Signal Reconstruction using Inter-microphone Time Difference for Bearing Fault Diagnosis

Anindita Adikaputri Vinaya
Engineering Management Department
Universitas Internasional Semen Indonesia
Gresik, Indonesia
anindita.vinaya@uisi.ac.id

Niswatun Faria
Engineering Management Department
Universitas Internasional Semen Indonesia
Gresik, Indonesia
niswatun.faria@uisi.ac.id

Abstract— **The separation of mixed signals is one of the focuses of research related to acoustic signal processing. Mixed signals can be obtained from unpredictable environments. In this research, we will use the inter-microphone time difference (ITD) approach to reconstruct the acoustical signal of machinery, which is the signal target of some interfering signal. We add artificial noise as interfering signals in our setup experiment. The signal reconstruction results of each interference signal are represented in the form of a spectrogram. Some of the reconstructed spectrograms have the highest amplitude, around 22x rpm. The harmonic frequency pattern shows that bearing fault diagnosis is found at 6x rpm, 8x rpm, 17x rpm, 20x rpm, 22x rpm, and 24x rpm. It corresponds to the baseline signal. The spectrogram of the reconstructed signal from the sine noise using ITD is closest to the signal target as a bearing fault signal with an LSD of 1.261.**

Keywords—mixed signal, bearing fault, ITD

I. Introduction

The human's ear is an acoustic sensor that can differentiate the incoming sound direction and separate a sound. It became the foundation of the research development on the localization and separation of signals [1-3]. The presence of interfering signals at specific frequencies may also interfere with the sound source as the signal target [4]. In 2016, Lopez tried to separate the signal of unmanned vehicles by considering noise variations. The interfering signals are colored noise, wind noise, and propeller noise [5]. The concept of acoustic signal separation has been widely used to identify machine conditions [4-6].

In their study, Gao Lixin *et al.* mentioned that some of the energy from engine vibration is transmitted through sound emission. A microphone detects machine failure instead of a vibration sensor [6]. Atmaja et al. identify the failure based on the estimation results of the mixed-signal separation [4]. The independent component analysis method is used to separate three engine sources. In their research, the signal interferes with stationary signals from other machines [4]. An increasing amount of research has been carried out with different methods and types of signals. In 2020, Ge *et al.* tried to reconstruct a machine's vibration signal with a broken rolling element bearing using an empirical ensemble of decomposition and entropy modes. The vibration signal with white noise will be reconstructed compared to the less noise signal or the original signal [7]. In 2020, Vinaya *et al.* used the non-negative method matrix factorization to separate the pump with bearing failure sound signal from the interference signal where the interfering signal is an acoustic signal from another machine [8]. Defects in bearings can be caused by several things such as cracks or breaks, lack of lubrication, rust, and excessive loads. The peak amplitudes classify defects in bearings at high frequencies in the vibrational and acoustic spectrum. Harmonic frequency charts are quickly spotted in this type of defect [7,8].

Inter-microphone time difference (ITD), Inter-microphone level difference (ILD), and inter-microphone phase difference (IPD) are some commonly used algorithms for localization sources. ILD has good results when the signal is at a high frequency [3,9]. A combination of ITD and IPD have commonly used because they provide exemplary performance in resource localization [3]. Smith *et al.* used the ITD method to determine the source location when the noise occurred [10]. There are two omnidirectional microphones to record mixed signals. ITD with cross-correlation is compared to spike signal.

Furthermore, the result shows the cross-correlation performs well when the microphone is perpendicular to the primary sound source [10,11]. By 2015, Lee et al. used the ITD method to localize the sound source that is applied in the interactions between humans and robots. The results show that this method can localize the Gaussian source well [11].

To get the source signal, we need a method to separate the mixed sound to be estimated and used to identify the damage. The ITD method with the phase difference approach has been used by Chanwoo et al. to separate the sound signal that interferes with other sound signals [10]. Farid used the combination method of ITD and ILD in 2016 for speech signal segregation [12]. Kantue *et al.* in 2020 also used the concept of the difference in the direction of arrival received by the sensor to detect rotor failure at unmanned vehicles [13]. In 2019, Pan used sound signals to detect rotor-stator friction failure. The concept of the difference in arrival time received by the sensor is also used in his research [14].

We propose the extraction of non-speech signals, i.e., acoustic signals from machines. It is a considerable challenge nowadays due to the limited variations of the alphabet and intonation compared to speech signals [1,8]. This study aims to reconstruct the acoustic signal from a machine with bearing defects. The phase difference information received by the two cardioid microphones was used to estimate the ITD. The method will be used to reduce the pump signal from noise, and a spectrogram will be used to represent the signal. The reconstruction results will be compared with the original signal, which will later be used to verify the machine's bearing defects.

II. Inter-Microphone Time Difference

Several processes must be passed to separate the signal. The recorded signal by two microphones will be transformed into the frequency domain to increase the sparsity. The ITD will be calculated based on the phase difference from the information received by the sensor. The ITD is then compared to the searched threshold based on the minimum function of the cross-correlation [9][11].

In order to estimate ITD, we use an approach to IPD calculation. This IPD is obtained based on the phase difference received by both sensors. In this study, we used two microphones. Microphones 1 and 2 will receive different information. The signals received by both microphones will be transformed into time-frequency domains with short-time Fourier transform (STFT). The following is the application of the Hamming window before the signal is transformed into the time-frequency domain. [9].

$$Y_1(n,m) = \sum_{n=f.L_f-(L_w-1)}^{m,L_f} w(f.L_f - n).y_1(n).e^{-j.2.\pi.n.k/N}$$
(1)

$$Y_2(n,m) = \sum_{n=f.L_f-(L_w-1)}^{m,L_f} w(f.L_f - n).y_2(n).e^{-j.2.\pi.n.k/N}$$
(2)

$Y_1(n,f)$ and $Y_1(n,f)$ are the signals received by the first microphone and a second microphone in the time-frequency domain of the STFT. Index frame is written with m, L_f denotes the number of samples between frames, L_w denotes frame length, and W indicates the window function used in all frames. The signal will be changed into the time-frequency domain through Fourier transform.

$$Y_1(m,f) = \sum_{s=0}^{S} Y_s(n,f)$$
(3)

$$Y_2(m,f) = \sum_{s=0}^{S} Y_s(n,f).e^{-j.2.\pi.\frac{f}{N}.\tau_s(m,f)}$$
(4)

If both sensors receive the number of sources of S, then the total source is represented as $0 \leq s \leq S$. Further ITD can be obtained by the equation:

$$|\tau_s(m,f)| \approx \frac{1}{\left|2.\pi.\frac{f}{N}\right|} \min |\theta_{Y1(m,f)} - \theta_{Y2(m,f)} - 2.\pi.r|$$
(5)

Based on equation 5, we can sort the ITD information received by the sensor. The target source's ITD is zero if the source is perpendicular to both sensors [3].

There are two masks in the ITD algorithm that complement each other [8]. One of the masks is to identify components in time-frequency domains that are signals source (target), while the other is to identify time-frequency components in interfering signals. The masks generate two different spectra in terms of power to show the signal target. [3] [9].

$$P_T(m|\tau_0) = \sum_{f=0}^{N-1} |Y_T(m,f)|^2$$
(6)

$$P_I(m|\tau_0) = \sum_{f=0}^{N-1} |Y_I(m,f)|^2$$
(7)

P_T and P_I represent the non-negative power of signal target and an interfering signal, while Y_T and Y_I are the signal target and interfering signal. The Y_T and Y_I spectra obtain through the equation:

$$Y_T(m,f|\tau_0) = \bar{Y}(m,f).\mu_T(m,f)$$
(8)

$$Y_I(m,f|\tau_0) = \bar{Y}(m,f).\mu_I(m,f)$$
(9)

$\bar{Y}(m,f)$ is the average spectrum function of the signal received by both $(Y_1(m,f), and\ Y_2(m,f))$, while $\mu_T(m,f)$ and $\mu_I(m,f)$ are functions masking signal targets and interfering signals. $\mu_T(m,f)$ is 1, if ITD threshold $|\tau_{TH}(m,f)|$ less than τ_0. It means the component in the time-frequency is the signal source (target). $\mu_I(m,f)$ is 1, if ITD threshold $|\tau_{TH}(m,f)|$ more than τ_0. It means the component in these time-frequencies is interfering signals. The selection of τ_0 parameters is obtained by minimizing the cross-correlation coefficients expressed in the following equation:

$$\hat{\tau}_0 = arg_{\tau_0}^{min}|C_{T,I}(\tau_0)|$$
(10)

$$C_{T,I}(\tau_0) = \frac{\frac{1}{N}\sum_{m=1}^{M}\gamma_T(m|\tau_0).\gamma_I(m|\tau_0)-\mu_{\gamma T}.\mu_{\gamma I}}{\sigma_{\gamma T}.\sigma_{\gamma I}}$$
(11)

where $C_{T,I}(\tau_0)$ is the cross-correlation coefficient, $\sigma_{\gamma T}$ and $\sigma_{\gamma I}$ are the standard deviation of power $\gamma_T(m|\tau_0)$ and $\gamma_I(m|\tau_0)$ derived from the decreasing of P_T and P_I power functions based on nonlinearity law on power signal obtained [9,11].

III. Research Method

The STFT implementation process consists of sampling, frame blocking, windowing, and FFT to calculate the covariance matrix coefficients of the information received by both microphones. This study used a sampling frequency of 16kHz, a frame length of 75ms, and an FFT size of 2048. Based on the phase difference of information received by both microphones, the ITD is known. The average spectrum function of the information received by both microphones will be calculated, which will be used to separate the interfering and target based on masking. The following process is masking interfering signals, and this target aims to identify the spectra in the time-frequency domain that is the source signal (target) and interfering signals through the ITD threshold and the obtained ITD. In order to distinguish interfering signals and targets, a masking function will be applied, and cross-correlation coefficients will be minimized to obtain optimal threshold ITD. In the final stage, the signal will be restored from the frequency domain into the time domain so that the source signal can be reconstructed again.

This study collected data by baseline signal recording and mixed-signal recording with a sampling rate of 16 kHz. The sound source of the machine used in this research as a target is the water pump running at 3000 rpm or 50 Hz with bearing defects. The recording process is done by convolutive mixing in the room with dimensions 3mx4mx3m. The recording time

is 4secs. At baseline signal recording, a microphone array is placed at a distance of 5cm from a speaker that emits the sound of a pump with a bearing defect. This baseline signal will be used as a comparison with the reconstructed signal based on ITD. Two cardioid microphones are placed at 90cm from speakers at mixed-signal recording, as seen in Fig.1. In this study, pink noise, white noise , sine noise and noise from another pump with unbalance conditions is used as interfering signals or noise. The colored noise level used is 0 dBFs. The volume of both speakers and the gain set on the Steinberg-usb audio interface is 50% of the maximum value.
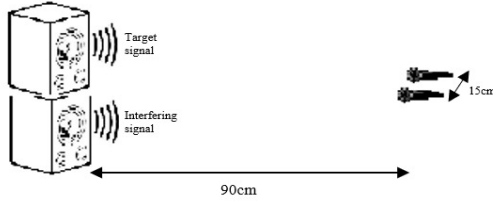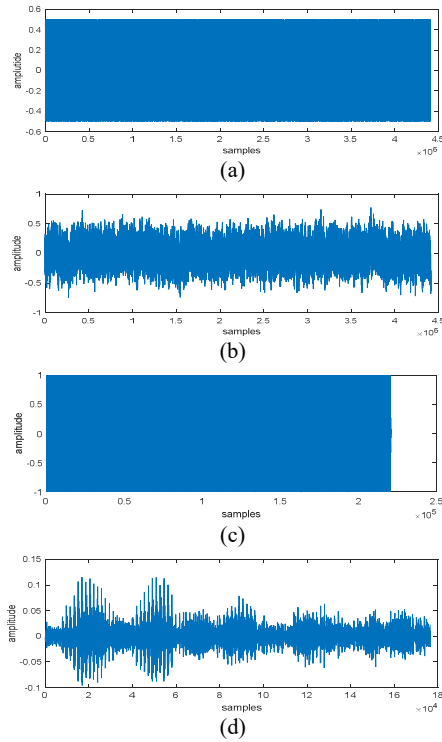

FIGURE 1. Mixed-signal recording

FIGURE 2. Interfering signals (a) white noise (b) pink noise (c) sine noise (d) pump noise

IV. RESULTS

In this study, we used a spectrogram to represent the signal. The dark red color in the spectrogram indicates a high signal amplitude. The highest signal amplitude of the baseline signal reaches -55.24dB at 1120Hz or about a frequency of 22x rpm. Based on Fig.3, peak amplitude at high frequencies is characteristic of bearing fault. Another characteristic of bearing faults in the presence of harmonic frequencies as indicated by horizontal lines all the time in multiples of the

engine rotational frequency (3000rpm or 1x rpm) which is seen at around 6x rpm, 8x rpm, 17x rpm, 20x rpm, 22x rpm, and 24x rpm. The harmonic frequencies in the baseline signal will be compared with the processed signal.
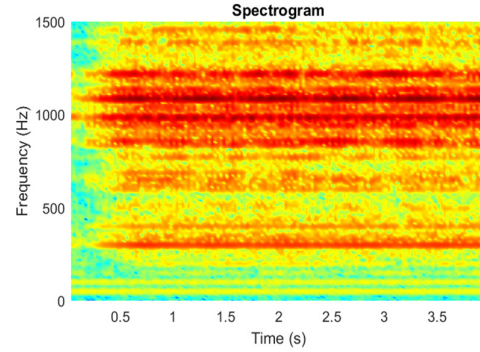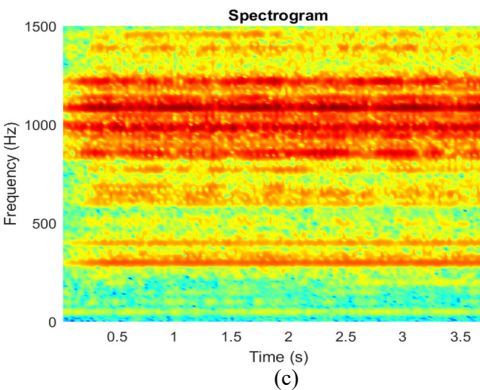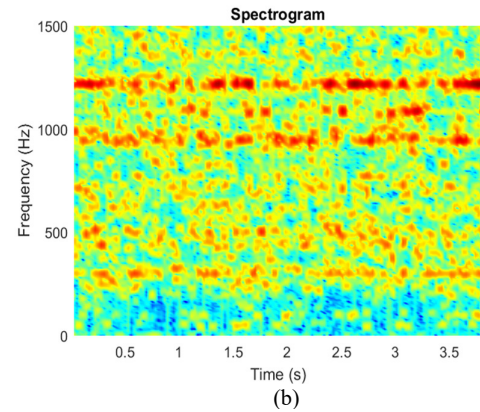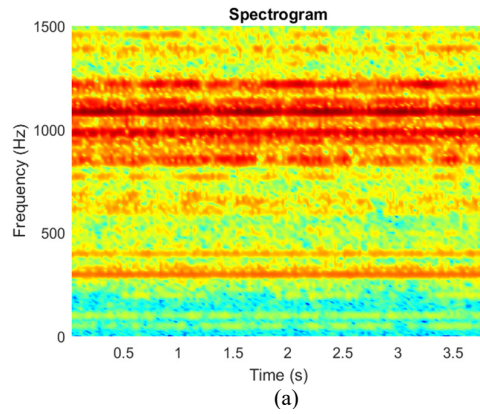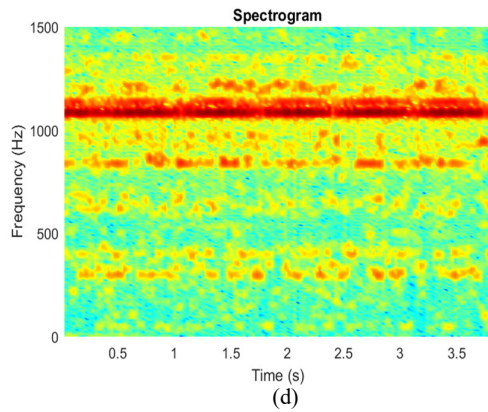

FIGURE 3. Spectrogram of baseline signal as a signal target

(a)

(b)

(c)

FIGURE 4. Spectrogram of a reconstructed signal target from (a) white noise (b) pink noise (c) sine noise (d) pump noise

The reconstructed signal can be seen in Fig4. In Fig. 4a, the spectrogram of the reconstructed signal from white noise has the highest amplitude located at about 22x rpm. The frequency corresponds to the characteristics of the baseline signal as the signal target. Spectrograms with similar characteristics are also found in the reconstructed signal from sine noise (figure 4. c) and pump noise (figure 4.d). Harmonic frequencies can also be known based on the resulting spectrogram, as shown in Table 1. Bearing defects in the baseline signal are indicated by harmonic frequencies around 6x rpm, 8x rpm, 17x rpm, 20x rpm, 22x rpm, and 24x rpm. The reconstructed signals from white noise and sine noise have a pattern that matches the signal target. This can be seen from the harmonic frequency pattern according to the baseline signal. Based on the results obtained, the reconstructed signal from white noise and sine noise spectrogram and harmonic frequency pattern is consistent with the baseline signal indicating bearing defects. The other two spectrograms (Figure 4. b and Figure 4.d) have different patterns of the signal target.

TABLE 1. Comparison of harmonic frequencies and LSD

Spectrum of	Harmonic frequencies of bearing fault						LSD
Baseline signal	6x	8x	17x	20x	22x	24x	-
Recostructed signal from white noise	6x	8x	17x	20x	22x	24x	1.749
Recostructed signal from pink noise	-	-	-	-	-	-	2.012
Recostructed signal from sine noise	6x	8x	17x	20x	22x	24x	1.261
Recostructed signal from pump noise	-	-	-	-	22x	-	2.215

This study will also evaluate a reconstructed signal to the signal target through the log spectral distance (LSD) parameter. This parameter is used to compare the spectral reconstruction signal to the signal target. The spectral reconstruction will get closer to the signal target when the LSD value gets smaller and vice versa. The most significant LSD is found in the reconstructed signal from pump noise with a value of 2.215, while the smallest LSD is found in the reconstructed signal from sine noise with a value of 1.261.

V. CONCLUSION

The inter-microphone time difference method has been used for the reconstruction of the signal target. In this research, spectrogram representation is used. The reconstructed signals from white noise and sine noise have been successfully used to diagnose bearing defects through spectrograms and harmonic frequency patterns. The spectral closeness between the signal target and the reconstructed signal was measured by LSD. The minimum LSD value is found in the reconstructed signal from sine noise with a value of 1.261

REFERENCES

[1]. M. Knaak, M. Kunter and D. Filberi, "Blind source separation for acoustical machine diagnosis," Digital Signal Processing, 2002. DSP 2002 14th International Conference on, 2002, pp. 159-162 vol.1

[2]. Kumar, D.N., & Singh, D.A. "Study of Microphone Array Characteristics and Noise Reduction." International Journal of Applied Engineering Research. 13(12). (2018) pp. 10826-10830

[3]. Aarabi, Parham, and Guangji Shi. "Phase-based dual-microphone robust speech enhancement." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 34.4 (2004): 1763-1773.

[4]. B. T. Atmaja and D. Arifianto, "Machinery fault diagnosis using independent component analysis (ICA) and Instantaneous Frequency (IF)," International Conference on Instrumentation, Communication, Information Technology, and Biomedical Engineering 2009, Bandung, 2009, pp. 1-5.

[5]. Lopez-Tello, Carlo. "Acoustic Detection, Source Separation, and Classification Algorithms for Unmanned Aerial Vehicles in Wildlife Monitoring and Poaching." (2016).

[6]. Gao, L., Zai, F., Su, S., Wang, H., Chen, P., & Liu, L. (2011). Study and application of acoustic emission testing in fault diagnosis of low-speed heavy duty gears. Sensors, 11(1), 599-611.

[7]. Ge, Jianghua, Tianyu Niu, Di Xu, Guibin Yin, and Yaping Wang. 2020. "A Rolling Bearing Fault Diagnosis Method Based on EEMD-WSST Signal Reconstruction and Multi-Scale Entropy" Entropy 22, no. 3: 290.

[8]. Vinaya, A. A., Nurmaulidah, F., Arifianto, D., & Arifianti, Q. A. M. O. (2020). Detection Of Pump Faults Based On Sound Signals Using Non Negative Matrix Factorization. Jurnal Teknologi, 82(2)

[9]. Kim, Chanwoo, et al. "Automatic selection of thresholds for signal separation algorithms based on interaural delay." Eleventh Annual Conference of the International Speech Communication Association. 2010.

[10]. Smith, Leslie S., and Steve Collins. "Determining ITDs using two microphones on a flat panel during onset intervals with a biologically inspired spike-based technique." *IEEE transactions on audio, speech, and language processing* 15.8 (2007): 2278-2286.

[11]. Lee, Sangmoon, Youngjin Park, and Youn-sik Park. "Three-dimensional sound source localization using inter-channel time difference trajectory." International Journal of Advanced Robotic Systems 12.12 (2015): 171.

[12]. M. N. Farid and D. Arifianto, Speech segregation based-on binaural cue: interaural time difference (ITD) and interaural level difference (ILD), Journal of Physics: Conference Series, vol. 776, p. 012067, Nov. 2016.

[13]. P. Kantue and J. O. Pedro, "Integrated Fault Detection and Diagnosis of an Unmanned Aerial Vehicle using Time Difference of Arrival," 2020 24th International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, Romania, 2020, pp. 336-342

[14]. Pan, Q., Zhou, R., Su, J. et al. Automatic localization of the rotor-stator rubbing fault based on acoustic emission method and higher-order statistics. J Mech Sci Technol 33, 513–524 (2019).

.