

Wood Defect Classification Based On Support Vector Machine

Hongbo Mu^{*1}, Yang Yang¹, Haiming Ni² and Dawei Qi^{*1}

¹College of Science, Northeast Forestry University, Hexing Road 26, Harbin, Heilongjiang Province, 150040, PR China

²College of Mechanical and Electrical Engineering, Northeast Forestry University, Hexing Road 26, Harbin, Heilongjiang Province, 150040, PR China

*Corresponding Author, E-mail: mhb-506@163.com, E-mail: qidw9806@126.com

Abstract

Effective identification of wood decay and crack defects by using support vector machine (SVM) theory. Extracting wood defect image through the image acquisition system, then processing the defect image by using gray level change, equalization, and median filtering et al. So as to achieve the purpose of improve the image quality and detection accuracy. Segment the target images, and then measures the area, perimeter, and diameter of the wood defects. Three eigenvalues, area and perimeter square ratio, length to diameter ratio and circumference and area ratio, which can be identified the wood defects was obtained. Separating these eigenvalues into training set and testing set. Training Support vector machine by using training set, establish a preliminary model of support vector machine. Using support vector machine model for accuracy test, if the test accuracy is low, repeatedly adjust the parameters of support vector machine for training and testing until reach the test accuracy. Making sure the kernel function and various parameters of support vector machine, constructing the support vector machine, which using for wood defect classification. The experimental results showed that this method has fast calculation speed, high precision, and helps to raise the utilization rate of wood.

Keywords: wood defects; image processing; support vector machine; classification

1. Introduction

Support vector machine (SVM) is an emerging classification technology that has many unique advantages to solve the small sample, nonlinear and high dimensional pattern recognition problems. Nowadays, the SVM has been widely used in wood nondestructive testing [1]. The traditional artificial wood test is difficult to meet the needs of the rapid non-destructive testing, so that the nondestructive testing methods get fast development and form a practical wood testing technology [2]. At present, the SVM algorithm has used in pattern recognition, regression estimation, the probability density function estimation et al. For example, the SVM algorithm has higher accuracy in the handwritten numeral pattern recognition, speech recognition, face image recognition, and article classification [3].

In this paper, we main detecting two kinds of typical wood defect types corrupt and crack by using support vector machine method. The wood defect types can be distinguished by using X-ray image. Through analysis the characteristic values of the X-ray image, the SVM which can recognized wood defect species can be established, and then the kinds of defects can be quickly and accurately recognized.

2. Wood Defects Image Pre-Processing

2.1. Histogram Equalization

The defective parts and the no defect parts of the timber have different X-ray absorption ability, when X-ray through the tested timber. The X-ray absorption of the defective parts always above (or below) the no defect parts, thus it can be used to detect the presence of wood defects. Histogram of natural images generally has greater frequency on the low gray zone, so that the details in darker areas are not clear. With the method of histogram dressing, gradation concentrated area of the original image can be made opened or equally distributed, thereby increasing the contrast, making clear image detail and achieving the purpose of image enhancement.

The collected images were histogram equalization processed with MATLAB, and the defect images and the histogram which shows the effects before and after the processing were shown in Figure 1 to Figure 8 respectively.

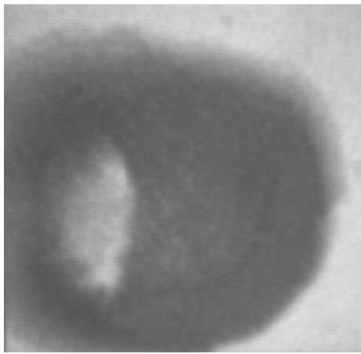


Figure 1. The Original Decay Image

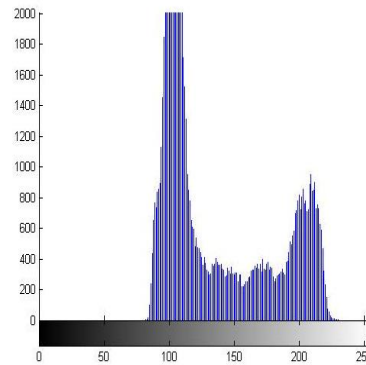


Figure 2. The Histogram of the Original Decay Image

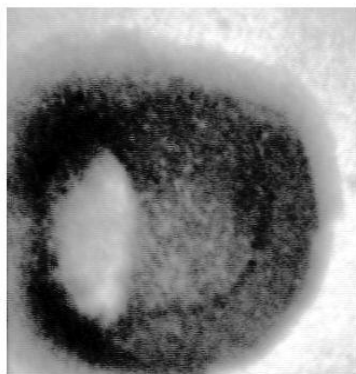


Figure 3. The Decay Image after Image Equalization

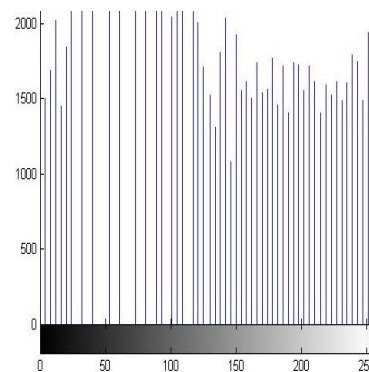


Figure 4. The Decay Image Histogram Equalization after Image Equalization

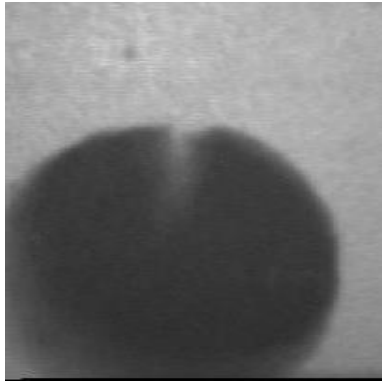


Figure 5. The Original Crack Image

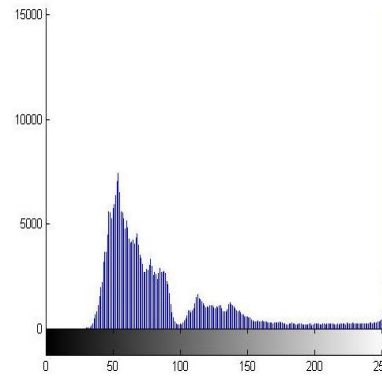


Figure 6. The Histogram of the Original Crack Image

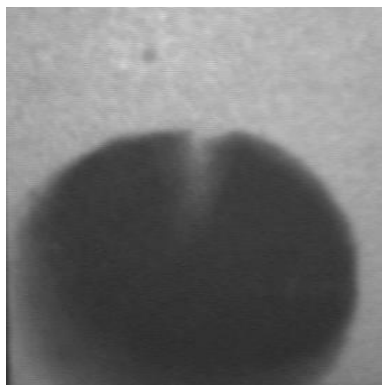


Figure 7. The Crack Image after Image Equalization

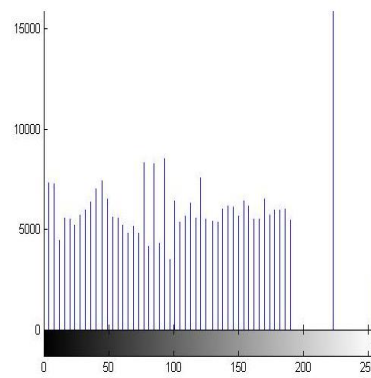


Figure 8. The Crack Image Histogram Equalization after Image Equalization

2.2. Median Filtering Process of Wood Defect Images

During the processing of image signal generation, transmission and recording, defect image acquisition is always affected by various noise, and image signal and noise are often intertwined, so that the details of the image itself such as boundary contours, lines are blurred unclear. Using Median filter process can not only remove noise but also preserve image detail, thereby improving image quality. The leveled decay and cracks images were processed with median filter using MATLAB, and the images obtained were shown in Figure 9 and Figure 10.

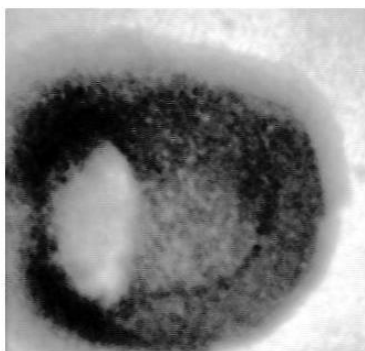


Figure 9. The Decay Image after Median Filtering

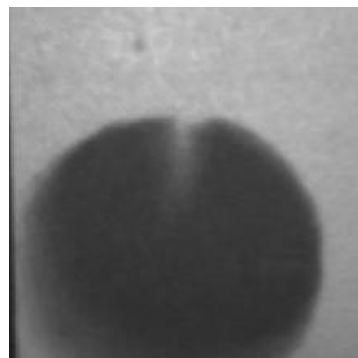


Figure 10. The Crack Image after Median Filtering

2.3. Wood Defect Image Segmentation

In order to extract the timber defects eigenvalues, the defective portion of wood should be separated from the entire image. For defects were decay and crack, first we preliminary separate the images with gray threshold segmentation method, and then use the edge detection method to segment the defective portion image of the wood, which used a Sobel edge detection operator for edge detection. Figure 11-14 shows the decay and cracks images, which were segmented by gradation threshold and edge detection.

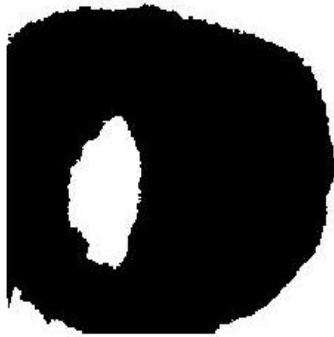


Figure 11. The Decay Image after Gray-Level Threshold Segmentation

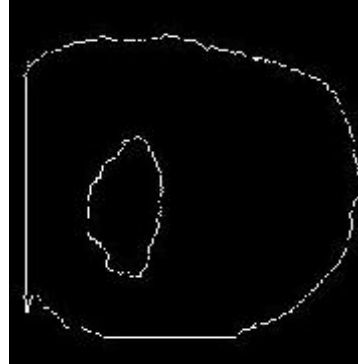


Figure 12. The Decay Image after Gray-Level Threshold Segmentation



Figure 13. The Crack Image after Gray-Level Threshold Segmentation



Figure 14. The Crack Image after Gray-Level Threshold Segmentation

3. The Selection and Extraction of Wood Defect Characteristic Value

3.1. The Selection of Wood Defect Characteristic Value

In order to get the multi-dimensional characteristic values [4], the defect feature was extracted. And then, put the obtained characteristic values into the SVM, complete the wood defect recognition. Therefore, how to choose a characteristic value of wood defect is especially important, which determines the SVM classification accuracy for wood defects. Usually, there are many wood defect image characteristics [5] can be used in the defect classification, such as the shape features, the color features and the texture features, *etc.* According to the related theory of wood defect classification, the shape feature is the most important and effective basis for identify the types of wood decay and crack defects. Therefore, the geometry feature is firstly considered to extracting for wood defect classification. While the crack defect usually slender, the decay defect always close to a rectangle or irregular oval shape. According to the shape features of these two kind of

defects, we can choose some shape characteristic value as follows: The ratio of perimeter square and area $\frac{C^2}{S}$; The ratio of long radius to short radius $\frac{D_1}{D_2}$; The ratio of area and perimeter $\frac{S}{C}$.

3.2. The Extraction of Wood Defect Characteristic Value

The image edge of wood defect becomes apparent after the image pretreatment. The perimeter, long radius, short radius, and area can be calculated by computing the pixels of the defect part and edge. Measured the above wood defect images by using the ImageJ software, the S, C, D1, D2 can be obtained, and the $\frac{C^2}{S}$, $\frac{D_1}{D_2}$, $\frac{S}{C}$ can be calculated. Part of the wood defect data extraction results as shown in Table 1.

Table 1. Part of S, C, D1, D2 and Some Characteristic Values of Wood Defect

	area(S)	perimeter(C)	Long radius(D_1)	Short radius(D_2)	The ratio of long radius and short radius($\frac{D_1}{D_2}$)	The ratio of area and perimeter ($\frac{S}{C}$)	The ratio of perimeter square and area ($\frac{C^2}{S}$)
cr ack	158.1	65.553	17.021	4.253	4.002	2.412	27.14
	32						2
	2196.	412.031	176.003	13.203	13.331	5.330	77.29
	302						8
	4714.	613.136	189.390	18.142	10.439	7.689	79.74
de cay	102						7
	181.2	60.825	20.693	4.210	4.915	3.011	20.44
	12						1
	6674.	353.642	94.896	78.438	1.209	18.87	18.73
	968					5	6
de cay	4776.	295.536	89.938	61.107	1.472	16.16	18.28
	152					1	7
	610.8	82.286	27.893	15.327	1.820	7.423	11.08
	36						5
	1850.	178.218	71.367	42.373	1.684	10.38	17.16
102					1	8	

4. Application of Support Vector Machine to Wood Defect Classification

4.1. The Support Vector Machine

The linear separable SVM earliest comes from the optimal classification plane in the linear separable cases. Firstly, consider a two kinds of pattern classification problems, assumes that the training data are:

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n), \quad x \in R^n, y \in \{+1, -1\} \quad (1)$$

where n is the number of samples, and also the dimensions of the input space. There will be a hyperplane make the two kinds of samples completely separate in the case of linear separable. The hyperplane is described as: $\omega x + b = 0$, ω is the normal direction of the hyperplane.

The classification is as follows:

$$\omega x_i + b > 1, \text{ if } y_i = 1, \quad \omega x_i + b \leq -1, \text{ if } y_i = -1 \quad (2)$$

As shown in Figure 15: If the training data can be divided no error. And the vector has the longest distance with the hyperplane, in which, the hyperplane was defined as the optimal hyperplane and the vector has the nearest distance between each kind of data and the hyperplane.

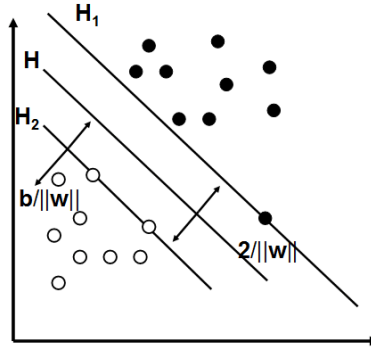


Figure 15. The Optimal Hyperplane

In Figure 15, H1 and H2 respectively are the sample planes, which have the nearest distance and parallel to the hyperplane in all kinds of samples. The distance between them is called classification interval. The optimal classification plane not only separated the two kinds of samples (guarantee the minimum experiential risk), but also make the classification interval reach the biggest (guarantee the minimum fiducial range). Make the actual risk of the classification function to the minimum. In linear separable cases, solving the optimal hyperplane, which can be seen as a quadratic programming problem to solve formula (3):

$$\min \Phi(\omega) = \frac{1}{2} \|\omega\|^2 \text{ s.t. } y_i (\omega \cdot x_i + b) \geq 1, \quad i = 1, 2, 3, \dots, l \quad (3)$$

The classification plane, which can satisfied the formula (3) is called the optimal classification plane. Use these training sample points, which can satisfied $y_i (\omega \cdot x_i + b) = 1$ as the support vector. From Figure 15 we can seen that the support vector are the sample points on the plane H1 and plane H2. The optimization function $\Phi(\omega)$ is quadratic form, the constraints is linear. And the typical of quadratic programming problem can be obtained from the following Lagrange function:

$$L(\omega, b, a) = \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^n a_i [y_i (\omega \cdot x_i + b) - 1] \quad (4)$$

The extreme value point in formula (4) is saddle point, which is the optimal solution. Take the minimum value $b = b^*$ of L relative b , the minimum value $\omega = \omega'$ of L relative ω and the maximum value a^* of L relative a . Calculating the derivative of L :

$$\frac{\partial}{\partial b} L(\omega, b, a) = \sum y_i a_i = 0, \quad \frac{\partial}{\partial \omega} L(\omega, b, a) = \omega - \sum y_i a_i x_i = 0 \quad (5)$$

Where $\frac{\partial L}{\partial \omega} = \left(\frac{\partial L}{\partial \omega_1}, \frac{\partial L}{\partial \omega_2}, \dots, \frac{\partial L}{\partial \omega_l} \right)$

After calculation formula (5), substitute this result into formula (3). The dual problem of the optimization problem formula $R(\omega) \leq R_{emp}(\omega) + \sqrt{\frac{h(\ln(2n/h) + 1) - \ln(\eta/4)}{n}}$ can

be obtained (h is the vc dimension of the function set, n is the number of samples, η is the parameter satisfied $0 \leq \eta \leq 1$). The maximize function was defined as:

$$\max : W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (6)$$

Which satisfy the constraint conditions: $\sum_{i=1}^l y_i \alpha_i = 0$ s.t. $\alpha_i \geq 0, i = 1, 2, \dots, l$.

This is a quadratic function optimization problem, which have unique solution in the inequality constraints. On the solve and analysis of the constrained optimization problem, the Kamsh, Kuhn-Tucker (KKT) conditions playing an important role. For the question like formula (6), the solution must satisfied:

$$\alpha_i \{y_i (\omega \cdot x_i + b) - 1\} = 0, i = 1, 2, \dots, l \quad (7)$$

As can be seen from the formula (7), only $\alpha_i > 0$ has effect on the classification problems, while $\alpha_i = 0$ useless to the classification results. These samples are called support vector. The optimal classification function can be got after solution:

$$f(x) = \text{sgn}\{y(\omega \cdot x + b)\} = \text{sgn}\left\{\sum_{i=1}^l \alpha_i^* y_i (x \cdot x_i) + b^*\right\} \quad (8)$$

Where, $y_i (\omega \cdot x_i + b) \geq 1 - \xi_i$.

In fact, the sum operation in formula (8) only implement into support vector. The α_i^* is the optimal solution of α_i ; And the b^* is the classification threshold. By the original constraints formula (3), we can know:

$$b^* = -\frac{1}{2} \left[\max_{y_i=-1} (\omega \cdot x_i) + \min_{y_i=1} (\omega \cdot x \cdot x_i) \right] \quad (9)$$

For the unknown categories test vector, it classification can be identified through enter its value into formula (8), which after pretreatment.

For nonlinear separable cases, the basic idea is through mapping the predetermined nonlinear input vector to a high dimensional feature space and constructs the optimal hyperplane in the high dimensional space. However, even if the optimal hyperplane has good expansion and can be found in the theory, there are still technical problems of how to deal with high dimensional feature space.

Afterwards, people found it doesn't need to consider the feature space in explicit forms, when structure the optimal classification hyperplane in the feature space [6]. Moreover, this kind of inner product operation can be implemented by using the original space function; we even didn't need to know the specific transformation form. According to the functional theory, it can be correspond the inner product in one certain transformation space, as long as a kernel function $K(x_i, x_j)$ satisfies the Mercer condition. For nonlinear problem, the SVM use the feature mapping method and the linear problem calculation framework to achieve.

So by introducing kernel function makes $\Phi(x) \cdot \Phi(x_i) = K(x, x_i)$, the objective function of quadratic programming problem change into:

$$\max : W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \{\Phi(x_i) \cdot \Phi(x_j)\} = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (10)$$

Looking for a maximum of the objective function, the constraint condition is:

$$\sum_{i=1}^l y_i \alpha_i = 0, \text{ s.t. } \alpha_i \geq 0 \quad i = 1, 2, \dots, l.$$

The corresponding classification function can be constructed after training. This function as shown in formula (11):

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^l \alpha_i^* y_i K(x, x_i) + b^* \right\} \quad (11)$$

The sample can be classified by calculate the value of classification function. The SVM schematic was shown in Figure 16:

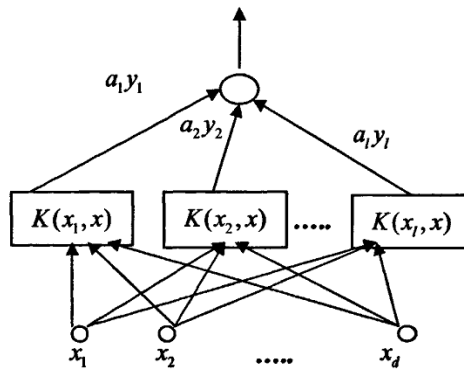


Figure 16. The SVM Schematic

In a nutshell, SVM realized the following ideas: Firstly, make use of nonlinear transform, which defined by the inner product function to transform the input space into a high dimensional space. In the form, the SVM classification function was similar to a neural network, the output is a linear combination of the intermediate node, and each intermediate node corresponds to a support vector [7].

Usually, one SVM can be determined just need to select the appropriate kernel function K , while doesn't need to know the $\Phi(x)$ and H . The necessary and sufficient conditions of kernel functions were given by the Mercer theorem [8]: For any function meet the formula (12), the formula (13) was established:

$$\int g(x)^2 dx < \infty \quad (12)$$

$$\iint K(x, y)g(x)g(y) \geq 0 \quad (13)$$

4. 2. The Design of the Support Vector Machine

The SVM flow chart [9], which used for wood defect recognition as shown in Figure 17.

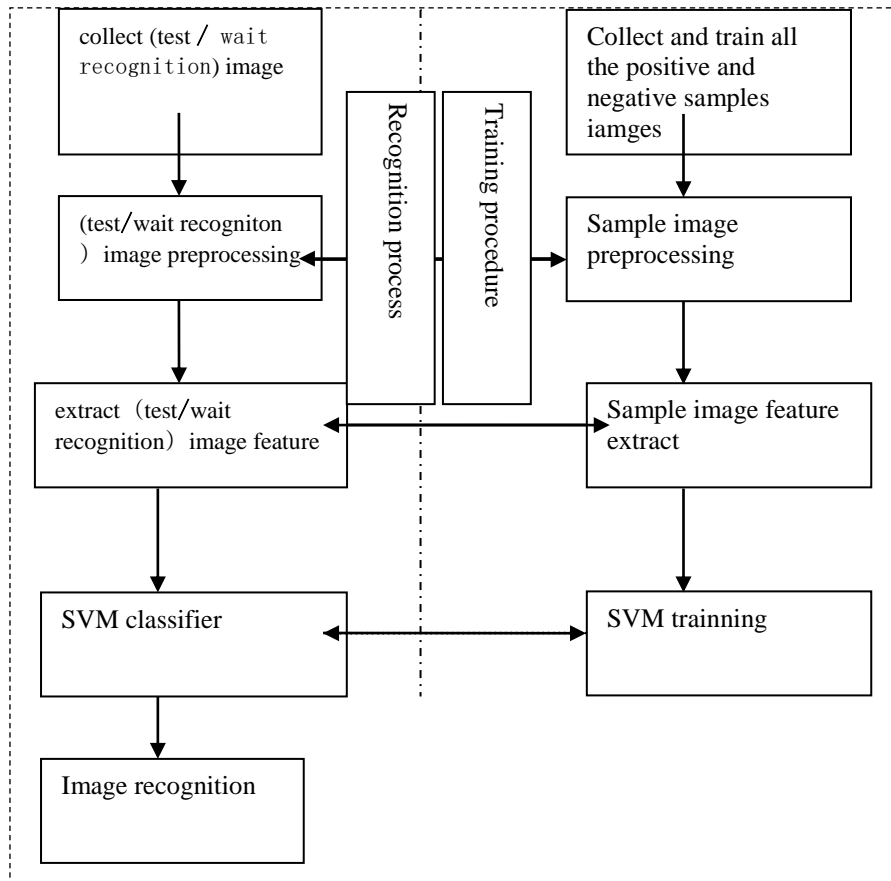


Figure 17. The Flow Chart of SVM Wood Defect Recognition

From Figure 17 we can see the main process of wood defect recognition by using SVM. The first step for image acquisition, image preprocessing and image feature extraction. The second step for training and testing. Finally, it can be used in image recognition. As shown in the above, the train set should be take into the SVM for training, and the preliminary model of SVM was obtained. Substitute the test set into the SVM to observe its classification effect. If the result is bad, the kernel function should be adjusted and tested repeatedly. Until reach the ideal test result, which can achieved wood defect recognition [10]. And then the design of the SVM was completed.

Divided the extracted characteristic values into two groups. One group as the train set, used for training the SVM. Another group as the test set, used to test the performance of the SVM [11]. Each group contains the eigenvalues denote crack and decay. Part of the group results are shown in Table 2 and Table 3.

Table 2. Part of Training Set

	Long radius and short radius	The ratio of area and perimeter	The ratio of perimeter square and area
crack (1)	4.002	2.412	27.142
)	13.331	5.330	77.298
decay (-	1.209	18.875	18.736
1)	1.472	16.161	18.287

Table 3. Part of Test Set

	The ratio of long radius and short radius	The ratio of area and perimeter	The ratio of perimeter square and area
crack (1)	1.209	18.875	18.736
)	1.472	16.161	18.287
decay (-)	1.820	7.423	11.085
1)	1.684	10.381	17.168

In the table above, 1 and -1 represents crack and decay defect, respectively.

After data standardization processing, these training and testing sets can be generation into the SVM for training and testing. For obtained different SVM model, we can only train these SVM respectively, which have different kernel function. And then, testing these kernel functions by using test set, chose the best kernel function, which have the best classification result. Finally, the SVM model can be determined.

From the above results, We know choose ' t^2 ', namely support vector machine with RBF kernel function, SVM test results right percentage is higher. So we can choose support vector machine (SVM) model which contain RBF kernel function to classify wood defects for better classification results. The completion of support vector machine (SVM) model construction make the new samples can be identified.

5. Results and Discussion

Support vector machine (SVM) method was used to identify the crack and decadent in wood. The ratio of perimeter and Square area, length diameter and short diameter, area and perimeter of defects were input SVM as the eigenvalue for Training, testing, and determining the kernel function. Support vector machine (SVM) model with good performance for identification was constructed which can used for wood defect detection.

Design of support vector machine (SVM) can classify defects for each image which contains one kind defect. When a X-ray image contains two or more than two defects, the image can be divided into several parts to identify the types of defects respectively.

Acknowledgement

This work is supported by the Fundamental Research Funds for the Central Universities (No. 2572014CB30), the Project of National Science Foundation of China (No.31570712), the Project of Natural Science Foundation of Heilongjiang Province of China (No.C201338) and the science and technology research project of education department of Heilongjiang province (No. 12543019).

References

- [1] Z.C. Wang and G.H. LI, "A method of wood defect identification and classifia-tion based on stress wave and SVM", Journal of Nanjing Forestry University (Natural Sciences Edition), vol. 39, (2015), pp. 130-136.
- [2] H.L. Feng, G.H. Li and Y.M. Fang, "Stress wave propagation modeling and application in wood testing", Journal of System Simulation, vol. 22, (2010), pp. 1490-1493.
- [3] G.Z. Li, M. Wang and J.H. Zeng, "An introduction of support vector machine", Beijing: Electronic Industry Press, (2004).
- [4] Y. Cui, "Digital image processing technology and application", Beijing: Electronic Industry Press.
- [5] X.P. He and Q.H. Zhang, "Image processing and analysis based on MATLAB", JOURNAL OF CTBU, vol. 20, (2003), pp. 123-128.
- [6] H. Guo, H.P. Liu and L. Wang, "Method for selecting parameters of least squares support vector machines and application", Journal of system simulation, vol. 18, (2006), pp. 2033-2037.

- [7] T.H. Wang, S.F. Tian and H.K. Huang, "Support vector machine based on weightiness of sample attribute", Journal of Beijing Jiaotong university, vol. 11, (2007), pp. 87-90.
- [8] M.D. Xue and L. Guo, "Image classification based on SVM algorithm", Computer engineering and applications, vol. 30, (2004), pp. 230-233.
- [9] Z.X. Lu and Y.D. Li, "Interactive support vector machine learning algorithm and its application", Journal of Tsinghua university science and thechnology, vol. 39, (1999), pp. 93-97.
- [10] J. Platt, N. Cristianini and T.J. Shawe, "Large margin DAGs for multiclassclassification", In: Advances in neural information processing systems, Massachusetts: MIT Press, (2000).
- [11] C.W. Reynold, "Flock, herds and schools : a distributed behavioral model", Computer Graphics", vol. 21, (1987), pp. 25-34.

