Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Study of engine noise based on independent component analysis

HAO Zhi-yong[†], JIN Yan[†], YANG Chen

(*School of Mechanical and Energy Engineering, Zhejiang University, Hangzhou 310027, China*)

[†]E-mail: haozy@zju.edu.cn; kingwend@163.com

Received Aug. 19, 2006; revision accepted Jan. 15, 2007

Abstract: Independent component analysis was applied to analyze the acoustic signals from diesel engine. First the basic principle of independent component analysis (ICA) was reviewed. Diesel engine acoustic signal was decomposed into several independent components (ICs); Fourier transform and continuous wavelet transform (CWT) were applied to analyze the independent components. Different noise sources of the diesel engine were separated, based on the characteristics of different component in time-frequency domain.

Key words: Acoustic signals, Independent component analysis (ICA), Wavelet transform, Noise source identification
doi:10.1631/jzus.2007.A0772 **Document code:** A **CLC number:** TB53

INTRODUCTION

Noise emission from diesel engine is a complicated acoustic signal with many different components mainly caused by combustion and mechanism operations. The rapid rise of pressure in the cylinder caused by combustion of fuel near the top dead center (TDC) transmits to the engine structure surface and forms an important part of the total noise emission. The combustion can also cause the vibration of cylinder head, connection rods and crankshaft, with the vibration being also an important source of engine noise. All The air-borne noise emission induced by combustion in cylinders is usually called combustion-induced noise. Otherwise, the movements of engine mechanical systems including the rotation of the crankshaft, the operation of valves, the injection of fuel, and the piston slap, are also main factors contributing to the noise radiation of the engine.

It is very important to separate the different sources of noise from the engine for the purpose of diagnosis and main noise source identification. Continuous wavelet transform (CWT) was applied to analyze the acoustic signals from diesel engine for the purpose of noise source identification (Hao and Han, 2004). The characteristic of noise signals in time-frequency domain can supply more information than FFT, although it is difficult to separate different noise sources distinctively due to the high speed operation of the engine and the superposition of different noise source in frequency domain.

Independent component analysis (ICA) separates the statistically independent sources by a finite set of observations recorded by sensors (Comon, 1994). Due to its blind source separation ability, ICA was applied to separate the artifacts in MEG data, reduce the noise in image, and etc. (Hyvärinen and Oja, 2000). Noise from diesel engine was separated into different components (Li *et al.*, 2001) by a sequential ICA model.

In this paper ICA was applied to analysis the acoustic signals of a one cylinder diesel engine. The diesel engine's noise signals were separated into several different independent components (ICs). Then FFT and CWT was used to process different ICs. From the characteristic of the ICs in both time and frequency domain, the ICs were identified to be different noise sources of diesel engine.

[*] Project (No. 50575203) supported by the National Natural Science Foundation of China

BASIC PRINCIPLES OF ICA

Mathematical model

The aim of ICA is to separate the source vectors from the observed vectors, when both the source vectors and the process of source vector being mixed are unknown. The linear ICA model without noise is formulated as follow:

$$\boldsymbol{x}=\boldsymbol{A}\boldsymbol{s}, \tag{1}$$

where $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, …, $\boldsymbol{s}_m$ are row vectors of unknown independent sources. $\boldsymbol{A}$ is an unknown $m\times n$ matrix, and $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, …, $\boldsymbol{x}_m$ are observation vectors got by sensors. The task of ICA is to find an estimation inverse matrix $\boldsymbol{W}$ of mixing matrix $\boldsymbol{A}$,

$$\boldsymbol{W}\approx\boldsymbol{A}^{-1}, \tag{2}$$

to make

$$\boldsymbol{y}=\boldsymbol{W}\boldsymbol{x}\approx\boldsymbol{A}^{-1}\boldsymbol{x}, \tag{3}$$

with $\boldsymbol{y}$ being the estimation of source vector. The starting point for ICA is a very simply assumption that the source components are statistically independent. It means that the probability density function (PDF) of independent sources must behave non-Gaussian distribution.

Numerical simulation

There are many different algorithms in use including negentropy, information maximization, maximum likehood and high-order cumulant in ICA. The FastICA algorithm is fixed point algorithm invented by Hyvärinen in 1997, for estimating signals Gaussianity with its four-order cumulant. Then fixed point algorithm was improved in 1999 by estimating the signals Gaussianity with negentropy instead of four-order cumulant (Hyvärinen, 1999). The algorithms are used to measure the Gaussianity of signals by its negentropy $J(\boldsymbol{y})$:

$$J(\boldsymbol{y})=H(\boldsymbol{y}_{\text{gauss}})-H(\boldsymbol{y}), \tag{4}$$

where, H is the entropy of $\boldsymbol{y}$, $\boldsymbol{y}_{\text{gauss}}$ is a Gaussian random variable with the same covariance vector as $\boldsymbol{y}$. A Gaussian variable has the largest entropy among all random variables of equal variance. A big Negentropy value of $\boldsymbol{y}$ means that PDF of $\boldsymbol{y}$ is far from Gaussian distribution. Negentropy can be estimated by higher-order moments and also usually by

$$J(\boldsymbol{y})\approx\sum_{i=1}^{p}[E\{G_i(\boldsymbol{y})\}-E\{G_i(\boldsymbol{v})\}]^2, \tag{5}$$

$\boldsymbol{v}$ is a Gaussian variable of zero mean and unit variance, and G is any non-quadratic function. The following functions had proved very useful:

$$\frac{1}{a_1}\log\cosh a_1\boldsymbol{y} \quad (1\le a_1\le 2),$$
$$-\exp^{-y^2/2},$$
$$\boldsymbol{y}^4,$$

here G is selected as $\boldsymbol{y}^4$, and based on fixed-point iterations method, the maximum $J(\boldsymbol{y})$ was found. Because of its high efficiency in calculation and convergence, FastICA is now widely used.

To validate the ability of ICA in blind source separation, a numerical simulation was given as follows. A set of artificial signals were chosen as source vectors, which include two sinuous signals of different frequency, and a sinuous signal multiplied by an exponentially decaying function. The last signal is a rectangular signal. The four signals' amplitude is nearly the same. Fig.1 shows the source vectors and artificial signals mixed by a random matrix $\boldsymbol{A}$ whose element is distributed randomly between 0 and 1.

Then FastICA was applied to separate the mixed vectors. Fig.2 shows the results of ICA. A property of ICA is that the separated signals have different variances from original ones which mean that the energy information of sources is distorted. Another property is the permutation of the source order.

From Fig.2 we can conclude that: ICA can recover well the source vectors from the mixed observations vectors, but with distorted source energy and a permutation.

ENGINE NOISE SOURCE SEPARATION

Noise measurement

The engine noise test was carried out in a semi-anechoic laboratory with acoustic wedges on all the surfaces except the ground. The background

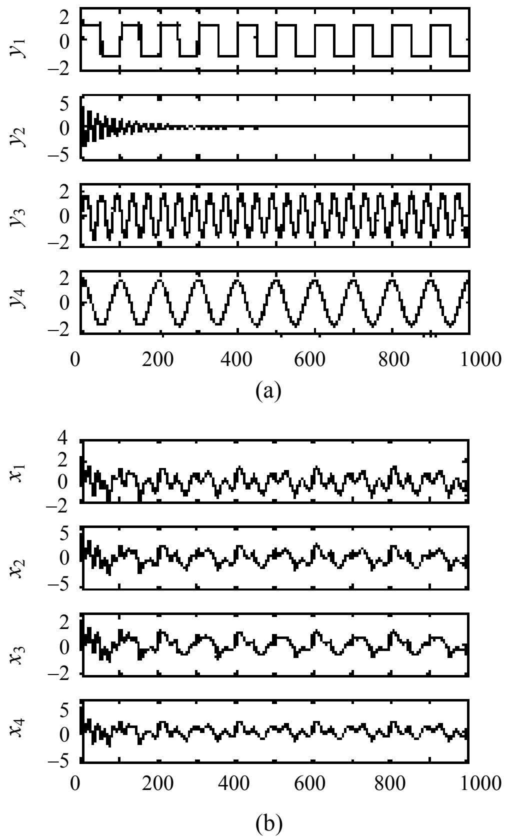

Fig.1 The source vectors and the mixed vectors. (a) Source vectors; (b) Mixed vectors

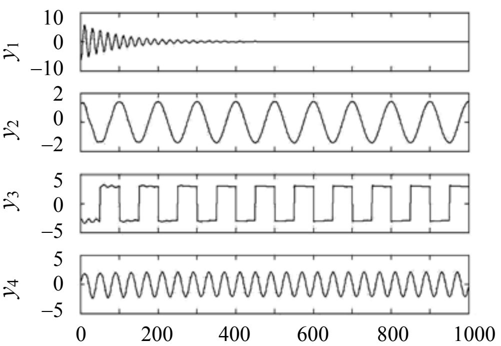

Fig.2 Separated source

sound pressure level is below 25 dBA. The tested engine is a single-cylinder four-stroke diesel engine, mounted in the test room center. The microphone and amplifier are made in Denmark by B&K Company. Four microphones were used to sample the signals, and three microphones were mounted 1 m away around except the rear end of the engine. Another microphone is put 10 cm over the cylinder head. The sampling frequency is 25 kHz. During the measurement, the revolution speed of the diesel engine was 2200 r/min, and the power was nearly 22 kW.

Engine acoustic signal statistical property

As mentioned, ICA requires that the source signals are non-Gaussian distributed. So it is necessary to study the distributions of acoustic signals from the tested diesel engine. In the linear ICA model, the mixing process was proposed as a linear transformation, the distribution of the observation signals has the same distribution as the source ones. So it is enough to measure the Gaussianity of observation signals to decide the distribution of source signals. A convenient way to estimate the Gaussianity of signals is to calculate the normalized kurtosis values. The normal kurtosis is defined as follow:

$$kurtosis(\boldsymbol{y}) = E(\boldsymbol{y} - \mu) / \sigma^4 - 3, \tag{6}$$

where $E(\cdot)$ represents the expected value of vectors. And y is the measured vector, μ is the mean of $\boldsymbol{y}$, σ^2 is the deviation of $\boldsymbol{y}$. The Gaussian distributed signal has a kurtosis value near 0, and the signal with kurtosis value greater than 0 is labelled as supper-Gaussian and if less than 0 the signal is called sub-Gaussian signal. The kurtosis values of 50 acoustic signals from tested engine were calculated, Fig.3 shows the statistically distribution of the acoustic signals' kurtosis values.

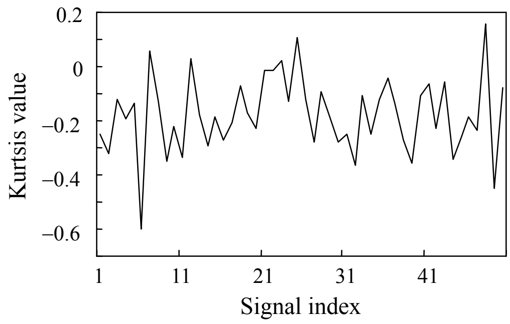

Fig.3 The kurtosis value of acoustic signals

Among the 50 kurtosis values, only 4 values were greater than 0, the other 46 values were less than 0. The result indicates that most of the acoustic signals from the tested engine are sub-Gaussian distributed. It means the tested engine acoustic data satisfy the basic requirement of ICA.

Engine noise separation

FastICA was applied to process the signals sampled several times by the 3 microphones around the engine with 3 independent components (ICs) separated from the raw acoustic emissions though the sequences of the 3 ICs were not always the same. Fig.4 shows the ICs in time domain. IC3 shows an obvious different characteristic in both time domain and frequency domain from the other two. The peaks of 36 Hz, 72 Hz obviously dominate in the whole frequency domain (Fig.5). As the revolution speed of the tested engine is 2200 r/min, the basic order frequency of inertia force of rotation components is nearly 36 Hz; the second order is 72 Hz. These frequencies strictly coincide with the frequencies of IC3. So we can conclude that IC3 is mechanical noise caused by the inertia force generated by the rotation of the crank shaft system and other rotation components.

The spectra of IC1 and IC2 overlap badly, the main energy of the two ICs is concentrated around 1000 Hz. In practice, a diesel engine has many resonance frequencies. Most of the resonance frequencies of engine components are between 800 Hz and 3000 Hz. Usually, the frequencies of excitation of both combustion and mechanical excitation are also mainly within 3000 Hz. Hence, both the combustion induced noise and mechanical noise have important components within the resonant band of the diesel engine. This makes it difficult to distinguish IC1 and IC2 in the frequency domain.

Wavelet transform

It is difficult to distinguish IC1 and IC2 in frequency domain. The information on IC1 and IC2 in time domain is difficult to discriminate. CWT has

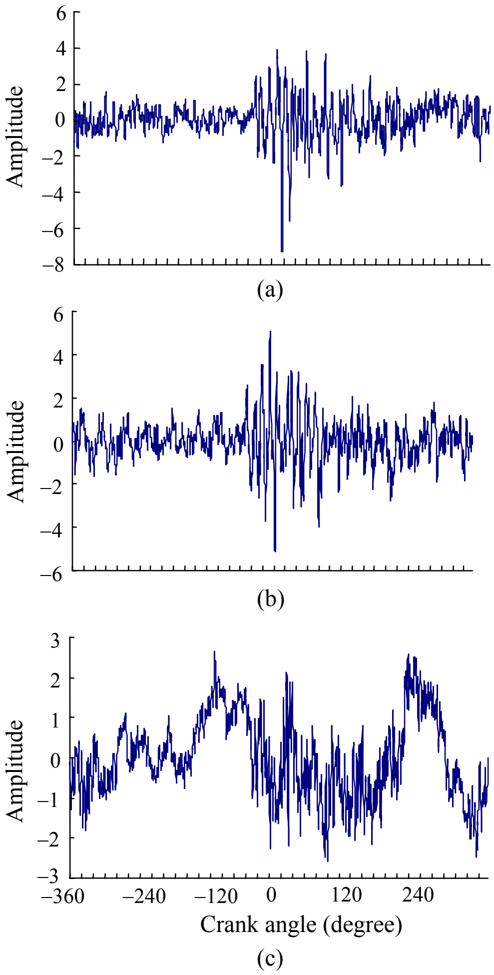

Fig.4 ICs in time domain. (a) IC1; (b) IC2; (c) IC3

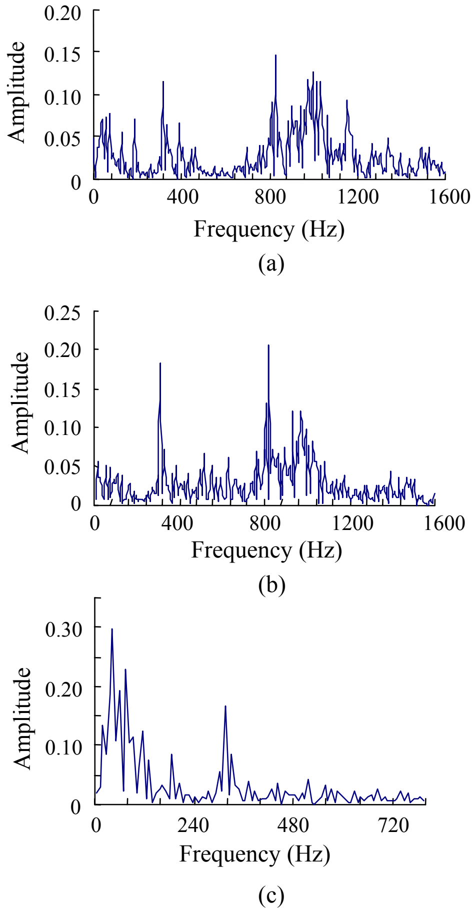

Fig.5 ICs in frequency domain. (a) IC1; (b) IC2; (c) IC3

been proved to be a powerful tool in time-frequency analysis and feature extraction (Lin and Qu, 2000) and is more suitable for analyze non-stationary acoustical signal from engine then FFT. The CWT of signals $x(t)$ is defined as:

$$W_x(a,t)=\int x(t)\psi_{a,b}^{*}\mathrm{d}t\,, \tag{7}$$

ψ is the mother wavelet, * stands for complex conjugate, a is the scales factor and b is the translation factor. It is important to choose a proper wavelet function as a mother wavelet of CWT. A guiding principle is to choose a wavelet whose shape in time domain is similar to the physical signals (Zheng and Li, 2002). Here complex Morlet wavelet, which is widely used in CWT because of its ability for time-frequency localization, was selected as a mother wavelet in CWT. Complex Morlet wavelet is defined by:

$$g(t)=\mathrm{e}^{\mathrm{j}\omega_0 t}\mathrm{e}^{-t^2/2}\,. \tag{8}$$

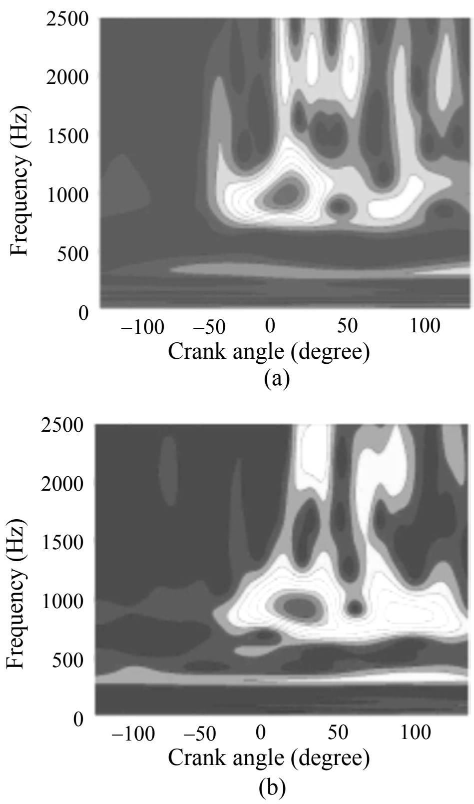

Fig.6 Cwt contour map of ICs; (a) IC1; (b) IC2

The Fourier transform of complex Morlet wavelet is

$$G(\omega)=\sqrt{2\pi}\mathrm{e}^{-(\omega-\omega_0)^2/2}\,. \tag{9}$$

The amplitude of wavelet coefficient represents the similarly of daughter wavelet with the physical signal. Fig.6 is the contour map of amplitude of the wavelet coefficient of IC1 and IC2. The wavelet coefficients in the white area of the graph are larger then the ones in the black area.

The results of CWT show the difference between IC1 and IC2. Both IC1 and IC2 have nearly the same distribution in frequency domain. But the main energy of IC1 centered around the top dead center and IC2 about 10° crank angle behind IC1 (Fig.7). Fuel combustion in cylinder occurs usually at nearly the top dead center. The rapid rise of pressure in cylinder causes vibration of the engine block and other components, and then the vibration transmits to the engine surface to form the air-borne noise. So the combustion noise usually also occurs around TDC. After combustion begins, the piston moves rapidly from one side of cylinder liners to the other side under the high pressure and speed. The excitation of piston slap can also induce an obvious noise radiation (Liu and

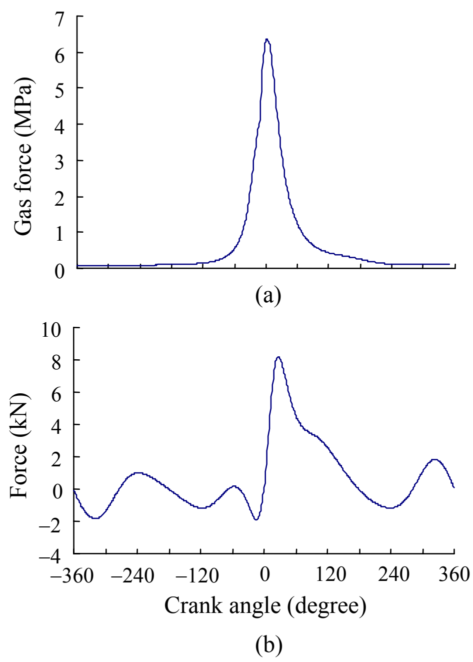

Fig.7 The gas force (a) and the piston slap (b)

Randall, 2005). The piston slap occurs after the combustion starts; the noise caused by the piston slap also follows the combustion induced noise emission. So IC2 is the noise component induced by the piston slap.

Correlation coefficient

Usually, the combustion-induced noise is the most important component in the diesel engine's overall acoustic signal, especially in the signal sampled over the cylinder head. So the signals over the cylinder head can seem to be "copies" of combustion-induced noise, though piston slap noise is included in the data. Correlation coefficient is the measurement of similarly of two samples. If two signals have a big correlation coefficient, they have more property in common. The correlation coefficients between three ICs and the acoustic signal sampled over the cylinder head are presented in Table 1.

Table 1 The correlation coefficient between ICs and the acoustic signals sampled over the cylinder

Type	Correlation coefficient
IC1	0.6808
IC2	0.1317
IC3	0.1446

Among the three ICs, IC1 has the maximum correlation coefficient with the signals sampled over the cylinder head. The result also proved that IC1 is the noise component induced by combustion.

CONCLUSION

Using ICA, the engine noise was effectively separated into several different ICs. The information from FFT and CWT results of ICs can reveal that the different IC is the different source of engine noise. ICA is a powerful tool for noise source separation and diagnosis of engine.

References

Comon, P., 1994. Independent component analysis, a new concept. *Signal Processing*, **36**(3):287-314. [doi:10.1016/0165-1684(94)90029-9]

Hao, Z.Y., Han, J., 2004. Identification of diesel front sound source based on continuous wavelet transforms. *Journal of Zhejiang University SCIENCE*, **5**(9):1069-1075. [doi:10.1631/jzus.2004.1069]

Hyvärinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, **10**(3):626-634. [doi:10.1109/72.761722]

Hyvärinen, A., Oja, E., 2000. Independent component analysis: algorithms and applications. *Neural Networks*, **13**(4-5):411-430. [doi:10.1016/S0893-6080(00)00026-5]

Li, W., Gu, F., Ball, A.D., 2001. A Study of the of the diesel engines using the independent component analysis. *Mechanical Systems and Signal Processing*, **15**(6):1165-1184. [doi:10.1006/mssp.2000.1366]

Lin, J., Qu, L.S., 2000. Feature extraction based on Morlet wavelet transform and its application for mechanical fault diagnosis. *Journal of Sound and Vibration*, **234**(1):135-148. [doi:10.1006/jsvi.2000.2864]

Liu, X.H., Randall, R.B., 2005. Blind source separation of internal combustion engine piston slap from other measured vibration signals. *Mechanical Systems and Signal Processing*, **19**(6):1196-1208. [doi:10.1016/j.ymssp.2005.08.004]

Zheng, H., Li, H., 2002. Gear fault diagnosis based on continuous wavelet transform. *Mechanical Systems and Signal Processing*, **16**(2-3):447-457. [doi:10.1006/mssp.2002.1482]