J. Sens. Sens. Syst., 14, 1–11, 2025
https://doi.org/10.5194/jsss-14-1-2025

JSSS Journal of Sensors and Sensor Systems Open Access

Suppression of polarization random noise in a two-dimensional force sensor based on random forest

Jian Liu

Jilin Province Economic Management Cadre College, Changchun, 130012, China

Correspondence: Jian Liu (13104499086@163.com)

Received: 13 September 2023 – Revised: 9 August 2024 – Accepted: 20 August 2024 – Published: 17 January 2025

Abstract. To improve the accuracy of polarization random noise signal recognition and the noise suppression effect, a method of polarization random noise suppression in a two-dimensional force sensor based on random forest is proposed. The polarization random noise signal model is constructed, the polarization random noise signal is enhanced and the polarization random noise power spectrum of a two-dimensional force sensor is calculated. The characteristic parameters of a polarization random noise signal are extracted and the polarization random noise signal recognition is completed by using a random forest model. Finally, fast Fourier transform is applied to suppress polarization random noise. Simulation results show that the method has a polarized random noise signal with a signal-to-noise range of −15 to 5 dB. The maximum recognition rate even reached 0.99. Therefore, the proposed noise suppression model has the advantages of high accuracy and a good noise reduction effect, which provides a new scheme for the polarization noise suppression problem in the sensor industry.

1 Introduction

In scientific research, production and engineering, it is often necessary to measure multiple data at the same time and sometimes multiple data with different properties (Rymarczyk and Sikora, 2020). If multiple sensors are used to complete the task, they are often limited by objective factors such as installation space and service conditions. Therefore, people have designed and developed multi-dimensional sensors. Among these multi-dimensional sensors, the development of multi-dimensional force sensors is particularly remarkable (Arbelaez et al., 2020). Two-dimensional force sensors are the most widely used sensors in mechanical measurement. They have the advantages of a compact structure, simple manufacture and convenient use (Kato and Yasuda, 2021; Kim et al., 2020). However, when using a two-dimensional force sensor for actual measurement, the polarization of the sensor produces random noise easily, which may affect the final measurement results. Therefore, in the application of a two-dimensional force sensor, solving the problem of polarization random noise has become the primary focus. Xie et al. (2019) propose a method for noise suppression based on short-time singular-value decomposition. This method intercepts the early signal segment through the short-time data window, simulates the actual partial-discharge signal through the typical partial-discharge pulse model, considers the influence of various signal-to-noise ratios and window lengths on the denoising results and uses singular-value decomposition to suppress the signal noise. The noise interference effect of the method is verified under oscillation and normal conditions, respectively. Ma and Zhang (2021) propose a method of noise suppression on the basis of an improved empirical wavelet transform (IEWT). This method focuses on IEWT decomposition. By decomposing the noisy signal into an empirical wavelet function (EWF) which is arranged in frequency order, the phenomenon of mode aliasing in complex signal noise decomposition can be effectively avoided. The decomposed EWF is filtered by the kurtosis rule. Finally, the useful EWF is reconstructed and denoised. Wu et al. (2018) propose a method of noise suppression based on gain self-compensation. Firstly, the transient noise caused by dead-time compensation is analysed. On this basis, the noise interference suppression method based on gain self-compensation is studied. According to the setting

requirements of the stability control platform for phase lag and transient noise, the noise suppression is realized by gain self-compensation. Westhausen and Meyer (2020) propose a dual-signal-conversion long short-term memory (LSTM) network (DTLN) for the real-time voice noise reduction problem. The method combines short-time Fourier transform and learning with fewer than 1 million parameters. The results show that the method has real-time noise reduction power and outperforms the deep noise suppression (DNS) challenge baseline, improving by 0.24 points in the average opinion score. Birnie et al. (2021) propose using a blind-spot network for self-supervised denoising for the random noise problem in seismic data processing. The results show that this method can effectively remove noise and reduce signal damage, which can improve image quality and subsequent task performance. Compared with traditional denoising techniques, this shows the potential of self-supervised learning in seismic applications. Dorst et al. (2023) propose an extended uncertainty-aware automated machine learning toolbox (UA-AMLT) for regression analysis to address the issue of machine learning models not directly handling measurement uncertainty in indoor air quality monitoring. The results show that this method improves the robustness of the model to random noise and provides an effective strategy for improving indoor air quality monitoring.

The above methods have their own characteristics in noise suppression, but there are still shortcomings. The method of short-time singular-value decomposition depends on the selection of a signal window and a signal-to-noise ratio and may not be suitable for non-stationary or atypical noise. Although the improved empirical wavelet transform can effectively avoid modal aliasing, its adaptability to complex and variable noise environments needs to be improved. The method based on gain self-compensation is mainly targeted at specific systems and has relatively weak universality. The noise suppression method based on random forest proposed in this study can more accurately identify and process the polarization random noise of two-dimensional force sensors by utilizing machine learning techniques. The random forest algorithm has strong classification and regression capabilities, which can adaptively learn noise features and effectively suppress random noise, thus solving the limitations of traditional methods in dealing with variable and complex noise.

2 Suppression of polarization random noise in a two-dimensional force sensor based on random forest

2.1 Calculation of the polarization random noise power spectrum of a two-dimensional force sensor

In the process of suppressing polarization random noise in a two-dimensional force sensor, the power spectrum of polarization random noise needs to be calculated first, assuming

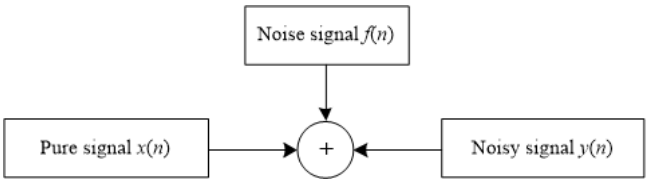

Figure 1. Signal model with polarization random noise.

that the polarization random noise signal model is

$$y(n) = x(n) + f(n). \tag{1}$$

In the equation, n is the index of discrete time points, which is an integer used to identify the value of the signal at a particular time point. $x(n)$ represents the pure vibration signal, $f(n)$ represents the polarization interference noise, $x(n)$ and $f(n)$ are independent and uncorrelated and $y(n)$ represents a signal with polarization random noise (Ullah et al., 2020).

In order to eliminate the truncation effect caused by signal framing, $y(n)$ is windowed first. In this study, a Hamming window is adopted and the polarization random noise signal $y_i(n)$ of frame i can be expressed as

$$y_i(n) = y(i \cdot L + n)\, w(n), 0 \le n \le L - 1. \tag{2}$$

In the equation $i = 0, 1, 2, \ldots, I-1$, I represents the number of frames that $y(n)$ can be divided into, and $w(n)$ represents the Hamming window containing L sampling points:

$$w(n) = \begin{cases} 0.5 - 0.5\cos\left(\frac{2\pi n}{L-1}\right), & 0 \le n \le L-1 \\ 0, & \text{else} \end{cases}. \tag{3}$$

Secondly, by performing a fast Fourier transform on each noise signal frame $y_i(n)$, the following short-time amplitude spectrum of the polarization random noise signal can be obtained:

$$Y(l,k) = \sum_{n=0}^{L-1} y(lQ+n) w(n) e^{-j2\pi kn/L}. \tag{4}$$

In the equation, Q represents the frame shift length, l represents the frame number and k represents the sub-band number. It can be seen from the above equation that, after short-time amplitude spectrum calculation, the noise signal is converted from a single time domain into a time–frequency domain (Huang et al., 2018; Imtiaz et al., 2019; Palagan, 2018).

The model with a polarization random noise signal is shown in Fig. 1.

The general process of a noise enhancement algorithm based on short-time spectrum estimation is as follows. Firstly, the noise power spectrum is estimated from the short-time power spectrum $|Y(l,k)|^2$ of the polarization random noise signal and the corresponding time–frequency gain factor $H(l,k)$ is calculated according to certain criteria. Then, the short-time amplitude spectrum $Y(l,k)$ is multiplied by

$H(l,k)$ to obtain the enhanced signal short-time amplitude spectrum $X(l,k)$. The phase-insensitive characteristic of the noise signal is then used and the enhanced signal phase is replaced by the phase of the polarization random noise signal so as to obtain the signal spectrum after noise elimination (Zhang, 2020). Finally, the enhanced time domain audio signal $x'(n)$ can be obtained by an inverse fast Fourier transform of the obtained signal spectrum, which is shown in Fig. 2.

Since the power spectrum transformation range of polarization random noise in a non-stationary environment is usually large (Maryami and Liu, 2024), it is necessary to perform the following first-order recursive smoothing on the power spectrum of a polarization random noise signal:

$$P(l,k) = \varsigma P(l-1,k) + (1-\varsigma)|Y(l,k)|^2 . \tag{5}$$

In this equation, $P(l,k)$ represents the smooth power spectrum of a polarization random signal and ς represents the constant smoothing factor set according to the actual situation.

The purpose of first-order recursive evaluation is to reduce the amplitude of those frequency points disturbed by noise to the same degree as the front and rear frequency points, rather than directly reducing them to 0. Therefore, the first-order recursive smoothing can not only effectively weaken the influence of background noise but also ensure that the useful foreground signal will not be lost too much, so that there is a better approximation between the enhanced signal and the original pure signal (Yao and Sun, 2019; Zhao et al., 2020; Zhu and Xie, 2019).

After obtaining the smooth power spectrum of the noisy signal, the next thing to do is to determine its local minimum at each frame. In this paper, the bi-directional path search method combining forward and backward searches is used to find the minimum value of the smooth power spectrum. That is, for the current processing frame, the forward search and backward search are performed simultaneously to obtain two local minimum values (Zhao et al., 2019). Finally, the larger of the two is taken as the local minimum value of the smooth power spectrum in each frequency band of the frame, which can be described as below:

$$P_{\mathrm{f}}(l,k) = \min\left\{P(l^*,k)\right\}, \; l \le l^* \le l+F-1 . \tag{6}$$

In Eq. (6), F represents the number of frames to consider after the current frame, which is used to determine the size of the search window. Specifically, when searching for the local minimum value of the smooth power spectrum, the algorithm will search within a determined frame range. This range starts from the current processed frame l and searches backward for $F-1$ frames to find the local minimum value. Calculate the probability of the existence of the foreground polarization signal. The probability of the existence $E(l,k)$ of the foreground polarization signal can be calculated with the following first-order recursive equation:

$$E(l,k) = \beta E(l-1,k) + (1-\beta)G(l,k) . \tag{7}$$

In this equation, β represents the constant smoothing factor set according to the situation. $G(l,k)$ represents the discrimination standard for the existence of a foreground polarization signal, which can be described as

$$G(l,k) = \begin{cases} 1, & \text{existent}, \quad D_{\mathrm{r}}(l,k) > \delta(k) \\ 0, & \text{non-existent}, \quad \text{else} \end{cases} . \tag{8}$$

In the equation, $\delta(k)$ represents the frequency correlation threshold set according to the actual situation and $D_{\mathrm{r}}(l,k)$ represents the ratio of the smooth power spectrum of the random noise signal with polarization to its local minimum, i.e.

$$D_{\mathrm{r}}(l,k) = \frac{P(l,k)}{P_{\min}(l,k)} . \tag{9}$$

In other words, when $D_{\mathrm{r}}(l,k)$ is greater than $\delta(k)$ in a certain sub-band, there is considered to be a foreground noise signal in the band. By contrast, there is considered to only be a polarization random noise signal in the band.

Calculate the smoothing factor of the time–frequency correlation. According to the existence probability $E(l,k)$ of the foreground noise signal obtained in the above steps, the smoothing factor $\rho(l,k)$ of the time–frequency correlation can be obtained:

$$\rho(l,k) = \rho_1 (1-\rho_1) E(l,k) . \tag{10}$$

In the equation, ρ_1 represents the constant smoothing factor set according to the actual situation and the value range of the smoothing factor $\rho(l,k)$ is $\rho_1 \le \rho(l,k) \le 1$.

Estimate the background polarization random noise power spectrum. When there is a foreground noise signal at a frame, i.e. when $\rho(l,k)$ approaches 1, the background polarization random noise power spectrum should be almost equal to the noise power spectrum at the previous frame. When there is almost no foreground noise signal at a frame, i.e. when $\rho(l,k)$ approaches 0, the background polarization random noise power spectrum should be approximately equal to the noisy signal power spectrum at the frame. Therefore, the polarization random noise power spectrum of a two-dimensional force sensor can be calculated as

$$L(l,k) = \rho(l,k)L(l-1,k) + (1-\rho(l,k))\,|Y(l,k)|^2 . \tag{11}$$

2.2 Polarization random noise recognition based on random forest

The key step in polarized stochastic noise suppression by a two-dimensional force sensor is noise recognition. As the random forest algorithm is very robust to noisy data, it reduces the miscalculation where a single decision tree may result from noisy data by building multiple decision trees and selecting or averaging their predictions. This property allows a random forest to perform well in its noise recognition task. The random forest algorithm randomly selects feature subsets to split when constructing the decision tree.

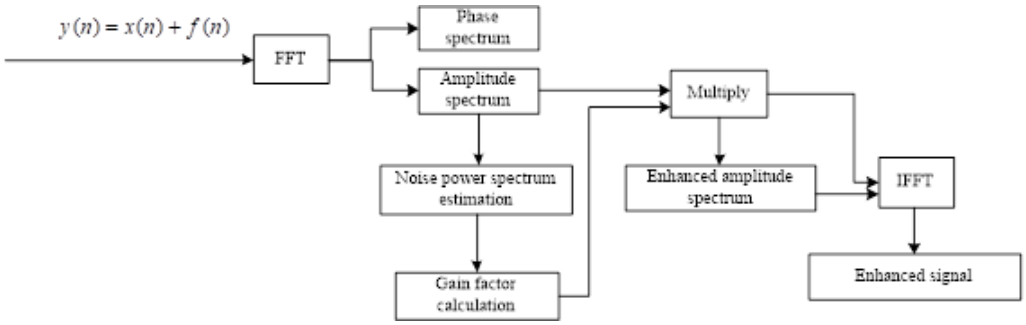

Figure 2. Noise signal enhancement process.

This stochasticity of feature selection not only increases the diversity of the models but also facilitates the discovery of features associated with noise, thus improving the accuracy of the noise identification. The characteristic parameters selected for noise recognition are the bandwidth occupied by the noise signal, kurtosis and the zero centre's largest normalized transient amplitude spectral density (Neumann et al., 2020). The calculation equations and definitions are as follows.

2.2.1 Occupied bandwidth

Occupied bandwidth refers to the proportion of broadband with 99 % of the total power in the total broadband. The power below or above the lower frequency limit of the occupied broadband is equal to 0.5 % of the total power. The occupied bandwidth of an amplitude spectrum can be used to distinguish between the frequencies of different signals, i.e. to distinguish noise signals from conventional signals.

2.2.2 Kurtosis

Kurtosis K indicates the number of characteristics of the peak value of the probability density distribution curve at the average value, which are used to distinguish the peak value of the amplitude spectrum of the polarization random noise signal. The kurtosis calculation equation of the amplitude spectrum of the polarization random noise signal is

$$K = \frac{E\left[(X(l,k) - E[X(l,k)])^4\right]}{\left\{E\left[(X(l,k) - E[X(l,k)])^2\right]\right\}^2}. \tag{12}$$

In Eq. (12), $X(l,k)$ represents the amplitude spectrum of the polarized random noise signal, where l and k may be indices related to time, frequency or other relevant parameters. $E[.]$ represents the expectation or mean operation used to calculate the average value of the expression in parentheses. The form of the entire equation is similar to the "standardized" fourth-order central moment, which measures the size of the peak of the data distribution relative to its variance. When the kurtosis is greater than 3, it indicates that the data distribution has heavy tails or more outliers than the normal distribution; when the kurtosis is less than 3, it indicates that the data distribution has lighter tails or fewer outliers than the normal distribution. In the analysis of polarized random noise signals, by calculating the kurtosis of the amplitude spectrum, we can obtain useful information about the peak characteristics of the signal, which is helpful for further processing and analysis of the signal.

2.2.3 Largest normalized transient amplitude spectral density at the zero centre

The largest value $\gamma_{\max}$ of the normalized transient amplitude spectral density at the zero centre represents the change in the transient amplitude of the polarization random noise signal, which is used to distinguish the strength of the signal. The calculation equation is

$$\gamma_{\max} = \frac{\max|\mathrm{DFT}(S_{\mathrm{cn}}(i))|^2}{N}, \tag{13}$$

where N indicates the sampling signal number and $S_{\mathrm{cn}}(i)$ indicates the normalized transient amplitude of the zero centre $S_{\mathrm{cn}}(i) = S_c(i) - 1$, where $S_c(i) = \frac{S(i)}{m_a}$. $S(i)$ represents the transient amplitude signal and $m_a = \frac{\sum_{i=1}^{N} S(i)}{N}$ represents the average value of $S(i)$ (Guo et al., 2020).

After calculating the characteristic parameters, the polarization random noise signal is recognized in the random forest. Based on bootstrap aggregation, random forest is a strong classifier that includes multiple weak classifiers (decision trees). First, put back sampling through the self-help method and set up training sub-datasets, with each set forming a decision tree. Then, randomly select characteristic parameters to train the decision trees. Each decision tree is irrelevant to the others. When making a decision, vote on the decision results made by multiple decision tree classifiers to obtain the final result (Drachenko and Mikhnyuk, 2020; Nishiyama et al., 2020). The recognition performance of random forest is higher than that of a single decision tree and overcomes the overfitting problem of decision trees. At the same time, due to random sampling, the variance of the training model is small, the generalization ability is strong and the model

is robust to missing data and unbalanced data. The random forest model is shown in Fig. 3.

The random forest algorithm here uses the method of feature selection on the basis of a Gini coefficient and classification of a regression tree and a decision tree. A Gini coefficient represents the chaotic degree of the model. A smaller Gini coefficient will lead to a smaller chaotic degree. The feature which has the smallest Gini coefficient is selected to ensure that all samples of each sub-node belong to one classification. The probability distribution's Gini coefficient is

$$\mathrm{Gini}(p) = 2p(1-p). \tag{14}$$

In Eq. (14), $\mathrm{Gini}(p)$ is the Gini coefficient as a measure of the degree of chaos in a dataset. The smaller the Gini coefficient, the less chaos in the dataset, i.e. the higher the purity of the dataset. p represents the probability of a certain category appearing in the dataset. In dichotomous problems, p usually indicates the probability of a positive class and $(1-p)$ indicates the probability of a negative class. In the multiclassification problem, the Gini coefficient is calculated separately for each category and considers the category distribution. As for sample set C, after traversing all the segmentation points of feature parameter A, it is separated into two parts by using the relationship between the feature parameter and threshold T_A (such as $A > T_A$), i.e. sample set C_1 satisfying $A > T_A$ and sample set C_2 not satisfying $A > T_A$. In the case of $A > T_A$, the sample set's Gini index is

$$\mathrm{Gini}(C, A) = \frac{|C_1|}{|C|}\mathrm{Gini}(C_1) + \frac{|C_2|}{|C|}\mathrm{Gini}(C_2). \tag{15}$$

In the equation, $\mathrm{Gini}(C_1)$ and $\mathrm{Gini}(C_2)$ represent the uncertainty of sample sets C_1 and C_2, respectively, and $\mathrm{Gini}(C, A)$ indicates set C's uncertainty after $A > T_A$ division.

The cart decision tree is a binary tree. In accordance with the various numerical characteristics of the dependent variables which are in the dataset, the regression tree and classification tree can be established. The specific steps are as follows:

1. Through the self-help method, sub-datasets are obtained. Characteristic parameters randomly selected are taken as the division characteristics of the classification tree node. For each selected characteristic parameter A, take all possible thresholds T_A, calculate the sub-dataset's Gini coefficient after $A > T_A$ division and select the characteristic parameter which has the smallest Gini coefficient as well as the corresponding threshold as the characteristic division point of the node.

2. If the node sample number or the tree depth meet the requirements, the construction of the classification tree ends and the successful cart decision tree is returned. Otherwise, step (1) is executed recursively for the two child nodes.

In this paper, the bandwidth occupied by three characteristic parameters, kurtosis and the largest normalized transient amplitude spectral density of the zero centre is used, and the polarization random noise signal is categorized and identified by a random forest classifier. First of all, normalize the signal power under various signal-to-noise ratios, extract the characteristic parameters of the noise signal and form the training dataset and test dataset of the random forest. Secondly, the training dataset is sampled by the self-help method to construct training sub-datasets. Each training sub-dataset establishes a decision tree. Characteristic parameters randomly selected are used to train the decision tree. Then, multiple irrelevant decision trees are generated, all of which form a random forest. Finally, the random forest is used to classify and identify the test dataset. The decision results are determined by the majority vote of each decision tree classifier. The obtained polarization random noise signal recognition model on the basis of the random forest is shown in Fig. 4.

2.3 Suppression of polarization random noise in a two-dimensional force sensor

The expression relationship between signal frequency f_t and time t of polarization random noise in a two-dimensional force sensor is as follows:

$$f_t = f_0 + \frac{\Delta V}{T_m}t,\ 0 < t < T_m. \tag{16}$$

In the equation, f_0 represents the carrier frequency of the noise signal, ΔV represents the largest frequency offset and T_m represents the signal period.

The expression relationship between polarization random noise signal echo frequency f_r and time t is as follows:

$$f_\mathrm{r} = f_0 + \frac{\Delta V}{T_m}\left(t - \frac{2R}{c}\right),\ 0 < t < T_m. \tag{17}$$

In the equation, R represents the distance of the two-dimensional force sensor and c represents the propagation speed of the polarization random noise signal. The equation describes the relationship between the echo frequency and the time of polarized random noise signals in a two-dimensional force sensor. The echo frequency is the sum of the carrier frequency and a frequency offset that varies over time. This frequency offset is determined by the maximum frequency offset, signal period and signal propagation time. Therefore, the equation actually describes a physical process: when a polarized random noise signal propagates in a two-dimensional force sensor, its echo frequency will change over time, which is determined by the carrier frequency, maximum frequency offset, signal period and distance and speed of the signal propagation. According to the above calculation equation, the relationship between differential frequency f_i and sensor distance R can be obtained:

$$f_\mathrm{r} = \frac{2\Delta V}{cT_m}R. \tag{18}$$

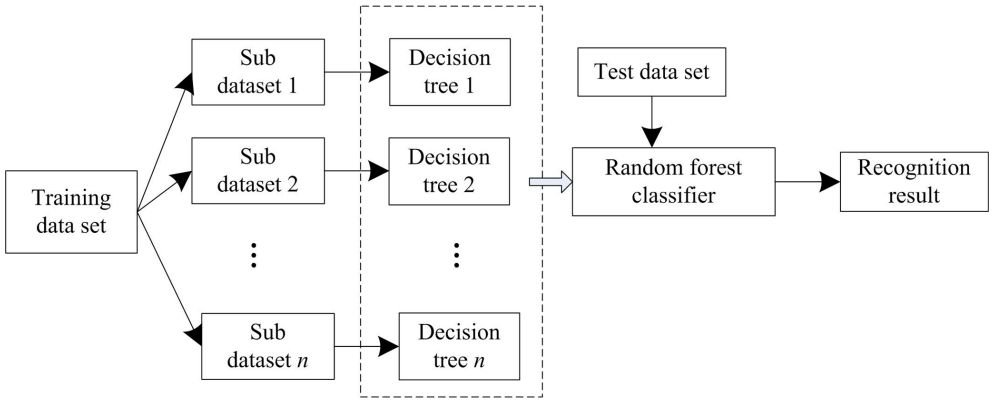

Figure 3. Random forest model.

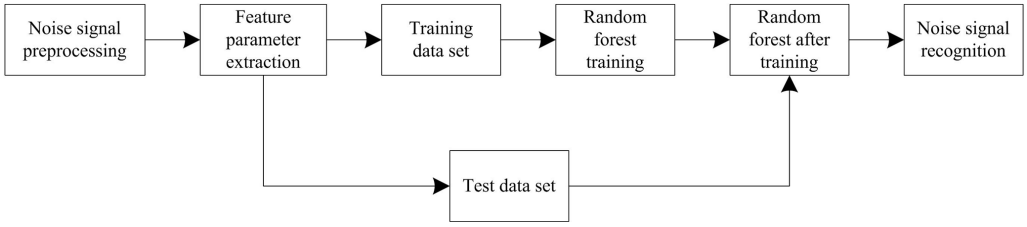

Figure 4. Identification model of polarization random noise based on random forest.

Linear frequency modulation selects the best suppression position R_0 according to the difference frequency signal between the polarization random noise signal and the ordinary polarization signal. When the optimal suppression position is given, the corresponding differential frequency f_{i0} can be obtained according to Eq. (19). A bandpass filter with centre frequency f_{i0} is set in the two-dimensional force sensor to let the required differential frequency signal pass through and suppress signals of other frequencies.

Due to the relative motion of the polarization random signal, the calculation equation of the frequency variation Δf_i of the differential frequency signal in one modulation period T_m is

$$\Delta f_i = \frac{2kv_jT_m}{c} = \frac{2v_j\Delta V}{c}. \tag{19}$$

In the equation, $k = \frac{\Delta V}{T_m}$ represents the frequency modulation slope and v_j represents the data change speed of the two-dimensional force sensor.

In a modulation period T_m, the ratio ε which is between the frequency variation of the differential frequency signal and the frequency of the differential frequency signal is

$$\varepsilon = \frac{\Delta f_i}{f_i} = \frac{T_m v_j}{R} \ll 1. \tag{20}$$

That is, if $\Delta f_i \ll f_m$, the frequency f_i of the differential frequency signal in one modulation period T_m can be considered the fixed value. Therefore, the polarization random noise interference can be suppressed by fast Fourier transform. The principal block diagram of noise interference suppression is shown in Fig. 5.

Because the signal in a period T_m can be considered a single frequency, its spectrum is an impulse function. The noise spectrum is uniform. Find the largest value of the polarization signal spectrum with noise interference, set it to zero and then perform fast Fourier transform to obtain noise interference. The interference noise signal is subtracted from the polarization signal containing noise interference to obtain the polarization signal after noise interference suppression.

3 Experimental verification

This article uses the simulation software MATLAB to generate a dataset and conduct a simulation experiment. The dataset to be tested contains ordinary signals and noisy signals. The signal parameters are set as follows. The symbol rate is 50 Hz. The carrier frequency is 1.55 MHz. The sampling frequency is 62 MHz. The m sequence with a period of 127 is used as the pseudo-random code sequence and the pseudo-random code rate is 635 kHz. The noise is zero mean Gaussian white noise and the signal-to-noise ratio ranges from -15 to 5 dB, with an interval of 1 dB. The training dataset and the test dataset are generated through MATLAB. Under different signal-to-noise ratio conditions, 2000 training data and 1000 test data are generated for ordinary signals and noisy signals.

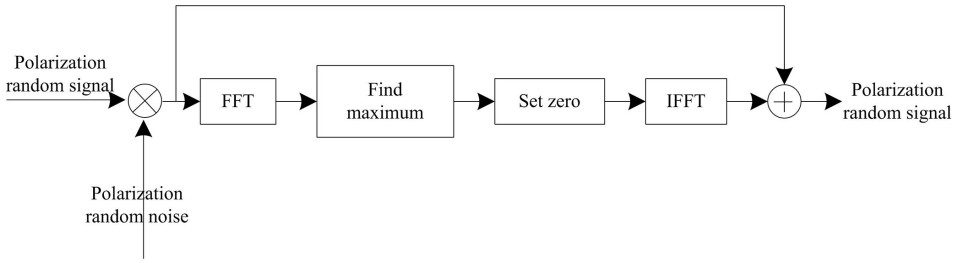

Figure 5. Schematic diagram of noise interference suppression.

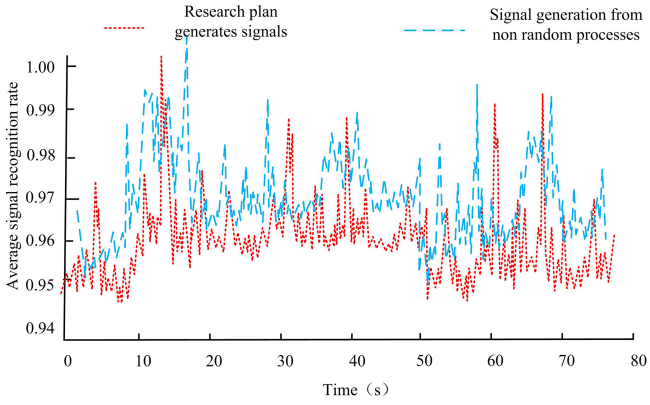

Figure 6. Comparison of signals between random walk signals generated by the research plan and data generated by non-random processes.

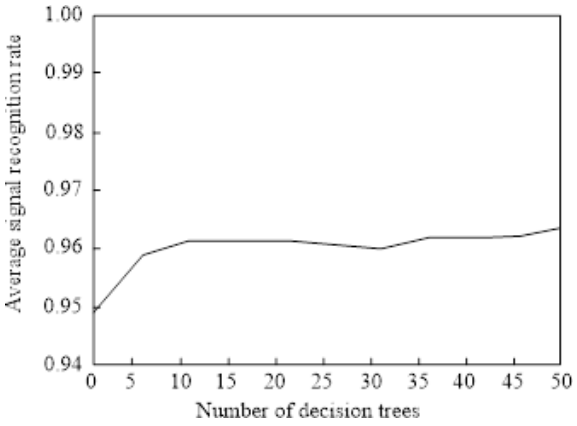

Figure 7. Average recognition rate of the signals.

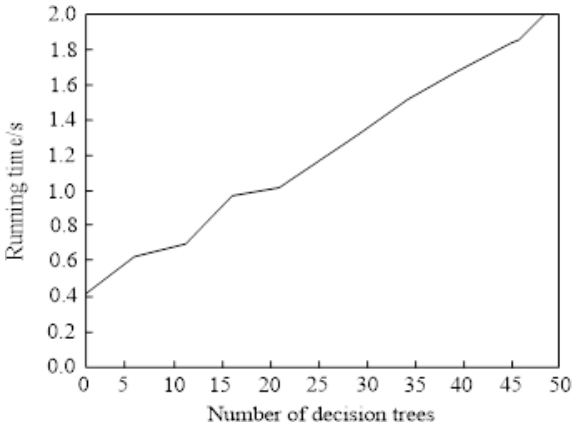

Figure 8. Running time.

3.1 Influence of the decision tree number on the performance of the method

In order to ensure that the generated random walk signals follow natural patterns in real life, in-depth research has been conducted on signals in real life to understand their statistical characteristics and variation patterns. Next, in MATLAB, these characteristics are utilized to simulate the signal, ensuring that the simulated signal is similar to the real signal in terms of amplitude, frequency and phase. In addition, the introduction of an *m* sequence as a pseudo-random code sequence not only increases the randomness of the signal but also makes it closer to the signal mode in actual communication. At the same time, by adding zero mean Gaussian white noise, the common noise interference in the real environment was simulated, thereby improving the fidelity of the signal. Finally, strict statistical analysis was conducted on the generated signal, which was compared with real-life signals to ensure that the statistical characteristics of the simulated signal were consistent with the real signal. These measures collectively ensure that the random walk signals generated in the experiment closely follow the natural patterns in real life. The comparison between the random walk signals generated by the research plan and the signals generated by non-random processes is shown in Fig. 6.

From Fig. 6, it can be seen that the random walk signals generated by the research scheme exhibit a high degree of similarity to the signals generated by non-random processes. The decision tree number is taken from 1 to 50 at an interval of 5. The average recognition rate of all the test data in the whole test set of this method is shown in Fig. 7 and the running time is shown in Fig. 8.

By analysing Fig. 7 in depth, it can be clearly observed that there is a positive correlation between the average correct recognition rate of the signal and the number of decision

trees. The figure clearly shows that, as the number of decision trees gradually increases, the average correct recognition rate of the signal also improves accordingly. This trend indicates that the integration of multiple decision trees helps improve the accuracy of the signal recognition. However, it is worth noting that when the number of decision trees exceeds 10, the growth trend of the recognition rate begins to stabilize, which means that further increasing the number of decision trees may not bring a significant improvement in the recognition rate. Meanwhile, another important factor to consider is the running time. The data in Fig. 8 show that the more decision trees there are, the longer the required running time will be. This is clearly a balancing issue, as although adding decision trees can improve recognition rates, it will also sacrifice operational efficiency. After considering both the recognition rate and the running time, the study selected 10 decision trees as the optimal balance point in the simulation experiment. This choice aims to ensure a high recognition rate while also controlling the running time, thereby achieving overall optimization of the performance of the method.

3.2 Comparison of the correct signal recognition rates

When the signal-to-noise ratio range was −15 to 5 dB, this method was compared with noise suppression methods based on short-time singular-value decomposition and improved empirical wavelet transform. Ensuring the improvement of the signal quality is crucial in signal conversion pipelines. In this process, adaptive filtering technology is first adopted to effectively remove noise components from the original signal and improve the signal-to-noise ratio of the signal. Then, through normalization processing, the signal is scaled to a unified scale and signal features are extracted using time domain analysis, frequency domain analysis or time–frequency domain analysis techniques. Use short-time Fourier transform (STFT) or wavelet transform to obtain the time–frequency characteristics of the signal. Transform the extracted features into a matrix data structure suitable for model processing. The comparison results of signal recognition accuracy are shown in Fig. 9.

According to the experimental results shown in Fig. 9, it can be clearly seen that the proposed method exhibits an excellent polarization random noise signal recognition ability in an environment with a signal-to-noise ratio range of −15 to 5 dB. This innovative method has a relatively high accuracy in identifying polarization random noise signals generated by two-dimensional force sensors. Compared with the two commonly used traditional methods, this method has significant advantages in the signal recognition rate. The data show that its average signal recognition rate is not only higher than these two traditional methods but also reaches an astonishing recognition rate of 0.99 under certain conditions. This achievement undoubtedly proves the effectiveness and superiority of the research method in identifying noise signals, providing a new solution to the noise suppression problem of two-dimensional force sensors.

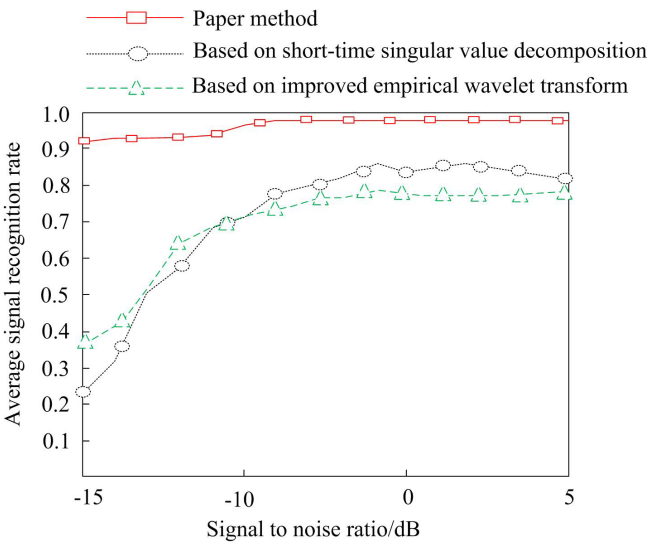

Figure 9. Comparison results of the correct signal recognition rate.

3.3 Noise suppression effect

As the key performance of the method of noise suppression, the noise suppression effect can fully reflect the suppression performance of the method. Therefore, the polarization random noise suppression effect of the method in this paper is verified based on the noise-free polarization random noise signal spectrum and by adding a signal-to-jamming ratio (SJR) = −10 dB interference noise to the basic signal to inspect the noise suppression effect.

The results in Fig. 10 show that the method in this paper has a good noise suppression effect and can remove the interference noise in the signal, and the signal after noise suppression is basically suppressed from the original signal.

3.4 Noise suppression efficiency

Due to the high workload of a two-dimensional force sensor and the large amount of data to be collected and processed, efficiency of the method of noise suppression is required. Therefore, it is necessary to compare and verify the noise suppression efficiency of this method. To further enhance the generalization ability of the model, the L1 and L2 regularization techniques were used. L1 regularization helps to generate a sparse-weight matrix, remove noise and perform feature selection. L2 regularization smoothens the weights and prevents the model from overfitting. Meanwhile, using the dropout technique, some of the neurons were randomly discarded during training to reduce the dependence of the model on the training data.

Analysing the comparison results of the noise suppression efficiency shown in Fig. 11, it can be seen that this article

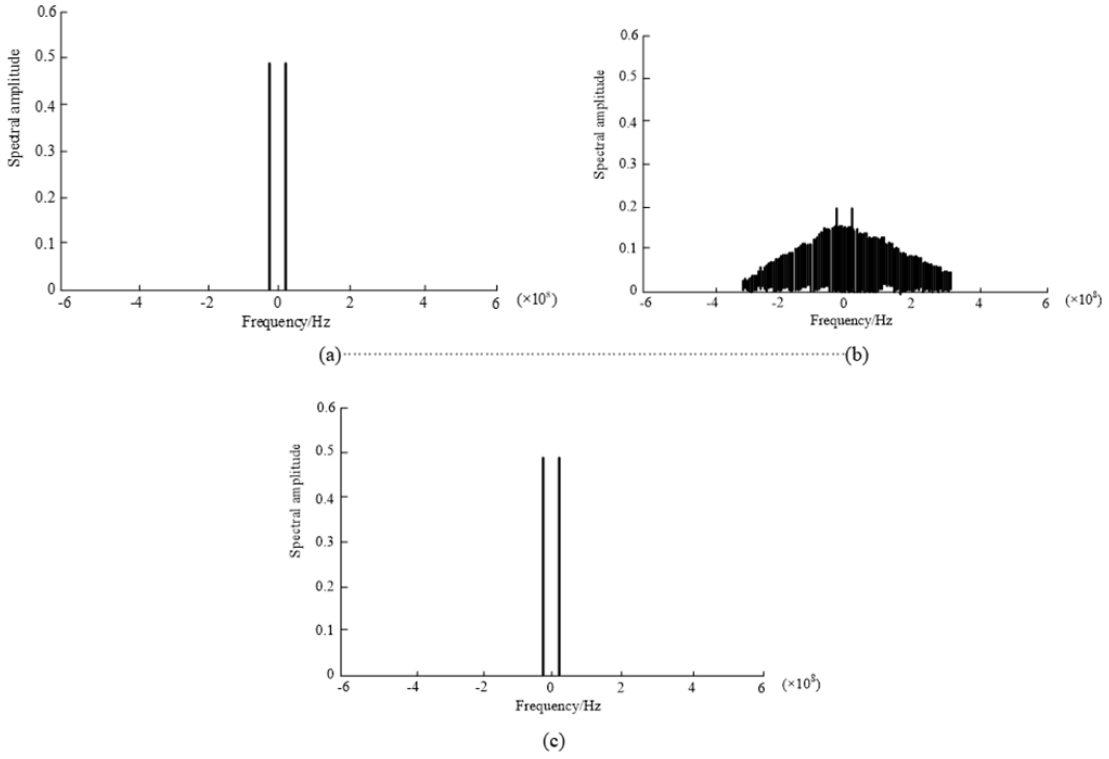

Figure 10. Noise suppression effect. **(a)** Signal spectrum without interference noise. **(b)** Noise-containing interference signal spectrum when SJR = −10 dB. **(c)** Signal spectrum after noise suppression.

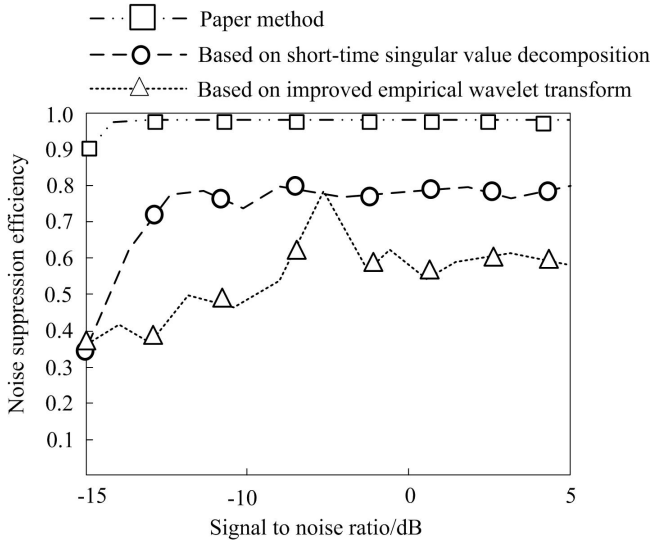

Figure 11. Noise suppression efficiency.

has high noise suppression efficiency and good noise suppression stability, which can ensure the stability of the two-dimensional force sensor operation and improve the reliability of the data transmission.

4 Conclusion

To improve the data transmission quality of a two-dimensional force sensor, a polarization random method of noise suppression in a two-dimensional force sensor based on random forest is proposed. The performance of the method is verified theoretically and experimentally. This method has a higher correct recognition rate of noise signals and a better noise suppression efficiency when suppressing polarization random noise in a two-dimensional force sensor. Compared with the method in terms of short-time singular-value decomposition and an improved empirical wavelet transform, the correct recognition rate of noise signals using this method is significantly improved and finally stabilized at 0.99. Moreover, the noise suppression effect of this method is good and the data after noise suppression are basically consistent with the original data. Therefore, the proposed method of noise suppression based on random forest can better meet the requirements of polarization random noise suppression in a two-dimensional force sensor.

Code and data availability. The author has confirmed that the data supporting the findings of this study are available within the article.

Competing interests. The author has declared that there are no competing interests.

Acknowledgements. The successful completion of this study was due to the funding by the Jilin Province Science and Technology Development Plan (project no. 20230203023SF), and we would like to express our sincere gratitude. Special thanks go to my supervisor and team members for their guidance and support during the research process. Their valuable opinions and selfless dedication enabled this study to proceed in depth. At the same time, we would like to express our gratitude to all the colleagues who participated in the discussion and provided assistance in this research, as well as all predecessors and peers who have made contributions to this field. Without everyone's joint efforts and support, this study would not have achieved such results. Finally, I would like to express my sincerest gratitude to all the experts and scholars who reviewed this paper. Your valuable feedback will be an important guide for my future research.

Financial support. This research has been supported by the Projects of Science and Technology Development Plan of Jilin Province (project no. 20230203023SF).

Review statement. This paper was edited by Daniel Platz and reviewed by two anonymous referees.

References

Arbelaez, A., Olvera-Cervantes, J. L., Corona, A., and Saavedra, C.: Compact closed-loop resonator filters with wide spurious free band and extended common-mode noise suppression, IET Microw. Antenna. P., 56, 1441–1449, https://doi.org/10.1049/iet-map.2019.0581, 2020.

Birnie, C., Ravasi, M., Liu, S. X., and Alkhalifah, T.: The potential of self-supervised networks for random noise suppression in seismic data, Artificial Intelligence in Geosciences, 2, 47–59, https://doi.org/10.1016/j.aiig.2021.11.001, 2021.

Drachenko, V. N. and Mikhnyuk, A. N.: Experimental verification of the efficiency of nonlinear processing using impulse noise suppression for a sonar station for detecting small-sized objects, Acoust. Phys.+, 66, 671–675, https://doi.org/10.1134/S1063771020050036, 2020.

Dorst, T., Schneider, T., Eichstädt, S., and Schütze, A.: Influence of measurement uncertainty on machine learning results demonstrated for a smart gas sensor, J. Sens. Sens. Syst., 12, 45–60, https://doi.org/10.5194/jsss-12-45-2023, 2023.

Guo, L., Kan, E., Wu, Y. X., Lv, H., and Xu, G. Z.: Noise suppression ability and its mechanism analysis of scale-free spiking neural network under white Gaussian noise, PLoS One, 15, 156–164, https://doi.org/10.1371/journal.pone.0244683, 2020.

Huang, Y. C., Cheng, C. H., and Wu, T. L.: A synthesized method for common-mode noise suppression filters with specified common-mode and differential mode response, IEEE T. Electromagn. C., 61, 893–902, https://doi.org/10.1109/TEMC.2018.2839264, 2018.

Imtiaz, W. A., Ahmed, H. Y., Zeghid, M., and Sharief, Y.: Analysis of noise suppression for OCDMA systems with fixed in-phase cross-correlation codes and single O/E converter, Optik, 18, 677–690, https://doi.org/10.1016/j.ijleo.2019.02.052, 2019.

Kato, T. and Yasuda, A.: A study of phase noise suppression in reference multiple digital PLL without DLLs, Analog Integr. Circ. S., 106, 441–447, https://doi.org/10.1007/s10470-020-01757-z, 2021.

Kim, Y., Park, G., Cho K., Raj, P. M., Tummala, R. R., and Kim, J. H.: Wideband power/ground noise suppression in low-loss glass interposers using a double-sided electromagnetic bandgap structure, IEEE T. Microw. Theory., 68, 19–25, https://doi.org/10.1109/TMTT.2020.3022009, 2020.

Ma, X. H. and Zhang, D. K.: Research on suppression of partial discharge noise of high voltage cable based on improved empirical wavelet transform, Transactions of China Electrotechnical Society, 36, 353–361, 2021.

Maryami, R. and Liu, Y.: Cylinder flow and noise control by active base blowing, J. Fluid. Mech., 985, A10, https://doi.org/10.1017/jfm.2024.261, 2024.

Neumann, N., Al-Husseini, Z., and Plettemeier, D.: Influence of dispersive element on phase noise suppression in Talbot effect based optical up conversion scheme, in: Photonic Networks, 21th ITG-Symposium, 24–25 November 2020, VDE, Print ISBN: 978-3-8007-5423-6, 2020.

Nishiyama, Y., Agarwal, V., and Zhang, R.: t1-noise suppression by γ-free recoupling sequences in solid-state NMR for structural characterization of fully protonated molecules at fast MAS, J. Phys. Chem. C, 124, 456–461, https://doi.org/10.1021/acs.jpcc.0c08828, 2020.

Palagan, C. A.: Joint noise suppression and dereverberation of separating speech signals by using prediction and separation matrix, in: 2018 3rd International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 15–16 October 2018, IEEE, 202-207, https://doi.org/10.1109/CESYS.2018.8724047, 2018.

Rymarczyk, T. and Sikora, J.: Optimization method and PCA noise suppression application for ultrasound transmission tomography, Prz. Elektrotechniczny, 96, 90–93, 2020.

Ullah, R., Islam, S., Ye, Z. F., and Asif, M.: Semi-supervised transient noise suppression using OMLSA and SNMF algorithms, Appl. Acoust., 170, 107–113, https://doi.org/10.1016/j.apacoust.2020.107533, 2020.

Westhausen, N. L. and Meyer, B. T.: Dual-signal transformation LSTM network for real-time noise suppression, in: Interspeech 2020, Shanghai, China, 25–29 October 2020, 2477–2481, https://doi.org/10.21437/Interspeech.2020-2631, 2020.

Wu, J. W., Miao, L. J., Shen, J., and Li, F.: Method for restraining transient noise of FOG based on gain self-compensation,

Journal of Chinese Inertial Technology, 26, 370–374, https://doi.org/10.13695/j.cnki.12-1222/o3.2018.03.015, 2018.

Xie, M., Zhou, K., Huang, Y., and He, M.: A white method of noise suppression for partial discharge based on short time singular value decomposition, Proceedings of the CSEE, 39, 915–922, https://doi.org/10.13334/j.0258-8013.pcsee.172747, 2019.

Yao, X. and Sun, Y.: Noise suppression for surface nuclear magnetic resonance data based on frame energy spectrum subtraction, in: 2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE), Xiamen, China, 18–20 October 2019, IEEE, 55–59, https://doi.org/10.1109/EITCE47263.2019.9095012, 2019.

Zhang, Y.: Simulation of noise suppression for high frequency signals in digital circuit remote network, Comput. Simulat., 37, 471–475, 2020.

Zhao, P., Jiang, Y. Z., Chen, B., Li, C. T., and Zhang, Y. Y.: SLF channel method of noise suppression based on adaptive blanking in local variance domain, Journal of Jilin University (Engineering Edition), 49, 1696–1705, https://doi.org/10.13229/j.cnki.jdxbgxb20180446, 2019.

Zhao, Y., Xin, T., Hu, M., and Pan, S. D.: Noise suppression research of fan shaped superfluid interference grating gyroscope, in: 2019 5th International Conference on Control Science and Systems Engineering (ICCSSE), Shanghai, China, 14–16 August 2019, IEEE, 63–67, https://doi.org/10.1109/ICCSSE.2019.00020, 2020.

Zhu, Y. and Xie, S.: Coherent noise suppression using adaptive homomorphic filtering for wideband electromagnetic imaging system, Sensors-Basel, 191, 56–69, https://doi.org/10.3390/s19204469, 2019.