filtering, IEEE Trans. Geosci. and Remote Sens. 52 (5) (2014) 2666–2677.

[5] H. Wu, S. Prasad, Semi-supervised dimensionality reduction of hyperspectral imagery using pseudo-labels, Pattern Recognit. 74 (2018) 212–224.

[6] H. Huang, Z. Li, H. He, Y. Duan, S. Yang, Self-adaptive manifold discriminant analysis for feature extraction from hyperspectral imagery, Pattern Recognit. 107 (2020) 107487.

[7] J. Zabalza, J. Ren, J. Zheng, J. Han, H. Zhao, S. Li, S. Marshall, Novel two-dimensional singular spectrum analysis for effective feature extraction and data classification in hyperspectral imaging, IEEE Trans. Geosci. and Remote Sens. 53 (8) (2015) 4418–4433.

[8] W. Sun, Q. Du, Hyperspectral band selection: A review, IEEE Geosci. and Remote Sens. Mag. 7 (2) (2019) 118–139.

[9] J. Feng, J. Chen, Q. Sun, R. Shang, X. Cao, X. Zhang, L. Jiao, Convolutional neural network based on bandwise-independent convolution and hard thresholding for hyperspectral band selection, IEEE Trans. Cyber. 51 (9) (2021) 4414–4428.

[10] R. N. Patro, S. Subudhi, P. K. Biswal, F. Dell'acqua, A review of unsupervised band selection techniques: Land cover classification for hyperspectral

earth observation data, IEEE Geosci. and Remote Sens. Mag. 9 (3) (2021) 72–111.

[11] S. Jia, G. Tang, J. Zhu, Q. Li, A novel ranking-based clustering approach for hyperspectral band selection, IEEE Trans. Geosci. and Remote Sens. 54 (1) (2016) 88–102.

[12] J. Feng, L. Jiao, F. Liu, T. Sun, X. Zhang, Unsupervised feature selection based on maximum information and minimum redundancy for hyperspectral images, Pattern Recognit. 51 (2016) 295–309.

[13] J. Tschannerl, J. Ren, P. Yuen, G. Sun, H. Zhao, Z. Yang, Z. Wang, S. Marshall, MIMR-DGSA: Unsupervised hyperspectral band selection based on information theory and a modified discrete gravitational search algorithm, Inform. Fusion 51 (2019) 189–200.

[14] W. Chen, Z. Yang, J. Ren, J. Cao, N. Cai, H. Zhao, P. Yuen, MIMN-DPP: Maximum-information and minimum-noise determinantal point processes for unsupervised hyperspectral band selection, Pattern Recognit. 102 (2020) 107213.

[15] G. Zhu, Y. Huang, J. Lei, Z. Bi, F. Xu, Unsupervised hyperspectral band selection by dominant set extraction, IEEE Trans. Geosci. and Remote Sens. 54 (1) (2016) 227–239.

[16] Q. Wang, F. Zhang, X. Li, Hyperspectral band selection via optimal neighborhood reconstruction, IEEE Trans. Geosci. and Remote Sens. 58 (12) (2020) 8465–8476.

[17] A. MartÍnez-UsÓMartinez-Uso, F. Pla, J. M. Sotoca, P. GarcÍa-Sevilla, Clustering-based hyperspectral band selection using information measures, IEEE Trans. Geosci. and Remote Sens. 45 (12) (2007) 4158–4171.

[18] Q. Wang, F. Zhang, X. Li, Optimal clustering framework for hyperspectral band selection, IEEE Trans. Geosci. and Remote Sens. 56 (10) (2018) 5910–5922.

600

33

[19] Q. Wang, Q. Li, X. Li, Hyperspectral band selection via adaptive subspace partition strategy, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 12 (12) (2019) 4940–4950.

[20] Q. Wang, Q. Li, X. Li, A fast neighborhood grouping method for hyperspectral band selection, IEEE Trans. Geosci. and Remote Sens. 59 (6) (2021) 5028–5039.

[21] H. Sun, J. Ren, H. Zhao, G. Sun, W. Liao, Z. Fang, J. Zabalza, Adaptive distance-based band hierarchy (ADBH) for effective hyperspectral band selection, IEEE Trans. Cyber. 52 (1) (2022) 215–227.

[22] W. Sun, L. Tian, Y. Xu, D. Zhang, Q. Du, Fast and robust self-representation method for hyperspectral band selection, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10 (11) (2017) 5087–5098.

[23] W. Sun, G. Yang, J. Peng, Q. Du, Hyperspectral band selection using weighted kernel regularization, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 12 (9) (2019) 3665–3676.

[24] Y. Cai, X. Liu, Z. Cai, BS-Nets: An end-to-end framework for band selection of hyperspectral image, IEEE Transactions on Geoscience and Remote Sensing 58 (3) (2020) 1969–1984.

[25] W. Sun, Q. Du, Graph-regularized fast and robust principal component analysis for hyperspectral band selection, IEEE Trans. Geosci. and Remote Sens. 56 (6) (2018) 3185–3195.

[26] M. Nickel, D. Kiela, Poincaré embeddings for learning hierarchical representations, in: NeurIPS, Vol. 30, 2017.

[27] I. Chami, A. Gu, V. Chatziafratis, C. Ré, From trees to continuous embeddings and back: Hyperbolic hierarchical clustering, in: NeurIPs, Vol. 33, 2020, pp. 15065–15076.

[28] J. Shi, Y. Wang, Hyperbolic wasserstein distance for shape indexing, IEEE Trans. Pattern Anal. Mach. Intell. 42 (6) (2020) 1362–1376.

[29] V. Khrulkov, L. Mirvakhabova, E. Ustinova, I. Oseledets, V. Lempitsky, Hyperbolic image embeddings, in: CVPR, 2020, pp. 6418–6428.

[30] N. Monath, M. Zaheer, D. Silva, A. McCallum, A. Ahmed, Gradient-based hierarchical clustering using continuous representations of trees in hyperbolic space, KDD, 2019, p. 714–722.

[31] L. Ma, M. M. Crawford, J. Tian, Local manifold learning-based k-nearest-neighbor for hyperspectral image classification, IEEE Transactions on Geoscience and Remote Sensing 48 (11) (2010) 4099–4109.

[32] M. Pal, G. M. Foody, Feature selection for classification of hyperspectral data by SVM, IEEE Trans. Geosci. and Remote Sens. 48 (5) (2010) 2297–2307.

[33] R. C. Wilson, E. R. Hancock, E. Pekalska, R. P. Duin, Spherical and hyperbolic embeddings of data, IEEE Trans. Pattern Anal. Mach. Intell. 36 (11) (2014) 2255–2269.

[34] J. Liang, J. Zhou, L. Tong, X. Bai, B. Wang, Material based salient object detection from hyperspectral images, Pattern Recognit. 76 (2018) 476–490.

[35] X. Yang, M. Dong, Z. Wang, L. Gao, L. Zhang, J.-H. Xue, Data-augmented matched subspace detector for hyperspectral subpixel target detection, Pattern Recognit. 106 (2020) 107464.

[36] W. Xie, J. Lei, S. Fang, Y. Li, X. Jia, M. Li, Dual feature extraction network for hyperspectral image analysis, Pattern Recognit. 118 (2021) 107992.

[37] F. Xiong, J. Zhou, Y. Qian, Material based object tracking in hyperspectral videos, IEEE Trans. Image Process. 29 (2020) 3719–3733.