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

MULTISCALE SPECTRAL SUBTRACTION METHOD AND SUPPORT VECTOR MACHINE FOR CLUSTERING HEAVY BACKGROUND NOISE CLASSIFICATION

Yu-Ting Tsai

Feng Chia University, Bachelor's Program in Precision System Design, Taichung, Taiwan
email: charles.brown@sheffield.ac.uk

Jin H. Huang

Feng Chia University, Department of Mechanical and Computer-Aided Engineering, Taichung, Taiwan,

For measuring the rotation engine vibration signals, the mechanical or acoustical sensors are difficult to get the clean signals from the heavy background noise of the machinery environment. The noise reduction algorithm may anticipate to decrease the background noise without disturbing the unpolluted signal. The spectral subtraction methods for noise attenuation case are reached for many years due to its simple and effectively. In this paper, the proposed algorithm incorporates the multiscale spectral subtraction method (MSS) with voice activity detector (VAD) and supports vector machine (SVM) for removing the background noise, since SVM can be required for distinguishing a multivariate signal into separated subcomponents. From the analysis case for the rotating machinery, MSS shows that the low frequency with their harmonics frequencies can be tracked through the clustering of the data from SVM. The present method is helpful in background noise and signal compartmentalization as its application can significantly lower the requirement for labelled training cases in both the inductive noise and wanted signals.
Keywords: spectral subtraction method, support vector machine, noise reduction

1. Introduction

Regarding to Industrial 4.0, numerous manufacturing machines are monitored by the signal data for enabling the cyber-physical system (CPS) technologies [1]. CPS vertically integrates the artificial intelligence system for useful data mining and logistics system for automatic machinery decision. Acoustical noise separation for the feature signal enhancement is an interesting study for improving the ability of data analysis. The voice activity detector (VAD) dominates an important role in increasing the signal capacity for noise subtraction. The VAD based on the statistical model such as smoothed likelihood ratio [2], minimum mean-square error (MMSE) [3], and noise power spectral density [4] can remove the noise from the given unclear noisy speech signal. Some of signal enhancement systems consist of the frequency domain spectral computation for minimum statics noise estimation [4, 5]. A lower false alarm rate and higher non-speech hit rate and for smooth spectrum VAD are proposed by Ramírez et al. [6]. An optimally-modified log-spectral amplitude speech estimator and a minima controlled recursive averaging (MCRA) noise estimation approach [7] for robust speech enhancement can achieve the noise suppression for weak speech components. Although the above methods can reduce the noisy signal significantly, the recently researches are still investigating the solutions for non-stationary noise and low clean component from the noisy signal.

Nearly years, machine learning approaches can illustriously evaluate the clean features using the supervised learning model. Fu et al. [8] proposed the signal to noise ratio convolutional neural network model for denoising, which may be much more efficient in disentangling noise and speech features than the transitional deep neural network model. The integration learning model of multi-style training and deep learning methods [9] relies on deep denoising autoencoders to emphasize on the most discriminative patterns for most learning test sets. Additionally, the most of speculate non-speech sections in the signal may require the noise classifier to construct the non-stationary environment sound. A noise robust speech processing [10] feature with two-class support vector machines (SVM) are proposed to train the different background noises for VAD. The likelihood ratio test for SVM proposed by Jo et al. [11] can adapt the decision rule for determining the threshold value of noisy spectrum for objective measurement.

Due to these advance researches, the noise reduction algorithm based on VAD and machine learning model may anticipate get the clean component of the vibration sound. In this study, we present a multiscale spectral subtraction method with the statistical model-based algorithm incorporated the multiclass SVM model attempt to reach the excellent noise suppression from the rotating machine signal with undesired noise. The paper is organized as follows. The proposed method is described in Section 2. In Section 3, we discuss the comparison results of the noise signal of the computer numerical controlled (CNC) machine from the experimental environment. Finally, the conclusion summaries the ability of clean components determination by proposed method.

2. Method description

Suppose a measure of the discrete cutting vibration signal $x(t_n)$, t_n is the time series. The spectrum of $x(e^{j\omega})$ from $x(t_n)$ can be obtained by discrete Fourier time-frequency conversion. The following equation shows the relationship between the estimated enhancement signal $y(e^{j\omega})$ and the inherent characteristic noise signal $n(e^{j\omega})$:

$$x(e^{j\omega}) = y(e^{j\omega}) + n(e^{j\omega}) \tag{1}$$

By considering the complex Gaussian probability density functions [11], the following hypotheses can be obtained remove the inherent noise signal $n(e^{j\omega})$:

$$p(x(e^{j\omega}) \,|\, n(e^{j\omega})) = \frac{1}{g_{n,e^{j\omega}}} e^{\left(-\frac{\left|x(e^{j\omega})\right|^2}{g_{n,e^{j\omega}}}\right)} \tag{2}$$

and

$$p(x(e^{j\omega}) \,|\, [y(e^{j\omega}) + n(e^{j\omega})]) = \frac{1}{g_{n,e^{j\omega}} + g_{y,e^{j\omega}}} e^{\left(-\frac{\left|x(e^{j\omega})\right|^2}{g_{n,e^{j\omega}} + g_{y,e^{j\omega}}}\right)} \tag{3}$$

where $g_{n,e^{j\omega}}$ and $g_{y,e^{j\omega}}$ are the variances of noise and estimated clean component. From equations (1) and (2), the desired amplitude estimator from likelihood ratio of the N_{th} frequency band to get the priori and a posteriori SNR [5] can be derived as follows:

$$\Lambda(e^{j\omega}) = \frac{p(x(e^{j\omega}) \,|\, n(e^{j\omega}))}{p(x(e^{j\omega}) \,|\, [y(e^{j\omega}) + n_N(e^{j\omega})])} = \frac{1}{1+\alpha} e^{\frac{\alpha \cdot \beta}{1+\alpha}} \tag{4}$$

where

$$\text{Priori SNR: } \alpha = \frac{g_{y,e^{j\omega}}}{g_{n,e^{j\omega}}} \tag{5}$$

and

$$\text{Posteriori SNR}: \ \beta = \frac{x(e^{j\omega})}{g_{n,e^{j\omega}}} \tag{6}$$

Then, the SVM-based VAD [11] is suggested by assuming the equation of a Gaussian type of hyper-plane is:

$$J(\Lambda(e^{j\omega})) = \sum_{i=1}^{K} \lambda_i b_i e^{\left(\frac{-0.5\left(\Lambda_i(e^{j\omega})-\Lambda(e^{j\omega})\right)^2}{\sigma^2}\right)} + \varepsilon_i \tag{7}$$

where ε is the real number. For SVM, the way is computationally simpler to solve the dual quadratic programming problem by take the Lagrange multipliers λ_i:

$$L(w,c,\lambda) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \lambda_i \left[b_i \left(J(\Lambda(e^{j\omega})) \right) - 1 \right] \tag{8}$$

Appear to meet the conditions of the extreme points of equation (8), it will change with the extreme area by accounting Karush-Kuhn-Tucker (KKT) condition, we get:

$$\max\{L\} = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{ij} \lambda_i \lambda_j b_i b_j \Lambda_i \Lambda_j \tag{9}$$

$$\sum_{i=1}^{N} \lambda_i b_i = 0, \ \lambda_i \geq 0 \tag{10}$$

In particular, the $b_j J(\Lambda_j) - 1 = 0$ would give the value of ε at the solution with nonzero λ_j. By substituting $\Lambda(e^{j\omega})$ to the SVM, the occurring feature noise sequence from the equation (2) can be estimated. Then, the estimated clean component value of $y(e^{j\omega})$ can be obtained by using the multiscale spectral subtraction method:

$$\hat{y}(e^{j\omega}) = \left[\left| x(e^{j\omega}) \right| - \mu E\{ \left| J(\Lambda(e^{j\omega})) \right| \} \right] e^{j\theta_x(e^{j\omega})} \tag{11}$$

where θ_x is the phase component of $x(e^{j\omega})$. The constant value of amplitude threshold μ is used to control the situation of $\left|\hat{y}(e^{j\omega})\right| \leq \max\{J(\Lambda(e^{j\omega}))\}$ by adjust the size of the approximate value $0.8 \leq \mu \leq 1$.

3. Result and discussion

3.1 Training cases setup

A rotating vibration signal is divided into learning (training samples) and testing (predict samples). In this study cases, Most of the learning and testing data are captured from the ideal CNC machine with no chatting vibration for assuming there are other non-stable cut the situation. Minimize the impact of fixed ambient noise on the cutting signal which may be very suitable for training cases. The measurement process is shown in Figure 1, and the structure of noisy clean signal pairs for training the data set are shown in Fig. 2.

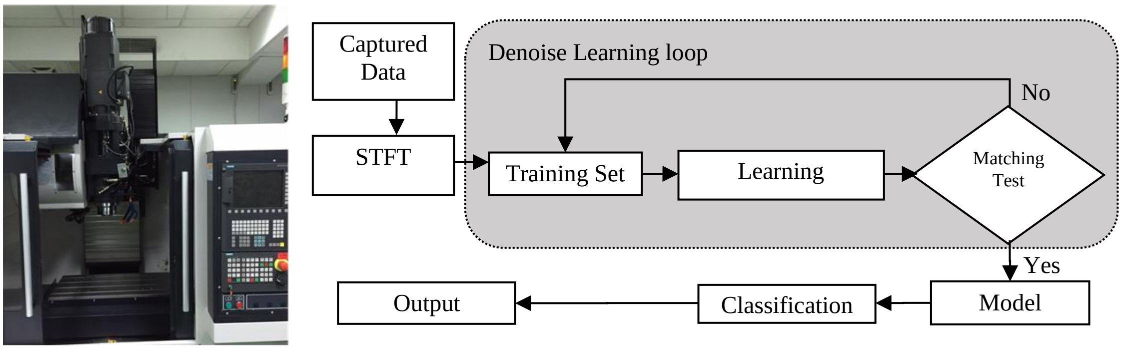

Fig. 1: The measurement setup and the denoise learning process.

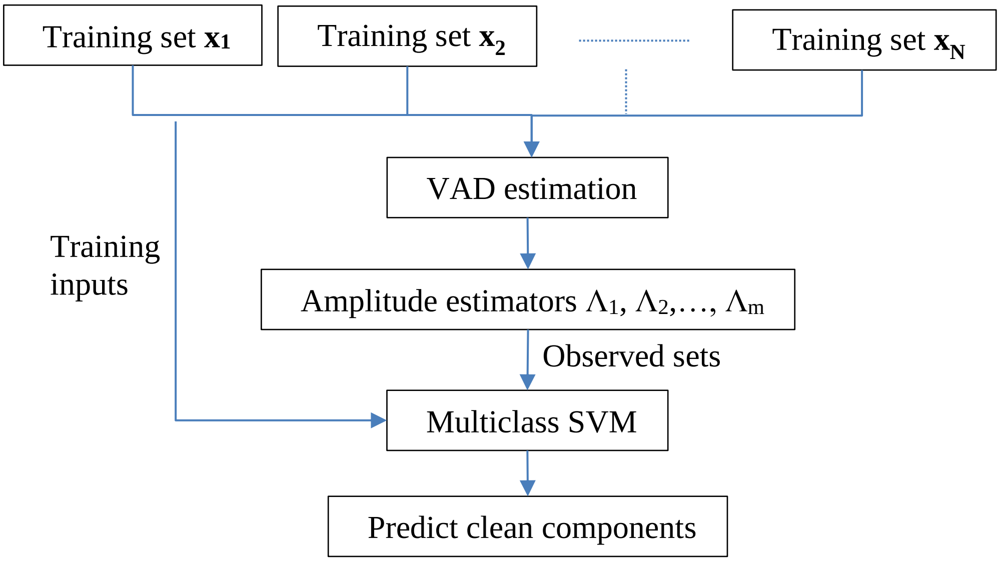

Fig. 2: The structure of proposed method for noise suspension.

3.2 Experimental results

In this Section, we investigate the original vibration signal which captured from the CNC machine in the noisy environment. The sample rate for the inputs signal is 16000 Hz. The length of total samples is 160000 as ten seconds for training test. Mean and variance normalization was used to the training input vectors to make learning process steadier. We set m = 50 according to the discretize indices of data were extracted from the original waveforms (Fig. 3) as the acoustic noise feature. The results of the proposed method are shown in Fig. 4(a) to 4(d). The estimation of the clean and noise components from the proposed method. As a result, the nonstationary acoustic noise components can be discriminated very well. The big differences are observed from time domain signals as shown in Fig. 4(a) and 4(c). The clean signal (in range of samples 1E4 to 3E4) indicates that the nonstationary situation of rotating spindle in start running , and it may not be reduce in noise estimation. The result of Welch PSD of clean component in Fig. 4(b) shows the better SNR which compared than noise component in Fig. 4(d). The validation results show that the proposed method has ability to distinguish the clean and noise components from the rotating CNC machine in the noisy environment.

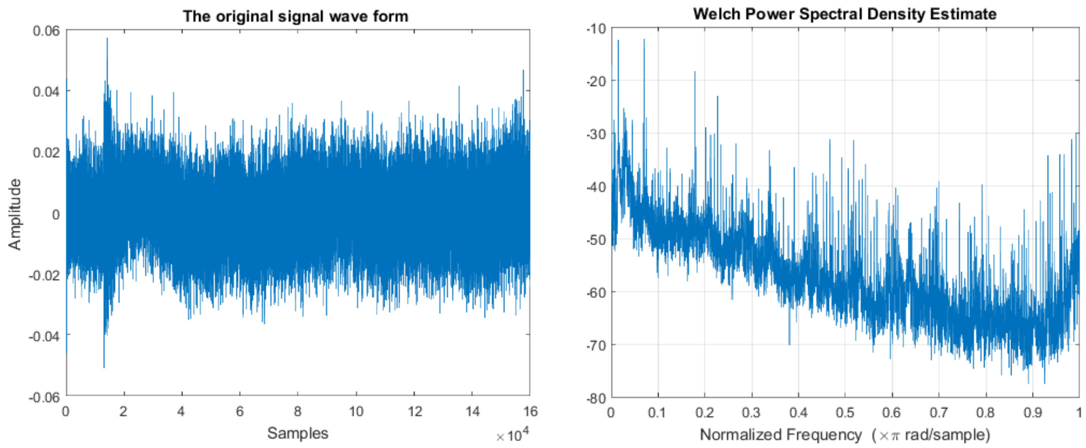

Fig. 3: The noisy signal for experimental test.

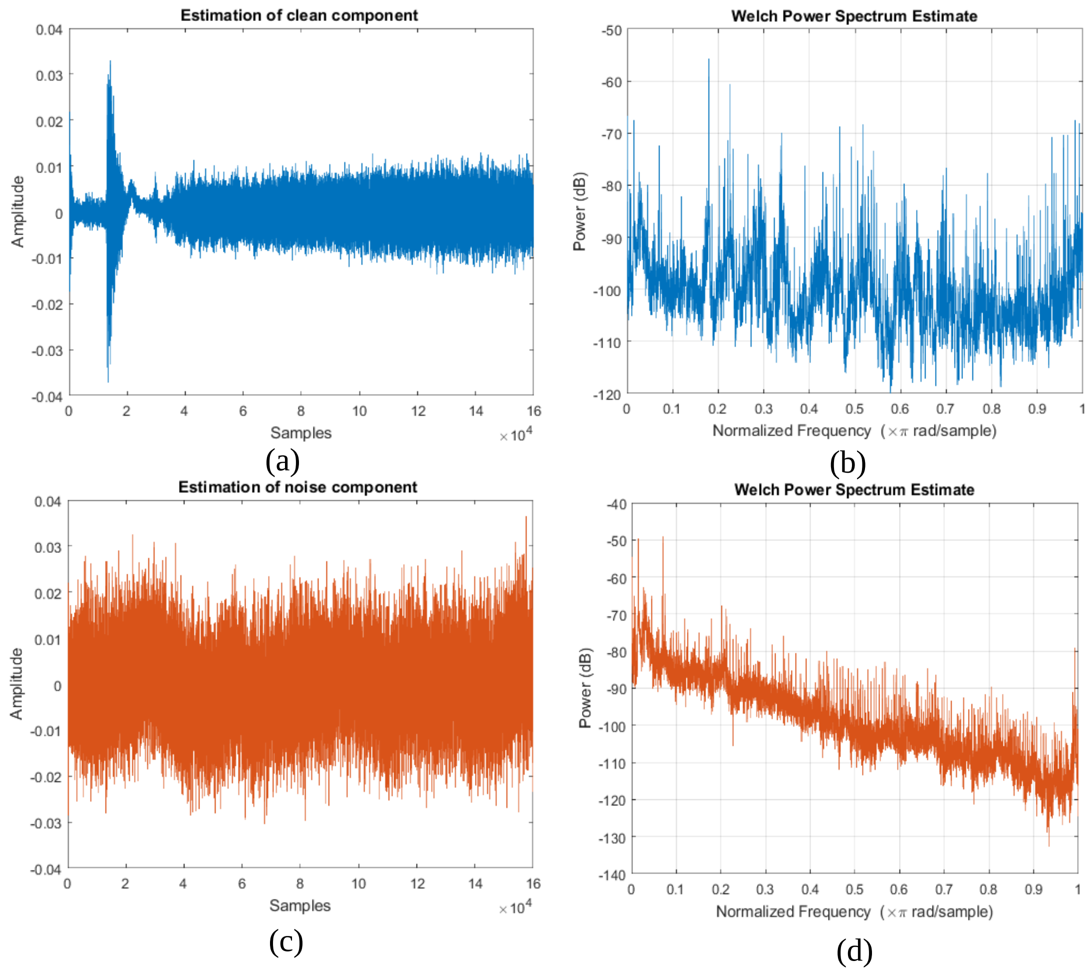

Fig. 4: The prediction of clean and noise components from the proposed method. (a) Time domain of the estimated clean component, (b) PSD frequency spectrum of the estimated clean component, (c) Time domain of the estimated noise component, (b) PSD frequency spectrum of the estimated noise component

4. Conclusions

In this study, the proposed method incorporates the multiscale spectral subtraction method (MSS) with voice activity detector (VAD) and supports vector machine (SVM) for removing the background noise. The case of results shows the sufficient noise signal can be determined from the low SNR signal. The contribution of this paper is that the SVM can effectively extract the features of clean components from the noisy environment. The presented results also show that the better SNR determined from clean components can be treated which compares to the noise component. In the future, we will discover more possible advance machine learning technique with the other objective functions for the task of noise reduction cases.

Acknowledgements

The authors would like to thank the Ministry of Science and Technology of Taiwan for its financial support specified under Contract Nos. MOST 105-2221-E-035 -034 -, MOST 105-2218-E-035 -005 -, and MOST 105-2511-S-035 -012 -MY3.

REFERENCES

1 Germany Trade & Invest, *Industrie 4.0: smart Manufacturing for the future*, [Online.] available: https://www.gtai.de/GTAI/Content/EN/Invest/_SharedDocs/Downloads/GTAI/Brochures/Industries/industrie4.0-smart-manufacturing-for-the-future-en.pdf

2 Yong, D. C. and Ahmet, K. Analysis and improvement of a statistical model-based voice activity detector, *IEEE Signal Processing Letters*, **8** (10), 276–278, (2001).

3 Ephraim, Y. and Malah, D. Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator, *IEEE Trans Acoustics Speech and Signal Processing*, **32** (6), 1109–1121, (1984).

4 Rainer, M. Noise power spectral density estimation based on optimal smoothing and minimum statistics, *IEEE Trans. Speech and Audio Processing*, **9** (5), 504–512, (2001).

5 Sohn, J. and Sung, W. A voice activity detector employing soft decision based noise spectrum adaptation, *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, **1**, 365–368, (1998).

6 Ramírez, J., Segura, J.C., Benítez, M.C., Torre, Á.D. and Rubio, A.J. An effective subband OSF-based VAD with noise reduction for robust speech recognition, *IEEE Trans. Speech Audio Process*, **13** (6), 1119–1129, (2005).

7 Israel, C. and Baruch, B. Speech enhancement for non-stationary noise environments, *Signal Processing*, 81, 2403–2418, (2001).

8 Fu, S. W., Tsao, Y. and Lu, X. SNR-aware convolutional neural network modeling for speech enhancement, ISCA Interspeech, San Francisco, 8–12 September 2016.

9 Lin, P., Lyu, D., Chen, F., Wang, S. S. and Tsao, Y. Multi-style learning with denoising autoencoders for acoustic modeling in the internet of things (IoT), *Computer Speech and Language*, 00, 1–15, (2017).

10 Saeedi, J., Ahadi, S. M. and Faez, K. Robust voice activity detection directed by noise classification, *Signal, Image and Video Processing*, **9** (3), 561–572, (2015).

11 Jo, Q. H., Chang, J. H., Shin, J. W. and Kim, N. S. Statistical model-based voice activity detection using support vector machine, *IET Signal Processing*, 3 (3), 205–210, (2009).