

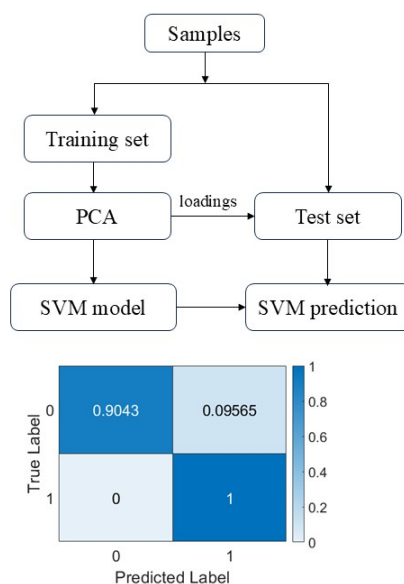
Identification of Genuine and Artificial Wood Grain Based on PCA-SVM

Jiaxin Wang,^a Liangze Pan,^{a,*} Zhen Wang,^b and Shangzhong Jin^{a,*}

* Corresponding authors: lzpan@cjlu.edu.cn, jinsz@cjlu.edu.cn

DOI: 10.15376/biores.21.3.6253-6266

GRAPHICAL ABSTRACT



Identification of Genuine and Artificial Wood Grain Based on PCA-SVM

Jiixin Wang,^a Liangze Pan,^{a,*} Zhen Wang,^b and Shangzhong Jin^{a,*}

This study presents an integrated method combining Principal Component Analysis (PCA) and Support Vector Machine (SVM) to distinguish genuine and artificial wood grains. The approach is based on acquiring nine-dimensional gloss data measured under varying angles and texture orientations, along with two surface roughness parameters: the arithmetic mean height (S_a) and maximum peak height (S_z). The dataset was divided into training and testing sets with a ratio of 7:3 after standardized. PCA was applied to the training set to extract the top k principal components. Then, these components served as input features for training an SVM classifier, whose discriminative performance was evaluated on the test set. Experimental results indicated that the proposed method achieved an accuracy of 96.76%, an F1-score of 0.9761, and a Matthews correlation coefficient (MCC) of 0.9285, substantially outperforming comparative models including standalone SVM, Logistic Regression (LR), Partial Least Squares (PLS), and Principal Component Regression (PCR). The method demonstrated high efficiency and robustness in distinguishing wood grain types, suggesting strong potential in practical engineering applications such as quality control and material authentication.

DOI: 10.15376/biores.21.3.6253-6266

Keywords: Glossiness; Roughness; Principal component analysis (PCA); Support vector machine (SVM)

Contact information: *a*: College of Optical and Electronic Technology, China Jiliang University, Hangzhou, China; *b*: Decotec Research Institute, Zhejiang Decotec Co., Ltd., Hangzhou, China;

* Corresponding authors: lzpan@cjlu.edu.cn, jinsz@cjlu.edu.cn

INTRODUCTION

With the increasing demand for both performance and aesthetics in decorative materials for modern home furnishing, architecture, and automotive interiors, wood grain surfaces have gained wide popularity due to their natural texture and aesthetic appeal (Zhang 2006; State Forestry Administration 2024). At the same time, various imitation wood grain products, such as wood grain films and wood grain plastic panels, are rapidly advancing toward high-end and personalized applications (Peng and Zhang 2016; Peng *et al.* 2021; Wang *et al.* 2022), making their appearance increasingly difficult to distinguish from that of real wood. Reliable and efficient identification of genuine versus artificial wood grain is therefore of great significance not only for wood quality supervision and market anti-counterfeiting, but also for providing scientific guidance to the development of imitation wood grain technologies, helping them more closely approximate the texture and expressiveness of natural wood.

Existing research has mainly focused on wood species identification and the classification of decorative paper surfaces. Methods employed in wood species identification include conventional wood identification, database-based retrieval, wood image recognition techniques, near-infrared spectroscopy, gas chromatography–mass

spectrometry (GC–MS), and DNA-based methods (Cheng *et al.* 2021; Liu *et al.* 2023). For decorative paper surface classification, *i.e.*, classification of imitation wood grain films, Li *et al.* (2018) quantified the surface chromatic and gloss parameters of decorative paper for wood-based panels. They employed these feature parameters in combination with a back-propagation neural network (BP neural network) to model and classify decorative paper, achieving an accuracy of 92.9%. Similarly, Zhang *et al.* (2015) applied visible spectroscopy combined with principal component analysis (PCA) to conduct pattern recognition and classification of different types of decorative paper, achieving a classification accuracy of 94% for the surface visual characteristics of various decorative papers.

Although significant progress has been made in wood species identification and decorative paper surface classification (Li 2019; Ye *et al.* 2019; China Industry Information Network 2020), there is still no systematic approach specifically designed for distinguishing between genuine and imitation wood grain. To address this issue, this study proposes a PCA–SVM-based method for the discrimination of genuine versus artificial wood grain. It is worth noting that this study focuses on distinguishing decorative natural wood surfaces from artificial wood-grain films, which are the two primary surface decorative materials in the targeted industrial application; therefore, substrate materials such as lacquered MDF and composite boards fall outside the classification scope and were not included in the experiments. The method first collects nine-dimensional glossiness data under three measurement angles and different texture orientations, along with two surface roughness parameters: arithmetic mean height (S_a) and peak height (S_z). The acquired data are standardized and then divided into training and testing sets at a ratio of 7:3. Subsequently, Principal Component Analysis (PCA) is applied to the training set to extract the top k principal components that yield the highest classification accuracy, thereby eliminating redundant information and suppressing noise. These principal components are then used as inputs to train a Support Vector Machine (SVM) classifier, whose performance is evaluated on the testing set. Since both the feature extraction and classification stages of this method explicitly account for dimensionality reduction and noise suppression, and because the dataset covers a wide variety of genuine and imitation wood grain samples, the approach demonstrates high efficiency, strong discriminative power, and wide applicability. Experimental validation confirms the feasibility of the proposed method, with results showing that it achieves an accuracy of 96.76%, an F1 score of 0.9761, and an MCC of 0.9285 on the testing set.

Method

The fundamental principle of the PCA-SVM-based genuine and artificial wood grain discrimination method proposed in this paper is illustrated in Fig. 1. First, n measurement points are selected on the sample. At each point, triaxial gloss data is collected from both genuine and artificial wood grain surfaces at different texture orientations, denoted as X_{ij} , where $i = 1, 2, 3$ represents the gloss measurement angles (20° , 60° , 85°), and $j = 1, 2, 3$ denotes the orientation of the gloss meter relative to the wood grain (parallel, 45° , perpendicular). Simultaneously, the arithmetic mean height S_a and peak height S_z of the sample surface roughness are obtained, collectively forming an 11-dimensional feature vector X . Second, the sample set is divided into training and test sets at a 7:3 ratio. Vector X undergoes normalization to eliminate dimensionality differences and scale effects across wood species. Principal Component Analysis (PCA) is then applied to the training set to linearly transform the feature vector into a set of mutually orthogonal

principal component vectors P_k . The first k principal components retain the maximum variance contribution, representing the reduced-dimensional feature subspace. Finally, P_k is used as input to train a Support Vector Machine (SVM) classifier. This approach not only filters out redundant and noisy features, but it also constructs a robust decision boundary in the low-dimensional space, ultimately enabling efficient differentiation between genuine and artificial wood grain.

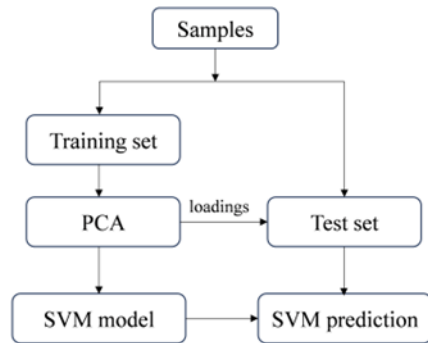


Fig. 1. Workflow of PCA-SVM

The PCA algorithm process follows these steps:

1. Construct the gloss-roughness matrix $X \in \mathbb{R}^{n \times 11}$ from the 9-dimensional gloss data and 2-dimensional surface roughness data obtained through n measurements.

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix} \quad (1)$$

2. The measured gloss-roughness matrix X undergoes zero-centering (Jiang *et al.* 2019; Cortes *et al.* 1995) *i.e.*, each column is subtracted by its corresponding column mean to obtain a zero-mean matrix $\bar{X} = X - \mu^T$.

- 3: Compute the covariance matrix C between each dimension of the zero-mean matrix $\bar{X}$, where $C = \frac{1}{n} \bar{X} \bar{X}^T$.

4. The next step is to perform an eigenvalue decomposition on the covariance matrix to extract the eigenvalues $\{\lambda_i\}$ and the eigenvectors $\{u_i\}$.

5. One sorts the eigenvalues in descending order and select the largest k eigenvalues. Then, use the corresponding k eigenvectors as row vectors to form the eigenvector matrix $P = [u_1, u_2, \dots, u_k] \in \mathbb{R}^{n \times k}$. Then one projects the original data onto the low-dimensional space constructed by these k eigenvectors, resulting in $Y = P^T \bar{X} \in \mathbb{R}^{k \times 11}$.

The resulting Y matrix represents the new feature matrix after dimensionality reduction to k dimensions, serving as input for subsequent support vector machine training.

6. To quantitatively assess the representativeness of retained principal components for the original data, explained variance and cumulative explained variance can be employed. The explained variance is defined as the proportion of the eigenvalue corresponding to the current principal component relative to the total sum of all eigenvalues, as shown in Eq. 2.

The cumulative explained variance is the sum of the contribution rates of the first k principal components, as shown in Eq. 3.

$$E_i = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j} \times 100\% \quad (2)$$

$$CE_k = \sum_{i=1}^k E_i = \frac{\sum_{i=1}^k \lambda_i}{\sum_{j=1}^n \lambda_j} \times 100\% \quad (3)$$

The SVM process is as follows. For the wood grain sample $(x_i, y_i), i=1, 2, \dots, n$ after PCA dimensionality reduction, x_i represents the reduced-dimensionality gloss-roughness data, where $x_i \in R^n$. y_i denotes the sample label, where $y_i \in \{0, 1\}$, with $y_i = 0$ for artificial wood grain samples and $y_i = 1$ for genuine wood grain samples. The objective of linear SVM is to find a decision boundary (Sheng *et al.* 2022; Zhang *et al.* 2025) that satisfies Eq. 4,

$$\omega^T x + b = 0 \quad (4)$$

where ω represents the weight vector and b represents the bias term, which determines the position of the decision boundary along the direction of the weight vector.

In the case of linear separability, the decision boundary is sought to maximize the margin between the two classes, *i.e.*, to maximize Eq. 5,

$$\begin{cases} \min_{\omega, b, \xi} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m \xi_i \\ y_i (\omega^T x_i + b) \geq 1 - \xi_i \end{cases} \quad (5)$$

where $\xi_i \geq 0$ is a slack variable, allowing for a small margin of error. $C > 0$ is a penalty coefficient, balancing the interval and misclassification.

Next, one constructs the Lagrangian dual to eliminate ω, b, ξ and obtains,

$$\begin{cases} \max_{\alpha} \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ s.t. \sum_i \alpha_i y_i = 0, 0 \leq \alpha_i \leq C \end{cases} \quad (6)$$

where α_i is the Lagrange multiplier. $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ is the inner product kernel function of the nonlinear mapping. In the optimal solution α^* , the samples of $\alpha^* > 0$ are the support vectors, denoted as SV .

The classification decision is shown in Eq. 7.

$$f(x) = \text{sign}(\sum_{i \in SV} \alpha_i^* y_i K(x_i, x) + b) \quad (7)$$

Extensive experimental results indicate that using Gaussian radial basis function (RBF) kernel functions yields superior classification performance for SVMs (Ke *et al.* 2021; Wainer and Fonseca 2021). The expressions are shown in Eqs. 8 and 9.

$$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \quad (8)$$

$$\text{Gamma} = \frac{1}{2\sigma^2} \quad (9)$$

The structure of SVM is shown in Fig. 2. It can be viewed as a three-layer architecture comprising the input layer, hidden layer, and output layer (Yang *et al.* 2022). The input layer receives feature vectors representing wood grain glossiness and roughness after PCA dimensionality reduction. Each node in the hidden layer “stores” a support vector, measuring its similarity to input samples *via* a kernel function. The output layer sums the similarity values from all hidden nodes, weighted by corresponding weights and biases, to produce a discriminative score. Based on the sign of this score, the model outputs the classification result of genuine or artificial wood grain.

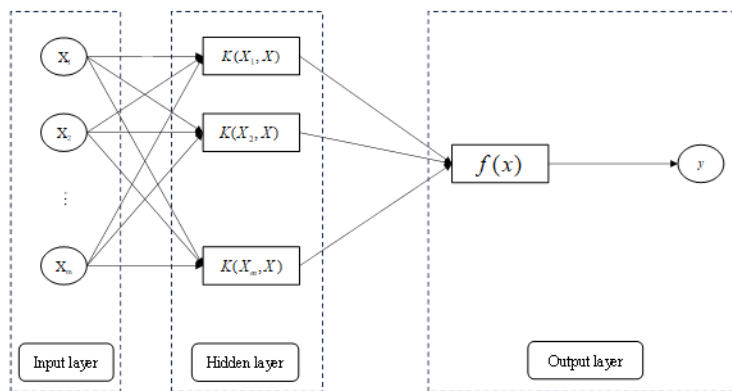


Fig. 2. Structure diagram of SVM

To quantitatively evaluate the quality and performance of this real-fake wood grain classification model, accuracy, F1 score, and MCC are adopted as evaluation metrics to comprehensively assess the model's classification effectiveness (Ferracina *et al.* 2025). Accuracy is defined as the proportion of correctly predicted samples in the entire dataset, calculated as shown in Eq. 10.

$$\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

Among these, TP denotes the true positive class, representing samples whose actual category is artificial wood grain and are correctly predicted as artificial. Similarly, TN denotes the true negative class, representing samples whose actual category is artificial wood grain and are correctly predicted as such. FP refers to the false positive class, where the actual category is genuine wood grain but the prediction is artificial, and FN refers to the false negative class, where the actual category is artificial wood grain but the prediction is genuine.

The F1 score is defined as the harmonic mean of precision and recall, as shown in Eq. 11. Its value ranges from 0 to 1, with higher values indicating better model performance.

$$\text{F1-score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (11)$$

where $precision = TP / (TP + FP)$, $recall = TP / (TP + FN)$

The Matthews Correlation Coefficient (MCC) is a comprehensive balanced metric for binary classification problems. The MCC ranges from -1 to 1, where 1 indicates perfect prediction, 0 indicates no better than random prediction, and -1 indicates complete disagreement between predictions and observations. On imbalanced datasets, MCC may provide more informative insights than accuracy or F1 score, as it offers a more balanced measure of the proportion of all correctly predicted classes (Boughorbel *et al.* 2017; Chicco *et al.* 2017). Its calculation formula is as follows.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (12)$$

EXPERIMENTAL

Materials and Instruments

The test samples used in this study included 7 common types of wood and 15 wood-grain decorative films (3M China Ltd.). The specific details of the samples are shown in Table 1.

Table 1. Test Sample List

No.	Name	Dimensions	Category
1	Red Oak	30 cm × 15 cm	Genuine Wood Grain
2	Chestnut	17 cm × 9 cm	Genuine Wood Grain
3	Teak	15 cm × 10 cm	Genuine Wood Grain
4	Black walnut	30 cm × 10 cm	Genuine Wood Grain
5	Beech	20 cm × 15 cm	Genuine Wood Grain
6	Camphor Tree	20 cm × 20 cm	Genuine Wood Grain
7	Maple	20 cm × 20 cm	Genuine Wood Grain
8	DW-1873MT	25 cm × 20 cm	Artificial Wood Grain
9	DW-1876MT	25 cm × 20 cm	Artificial Wood Grain
10	DW-1897MT	25 cm × 20 cm	Artificial Wood Grain
11	DW-1899MT	25 cm × 20 cm	Artificial Wood Grain
12	DW-2198MT	25 cm × 20 cm	Artificial Wood Grain
13	DW-2209MT	25 cm × 20 cm	Artificial Wood Grain
14	DW-2213MT	25 cm × 20 cm	Artificial Wood Grain
15	DW-2216MT	25 cm × 20 cm	Artificial Wood Grain
16	FW-1125	25 cm × 20 cm	Artificial Wood Grain
17	PW-2307MT	25 cm × 20 cm	Artificial Wood Grain
18	PW-2310MT	25 cm × 20 cm	Artificial Wood Grain
19	PW-2315MT	25 cm × 20 cm	Artificial Wood Grain
20	WG-1141	25 cm × 20 cm	Artificial Wood Grain
21	WG-1841	25 cm × 20 cm	Artificial Wood Grain
22	WG-833	25 cm × 20 cm	Artificial Wood Grain

Measurement of material surface gloss was conducted in accordance with ASTM D523 requirements using a multi-angle precision gloss meter (CS-380SE, Hangzhou CHNSpec Technology Co., Ltd) to measure the gloss of real wood and simulated wood grain film surfaces, as shown in Fig. 3(a). This gloss meter simultaneously obtained gloss data at 20°, 60°, and 85° angles for the sample surface. Surface roughness measurements

were performed using a confocal microscope (VK-X3000, Keyence) to analyze the surface topography of the samples, as shown in Fig. 3(b). An objective magnification of 5X and an eyepiece magnification of 10X were selected.

To ensure the repeatability and comparability of gloss measurements, all experiments were conducted under controlled laboratory conditions. The experimental environment was maintained at 23 ± 1.5 °C and $48\% \pm 4\%$ relative humidity.

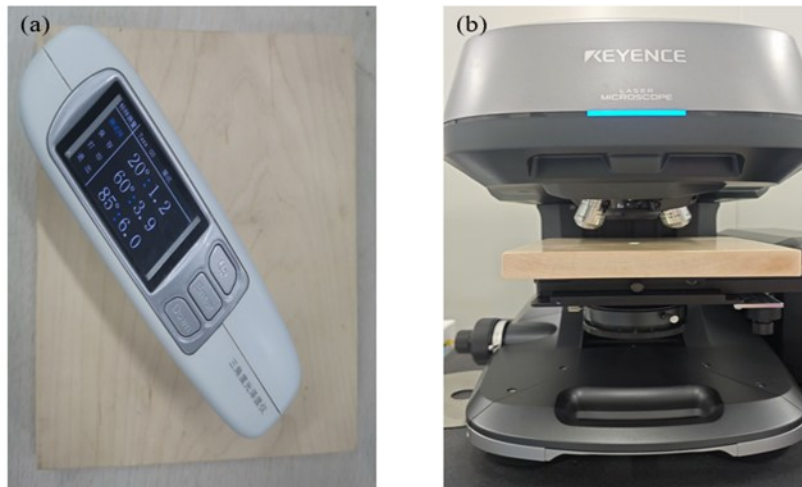


Fig. 3. Test diagram: (a) shows the gloss test diagram for the wood surface. (b) shows the surface roughness test diagram

Sanding

To ensure that the established model applies to natural wood at any processing level, this study employed 11 different grits of sandpaper (120, 320, 400, 600, 800, 1000, 1200, 1500, 2000, 3000, and 5000 grit) mounted on an automatic sanding machine for sequential grinding. Throughout the grinding process, constant pressure was maintained between the sandpaper and the specimen surface. The grinder was set to rotate at 3000 rpm to ensure consistent grinding methods and energy input across all grit sizes. To maximize the grinding efficiency of each grit size, grinding time for each grit was no less than 5 minutes. Additionally, each sheet of sandpaper was replaced after 10 minutes of continuous grinding to prevent efficiency loss due to abrasive wear. Following completion of grinding with a single grit size, the specimen surface was immediately cleaned and dusted. This involved wiping the surface with ethyl alcohol and allowing it to rest for 10 minutes under identical environmental conditions. This step eliminated surface variations caused by heat accumulation and humidity changes during grinding, thereby ensuring the reliability of subsequent gloss and surface roughness measurements.

No varnish or coating was applied to the specimens in this study. The purpose was to establish a baseline model based on the intrinsic surface morphology and optical response of natural wood. Varnish layers may introduce additional variables, such as coating thickness, curing conditions, and surface leveling, which could affect gloss and roughness measurements.

Data Collection

For natural wood samples, after sandpaper pre-treatment, nine sampling points were uniformly distributed across the surface to ensure effective characterization of the

overall surface features. Preliminary tests indicated that when the sampling density exceeded approximately six to eight points, the measured gloss distribution became stable. Therefore, nine representative sampling points were selected, covering visually distinct regions of the surface, to balance measurement reliability and experimental efficiency. For each sampling point, a gloss meter recorded gloss values at three observation angles (20° , 60° , and 85°) under three orientations relative to the wood grain direction: 0° , 45° , and 90° . This forms a gloss feature vector with a length of 9 at the current sampling point location. To minimize the impact of random errors on the discrimination results, three independent measurements were performed at each sampling point. Simultaneously, under identical conditions, a confocal microscope recorded the three-dimensional topographical distribution near each sampling point, as shown in Fig. 3(a)-(d).

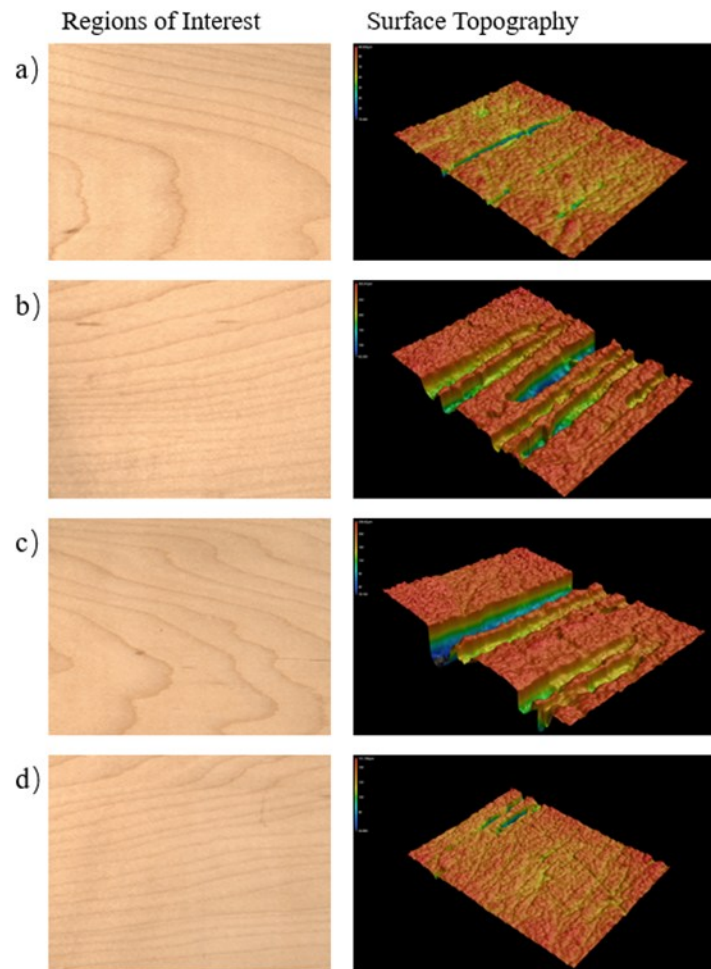


Fig. 4. Regions of Interest and 3D Surface Topography: a-d show the ROI (left) and the surface topography of the corresponding region (right)

Algorithms were then used to extract two surface roughness parameters: the arithmetic mean height (S_a) and the maximum height (S_z) near the current sampling point. Under a single sandpaper grit condition, each sample yields 9 sets of 9-dimensional gloss data and 2-dimensional surface roughness data, totaling 9 sets of 11-dimensional data. Combining unpolished samples and those treated with 11 different grits of sandpaper, a total of 756 sets of 11-dimensional gloss-surface roughness data were collected for natural

wood. The method for obtaining gloss and surface roughness data for simulated wood grain is fundamentally consistent with that for natural wood grain, with two key differences: First, artificial wood grain requires no pre-treatment such as sandpaper abrasion. Second, 25 sampling points are uniformly distributed across a single sheet of artificial wood grain paper to ensure the collected data volume remains comparable to that of natural wood grain. Following the aforementioned collection process, a total of 375 sets of 11-dimensional gloss-roughness data were ultimately collected for the simulated wood grain.

Selected gloss data are presented in Table 2, while selected surface roughness data are shown in Table 3. The two top rows in Table 2 represent different angular parameters involved in the gloss measurement process. The first row (20°, 60°, and 85°) corresponds to the standard specular measurement angles of the multi-angle gloss meter, which define the reflection geometry according to gloss measurement standards. The second row (0°, 45°, and 90°) denotes the relative orientation angle between the measurement direction and the wood grain direction on the sample surface, which was introduced to evaluate the anisotropic optical behavior of wood textures.

It should be noted that the acquisition time for a single gloss measurement was approximately 10 s, and surface roughness measurement took about 60 s. As a result, the total data acquisition time per sample was approximately 70 s, indicating that the proposed approach is practically feasible for industrial inspection scenarios.

Table 2. Gloss Data for Samples

Name	20°			60°			85°		
	0°	45°	90°	0°	45°	90°	0°	45°	90°
Camphor Tree	0.7	0.7	0.7	2.9	2.9	3	3.5	4.6	3.1
Maple	1.1	1.1	1.1	3.1	3.1	3.7	2.4	1.8	3.2
Red oak	0.8	0.8	0.7	3.2	3.2	2.9	5.6	5.3	4.6
Chestnut	1	0.9	0.9	5.5	5	3.9	16.8	14.6	11.4
DW-1899MT	0.2	0.2	0.2	0.6	0.5	0.5	1.1	1.1	1
DW-2213MT	0.2	0.2	0.2	0.4	0.4	0.4	1	1	1
WG-1841	0.9	0.9	0.9	8.8	8.2	8.3	23.4	19.6	18.8
WG-833	0.2	0.2	0.2	0.6	0.5	0.6	2.3	1.9	1.9

Table 3. Roughness Data Table for Samples

Name	Arithmetic Mean Height (Sa)	Peak Height (Sz)
Black walnut	10	182
Teak	13	196
Beech	11	95
DW-1876MT	20	117
DW-2198MT	9	74
FW-1125	4	48

RESULTS AND DISCUSSION

Model Construction

To determine the optimal number of principal components to use as model discrimination criteria, the test set data were input into the constructed PCA-SVM genuine/artificial wood grain recognition model. The explained variance and cumulative explained variance of each principal component vector were calculated according to Equations (2) and (3), as shown in Table 4.

Table 4. Explained Variance and Cumulative Explained Variance of Principal Components

PC No.	Explained Variance (%)	Cumulative Explained Variance (%)
1	68.9842	68.9843
2	29.0806	98.0649
3	0.7376	98.8025
4	0.6567	99.4592
5	0.3442	99.8034
6	0.1046	99.9080
7	0.0491	99.9570
8	0.0238	99.9809
9	0.0153	99.9962
10	0.0023	99.9984

Based on the calculated explained variance and cumulative explained variance, the first principal component accounted for 68.9842% of the output data, while the cumulative explained variance of the top 10 principal components reached 99.9984%. To determine the optimal number of principal components for model input, different numbers of principal components were selected in descending order and applied to the model. The generated models were then tested on the test dataset. The classification accuracy of the test set is shown in Table 5. Based on the test set classification results, it can be observed that when 6 principal components were selected as inputs, the classification accuracy of the test set output no longer changed. Therefore, it can be concluded that for this classification model, the first 6 principal components had extracted most of the information useful for classification. To reduce noise and redundancy while lowering model complexity, selecting 6 principal components as inputs ensures sufficient information retention while avoiding unnecessary dimensions and overfitting.

Table 5. Accuracy Rates for Different Principal Component Inputs on the Test Set

PC No.	Test set accuracy (%)
1	68.53
2	94.12
3	95.59
4	96.18
5	96.47
6	96.76
7	96.76
8	96.76
9	96.76
10	96.76

Analysis

After applying PCA to the original test dataset, the top 6 principal component vectors were selected as input features to construct a PCA-SVM classification model. The resulting confusion matrix is shown in Fig. 5(b). The results demonstrate that this method achieved a TP value of 0.9043 and a TN value of 1 for wood grain classification, indicating excellent classification performance. To further validate the effectiveness of PCA-based dimensionality reduction in feature extraction, the original untransformed test dataset was directly input into the SVM model for classification. The resulting confusion matrix is shown in Fig. 5(a). The TP value and TN value decreased to 0.8522 and 0.8259, respectively, which is significantly lower than the wood grain classification performance achieved by the PCA-SVM model. This indicates that PCA effectively eliminated redundant and noisy features (such as partial anomalies or error data) during dimension reduction, thereby enhancing the overall accuracy and robustness of wood grain recognition.

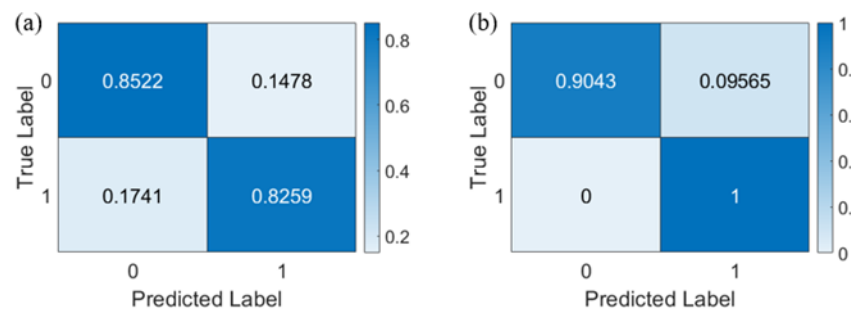


Fig. 5. The confusion matrix for (a) SVM and (b) PCA-SVM

Additionally, the original gloss-roughness test dataset recorded in this paper was input into logistic regression (LR), partial least squares regression (PLS), and principal component regression (PCR) models for wood grain recognition. The confusion matrix results are shown in Fig. 6. To evaluate the overall performance of different models in wood grain recognition, the accuracy, F1-Score, and MCC values for each model were calculated, with results shown in Table 6. Compared with the standalone SVM, LR, PLS, and PCR models, the proposed PCA-SVM model demonstrated superior performance across multiple metrics including accuracy, F1-Score, and MCC. This validates its effectiveness and advantages in the wood grain recognition task.

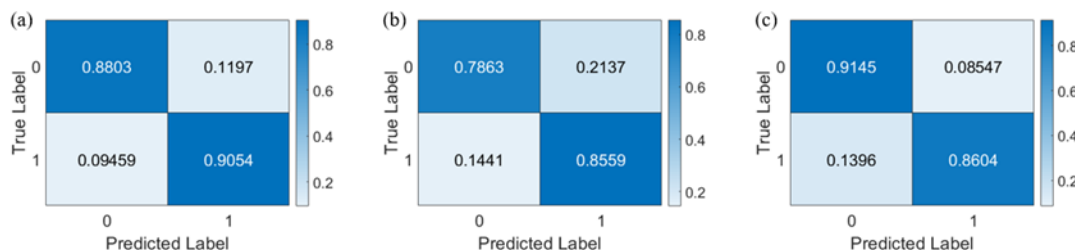


Fig. 6. Confusion matrix for (a) LR, (b) PCR, and (c) PLS

Table 6. Performance Metrics of Different Models

Model	Accuracy/%	F1 score	MCC
SVM	83.48	0.8685	0.6542
PCA-SVM	96.76	0.9761	0.9285
LR	89.68	0.9251	0.7764
PCR	83.19	0.8743	0.6415
PLS	87.91	0.9	0.7266

CONCLUSIONS

1. This paper proposes a principal component analysis – support vector machine (PCA-SVM)-based method for authenticating genuine and artificial wood grain. By capturing three-angle gloss data of genuine and artificial wood grain at different texture orientations, a 9-dimensional gloss feature set is obtained. This feature set is combined with surface roughness parameters S_a and S_z to construct the feature space. After standardized preprocessing, the data is divided into training and test sets at a 7:3 ratio.
2. During feature extraction, Principal Component Analysis (PCA) identifies the top k most discriminative principal components to achieve dimensionality reduction and noise suppression. In the classification stage, a Support Vector Machine (SVM) model is trained to distinguish test samples. Experiments validated the feasibility of this method for genuine/artificial wood grain recognition and systematically evaluated its classification performance.
3. The results demonstrated an identification accuracy of 96.76% on the test set, with an F1 score of 0.9761 and a MCC value of 0.9285—all significantly outperforming standalone classification methods such as SVM, LR, PLS, and PCR.
4. The proposed method integrates dimensionality reduction and noise suppression in feature extraction and classification modeling, enabling stable and efficient discrimination within the sample range examined in this study. The results demonstrate the feasibility of using gloss and surface roughness features to distinguish natural wood from the selected artificial wood grain films. However, the present model was trained and validated using a limited set of commercially available simulated wood grain products. Its applicability to new artificial wood grain products from different manufacturers or with different surface finishing processes still requires further validation. Therefore, future work will expand the sample database to include more sources, materials, textures, and finishing conditions, so as to improve the generalization ability and practical applicability of the model.

ACKNOWLEDGMENTS

The authors are grateful for continued financial support from National Science and Technology R&D Special Project: Research on Key Metrology and Testing Technologies for Multimodal Perception in Artificial Intelligence (No. 2021YFF0600203), National Science and Technology R&D Special Project: Complex Motion Grating Measurement Calibration Model, Error Compensation Technology, and Validation (No. 2023YFF0616200).

REFERENCES CITED

- Boughorbel, S., Jarray, F., and El-Anbari, M. (2017). "Optimal classifier for imbalanced data using Matthews correlation coefficient metric," *PLOS ONE* 12(6), article e0177678. <https://doi.org/10.1371/journal.pone.0177678>
- Cheng, Y. Z., Zhong, L. H., and Sun, X. K. (2021). "Review of wood identification technology," *Agricultural Technology & Equipment* (01), 125-128.
- Chicco, D. (2017). "Ten quick tips for machine learning in computational biology," *BioData Mining* 10(1), article 35. <https://doi.org/10.1186/s13040-017-0155-3>
- China Industry Information Network (2020). "Analysis of application fields and supply-demand status of decorative paper in China in 2019," *Paper and Papermaking* 39(6), article 52.
(<https://d.wanfangdata.com.cn/periodical/CiFQZXJpb2RpY2FsQ0hJU29scjI1TMjA5NTEwMjEwOTUwNDYSDXpoenoyMDIwMDYwMjAaCGVxbGM4dGl3>)
- Cortes, C., and Vapnik, V. N. (1995). "Support vector networks," *Machine Learning* 20(3), 273-297. <https://doi.org/10.1023/A:1022627411411>
- Ferracina, F., Krishnamoorthy, B., Halappanavar, M., Shengwei, H., and Vidyasagar, S. (2025). "Predictive analytics of selections of russet potatoes," *Crop Science* 65, article e21432. <https://doi.org/10.1002/csc.2.21432>
- Jiang, C., Zhang, X. H., Zou, X. L., Lu, G. Y., and Liu, Y. C. (2019). "Identification of *Salmonella* and *Listeria monocytogenes* by near-infrared spectroscopy and PCA," *Meteorological and Environmental Research* 10(4), 77-80.
<https://doi.org/10.59547/j.issn2152-3940.9019.54.518>
- Ke, M., Zhao, Y. D., Dian, H. W., and Kit, P. W. (2022). "A self-adaptive RBF neural network classifier for transformer fault analysis," *IEEE Transactions on Power Systems* 25(3), 1350-1360. <https://doi.org/10.1109/TPWRS.2010.2040491>
- Li, K., Zhang, M. M., Yang, Z., and Lv, B. (2018). "Classification for decorative papers of wood-based panels using color and glossiness parameters in combination with neural network method," *Journal of Forestry Engineering* 3(01), 16-20.
<https://doi.org/10.13360/j.issn.2096-1359.2018.01.003>
- Li, X. B. (2019). "Discussion on decorated materials in custom furniture," *China Wood-Based Panels* 26(4), 5-8.
- Liu, S. J., Wang, J. J., Lu, Y., Zheng, C., Jiao, L. C., Guo, J., Yin, Y. F., and He, T. (2023). "Application of machine learning approaches to wood species identification," *World Forestry Research* 36(05), 70-75.
<https://doi.org/10.13348/j.cnki.sjlyyj.2023.0070.y>
- Peng, X. R., and Zhang, Z. K. (2016). "A review of manufacturing technology of flexible decorative veneers," *China Wood Industry* 30(6), 23-26.
<https://doi.org/10.19455/j.mcgy.20160605>
- Peng, X. R., Lyu, B., Wang, C., Zhao, L. Y., and Zhang, Z. K. (2021). "New products and technologies of decorative surfaces for wood products," *Chinese Journal of Wood Science and Technology* 35(02), 6-11.
- Sheng, J. Q., Fang, Y., Zhang, L., and Wu, Q. Y. (2022). "Wireless network security state detection method based on particle swarm optimization support vector machine," *Techniques of Automation and Applications* 41(12), 55-58.
[https://doi.org/10.20033/j.1003-7241.\(2022\)12-0055-04](https://doi.org/10.20033/j.1003-7241.(2022)12-0055-04)
- State Forestry Administration (2014). *China Forestry Development Report 2014*, China Forestry Publishing House, Beijing.

- Wainer, J., and Fonseca, P. (2021). "How to tune the RBF SVM hyperparameters? An empirical evaluation of 18 search algorithms," *Artificial Intelligence Review* 54(6), 4771-4797. <https://doi.org/10.1007/s10462-021-10011-5>
- Wang, J., Long, J., Luo, S., and Gou, X. B. (2022). "Development status and new technologies of paper-based decorative materials," *Chinese Journal of Wood Science and Technology* 36(03), 10-14. <https://doi.org/CNKI:SUN:MCGY.0.2022-03-002>
- Yang, Z., Zhang, P. P., Wang, G. Q., and Du, B. X. (2022). "Research on mixed gas classification based on PCA-SVM," *Journal of Natural Science of Heilongjiang University* 39(03), 345-354. <https://doi.org/10.13482/j.issn1001-7011.2022.03.007>
- Ye, J. Y., Xie, X. Q., Zhuo, Y., Chen, B. Q., Shen, J. X., and Liu, Y. Q. (2019). "Research on decorative wood-based panel for use in furniture," *Furniture* 40(1), 35-37. <https://doi.org/10.16610/j.cnki.jiaju.2019.01.006>
- Zhang, M. M., Yang, Z., Lyu, B., Liu, Y. N., and Sun, X. D. (2015). "Pattern recognition of decorative papers with different visual characteristics using visible spectroscopy coupled with principal component analysis (PCA)," *Spectroscopy and Spectral Analysis* 35(2), 443-446.
- Zhang, Q. L. (2006). "The application of decorative paper in wood-based panel surface decoration," *China Wood-Based Panels* 13(11), 1-4, 28.
- Zhang, S. Z., Wang, R., Zhao, Y. X., and Zhang, M. J. (2025). "Discrimination analysis of Qingke liquor by using ultra-violet spectroscopy combined with chemometrics," *Journal of Instrumental Analysis* 44(06), 1107-1114.

Article submitted: October 22, 2025; Peer review completed: December 1, 2025; Revised version received: December 14, 2025; Accepted: February 1, 2026; Published: May 21, 2026.

DOI: 10.15376/biores.21.3.6253-6266