Condition assessment of timber utility poles based on a hierarchical data fusion model

Yang Yu[1], Ulrike Dackermann[2], Jianchun Li[3] and Mahbube Subhani[4]

Abstract: This paper proposes a novel hierarchical data fusion technique for the non-destructive testing (NDT) and condition assessment of timber utility poles. The new method analyses stress wave data from multi-sensor multi-excitation guided wave testing using a hierarchical data fusion model consisting of feature extraction, data compression, pattern recognition and decision fusion algorithms. The proposed technique is validated on guided waved at a of in-situ timber poles. The actual health states of these poles are known from autopsies conducted after the testing, forming a ground-truth for supervised classification. In the proposed method, a data fusion level extracts the main features from the sampled stress wave signals using power spectrum density (PSD) estimation, wavelet packet transform (WPT) and empirical mode decomposition (EMD). These features are then compiled to a feature vector via real-number encoding and sent to the next level for further processing. Principal component analysis (PCA) is also adopted for feature compression and to minimise information redundancy and noise interference. In the feature fusion level, two classifiers based on support vector machine (SVM) are applied to sensor separated data of the two excitation types and the pole condition is identified. In the decision making fusion level, the D-S evidence theory is employed to integrate the results from the individual sensors obtaining a final decision. The results of the in-situ timber pole testing show that the proposed hierarchical data fusion model was able to distinguish between healthy and faulty poles demonstrating the effectiveness of the new method.

Author Keywords: Condition assessment, timber poles, hierarchical fusion, support vector machine, D-S evidence theory, non-destructive testing

[1]Technical Officer, School of Civil Environmental Engineering, Faculty of Engineering and Information Technology, Univ. of Technology Sydney, NSW 2007, Australia. E-mail: yang.yu@uts.edu.au
[2]Lecturer, School of Civil Environmental Engineering, Faculty of Engineering and Information Technology, Univ. of Technology Sydney, NSW 2007, Australia. E-mail: ulrike.dackermann@uts.edu.au
[3]Professor, School of Civil Environmental Engineering, Faculty of Engineering and Information Technology, Univ. of Technology Sydney, NSW 2007, Australia. E-mail: jianchun.li@uts.edu.au
[4]Research Fellow, School of Engineering, Faculty of Science, Engineering and Built Environment, Deakin Univ., VIC 3220, Australia. E-mail: mahbube.subhani@deakin.edu.au

Introduction

Timber utility poles are used all over the world for power and communication distribution due to their low cost and practicality. In Australia, an estimated number of more than seven million timber poles form part of the country's infrastructure (Nguyen et al. 2004). Among these, about five million poles are used for communication and power supply and are worth more than $10 billion. For pole management and maintenance, the Australian government spends every year approximately $50 million to ensure reliability of the network and to avoid potentially disastrous pole failures. The conventional methods for pole assessment are sounding, visual inspection and core drilling (Tansasoiu et al. 2002). These methods, however, are subjective techniques, which highly depend on the experience and skills of the inspector. More importantly, none of these methods are able to provide reliable assessment of the underground sections of the poles, which are indeed the most critical and vulnerable in terms of structural safety. These drawbacks severely compromise the inspection and maintenance management of timber poles. According to (Nguyen et al. 2004), around 30,000 electricity poles are annually substituted in the Eastern States of Australia, although more than 80% of the substituted poles still maintain in a healthy condition, resulting in a large waste of natural resources and money. In addition, research studies have shown that while the present averaged serviceable lifespan of timber utility poles is around 35 years, the life expectancy can be prolonged to more than 75 years if appropriate maintenance and inspection techniques are adopted (Stewart 1996).

In the last two decades, various NDT techniques have been developed to evaluate the integrity and condition of pile structures such as deep foundation piling. Among these techniques, stress wave methods based on guided wave testing such as the impulse response (IR) method (Davis and Dunn 1974), the sonic echo (SE) method (Paquet 1968; Steinbach and Vey 1975; Lin et al. 1991; Van and Middendorp 1980) and the bending wave (BW) method (White and Ross 2014; Qian and Mita 2005) are well established in the pile testing industry. For these tests, an excitation force is applied and the structural responses are measured using a sensor installed on the top of pile structure. By analysing the reflexogenic signals, the health condition of the pile including the underground section can be evaluated. Although some of these stress wave methods have been adopted for the condition assessment of timber poles, most of this work is research-based and many challenges still have to be overcomed before accurate and reliable assessment can be achieved (Dackermann et al. 2014a; Krause et al. 2014; Li et al. 2012; Subhani et al. 2013).These challenges include complicated wave propagation in timber, related to the complexity of the timber material with anisotropic characteristics and uncertainties from natural defects and deterioration, as well as unknown soil conditions and environmental factors such as temperature and moisture fluctuations influencing the wave propagation. As a result, traditional and newly developed signal processing methods often fail to fully interpret wave patterns and to produce accurate and reliable condition assessment. Challenges further stem from practical field testing conditions, where sensor measurements are influenced by noise submerging the actual wave propagation information. Furthermore, if testing data originates only from one source, it is non-inclusive and may by subject to errors from operators and environmental factors. Hence, to achieve comprehensive and reliable pole assessment, multiple types of signals should be recorded and analysed. Furthermore, using a multi-sensor system instead of only a single sensor or device can facilitate the higher identification accuracy. In such case, however, sensors installed at different positions of the pole may result in conflicting assessment results, making a final decision difficult.

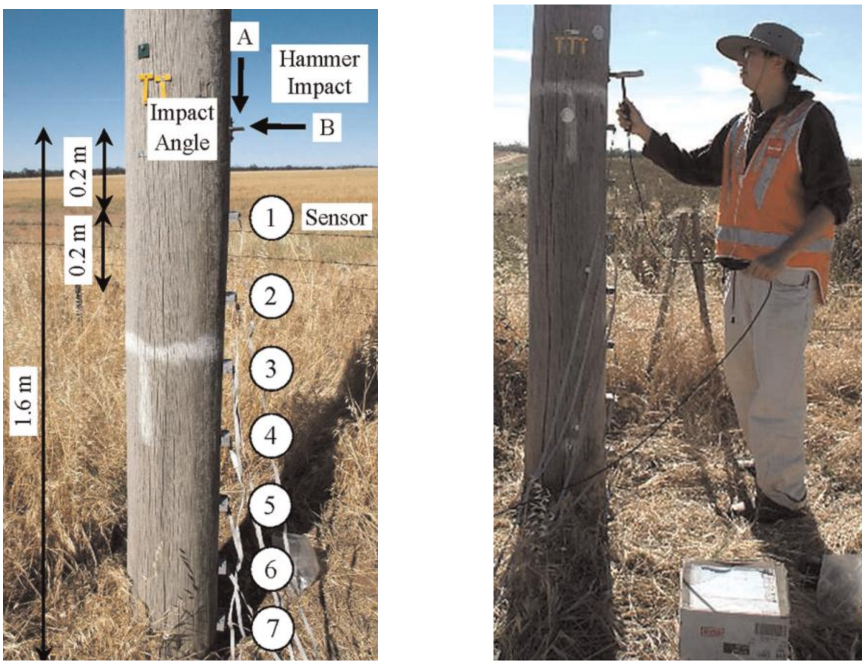
Aiming at solving some of the challenges above, this paper proposes a novel hierarchical data fusion model, combined with multi-sensor and multi-excitation guided wave testing, to analyse stress wave signals for the condition assessment of timber poles. In the data analysis level, power spectrum density (PSD) estimation, wavelet packet transform (WPT) and empirical mode decomposition (EMD) are used to extract signal features from the stress wave data. Next, a feature vector is formed, which is fed into a support vector machine (SVM) classifier for pattern recognition. In order to enhance the classification accuracy and prevent a slow convergence rate, a genetic algorithm (GA) is employed to optimize the classifier parameters. Furthermore, the sigmoid function is adopted to transform the standard outputs of SVM into the confidence probability, realizing the objective assignments of basic probability assignment function. In the decision level, the D-S evidence theory is adopted to fuse the initial identification results from different sensors in the testing system and to make a final decision. In-situ field testing data is used to verify the feasibility and performance of the proposed hierarchical model.

The remainder of this paper is organized as follows: the second section gives background information on the field testing and the hierarchical data fusion model; in the third section, the related data analysis algorithms are described in detailed together with feature analysis and model setup of the stress wave signals. The fourth section presents the performance results of the proposed model for condition assessment of timber poles using the experimental field data. Finally, a conclusion is drawn in the last section.

Guided wave testing and hierarchical data fusion model

Guided wave testing

To validate the proposed method, guided wave testing was conducted on eight in-situ timber poles that were scheduled for decommissioning. After the testing, the poles were dismembered to determine their actual health states. For the guided wave testing, an impact hammer, seven accelerometers and a data acquisition system were employed. The experimental field setup is depicted in Fig. 1. For the testing, a hammer impact was induced at a height of 1.6 m in either longitudinal direction (using an impact angle) to generate primarily longitudinal waves or in transverse direction to generate bending waves. The impact hammer used was a PCB model HP 086C05 of sensitivity 0.24mV/N. The responses of the pole were captured by an array of sensors, which were installed on the side of the pole with a uniform spacing of 0.2 m as illustrated in Fig.1 (a). The sensors were low cost dual-axis piezoresistive accelerometers of model ADXL320, with a sensitivity range between 154 and 194 mV/g and a frequency bandwidth from 0.5 to 2.5 kHz. All the sensors were embedded within the plastic cases, which were drilled to the tested poles. The data acquisition system was a mid-range 8 channel system with 13-bit 4M sample/second per channel of model NI PCI-6133. For the testing, the sampling frequency was set to 1 Hz with 0.2 s sampling time. Each test was repeated five tests, i.e. for each pole five hammer strikes were executed in longitudinal direction and five in transverse direction. More details on the conducted experimental setup and testing can be found in (Dackermann et al. 2014b).



(a) Experimental field setup (b)Testing execution

Fig. 1.Experimental field setup and guided wave testing execution

Hierarchical data fusion model

The proposed hierarchical data fusion model is shown schematically in Fig.2. This model is divided into three analysis levels, i.e. a data level, a feature level and a decision level. In the first level, dominate signal features are extracted minimising information redundancies and disturbances such as environmental noise and wave dispersion. In this level, three different signal processing methods (PSD, WPT and EMD) and PCA are employed to isolate information sensitive to the health condition of the pole structure. In the second level, a state-of-the-art SVM classifier is constructed to intelligently analyse these feature parameters, facilitating the initial evaluation of the pole condition. In the third level, influences from different sensor locations, excitation types and training samples on the condition assessment are addressed. Here, according to the probability outputs of the SVM, the basic probability assignment (BPA) of each proposition of the pole condition discernment frame is obtained. The initial recognition result can be regarded as independent evidence and all results including conflicting evidences are combined using evidence combination rule, solving the problem of conflicting and inaccurate identification of the pole condition using only one SVM classifier. The entire model employs a hierarchical structure, in which the results from the former level are used as the inputs for the next level. Accordingly, this method can separately realize multiple levels of information processing, guaranteeing the robustness and accuracy of assessment results.

Methodology

Data level fusion based on multi-feature extraction

Stress-wave-based condition assessment of timber poles is essentially a pattern recognition problem. Hence, signal feature extraction is a key element in the data processing and its performance is closely related to the accuracy of the final assessment results. While there are many feature extraction methods available, each approach has its limitations. Therefore, in the proposed method, three different analytical methods are employed in parallel to extract key features of the stress wave signals, i.e. PSD (Gangopadhyay et al. 1989), WPT (Sun and Chang 2002) and EMD (Huang et al. 1998).

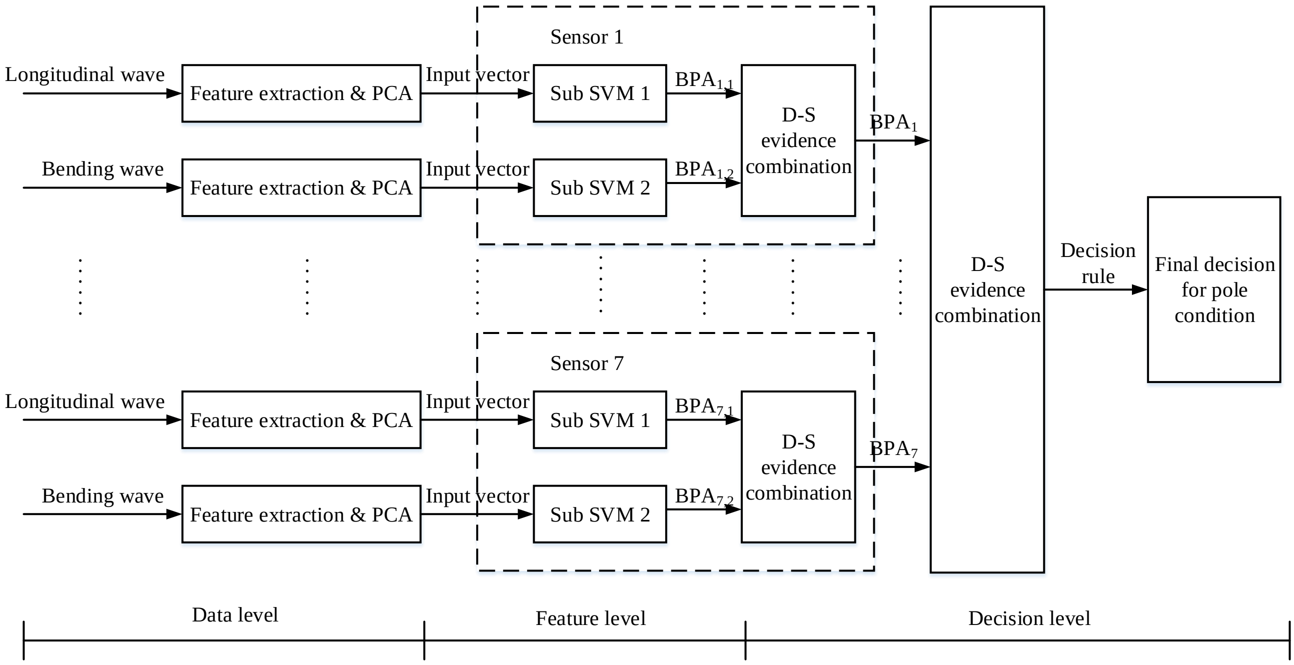


Fig. 2. Hierarchical data fusion model for pole condition assessment

Power spectrum density (PSD) feature extraction

Typical stress wave signals captured from guided wave testing have nonlinear and random characteristics and are polluted with noise. These characteristics can be seen in Fig. 3, which displays stress wave signals of one intact pole from sensor 1 with longitudinal and bending wave excitations. Generally, PSD estimation is an effective way to analyse this type of signals. Up to present, various signal PSD estimation techniques have been reported including parametric and nonparametric methods. Compared with nonparametric methods, parametric ones are apt to provide better results especially when the analysed signal length is relatively short, which is always regarded as short quasi-stationary sequence. In this work, a parametric approach using autoregressive (AR) coefficient estimation is utilised to transform the time-series stress wave signals into a series of real-valued variables based on the assumption that signals can be obtained from a time-series model of a random process. As well, AR model is able to model stress wave signals as the output of a linear m-order AR model combined with zero-mean Gaussian white noise and its specific expression is given by (Fugate et al. 2001):

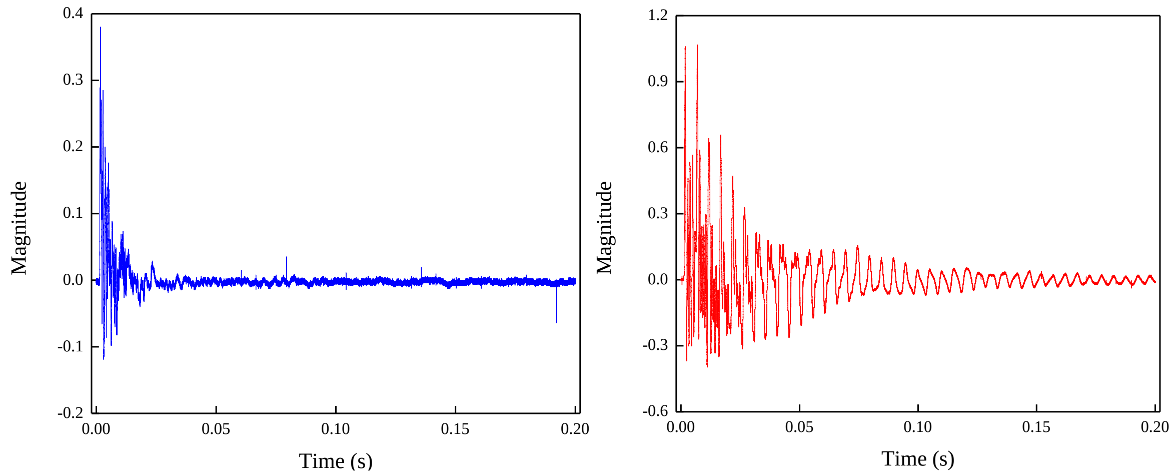
$$y(n) = -\sum_{i=1}^{m} a(i) y(n-i) + e(n) \tag{1}$$

where $y(n)$ is the discrete time-history response that model the stress-wave signals, $a(i)$ is an AR coefficient and $e(n)$ is the time-series sequence of a Gaussian white noise process.

The Burg method is utilized here to estimate the AR parameters of the signals, which is mainly based on Levinson-Durbin recursion and least square criterion. The order m in the AR model is an important parameter to be identified from the signals and describes a trade-off relationship between adding resolution and declining model accuracy. Here, the trial-and-error method is used to identify the optimal model order via minimizing the Akaike information criterion (AIC) function (Marple 1987):

$$\min \quad AIC[m] = N \cdot \ln(\tau) + 2m \tag{2}$$

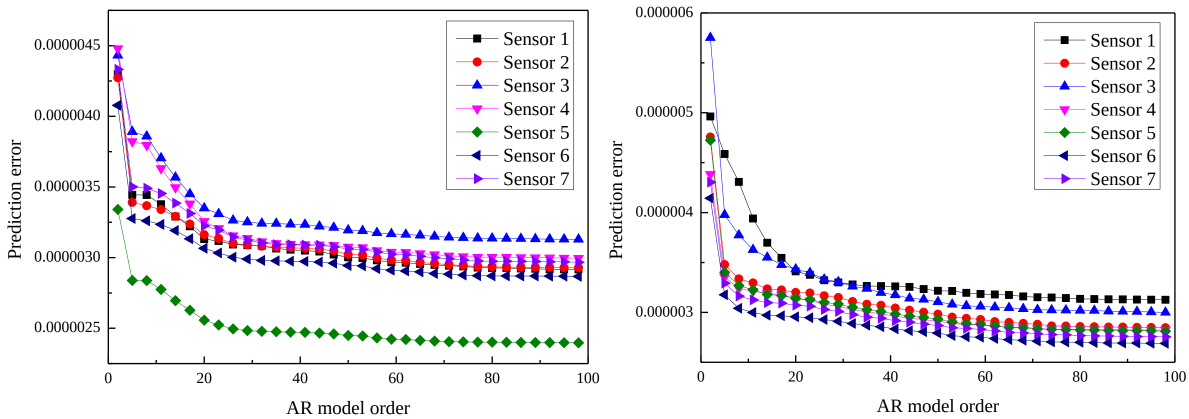
where m denotes the model order of the AR system, N denotes the total number of sampling points, and τ denotes the variance estimation of the white noise input to the AR model for order m.



(a) Longitudinal wave excitation (b) Bending wave excitation

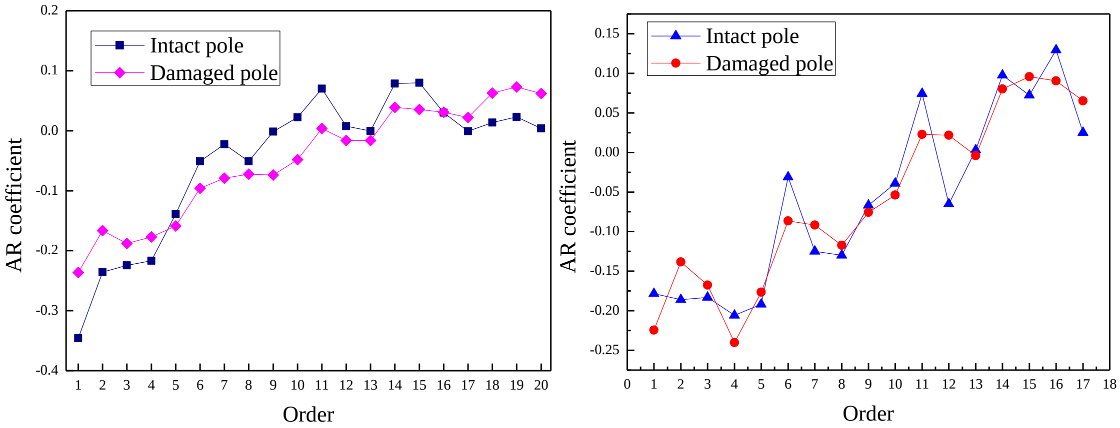
Fig. 3. Typical stress wave signals of an intact pole from sensor 1.

Fig. 4 shows the relationship between variance estimation and AR model order for both longitudinal and bending waves obtained from the sensor array. It can clearly be observed that the variance estimation shows a decreasing trend with an increasing model order. This decreasing trend reduces once the order exceeds the knee points. In general, a higher order will lead to a high-dimensional feature vector which may influence the classification accuracy of the SVM. Therefore, the knee points of AIC curves are selected as the optimal order in this work, i.e. *m*=20 for longitudinal wave and *m*=17 for bending wave excitation.



(a) Longitudinal wave data (b) Bending wave data

Fig. 4. Mean AIC responses



(a) Longitudinal wave data (b) Bending wave data

Fig. 5. AR coefficients for both intact and damaged poles

Fig. 5 showsa typical case of the resulting AR coefficients calculated using the parametric method for longitudinal and bending wave signals for an intact and a damaged timber pole. These results show that only tiny AR coefficient differences can be found between intact and damaged poles, which verifies the importance and necessity of multi-sensor extraction for condition assessment of timber poles.

Wavelet packettransform(WPT) feature extraction

The second feature extraction method employed is WPT, which is a further development of the wavelet transform and can provide a broader scope of probabilities for the analysis of the sampling signals(Sun and Chang 2002). In WPT, both low frequency and high frequency elements are separatedbased on successive time fixation of frequency sub-bands produced by a tree structure of high-pass and low-pass filtering operation. The process of WPT is schematically depicted in Fig. 6, in which an original time series signal $S_{0,0}(n)$ with length N in the first level is split into a low frequency and a high frequency signal as follows (Sun and Chang 2002):

$$S_{1,0}(n) = \sum_{i=0}^{N-1} g(i) S_{0,0}(n-i) \tag{3}$$

$$S_{1,1}(n) = \sum_{i=0}^{N-1} h(i) S_{0,0}(n-i) \tag{4}$$

where$S_{1,1}(i)$ and $S_{1,0}(i)$ denote the details and approximations in the first level, and $h(i)$ and $g(i)$ denote the response functions of high-pass and low-pass filters, which are dependent on the selection of the wavelet function. To guarantee the time fixation of every frequency band and improve the frequency resolution, the outputs of the filters are reduced to half of the signal length in the upper level at the end of every filtering phase (Yen and Lin 2000).

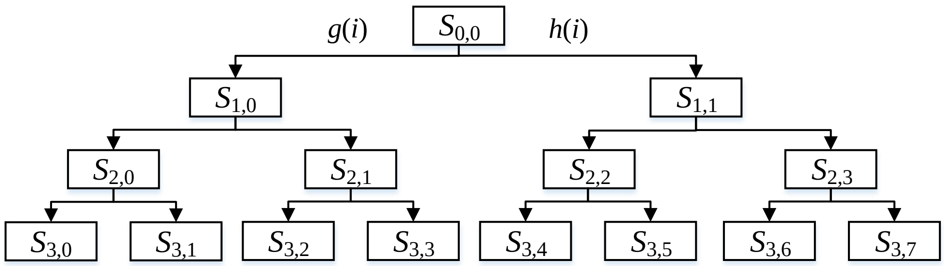
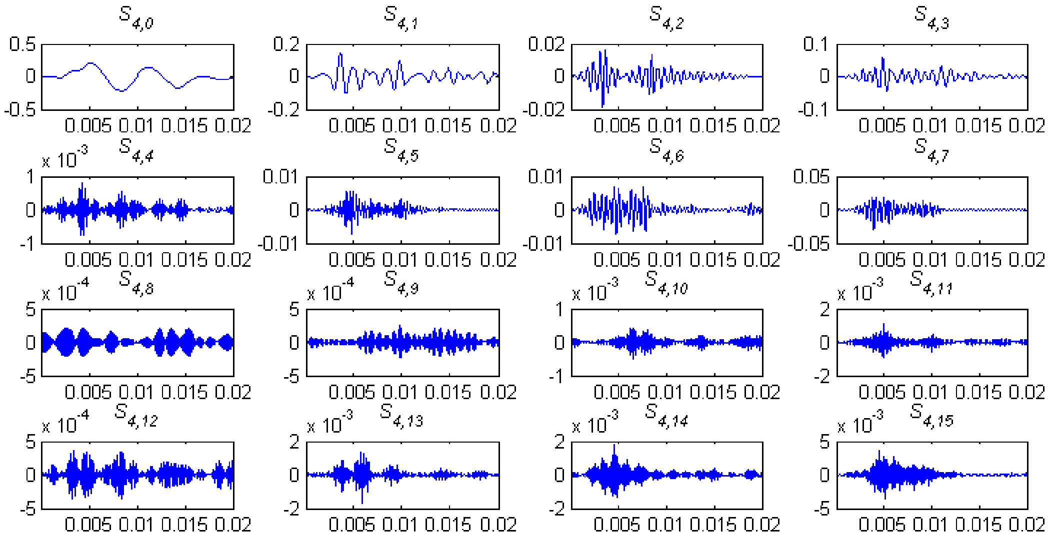


Fig. 6. Tree structure of WPT

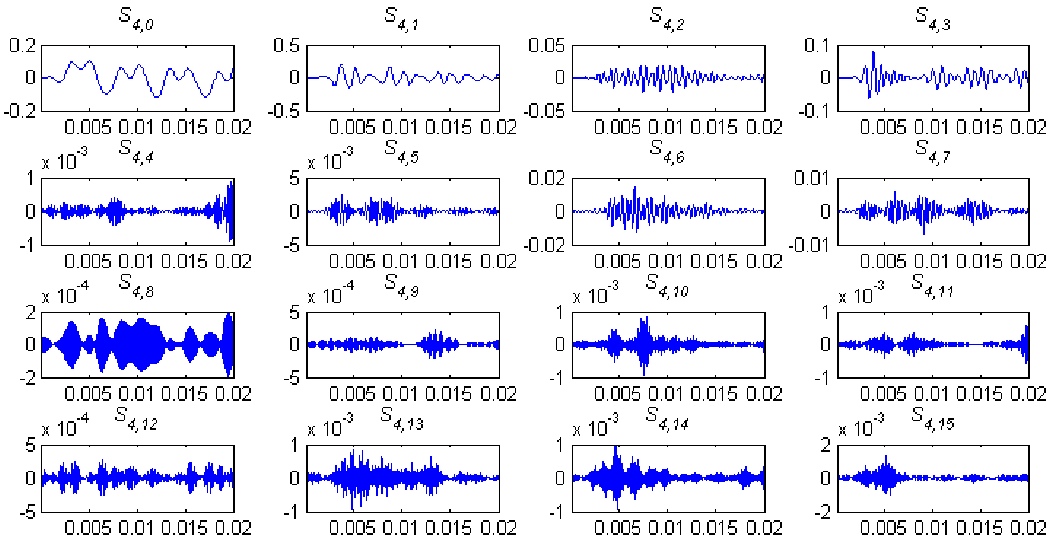
Previous studies haveshown that very good classification results can be achieved if the BiorSplines 6.8 wavelet function is selected. This is mainly due to its good orthogonality and symmetry together with high vanishing moments. Fig.7 shows the WPT reconstructed bending wave signals of the first 16 frequency bands after 10 layers of decomposition.It is noticed that in the first 2 bands there is no remarkable amplitude difference between intact and damaged poles. However, starting from the third frequency band, clear differences are present between the stress wave signal frequency compositions of the different transform scales. This characteristic is mainly reflected in energy and can be expressed by:

$$E_{k,j} = \sum_{i=1}^{N} [S_{k,j}(i)]^2 \tag{5}$$

where j=3,…,16 , and N is the length of reconstructed signal. Fig. 8 shows the sectional energy features of WPT signals from intact and damaged poles. It is noticeable that the signal energy of the fourth frequency band from the intact pole is clearly higher than that of the damaged pole. Hence, these energy features of WPT scale space can be used as feature parameters for timber pole condition assessment.

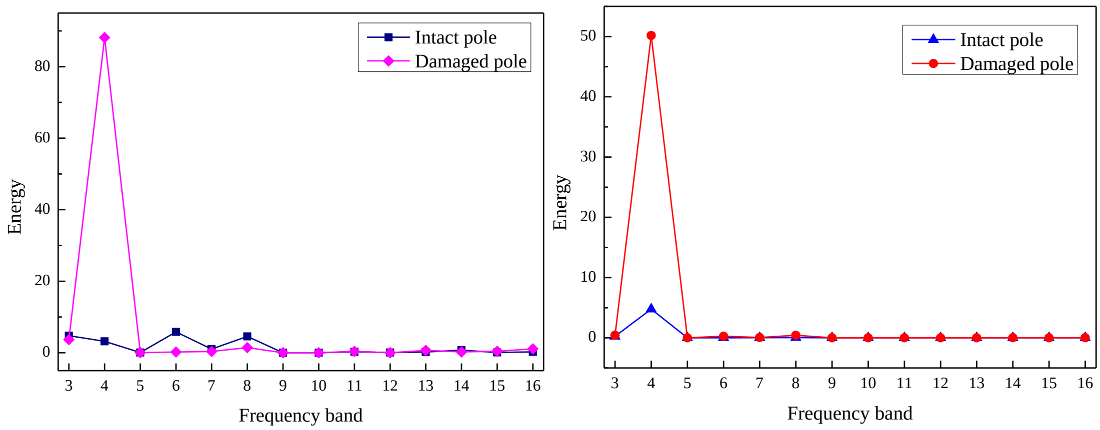


(a) Intact pole



(b) Damaged pole

Fig. 7. WPTof bending wave signals



(a) Longitudinal wavedata (b) Bending wave data

Fig. 8. Comparison of WPT signal energy features between intact and damaged pole

Empirical mode decomposition (EMD) feature extraction

EMD is an effective and adaptive method for analysing non-stationary and nonlinear signals and was first proposed by Huang (1998). Using this method, a complex signal can be decomposedinto a collection of intrinsic mode functions (IMFs) and a residue.Since the decomposition of the signal is mainly dependent on the information contained in the signal, the IMF number is always limited and its components gained via decomposition are steady (Reddy et al. 2015; Rezaei and Taheri 2010; Yang and Tavner 2009). The IMFs generally satisfy two conditions: 1) in the entire signal sequence, the number of extreme points and crossing through zero points should be the same or the difference no more than 1; 2) at any time point, the averaged value of the envelop constructed by the local maximum and minimum points is 0. The procedure of EMD can be described as follows.First, identify all extreme points of the signal $x(t)$ and interpolate these minimum and maximum points into the lower and upper envelop of the signal using the cubic spline method. Second, calculatethe averaged value of the lower and upper envelop, which is denoted as $m_1(t)$, thus the detail $d_1(t)$ can be extracted by (Huang 1998):

$$d_1(t) = x(t) - m_1(t) \tag{6}$$

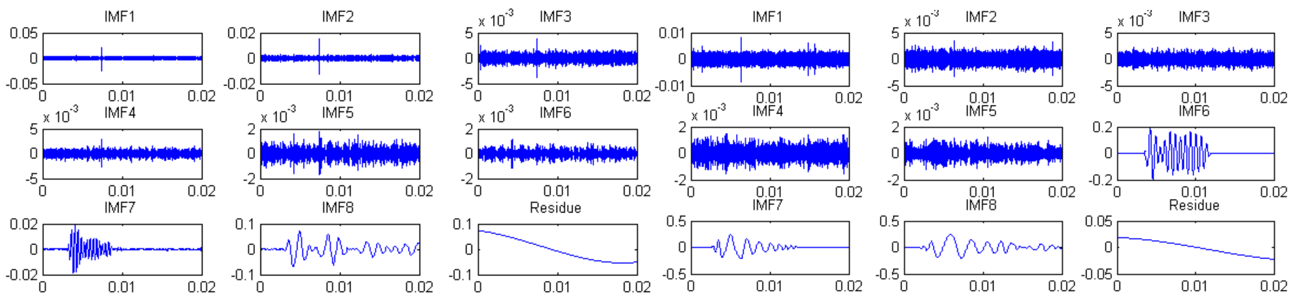
If $d_1(t)$ dissatisfies the conditions of IMF, it will be regarded as a new signal and all the steps are repeated until the condition is satisfied. At this point, the detail $d_1(t)$ is an effective IMF, denoted as $c_1(t)$ that is the highest frequency component in theoriginal signal. Third, the difference $r_1(t)$ is obtained by subtracting $c_1(t)$ from$x(t)$ (Huang 1998):

$$r_1(t) = x(t) - c_1(t) \tag{7}$$

Then take $r_1(t)$ as a new signal and repeat all the steps to obtain $c_2(t)$, $c_3(t)$,…,$c_n(t)$ and $r_2(t)$, $r_3(t)$,…,$r_n(t)$. If $r_n(t)$ or $c_n(t)$ meets the pre-set stopping criterion, and $r_n(t)$ becomes a monotone function, the cycle is terminated. Finally, the following expression is obtained (Huang 1998):

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{8}$$

where $c_i(t)$ is the ith IMF and $r_n(t)$ is the residue. $c_1(t)$ contains the highest frequency while $c_n(t)$ has the lowest frequency, representing the tendency of $x(t)$.



(a) Intact pole (b) Damaged pole

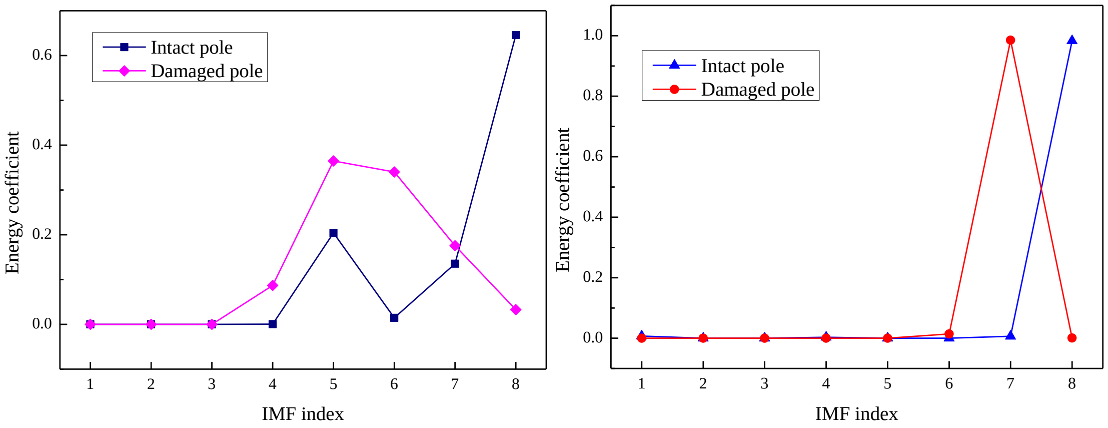
Fig. 9. EMD of bendingwave signals

After EMD decomposition, the obtained IMFs represent a group of steady signals on the feature scale and every IMF has its own unique energy information. Moreover, it has been found that the signal energy mainly focuses on the first eight layers, which means that the first eight IMFs can completely portray the signal energy feature. Fig.9 shows the first eight layers and the residue of EMD decomposition of bending wave signals for anintact and a damaged timber pole.Based on these results, the energy features of the EMD decomposition areobtained and the energy coefficient of each IMF is given by:

$$Q_i = \frac{E_i}{\sum_{i=1}^{8} E_i} \tag{9}$$

$$E_i = \sum_{n=1}^{N}[c_i(n)]^2 \tag{10}$$

where N is the signal length. The comparative results of energy coefficients of the first eight IMFs between an intact and a damaged pole are displayed in Fig.10. It is clearly seen that for both signal types, the last three IMFs have obvious distinguishable patterns for the intact and the damaged case. This shows the feasibility of usingthe energy coefficients of IMFs as a feature vector for pole condition identification.



(a) Longitudinal wave data (b) Bending wave data

Fig. 10. Comparison of energy coefficients of IMFs between intact and damaged poles

Construction of feature vector

After feature extraction of the stress wave signals, the final step in the data levelinvolves the construction of a single vector containing the signal state information to be used as input for the SVM for classification learning and testing. Here, three types of signal features (AR model coefficients, energy values of WPT and energy coefficients of EMD) are combined together with a class label by real-number encoding. Fig. 11 shows the specific encoding process. According to the analysis above, there is a total of 44 indicators in the vector for longitudinal wave signals and 41 indicators for bending wave signals.

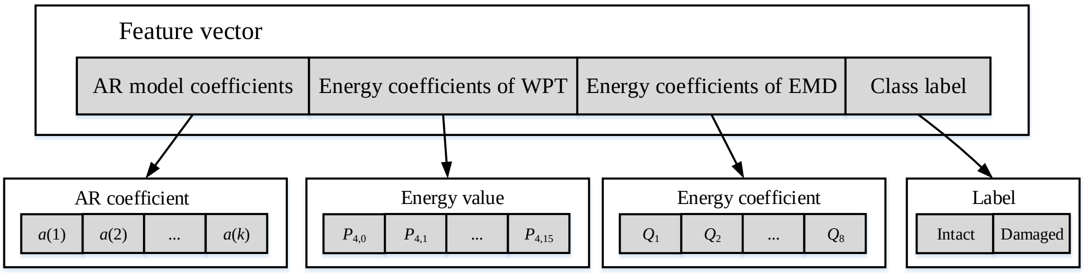
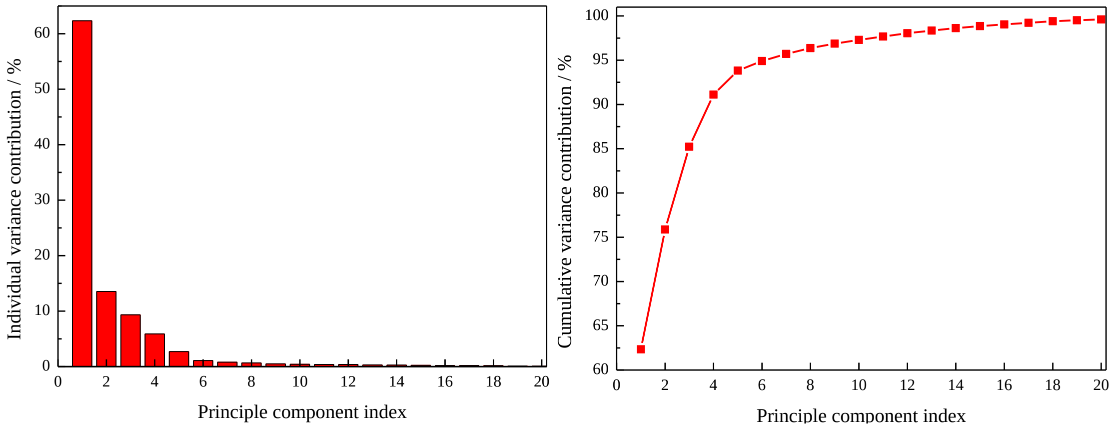


Fig. 11. Encoding for single feature vector

Although multiple feature extraction is able to provide comprehensive information on the signal features, some of these features may be redundant or polluted with noise. If the full feature vector is directly usedas input for the SVM classifier, the classification accuracy and generalization capacity will be greatly affected. As a result, the number of indicators should be reduced to obtain an optimal classification model.

In this work, before beingused for training in the SVM classifier, the extracted feature vector is pre-processed adopting principal component analysis (PCA), which is considered as an effective statistical multi-variable data processing method for reducing feature dimensionality. This method is based on the linear transformation of the original data set into a new data set with fewer components, which are termed principal components (PCs) (Kuzniar and Waszczyszyn 2006). Every PC is a linear combination of the components in the original data set, and all are orthogonal to each other, thus setting up an orthogonal basis of data space. By disregarding PCs of small contribution, the dimension of the feature vector is decreased without remarkably influencing the signal information.

In this case, there are a total of 44 components in the vector for longitudinal wave signals and 41 components for bending wave signals, corresponding to 44 and 41 SVM input nodes, respectively. However, such large number of input nodescan cause problems in calculation efficiency and training convergence. Hence, PCA is employed to reduce the dimension size. Here, for each sensor in the array, there are 40 samples of longitudinal wave signals and 30 samples of bending wave signals. As presented above, after projection, there are 44 and 41 PCs corresponding to longitudinal and bending wave signals, respectively. Fig.12 showsthe individual and cumulative contributions of the first 20 PCs of the feature vector of sensor 1of bending wave signals. It can be seen thatthe first 8 PCs make up more than 95% contribution of the original information (first 15 PCs for longitudinal wave data). So with a 5% loss ofinformation, the vector dimensions can be significantly reduced, which is greatly beneficial toSVM training. The resulting input vector dimensions for the SVMs for the remaining sensors are listed in Table 1.



(a) Individual contribution (b) Cumulative contribution

Fig. 12. PCA results of bending wave signals from sensor 1

Table 1. Feature vector dimensions of other six sensors

Signal type	Feature vector dimension					
	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Sensor 6	Sensor 7
Longitudinal wave	14	16	17	15	12	15
Bending wave	10	5	7	8	12	7

Feature level fusion based on SVM

Theoretical background

SVM is a machine learning algorithm, which is based on the principle of structural risk minimization and kernel-based method (Qu et al. 2013; Ao et al. 2014). SVM utilizes afinite number of samples to train the model toexplore the optimal compromise between generalization performance and classification accuracy, and reveals the distinct benefits in dealing with problems of nonlinearity, small samples and high dimension. Hence it is considered as one of the most effective machine learning algorithms. The principle of SVM for classification is to seek anoptimal line or hyperplane between two groups of data with maximum margin. For a given training set $\{(x_i, y_i)_{i=1}^{l}\}$, where x_i denotes the input vector and y_i is the class label, the characteristic space of SVM is represented by (Vapnik 1998):

$$f(x) = \omega \cdot x + b \tag{11}$$

where ω is the vector of connecting weight and b is a bias. The optimal classifier can be obtained by calculating the following minimization optimization problem (Vapnik 1998):

$$\begin{aligned} &\textit{Minimize} \quad \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{l}\xi_i \\ &s.t. \quad y_i(\omega \cdot x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \; and \; i = 1, 2, ..., l \end{aligned} \tag{12}$$

where C represents the penalty factor to adjust the balance between the classifier complexity and to minimize the training error, and ζ_i denotes the classification errors. By incorporating a Lagrange function, the above optimization problem canbe written as a dual optimization problem, expressed by (Vapnik 1998):

$$\begin{aligned} &\textit{Maximize} \quad \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{l}\alpha_i\alpha_j y_i y_j (x_i \cdot x_j) \\ &s.t. \quad \sum_{i=1}^{l}\alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C \; and \; i = 1, 2, ..., l \end{aligned} \tag{13}$$

Therefore, the optimal nonlinear decision function can be obtained by solving the above problem, which is expressed by:

$$f(x) = \mathrm{sgn}[\sum_{i=1}^{l}\alpha_i y_i K(x_i, x) + b] \tag{14}$$

where $K(x_i,x)$ is the kernel function, which is used in SVM for nonlinear classification. In this paper, the radial basis function (RBF) is employed as the kernel function due to its excellent performance on nonlinear classification. The expression of RBF is

$$K(x_i, x) = \exp(-\frac{\|x_i - x\|^2}{2\sigma^2}) \tag{15}$$

where σ is the width of RBF.

Construction of SVM classifier

To identify the health condition of the timber poles from the extracted feature vector, fourteen sub SVM classifiers, corresponding to the two signal types (longitudinal and bending waves) and seven sensors, are built. The PCs selected from the feature vector in Section 3.1.4 are used as the inputs of the SVM classifier while the output of the model is the pole condition. There are two condition types denoted as 0 and 1, where 0 means ‘damaged’ and 1 represents ‘intact’. The classifier parameters are also selected, which are related to the model generalization performance and classification accuracy. Here, the genetic algorithm (GA) is adopted to optimize the penalty factor C and the kernel function parameter σ in the classifier model. The classification accuracy is evaluated by leave-one-out cross-validation using the fitness function for parameter optimization. The optimization process can be regarded as solving the following maximization optimization problem:

$$\begin{aligned} &\underset{C,\sigma^2}{Maximize} \quad R_{cv}(C,\sigma^2) \\ &s.t. \quad C_{\min} \le C \le C_{\max} \\ &\qquad\quad \sigma^2_{\min} \le \sigma^2 \le \sigma^2_{\max} \end{aligned} \tag{16}$$

where $R_{cv}(C,\sigma^2)$ denotes the ratio between correct classification samples and the total samples. For each sub SVM classifier, the optimization process can be divided into the following steps:

Step 1. Initialize the chromosome number N, maximal iteration number$T$, search range of parameter (C,σ^2), crossover probability P_c and the mutation probability P_m. In this case, N=20, T=100, C_{min}=0, C_{max}=100, σ^2_{min}=0, σ^2_{max}=100, P_c=0.7 and P_m=0.01.

Step 2. Chromosome encoding for parameter (C,σ^2). Randomly generate initial chromosome and set initial iteration t=0.

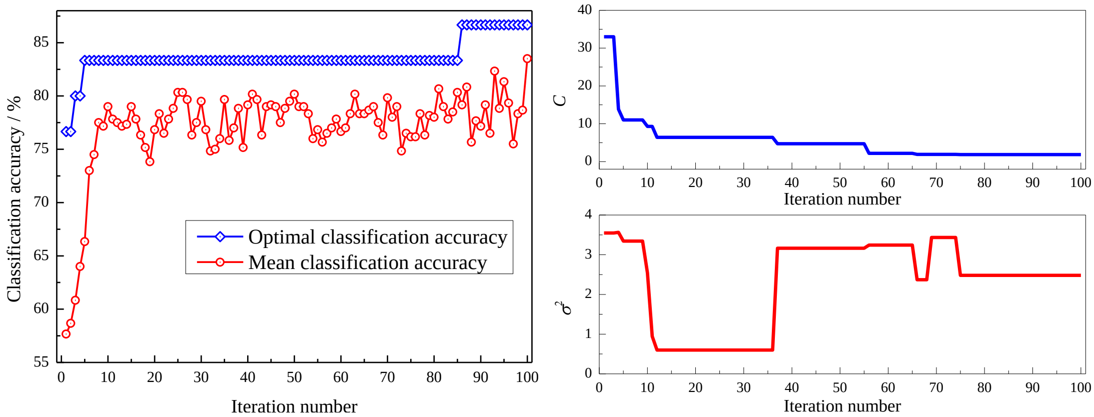
Step 3. Calculate the individual fitness, i.e. $R_{cv}(C,\sigma^2)$.

Step 4. Select the part of the chromosome to produce the new chromosome by roulette wheel strategy.

Step 5. Carry out the crossover and mutation operation to generate the new chromosome.

Step 6. If the termination rule is not satisfied and $t \le T$, go to Step 2.

Fig.13 shows the parameter optimization process of sub SVM of bending wave signals from sensor 1. Fig. 13(a) showsthe algorithm convergence during the iteration while Fig. 13(b) gives the variance trend of each parameter. It is apparent that the best fitness gradually increases with the addition of iteration number though some fluctuations still exists in the average fitness (classification accuracy). Moreover, it is found that the parameter C can quickly arrive at its optimum compared with σ^2, which requires more iteration. Table 2 gives the classification results for both wave types of all sub SVMs usingleave-one-out cross-validation. It is observed that the classification accuracies of sub SVMs based on bending wave data are higher than those of longitudinal wave data. This may bedue to two reasons: 1) due to the shorter wavelength in the induced frequency band, the bending wave is more susceptible to smaller damage compared to the longitudinal wave; 2) the bending wave has more prominent radial and longitudinal displacement components on the surface of the tested pole. Therefore, the bending wave is able to reflect the damage scenarios more accurately compared to the longitudinal wave since the radial component is more susceptible to the real damages.



(a) Convergence rate (b) Parameter optimization

Fig. 13. Parameter optimization process of sub SVM of bending wave from Sensor 1

Table 2. Classification results of all sub SVMs using leave-one-out cross-validation

Sensor Number	Classification accuracy		
	Longitudinal wave	Bending wave	Mean
Sensor 1	80% (32/40)	86.67% (26/30)	82.86% (58/70)
Sensor 2	87.5% (35/40)	90% (27/30)	88.57% (62/70)
Sensor 3	87.5% (35/40)	96.67% (29/30)	91.43% (64/70)
Sensor 4	77.5% (31/40)	83.33% (25/30)	80% (56/70)
Sensor 5	62.5% (25/40)	80% (24/30)	70% (49/70)
Sensor 6	92.5% (37/40)	93.33% (28/30)	92.86% (65/70)

Sensor 7	82.5% (33/40)	86.67% (26/30)	84.29%(59/70)
Mean	81.43% (228/280)	88.1% (185/210)	

Probability output of SVM classifier

The standard outputs of SVM classifiers are based on hard decision and are dependent on avoting method. However, for practical engineering applicationswith nonlinear classification problems, a soft decision with probability outputs is required to provide an objective evaluation with different categories. In this work, the sigmoid function is adopted to map the outputs of SVM to the range of [0, 1] to obtain posterior probability outputs. The specific expression is given by:

$$p(x) = \frac{1}{1 + \exp[Af(x) + B]} \tag{17}$$

where *A* and *B* are two factors to adjust the flexibility of the sigmoid function and their optimal values can be obtained by solving the following minimization problem:

$$min \qquad -\sum_{i=1}^{l}[t_i \ln(p_i) + (1 - t_i)\ln(1 - p_i)] \tag{18}$$

where $p_i = \frac{1}{1 + \exp[Af(\boldsymbol{x}_i + B)]}$ and $t_i = \frac{y_i + 1}{2}$, y_i is the classification label of sample *i*.

Decision level fusion based on D-S evidence theory

The evidence theory is usedto manipulate and model incomplete, inaccurate, uncertain and even conflicting information. It was first presented by Dempsterbased on the notion of upper and lower probabilities, and then consummated by Shafer (Dampster 1967; Shafer 1976). This method has been successfully applied in many application fields such as data fusion, image analysis, pattern recognition and decision making (Talon et al. 2014). For probability estimation, the evidence theory adopts a belief function, which is built usingevent probability and constraint expression withthe novel percept notions, i.e. unknown or uncertainty.

Frame of discernment

In the evidence theory, a non-null set, which contains N exhaustive and exclusive hypotheses, is defined as the frame of discernment. The power set is denoted by 2^{θ}. In this work, the frame of discernment for pole condition assessment is $\theta = \{A_1, A_2\}$, where A_1 and A_2 represent the damaged and intact conditions, respectively.

Basic Probability Assignment

The basic probability assignment function (BPA) of the proposition A is a mappingfrom 2^{θ} to the interval [0, 1], which meets the following relationship (Dampster 1967; Shafer 1976):

$$\begin{cases} m(\Phi) = 0 \\ \sum_{A \subseteq \theta} m(A) = 1 \end{cases} \tag{19}$$

where $m(A)$ is called the BPA of event A, which denotes a certain piece of evidence. Generally, the probability output of sub SVM classifier can be regarded as the BPAs, which means that $p(x)$ denotes the BPA value of the damaged condition, $1-p(x)$ denotes the BPA value of the intact condition. Also, how to assign the uncertainty is an important issue to be solved. Here, the error upper bound of SVM classification identification is introduced and adopted: if a group of training samples are able to be separated by an optimal hyperplane, the classification error upper bound of testing samples is the ratio of mean number of support vectors in the training set to the total training sample number:

$$E(P_{error}) = \frac{N_{sv}}{N_t - 1} \tag{20}$$

where N_t and N_{sv} denote the total training sample number and mean number of support vectors, respectively. The above expression represents the sample uncertainty by SVM classification, corresponding to the uncertaintiy degree Θ in the frame of discernment θ. In order to satisfy the requirement that the summation of BPAs equals to 1, the probability outputs of two categories (damaged and intact) should be multiplied by the coefficient $1-E(P_{error})$. Accordingly, the mathematical expression of BPAs by SVM classifier is given as follows:

$$\begin{cases} m(A_1) = p(x)(1 - \dfrac{N_{sv}}{N_t - 1}) \\ m(A_2) = [1 - p(x)](1 - \dfrac{N_{sv}}{N_t - 1}) \\ m(\Theta) = \dfrac{N_{sv}}{N_t - 1} \end{cases} \tag{21}$$

Evidence combination rule

Suppose m_1 and m_2 are BPA functions from different evidence sources in the same frame of discernment, and its focal elements are B_i and C_j. The rule of combination of B_i and C_j is given as follows (Dampster 1967; Shafer 1976):

$$\begin{cases} m(A) = 0, & A = \varnothing \\ m(A) = m_1 \oplus m_2 = \dfrac{\sum_{B_i \cap C_j = A} m_1(B_i) m_2(C_j)}{1 - K} \end{cases} \tag{22}$$

where $K = \sum_{B_i \cap C_j = \varnothing} m_1(B_i) m_2(C_j)$ denotes the conflict degree among different information sources. Generally, the combination result of nBPA functions ($m_1,\ldots,m_n$) in the same frame of discernment is given by (Dampster 1967; Shafer 1976):

$$m = m_1 \oplus m_2 \oplus \cdots \oplus m_n \tag{23}$$

Decision making

In this paper, the maximum trust degree approach is used to make a final decision in accordance with the evidence combination results:

$\exists A_1$ and $A_2 \subset \theta$, $m(A_2)$=max{$m(A_k)$, $A_k \subset \theta$ and $A_k \neq A_1$}. If the following expression is satisfied:

$$\begin{cases} m(A_1) - m(A_2) > \delta_1 \\ m(\Theta) < \delta_2 \\ m(A_1) > m(\Theta) \end{cases} \tag{24}$$

A_1 is the final result, where δ_1 and δ_2 denote the thresholds. In this work, δ_1=0.3 and δ_2=0.3.

Case study

To evaluate the performance of the hierarchical data fusion model on the condition assessment of timber utility poles, all cases of timber poles are utilized to testing the model using leave one pole out method, in which five cases related to the same pole are taken out in turn as the unknown validation samples and the rest are used as the training set. Tables 3 and 4 show the initial estimations of all sub SVM classifiers of the two randomly selected cases. According to the description in 3.3.4, when $m(A_1)-m(A_2) > 0.3$, $m(\Theta)< 0.3$ and $m(A_1) > m(\Theta)$, the assessment result is $A_{1;}$ when $m(A_2) - m(A_1) > 0.3$, $m(\Theta) < 0.3$ and $m(A_2) > m(\Theta)$, the assessment result is A_2. Otherwise, there is no recognition result from the system. Consequently, it can be seen thatforthe damaged pole case, the assessment results from the sensors 2, 4, 5, 6 and 7 dissatisfy the decision rule. The same problem exists for the sensors 1, 3 and 5 for the intact pole case. It is believed that there are two possible reasons contributing to this phenomenon. One reason may be linked to noise sensitivity, sensor or measurement errors, which cause uncertainties in the input feature vectors of the sub SVM classifiers. The other reason is that the training samples are so limited that they are not able to include all the possible conditions, which may lead to the identification errors in sub SVM classifiers.

Tables 5 and 6 give the first-layer evidence combination results of two sub SVM classifiers of the same sensor for the two cases. For the damaged pole case, the support probability of the proposition A_1 (damaged condition) has increased to 0.8491 at sensor 1 while the support probability of A_2 (intact condition) and uncertainty degree declined to 0.0372 and0.1137, respectively. This result meets the decision rule, which means that the assessment result of sensor 1 is A_1 'damaged pole'. Similarly, the other six sensors have the same assessment results. Therefore, after first-layer evidence combination, all the sensors in the testing system give the same result forA_1, which conforms to the practical condition of the pole. However, for the intact pole, because $m_1(A_1) > m_1(A_2)$ and $m_3(A_1) > m_3(A_2)$, the assessment results from sensors 1 and 3 are still in contradiction with that of the other five sensors. Therefore, it is difficult for the system to make a final decision.

Table 3. Initial recognition result by SVM classifiers for the damaged pole case

BPA	A_1	A_2	Θ
$m_{1,1}$	0.5734	0.1702	0.2564
$m_{1,2}$	0.6911	0.1020	0.2069
$m_{2,1}$	0.2940	0.4624	0.2436
$m_{2,2}$	0.7708	0.0395	0.1897
$m_{3,1}$	0.5858	0.1578	0.2564
$m_{3,2}$	0.7696	0.0063	0.2241
$m_{4,1}$	0.4170	0.3138	0.2692
$m_{4,2}$	0.7860	0.0243	0.1897
$m_{5,1}$	0.2747	0.4689	0.2564
$m_{5,2}$	0.6806	0.1125	0.2069
$m_{6,1}$	0.2770	0.4666	0.2564
$m_{6,2}$	0.7886	0.0217	0.1897
$m_{7,1}$	0.3245	0.4319	0.2436
$m_{7,2}$	0.7491	0.0440	0.2069

Table 4. Initial recognition result by SVM classifiers for the intact pole case

BPA	A_1	A_2	Θ
$m_{1,1}$	0.4836	0.2600	0.2564
$m_{1,2}$	0.5158	0.2773	0.2069
$m_{2,1}$	0.0203	0.7361	0.2436
$m_{2,2}$	0.0241	0.7862	0.1897
$m_{3,1}$	0.1292	0.6144	0.2564
$m_{3,2}$	0.7225	0.0534	0.2241
$m_{4,1}$	0.1816	0.5492	0.2692
$m_{4,2}$	0.0506	0.7597	0.1897
$m_{5,1}$	0.4625	0.2811	0.2564
$m_{5,2}$	0.0437	0.7494	0.2069
$m_{6,1}$	0.2144	0.5292	0.2564
$m_{6,2}$	0.0033	0.8070	0.1897
$m_{7,1}$	0.0646	0.6918	0.2436
$m_{7,2}$	0.0191	0.7740	0.2069

Tables 7 and 8 show the second-layer evidence combination results of all seven sensors for the two cases. From the tables it can be seen that after the second-layer combination, the uncertainty degree declines to 0 and the support probabilities of the propositions as the final decisions for the two cases ascend to 100%, which is in agreement with the practical conditions of the cases. The results verify that compared to theresults from the single SVM classifier, the confidence probability of the final decision is greatly improved through the two-layer evidence combination.

Table 5. First-layer evidence combination for the damaged pole case

BPA	A_1	A_2	Θ
m_1	0.8491	0.0372	0.1137
m_2	0.7785	0.0627	0.1588
m_3	0.8852	0.0020	0.1128
m_4	0.8481	0.0198	0.1321
m_5	0.6386	0.1802	0.1812
m_6	0.7880	0.0365	0.1755
m_7	0.7779	0.0608	0.1613

Table 6. First-layer evidence combination for the intact pole case

BPA	A_1	A_2	Θ
m_1	0.6659	0.1925	0.1416
m_2	0.0008	0.9253	0.0739
m_3	0.5084	0.1787	0.3129
m_4	0.0192	0.8738	0.1070
m_5	0.0712	0.7420	0.1868
m_6	0.0015	0.8964	0.1021
m_7	0.0022	0.9120	0.0858

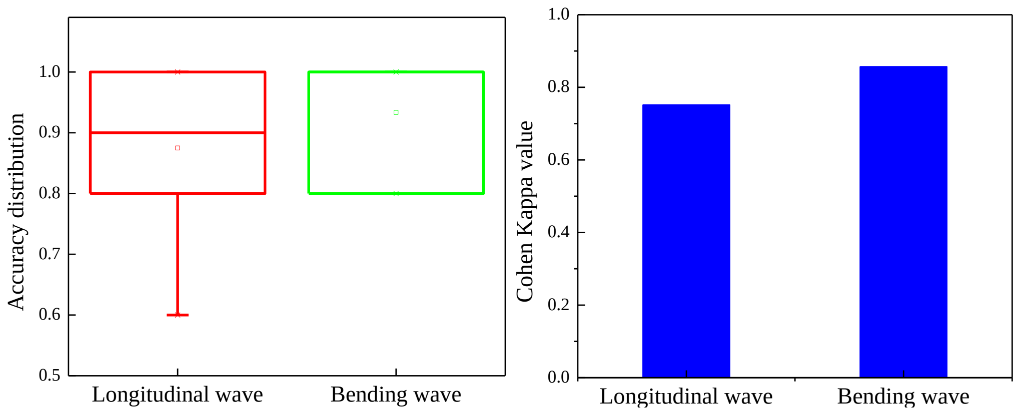
Table 7. Second-layer evidence combination for the damaged pole case

BPA	A_1	A_2	Θ
m	1.0000	0	0

Table 8. Second-layer evidence combination for the intact pole case

BPA	A_1	A_2	Θ
m	0	1.0000	0

Fig. 14 gives the statistical accuracy analysis results of the proposed model for timber poles with longitudinal and bending wave excitations. Fig. 14 (a) displays theclassification accuracy distributions for two excitation cases. The results clearly illustrate that bending wave excitation outperforms longitudinal wave in the respect of classification accuracy with average value of 93.33% although the latter also could arrive at 87.5%. Fig. 14 (b) shows the related Cohen Kappa values of the data fusion model with different excitation types. Generally, Kappa value is used to estimate the distribution of the forecast labels, which could not be expressed by the classification accuracy. Kappa value always changes between 0 and 1 with the maximal value representing the best forecast with all values located at the diagonal line in the confusion matrix (Witten et al. 2011). Similar to results in Fig. 14 (a), Fig. 14 (b) also verifies that bending wave cases also present a better result than the longitudinal wave cases, which accords with the previous analysis that captured signal features from bending wave are more sensitive to the damage scenario than that of longitudinal wave.



(a) Classification accuracy distribution (b) Cohen Kappa value

Fig. 14. Statistical indicators of the proposed model with different wave signal excitations

Conclusion

This paper presented a novel method for the health condition assessment of in-situ timber utility poles based on a hierarchical data fusion model. In the proposed method, first, stress wave signals were recorded in a sensor array using two types of wave excitation (longitudinal and bending wave). Second, for each sensor, the stress wave data was analysed using AR coefficients, wavelet packet energies and energy coefficients of IMFs. The derived parameters were then used to form a feature vector and PCA was applied for data compression. Third, for each sensor, two sub SVM classifiers were built upfor initial estimation of the pole condition. Fourth, to improve the identification accuracy of the classifierand to obtain a final decision, GA was employed to optimize the two main parameters in the model. The experimental results from in-situ timber pole testing demonstrated that the proposedhierarchical model is able to greatly improve the identification accuracy.The support probability of right proposition was increased from 0.7886 to 1. This prevents the problem of difficult decision making and satisfies the requirement of condition assessment of timber utility poles in engineering application and management. In this work, the SVM model is set up based on the off-line learning, which is not suitable for real-time structural health monitoring of timer poles. In the future work, the model will be realised in the on-line way to update the model in real-time for the higher accuracy. Furthermore, except the application in the pole condition assessment, the proposed hierarchical model can also be developed for leakage detection of pipeline networks, fault diagnosis of machines and damage detection of bridge and building structures, in which the multi-sensor system is essential for capturing multi-source data information to implement the monitoring tasks. Moreover, sometimes the information provided by only one-type sensor may be incomplete or inaccurate due to the self-deficiency and environmental factors. Therefore, multi-type-sensor system and multi-source information fusion algorithm will be also investigated to realise their application in civil engineering in future.

Acknowledgement

This work is supported by the Australian Research Council and our industry partner Ausgrid, through the Linkage Project (LP110200162). The authors are also thankful to the anonymous reviewers for their constructive comments.

References

Ao, H., Cheng J., Zheng J., and Truong, T. K. (2014).“Roller bearing fault disgnosis method based on chemical reaction optimization and support vector machine.”*J. Comput. Civ. Eng.*, 10.1061/(ASCE)CP.1943-5487.0000394, 04014077.

Dackermann, U., Crews, K., Kasal, B., Li, J., Riggio, M., Rinn, F., and Tannert, T. (2014a).“In situ assessment of structural timber using stress-wave measurements.”*Mater.Struct.*, 47(5), 787-803.

Dackermann, U., Skinner, B., and Li, J. (2014b). “Guided wave-based condition assessment of in situ timber utility poles using machine learning algorithms.”*Struct.Health.Monit.*, 13 (4), 374-388.

Davis, A. G., and Dunn, C. S. (1974).“From theory to field experience with the non-destructive vibration testing of piles.”*Proc.Instn. Civ.Engrs.*, 57(4), 571-593.

Dempster, A. P. (1967). “Upper and lower probabilities induced by a multi-valued mapping.”*Ann. Math. Statist*,38(6), 325-339.

Fugate, M. L., Sohn, H., and Farrar C. R. (2001).“Vibration-based damage detection using statistical process control.”*Mech. Syst. Signal Pr.*, 15(4), 707-721.

Gangopadhyay, A., Comillon, P., and Jackson, L. B. (1989).“Autoregressive modeling for the spectral analysis of oceanographic data.”*J.Geophys. Res.*, 94(C11), 16215-16226.

Huang, N. E., Shen, Z., Long, S. R., Wu, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N., Tung, C. C., and Liu., H. H. (1998). “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis.”*Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Science*, 454(1971), 903-995.

Krause, M., Dackermann, U., and Li, J. (2014).“Elastic wave modes for the assessment of structural timber: ultrasonic echo for building elements and guided wave modes for pole and pile structures.”*J. Civil. Struct. Health Monit.*, 5(2), 221-249.

Kuzniar, K. and Waszczyszyn, Z. (2006).“Neural networks and principal component analysis for identification of building natural periods.”*J. Comput. Civ. Eng.*, 20(6), 431-436.

Li, J., Subhani, M., and Samali, B. (2012).“Determination of embedment depth of timber poles and piles using wavelet transform.”*Adv.Struct. Eng.*, 15(5), 759-770.

Lin, Y., Sanasalone, M., and Carino, N. J. (1991).“Impact-echo response of concrete shafts.”*Geotech. Test. J.*, 14(2), 121-137.

Marple, S. L. (1987). *Digital spectral analysis with applications*, 1st ed, Englewood Cliffs, NJ: Prentice Hall.

Nguyen, M. N., Foliente, G. C., and Wang, X. M. (2004).“State-of-the-practice & challenges in non-destructive evaluation of utility poles in service.”*Key. Eng. Mat.*, 270-273, 1521-1528.

Paquet, J. (1968). “Etude vibratorie des pieux en beton: reponseharmonique et impulsionelle, application au controle.”*Annales de L'Institut Technique du Batimentet Des Travaux Publics*,21(245), 789-803.

Qian, Y., and Mita, A. (2005).“Damage identification of structures based on pattern classification using limited number of sensors.”*Proc. of the 5th International workshop on structural health monitoring*, Stanford, CA, USA.

Qu, D., Li, W., Zhang, Y., Sun, B., Zhong, Y., Liu, J., Yu, D., and li, M. (2013).“Support vector machines combined with wavelet-based feature extraction for identification of drugs hidden in anthropomorphic phantom.”*Measurement*, 46(1), 284-293.

Reddy, D. M., Krishna, P., and Sathesa.(2015). “Innovative method of empirical mode decomposition as spatial tool for structural damage identification.”*Struct.Control.Hlth.*, 22(2), 365-373.

Rezaei, D., and Taheri, F. (2010).“Damage identification in beams using empirical mode decomposition.”*Struct.Health.Monit.*, 10(3), 261-274.

Shafer, G. (1976). *A mathematical theory of evidence*, Princeton University Press, London.

Steinbach, J., and Vey, E. (1975).“Caisson evaluation by stress wave propagation method.”*J.Geotech. Eng.*, 101(4), 361-378.

Stewart, A. H. (1996). *Wood poles life span: What you can expect*, IEEE/PES Expo (Edition).

Subhani, M., Li, J., Samali, B., and Yan, N. (2013).“Determination of the embedded lengths of electricity timber poles utilizing flexural wave generated from impacts.”*Aus. J.Struct. Eng.*, 14(1), 85-96.

Sun, Z., and Chang, C. C. (2002).“Strcutural damage assessment based on wavelet packet transform.”*J.Struct. Eng.*, 128(10), 1354-1361.

Talon, A., Curt, C., and Boissier, D. (2014).“Performance assessment based on evidence theory and fuzzy logic: application to building and dam performance.”*J. Comput. Civ. Eng.*, 28(1), 124-133.

Tansasoiu, V., Miclea, C., and Tanasoiu, C. (2002) “Nondestructive testing techniques and piezoelectric ultrasonics transducers for wood and bult in wooden structures.”*J.Optoelectron. Adv. M.*, 4(4), 949-957.

Van, K. H., and Middendorp, P. (1980).“Interpretation of results from integrity tests and dynamic load tests.” Proc., of the International Seminar on the Application of Stress Wave Theory in Piles, Stockholm, Sweden, 217-223.

Vapnik, V. N. (1998). *Statistical learning theory*, John Wiley & Sons, Inc.,New York.

White, R. H., and Ross, R. J. (2014).*Wood and timber condition assessment manual: second edition*, General Technical Report FRL-GTR-234. Madison, WI: U.S. Department of Agriculture. Forest Service, Forest Products Laboratory.

Yang, W., and Tavner, P. J. (2009).“Empirical mode decomposition, an adaptive approach for interpreting shaft vibratory signals of large rotating machinery.”*J. Sound Vib.*, 321(3-5), 1144-1170.

Yen, G. G., and Lin, K. (2000).“Wavelet packet feature extraction for vibration monitoring.”*IEEE T. Ind. Electron.*, 47(3), 650-667.