24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

AN ACOUSTIC FAULT SAMPLE EXPANSION METHOD BASED ON NOISE DISTURBANCE

Na Wei

Nanjing University, Key Laboratory of Modern Acoustics and Institute of Acoustics, Nanjing, China

Linke Zhang*

Wuhan University of Technology, School of Energy and Power Engineering, Wuhan, China
email: lincol@whut.edu.cn

Xiaojun Qiu

University of Technology Sydney, Centre for Audio, Acoustics and Vibration, Faculty of Engineering and IT, Sydney, Australia

Sample expansion is an effective approach to resolve the incomplete sample problem in ship acoustics fault source identification. A novel sample expansion method based on noise disturbance is proposed. Firstly, the expanded samples are generated by adding different signal-to-noise ratio noise to the real acoustics fault samples. Then, the training collection is enhanced by combining the original training samples with the expanded samples. Experimental results on both simulated data and measured data demonstrate that the classifier performance can be improved by the expanded training sample set.
Keywords: acoustic fault, sample expansion, noise disturbance

1. Introduction

In actual ship acoustic fault source identification, it is strongly difficult to obtain fault sample signals due to the background noise, the nonlinearity of structure vibration propagating and the complexity of transmission path and so on, which leads to fault samples deficiencies problem [1].The information deficiencies or small samples phenomenon may result in acoustic fault source identification performance decline greatly. .Various priori information is generally considered to generate expanded samples, which are the effective supplements of actual noise sources [2].

Generally, there are two main methods to construct expanded samples. One is to generate expanded samples by introducing priori knowledge. This approach demands a deep understanding of prior knowledge in this field from researchers, such as Niyogi's method [3], which creates virtual samples by the prior knowledge obtained from a given small training data to improve recognition performance. Similarly, Scholkopf et al. [4] apply prior information in handwritten number recognition to improve the learning machines performance. The other is to expand samples by the idea of disturbance. The essence of this method is to construct the sample directly from the original training samples without prior knowledge. So this method has strong adaptability. Such methods as Bootstrapping [5], Cross-Validation [6], Monte-Carlo simulation [7] and so on has been highly applied to different research areas. Unfortunately, for our ship acoustic fault source expansion problem, such methods may contaminate real sample phase information, which can lead to distortion of signals frequency domain information. Sietsma et al. [8] show that adding noise to training samples can improve the generalization ability of learning methods. Also Bishop [9] proves theoretically that the noise injection is equivalent to the smooth regularization operation. So the noise disturbance method is introduced to expand the ship acoustics fault sample numbers by adding noise into real fault source samples in this paper.

2. Sample expansion based on noise disturbance

The ship acoustic fault signals are collected through multiple channel acceleration sensors arranged on different parts of the ship. The real noise source samples can be obtained by these collected signals. However, in practice the typical collected fault sample number is limited. So it is quite necessary to make full use of these limited real samples. This paper proposes a sample expansion method based on noise disturbance by adding noise into the real fault source samples.

Let $x(t)$ be an N-dimensional real fault signal, the expanded sample $y(t)$ can be generated as

$$y(t) = x(t) + \alpha n(t) \tag{1}$$

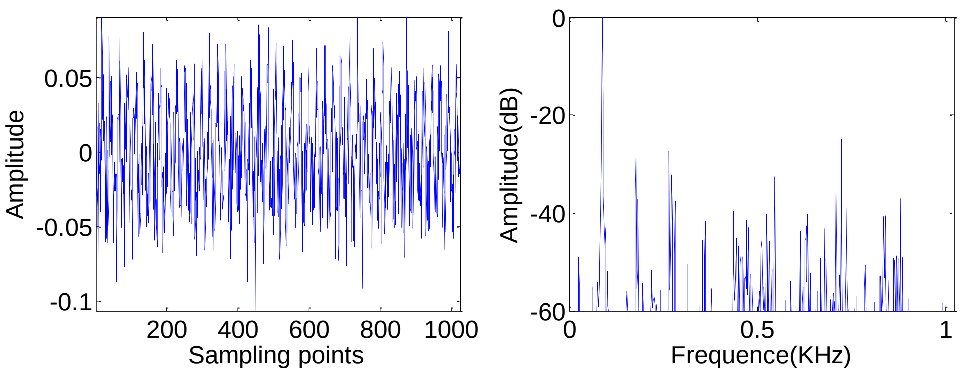
where $n(t)$ is the disturbed noise, and α is a constant, which can be determined by the signal-to-noise ratio (SNR) [10] as

$$\mathrm{SNR} = 10 * \log 10 \left\{ \frac{E\left[\left|x(t)\right|^2\right]}{E\left[\left|n(t)\right|^2\right]} \right\} \tag{2}$$

where $E[\cdot]$ denotes mathematical expectation. The greater the SNR is, the more energy the signal has, and the less energy the noise has conversely. Multi-expanded samples can be generated by disturbing the real sample $x(t)$ with different SNR noise.

Theoretically, noise with arbitrary SNR is can be added to real samples. However, much real sample information will be lost in the case of high-energy injected noise, while the real sample information cannot be supplemented effectively or training data redundancy with unduly low-energy noise.

Figure 1 is one acoustic fault sample acquired by ship 1:1 cabin model, whose operating mode is half-open pump, open motor, vibration exciter with vibration frequency 360Hz, working current 150mV. Fig. 2 shows its expanded samples power spectra with different SNR disturbing noises. It is evident that the fault source signal is buried in the injected noise when the SNR falls on -20 dB, and appears gradually as injected noise energy weakens. Hence, the training data can be effectively supplemented by noise disturbance method for better adjusting to the changing surrounding noise interference, which can be demonstrated by subsequent experimental results.



(a)Time signal (b)Power spectrum

Figure 1: An acoustic fault sample

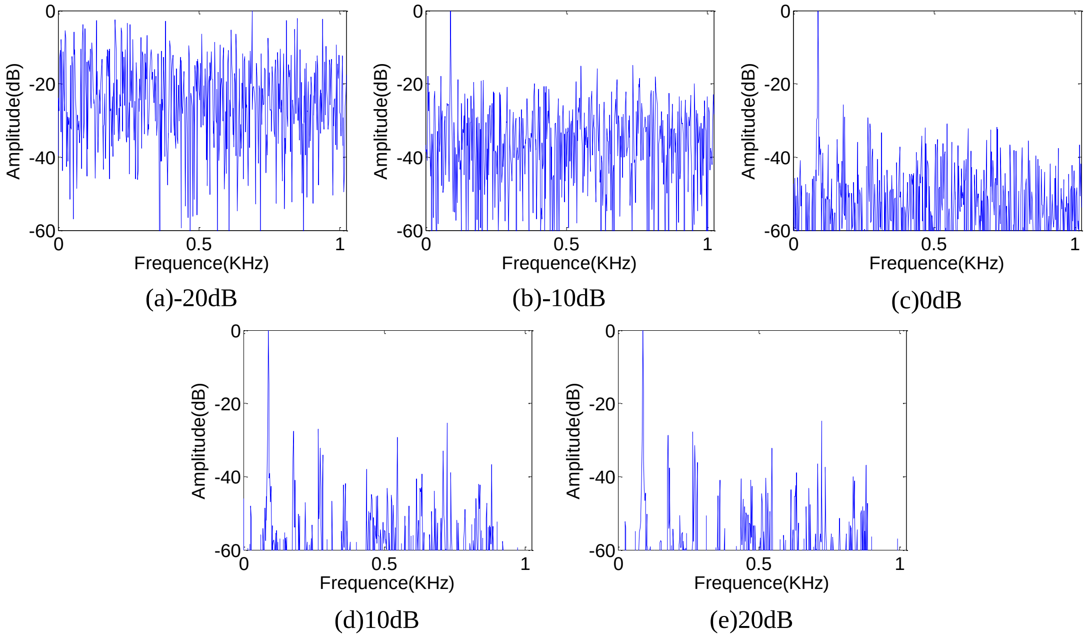


Figure 2: Expanded samples power spectrum with different SNR

3. Performance assessment

We present several experiments both on the simulation dataset Data 1 and measured dataset Data 2 to demonstrate the proposed method effectiveness.

(1) Dataset

Data 1 is simulated time series derived from $x(t) = \exp(\mathrm{j}2\pi ft) + n(t)$, where f is signal frequency, t is sampling time and $n(t)$ is random noise. Two classes of signals are $x_1(t) = \exp(\mathrm{j}2\pi 200t) + \exp(\mathrm{j}2\pi 300t) + n_1(t)$ and $x_2(t) = x_1(t) + 0.1\exp(\mathrm{j}2\pi 350t) + n_2(t)$. Both sampling length is 5120. 512 sampling points are continuous selected along the sampling time sequence as an observation sample. Each class signal generates 10 real samples.

Data 2 is the measured dataset and the information is demonstrated in Tab.1. The length of data sampling is 16384. Along the direction of sampling time, 1024 sampling points are consecutively selected with 50% coverage as an observation sample. Hence, 30 samples are generated in each class. Also 10 samples are selected for each class as real samples.

The positive frequency bands of signal power spectrum are abstracted as the feature vector. So the feature vector dimension is 256 and 512 respectively for Data 1 and Data 2. Fig.3 shows the power spectrum of an actual sample belonging to these two data sets.

Table 1: The information of measured dataset Data 2.

Operating Mode	Measuring Point	Class
Pump (Half Open), Generator, Vibration Exciter-360 Hz (150mv)	8	1
Pump (Half Open), Vibration Exciter-360 Hz (150 mv)	8	2

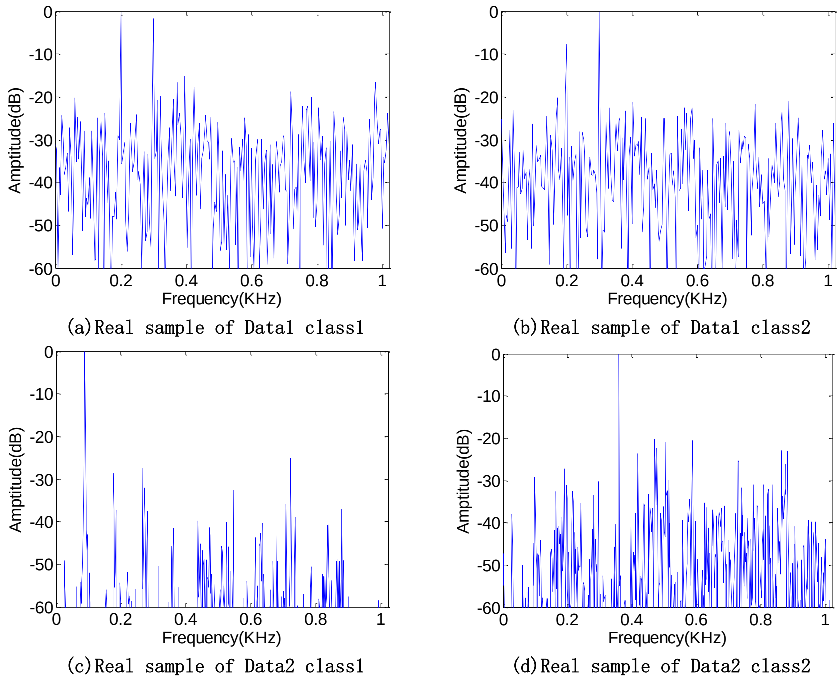


Figure 3: Real samples power spectrum.

(2) Sample expansion

Expanded samples are generated by injecting noise with determinate SNR in various real samples which expand 400 samples. Figs. 4-5 show the projection of the first two main components of expansion samples in Data 1 and Data 2 respectively. Obviously, the two-dimensional projection distribution of the expansion sample is more and more approximate to the real samples with the decrease of injected noise energy.

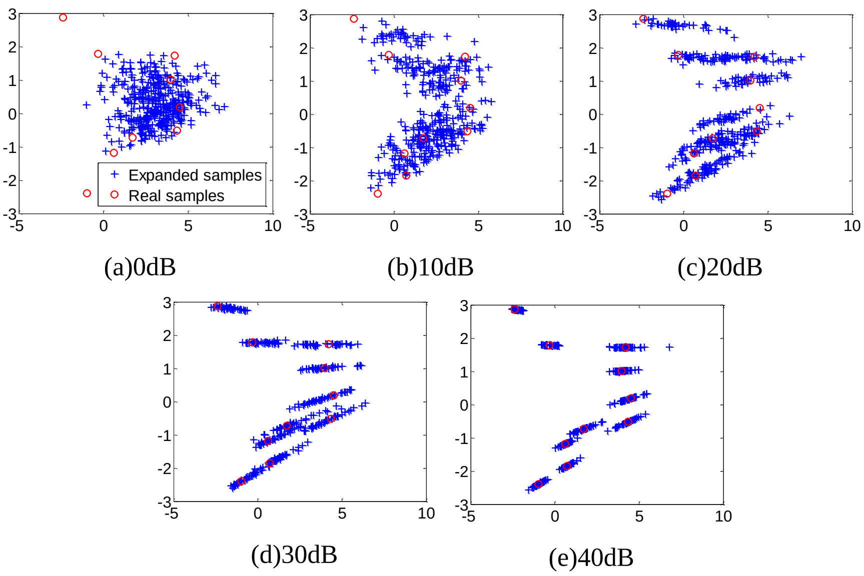


Figure 4: Expanded sample projection of Data 1 Class 1.

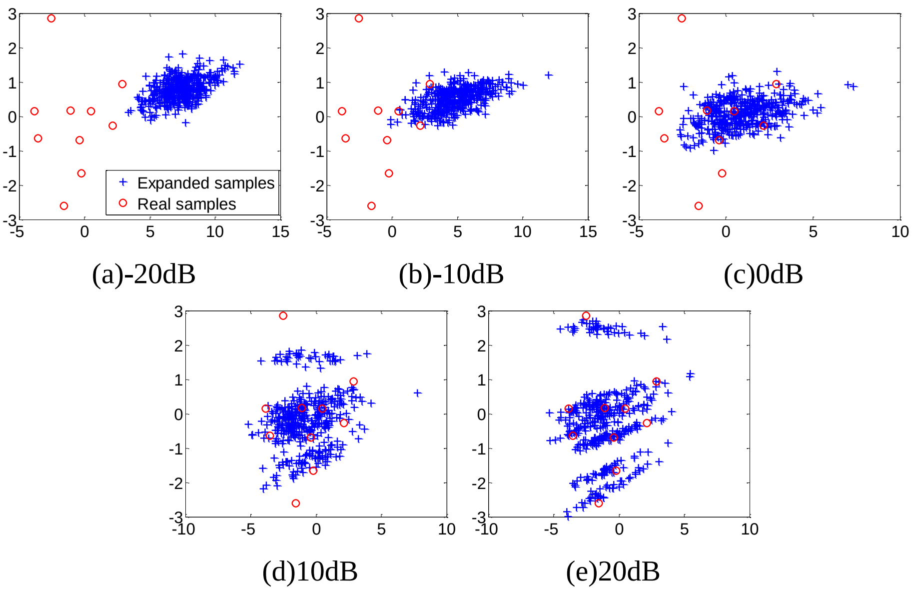


Figure 5: Expansion sample projection of Data 2 Class 1.

(3) Recognition accuracy rate

Furthermore, this section is devoted to the performance analysis of the proposed expansion techniques in correspondence of different numbers for the expanded samples. The recognition accuracy rate is under consideration for artificial neural networks classifier. The training samples are selected as the following three programs:

- Program 1: choose 5 real samples;
- Program 2: choose all 10 real samples;
- Program 3: choose 5 real samples, and also expanded samples of various SNR with an interval of 20;

Testing samples are selected from the residual 5 real samples of program 1. The classifier recognition accuracy is shown in Fig.6.

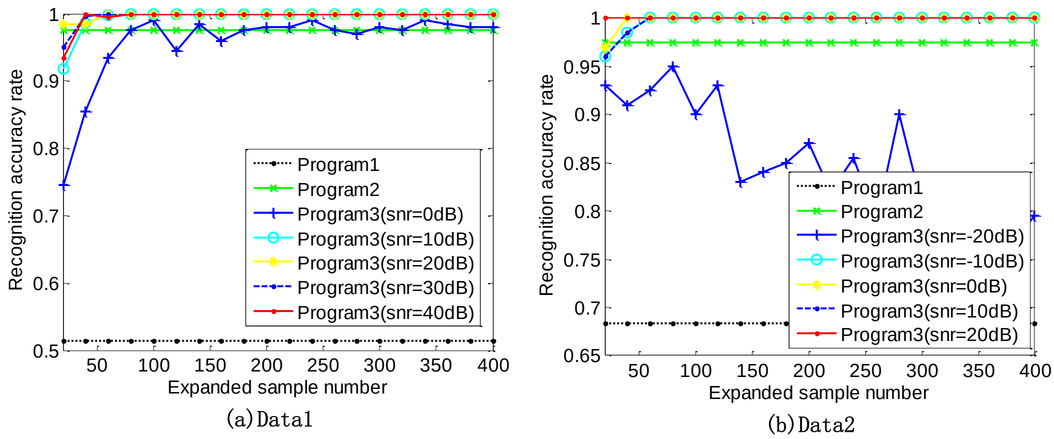


Figure 6: Recognition accuracy rate.

As demonstrated in Fig. 6, for Data 1 the recognition accuracy rates of program 1 and program 2 are 0.5146 and 0.9750 separately. In the case of 0 dB SNR, the recognition accuracy rate fluctuates around 0.9750 as the expansion samples increase, and when the SNR is over 10 dB, the recognition accuracy rate remains 1 when the expanded samples number is greater than or equal to 80. For measured Data 2, the recognition accuracy rate presents a downward trend as the expanded samples number increases when SNR is -20 dB. This behaviour can be explained that there exists great deviation betweenreal samples and expanded samples with -20 dB noise distributed, which have interfered the classifier performance. When SNR is higher than -10 dB, the recognition accuracy rateremains 1 when the sample number is greater than or equal to 60 and especially, when SNR is 20 dB, the recognition accuracy rate have already arrived 1 while expansion sample number is 20.

4. Conclusion

The noise disturbance is applied for acoustic fault sample expansion. The proposed method sufficiently makes use of the real sample information, which addresses the issue of incomplete samples in ship acoustics fault source identification. It is shown experimentally the classifier recognition rate is related to the SNR of injected noise and the number of expansion samples. Under a determinate SNR, high rate of classifier recognition is potentially obtained through small numbers of expansion samples.

REFERENCES

1 Zhang, L. K. *Study on the technique of acoustic fault identification and its application*, PhD thesis, Wuhan, China, Naval University of Engineering, (2006).

2 Zhang, L. K., Cui, L. L. Review on mechanical noise source classification based on small sample data set. *Journal of Ship Mechanics*, **15**(8): 940-947, (2011).

3 Niyogi P., Girosy F., Poggio T. Incorporating Prior Information in Machine Learning by Creating Virtual Examples. *Proceedings of IEEE*, **86**(11): 2196-2209, (1998).

4 Schölkopf B., Simard P., Vapnik V., et al. Prior knowledge in support vector kernels. *Proceedings of Conference on Advances in Neural Information Processing Systems*, 640-646, (2003).

5 Hu Z. D., Cao Y., Zhang S. F., et al. Bootstrap method for missile precision evaluation under extreme small sample test. *Systems Engineering & Electronics*, **30**(8): 1493-1497, (2008).

6 Lachenbruch P.A, Mickey R. M. Estimation of error rates in discriminant analysis. *Technometrics*, **10**: 1-1, (1968).

7 Vorechovsky M., Novak D. Correlation control in small-sample Monte Carlo type simulations I: A simulated annealing approach. *Probabilistic Engineering Mechanics*, **24**(3): 452-462, (2009).

8 Sietsma J., Dow R. J. F. Creating artificial neural networks that generalize. *Neural Networks*, **4**(1): 67-79, (1991).

9 Bishop C. M. Training with noise is equivalent to tikhonov regularization. *Neural Computation*, **7**(1): 108-116, (1995).

10 Bell M. R. Information theory and radar waveform design. *IEEE Trans. Inform. Theory*, **39**(3): 1578-1597, (1993).