

## Research Article

# Hardwood Species Classification with Hyperspectral Microscopic Images

Peng Zhao  and Cheng-Kun Wang 

Information and Computer Engineering College, Northeast Forestry University, Harbin City 150040, China

Correspondence should be addressed to Peng Zhao; bit\_zhao@aliyun.com

Received 1 February 2019; Revised 29 April 2019; Accepted 3 June 2019; Published 27 June 2019

Academic Editor: Wee Chew

Copyright © 2019 Peng Zhao and Cheng-Kun Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a hardwood species identification method based on wood hyperspectral microscopic images. A SOC710VP hyperspectral stereomicroscope was used to acquire microscopic images of a hardwood cross section. In these microscopic images, each part's spectral features are discussed. We found that the spectral divisibility of wood vessels' peripheral and central regions in the hyperspectral microscopic images can be used for hardwood species recognition. Mathematical morphological operation and K-L divergence were used to extract spectral features at the wood vessels' peripheral regions and central regions, respectively. By comparing wood vessels' spectral similarity across wood species samples, we found that wood vessels' peripheral spectral divisibility is larger than its central. Finally, the spectral information from randomly selected regions of interest (i.e., ROI) and that of wood vessels' peripheral and central regions have value as a classification basis. In our hardwood species classification experiments, three dimensionality reduction algorithms, principal component analysis (PCA), kernel principal component analysis (KPCA), and multidimensional scaling (MDS), and the three classifiers, BP neural network, support vector machine (SVM), and Mahalanobis distance (MD), are combined to perform hardwood species classification work. Experimental results indicate that the best recognition effect can be achieved at the peripheral region of wood vessels using PCA or MDS with the MD algorithm.

## 1. Introduction

Wood is an important renewable natural resource. Different tree species result in many different wood products circulating in the market. In particular, many hardwood species are widely used to make furniture, and they cannot be identified by visual observation or odor discrimination. Hardwood misclassification may result in significant economic losses due to the differences in quality and price of different hardwood species.

With the rapid development of machine vision, it has become possible to identify the processed hardwood species by using intelligent methods. Pan and Kudo [1] and Yusof et al. [2] used microscopic images of wood cross sections to discuss the characteristics of different hardwood species. They showed that different hardwood species have differences in their wood's vessel structure. Furthermore, some textural or statistical properties for wood vessels were

extracted from the 2D wood images to perform hardwood species classification [3, 4]. Macroscopic wood features are also used in wood species recognition. For example, Hu et al. used a deep learning scheme for lumber grade classification in terms of lumber defects, textures, and species [5].

In recent years, spectral analysis schemes have been proposed for wood species classification [6–8]. These schemes are in essence spectrum-based classifications which mainly use 1D spectral curves. They usually have low computational complexity with fast speed due to the 1D data but may suffer from spectral noise and disturbance from the surrounding site. Therefore, wood species recognition with 3D hyperspectral imaging which includes both visual and spectral information of wood tissues may be a good choice. In fact, 3D hyperspectral imaging schemes have been used in agricultural engineering for food quality evaluation [9] and crop content predictions [10, 11].

In this paper, the microscopic hyperspectral images of hardwood cross sections were recorded using a SOC710VP hyperspectral stereomicroscope to investigate the possibility of hardwood species identifications. The spectral features of different wood tissues such as wood vessel, wood ray, and background are discussed. It was found that spectral information from the peripheral and central regions of wood vessels can be used for hardwood species identification.

## 2. Materials and Methods

**2.1. Wood Samples.** Six hardwood species were studied (Table 1). One thousand four hundred and forty wood samples were prepared from six species, and each sample was cut into a cube with length, width, and height of 2 cm, 2 cm, and 3 cm, respectively. Nine hundred samples were randomly selected as the training set, and the other 540 samples are used as the test set. Each sample's cross sections (2 cm  $\times$  2 cm) were cut using a disk saw to make the two surfaces as smooth as possible, while the radial sections and tangential sections (3 cm  $\times$  2 cm) were cut using an electric plane. The wood sample's cross sections were used for hyperspectral imaging. In these cross sections, the wood growth rings were as close to parallel with the section's two edges as possible. *Quercus rubra* and *Fraxinus americana* had similar colors, while cross sections for the other four species had different colors.

The intraclass differences of wood samples are caused by the tree's age, cutting position, and place of growth. For each wood species, our wood blocks are taken from different individual trees at random position in the cross section of the timber to include these intraclass differences. In this way, the reliability of our proposed scheme can be ensured.

**2.2. Spectral Data Acquisition.** The spectral acquisition instrument used for this research was a SOC710VP hyperspectral stereomicroscope, with 128 spectral bands and a spectral range of 372–1038 nm. To improve spectral acquisition, all wood sample surfaces were flat, without burrs, and the spectral acquisition environment was consistent and stable. The dimension of captured microscopic hyperspectral images was 520  $\times$  696  $\times$  128. To eliminate the influence of dark current, it was also necessary to perform black and white correction processing on the detected samples. The wood image's enlargement factor was 45.

To control costs, we chose the visible/near infrared hyperspectral stereomicroscope as the experimental instrument. Spectral data in the visible region may vary because wood sample color may change with environmental variation. Therefore, we maintained uniform environmental conditions during the spectral experiments and wood sample preservation. In addition, 780–900 nm is the near infrared band, which best characterizes the different wood properties. A stereomicroscope was used to capture wood microscopic images conveniently for each wood sample. The wood slice is not used as its production is complex and time consuming.

TABLE 1: Species studied.

Number	English name	Scientific name
1	Red oak	<i>Quercus rubra</i>
2	Ash	<i>Fraxinus americana</i>
3	Merbau	<i>Intsia bijuga</i>
4	Golden silk teak	<i>Tectona grandis</i>
5	Rose wood	<i>Guibourtia tessmannii</i>
6	Limba	<i>Terminalia neotaliala capuron</i>

**2.3. Spectral Analysis of Different Tissues of Hardwood Microscopic Hyperspectral Images.** The microscopic hyperspectral images of cross sections mainly include wood vessels, wood rays, and axial parenchyma, as shown in Figure 1. The closed holes are wood vessels which are important features for species recognition [12, 13]. The dense, thin, approximately parallel lines are called wood rays, while the thick lines approximately perpendicular to the wood rays are the axial parenchyma.

Hyperspectral images are 3D cubes which have both spatial information and spectral information, and the spectral curves of different wood tissues for the same hardwood species are different. For example, Figure 2 illustrates the spectral curves of *Intsia bijuga* in the central and peripheral regions of vessels, in the regions of wood rays, axial parenchyma, and the background. Since vessels are the most important features in hardwood species identification [1, 3, 4, 12], the spectral information in the vessels' central and peripheral regions is extracted and discussed.

**2.4. Spectral Analysis in Vessel Peripheral Regions with Mathematical Morphology.** Mathematical morphology is widely used in image processing, which can be used to extract useful image components to enhance ROI and suppress background regions. We denote the original image, background image, and ROI image as  $\mathbf{H}(x, y)$ ,  $\mathbf{B}(x, y)$ , and  $\mathbf{F}(x, y)$ , respectively. The structural element with a specified shape is denoted as  $\mathbf{A}$ . The background image and ROI image can be expressed as follows:

$$\begin{aligned} \mathbf{B}(x, y) &= \mathbf{H}(x, y) \circ \mathbf{A}(x, y), \\ \mathbf{F}(x, y) &= \mathbf{H}(x, y) - \mathbf{B}(x, y), \end{aligned} \quad (1)$$

where  $\circ$  represents the morphological opening operation.

In order to eliminate noise, a corrosion operation is required for the obtained  $\mathbf{F}(x, y)$ . Then, the binary image is formed by using threshold processing to  $\mathbf{F}(x, y)$ . The gray-level image for each band of hyperspectral image is processed using the abovementioned procedures to obtain the 128 binary images which form a cube of  $\mathbf{M} \times \mathbf{N} \times 128$  dimensions (i.e.,  $\mathbf{M}$  and  $\mathbf{N}$  are the width and height of hyperspectral image).

This cube can be regarded as  $\mathbf{M} \times \mathbf{N}$  vectors of 128 dimensions, and the element in every vector is either 0 or 1. The number of element 1 in every vector is registered so as to form a matrix  $\mathbf{D}$  of  $\mathbf{M} \times \mathbf{N}$  dimensions. A threshold  $k$  is defined to convert the matrix  $\mathbf{D}$  into a binary image  $\mathbf{E}$  as follows, which is illustrated in Figure 3.

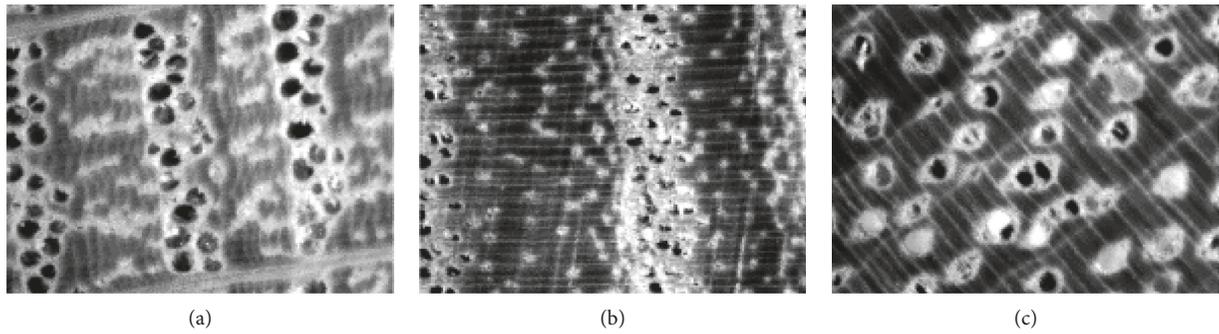


FIGURE 1: Microscopic hyperspectral images of cross sections of different hardwood species. (a) *Quercus rubra*. (b) *Fraxinus americana*. (c) *Intsia bijuga*.

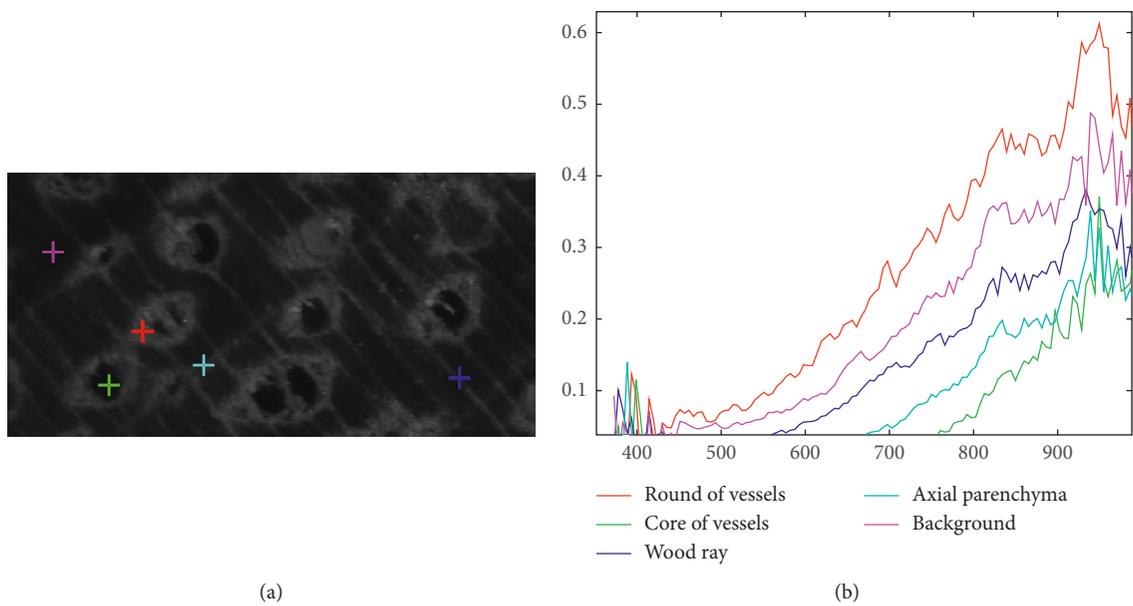


FIGURE 2: Spectral curves for different tissues of *Intsia bijuga*: (a) the 5 cross marks indicate different tissues; (b) corresponding spectral curves for each tissue.

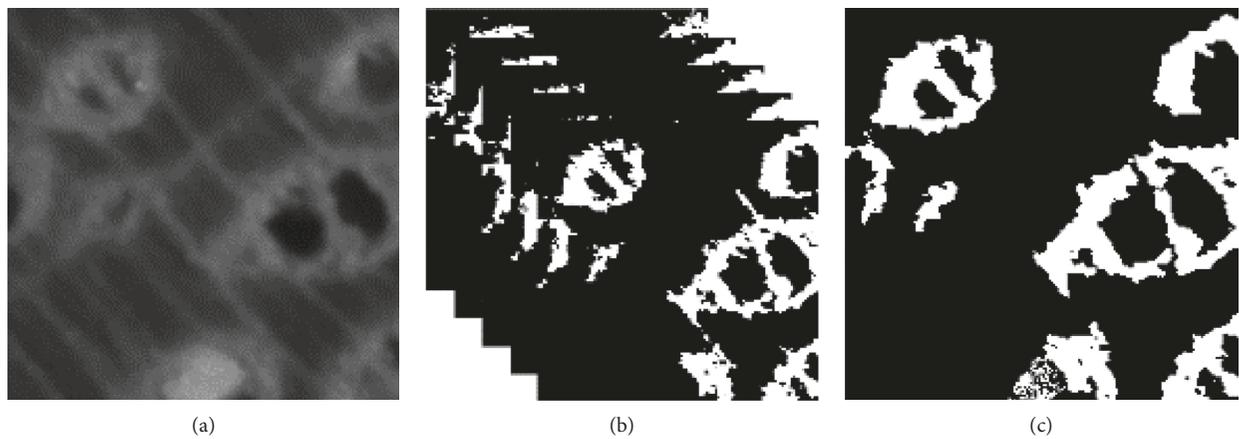


FIGURE 3: The wood vessels' peripheral region extraction with mathematical morphology: (a) the hyperspectral 56<sup>th</sup> band image; (b) the binary image cube of  $M \times N \times 128$  dimensions; (c) the binary image  $E$ .

$$\mathbf{E}(x, y) = \begin{cases} 1, & \text{if } \mathbf{D}(x, y) \geq k, \\ 0, & \text{if } \mathbf{D}(x, y) < k. \end{cases} \quad (2)$$

For the binary image  $\mathbf{E}$ , the spectra are extracted whose corresponding pixel value is 1 so as to form a new matrix  $\mathbf{X}$  of  $C \times 128$  dimensions, while  $C$  represents the number of element 1 in binary image  $\mathbf{E}$ .

The abovementioned image and spectral processing operations were performed to a ROI in a hyperspectral image of one sample. In this work, the size of ROI is set as  $100 \times 100$ . In a wood hyperspectral image of  $520 \times 696 \times 128$  dimensions, the initial ROI is placed in the upper-left position in the image to obtain the computed matrix  $\mathbf{X}$ . Then, the ROI moves gradually from left to right and from top to bottom to traverse the whole hyperspectral image to obtain multiple matrices  $\mathbf{X}_i$ ,  $i = 1, 2, \dots, n$ . These matrices  $\mathbf{X}_i$ ,  $i = 1, 2, \dots, n$  are connected by columns to obtain a matrix  $\mathbf{Y}$  with a dimension of  $(C_1 + \dots + C_n) \times 128$ , and the mean value for each column is computed to form a new vector  $\mathbf{y}$  of 128 dimensions. For a hardwood species with  $m$  hyperspectral image specimens, we may obtain multiple vectors  $\mathbf{y}_j$ ,  $j = 1, 2, \dots, m$  and calculate the mean vector  $\bar{\mathbf{y}}$ . As an example, Figure 4 illustrates the mean spectral vector  $\bar{\mathbf{y}}$  for the six hardwood species in our measurements in the wood vessels' peripheral regions. The training samples are used to calculate  $\bar{\mathbf{y}}$  with the ROI of  $100 \times 100$  and the structural element  $\mathbf{A}$  of disk shape with the size  $r$ . In Figure 4, we can see that the mean spectral curves for training sets in wood vessels' peripheral regions are different for the six hardwood species. Therefore, this spectral difference can be used to identify hardwood species. Since different hardwood species have different vessel sizes, the structural element  $\mathbf{A}$  with different size  $r$  is used for the 6 hardwood species. This is explained in detail in Section 3.4.

**2.5. Spectral Analysis in Wood Vessels' Central Regions with K-L Divergence.** K-L divergence is widely used in information theory as a signal similarity measuring tool [14]. We suppose the probability distributions of two discrete random signals as  $\mathbf{X} = [x_1 \ x_2 \ \dots \ x_n]^T$  and  $\mathbf{Y} = [y_1 \ y_2 \ \dots \ y_n]^T$ , respectively. The K-L divergence for  $\mathbf{Y}$  to  $\mathbf{X}$  is then defined as follows:

$$D_{KL}(X||Y) = \sum_{i=1}^n x_i \log_a \frac{x_i}{y_i}, \quad (3)$$

$$\left( \sum_{i=1}^n x_i = \sum_{i=1}^n y_i = 1 \right).$$

Therefore, the difference between two band images of a hyperspectral image cube can be measured using the K-L divergence equation (3).

Suppose a hyperspectral image cube  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L]^T \in R^{L \times N}$ , where  $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{iN}]$  and  $L, N$  represent the number of spectral bands and pixels of a 2D band image, respectively. After normalization of data in  $\mathbf{X}$ , the K-L divergence of the  $j^{\text{th}}$  band image relative to the  $i^{\text{th}}$  band image can be defined as follows:

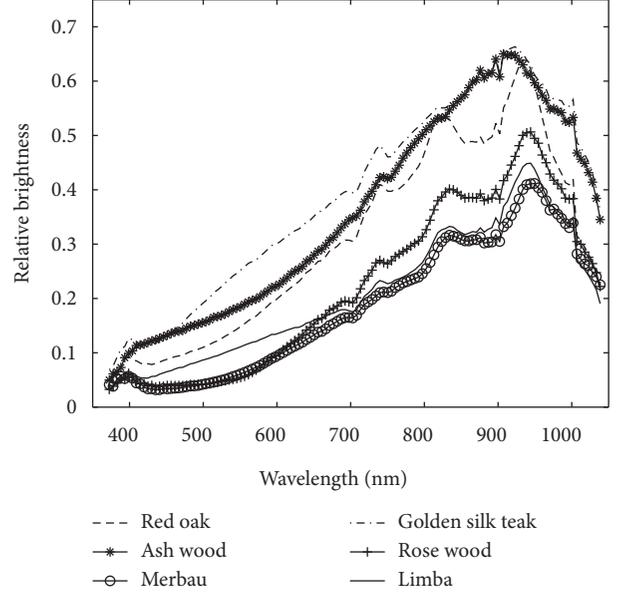


FIGURE 4: Spectral difference in wood vessels' peripheral region for six hardwood species.

$$D_{KL}(x_i||x_j) = \sum_{k=1}^N x_{ik} \log \frac{x_{ik}}{x_{jk}}. \quad (4)$$

In this way, we can calculate the K-L divergence of two band images to obtain a symmetric matrix  $\mathbf{D}$ . The  $i^{\text{th}}$  row of this matrix is defined as  $\mathbf{a}_i = [\mathbf{D}(i, 1) \ \dots \ \mathbf{D}(i, L)]$ , and the mean value of its  $L$  elements represents, to some extent, the information entropy of the  $i^{\text{th}}$  band image. Therefore, we may use the 10 spectral bands with the 10 largest entropies to form an image cube  $\tilde{\mathbf{X}} = [\mathbf{x}_{\lambda 1}, \mathbf{x}_{\lambda 2}, \dots, \mathbf{x}_{\lambda 10}]^T \in R^{10 \times N}$  (Figure 5(a)) and its complementary image cube  $\tilde{\mathbf{X}}$  (Figure 5(b)). This cube  $\tilde{\mathbf{X}}$  is classified into 4 parts by use of a K-means clustering algorithm so that the spatial pixels in the original wood hyperspectral image cube are classified into 4 parts, as illustrated in Figures 5(c)–5(f).

As shown in Figure 5, the white part in Figure 5(d) is the wood vessels' central region, which is segmented accurately. Figures 5(c), 5(e) and 5(f) indicate the wood vessels' peripheral and background regions. In fact, the wood vessels' peripheral region is clustered separately into different parts, as illustrated in Figures 5(c), 5(e) and 5(f). Therefore, this classification is not suitable for accurately extracting the wood vessels' peripheral region. In order to more accurately segment the vessels' central region, it is necessary to perform a morphological corrosion operation on the white sections in Figure 5(d).

The clustering process in Figure 5 has two problems. Firstly, selecting the ROI randomly in a hyperspectral image does not ensure location of the wood vessels' central region. Secondly, as can be seen from Figure 5, the K-means cluster results in 4 parts (i.e., Figures 5(c)–5(f)), but only Figure 5(d) represents the wood vessels' central region.

To solve these problems, we propose an automatic identification scheme for the wood vessels' central region. For a ROI  $\mathbf{W}$  of size  $S \times S$ , the binary images produced by

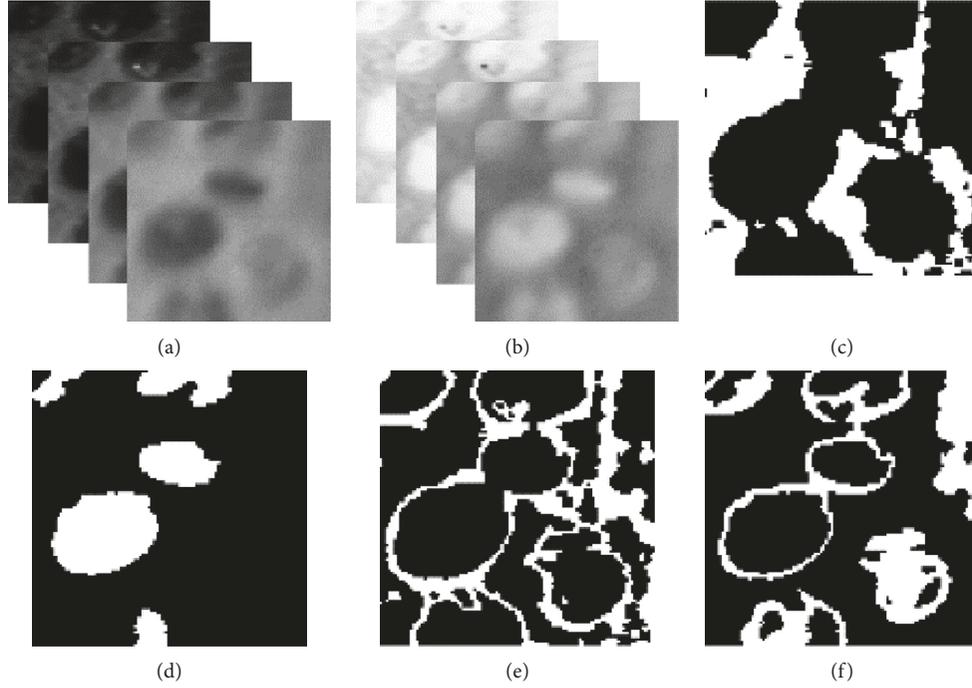


FIGURE 5: Spatial clustering results with spectral information and  $K$ -means algorithm: (a) image cube with selected 10 bands; (b) the complement image cube; (c-f) the classified 4 parts.

$K$ -means cluster are  $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \mathbf{B}_4$ . The selected 10 spectral bands obtained from K-L divergence are  $\lambda_1, \lambda_2, \dots, \lambda_{10}$ , and the corresponding binary images for these 10 band images are denoted as  $\mathbf{X}_n, n = 1, 2, 3, \dots, 10$ . Here  $C_p(\mathbf{X})$ ,  $p = 0, 1$  represents the occurrence number of  $p$  in matrix  $\mathbf{X}$ , and  $C(\mathbf{X}, \mathbf{W})$  represents the number of identical elements in the same position of matrices  $\mathbf{X}$  and  $\mathbf{W}$ . The 4 measurement parameters are defined as follows:

$$b1_k = \frac{(1/10)\sum_{i=1}^{10} C_0(\mathbf{X}_i)}{C_0(\mathbf{B}_k)}, \quad k \in \{1, 2, 3, 4\}, \quad (5)$$

$$b2_k = \frac{(1/10)\sum_{i=1}^{10} C_1(\mathbf{X}_i)}{C_1(\mathbf{B}_k)}, \quad k \in \{1, 2, 3, 4\}, \quad (6)$$

$$b3_k = \frac{(1/10)\sum_{i=1}^{10} C(\mathbf{X}_i, \mathbf{B}_k)}{S \times S}, \quad k \in \{1, 2, 3, 4\}, \quad (7)$$

$$b_k = \frac{1}{3} (\|b1_k\| + \|b2_k\| + \|b3_k\|). \quad (8)$$

Here  $\|x\|$  is defined as follows:

$$\|x\| = \begin{cases} x, & x \leq 1, \\ \frac{1}{x}, & x > 1. \end{cases} \quad (9)$$

We calculated  $b_k$  for the binary image  $\mathbf{B}_k, k \in \{1, 2, 3, 4\}$  with equation (8) and define  $g = \max b_k$  and a threshold  $T$ . When  $g \geq T$ , the corresponding binary image  $\mathbf{B}_k$  includes the wood vessels' central region. Using this approach, the ROI moves gradually from left to right and from top to bottom

with a fixed step length  $S$  to traverse the whole hyperspectral image. In this way, we locate all binary images  $\mathbf{K}_i, i = 1, 2, \dots, l$  and include the wood vessels' central region. These image matrices  $\mathbf{K}_i, i = 1, 2, \dots, l$  are connected by columns to obtain a matrix  $\mathbf{Z}$  with a dimension of  $(C_1 + \dots + C_l) \times 128$ , and the mean value of each column is computed to form a new vector  $\mathbf{z}$  of 128 dimensions. For a hardwood species with  $m$  hyperspectral image specimens, we derive multiple vectors  $\mathbf{z}_j, j = 1, 2, \dots, m$  and calculate the mean vector  $\bar{\mathbf{z}}$ . Figure 6 illustrates the mean spectral vector  $\bar{\mathbf{z}}$  for the six hardwood species in our studies of the wood vessels' central regions. The training samples are used to calculate  $\bar{\mathbf{z}}$  with a ROI size of  $100 \times 100$ .

### 3. Results and Discussion

**3.1. Spectral Difference Analysis for Wood Vessels' Peripheral Region, Central Region, and Randomly Selected ROI.** We may use spectral properties of a specific wood region or tissue to guide wood species identification using both interclass and intraclass differences. It is obvious that large interclass differences and small intraclass differences will produce more reliable results. These two differences can be represented by the spectral curve's interclass and intraclass distances. Therefore, here, we use the divergence matrix to indicate the spectral curve's interclass and intraclass differences. We define the intraclass divergence matrix for the  $j^{\text{th}}$  wood species as follows:

$$\mathbf{S}_j = \frac{1}{n_j} \sum_{k=1}^{n_j} (\mathbf{x}_k^j - \mathbf{m}_j)(\mathbf{x}_k^j - \mathbf{m}_j)^T. \quad (10)$$

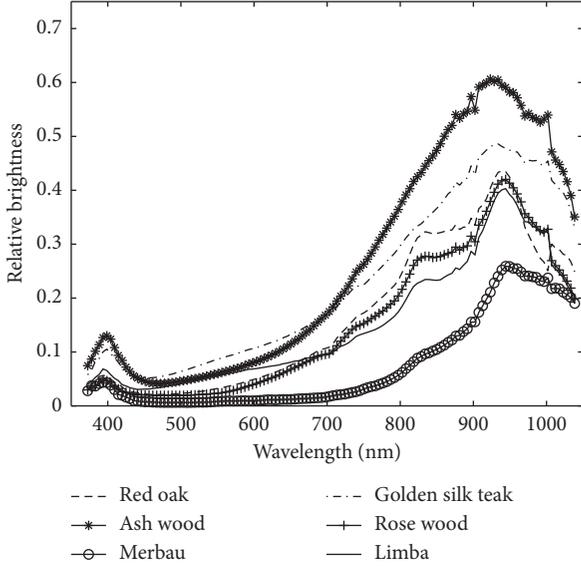


FIGURE 6: Spectral difference in wood vessels' central region for six hardwood species.

Here  $j = 1, 2, \dots, c$  ( $c = 6$ );  $\mathbf{X}_k^j$  represents the  $k^{\text{th}}$  sample for the  $j^{\text{th}}$  species,  $\mathbf{m}_j$  represents the mean value of all samples for the  $j^{\text{th}}$  species, and  $n_j$  represents the number of all samples for the  $j^{\text{th}}$  species. The overall intraclass divergence matrix can be defined as follows:

$$\mathbf{S}_w = \sum_{j=1}^c P_j \mathbf{S}_j, \quad (11)$$

where  $P_j$  represents the prior probability for the  $j^{\text{th}}$  wood species, and it can be estimated by  $P_j = n_j/n$ , where  $n$  represents the overall number of all species samples. Similarly, the interclass divergence matrix is defined as follows:

$$\mathbf{S}_b = \sum_{j=1}^c P_j (\mathbf{m}_j - \mathbf{m})(\mathbf{m}_j - \mathbf{m})^T, \quad (12)$$

where  $\mathbf{m}$  represents the mean value of all species samples. The overall divergence matrix is then defined as follows:

$$\mathbf{S}_t = \mathbf{S}_b + \mathbf{S}_w. \quad (13)$$

Finally, the pattern divisibility rule is defined as follows:

$$J = \frac{\text{tr}[\mathbf{S}_b]}{\text{tr}[\mathbf{S}_w]}, \quad (14)$$

where  $\text{tr}[\mathbf{X}]$  represents the trace of matrix  $\mathbf{X}$ . It is obvious that a higher  $J$  value will produce a more reliable identification.

In terms of a specific wood region or tissue, we considered the wood vessels' peripheral region, central region, and randomly selected ROI. Spectral extractions in wood vessels' peripheral and central regions have been discussed in Sections 2.4 and 2.5. The size of the randomly selected ROI was  $300 \times 300$ ; the mean vector of all spectral vectors in this ROI is the final spectral extraction for a hyperspectral cube sample.

The smoothing procedure for the spectral preprocessing and the SNV calibration procedure are usually required before spectral extraction and pattern recognition as they cancel the noise disturbance and waveform offset. This preprocessing process is illustrated in Figure 7.

Table 2 gives the  $J$  values for the wood vessels' peripheral region, central region, and randomly selected ROI in the training samples. It indicates that the spectral divisibility for the wood vessels' peripheral region was most useful, while that for the wood vessels' central region was least useful. Since the wood vessels' central region is hollow, with some resin and probably some filling material, its spectral curve is variable and thus has a large intraclass difference with a small  $J$  value.

Finally, the time required was compared for spectral extractions in the wood vessels' peripheral region, central region, and random ROI. The computer configuration used was as follows: CPU Intel Core I7-6700, internal memory 8G, hard disk 1T, display card AMD Radeon R7200 Series. In these three spectral extractions, one hyperspectral image was cut into 30 ROI of  $100 \times 100$ . Time required for processing one ROI is given in Table 3. We can see that the processing speed of random ROI was the fastest while that for the vessels' central region was the slowest.

**3.2. Comparisons of Identification Accuracy.** Before sending the extracted spectral vectors into classifiers, spectral dimensionality reduction was usually required. In this work, three algorithms, PCA, KPCA, and MDS, were used to reduce dimensionality. In spectral classification, three classifiers, BP neural networks, SVM, and MD, were used for identification comparisons.

Firstly, we used the MD classifier to calculate the MD between a test sample and the training sample sets of the 6 hardwood species. The usual nearest neighbor classification was used to judge the test sample's wood species. The classification results are illustrated in Figure 8. We can see that the classification correct ratio (i.e., CCR) was influenced by spectral dimensionality reduction. For instance, if MDS was used, the CCR stabilized only when the spectral dimensionality was larger than 7. On the other hand, if PCA or KPCA was used, the CCR stabilized only when the spectral dimensionality was larger than 3. Moreover, we can see that the CCR was also influenced by the wood region of the spectral collection. In fact, the CCR of the wood vessels' peripheral region was the highest, approaching 100%, while the CCR of the wood vessels' central region was the lowest, and its mean value was approximately 70%–80%. The CCR of random ROI fell between the peripheral and central vessel regions and was approximately 90%.

The classification accuracy was the calculated CCR, which was the overall classification ratio (%) for the 540 test set samples. This classification accuracy is illustrated in Figures 8–11 and Table 4.

Secondly, we used the BP neural network classifier. Since the threshold and weight for the network initialization were selected randomly, the classification accuracy of every trained network may fluctuate slightly. Figure 9 illustrates

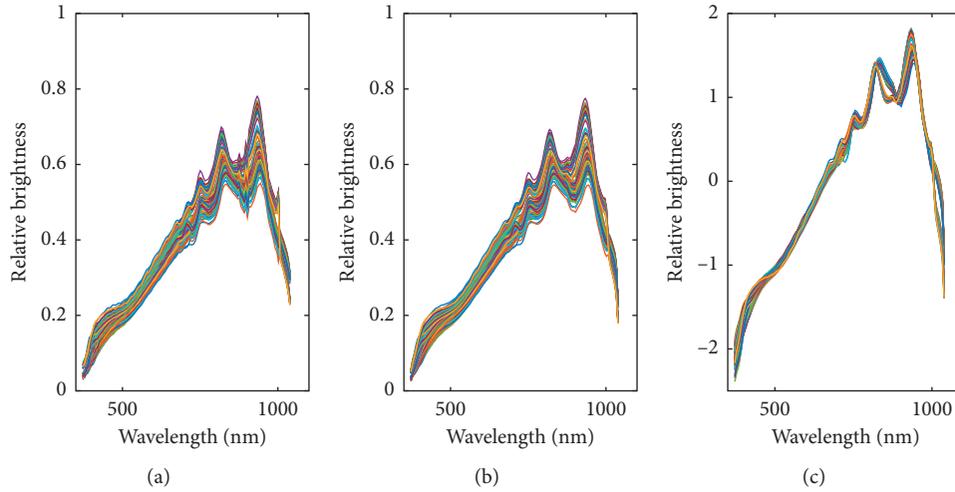


FIGURE 7: Spectral preprocessing procedure: (a) original spectral curve; (b) smoothed spectral curve; (c) spectral curve with SNV processing.

TABLE 2:  $J$  values for different regions of the vessels.

Item	Vessels' central region	Vessels' peripheral region	Random ROI
$J$	1.6506	9.0970	6.6387

TABLE 3: Time requirement comparisons for 3 spectral extractions.

Method	Process	Computing time (s)
Vessels' peripheral region	$G(x, y)$	0.6095
	$D(x, y)$	0.0056
	$E(x, y)$	0.0008
	Total	0.6199
Vessels' central region	$D$	0.7714
	$K$ -means cluster	0.2252
	Central region selection	0.0087
Random ROI	Total	1.0181
	Total	0.0874

the test CCR for the 25 trained networks, while the spectral dimensionality was reduced to 5D by PCA. We observed that the overall CCR was stable, though there were slight differences between each CCR across the 25 networks. Therefore, the mean CCR for 25 networks was calculated and used as the final CCR. Figure 10 illustrates the mean CCR for different spectral dimensionality reductions and different wood regions of spectral collection. We can see that CCRs from vessel central regions were lowest, with mean values of approximately 70%–80%, while those from peripheral regions and random ROI were almost identical, approaching 100%.

Thirdly, the SVM classifier was used. The predictive power of SVM depends on the kernel function used. The radial basis kernel function was usually used. When determining the penalty factor  $C$  and  $\gamma$  value, the grid search method can be used to find the best parameters. Figure 11 illustrates the identification comparisons with the best

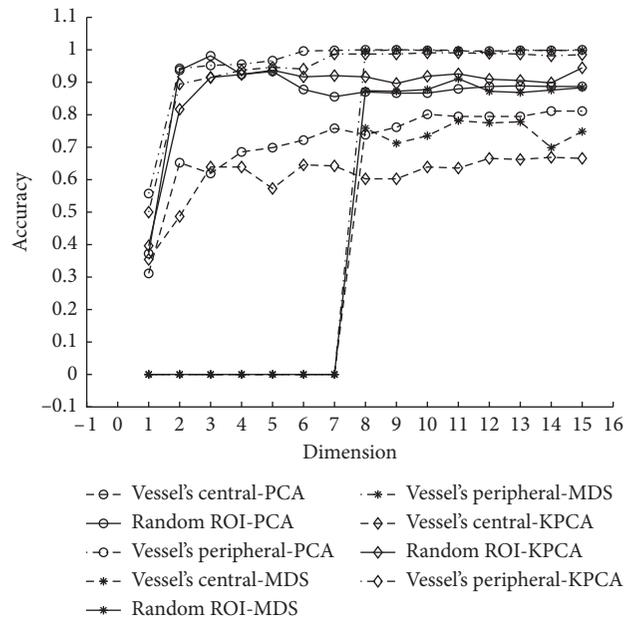


FIGURE 8: Spectral classification comparisons with MD classifier.

parameters  $C$  and  $\gamma$ . We can see again that the CCR of central regions was lowest, with a mean of approximately 80%, while those of peripheral regions and ROI were almost identical, approaching 100%. The CCR of random ROI decreased with the application of KPCA.

The best identifications with different classifiers and spectral dimensionality reductions are given in Table 4.

Figures 8, 10, and 11 and Table 4 indicated that the CCR of vessels' central regions was the lowest in all cases and that the CCR of vessels' peripheral regions was slightly higher than that of random ROI in the case of BP network and SVM. In the case of MD classifier, the CCR of vessels' peripheral regions was much higher than that of random ROI. Moreover, we obtained good CCR of peripheral regions with a low spectral dimension. For example, the best CCR (100% of vessels' peripheral region) was achieved using PCA and

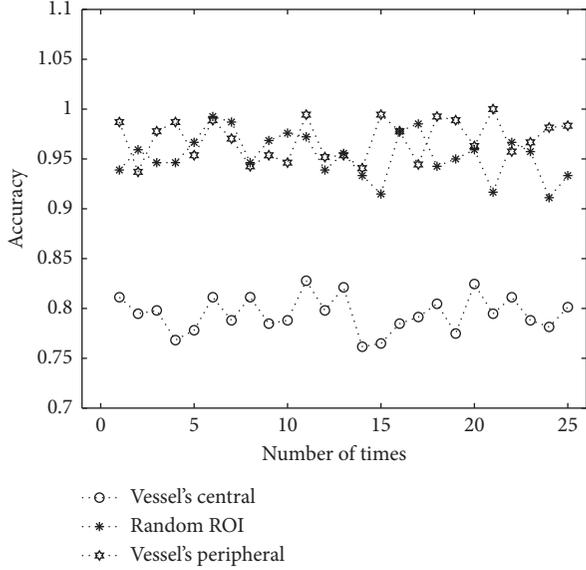


FIGURE 9: Identification accuracy of BP network for 25 training times (i.e., 25 networks).

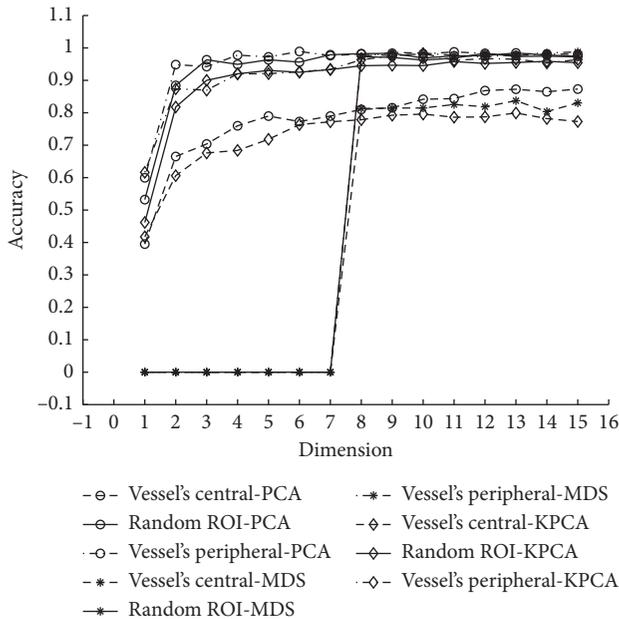


FIGURE 10: Spectral classification comparisons with BP neural network classifier.

MD classifier with a spectral dimension of 8 or by using MDS and MD classifier with a spectral dimension of 9.

Concerning the spectral dimensionality reductions, good results were obtained using PCA and KPCA with a low spectral dimension. However, if MDS was used, the CCR stabilized only when the spectral dimensionality was larger than 7. In general, PCA was the best choice in terms of identification accuracy and spectral dimension.

**3.3. Comparisons of Identification Robustness.** Wood hyperspectral image acquisition may be influenced by external environmental factors such as temperature, humidity,

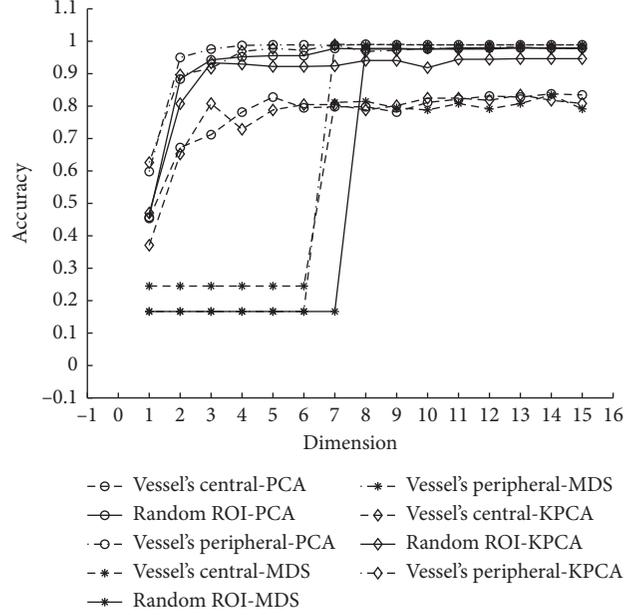


FIGURE 11: Spectral classification comparisons with SVM classifier.

and illumination. In this section, a simulation experiment on noisy wood hyperspectral image classification was performed. Since the training set of wood images can be acquired in advance in a controlled environment, we only added white noises to the test set. The signal-to-noise ratio (SNR) was used to measure the test set's image quality.

The robustness of our three spectral extraction schemes (i.e. vessels' central, vessels' peripheral, random ROI) was compared using an error rate defined in terms of SNR.

$$\text{Error rate} = (\text{CCR}_{\text{ideal}} - \text{CCR}_{\text{noisy}}), \quad (15)$$

where  $\text{CCR}_{\text{ideal}}$  and  $\text{CCR}_{\text{noisy}}$  refer to the classification accuracy for ideal test images and noisy test images, respectively. The falling trend of "error rate" (Figure 12) indicates that the faster the curve falls, the more robust the corresponding spectral extraction scheme. Therefore, we conclude that in the case of BP and MD classifiers, the vessels' peripheral extraction scheme was the most robust, while for SVM, the vessels' peripheral scheme was slightly more robust than random ROI and much more robust than the vessels' central scheme.

**3.4. Discussions of Parameters.** In the wood vessels' peripheral region extraction, the two key parameters are the size  $r$  of structure element  $A$  and the threshold  $k$  in equation (2). Here we explain the effect of these two parameters and give the best values of  $r$  and  $k$  for each species.

Figure 13 illustrates the wood vessels' peripheral region extraction comparisons in a ROI with  $r = 30$  and different  $k$  for wood species *Intsia bijuga*. The value interval of  $k$  is  $[2, 47]$  with a step length of 5. We can see that a small  $k$  will make the wood rays or axial parenchyma remain in vessels' peripheral region, while a large  $k$  will make this region contract.

Figure 14 illustrates the extraction comparisons with  $k = 43$  and different  $r$  values. The value interval of  $r$  is  $[1, 46]$  with

TABLE 4: Best classification results with different classifiers and spectral dimensionality reductions.

Classifiers	Dimensionality reductions Spectral region	PCA		MDS		KPCA	
		CCR (%)	Spectral dimension	CCR (%)	Spectral dimension	CCR (%)	Spectral dimension
BP	Central vessel region	81.78	13	80.42	13	79.4	13
	Random ROI	97.80	9	97.57	9	94.25	9
	Peripheral vessel region	98.67	7	98.14	8	97.65	9
MD	Central vessel region	81.29	14	78.14	11	66.48	12
	Random ROI	88.88	13	91.11	11	94.44	15
	Peripheral vessel region	100	8	100	9	98.70	7
SVM	Central vessel region	83.70	15	83.33	14	83.33	13
	Random ROI	97.78	7	97.78	8	94.62	12
	Peripheral vessel region	99.07	8	99.07	7	98.88	7

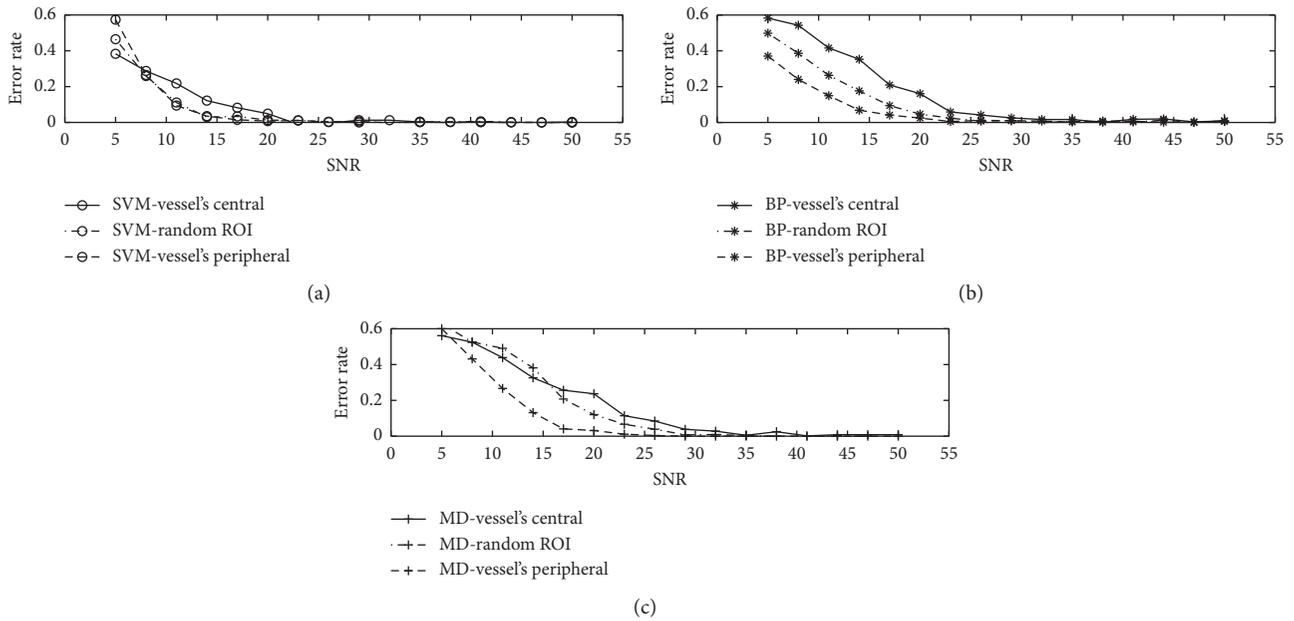
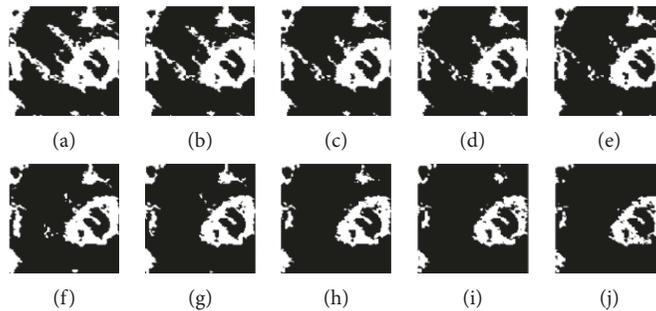


FIGURE 12: Different SNR with error rates: (a) SVM; (b) BP; (c) MD.

FIGURE 13: Wood vessels' peripheral region extraction with different  $k$  values. (a)  $k = 2$ , (b)  $k = 7$ , (c)  $k = 12$ , (d)  $k = 17$ , (e)  $k = 22$ , (f)  $k = 27$ , (g)  $k = 32$ , (h)  $k = 37$ , (i)  $k = 42$ , and (j)  $k = 47$ .

a step length of 5. We can see that a small  $r$  will make this region disappear and make wood rays appear, while large  $r$  will make this region contract again.

In summary, we conclude that small  $r$  and  $k$  produce more wood rays or axial parenchyma, while large  $r$  and  $k$  make these tissues disappear while causing vessel peripheral regions to contract. Table 5 gives the best values of  $r$  and  $k$  for

our 6 experimental species, and the corresponding extraction results are illustrated in Figure 15.

**3.5. Incremental Wood Species Identification.** To test the incremental identification performance of our scheme, two new wood species were added (i.e., Manchurian Ash

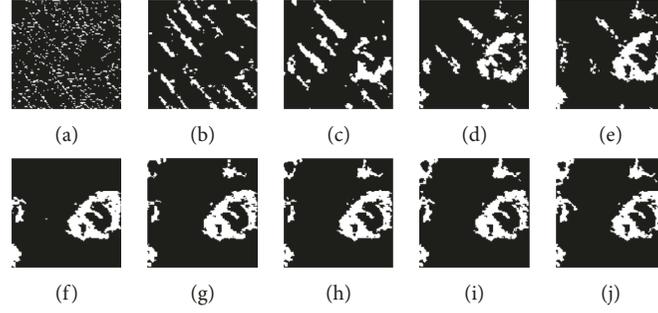


FIGURE 14: Wood vessels' peripheral region extraction with different  $r$  values. (a)  $r = 1$ , (b)  $r = 6$ , (c)  $r = 11$ , (d)  $r = 16$ , (e)  $r = 21$ , (f)  $r = 26$ , (g)  $r = 31$ , (h)  $r = 36$ , (i)  $r = 41$ , and (j)  $r = 46$ .

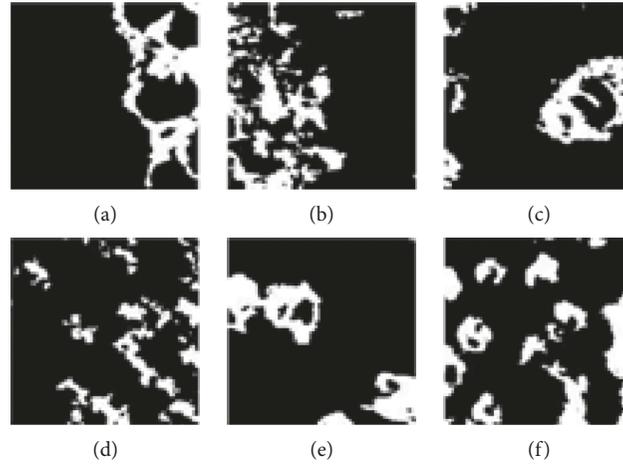


FIGURE 15: Vessel peripheral region extraction with best  $r$  and  $k$  for 6 species. (a) Red oak. (b) Ash wood. (c) Merbau. (d) Golden silk teak. (e) Rose wood. (f) Limba.

TABLE 5: Best values for parameters  $r$  and  $k$ .

Tree	$K$	$r$
Red oak	45	26
Ash wood	30	35
Merbau	43	30
Golden silk teak	45	30
Rose wood	50	45
Limba	20	30

(*Fraxinus mandshurica* Rupr.) and Chinese Oak (*Quercus mongolica* Fisch.)). It was usually difficult to discriminate these two wood species since they have similar external appearance. The identification results with PCA and three classifiers, BP neural network, SVM, and MD, are illustrated in Figure 16. Again, we can see that the CCR of vessels' central regions was the lowest in all cases, while the CCR of vessels' peripheral regions was higher than that of random ROI. In fact, the CCR of vessels' peripheral regions approaches 90% approximately. Therefore, we conclude that the CCR of vessels' peripheral regions was stable and best when new wood species samples were added.

3.6. *Comparisons with Other Works.* Finally, the proposed scheme is compared with the two algorithms described in references [15, 16]. Hyperspectral images must be converted

to grayscale images before comparison. To this end, two methods can be applied. The first method consists of selecting the waveband of the hyperspectral images using the procedure described in Section 2.5, then performing wavelet-based image fusion on selected images, and finally saving grayscale images after image fusion. This method is detailed below.

The retained images after waveband selection were defined as  $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_j, \dots, \mathbf{X}_m\}$ ,  $m \in \{1, 2, 3, \dots, 10\}$ , where the image size of  $\mathbf{X}_j$  was  $M \times N$ . Given that each image was derived from a hyperspectral image, image registration of the above images was not necessary. Each image was processed by a wavelet transform based on Haar wavelets. The number of layers for decomposition was 4. The low-frequency wavelet coefficients  $\{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_j, \dots, \mathbf{A}_m\}$  and high-frequency wavelet coefficients  $\{\mathbf{H}_1^k, \mathbf{H}_2^k, \dots, \mathbf{H}_j^k, \dots, \mathbf{H}_m^k\}$ ,  $k \in \{1, 2, 3\}$  of each image were then obtained. The high-frequency wavelet coefficients include horizontal, vertical, and diagonal high-frequency coefficients. The new wavelet coefficients  $\mathbf{A}'$  and  $\mathbf{H}'$  were then calculated by fusion as

$$\begin{cases} \mathbf{A} = \frac{1}{m} (\mathbf{A}_1 + \mathbf{A}_2 + \dots + \mathbf{A}_j + \dots + \mathbf{A}_m), \\ \mathbf{H}^{k'} = \mathbf{H}_j^k, \text{ s.t. } |\mathbf{H}_j^k| = \max(|\mathbf{H}_1^k|, |\mathbf{H}_2^k|, \dots, |\mathbf{H}_j^k|, \dots, |\mathbf{H}_m^k|). \end{cases} \quad (16)$$

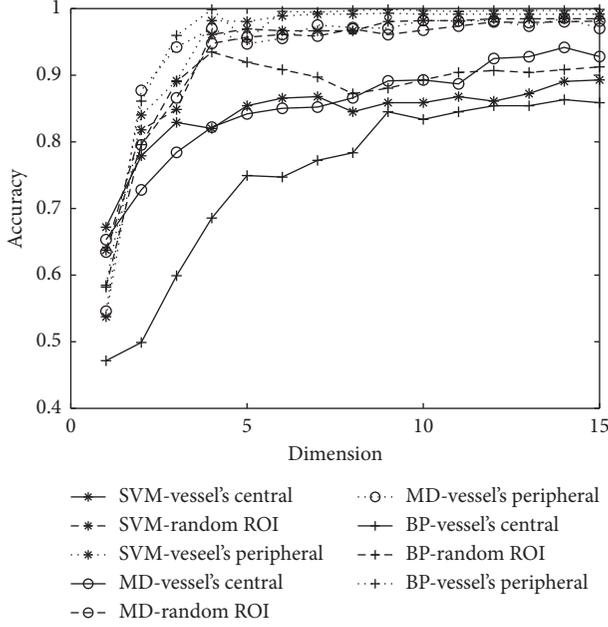


FIGURE 16: Spectral classification comparisons for 8 wood species.

These new wavelet coefficients were then used for inverse wavelet transform to obtain the fused digital image  $X^f$ . The results of image fusion are presented in Figure 17. Note that the fused image includes the texture features of all of the waveband images.

The second aforementioned method consists in applying principal component analysis (PCA) to reduce the dimension of hyperspectral images. The first four principal components are presented in Figure 18, but we only retained the first principal component with clear texture.

The method described in reference [15] is a traditional approach to image texture classification applied for the identification of wood species based on GLCM (Gray-Level Co-Occurrence Matrix), which will not be described here. The method described in reference [16] also aims to image texture classification for wood species but based now on IBGLAM (Improved Basic Gray Level Aura Matrix). The corresponding specific steps are described below.

According to the descriptions in reference [16], an image needs to be first compressed into a 4-bit grayscale image, which is processed with equation (17).  $F'(x, y)$  refers to an 8-bit grayscale image; floor represents downward rounding;  $F(x, y)$  refers to a 4-bit grayscale image. The result of grayscale image compression is shown in Figure 19. Note that the image textures are still clear despite the compression of the pixel values of the grayscale image.

$$F(x, y) = \text{floor}\left(\left(\frac{F'(x, y)}{\max(F'(x, y)) + 1}\right) * 16\right). \quad (17)$$

Then, the feature matrix  $\mathbf{B}$  was calculated with the expressions in equation (19). The grayscale image was defined as  $F(i, j)$ ,  $i \in \{2, 3, \dots, M-1\}$ ,  $j \in \{2, 3, \dots, N-1\}$ , where  $M$  and  $N$  represent the length and width of the grayscale image, respectively. To make the subscripts of  $\mathbf{B}$  valid, it is

necessary to add 1 to each pixel value of  $F(i, j)$  before calculation. In this way, we avoid the occurrence of 0, with  $F(i, j) \in \{1, 2, \dots, 16\}$ . The matrix  $\mathbf{B}$  calculated with equation (17) is a  $16 \times 16$  square matrix.

$$\begin{cases} \mathbf{B}(F(i, j), F(i-1, j)) = \mathbf{B}(F(i, j), F(i-1, j)) + 1, \\ \mathbf{B}(F(i, j), F(i-1, j+1)) = \mathbf{B}(F(i, j), F(i-1, j+1)) + 1, \\ \mathbf{B}(F(i, j), F(i, j+1)) = \mathbf{B}(F(i, j), F(i, j+1)) + 1, \\ \mathbf{B}(F(i, j), F(i+1, j+1)) = \mathbf{B}(F(i, j), F(i+1, j+1)) + 1. \end{cases} \quad (18)$$

To reduce the dimension of the feature vector, the square matrix  $\mathbf{B}$  was calculated with the following expression:

$$\mathbf{B}' = (i, j) = \begin{cases} \mathbf{B}(i, j) + \mathbf{B}(j, i), & \text{if } (i \neq j), \\ \mathbf{B}(i, j), & \text{if } (i = j). \end{cases} \quad (19)$$

Finally, the square matrix  $\mathbf{B}'$  was converted according to equation (19), where  $\cup$  represents the elements in series. The converted vector  $\mathbf{x}$  was the feature vector with a length of 136. Figure 20 shows the schematic diagram of features. Table 6 shows the processing time for a hyperspectral image when applying the above method.

$$\mathbf{x} = \cup \mathbf{B}'(i, j), \quad i \in \{1, 2, \dots, 16\}, \quad j \in \{i, i+1, \dots, 16\}. \quad (20)$$

Figure 21 shows the CCR of the vessel peripheral spectral extraction method proposed in this paper in comparison with the IBGLAM and GLCM methods based on support vector machines (SVM) for the identification of eight tree species, including the six tree species in Table 1 and the two tree species in Section 3.5.

It can be seen from Figure 21 that the CCR of the two methods described in references [15, 16] were only around 80% for the samples used in this paper, whereas the CCR of our method proposed in this paper was much higher than that of the other two methods. In addition, when PCA was used for generating grayscale images, the accuracy of image classification with GLCM and IBGLAM was only 70.93% and 73.70%, respectively.

## 4. Conclusions

We have proposed a novel hardwood species identification scheme using hardwood microscopic hyperspectral images. The spectral feature of wood vessels is especially considered to investigate the possibility of tree identification to the species level. Specifically, the spectral features of vessel peripheral regions are acquired using morphological operations and those of central regions by use of K-L divergence and an automatic identification scheme. The spectral divisibility rules in Section 3.1 and experimental comparisons in Section 3.2 indicate uniformly that the CCR of peripheral regions is best in all cases of spectral dimensionality reduction. Moreover, we can also conclude that wood vessels' peripheral extraction scheme is robust when wood images are disturbed by white noise.

Our work attempted to identify hardwood trees to species level using spectral analysis. Considering the

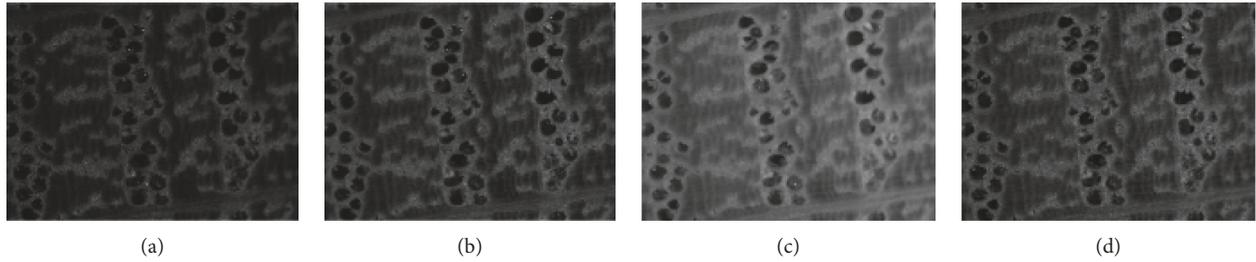


FIGURE 17: Schematic of image fusion with wavelet transform. (a) Band  $X_1$ , (b) Band  $X_2$ , (c) Band  $X_3$ , and (d) fused image, taking a three-waveband image as an example.

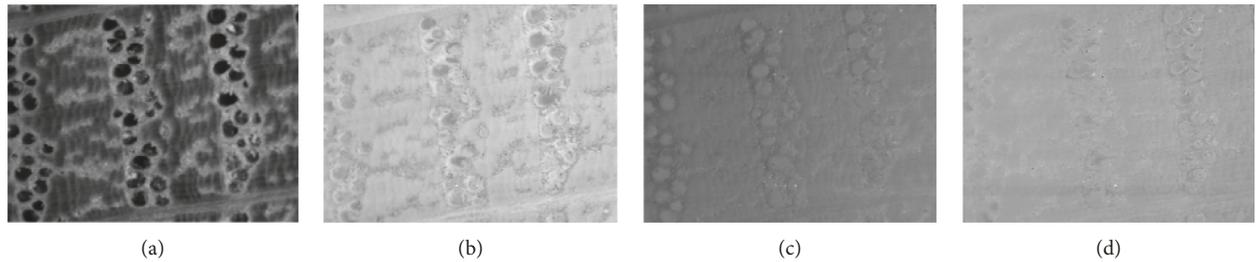


FIGURE 18: Results of dimensionality reduction from PCA ((a) PC1, (b) PC2, (c) PC3, and (d) PC4).

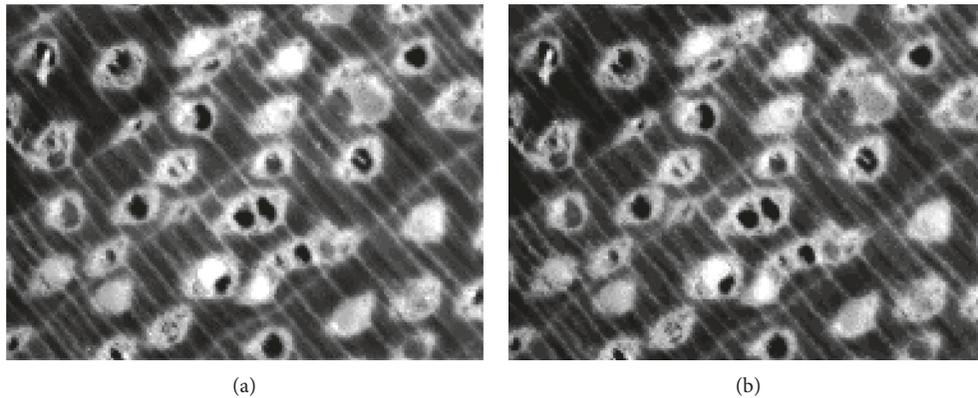


FIGURE 19: Result of compression of pixel values ((a) original 8-bit image; (b) compressed 4-bit image).

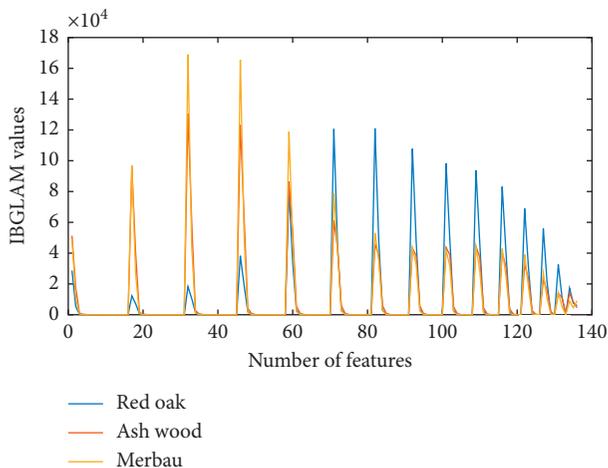


FIGURE 20: IBGLAM feature vector.

TABLE 6: Processing time when applying the methods described in references [15, 16].

Spectral dimension reduction	Texture feature computation		
	PCA	IBGLAM	GLCM
K-L divergence			
42.9767 (s)	0.6414 (s)	0.1698 (s)	0.0329 (s)

instrument's cost, we chose the visible/near infrared SOC710VP hyperspectral imager with 128 spectral bands and a spectral range of 372–1038 nm. Spectral data in the visible region may vary due to changes in wood color produced by environmental variation. Therefore, we kept the environment constant during spectral work and wood sample preservation. In addition, 780–900 nm is the near infrared band, which best characterizes different wood properties.

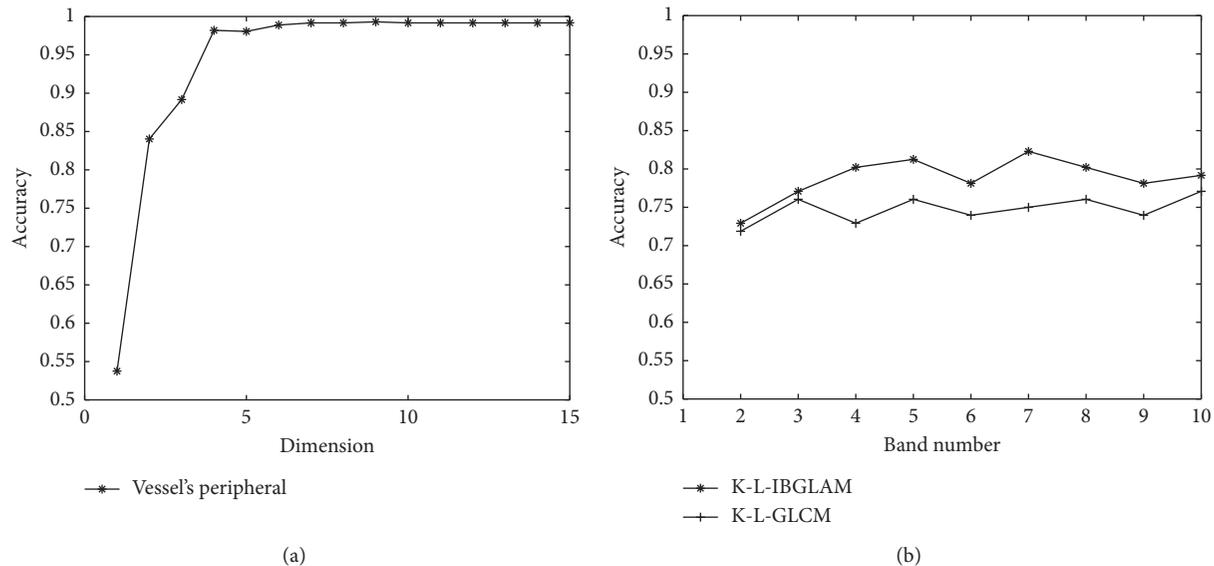


FIGURE 21: Comparison of proposed method vs. other methods: (a) vessel peripheral spectral extraction method; (b) IBGLAM and GLCM methods.

The method described in reference [15] is suitable for wood species with large different textures, but the highest CCR was only 75.92% when using the samples provided in this paper. Zamri et al. [16] classified 52 tropical tree species and obtained the highest CCR of 97.01%. However, the main research objects of reference [16] were tropical tree species. These species tend to have large evenly distributed vessels, and most of them are diffuse-porous species that often have clear texture features. The highest CCR was only 82.03% with the method of reference [16] using wood samples provided in this paper. The research objects in this paper include not only tropical tree species, but also subtropical tree species. Moreover, the vessels of these subtropical tree species present different sizes and they are not evenly distributed. Furthermore, the method proposed in this paper does not require rich textures in the cross sections of wood species. As long as the vessels of tree species are clear, the tree species can be classified according to the spectral information of structures around the vessels.

### Data Availability

Access to research data is restricted. The experimental instruments and wood species samples are owned by other academic institutions; therefore, we cannot publicly release the research spectral wood data.

### Conflicts of Interest

The authors declare that there are no conflicts of interest.

### Acknowledgments

This research was supported by the National Natural Science Foundation of China (grant no. 31670717), the Fundamental Research Funds for the Central University (grant no.

2572017EB09), and the Heilongjiang Province Natural Science Foundation (grant no. C2016011).

### References

- [1] S. Pan and M. Kudo, "Segmentation of pores in wood microscopic images based on mathematical morphology with a variable structuring element," *Computers and Electronics in Agriculture*, vol. 75, no. 2, pp. 250–260, 2011.
- [2] R. Yusof, M. Khalid, and A. S. Khairuddin, "Application of kernel-genetic algorithm as nonlinear feature selection in tropical wood species recognition system," *Computers and Electronics in Agriculture*, vol. 93, no. 1, pp. 68–77, 2013.
- [3] F. Ruffinatto, A. Crivellaro, and A. C. Wiedenhoeft, "Review of macroscopic features for hardwood and softwood identification and a proposal for a new character list," *IAWA Journal*, vol. 36, no. 2, pp. 208–241, 2015.
- [4] U. Khairuddin, R. Yusof, M. Khalid et al., "Optimized feature selection for improved tropical wood species recognition system," *ICIC Express Letters, Part B: Applications*, vol. 2, no. 2, pp. 441–446, 2011.
- [5] J. F. Hu, W. L. Song, W. Zhang, Y. Zhao, and A. Yilmaz, "Deep learning for use in lumber classification tasks," *Wood Science and Technology*, vol. 53, no. 2, pp. 505–517, 2019.
- [6] B. K. Lavine, C. E. Davidson, A. J. Moores, and P. R. Griffiths, "Raman spectroscopy and genetic algorithms for the classification of wood types," *Applied Spectroscopy*, vol. 55, no. 8, pp. 960–966, 2001.
- [7] A. K. Moore and N. L. Owen, "Infrared spectroscopic studies of solid wood," *Applied Spectroscopy Reviews*, vol. 36, no. 1, pp. 65–86, 2001.
- [8] V. Piuri and F. Scotti, "Design of an automatic wood types classification system by using fluorescence spectra," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 3, pp. 358–366, 2010.
- [9] X. Fu and J. Chen, "A review of hyperspectral imaging for chicken meat safety and quality evaluation: application, hardware, and software," *Comprehensive Reviews in Food Science and Food Safety*, vol. 18, no. 2, pp. 535–547, 2019.

- [10] S. T. Monteiro, Y. Minekawa, Y. Kosugi, T. Akazawa, and K. Oda, "Prediction of sweetness and amino acid content in soybean crops from hyperspectral imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 1, pp. 2–12, 2007.
- [11] P. K. Goel, S. O. Prasher, R. M. Patel, J. A. Landry, R. B. Bonnell, and A. A. Viau, "Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn," *Computers and Electronics in Agriculture*, vol. 39, no. 2, pp. 67–93, 2003.
- [12] I. Ibrahim, A. S. M. Khairuddin, M. S. Abu Talip, H. Arof, and R. Yusof, "Tree species recognition system based on macroscopic image analysis," *Wood Science and Technology*, vol. 51, no. 2, pp. 431–444, 2017.
- [13] M. Khalid, L. Eileen, R. Yusof et al., "Design of an intelligent wood species recognition system," *International Journal of Simulation: Systems, Science & Technology*, vol. 9, no. 3, pp. 9–19, 2008.
- [14] H. Q. Wei, L. Chen, and L. K. L. Guo, "KL Divergence-based fuzzy cluster ensemble for image segmentation," *Entropy*, vol. 20, no. 4, p. 273, 2018.
- [15] R. Bremananth, B. Nithya, and R. Saipriya, "Wood species recognition using GLCM and correlation," in *Proceedings of the International Conference on Advances in Recent Technologies in Communication & Computing (IEEE)*, pp. 615–619, Las Vegas, NV, USA, July 2009.
- [16] M. I. P. Zamri, F. Cordova, A. S. M. Khairuddin et al., "Tree species classification based on image analysis using Improved-Basic Gray Level Aura Matrix," *Computers and Electronics in Agriculture*, vol. 124, pp. 227–233, 2016.



**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

