An Effective Feature Extraction Method of Quantitative Recognition of Ultrasonic Echo for Bonding Quality of Composite Sheet

Dong Li[1,*], Bateer[2], Bingqing Gao[3]

[1] *College of information engineering, Inner Mongolia University of Technology, China*

[2] *SRA, Inner Mongolia University of Technology, China*

[3] *Hohhot power supply section of China Railway Hohhot Bureau Group Co., Ltd., China*

* Corresponding author

Abstract

In this paper, a set of effective ultrasonic echo feature extraction method is proposed for the quantitative identification of thin plate composite bonding quality, which opens up a new research idea for the research of ultrasonic echo feature extraction for quantitative identification of thin plate bonding quality, and also lays a foundation for further quantitative identification of bonding structure debonding degree of thin plate. In addition to the time and frequency domain analysis, the empirical mode time-frequency analysis method is used to highlight the time-frequency characteristics of the echo signal, improve the resolution of the signal analysis, and expand the space for the echo signal feature extraction. The feature validity evaluation theory of fault diagnosis is introduced to evaluate the feature validity of high-dimensional signal parameters, and the invalid signal parameters are eliminated to obtain an effective echo signal feature set with low dimension. Using the exhaustive search strategy of feature selection theory, geometric distance measurement and information measurement evaluation criteria, and through the classification verification of support vector machine classifier, the effective feature vector for quantitative identification of adhesive quality is obtained.

Keywords: Ultrasonic echo; feature extraction; feature selection

I. Introduction

Ultrasonic testing plays a key role in improving the product quality of high-tech industries such as electric power, metallurgy, automobile, railway and modern architecture [1,2]. It is not only an important means to realize quality control, save raw materials, improve technology and increase productivity in industrial production, but also an important detection means for safe and efficient operation of equipment [3,4]. At present, there are many researches on ultrasonic testing, but most of them focus on the qualitative identification and evaluation of defects. The research on how to realize the quantitative identification of defects is still in the initial stage. Elena et al. developed a new signal post-processing algorithm for reconstructing joint region in order to overcome the problems related to rough surface and non-parallel layer for layered titanium / carbon fiber composite, and realized defect detection using the proposed technology [5]. Pen et al. Used continuous wavelet transform to realize defect detection and three-dimensional reconstruction of high silica phenolic resin cladding, and realized dynamic thermal tomography visualization. It shows that the echo frequency components reflected from different multi-layer interfaces are different. It is necessary to use time-frequency domain analysis algorithm to realize echo tomography [6]. Weiwei et al. proposed an improved empirical change mode decomposition (Ecmd) method for ultrasonic echo signal processing. Ecmd was used to preprocess the coal rock interface echo, and the weight coefficient optimization iterative method was introduced. The asymmetric Gaussian model was used to fit the partially distorted echo envelope, which can effectively reduce the calculation error. The multi-layer interface of coal and rock can be effectively identified [7]. Chen proposed a denoising algorithm based on the combination of self-adaptive complete set empirical mode decomposition (Ceemdan) and fast independent component analysis (Fast-ICA) algorithm for ultrasonic detection of transformer bushing lead, which can better retain the effective information such as the starting position of echo, and better extract the state of the lead in the bushing, which has

certain practical significance [8]. Li et al. Proposed an adaptive filtering modeling method of snow ultrasonic echo signal. According to the characteristics of ultrasonic echo signal, combining modeling method and modern digital signal processing technology [9]. In this paper, aiming at the quantitative identification of the bonding quality of the thin plate bonding structure, the ultrasonic echo method is used to build the experimental system for the ultrasonic detection of the bonding quality of the thin plate. A set of scientific ultrasonic echo feature extraction method is put forward, which opens up a new research idea for the research of ultrasonic echo feature extraction for quantitative identification of thin plate bonding quality, and also lays a foundation for further realizing the quantitative identification of ten grades of adhesive structure of thin plate.

II. Ultrasonic Testing Method for Bonding Quality

2.1 Ultrasonic echo detection method

Ultrasonic bonding quality testing methods mainly include the following: pulse echo method, Lamb wave method, ultrasonic oblique incidence method, acoustic ultrasonic method, electromagnetic ultrasound, laser ultrasound etc [10-15]. Pulse echo method is a mature ultrasonic propagation detection method. This method is usually vertical incidence, mainly using the reflection of ultrasonic short pulse on the debonding interface to detect defects. When the pulse echo method is used to detect the debonding structure, the ultrasonic wave reflects and transmits many times between the layers of medium, and the echo contains a lot of defect information. Ultrasonic pulse echo multiple reflection method is widely used in aerospace industry to detect the debonding interface quality of solid rocket motor. The principle diagram of echo method is shown in the following figure. Due to the unique advantages of pulse echo method in engineering application, this paper studies the pulse echo method.
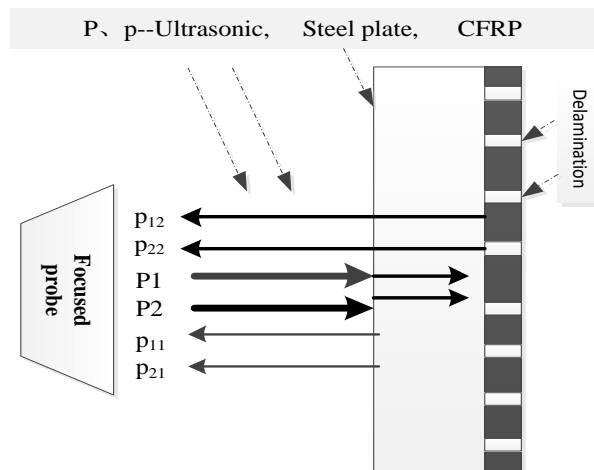

Fig 1: Schematic diagram of ultrasonic testing

2.2 Analysis and design of ultrasonic testing experimental system

In order to adjust the ultrasonic testing system at any time, the ultrasonic testing system is designed and built in the form of module circuit. The system is mainly composed of ultrasonic probe, ultrasonic transmitting circuit module, ultrasonic receiving circuit module, high precision data acquisition module and PC. The hardware structure of the system is shown in Figure 2.
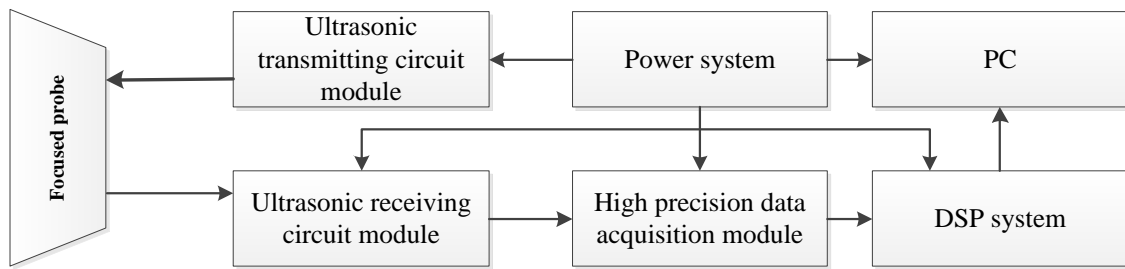
Fig 2: Block diagram of ultrasonic detection system

The 2.5z14n and 5z14n ultrasonic straight probes with nominal frequencies of 2.5MHz and 5MHz in Shantou Institute of ultrasonic instruments are compared and analyzed.After analyzing the spectrum of the excitation pulse, it is found that the pulse excitation signal with the same amplitude and pulse width can stimulate the probe with 2.5m low frequency more fully, but the attenuation of the transmitted wave is slow, which has great interference on the echo signal. However, the high-frequency probe attenuates rapidly, and the echo has less suppression, and the echo carries more information. Therefore, the ultrasonic probe with 5m center frequency is selected for the follow-up experiments.

In addition, this paper selects a transmitting circuit which uses the high-speed switching characteristics of transistors and capacitor energy storage to generate high-power pulses to generate the sharp pulse excitation needed by the ultrasonic probe. At the same time, in order to avoid damaging the ultrasonic system caused by high-voltage transmitting pulse, the necessary receiving isolation stage or limiting circuit is inserted in front of the echo signal acquisition system.The data acquisition module uses a new high-speed acquisition module techv developed by Beijing Dasheng Technology Co., Ltd_AD212H_The digital signal processing system adopts el-dsp-e300 DSP experimental system.In order to meet the needs of research, ten samples with different debonding grades were made in Hexi chemical machinery company of China90% debonding sample block. As shown in Figure 3. The Measured echo of 2.5M\5M probe on No-load load and the state of 50% debonding was obtained by the ultrasonic testing system.
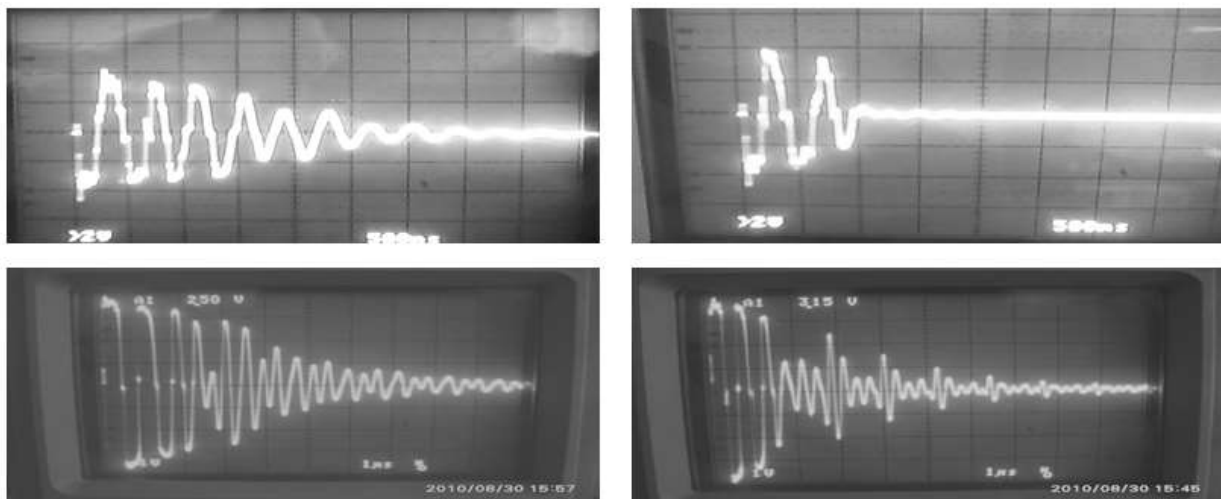

Fig 3: Measured echo of 2.5M\5M probe on No-load load and the state of 50% debonding

III. Feature Analysis and Extraction of Ultrasonic Echo Signal

The ultrasonic testing experimental system is manually operated. The hand shaking will affect the coupling degree between the ultrasonic probe and the surface of the tested block, which directly affects the ultrasonic projection effect and has a great impact on the echo signal. Therefore, before the ultrasonic echo feature analysis and extraction, in order to effectively balance the impact of hand shaking on the echo signal each sampling, In this

paper, a large number of samples are obtained by equal interval mean. This method can effectively get the real estimation of samples, and provide more reliable data information for the subsequent echo signal feature analysis. The specific method is that if each block sample collects n sample data, the average value of n sample data is taken as an effective sample data, n * m data are collected in turn, and M mean data are obtained as effective original signal data for subsequent feature analysis and extraction. In order to further eliminate the correlation of random interference of continuous sample data, the following formula (1) is shown, where x Is the original sample data, X To get the effective sample estimation.

$$X = \sum_{i=0}^{M-1} x(iN) \tag{1}$$

3.1 time domain and frequency domain ultrasonic echo feature analysis and extraction

Time domain and frequency domain ultrasonic echo feature analysis methods are indispensable tools for quantitative or automatic detection of ultrasonic echo method. The thickness of the steel layer of the thin plate test block is obtained by measuring the thickness of the bonding block d =3.8mm. Check the propagation speed of ultrasonic wave in steel medium $c \approx 6080 m/s$ The echo period is calculated $T = 1.25 us$ Because the 2.5us signal before the echo is affected by the circuit containing the excitation signal and the limited amplitude circuit, and the echo signal attenuation is nearly zero after 10us, the echo data from 2.5us to 10us is clipped for analysis.1000 groups of waveform data are taken from 10 steel rubber thin plate samples, and 100×10 groups of effective standard sample data are obtained by means of sample mean denoising method for subsequent feature analysis.

3.1.1 Time domain ultrasonic echo feature analysis and extraction

The so-called time-domain echo analysis is called the analysis of ultrasonic waveform. The amplitude of dynamic signal is analyzed with time as variable. Through the waveform analysis, the amplitude relations of various physical quantities are given, such as the magnitude of amplitude, the envelope area of waveform, etc. Combined with the mechanism of ultrasonic echo transmission of thin plate bonding structure, four characteristic parameters of echo signal in time domain are defined, namely signal energy error ΔW Signal energy attenuation ratio (SAR), signal zero crossing point NZ and signal positive and negative area difference SS.

Signal energy error ΔW it is described as follows:

$$\Delta W(i) = W(i) - W_b \tag{2}$$

$$W(i) = \sum_{i=1}^{NT} x(i)^2 \tag{3}$$

$$W_b = \sum_{i=1}^{NT} x_b(i)^2 \tag{4}$$

Among $W(i)$ For the first time i Sample data energy, W_b Is the energy value of the standard fully viscous state sample, $x(i)$ Sample the amplitude of the sample echo signal, $x_b(i)$ It is the amplitude sampling of echo signal of standard full viscosity state sample.

Because the ultrasonic probe is used to transmit and receive simultaneously, the ultrasonic echo obtained is essentially a process of repeated excitation and reception by the ultrasonic probe. If ti-j is assumed to be a transmitted wave, then the signal of Tj-l period is the echo signal of Ti-j transmitted wave. Therefore, the signal energy attenuation ratio (SAR) is defined, as shown in the following formula (5):

$$Sar = \sum_{T_1 - T_3} x(k)^2 / \sum_{T_4 - T_6} x(k)^2 \tag{5}$$

As shown in Fig.4 to 7, data1-data10 of each subgraph is the data of ten quantitative grades (100×10groups) of adhesion degree of thin plates ΔW Characteristic parameter data of SAR, NZ and SS signals, y It is the first-order linear fitting result of the signal parameters. It can be seen from the figure that these signal characteristic parameters have certain potential for quantitative identification of debonding grades, but whether they can be used as effective features for quantitative identification of ten debonding grades still needs to be analyzed.
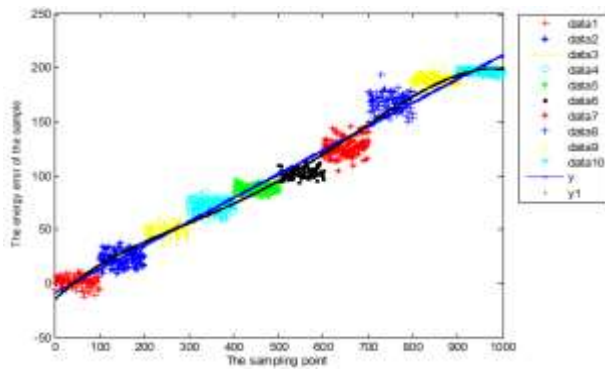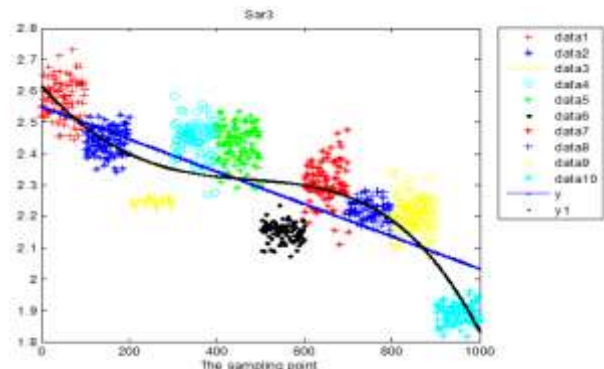

Fig 4: The energy error
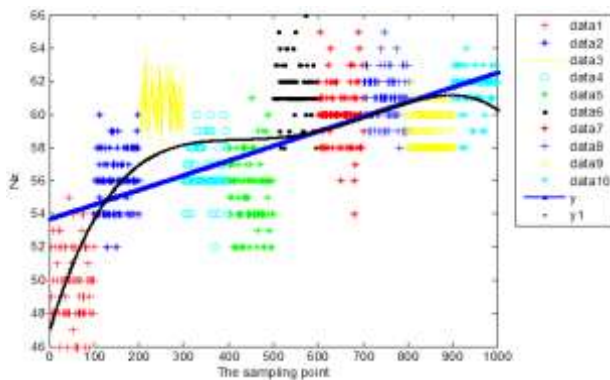

Fig 5: Attenuation of Signal

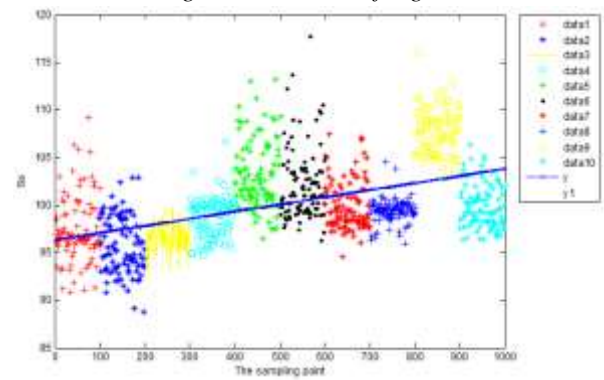*Fig 6: The zero-crossing number of signal*

Fig 7: The difference of signal area

3.1.2 Frequency domain ultrasonic echo feature analysis and extraction

Frequency domain characteristic analysis of ultrasonic echo signal is to apply spectrum analysis to ultrasonic testing technology. Spectrum analysis of ultrasonic echo signal is to analyze various dynamic signals through Fourier transform to frequency domain, which can be obtained more useful information and rules will guide the research work and engineering practice.

Based on the direct analysis of power spectrum curves of ten debonding degree test blocks, combined with the detection mechanism of ultrasonic spectrum analysis technology, four signal parameters, namely, center frequency energy, central frequency value, frequency energy decay factor and low lobe harmonic frequency band energy, are extracted. The specific definitions are shown in table 1

Table 1 definition of frequency domain signal characteristic parameters

Characteristic parameter	definition
Center frequency value	Frequency corresponding to the maximum amplitude of power spectrum
Central band energy	Taking the center frequency as the center, the 0.5m bandwidth signal energy is taken as the central frequency band energy characteristic parameter
Frequency energy decay factor	The ratio of high-frequency band energy to low-frequency band energy with the central frequency as the boundary
Low lobe harmonic	The maximum amplitude of the specific power spectrum corresponds to the next

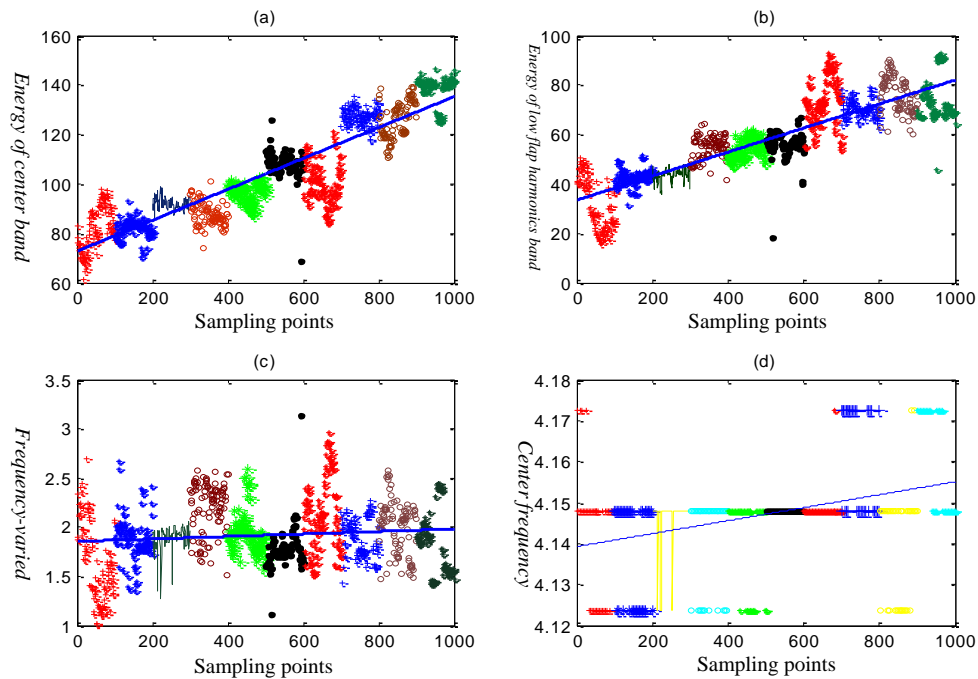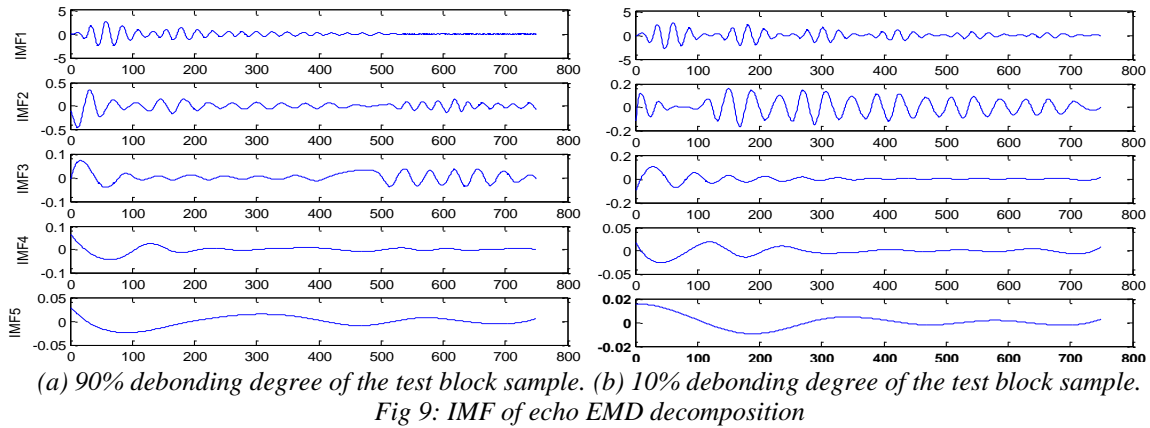| band energy	maximum value of the power spectrum on the lower side, and the corresponding frequency is the energy value of the center 0.5m bandwidth signal

Fig 8: $(a) - (d)$ Energy of center band\Low flap harmonics band\Frequency-varied \Center frequency

As shown in Figure 8 (a) - (d) This is the frequency domain signal characteristic parameter data diagram of ten debonding level samples of central frequency band energy, low lobe harmonic frequency band energy, frequency energy decay factor and center frequency value. The data points are the same as those in time domain analysis.

3.2 Analysis and extraction of ultrasonic echo characteristics in time-frequency domain

In the process of time-frequency signal analysis, this paper introduces empirical mode decomposition (EMD) method, which is a new time-frequency signal analysis method proposed by Norden E. Huang and others in 1998.This method is rarely used in ultrasonic echo defect diagnosis. Ultrasonic echo signals are unstable and nonlinear in nature. EMD decomposition method has good processing effect on nonlinear signals and can highlight the important time-frequency information of ultrasonic echo signals adaptively. The echo signals of 10 sample levels are decomposed by EMD with iterative threshold SD of 0.01, and $100 \times 8 \times 10$ sets of IMF data are obtained. It is found that the fluctuation information of ultrasonic echo signals is mainly concentrated in IMF (intrinsic mode function) 1-5, so only the first five modes are processed and analyzed. As shown in Fig.9 (a)-(b) are the time-domain wave diagrams of the first five IMF components after EMD decomposition of the total viscosity state and 90% debonding degree of the test block sample.

(a) 90% debonding degree of the test block sample. (b) 10% debonding degree of the test block sample.
Fig 9: IMF of echo EMD decomposition

The zero crossing point, signal area, maximum amplitude in time domain and the maximum amplitude, center frequency and signal energy of frequency spectrum are extracted respectively to obtain a 5×6 =30 dimensional eigenvector space. In addition, four signal parameters of Hilbert marginal spectrum, including center frequency, central spectrum amplitude, central frequency band energy and marginal spectrum energy, are calculated in time and frequency domain, and the subsequent feature validity analysis is carried out.

So far, a 42 dimensional set of super echo signal characteristic parameters is extracted in time domain, frequency domain and time-frequency domain. If the feature parameter space of the echo signal is identified quantitatively or searched exhaustively to find the optimal feature subset, the computation is very large. Therefore, the feature validity evaluation theory of fault diagnosis theory is introduced.

3.3 Evaluation and selection of ultrasonic echo characteristics

In order to evaluate the effectiveness of fault diagnosis, it is necessary to define a method to evaluate the effectiveness of fault diagnosis parameters, Analysis of the effectiveness of engineering classification. Three indexes are defined in the paper, i.e. sensitivity, stability, as shown in the table 2.

Table 2 evaluation parameters of ultrasonic echo characteristics

Evaluate index	Definition	Calculation formula
Sensitivity	It is used to reflect the sensitivity of characteristic parameters to the change of debonding degree. Sensitivity of characteristic parameter x to defect grade or defect degree y	$\xi(Y \mid X) = \left| \dfrac{\partial Y}{\partial X} \right|$
Stability	The characteristics of the test equipment and its working conditions are affected. In general, the more stable the characteristic parameters, the better.	$S = \dfrac{1}{n} \sum\limits_{i=1}^{n} \sigma_i$
Recognition rate	The recognition probability of the characteristic quantity to the quantitative debonding grade.	$DR = \dfrac{1}{n} \sum\limits_{i=1}^{n-1} \left| \mu_{i+1} - \mu_i \right| / (\sigma_{i+1} + \sigma_i)$

The above three evaluation indexes can reflect the effectiveness of signal parameters for quantitative identification to a certain extent. As shown in table 3, the data set of effectiveness evaluation indexes of ultrasonic echo signal parameter characteristics are given. With the different recognition methods, the evaluation index can not get the best characteristic parameters strictly, but it can give a good trend of signal parameters to the recognition results. When the invalid signal parameters are eliminated according to the evaluation index of feature effectiveness, it is necessary to combine the specific identification problems and analyze various factors.

Feature validity evaluation method is introduced into the analysis process of echo signal parameters, which can effectively evaluate the effectiveness of signal parameters for quantitative identification of debonding grade. Then, the feature set of echo signal for quantitative identification of debonding grade is obtained. It provides a good data base for subsequent feature subset selection.

IV. Feature selection algorithm

Feature selection can be divided into filter, wrapper and embedded according to whether the evaluation criteria are related to the classifier. Filtering feature selection algorithm is mainly based on subset evaluation criteria for feature selection. In this paper, the filtering feature selection algorithm is used, which will be introduced in detail below. Encapsulation feature selection needs to determine the classifier algorithm in advance, and use the classifier to evaluate the feature subset. The feature of the algorithm is that it is not universal, but has high recognition rate. In addition, it has the disadvantages of high computational cost and long execution time. Embedded feature selection, also known as hybrid feature selection, combines the characteristics of the previous two models and uses different evaluation criteria in different search stages. The most typical algorithms are decision tree algorithm.

4.1 Filtering feature selection

Filtering feature selection mainly depends on the attributes of the feature space itself, and has nothing to do with the selected classifier and learning algorithm. Therefore, the early feature selection mostly belongs to the filtering type, mainly including the first three sub sets evaluation criteria introduced before. Because it obtains feature subsets through the intrinsic properties of feature sets, those feature subsets with higher correlation can obtain better classification results. Moreover, in the process of feature subset selection, the training steps of classifier for feature subset are omitted. Therefore, for feature sets with larger dimensions, it has less computation, but because it is independent of learning algorithm, its accuracy is low, the model is shown in Figure 10.

Table 3 The index of feature validity table of Ultrasound echo

NO	1	2	_3_	_4_	5	6	_7_
ξ	0.0009	0.0003	0.0000	0.0001	0.0006	0.0005	0.0000
S	0.0191	0.0932	0.1123	0.0452	0.0505	0.0619	0.0973
DR	2.0311	0.5166	0.1717	0.1927	0.9493	0.5231	0.4278
NO	8	_9_	10	11	_12_	_13_	14
ξ	0.0003	0.0001	0.0001	0.0001	0.0000	0.0000	0.0002
S	0.1474	0.1087	0.0854	0.1008	0.0773	0.0680	0.1129
DR	0.4303	0.0742	0.1345	0.1231	0.0569	0.1111	0.1730
NO	15	16	17	18	19	20	21
ξ	0.0003	0.0002	0.0002	0.0001	0.0002	0.0003	0.0001
S	0.1170	0.0952	0.1013	0.1114	0.1405	0.1202	0.0886
DR	0.3298	0.3080	0.2306	0.1626	0.1186	0.1392	0.1275
NO	_22_	_23_	_24_	_25_	_26_	_27_	_28_
ξ	0.0001	0.0001	0.0001	0.0002	0.0000	0.0000	0.0000
S	0.1080	0.1077	0.0829	0.1053	0.0507	0.0776	0.0615
DR	0.0808	0.0754	0.0795	0.1000	0.1256	0.0666	0.0581
NO	29	30	_31_	_32_	_33_	_34_	_35_
ξ	0.0002	0.0002	0.0001	0.0001	0.0001	0.0000	0.0000
S	0.1442	0.1249	0.0986	0.1190	0.1062	0.0490	0.0439
DR	0.1215	0.1582	0.1004	0.0618	0.0680	0.0888	0.1626
NO	_36_	_37_	_38_	39	40	41	_42_
ξ	0.0001	0.0000	0.0000	0.0003	0.0001	0.0004	0.0000

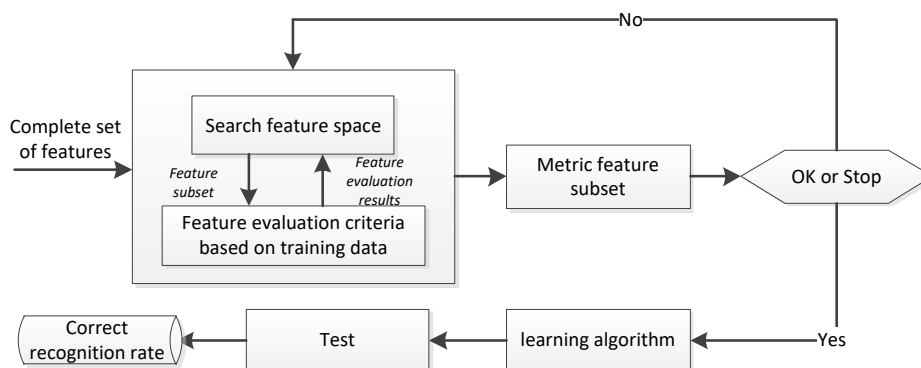| S	0.1572	0.1180	0.0941	0.0725	0.0594	0.0555	0.1718
DR	0.0764	0.0828	0.0869	0.2073	0.2534	0.6250	0.1004

Fig 10: The filter model of feature selection

4.2 Experimental analysis of feature selection

A 20 dimensional feature set is obtained after the effectiveness evaluation of ultrasonic echo signal in Section 3.3.At this time, the characteristic parameters are 1 256 8 10 11 14 15 16 17 18 19 20 21 29 30 39 40 41,They are signal energy error, signal energy attenuation ratio, central frequency band energy, low lobe harmonic band energy, center frequency value, maximum amplitude of imf2-3 fluctuation, zero crossing point of imf1-5, signal area of imf1-3, maximum amplitude of imf1-2 spectrum, marginal spectrum energy, central spectrum amplitude and central frequency band energy of Hilbert marginal spectrum.

Based on the filtering feature selection model, using the exhaustive search method, the evaluation criteria of feature subset select the commonly used geometric distance measurement and information measurement, write a program, can select two groups of feature subsets, through the specific classifier verification, get a set of effective feature subsets, carry on the follow-up recognition research of the subject.

Feature selection based on geometric distance metric.

In an n-dimensional space, there are m categories (W_1 , W_2 ... W_M)Sample set of $\left\{ X_1^{(i)}, X_2^{(i)},........X_N^{(i)} \right\}$, then W_i The intra class distance of a class is defined as:

$$S_w^{(i)} = \frac{1}{N_i} \sum_{k=1}^{N_i} (X_k^{(i)} - m^{(i)})^2 \tag{7}$$

Among them, $m^{(i)}$ express W_i The average value of the sample within the class.

From this, we can know that the intra class distance of the total sample set is:

$$S_w = \sum_{i=1}^{M} P(W_i) S_w^{(i)}$$
$$= \sum_{i=1}^{M} P(W_i) \frac{1}{N_i} \sum_{k=1}^{N_i} (X_k^{(i)} - m^{(i)})^2 \tag{8}$$

In a multidimensional space, there are two categories i and j In the case of Euclidean distance, there are many distance measures between them

$$S_B^{(ij)} = (m^{(i)} - m^{(j)})^2 \tag{9}$$

Then the total inter class distance of all kinds of samples is defined as:

$$S_B = \frac{1}{2} \sum_{i=1}^{M} P(W_i) \sum_{j=1}^{M} P(W_j) S_B^{(ij)} \tag{10}$$

Ruoling $m = \sum_{i}^{M} P(W_i) m^{(i)}$ And M is the population mean value of the sample, then there are:

$$S_B = \sum_{i}^{M} P(W_i)(m^{(i)} - m)^2 \tag{11}$$

Using the above formulas 8 and 11, the intra class distance and inter class distance of feature subset combination of all feature parameters are calculated respectively. Based on the distance measurement criterion of formula 6-7, the best feature subset is selected. The feature subset is a combination of the characteristic parameters of energy error No. 1, the characteristic parameters of zero crossing point of imf1 No. 14 and the characteristic parameters of spectrum maximum value of imf1 No. 29. As shown in Figure 11:
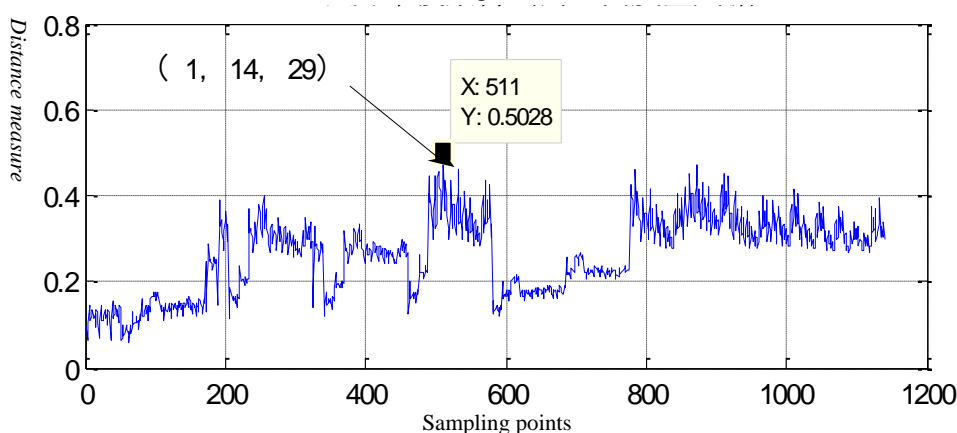

Fig 11: Distance metric of exhaustive search from c20-3

Feature selection based on information measurement maximum correlation and minimum redundancy

When building a classifier, the more features are used, the better. If too many features are used to build the classifier, it will increase the complexity of the classifier and take longer time to train the classifier. Moreover, if the irrelevant or redundant feature quantity is introduced into the classifier, it will affect the recognition accuracy and produce adverse effects, It is necessary to select feature subset with high correlation degree and low redundancy degree from feature set to realize classification. At present, most of the feature selection algorithms only consider the correlation between features and categories, and less attention is paid to the redundancy between features. In this paper, based on the maximum correlation feature selection algorithm, the feature correlation and redundancy are combined, and the feature subset is an effective feature subset filtering out redundant information.

The maximum correlation minimum redundancy algorithm is a correlation measure factor which is characterized by nonlinear correlation mutual information. The judgment of feature redundancy is essentially the consideration of the correlation between two features. If the correlation between the two features is very large, it means that the two features have a lot of repeated information for the classification. The purpose of feature selection is to select the feature subset with the minimum correlation between the features. For those features that are completely related, we call them redundant features. The specific algorithm is as follows:

For n-dimensional feature vector space, feature selection is a feature subset S of M features. For category C, it has maximum dependence on these m features

$$Max\, D(S,c), \quad D = I(\{x_i, i = 1,2....m\}; c) \tag{12}$$

The principle of maximum correlation is essentially the maximization of mutual information between features and target categories, which can be expressed by the following formula:

$$MaxD(s), D = \frac{1}{|S|} \sum_{x_i \in S} I(c, x_i) \tag{13}$$

Among them, D is the mutual information value between features and categories, C is the category of data set, and | s | is the number of feature sets.

Minimum redundancy is a description of the dependency relationship between features. It requires the minimum correlation between each feature attribute, that is, the minimum redundancy principle

$$MinR(S), R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i, x_j) \tag{14}$$

Where R is the mutual information value between features.
The main idea of the maximum correlation and minimum redundancy algorithm is to combine the selection criteria of maximum correlation between features and categories and the selection criteria of minimum redundancy between features. The operator is defined in the algorithm ϕ the formula can be expressed as follows:

$$Max\phi = D / R \tag{15}$$

Finding the largest feature subset of the operator is the allowable subset with maximum correlation and minimum redundancy. As shown in Figure 12 is the D, R data graph of 20 characteristic parameters. The specific calculation method is to first use histogram and empirical data to obtain sample myopia distribution, and then calculate $p(x)$, $p(y)$ and $p(x, y)$ then the mutual information entropy D, R of the required random variables is obtained. As shown in Fig.13, the data graph of information measurement criterion is shown. From the graph, it can be clearly distinguished that the energy error of signal No. 1, 6 and 18, i.e. the corresponding signal energy error of No. 1, 6 and 29, the energy of low lobe harmonic frequency band and the characteristic parameter of maximum value of imf1 spectrum have large correlation and small redundancy.
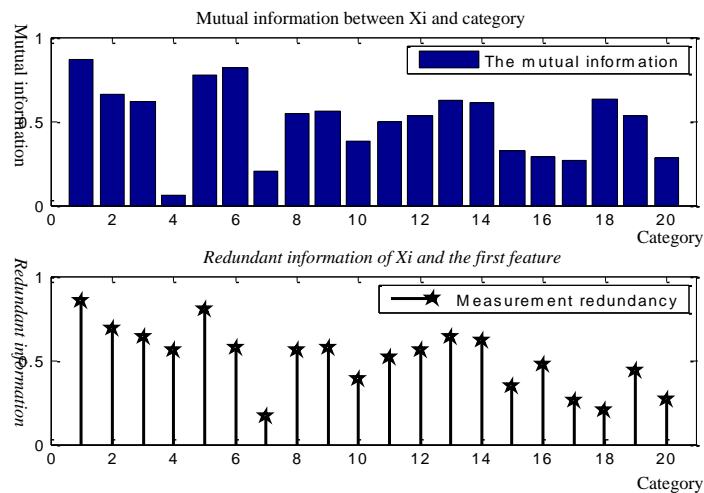

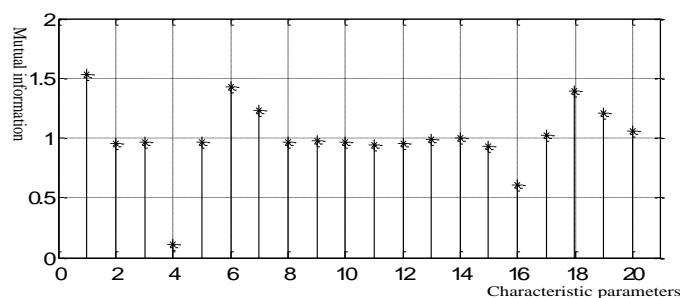
Fig 12: Data of mutual information

Fig 13: Criterion of information measurement

V. Result Discussion

In this paper, the time-domain, frequency-domain and time-frequency-domain feature sets are given, and some better feature subsets are given by using the filtering feature selection algorithm. Through the subsequent classification verification based on support vector machine classifier (SVM), the recognition rate of feature subset obtained by information measurement can reach more than 80%, which basically meets the needs of the project, while the recognition rate of feature subset obtained by geometric distance measurement can only reach 60%. This paper has basically completed the research goal.

In theory, feature selection can search all possible feature combinations by exhaustive method, and select the best feature subset in the evaluation criteria as the final output. However, in the actual search process, n features can have 2N-1 different feature combinations. Therefore, although the exhaustive method is simple in theory, it is not practical in practice because its computation and running time increase exponentially with the increase of feature dimension, and hundreds or even thousands of features are often encountered in practical application. Therefore, exhaustive method is not the best choice in real life. In addition to exhaustive search, there are other search methods, such as heuristic search and random search. These search strategies focus more on operational efficiency, and do not have to search for the best feature subset, but they can find a better balance between the efficiency and the quality of feature subset, which is also the goal of many feature selection algorithms.

A basic framework of feature selection is proposed. A typical feature selection algorithm consists of four basic steps shown in Fig.14.
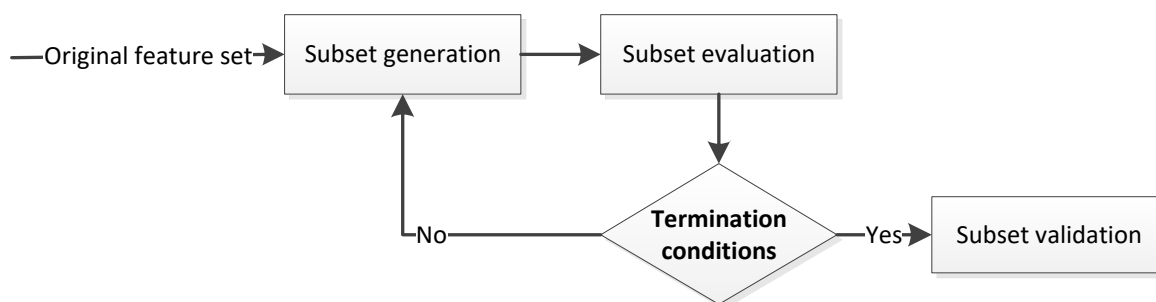

Fig 14: Based steps of feature selection

(1) Subset generation: according to a certain search strategy, candidate feature subsets are generated along different search directions. The generation of feature subset is a search process, mainly including the selection of starting point, search direction and search strategy. Among them, the selection of search strategy in feature selection process has a great impact on feature subset results.
(2) Subset evaluation: used to evaluate the advantages and disadvantages of feature subset. In the process of feature selection, there are different feature subsets due to the difference of algorithms used. How to determine which feature subsets are optimal needs to be evaluated by some evaluation rules.

(3) Termination condition judgment: the feature selection algorithm needs to give a preset stop condition, otherwise the algorithm will continue to cycle, and it needs to determine the termination condition when the algorithm stops.

(4) Subset verification: the process of subset validation is to test whether the subset selected by the algorithm is good or bad. Usually, the original feature set and the feature set after feature selection are trained and tested by classifier, and the advantages and disadvantages of the selected feature subset are finally verified by comparing the results.

The generation of feature subset is a process of continuous search, which mainly includes the search direction and search strategy of feature subset to be evaluated. The search direction mainly refers to which direction the feature selection algorithm uses to search subsets. The common search directions for generating subsets include forward search, backward search, bidirectional search and random generation.

According to the search direction, search strategies can be divided into forward, backward and bidirectional. For small-scale feature sets, the optimal subset can be obtained by exhaustive search. For the medium-sized feature set, when the evaluation function satisfies the monotonicity of the feature dimension, the branch and bound method (BB) proposed by Narendra et al. [16] can be used to solve the optimal feature subset. But in practical problems, the evaluation function usually does not have monotonicity, and the relationship between the complexity of BB algorithm and the number of features is exponential. Therefore, people are looking for search algorithms that can get better suboptimal solutions.

The ultimate goal of feature selection is to minimize the error rate of the classifier, and the feature subset obtained after feature selection is ultimately used for the subsequent learning algorithm. Therefore, the most intuitive evaluation standard is to use the classifier error rate that is to select the feature set with the minimum error probability of the classifier. However, this method needs to calculate the combination of all possible feature subsets, so it is not practical in real life. Even if the class conditional distribution density is known, the calculation process of the classifier's error rate is very complex. In practice, the class conditional distribution density is also unknown, which makes it more difficult to calculate. Therefore, feature selection based on evaluation function is also commonly used evaluation criteria, mainly including distance measurement, information measurement, correlation measurement, consistency measurement and classifier measurement.

The comparison of search strategies and evaluation criteria for feature selection is shown in table 4 [17-19] and table 5 [20-25].

Table 4 The comparison between different measures

evaluation criterion	Classification of computational complexity	Accuracy	generality	Processing data types
Distance measure	low	uncertain	strong	Continuous / discrete
Information measurement	higher	uncertain	strong	Continuous / discrete
Correlation measure	low	uncertain	strong	Continuous / discrete
Classification error rate measurement	Very high	Very high	weak	Continuous / discrete

Table 5 Summary of search algorithm for feature selection

SFS SBS	Sequential forward search (SFS) is a sequential backward search algorithm. starts from the empty set, first selects the best feature K, and then selects the two best evaluated features including K. in this way, the feature selected will not be deleted. SBS starts from the complete set and continuously deletes the features. The feature is that once deleted, it will not be selected again
GSFS(g)	The generalized sequential forward / backward search algorithm and the generalized sequential

GSBS(g)	forward / backward search algorithm evaluate the feature subset with the size of g each time, and each step is to add (forward search) or delete (back search) for g features
PTA(l,r) GPTA(l,r)	Add L and subtract R method plus l take away R, go forward 1 step (add 1 feature by SFS method), backward r step (delete r feature by SBS method), gpta (L, R) select GSFs (L) method to add features, GSBS (R) method to subtract features
SFFS SBFS	The floating mode of PTA (L, R) method. Different from PTA (L, R), SFFS / SBFS does not limit the number of forward / backward steps. SFFS and SBFS methods can backtrack unlimited, as long as backtracking can find better features of pen current features. The difference between SFFS and SBFS is that the starting point of search space is different. SFFS starts with empty sets, while SBFS starts with full sets
BAB,BAB+ Bab + +	The branch and bound method and its extended series usually need to evaluate the monotonicity of the function and find the optimal solution
GA	Genetic algorithm can't guarantee to find the optimal solution, but the search space is large, the speed is fast, and it is not easy to fall into local extremum
SA	The simulated annealing search algorithm has a large amount of calculation, and it is difficult to set the initial temperature and the number of iterations under each temperature value

VI. Conclusion

Ultrasonic testing is widely used in the field of industrial nondestructive testing. At present, there are many researches on ultrasonic testing, but most of them focus on the qualitative identification and evaluation of defects. The research on how to realize the quantitative identification of defects is still in the initial stage. In this paper, aiming at the quantitative identification of the bonding quality of thin plate bonding structure, the ultrasonic echo method is used to build an experimental system for the ultrasonic detection of the bonding quality of thin plate. In the process of signal characteristic parameter analysis, in addition to time-domain and frequency-domain analysis, the paper also uses the time-frequency analysis method of empirical mode decomposition signal, which effectively highlights the time-frequency characteristics of the echo signal, improves the resolution of signal analysis, and expands the space for echo signal feature extraction. It is difficult to find the signal characteristics which can quantitatively identify the bonding quality through signal processing methods. It is necessary to study the transmission mechanism of ultrasonic in the thin plate bonding structure. In this paper, a large number of echo signal parameters in time domain, frequency domain and time-frequency domain are analyzed. The feature validity evaluation theory of fault diagnosis is introduced to evaluate the feature validity of high-dimensional signal parameters. The invalid signal parameters are eliminated, and an effective echo signal feature set with low dimension is obtained. Using the exhaustive search strategy of feature selection theory, geometric distance measurement and information measurement evaluation criteria, and through the classification verification of support vector machine classifier, the effective feature vector for quantitative identification of adhesive quality is obtained. In this paper, a set of scientific ultrasonic echo feature extraction method is proposed, which opens up a new research idea for the research of ultrasonic echo feature extraction for quantitative identification of thin plate bonding quality, and also lays the foundation for further realizing the quantitative identification of ten grades of adhesive structure of thin plate.

Acknowledgments

References

[1] R.A. Ponram, B.H. Prasad, S.S. Kumar, et al., "Thickness mapping of rocket motor casing using ultrasonic thickness gauge," Mater Today Proc, vol. 5, pp. 11371-11375, 2018.

[2] K. Zhang, Shicheng L, Shicheng Z, "Detection of disbonds in multi-layer bonded structures using the laser ultrasonic pulse-echo mode," Ultrasonics. vol. 94, no. 94, pp. 411-418, 2019.

[3] K. Zhang, Shicheng L, Z. Zhou, "Quantitative analysis of voids in multi-layer bonded structures based on transmitted laser ultrasonic waves," Arch. Civil. Mech. Eng, vol. 19, no. 19, pp. 79-90, 2019.

[4] K.A. Tiwari, R. Raisutis, V. Samaitis, "Signal processing methods to improve the Signal-to-noise ratio (SNR) in ultrasonic non-destructive testing of wind turbine blade," Procedia. Struct. Integr, vol. 5, no. 5, pp. 1184-1191, 2017.

[5] J. Elena, M. Liudas, S. Vykintas, et al., "Ultrasonic non-destructive testing of complex titanium/carbon fibre composite joints," Ultrasonics, vol. 95, pp. 13-21, 2019.

[6] W. Peng, F. Wang, X.L. Meng, et al., "Dynamic thermal tomography based on continuous wavelet transform for debonding detection of the high silicon oxygen phenolic resin cladding layer," Infrared Phys Technol, vol. 92, no. 92, pp. 115-121, 2018.

[7] W. Wei, L. Li, W.F. Shi, et al., "Ultrasonic imaging recognition of coal-rock interface based on the improved variational mode decomposition," vol. 170, pp. 108728, 2020.

[8] G. Chen, Y. Wang, X.F. Wu, et al., "Application of CEEMDAN-FastICA in ultrasonic testing of transformer bushing leads," Nondestructive Examination, vol. 42, no. 8, pp. 8-14, 2020.

[9] P. Li, B.B. Gu, Q. Chen, L. Jiang. "Adaptive filtering modeling method for ultrasonic snow echo signals," Computer Simulation, vol. 37, no. 4, pp. 224-228, 2020.

[10] A. Katunin, A. Wronkowicz-Katunin, K. Dragan, "Impact damage evaluation in composite structures based on fusion of results of ultrasonic testing and X-ray computed tomography," Sensors, vol. 20, pp. 1867, 2020.

[11] N. Zhao, A. Basarab, D. Kouame, "Joint segmentation and deconvolution of ultrasound images using a hierarchical bayesian model based on generalized gaussian priors," IEEE Trans. Image Process, vol. 25, pp. 3736-3750, 2016.

[12] Y. Park, A. Choi, K. Kim, "Monaural sound localization based on reflective structure and homomorphic deconvolution. Sensors, vol. 17, pp. 2189, 2017.

[13] Y. Chang, Y. Zi, J. Zhao, et al., "An adaptive sparse deconvolution method for distinguishing the overlapping echoes of ultrasonic guided waves for pipeline crack inspection," Meas. Sci. Technol, vol. 28, pp. 035002, 2017.

[14] Abdessalem B, Farid C. Resolution improvement of ultrasonic signals using sparse deconvolution andvariational mode decomposition algorithms. Russ. J. Nondestruct. Test. vol. 56, pp. 479-489, 2020.

[15] Y.Q. Wang, "Preliminary study on ultrasonic echo signal model using digital signal processing technology," Dalian master of technology thesis, 2017.

[16] J.J. Lin, "Development of ultrasonic automatic detection system based on hsd4 acquisition card," Master's thesis of Harbin Institute of technology, 2017.

[17] M. Xu, "Ultrasonic testing of bonding quality of multilayer metal bonded structures," Master's thesis of China Academy of Engineering Physics, 2017.

[18] M.Z. Tang, J.L. Zhang, "Wavelet decomposition and feature extraction of ultrasonic signal of adhesive structure," Journal of testing technology, vol. 17, no. 4, pp. 363-366, 2013.

[19] Y.G. Zhao, "Research on ultrasonic testing of thin plate bonding defects based on time-frequency analysis and neural network," Master of Inner Mongolia University, 2018.

[20] M. Dash, H. Liu, "Feature selection for classifications. Intelligent Data Analysis," An International Journal, vol. 16, no. 21, pp. 131-156, 1997.

[21] D. Koller, M. Sahami. "Toward optimal feature selection," Proceedings of International Conference on Machine Learning, 2006,(33):237-250

[22] H.J. George, K. Ron, P. Karl, "Irrelevant features and the subset selection problem," Machine Leaming in: Proceedings of the llth International Conference, vol. 11 no. 5, pp.121-129, 2004.

[23] W. Jason, T. Mike, "Use of the zero-norm with linear models and kemel methods," Journal of Machine Learning Research, vol. 3, no. 19, pp. 1439-1461, 2003.

[24] Y. Mao, X.B. Zhou, Z. Xia, "A survey of feature selection algorithms," Pattern Recognition and Artificial Intelligence, 2007.

[25] J. Cao, "Research on feature selection algorithm based on maximum correlation and minimum redundancy," Master Thesis of Yanshan University, 2010.

[26] R.M. Wu, "Research on digital ultrasonic testing system and key technologies," Doctoral Dissertation of Zhejiang University, 2004.

[27] T. Yang, "Research on tubing nondestructive testing and defect quantification technology based on multi-sensor fusion," Doctoral Dissertation of Tianjin University, 2004.

www.converter-magazine.info