Feature extraction and classification of hyperspectral images based on improved two-dimensional compact variational mode decomposition

Renxiong Zhuo[1], Cui Mao[2], Ruiying Kuang[3]
Changsha Institute of Technology, ChangSha, China
[3]Corresponding author
E-mail: [1]*rxzhuo_edu@163.com*, [2]*373126004@qq.com*, [3]*kuangruiying@xtu.edu.cn*

Check for updates

Abstract. In order to solve the problem of noise interference in hyperspectral image (HSI) feature extraction, an improved two-dimensional compact variational mode decomposition (2D-C-VMD) method is proposed to extract the spatial dimension feature information of HSI. By smoothing the low-frequency mode components obtained from the first decomposition of 2D-C-VMD, the method is called L-2D-C-VMD, which can further eliminate the adverse effects of spatial domain noise on the classification of different ground objects to a certain extent. Through simulation test analysis on the public datasets Indian Pine and Pavia University. When the Indian Pine training sample takes 4 %, the OA value of the proposed classification method can be increased by about 2 %. When the Pavia University training sample is taken at 0.5 %, the OA value of the proposed classification method can be increased by nearly 3 %, and the proposed method has better competitiveness with fewer training samples.

Keywords: hyperspectral image, spatial domain noise, 2D-C-VMD, classification.

1. Introduction

Hyperspectral imaging (HSI) technology can capture three-dimensional (3D) scene data of ground objects, including 1D spectral dimension and 2D spatial dimension feature information [1]. The spectral and spatial dimensional feature information in the data can be used to classify ground objects [2]. HSI technology, with its fine spectral resolution advantages, is widely used in information security, Marine surveillance, precise classification of crops, biomedicine and resource exploration and other fields [3]. An obvious advantage of HSI is that it has very high spectral resolution, allowing fine differentiation of different targets based on the spectral response of each narrow band. However, the existence of noise interference in HSIs has a great impact on the classification of different features in HSIs. [4-5].

The effect of feature extraction has a great influence on the classification performance of different ground objects. From the data form of feature extraction processing, the feature extraction of HSI can be carried out from the perspective of spectral dimension, spatial dimension and space spectrum union respectively. The research literature shows that the feature extraction from the perspective of combining spatial dimension and space spectrum is more effective. Literature [6] proposes to apply two-dimensional empirical mode decomposition (2D-EMD) to the extraction of spatial dimensional feature information of HSIs. Using the processing ability of EMD to nonlinear and non-stationary data, the 2D images of each band of HSIs are decomposed into mode components of different scales. From the perspective of decomposed pattern components, the lower order pattern components are more conducive to classification. The low order mode component has better suppression effect on spatial noise. Literature [7] proposes to decompose each band of HSI by two-dimensional empirical wavelet transform (2D-EWT) to obtain several wavelet approximation coefficients and detail coefficients of different scales. The approximation coefficient corresponds to the low frequency effective feature information in the original image, while the detail coefficient corresponds to the high frequency detail feature

information in the original image. In the literature, the coefficients of the first four different scales are reconstructed to obtain a new spatial dimension feature information. The new spatial dimension feature information is introduced into SVM classifier for classification and recognition, and the classification performance of SVM can be greatly improved through the process of feature extraction. In addition to whether the classification effect contains effective feature information, the performance of the classifier is also one of the important factors affecting the classification effect. Therefore, in literature [8], optimization strategies for SVM classifiers are analyzed. However, using different bionic algorithms to optimize the parameters of support vector machine network will greatly increase the operation time of the network. In reference [9], the comparative analysis of feature extraction time of different spatial dimension feature extraction methods is also discussed, and the time cost of different spatial dimension feature extraction methods is also different. The spatial dimension noise of HSI has a great influence on the effect of feature extraction. The 2D-EMD can decompose the 2D image of each band of HSI into different scale mode components from high frequency to low frequency. The spatial dimension feature information of the extracted low frequency component has a good recognition effect and has a good inhibition effect on the spatial dimension noise. However, 2D-EMD lacks the theoretical support of mathematics. The 2D-EWT combines the adaptability of EMD and the theory of wavelet decomposition, and provides the theoretical explainability in mathematics. Compared with 2D EMD, this method has lower computational complexity in HSI feature extraction. However, when fewer labeled samples are trained, the recognition effect is worse than that of 2D-EMD. HSIs are decomposed into sub-band images of different scales by two-dimensional variational mode decomposition (2D-VMD). The spatial dimension information extracted from sub-band images of different scales can improve the classification effect of different ground objects, but the interpretability of sub-band images of different scales is not strong.

Aiming at the noise interference in HSI feature extraction, literature [10] proposed a new two-dimensional compact variational mode decomposition (2D-C-VMD) strategy to extract the spatial domain feature information of HSI and further improve the classification performance of different ground objects in HSI. However, for processing HSIs with only one 2D-C-VMD decomposition, whether the spatial domain noise on each band can be effectively removed is a topic that we pay further attention to.

2. Two-dimensional compact variational mode decomposition

The 2D-VMD method can decompose the input image into several sub-band images. In order to reproduce the original image, 2D-VMD explores the reconstruction characteristics of the original image with different mode components. The 2D-VMD is an adaptive and non-recursive decomposition technique for non-stationary and nonlinear signals, which has a more complete mathematical theoretical basis than 2D-EMD. At the same time, based on 2D-VMD, 2D-C-VMD is proposed [11].

On the basis of 2D-VMD, the spatial and spectral compactness of the mode components obtained by the decomposition is improved by introducing binary support function A_k. The mode decomposition problem can be formalized as [11]:

$$find\ u_k, A_k\ s.t\ f = \sum_k A_k \cdot u_k. \tag{1}$$

The original input signal f can be decomposed into the product form of each mode component and the corresponding binary support function. Introducing sparsity to promote regularity constraint on support function is an important means to achieve reasonable compact local support. By utilizing Total Variation (TV) and L^1 penalty functions, the support area and boundary length are effectively penalized. Combining the bivariate support function A_k and its regularization function into the 2D-C-VMD functional can be expressed as [11]:

$$\min_{u_k:\mathbb{R}^2\to\mathbb{R},A_k:\mathbb{R}^2\to\{0,1\},\vec{\omega}_k\in\mathbb{R}^2}\left\{\sum_k \alpha_k\left\|\nabla\left[u_{AS,k}(\vec{x})e^{-j<\vec{\omega}_k,\vec{x}>}\right]\right\|_2^2 + \beta_k\|A_k\|_1 + \gamma_k\,\mathrm{TV}(A_k)\right\},$$

$$\mathrm{s.\,t.}\,\forall\vec{x}\in\mathbb{R}^2:\sum_k A_k(\vec{x})\cdot u_k(\vec{x}) = f(\vec{x}). \tag{2}$$

In the above formula, the total variation (TV) is defined as [11]:

$$\mathrm{TV}(A_k) \coloneqq \underbrace{sup\langle A_k, div\phi\rangle}_{|\phi\leq 1|},\quad \phi:\mathbb{R}^2\to\mathbb{R}^2. \tag{3}$$

L^1 is defined as [11]:

$$\|A_k\|_1 \coloneqq \int_{\mathbb{R}^2}|A_x(x)|dx, \tag{4}$$

where the penalty function L^1 and the total variational TV penalty are used to ensure that the obtained mode components have limited spatial support and regular contours. In the above formula, $\mathbb{R}^2$ is the 2D real number space, and $\vec{\omega}_k$ and A_k are u_k, respectively, corresponding to the center frequency and the binary support function A_k ($k = 1,2,...,K$), where K represents the number of sub-band mode components obtained by decomposition of the original 2D signal. The $u_{AS,k}(\vec{x})$ is the 2D analytic signal of u_k, α_k is the mode bandwidth coefficient, and β_k and γ_k coefficients are the weight coefficients of the support function and the total variable component, respectively. The div means to find the divergence, and sup means to find the upper bound [11].

The 2D-C-VMD can decompose the original HSIs into smooth and clear boundary sub-band images. According to this feature, the mode component obtained from 2D-C-VMD is used as the feature extraction for HSIs [10].

3. Experimental process and analysis

3.1. Datasets

The data set verified by the experiment is tested and analyzed using two sets of public data sets, which can be obtained [12]. The first Indian Pine Scene data set was collected at the AVIRIS sensor at Indian Pine Proving Ground in northwestern Indiana. The spatial dimension of Indian Pine scene data is 145×145, the spatial resolution is about 20 m, and it is composed of 224 spectral reflection bands. The Indian Anson data space contains two-thirds of agricultural woodland and one-third of forest vegetation.

The available ground features are divided into 16 categories, removing spectral bands containing atmospheric noise and water absorption (bands 104-108, 150-163, 220), reducing the number of original bands to 200. Indian Pine scene data is one of the most commonly used HSIs, and it is also considered a challenging data set. Because the spectral features of some feature classes are very similar, the pixels in the images are heavily mixed. The second dataset of the Pavia University scene was captured by the ROSIS 03 sensor during a flight battle in Pavia, northern Italy. The spatial dimension of the data is 610×340, the spatial resolution is 1.3 m, and the number of spectral bands is 115. After removing 12 noise bands, the remaining 103 bands are processed, and the available ground objects are labeled into 9 categories.

3.2. Metric

There are many indexes to measure the classification Accuracy of different types of ground objects in HSIs, among which the most commonly used indexes include Overall Accuracy (OA),

Average Accuracy (AA) and Kappa coefficient [13-15]. The realization of these three indexes is derived from the confusion matrix. The confusion matrix is used to judge whether the classification output is consistent with the real label, as shown below [13]:

$$C = \begin{bmatrix} c_{11} & \cdots & c_{1i} & c_{1j} & \cdots & c_{1r} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c_{i1} & \cdots & c_{ii} & c_{ij} & \cdots & c_{ir} \\ c_{j1} & \cdots & c_{ji} & c_{jj} & \cdots & c_{jr} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c_{r1} & \cdots & c_{ri} & c_{rj} & \cdots & c_{rr} \end{bmatrix}. \tag{5}$$

Table 1. Ground-truth classes for the Indian Pine and Pavia University scene

Indian Pine			Pavia University		
Class	Class name	NoS	Class	Class name	NoS
1	Alfalfa	46	1	Asphalt	6631
2	Corn-notill	1428	2	Meadows	18649
3	Corn-mintill	830	3	Gravel	2099
4	Corn	237	4	Trees	3064
5	Grass-pasture	483	5	Painted metal sheets	1345
6	Grass-trees	730	6	Bare Soil	5029
7	Grass-pasture-mowed	28	7	Bitumen	1330
8	Hay-windrowed	478	8	Self-Blocking Bricks	3682
9	Oats	20	9	Shadows	947
10	Soybean-notill	972			
11	Soybean-mintill	2455			
12	Soybean-clean	593			
13	Wheat	205			
14	Woods	1265			
15	Buildings-Grass-Trees-Drives	386			
16	Stone-Steel-Towers	93			

In the confusion matrix, diagonal elements represent the number of samples accurately identified for each category, while non-diagonal elements show misjudgments in the classification process, where samples from one category are mistakenly assigned to another. Through the confusion matrix, the number of correct classification and misclassification of each type can be visually observed. However, it is not intuitive to judge the classification effect directly from the confusion matrix, so a variety of indexes to measure the classification accuracy have been derived, such as overall accuracy (OA), average accuracy (AA) and Kappa coefficient, etc. These indexes are widely used to evaluate the classification performance. Where, c_{ii} represents the number of class i samples that are correctly classified, c_{ij} represents the number of class i samples that are classified to class j, and r is the total number of categories of samples. The overall classification accuracy refers to the percentage of the number of correctly classified pixels in the entire image. Through this index, the quality of the classification results can be intuitively seen. The definition is as follows [9]:

$$\text{OA} = \frac{\sum_{i=1}^{r} c_{ii}}{N} \times 100\ \%. \tag{6}$$

Although the overall classification accuracy can effectively reflect the overall accuracy of classification results, it shows certain limitations in the face of multi-class cases where the number of samples varies greatly between classes. Specifically, its assessment results may be dominated by those categories with large sample sizes, which may not accurately reflect the classification performance of all categories. The average accuracy, the arithmetic average of the accuracy of

each category, provides a way to evaluate the impact of different categories on the overall classification effect, and more comprehensively measures the performance of the classifier on each category. Its definition is as follows [9]:

$$AA = \frac{\sum_{i=1}^{r} \frac{c_{ii}}{c_{+i}}}{N} \times 100\%, \tag{7}$$

where c_{ii} represents the number of class i samples that are correctly classified, c_{+i} represents the pixel number classified into class i at the time of prediction.

The Kappa coefficient fully considers the impact of uncertainty on classification results, is used to test the degree of consistency, and can objectively measure the overall result of classification. Its definition is as follows [9]:

$$\text{Kappa} = \frac{N \sum_{i=1}^{r} c_{ii} - \sum_{i=1}^{r}(c_{i+} \times c_{+i})}{N^2 - \sum_{i=1}^{r}(c_{i+} \times c_{+i})}, \tag{8}$$

where, N represents the total number of samples, c_{i+} represents the real pixel number of class i, and c_{+i} represents the pixel number classified into class i at the time of prediction.

For the classification effects of different HSI ground object categories, we conducted a comparative analysis using the measurement indicators OA, AA and Kappa. OA represents the total accuracy rate of classification. Referring to Eq. (6), the larger the OA value, the better the overall classification effect of the ground object category. AA is a classification measurement index for each type of ground object. Referring to Eq. (7), the larger the AA value, the better the classification effect of a certain type of ground object. Referring to Eq. (8), the Kappa coefficient is a ratio that represents the reduction in errors between a classification and a completely random classification.

3.3. Experimental flow chart

The flow chart of the experimental simulation is shown in Fig. 1. The 2D-C-VMD of each band of the original HSI is carried out separately. The 2D-C-VMD can decompose each 2D band image of HSI into sub-band images of different scales. It is worth noting that the number of sub-band images obtained by decomposition is what we first need to give, that is, the prior parameter K value. In literature [16], we explained the prior parameter K value in detail. The sub-band images of different scales obtained from each 2D image decomposition are stacked respectively, and several feature samples of different scales are obtained. Also known as in Fig. 1 is first decomposed Mode 1, Mode 2 and Mode K respectively to carry on the stack, respectively get $B_{Mode1}^{f_{(i,j)}^{l}}$, $B_{Mode2}^{f_{(i,j)}^{l}}$ and $B_{ModeK}^{f_{(i,j)}^{l}}$. The $f_{(i,j)}^{l}$ is used to represent each 2D band image of the HSI, l represents the number of bands, and (i,j) represents the spatial position of each band. Then, the first mode component obtained by decomposition is smoothed again to obtain the smoothed feature information sample. Finally, the feature information sample is classified by support vector machine and compared according to the analysis result.

Below we give a detailed description of the flow chart in Fig. 1.

Step 1: Two original data sets, Indian Pine and Pavia U, were selected to remove the band images with heavy noise pollution. HSI data is expressed as $R^{\{W \times H \times L\}}$, W represents the width of spatial dimensional data, H represents the height of spatial dimensional data, and L represents the total number of bands of HSI.

Step 2: Each spatial image $(f_{(i,j)}^{l})$ is decomposed by 2D-C-VMD, and several mode components are obtained respectively, and then each mode component is re-stacked, respectively get $B_{Mode1}^{f_{(i,j)}^{l}}$, $B_{Mode2}^{f_{(i,j)}^{l}}$ and $B_{ModeK}^{f_{(i,j)}^{l}}$.

Step 3: Repeat step 2 to decompose the decomposed $B_{Mode1}^{f_{(i,j)}^l}$ again by 2D-C-VMD, and stack the first pattern component obtained by decomposition, then get respectively $\bar{B}_{Mode1}^{f_{(i,j)}^l}$.

Step 4: The SVM classification and effect analysis were carried out on the pattern component features obtained from the third step decomposition.

It is worth noting that the influence of the prior parameter K value should be considered when performing the 2D-C-VMD decomposition in step 3. Here, we refer to the scheme proposed in reference [16] for the optimal K-value screening process.
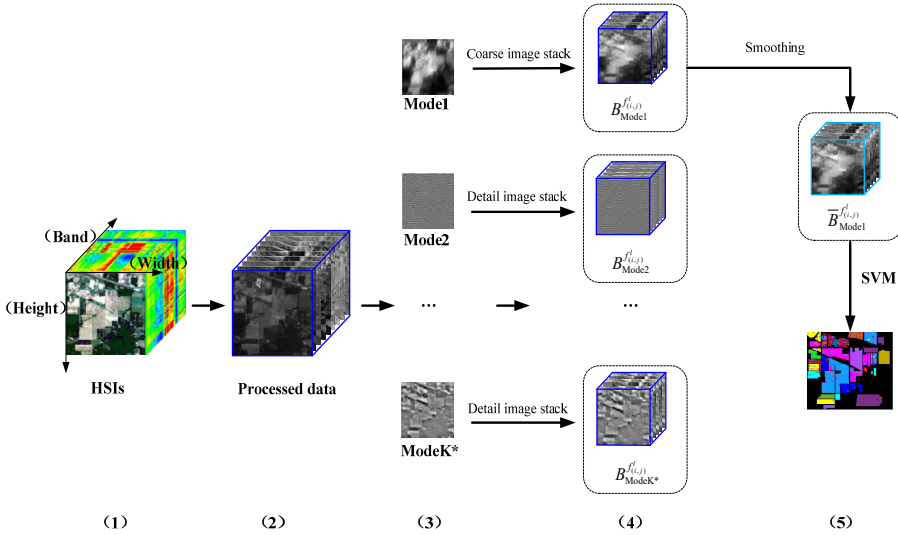

Fig. 1. Feature extraction and classification of HSIs with improved 2D-C-VMD

3.4. Analysis of experimental results

In Fig. 2, sub-band images of Indian Pine and Pavia University at 4 different scales were obtained by 2D-C-VMD in the 23rd band and the 30th band, respectively. In the literature [16], we discussed and analyzed the prior parameter K values of different HSIs. The original 23th-band and 30th-band images of the two data sets are shown in subgraphs a) and b) of Fig. 2. From the decomposed sub-band images of different scales, we can clearly see that the first sub-band obtained from the decomposition of the two sets of data contains the low-frequency contour component feature information in the original image. Subsequent sub-band images contain high frequency detail feature information of edges and textures in the original image.

According to the preliminary classification results of different features, the results of the first mode component of the two sets of data are the best. It shows that the first low-frequency mode component obtained by decomposition can suppress the adverse effect of spatial domain noise on classification. At the same time, we also compare and analyze the feature information of different scale sub-band images. In order to further mine the effective feature information of the first mode component obtained by decomposition, we perform the two-dimensional compact variational mode decomposition on the first mode decomposition again to extract more effective feature information conducive to classification.

In Fig. 3, we again use 2D-C-VMD for the first optimal mode component from the Indian Pine and Pavia University decomposition. Through comparative analysis of several experiments, we took 3 and 4 prior parameters of the data sets of Indian Pine and Pavia university respectively. From the decomposition of the 23 band and 30 band in Fig. 3 to obtain sub-band images of different scales, we can find that the first mode component obtained by decomposition is smoother.

| a) IP Band 23 | a1) Mode 1 | a2) Mode 2 | a3) Mode 3 | a4) Mode 4 |

| b) PU Band 30 | b1) Mode 1 | b2) Mode 2 | b3) Mode 3 | b4) Mode 4 |

Fig. 2. Sub-band images of different scales obtained by 2D-C-VMD decomposition of Indian Pine and Paiva University

| a) Original IP image | a1) Mode 1 | a2) Mode 2 | a3) Mode 3 |

| b) Original PU image | b1) Mode 1 | b2) Mode 2 | b3) Mode 3 | b4) Mode4 |

Fig. 3. Sub-band images of different scales obtained by 2D-C-VMD of the optimal low-frequency component
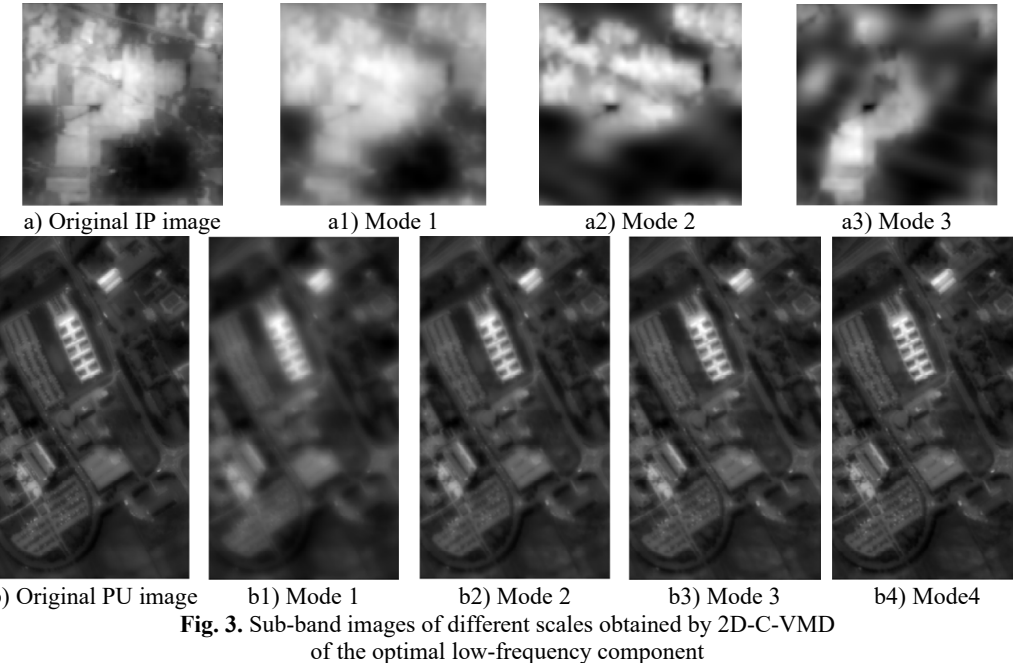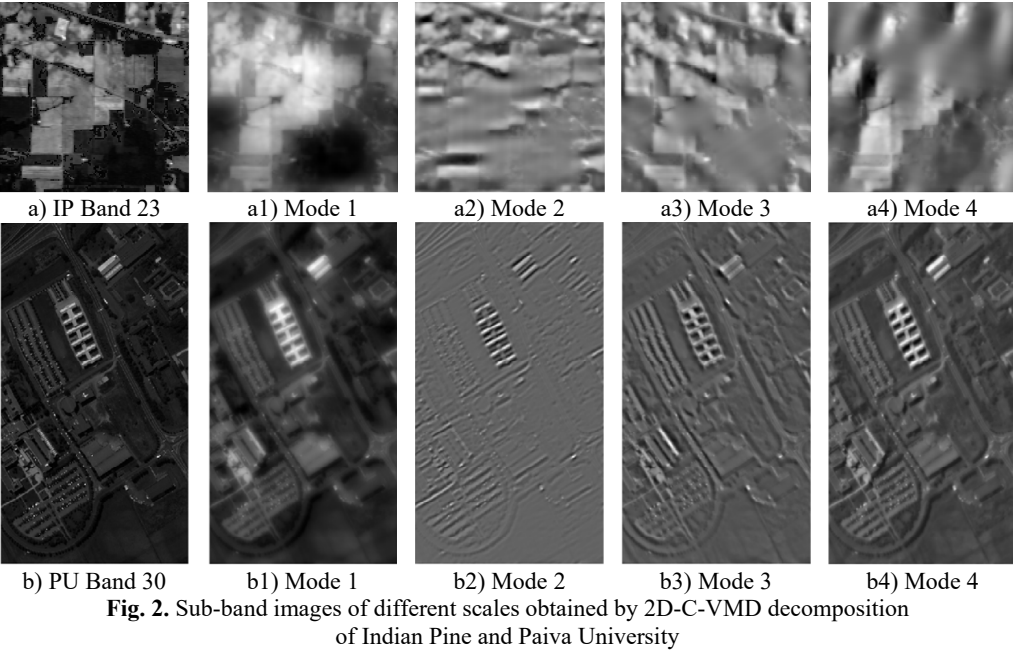
Table 2 and Table 3 show the classification of Indian Pine and Pavia University data sets under different schemes. The measurement indexes were represented by AA, OA and Kappa respectively. Select 10 % and 2 % of different feature information as the training set, and the rest as the test set. We choose the support vector machine network with good classification performance for small sample data, and the network adopts cross-validation for parameter optimization. The Raw in Table 2 and Table 3 represent the results of support vector machine classification directly on the original data set. L-2D-C-VMD means that the optimal

low-frequency component obtained from the first decomposition is decomposed again by two-dimensional compact variational modes. Let's focus first on the OA metric, because from this metric we can see the overall performance of a category in the data set. It can be seen from the classification effect that the OA value of the original Indian Pine was 78.70 % directly identified by classification, and after the first 2D-C-VMD decomposition, the OA value was increased to 97.57 %.

Table 2. The AA, OA and Kappa values of Indian Pine under different schemes
for different feature classifications

Class	Raw	2D-C-VMD	L-2D-C-MVD
1	62.44±15.48	90.24±7.80	98.54±1.26
2	75.05±3.56	95.14±2.12	97.37±0.99
3	64.10±6.75	97.56±1.47	98.63±0.30
4	50.70±5.05	95.87±2.15	97.56±0.85
5	90.34±2.74	99.36±0.36	99.12±0.89
6	94.79±0.78	99.67±0.28	98.90±0.78
7	84.80±5.59	95.20±1.69	99.20±1.69
8	97.02±1.10	99.95±0.10	99.86±0.29
9	38.89±12.28	86.67±12.06	62.22±22.65
10	68.37±1.52	96.34±1.86	97.39±1.62
11	79.33±3.08	97.20±1.35	98.04±0.58
12	69.21±3.34	96.70±1.16	96.70±1.07
13	94.89±3.43	99.35±0.84	99.24±1.06
14	95.03±1.37	99.86±0.13	99.40±0.40
15	48.36±13.06	98.56±1.57	99.48±0.65
16	90.00±4.39	96.90±0.61	92.62±3.92
AA	75.21±0.91	96.54±0.96	95.89±1.53
OA	78.70±0.76	97.57±0.48	98.17±0.19
Kappa×100	75.64±0.96	97.23±0.55	97.92±0.21

After the first filtering, the classification performance of different feature categories is greatly improved. However, after we smoothed the optimal low-frequency component again, there was also a small increase in the value of OA in the dataset Indian Pine. In the OA index, the overall OA decreases due to the influence of category 9. In the dataset Pavia university, the OA value of the original Raw was 87.03 %, and after the first 2D-C-VMD processing, the OA value became 95.21 %, which suppressed the adverse impact of spatial dimension noise on classification to a certain extent. Similarly, after smoothing and filtering the first mode component, OA value increased by about 1.7 %, Kappa increased by about 2 %, and AA increased by about 3 %.

In order to avoid the contingency of the results of a single training sample, we selected different proportions of training samples from Indian Pine and Pavia University as the training set, and the rest as the test set for comprehensive comparative analysis. Here, OA and Kappa indexes are selected for metric analysis. Support vector machine classifier is also used to evaluate the classification of ground objects under different training samples. In Fig. 4, the OA and Kappa values of the recognition results of Indian Pine training samples at 4 %, 7 %, 10 % and 20 % are given. From the OA and Kappa values, it is obvious that after 2D-C-VMD spatial filtering, the classification performance of support vector machine network can be further improved while overcoming spatial noise interference. In Fig. 5, the OA and Kappa values of Pavia University training samples under 0.5 %, 1 %, 2 % and 5 % are given.

By combining the recognition effect of Fig. 4 and Fig. 5, we can conclude that the classification performance of different ground objects can be further improved after the quadratic 2D-C-VMD of the original HSI, and the adverse effects of spatial domain noise on classification can be suppressed to a certain extent. At the same time, in the case of less training samples, the recognition effect advantage is more obvious through the secondary use of 2D-C-VMD.
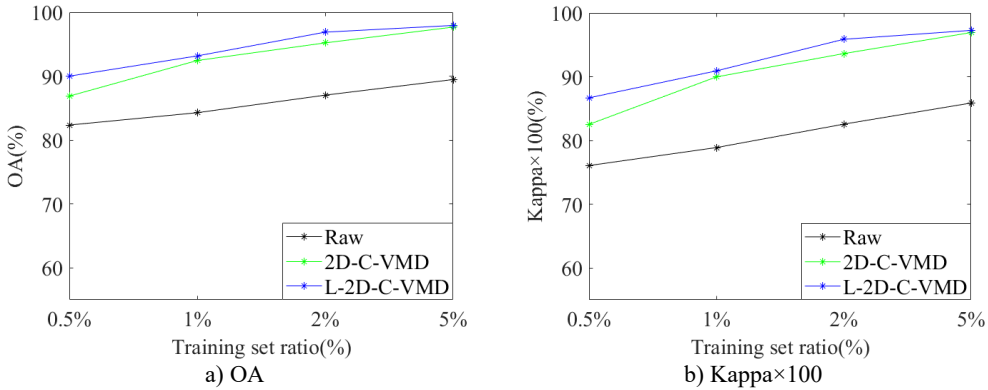
Table 3. The AA, OA and Kappa values of Pavia University under different schemes for different feature classifications

Class	Raw	2D-C-VMD	L-2D-C-MVD
1	87.31±2.38	95.30±0.64	97.21±1.12
2	96.31±0.98	99.57±1.53	99.44±0.48
3	65.25±1.50	91.76±3.88	90.20±3.24
4	89.31±2.19	86.71±2.89	88.34±0.98
5	99.38±0.16	95.86±5.90	98.96±0.64
6	63.03±5.86	99.41±0.46	98.90±0.56
7	72.80±3.98	97.24±1.70	97.21±2.73
8	81.90±3.85	84.04±3.54	94.96±1.73
9	93.19±4.70	61.16±8.13	80.29±5.81
AA	83.16±1.26	90.12±0.39	93.95±0.98
OA	87.03±0.74	95.21±0.25	96.89±0.45
Kappa×100	82.56±0.72	93.63±0.33	95.87±0.60

a) OA   b) Kappa×100
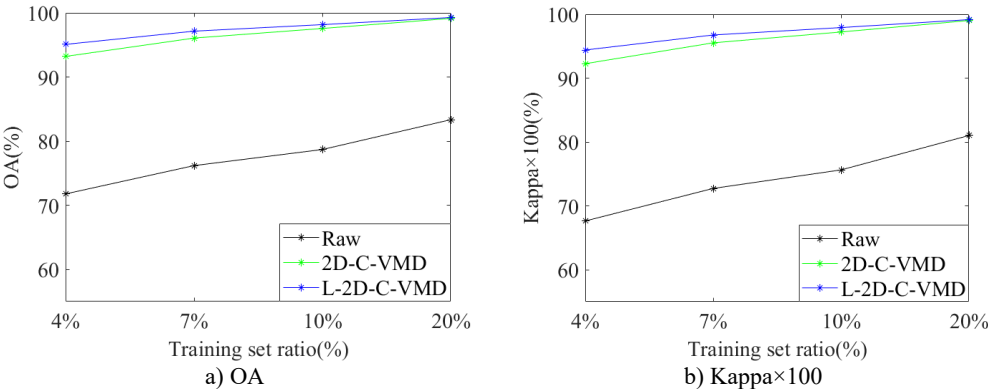
Fig. 4. Classification of OA and Kappa values of Indian Pine data under different training samples

a) OA   b) Kappa×100

**Fig. 5.** Classification of OA and Kappa values of Pavia University data under different training samples

In Table 4, we make a comparative analysis of different spatial domain methods. Indian Pine and Pavia University datasets selected 10 % and 2 % training samples respectively as training sets, and the rest were used as test sets. The classifier still uses the support vector machine as the classifier.

Fig. 6 and Fig. 7 respectively show the color maps of different feature classification effects of Indian Pine and Pavia University data sets under different spatial domain feature extraction methods. Subfigure a) is the color map label for the different features of the original data set. Raw represents a rendering of the original data set directly classified by support vector machine. The

closer the classification color image effect is to subfigure a), the better the feature extraction effect is. From Fig. 6, you can see that there are between 1 and 16 different color labels, and each color represents a feature category. The details of each feature category can be combined with the list given in Table 1. It can be seen from the color map of classification that the feature information extracted by 2D-EWT in Indian Pine data is less effective than 2D-EMD and 2D-C-VMD. The effect of 2D-EMD is close to that of 2D-C-VMD. After using 2D-C-VMD twice, the classification effect is improved to some extent. Improved from the original OA value of 78.7 % to 98.17 %. There are a total of 9 ground object categories in the data of Pavia University, and the reference value OA given is 87.03 %. After the extraction of spatial domain features, the classification effect has been improved to a certain extent.

Table 4. Comparison of feature extraction effect of different spatial domain methods

Indian Pine						
Class	Raw	2D-EMD	2D-LP-EWT	2D-T-EWT	2D-C-VMD	L-2D-C-MVD
OA	78.70±0.76	97.74±0.18	93.99±0.72	86.48±0.72	97.57±0.48	98.17±0.19
Kappa×100	75.64±0.96	97.43±0.21	93.14±0.82	84.57±0.83	97.23±0.55	97.92±0.21
Pavia University						
Class	Raw	2D-EMD	2D-LP-EWT	2D-T-EWT	2D-C-VMD	L-2D-C-MVD
OA	87.03±0.74	93.67±0.59	90.38±0.38	92.76±0.28	95.21±0.25	96.89±0.45
Kappa×100	82.56±0.72	91.60±0.79	87.15±0.53	90.35±0.40	93.63±0.33	95.87±0.60

a) Classified color chart

b) Raw

c) 2D-EMD

d) 2D-LP-EMD

e) 2D-T-EMD
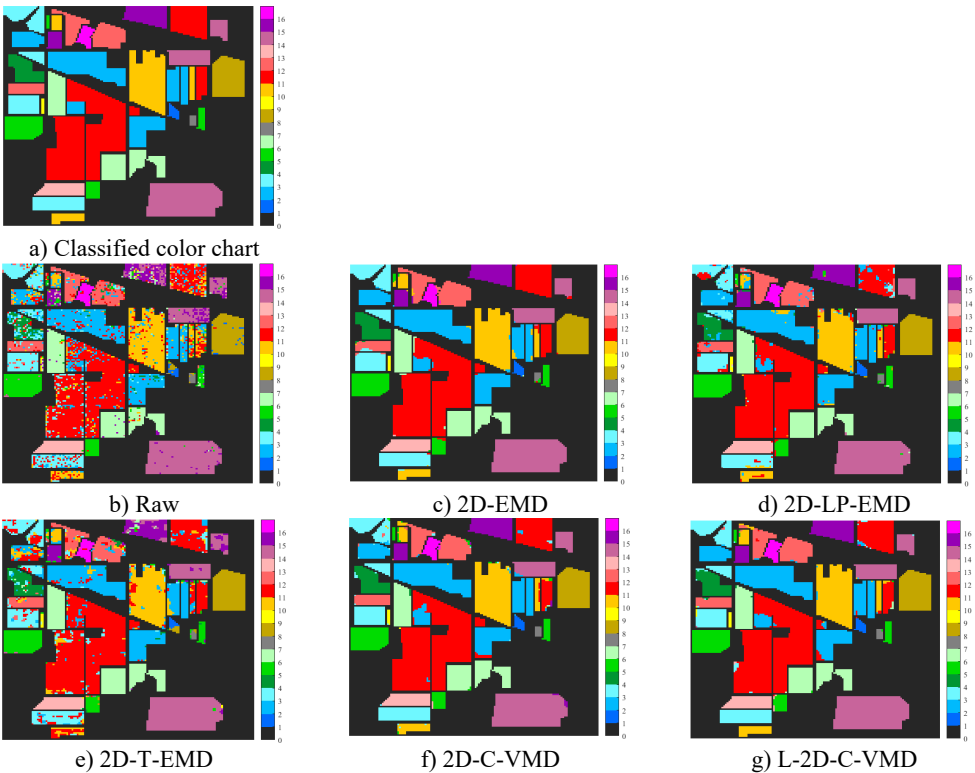
f) 2D-C-VMD

g) L-2D-C-VMD

Fig. 6. Classification results of Indian Pine dataset under different spatial domain methods

From the classification effect, 2D-C-VMD spatial domain filter has the best OA value enhancement effect. When 2D-C-VMD is used again to filter the optimal low-frequency component obtained from the first decomposition, it can be seen from the OA and Kappa indexes that there is also some room for improvement. In addition, after the previous discussion of different spatial domain methods, 2D-EMD uses iterative decomposition to obtain several inherent

mode components, and the algorithm convergence speed is slow and the computational efficiency is low, so it often requires more computational costs when processing HSIs containing multiple bands. 2D-EMD also lacks theoretical support in mathematical theory. There are many sub-methods of 2D-EWT, so 2D-LP-EWT and 2D-T-EWT are selected here for comparative analysis. The 2D-EWT combines the adaptability of EMD and the theoretical completeness of wavelet, and can be used for feature extraction of hyperspectral images. From the classification effect of feature extraction, there is still a large room for improvement. Using the variational idea in mathematics, the 2D-C-VMD method can decompose the original image into sub-band images of different scales. The variational idea can screen several sub-band images of different scales at the same time. The decomposition of each sub-band image has a good timeliness, which can greatly save the time of feature extraction. In literature [16], we carried out relevant research and analysis.

According to the classification performance of Indian Pine and Pavia University data sets under different spatial domain methods, the smooth classification effect of 2D-C-VMD in the two sets of data is better. The method is simple and easy to operate. It smoothed the first mode component obtained by decomposition. According to the classification effect, 2D-C-VMD has good robustness in smoothness.
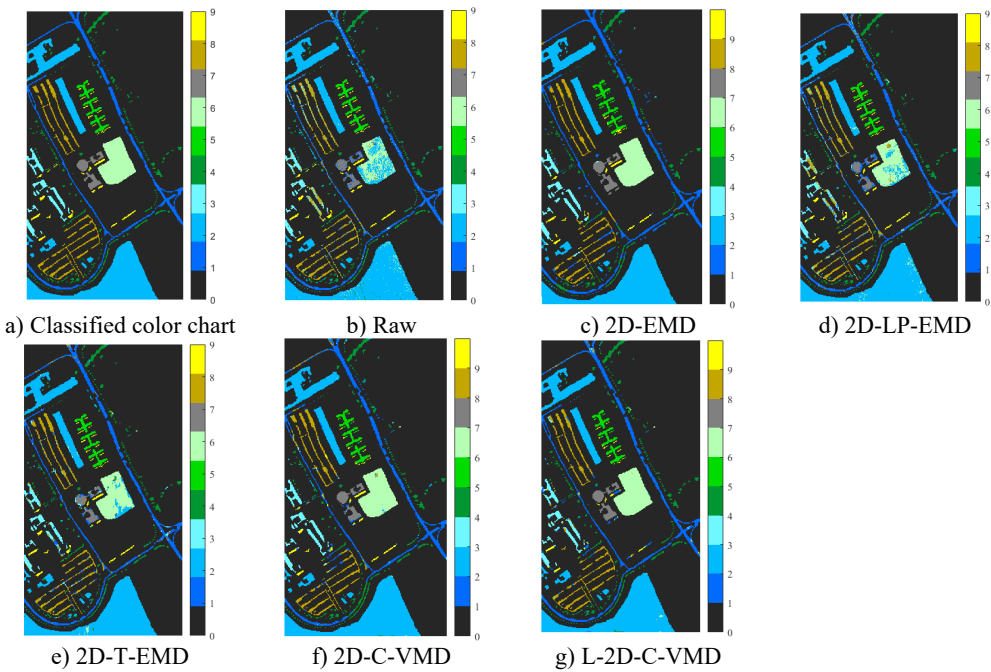

a) Classified color chart　　b) Raw　　c) 2D-EMD　　d) 2D-LP-EMD

e) 2D-T-EMD　　f) 2D-C-VMD　　g) L-2D-C-VMD

Fig. 7. Classification results of Pavia University dataset under different spatial domain methods

4. Conclusions

In this paper, we explore the feature extraction and classification of HSIs by the L-2D-C-VMD. We use 2D-C-VMD to smooth the first mode component obtained from the initial decomposition. From the simulation results of the two sets of HSIs experimental data, it can be seen that after smoothing the first mode component and suppressing the spatial dimension noise, the classification performance for different ground objects can be further improved.

However, it is worth noting that the feature information scale of the image includes the low frequency contour information and the high frequency detail feature information such as the edge texture in the image. In the follow-up research process, we will further consider the fusion analysis

of the smoothed feature information and the high-frequency detail feature information.

Acknowledgements

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

Renxiong Zhuo: methodology, writing and funding acquisition. Cui Mao: formal analysis. Ruiying Kuang: funding acquisition and formal analysis.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] M. E. Grotte et al., "Ocean color hyperspectral remote sensing with high resolution and low latency-the HYPSO-1 CubeSat mission," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 60, pp. 1–19, Jan. 2022, https://doi.org/10.1109/tgrs.2021.3080175

[2] M. Wang, Y. Xu, Z. Wang, and C. Xing, "Deep margin cosine autoencoder-based medical hyperspectral image classification for tumor diagnosis," *IEEE Transactions on Instrumentation and Measurement*, Vol. 72, pp. 1–12, Jan. 2023, https://doi.org/10.1109/tim.2023.3293548

[3] C.-I. Chang, "Hyperspectral anomaly detection: a dual theory of hyperspectral target detection," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 60, pp. 1–20, Jan. 2022, https://doi.org/10.1109/tgrs.2021.3086768

[4] Q. Shi, X. Tang, T. Yang, R. Liu, and L. Zhang, "Hyperspectral image denoising using a 3-D attention denoising network," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 59, No. 12, pp. 10348–10363, Dec. 2021, https://doi.org/10.1109/tgrs.2020.3045273

[5] P. V. Arun, K. M. Buddhiraju, A. Porwal, and J. Chanussot, "CNN-based super-resolution of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 58, No. 9, pp. 6106–6121, Sep. 2020, https://doi.org/10.1109/tgrs.2020.2973370

[6] B. Demir and S. Erturk, "Empirical mode decomposition of hyperspectral images for support vector machine classification," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, No. 11, pp. 4071–4084, Nov. 2010, https://doi.org/10.1109/tgrs.2010.2070510

[7] T. V. N. Prabhakar and P. Geetha, "Two-dimensional empirical wavelet transform based supervised hyperspectral image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 133, pp. 37–45, Nov. 2017, https://doi.org/10.1016/j.isprsjprs.2017.09.003

[8] J. Tang et al., "Feature extraction of hyperspectral images based on SVM optimization of 2D-EMD and GWO," *Journal of Measurements in Engineering*, Vol. 12, No. 4, pp. 548–561, Dec. 2024, https://doi.org/10.21595/jme.2024.23844

[9] J. Tang et al., "Research on the relationship between feature extraction time and training samples of hyperspectral image based on spatial domain," *Journal of Measurements in Engineering*, Vol. 13, No. 1, pp. 73–88, Mar. 2025, https://doi.org/10.21595/jme.2024.24249

[10] R. Zhuo, Y. Guo, and B. Guo, "A hyperspectral image classification method based on 2-D compact variational mode decomposition," *IEEE Geoscience and Remote Sensing Letters*, Vol. 20, pp. 1–5, Jan. 2023, https://doi.org/10.1109/lgrs.2023.3268776

[11]   D. Zosso, K. Dragomiretskiy, A. L. Bertozzi, and P. S. Weiss, "Two-dimensional compact variational mode decomposition," *Journal of Mathematical Imaging and Vision*, Vol. 58, No. 2, pp. 294–320, Feb. 2017, https://doi.org/10.1007/s10851-017-0710-z

[12]   M. Graña, M. A. Veganzons, and B. Ayerdi, "Hyperspectral remote sensing scenes," Computational Intelligence Group, 2024.

[13]   W.-T. Zhang, Y.-B. Li, L. Liu, Y. Bai, and J. Cui, "Hyperspectral image classification based on spectral-spatial attention tensor network," *IEEE Geoscience and Remote Sensing Letters*, Vol. 21, pp. 1–5, Jan. 2024, https://doi.org/10.1109/lgrs.2023.3332600

[14]   A. K. Singh, R. Sunkara, G. R. Kadambi, and V. Palade, "Spectral-spatial classification with naive bayes and adaptive FFT for improved classification accuracy of hyperspectral images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 17, pp. 1100–1113, Jan. 2024, https://doi.org/10.1109/jstars.2023.3327346

[15]   Y. Su, L. Gao, M. Jiang, A. Plaza, X. Sun, and B. Zhang, "NSCKL: normalized spectral clustering with kernel-based learning for semisupervised hyperspectral image classification," *IEEE Transactions on Cybernetics*, Vol. 53, No. 10, pp. 6649–6662, Oct. 2023, https://doi.org/10.1109/tcyb.2022.3219855

[16]   R. Zhuo, Y. Guo, B. Guo, B. Liu, and F. Dai, "Two-dimensional compact variational mode decomposition for effective feature extraction and data classification in hyperspectral imaging," *Journal of Applied Remote Sensing*, Vol. 17, No. 4, Dec. 2023, https://doi.org/10.1117/1.jrs.17.044517

Renxiong Zhuo received his Ph.D. degree from Hangzhou Dianzi University in 2024. Now he works at Changsha Institute of Technology. His main research interests include hyperspectral image processing, data mining, machine learning, and deep learning.

Cui Mao received her M.A. degree in design from the Fine Arts Academy at Hunan Normal University and is currently employed at Changsha Institute of Technology. Her main research areas include interior space design, narrative space application, illustration design, machine learning, and children's picture books.

Ruiying Kuang obtained her bachelor's degree from Changsha University of Science and Technology and master's degree from Hunan Normal University. Her main research interests include protection and digitalization on intangible cultural heritage. She had been an Associate Professor at the Art Institute of Xiangtan University from 2014 to 2023. Now she is an Associate Professor at the Changsha Institute of Technology.