Identification of Diesel Sound Source Based on The Independent Component Analysis

Han Jun¹ and Hao Zhiyong²

¹. School of Mechanical Engineer, Tianjin University, Tianjin 300072, China, E-mail: junhan01.student@sina.com
Tel :+86-02227890920; Fax: +86-02227404244
². College of Mechanical and Energy Engineering, Zhejiang University, Hangzhou 310027, China,
E-mail: haozy@zju.edu.cn

Abstract: As a new approach of blind source separation (BSS), independent component analysis (ICA) has attracted extensive attention of researchers in the field of information processing. In this paper, the basic theory and algorithm of ICA are briefly introduced, and then ICA is used for the preprocessing of engine acoustic signals to identify the engine noise sources. The ICA decomposes the signals into a number of independent components (ICs) so the individual engine acoustic sources can be studied separately. The theoretical description of the characteristics of the diesel engine noise sources is introduced. A numerical example is presented to verify the separation efficiency of the ICA. The numerical example shows that the ICA can effectively separate the embedded low-level sources. Although acoustic signals from diesel engines contain useful information, it is still difficult to resolve them accurately based on noise measurement. This is because the occurrences of each noise source are too close together. The low-level noise sources cannot be identified successfully. The ICA model is used to extract the above engine noise sources of two types of engines. The continuous wavelet transform (CWT) is applied to represent the ICs in the time-frequency domain. The source separation results from the recorded acoustic signals are accepted.

Keywords: Acoustic Signals  Independent Component Analysis  Diesel Engine  Sound Source Identification

1. Introduction

The acoustic signals obtained from a diesel engine contain useful information to reflect the engine operating conditions. Such information often exists in the form of short duration transients. These non-stationary transients are further contaminated by a variety of noise sources such as firing processes, overlap of events and background noise. Hence, it is difficult to use conventional methods to extract the information from the noisy acoustic signals for condition monitoring.

A variety of signal processing methods including statistical analysis, spectral analysis, time-frequency analysis and wavelet transform have been used to analyse engine noise[1,2]. These methods are applied to investigate the noise-generation mechanisms and to reveal the individual features of the sources. Each method is based upon component energy contributions to retrieve information about engine noise. Firstly, the noise signals are represented by using either the time domain, frequency domain or the joint time-frequency domain. The noise sources are then identified by the energy variations of the represented signals. As such methods are all based on energy conservation, they are useful for finding predominant information such as combustion peaks.

The other low-level noise sources cannot be identified successfully, using the methods mentioned above. This is because these methods retain the signal energy information from one domain to another. The low-energy noise sources are either buried by the combustion events or too small to be recognized. Hence, these signal energy conservation based methods are unable to recognise.

The independent component analysis (ICA) brings a different strategy in dealing with the problems of blind sources separation (BSS). In the ICA it is assumed that the measured data is a linear combination of the indirectly observed latent sources. As long as the latent signals are statistically independent the ICA should be able to decompose them into independent components (ICs) successfully. It has been reported recently that the ICA is an effective approach for analyzing brain electroencephalogram (EEG) data, image processing, feature extraction telecommunications and financial applications[3,4]. This paper focuses mainly on the study of the acoustic signals generated from a test diesel engine using the ICA in an effort to identify its noise sources. The source separation results from the recorded acoustic signals are accepted.

2. Theoretical Modeling

2.1 Background of the ICA

The idea of the ICA derives from the well-known cocktail-party problem. Imagine that you are in a room where two people are speaking simultaneously. You have two microphones, which you hold in different locations. The microphones give you two recorded time
signals, which we could denote by \(x_1(t)\) and \(x_2(t)\), with \(x_1\) and \(x_2\) the amplitudes, and \(t\) the time index. Each of these recorded signals is a weighted sum of the speech signals emitted by the two speakers, which we denote by \(s_1(t)\) and \(s_2(t)\). We could express this as a linear equation:

\[
x_1(t) = a_{11}s_1(t) + a_{12}s_2(t) \quad (1)
\]

\[
x_2(t) = a_{21}s_1(t) + a_{22}s_2(t) \quad (2)
\]

where \(a_{11}, a_{12}, a_{21}, a_{22}\) are some parameters that depend on the distances of the microphones from the speakers. It would be very useful if you could now estimate the two original speech signals \(s_1(t)\) and \(s_2(t)\), using only the recorded signals \(x_1(t)\) and \(x_2(t)\). This called the cocktail-party problem. For the time being, we omit any time delays or other extra factors from our simplified mixing model. This classic cocktail-party problem implies the idea of blind source separation (BSS): recovering the original latent source signals from the measured mixtures. The term blind means both the mixing process and the latent sources are unknown, that is true in the measurement of the diesel engine acoustic signals in which both the noise sources and their mixing process could not be determined mathematically. The only a priori knowledge is the assumption that the sources are statistically independent.

### 2.2 ICA model

We suppose that the observed \(k\) time series \(x_1(t), x_2(t), \ldots, x_k(t)\) are an instantaneous linear mixture of unknown mutually independent components \(y_1(t), y_2(t), \ldots, y_k(t)\). The observed series can therefore be modeled in a matrix form:

\[
x(t) = Ay(t), \quad 1 \leq t \leq N \quad (3)
\]

where \(y(t) = [y_1(t), y_2(t), \ldots, y_k(t)]^T\), \(x(t) = [x_1(t), x_2(t), \ldots, x_k(t)]^T\), and \(A\) is an unknown \(k \times k\) mixing matrix.

### 2.3 Extraction of independent components

Many existing ICA techniques, e.g. INFORMAX[4], MMI[5], etc., can recover those independent components \(y_i(t)\) in Eq. (3) up to an Unknown constant and a permutation of indices through a de-mixing matrix \(W\) with

\[
\hat{y}(t) = Wx(t) = WAy(t), \quad 1 \leq t \leq N \quad (4)
\]

where \(\hat{y}(t) = [\hat{y}_1(t), \hat{y}_2(t), \ldots, \hat{y}_k(t)]^T\) is an estimate of \(y(t)\). This paper choose Informax algorithm based upon the fact that it can separate any combination of super Gaussian and sub-Gaussian signals. The details of Informax can be referred to in the relevant papers.

### 2.4 Representation of the ICA signals

Theoretically, each decomposed IC should contain at least one component corresponding to one of the physical sources. For diesel engines, this component may be one of the impulsive sources. These types of noise sources produce sound signals of short duration and wide frequency bands. Such signals then need time-frequency transform to resolve them properly. It has been shown that the continuous wavelet transform (CWT) has better time-frequency localization, higher resolution and is faster in computation. To investigate the features of engine noise sources, the CWT is used in this study to post-process the ICs. The IC is represented as

\[
Y_i(a, b) = \frac{1}{a} \int y_i(t) \psi\left(\frac{t - b}{a}\right) dt, \quad i = 1, \ldots, m \quad (5)
\]

where \(\psi(\bullet)\) is called mother wavelet with dilation and translation parameters \(a\) and \(b\). The Morlet wavelet is used here due to its high localization in both the time and the frequency domains.

### 3. Results and Discussion

#### 3.1 Experiment procedure

Measurement configuration consisted of a microphone in conjunction with the charge amplifier, the FFT analyzer and PC for additional computation. The sampling frequency was 40 kHz.

The acoustic signals were obtained from a single-stroke engine and a four-stroke diesel engine in a semi-anechoic laboratory with acoustic wedges on all surfaces except on the floor respectively. The engine is installed in the center of the room. A microphone was positioned 1 m away from the engine front. Since there was no sound reflection to the engine sound radiation surfaces in the laboratory, the sound signals from other side of the engine were correspondingly small at the measurement position of the engine front. Therefore, the measured signals should mainly mirror the sound energy. This has been confirmed in the sound intensity experiments[6]. The sound signals were adopted into a computer through the microphone and then fed into the ICA model to estimate the independent noise sources. Afterwards, the results from the ICA were then post-processed using the CWT and presented in two-dimensional contour plots to show the more accurate time-frequency locations of the events. The real part of the wavelet coefficients is displayed, where the abscissa is time and the ordinate is frequency. Light gray and white regions of the graph present the positive wavelet coefficients, dark gray and black regions the negative values. During the measurements, the revolution speed...
of the diesel engine was 3600 rpm and the load was 577 Nm in four-stroke diesel and 1800 rpm in single-stroke in diesel.

3.2 Results and Discussion

The acoustic wave signals from the single-stroke engine were recorded in Fig. 1a. The result of continuous wavelet transform to the acoustic signal is presented in Fig. 1b. The contour plot of the CWT presents the sound energy distribution of the engine front in time-frequency domain. The time-frequency distribution describes simultaneously when a signal component occurs and how its frequency spectrum develops with time. As the positions of ridges represent the distribution of major sound energy, it is evident from the map of Fig. 1b that the major sound energy is around frequency 100 Hz~500 Hz, periodically and respectively. Also, some scattered spots appeared between 500 Hz and 3000 Hz in it. These features suggest the presence of transient characters, which may be due to non-linearity in the acoustic signals. It confirms the findings in previous theoretical and experimental studies that engine acoustic signals are periodic and unsteady.

![Figure1 Spectrum analysis of acoustic signals](image1)

(a) Sound pressure signals of single-stroke

(b) Continuous wavelet transform of signals

Figure1 Spectrum analysis of acoustic signals

In Fig. 1, the other low-level noise sources cannot be identified successfully, using the CWT. This is because the traditional signal processing methods including statistical analysis, spectral analysis, time-frequency analysis and wavelet transform retain the signal energy information from one domain to another. The low-energy noise sources are either buried by the combustion events or too small to be recognized. Hence, these signal energy conservation based methods are unable to recognise such noises induced by fuel injections or valve movements, which contain relatively small energy.

Fig. 2 illustrates the one result of ICA. It illustrates the combustion-related noise source in the time-frequency plane. Fig. 3 is one peak of the result. The result is also presented in the form of 2-D and 3-D graphs. Judging from the time location of the main peak that occurs around the TDC in Fig. 2 and Fig. 3, it can be deduced that this event is caused by the combustion of the single-stroke engine. Its frequency response as shown in the contour plot is around 300 Hz, periodically and respectively, which is within the engine structure resonance frequency band.

![Figure2 The ICA of level-frequency domain](image2)

Figure2 The ICA of level-frequency domain

![Figure3 The ICA of the combustion noise](image3)

Figure3 The ICA of the combustion noise

There are also a number of other small components as shown in the Fig. 2. This is due to the fact that the
ICA focuses only on the separation of the major components by reaching the maximum independence. As has been demonstrated in Li[7], when a signal is far too small than the others, it might not be separated but mixed together with others. In fact, even in most medical applications, in which the sources are rather simple, the ICA is unable to separate them completely.

Figure 4 The ICA of high frequency domain

The results of ICA in high frequency domain are presented in Fig.4. These events happen soon after the combustion after the TDC. The frequency positions are around 1000 Hz, 2000 Hz and 3500 Hz, periodically and respectively, which are close to the combustion event. Under the conventional time-frequency analysis, these are difficult to identify these events to these close to time-frequency position to that of the combustion event. Here the ICA separates these events although it possesses rather small energy level.

The acoustic signals were obtained from a four-stroke diesel engine in a semi-anechoic laboratory with acoustic wedges on all surfaces except on the floor.

The acoustic wave signals from the diesel engine were recorded in Fig.5(a). The result of continuous wavelet transform to the acoustic signal is presented in Fig.5(b). The contour plot of the CWT presents the sound energy distribution of the engine front in time-frequency domain. As the positions of ridges represent the distribution of major sound energy, it is evident from the map of Fig.4b that the major sound energy is around frequency 150 Hz, 250 Hz and 400 Hz~500 Hz, periodically and respectively. It can be found in Fig.5 that the sound energy periodicity (0.033 s) is related to work cycle and rotation speed (3600 rpm) of the engine.

On the other hand, because the rotation speed is 3600 rpm, the work cycle excitation frequency is 120 Hz. So the major energy lays in 120 Hz. And the result can be found in Fig.5(b).

Fig.6 illustrates the one ICA result of acoustic signals. Its frequency response as shown in the contour plot is around 120 Hz, periodically. The result extracts the major energy form the acoustic signals.

(b) Continuous wavelet transform of signals

Figure 5 Spectrum analysis of acoustic signals

The event of two times of excitation frequency is represented in Fig.7. This event happens soon after the
combustion after the TDC. Under the convention time-frequency analysis, it is difficult to identify respectively. And the event occurs periodically.

The event related high frequency of signals is shown in Fig. 8. The frequency range of this event is located in 1000 Hz~1500 Hz. The result is difficult to extract in Fig. 5 using traditional time-frequency method. The ICA separates these events although it possesses rather small energy level.

However, it has been observed that there are events that could not be attributed to any known sources. This suggests that there are other engine noise sources yet to be identified. It is also noted from the above discussion that most separated noise sources have frequency response range below the 3 kHz.

4. Conclusions
The ICA was applied to the acoustic signal analysis in order to identify the sound sources of engines. The ICA is different from the conventionally used acoustic signal processing methods in that is separates individual sources based on their statistical independence.

The ICA is very helpful in identifying embedded low-level events such as transients which are very common in diesel engines. As a practical application to diesel sound sources, the results of tests, together with the CWT, were analyzed. The main objectives of two engines were to identify the sound sources of diesel engines. Nonetheless, it is to note that the ICA is also limited due to its restriction on the sources.

References