



Conference Paper

Application of Short Time Fourier Transform and Wavelet Transform for Sound Source Localization Using Single Moving Microphone in Machine Condition Monitoring

Meifal Rusli, Lovely Son, Mulyadi Bur, and Agus Arisman

Department of Mechanical Engineering, Faculty of Engineering, Andalas University, Kampus Limau Manis, Padang 25163- Indonesia

Abstract

The paper discusses means to predict sound source position emitted by fault machine components based on a single microphone moving in a linear track with constant speed. The position of sound source that consists of some frequency spectrum is detected by time-frequency distribution of the sound signal through Short Time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT). As the amplitude of sound pressure increases when the microphone moves closer, the source position and frequency are predicted from the peaks of time-frequency contour map. Firstly, numerical simulation is conducted using two sound sources that generate four different frequencies of sound. The second case is experimental analysis using rotating machine being monitored with unbalanced, misalignment and bearing defect. The result shows that application of both STFT and CWT are able to detect multiple sound sources position with multiple frequency peaks caused by machine fault. The STFT can indicate the frequency very clearly, but not for the peak position. On the other hand, the CWT is able to predict the position of sound at low frequency very clearly. However, it is failed to detect the exact frequency because of overlapping.

Corresponding Author: Meifal Rusli; email: meifal@ft.unand.ac.id

Received: 1 August 2016
Accepted: 18 August 2016
Published: 6 September 2016

Publishing services provided by Knowledge E

© Meifal Rusli et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Selection and Peer-review under the responsibility of the ICoSE Conference Committee.

Keywords: Short time Fourier transform, Continuous wavelet transform, Sound source localization, Condition monitoring

1. Introduction

Sound source localization is a complex work that acoustic engineers face today. Some standards based on microphone arrays are used to analyze the noise source. In general, the methods are categorized into three: near-field acoustic holography, beamforming, and inverse method. Most of these methods were developed using 20 channels or more of microphones and data acquisition [1]. Meanwhile, other researchers have been trying to develop sound source localization method with minimal number of microphone, like binaural and inter-aural sound localization [2].

Likewise, the vibration and acoustic signal analysis are very important methods in condition monitoring and fault diagnostics of machine components. With the rapid development of the signal processing techniques, the analysis of stationary and transitory signals has largely been based on well-known spectral techniques such as: Fourier



Transform and wavelet transform [3,4,5]. Then, to localize sound sources using acoustic signals, some methods also have been developed. For instance, sound source localization using time frequency histogram by two microphones [6], through the fusion between visual reconstruction with a stereoscopic camera pair with several microphones [7], and the application of envelope and wavelet transform to enhance the resolution of the received signals through the combination of different time-frequency contents [8].

This paper discusses about the implementation of a single microphone that is moving in a linear track and constant speed to detect sound source position. The sound position and frequency peaks are detected by the peak of the time-frequency distribution of the sound signal using short time Fourier transform (STFT) and continuous wavelet transform (CWT). Firstly, numerical simulation is conducted by two loudspeakers as the sound sources that generate four different frequencies of sound signals. The second case is the sound generated by a rotating machine with unbalanced, misalignment and bearing defect. Each defect generates specific frequency of vibration and sound. This single moving microphone method has benefit that could be developed for autonomous or robotic condition monitoring system with simpler and cheaper devices and analysis rather than multi-channel microphones array.

1.1. Short Time Fourier Transform and Continuous Wavelet Transform in Moving Microphone

The acoustic signals from faulty components are basically non-stationary. There are two basic approaches to analyze a non-stationary vibration and acoustic signal in time and frequency domain simultaneously. One approach is short time Fourier transform (STFT) by splitting an acoustic signal into segments in time domain by proper selection of a window function and then to carry out a Fourier transform on each of these segments separately and to deliver an instantaneous spectrum. Another approach is the continuous wavelet transform (CWT), where the non-stationary acoustic signal to be analyzed is filtered into different frequency bands, which are split into segments in time domain and their frequency contents and energy are analyzed. Wavelet analysis overcomes the disadvantage of STFT since CWT uses a windowing technique with variable sized regions. Wavelet analysis allows the use of long time intervals where we want more precise low-frequency information, and shorter regions where we want high-frequency information. In this analysis Morlet function is applied for basic wavelet function.

Furthermore, when a fixed source is measured by a moving microphone over a period of time, the distance between source and microphone is no longer constant, but as a function of time. When the moving microphone is approaching closer the sound source, the emitting sound needs less time to reach the microphone, the waveform is compressed resulting in an increase of frequency, especially when the microphone moves at very high speed. This phenomenon is known as the Doppler Effect [9]. Likewise, because the distance decreases, the amplitude increases and becomes maximum when the microphone in the closest position to the sound source. The highest peak in

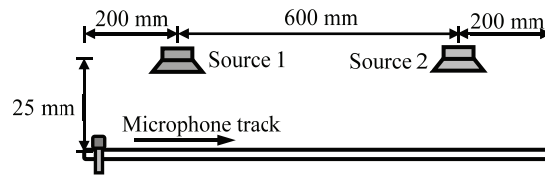


Figure 1: The numerical simulation model.

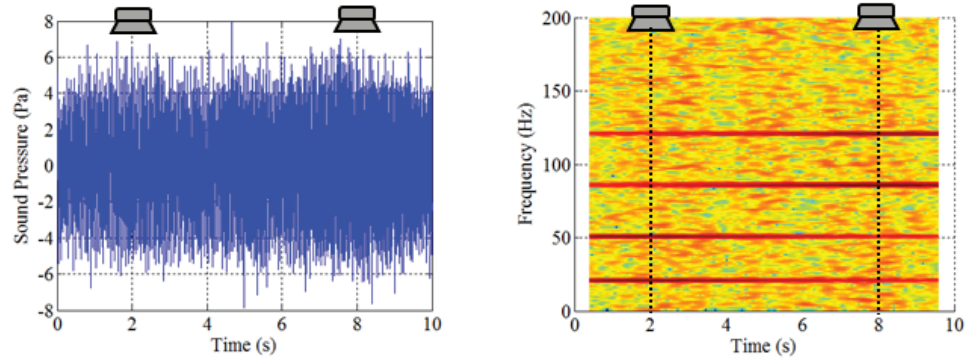


Figure 2: The time waveform (left) and the time-frequency spectrogram using STFT (left) of the sound signal in numerical model.

time-frequency distribution of the sound signal is then detected as the sound source position in each frequency.

2. Numerical Simulation

Firstly, a numerical simulation is conducted using two loudspeakers generating four different frequencies of sound like illustrated in Fig. 1. The first loudspeaker generates sinusoidal sound at frequencies of 25 and 50 Hz, and the second one generates 85 and 125 Hz of signal. The microphone moves on the track one meter length with low speed of 0.1 m/s. Both sound sources generate random noise with the signal to the noise ratio of 5. The sound signal from the microphone in time domain and time-frequency distribution by STFT is depicted in Fig. 2, while time-frequency wavelet energy distribution is shown in Fig. 3.

The result of wavelet transform is not easy to be understood compared with the FFT result. In a broad sense, and it is better to speak about the approximated frequency corresponding to a scale by the relationship of $F_a = F_c / (a \cdot T_d)$. Where F_c is the center frequency of wavelet, a and T_d are the scale and the sampling period, respectively. Fig. 3 shows that by STFT, the sound at frequency of 20 Hz and 50 Hz are detected at the first loudspeaker position. Meanwhile the sound at frequency of 85 and 125 Hz are detected at the second loudspeaker position. However, it is not easy to distinguish the peak from the frequency line.

On the other hand, in time-frequency wavelet energy distribution like shown in Fig. 3, the sound position at low frequency can be detected more clearly at the first sound source. However, two higher frequencies at the second sound source are not detected well. Although zooming in order to increase frequency resolution is done at 100 Hz

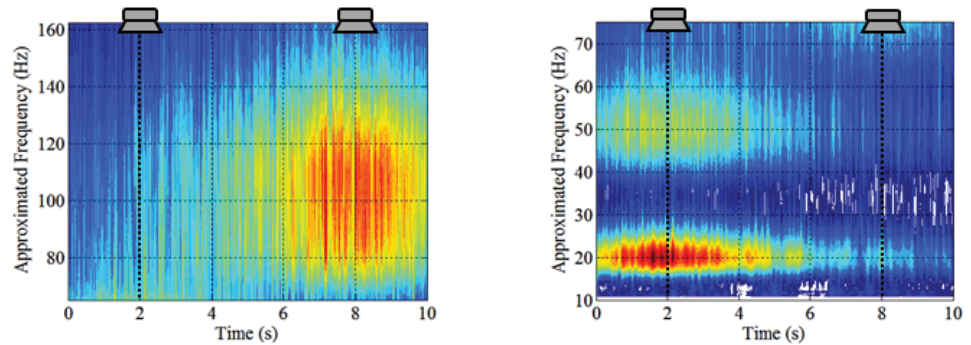


Figure 3: Time-frequency wavelet energy spectrum at different frequency range.

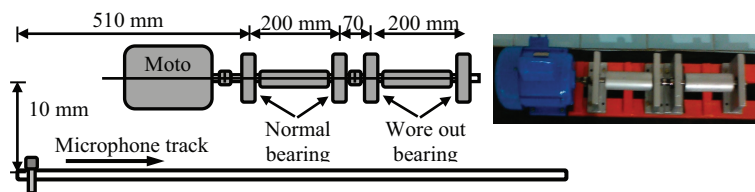


Figure 4: The rotor dynamics model.

to 160 Hz of frequency range, the sound at frequency of 85 Hz and 125 Hz are found merging into a single peak distributed from 85 Hz to 120 Hz. One of the drawbacks of CWT is the overlapping occurred at higher frequency [4]. The spectrum peaks at frequencies of 85 Hz and 125 Hz are too close each other in both position and frequency distance, so the overlapping makes both frequency peaks merge in a single distributed peak. The resolution of the CWT is very good at high scale, which means it is very satisfied at low frequency. Therefore, the CWT will contribute well to detect the sound source at relatively low frequency or longer distance of frequency peaks of spectrum.

3. The Rotor Dynamics Model

The second case is the application of both methods in a simple rotating machine consists of an electric motor, two pairs of ball bearing with two thick rotor and shaft like illustrated in Fig. 4. The right pair has fault wore out bearings, so they have more clearance or looseness between inner ring, outer ring and the ball. The shaft between the electrical motor and both pairs of bearing-shaft-rotor are connected by flexible coupling. Moreover, aligning and balancing process are unavailable in this case, so the unbalance and the misalignment signal may appear in the sound analysis. The microphone moves at very low speed of 0.024 m/s. The timewave of sound signal and time-frequency spectrogram using STFT is shown in Fig. 5, while the time-frequency wavelet energy spectrum are shown in Fig. 6.

The sound signal in time domain presented in Fig. 5 shows that higher pressure level of sound generated by wore out pair of bearing part. Most energy is emitted at the first harmonics rotation frequency about 48 Hz like confirmed by the time-frequency spectrogram using STFT. It is shown that pairs of bearing-shaft-rotor part generate the sound at the first harmonic. However the wore out part generates the highest energy

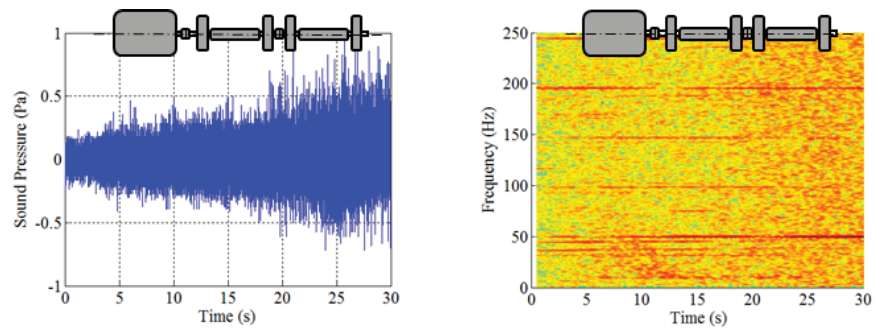


Figure 5: Time wave of sound signal (left), and time-frequency spectrogram by STFT (right).

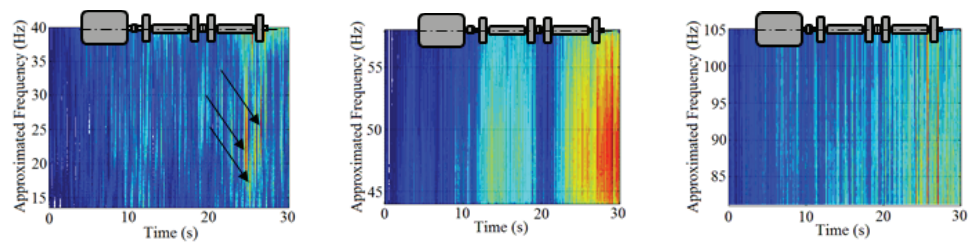


Figure 6: The time-frequency spectrogram using CWT in various frequency ranges.

of sound. It is possibly caused by unbalanced and misalignment occurred in both part. Moreover, the second, the third, fourth, and fifth harmonic are also generated along the system that is possibly caused by the bearings looseness or clearance. In addition, some signal at sub-harmonics frequency also observed to be omitted by the electrical motor.

Fig. 6 shows the time-frequency spectrogram using CWT in various frequency ranges. At sub harmonics frequency range, it is found that some transient sound emitted by last wore out bearing around the FTF of bearing frequency. Furthermore, CWT also shows that the highest energy emitted from wore out part at rotation frequency at 48 Hz, higher than generated by the normal one. In addition, higher energy also emitted by the last bearing at the second and the third frequency harmonics. The same problem is found when using CWT at higher frequency. The application of CWT is difficult to observe the frequency of signal clearly because of the overlapping phenomenon. The overlapping increases significantly by longer time period of sampling.

4. Conclusion

This paper has presented the application of single moving microphone to predict the sound source position for machine condition monitoring. It is observed qualitatively that the time-frequency distribution of sound signal by STFT and CWT are possible to be used to predict the sound source position. The STFT time-frequency distribution shows the frequency clearly, although not so clearly for the position of the peak. On the other hand, the CWT could predict the position of sound at low frequency very clearly, but in contrast, it failed to shows the exact frequency because of overlapping. Combination of STFT and CWT could be applied to overcome the both drawbacks in

machine condition monitoring. Furthermore, the implementation of band pass filter could be proposed to improve the CWT in higher frequency.

5. Acknowledgement

The authors gratefully acknowledge the financial support from Fundamental research grant, Andalas University by contract number 41/H.16/Fundamental/LPPM/2015. We would like to express profound gratitude to colleagues in Structural Dynamics Laboratory and Department of Mechanical Engineering for the guidance, help, and cooperation.

References

- [1] J. Lanslots, F. Deblawe, and K. Janssens, Selecting Sound Source Localization Techniques for Industrial Applications, *Sound Vibrat*, 6–9, (2010).
- [2] C. Faller and J. Merimaa, Source localization in complex listening situations: selection of binaural cues based on interaural coherence, *J Acoust Soc Am*, **116**, 3075–3089, (2004).
- [3] K. Shibata, A. Takahashi, and T. Shirai, Fault diagnosis of rotating machinery through visualization of sound signal, *J Mech Syst Signal Process*, **14**, 229–241, (2000).
- [4] P. W. Tse, W. Yang, and H. Y. Tam, Machine fault diagnosis through an effective exact wavelet analysis, *J Sound Vibrat*, **277**, 1005–1024, (2004).
- [5] H. Bendjama, S. Bouhouche, and M. Boucherit, Application of Wavelet Transform for Fault Diagnosis in Rotating Machinery, *International Journal of Machine Learning and Computing*, **2**, no. 1, 82–87, (2012).
- [6] P. Aarabi and S. Mavandadi, Robust sound localization using conditional time-frequency histograms, *Inf Fusion*, **4**, 111–122, (2003).
- [7] J. Cech, R. Mittal, A. Deleforge, and J. Sanchez-Riera, Active-Speaker Detection and Localization with Microphones and Cameras Embedded into a Robotic Head (2013), IEEE-RAS International Conference on Humanoid Robots, Atlanta, United States, Oct 2013.
- [8] O. M. Bouzid, G. Tian, J. Neasham, and B. Sharif, Envelope and Wavelet Transform for Sound Localization at Low Sampling Rates in Wireless Sensor Networks, *Journal of Sensors*, **2012**, (2012).
- [9] B. Ginn, J. Gomes, and J. Hald, Recent advances in Rail Vehicle Moving Source Beamforming, *Internoise*, 15-18 September 2013, Innsbruck-Austria.