Machinery Fault Diagnosis Using Independent Component Analysis (ICA) and Instantaneous Frequency (IF)

B. T. Atmaja, D. Arifianto
Department of Engineering Physics
Sepuluh Nopember Institute of Technology
Sukolilo Campus, Surabaya 60111 Indonesia

Abstract-Machine condition monitoring plays an important role in industry to ensure the continuity of the process. This work presents a simple and yet, fast approach to detect simultaneous machinery faults using sound mixture emitted by machines. We developed a microphone array as the sensor. By exploiting the independency of each individual signal, we estimated the mixture of the signals and compared time-domain independent component analysis (TDICA), frequency-domain independent component analysis (FDICA) and Multi-stage ICA. In this research, four fault conditions commonly occurred in industry were evaluated, namely normal (as baseline), unbalance, misalignment and bearing fault. The results showed that the best separation process by SNR criterion was time-domain ICA. At the final stage, the separated signal was analyzed using Instantaneous Frequency technique to determine the exact location of the frequency at the specific time better than spectrogram.

Keywords: Machinery Fault Diagnosis, Independent Component Analysis, Natural Gradient, Spectrogram, Instantaneous Frequency.

I. INTRODUCTION

Machinery fault diagnose techniques have been developed for several years. Generally, the technique to monitor machine condition involves direct contact measurement. An operator attaches a sensor on the body of a machine to measure its vibration. Mechanical vibration of a rotating machine emits specific acoustical pattern of sound. This sound can be used as diagnosis tools. When a machine will fail, it gradually emits a different pattern of sound that indicates its worn-out state; the worse the emitted sound, the easier for human ear to identify. However, it is not practical for an operator for continuous monitoring. Besides in real plant, it is common to find a cluster of machines located in one location.

The sound of machine cluster is naturally a mixing from the machines, background noise, reverberation and others. Therefore to analyze a mixture of machine sound requires a separation. Each part of separation component then can be analyzed further to characterize the condition. Blind source separation is a problem that represents separation of sound mixture. The methods to solve that problem generally call as independent component analysis because it separates source component by statistical independent of each component to others.

The machine fault conditions that often occur in industry are unbalance, misalignment, mechanical looseness and bearing fault. Characteristic of normal condition can be used as control or baseline data. Then any alteration of parameters can be used to indicate a fault condition of a machine. This technique may be useful for predictive maintenance.

In this paper, we propose ICA-based techniques for multiple machine sound separation. Time-domain ICA, Frequency-domain ICA and Multi-stage ICA techniques were evaluated to separate 2, 3 and 4 pair of machines and sensors configuration. The separated sound was then analyzed further using the Instantaneous Frequency technique to estimate the fundamental frequency that shows each individual machine characteristic. In the next part of the paper we will briefly present ICA-based techniques, the experiment and followed by discussion of the results.

II. THEORY

A. Machinery Fault Type

Most machines used in industry are rotating machinery. For example: motor, pump, generator, turbine, diesel and others. This implies that this class of machinery has a part that moves periodically over time. We used an electrical motor whose configuration is shown in figure 1.

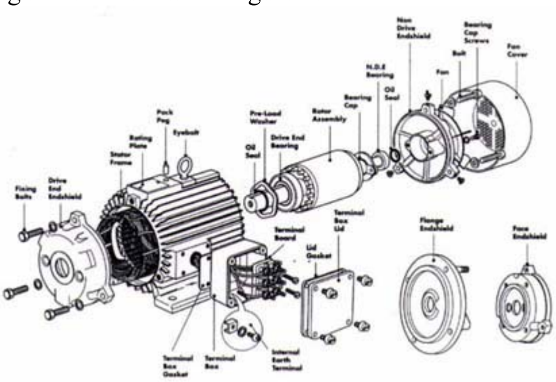


Figure 1. Diagram of Rotating Machine (electrical motor)[1]

An electrical motor or rotating machine usually consists of parts such as motor, stator, bearing, shaft, impeller/blade, etc. Fault condition of machine is typically caused by a failure from one or more components. Types of some fault condition typically occur in machine operation are:

- Unbalance, this is the most common failure because of uneven or unbalance mass at the rotor or rotating part as the centre of rotation.
- Misalignment, this due to asymmetry of shaft when aligning two or more machines shaft to be coupled to be an integrated machine.
- Bearing Fault, it can be caused by failure of ball bearing, inner race, outer race, ball cage or bearing house.

These fault conditions can be characterized by analyzing many parameters. In this research, sound pattern analysis using Independent Component Analysis and Instantaneous Frequency (IF) were used as a diagnosis tool.

B. Independent Component Analysis (ICA)

Independent events occur when an event does not have any relation with another event. This statistical property is stronger then uncorrelated. Mathematically, probability function of those events can be formulated as follows,

$$p(y_1, y_2) = p(y_1).p(y_2) \qquad ...(1)$$

The probability functions of those events are the multiplication of each other. Lets define x_j is measured signal by *i*-th sensor. The s_i are the multiple sources that statistically independent. The measured signal can be expressed using the following equation,

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \ldots + a_{jn}s_n \qquad ...(2)$$

x_j symbolizes mixed signal from sources. The mixing matrix a_i is not known before, as well as the source s_i. The task of ICA is to find out the estimation signal y from unknown mixing process.

Problem in ICA is to determine separation filter, $\boldsymbol{W}$, where $\boldsymbol{W}=\boldsymbol{A}^{-1}$. The more ideal separating filter, the better estimation signal we get. The mixing and problem formulation of ICA can be illustrated in fig.2 below.

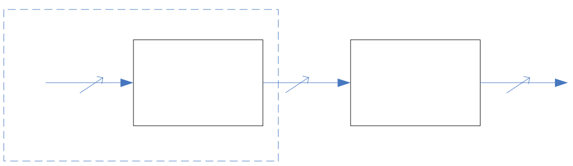


Figure 2. ICA Process

From the picture and as explained before, ICA can be modeled as follows,

$$\mathbf{x} = \mathbf{As} \qquad ...(3a) \qquad \text{and} \qquad y = W\mathbf{x} \qquad ...(3b)$$

Estimation signal from the mixture can be obtained if we have a separating filter $\boldsymbol{W}$. To determine this separating filter they are many choices depend on algorithm and methods. Amari, Chicocky and Yang [6] used natural gradient algorithm to reach ideal separation filter $\boldsymbol{W}$. The weighting factor of separation filter expressed as follows,

$$\begin{aligned} W_{k+1} &= W_k - \mu \frac{\partial \mathbf{J}(W_k)}{\partial W} W_k^T W_k \\ &= W_k + \mu \left\lfloor I - f(\mathbf{u}_k)\mathbf{u}^T \right\rfloor W_k \qquad ...(4) \end{aligned}$$

The algorithm approached ICA by information maximization of mutual information. In eq.(4), μ is a positive learning rate of *f(u)=[f(u1).f(u2)…f(un)]*, where *f(u)* is derivative of sigmoid function and $u=\boldsymbol{W}.x$ is a function before non-linearity.

Hyvarinen *et. all* introduced FastICA approach from fixed-point iteration of weighting factor [9]. The algorithm is based on non-gaussianity and negentropy of the input signal. The algorithm can be formulation as below,

$$\begin{aligned} \mathbf{w}^+ &= \mathbf{w} - \frac{E\{\mathbf{x}g(\mathbf{w}^T\mathbf{x})\} - \beta\mathbf{w}}{E\{g'(\mathbf{w}^T\mathbf{x})\} - \beta} \\ &= E\left\{\mathbf{x}g(\mathbf{w}^T\mathbf{x})\right\} - E\left\{g'(\mathbf{w}^T\mathbf{x})\right\}\mathbf{w} \qquad ...(5) \end{aligned}$$

Where g is the derivation of contrast function as approach non-gaussianity. FastICA algorithm may converge faster than those of other algorithms, but assumption of source is the main factor to choose the best algorithm. There are many algorithm in ICA that can be used depend on characteristic of the signal, the assumption and a prior knowledge of sources.

C. Instantaneous Frequency (IF)

Instantaneous frequency (is) is defined as derivative of phase with respect to time. The value of IF can be approached from two ways, analytic signal and short time Fourier transform (STFT). Windowed or STFT of signal *x(t)* can be formulated as follows,

$$STFT_s(\omega,t) = \int_{-\infty}^{\infty} w(\tau - t)x(\tau)e^{-j\omega t}d\tau \qquad ...(6)$$

Using derivative phase of that signal to time, instantaneous can be obtained by,

$$\frac{\partial}{\partial t}X(\omega,t) = \int_{-\infty}^{\infty} -w'(\tau - t)e^{-j\omega\tau}x(\tau)d\tau \qquad ...(7)$$

where *w'(t)* is a derivation of window function to time. The IF calculation can be seen in fig. 3 (right panel), shown IF from amplitude spectrum. IF technique accommodates the weakness of spectrogram from STFT, which Heisenberg uncertainty occurs between time and frequency resolution. The higher frequency resolution we set, the lower time resolution we get and vice versa. The IF, however, can be used to determine exactly both time and frequency at the same time without suffering time-frequency resolution trade-off.

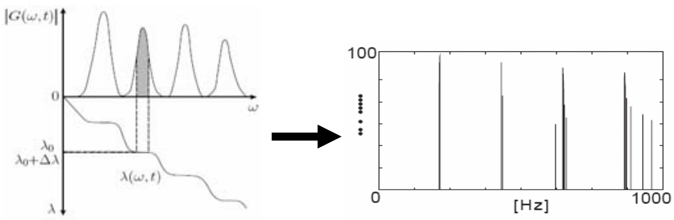


figure 3. IF Amplitude Spectrum[7]

III. EXPERIMENTAL SET-UP

We developed microphone array as the sensor. The sensor itself consists of 4 equally spaced microphones to avoid spatial aliasing. These microphones are connected to 24-bit, four-channel M-Audio PCI soundcard for simultaneous data acquisition. Fig. 4 and fig. 5 show the 3-source-3-sensor configuration used in this research. We also implemented 2-sensor-2-source and 4-sensor-4-source experimental set-ups to investigate the performance of the algorithm when the system is scaled up.

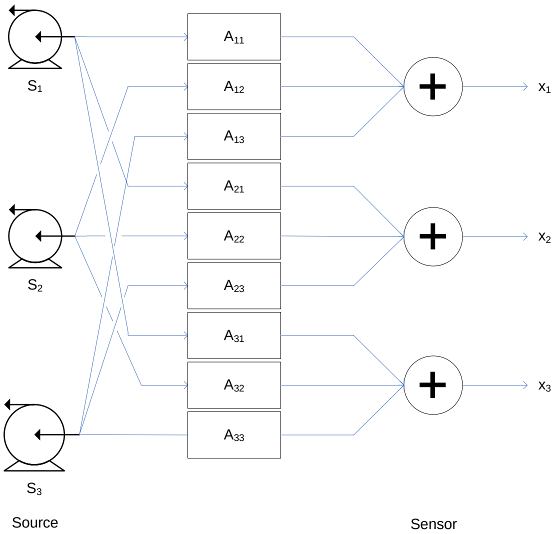


Figure 4. ICA model for 3 source – 3 sensor

The hardware configuration and simulation were shown in fig.5. First, we measured each individual motor whose condition is normal, misaligned, unbalanced and ball bearing failure to obtain baseline data. This step was necessary for result validation. Starting from 2-2-configuration, we set one normal machine and the other was abnormal, then abnormal-abnormal pair until all these four machines ran simultaneously.

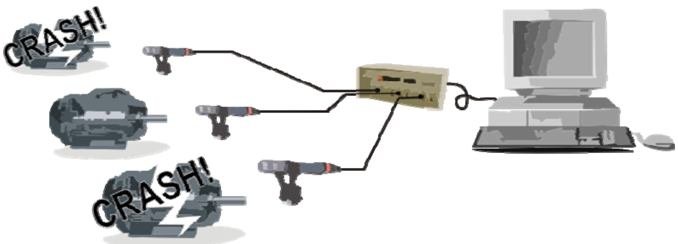


Figure 5 Hardware Configuration of ICA Model

After off-line separation, we obtained estimated separated signals from the mixture that represents number of motors with different conditions. The estimated signal was then estimated its fundamental frequency using instantaneous frequency to determine each machine condition. These steps are shown in fig.6.

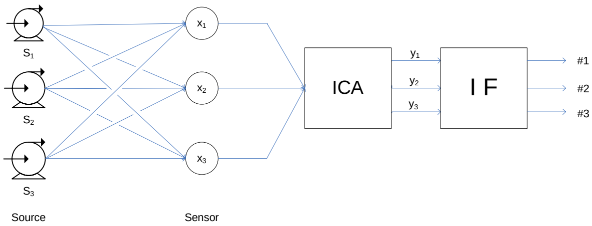
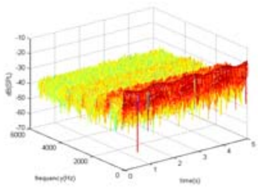


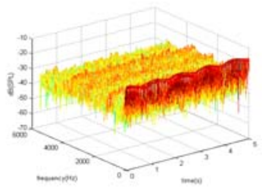
Figure 6. Global System

IV. RESULTS

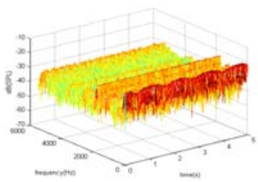
The first data processed in this research are baseline data as a control. The baseline data come from single-channel recording of each machine condition: normal, unbalance, misalignment and bearing fault. The spectrogram of baseline data that shows machine sound characteristic in time, frequency and amplitude are depicted in figure 7 below. The x-axis is time, y-axis is frequency and z-axis is amplitude. The color of spectrogram also represents sound pressure level in each time and frequency. The red color indicates a high frequency and the yellow is low frequency. The spectrogram plot comes from equation of STFT as in (6).



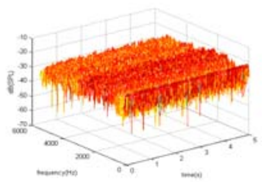
a. Normal



b. Unbalance



c. Misalignment

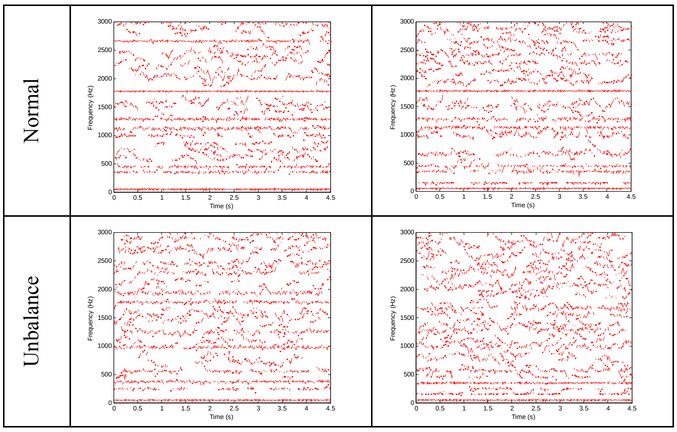


d. Bearing Fault

Figure 7. Spectrogram of Source Signal

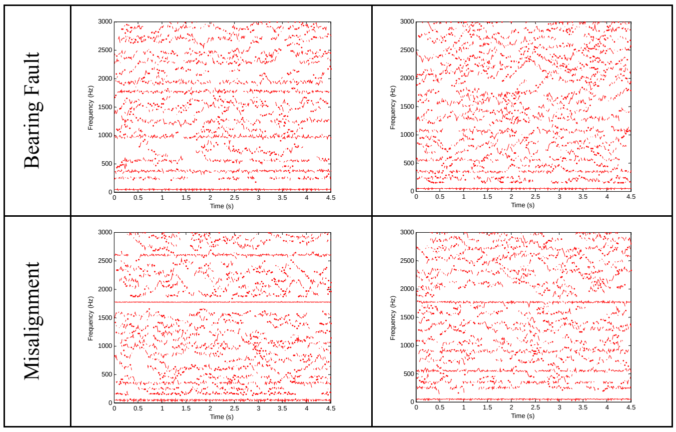
From fig.7(a), it can be clearly seen that on the normal condition the machine emits low frequency. The red color dominated in low frequency region and yellow in the upper frequency region. In this case, the normal machine has frequency at 51 Hz, between 1000-1300 Hz, 1770 Hz. Fig.7 (b) is spectrogram of unbalance machine. It shows that there are increment frequency around centre of the spectrogram (more red color exists). The unbalance machine in this case has fundamental frequency at normal plus 1990 Hz. Also the amplitude increase at frequency 1000 Hz and 1770 Hz.

Fig.7(c) shows misalignment condition. The red color appear at high frequency, as well as at the centre frequncy. In this condition the fundamental frequency were increased at plus 2500, 3400, 5100 and 5300 Hz. Fig.7(d) depicts bearing fault. The red color dominated at the all frequency that indicated the amplitude increase near to all frequency. The highest amplitude at this case also from instantaneous frequnecy information give fundamental frequency at 73 Hz, 250 Hz dan 350 Hz.



a. Source Signal b. Estimate Signal

Figure 8. Source and Estimated Signal of Normal and Unbalance



a. Source Signal b. Estimate Signal

Figure 9. Source and Estimated Signal of Bearing Fault and Misalignment

The second data is mixing machine sound to be separated using ICA methods. Each microphone will receive different signal intensity depending on distance and angle from motor to microphone. This difference use to separate signal from each source by statistically independent between each source. As an example, three motor condition: normal, unbalance and misalignment working together and the machine sounds captured by three microphone. Fisrt microphone closed to normal motor will receive normal motor signal stronger than other, but it mixed with unbalance and misalignmnet signal. It also happens on second and third microphone. The task of ICA is got the origin signal: normal, unbalance and misalignment from mixed signal at each microphone. The result of estimated signal analyzed by Instantaneous frequency, from in (7), to find simmilarity between estimated signal and source signal. Estimated signal of normal condition and unbalance condition by instantaneous diagram can be seen in figure 8.

It can be seen from figure above that estimated signal from both normal and unbalance condition very similar to the original source signal. On normal estimation signal, they are increasing fundametal frequency around 3500 Hz, the mixing signal known not yet separate clearly. The adding frequency maybe come from misalignment signal which have that frequency. On second figure, the estimated of unabalance signal also very simmilar and the amplifying occur at 1770 and 1990 Hz, it can caused by signal that have some same frequency with unbalance condition i.e normal condition. The next figure are from others condition, misaligmnent and bearing fault condition. The instantaneous frequency pattern of both signal from source and estimated can be seen from figure 9. The bearing fault condition was obtained by breaking the ball cage of bearing. This treatment caused bearing home loose and ball bearing can move generously than before. The estimated signal got not as well as got before. The sound pattern of estimated signal blurred so it can identify signal clearly. But more detailed, the lowest frequency still can be viewed. Because all of condition have fundamental frequency at low frequency so it too difficult to diagnosis that condition in real world application later.

The last figure come from misalignment condition. The estimated signal shows better than before, it seems that estimated signal simmilarity again with source signal. But some fundamental frequency attenuated due to separation process. The fundamental frequency at 2500 Hz adn 3400 Hz lost in estimated signal. At high frequency, 5100 and 5300 Hz still can be viewed although not as clear as source signal. Generally, the estimated signal still can be identify as misalignment condition.

The separation prosess in this research have done by ICA method in time domain. The algortihm used is natural gradient from Amari, Chihochky and Yang [2]. The result of separation process can says good for early research. Attenuation and amplifying occured in estimated signal so the signal can be separated perfectly. In the signal processing field, separation methods to solve blind source separation problem still establishing untill now. ICA is one popular to solve BSS and they are many algorithm in ICA. The separation process of machine sound can followed the ICA development. Certain algorithm maybe usefull for machine sounds separation, and instantaneos frequency proven efective to identify signal than time domain signal. The IF technique also affective as diagnosis tools of fault condition that accomadate Heissenberg's uncertainty shows in spectrogram.

IV. Conclusions

According to the result of this research on machinery fault diagnosis using ICA and IF, it can be concluded as:

1. Machinery sounds can be separated by independent component analysis (ICA) using assumption that each machine statistically independent with the others.
2. Machinery fault diagnosis can be obtained by sound pattern analysis using instantaneous frequency to determine fundamental frequency each condition. In this research, normal motor condition have frequency at 51 Hz, between 1000-1300 Hz, 1770 Hz and 2650 Hz, unbalance at 46 Hz, 1000 Hz, 1770 Hz and 1990 Hz, misalignment at normal frequency plus 3400 Hz, 5100 and 5300 Hz. The last, bearing fault at 73 Hz, 250 Hz dan 350 Hz.

3. The estimated signals as a result of separation process have similarity with source signal, but it was attenuated and amplified in term of frequency on the estimated signal. In some cases, it causes difficulties on identifying the signal estimates.

The on-going research is we are implementing and evaluating the other algorithm such as time-frequency analysis based on instantaneous frequency without ICA for direct separation. We investigate of using less sensors compared to the sources which is more complex due to the size of the mixture matrix. The goal of this work is a remote machine monitoring system for a real plant application.

ACKNOWLEDGMENT

In this research we would like to acknowledge the Ministry of Education, Republic of Indonesia for funding this research on a PKM-P 2009 research grant. The authors would like to acknowledge the head of acoustic laboratory of Engineering Physics – ITS for admission to conduct this research.

REFERENCES

[1] http://www.scardana.ca/pages/Motors.html

[2] Amari, S. 1997. “Natural Gradient Works Efficiently in Learning”. Saitama: RIKEN, Japan.

[3] Furukawa, M., Hioka, Y., Ema, T., Hamada, N., “Introducing New Mechanism in the Learning Process of FDICA-based Speech Separation”, Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC2003), pp.291-294, Sept. 2003, Kyoto, Japan.

[4] Nishikawa, T., Saruwatari, H., Shikano, K., Araki, S., Makino, S., “Multistage ICA for Blind Source Separation of Real Convolutive Mixture”, Proc. Int. Symp. On Independent Component Analysis and Blind Signal Separation (ICA2003), pp. 523-528, April 2003, Nara, Japan.

[5] Nishikawa, T., Saruwatari, H., Shikano, K., “Stable Learning Algorithm for Low-Distortion Blind Separation of Real Speech Mixture Combining Multistage ICA and Linear Prediction”, Proc. NOLISP 2008.

[6] Amari, S., Cichocki, A., Yang, H., A New Learning Algorithm for Blind Signal Separation”, Advances in Neural Information Processing System 8, MIT Press, 1996.

[7] Arifianto, D. “Fundamental Frequency Estimation and Voiced/Unvoiced Determination Using Instantaneous Frequency”, PhD Thesis, Tokyo Institute of Technology, 2005.

[8] Bell, A J and Sejnowksi T J. 1996. “The ‘Independent Components’ of Natural Scene are Edge Filters”. CNL: The Salk Institute (submitted to vission research).

[9] Hyvärinen, A., Karhunen, J., Erkki, O “Independent Component Analysis” John Wiley and Sons, 2001.

[10] Amari, S., Cichocki A., Adaptive Blind Signal and Image Processing, John Wiley and Sons, 2002.

[11] Rahmadana, Y. “Sound Pattern determination of Rotating Machinery Faults using Instantaneous Frequency”, Proc. of ICTS ITS, 2007.

[12] Widyotriatmo, A and Prasetiyo, I. “Speech Signal Separation from Noise using Blind Source Separation”. Proceeding of VibrAstic Surabaya: Teknik Fisika-ITS, 2005.