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

Comparison of different techniques for time-frequency analysis of internal combustion engine vibration signals*

Yang JIN[†1,2], Zhi-yong HAO[†1], Xu ZHENG[1]

([1]*Department of Energy Engineering, Zhejiang University, Hangzhou 310027, China*)

([2]*Department of Automotive Engineering, Hubei University of Automotive Technology, Shiyan 442002, China*)

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

Received Aug. 25, 2010; Revision accepted Mar. 1, 2011; Crosschecked June 8, 2011

Abstract: In this study, we report an analysis of cylinder head vibration signals at a steady engine speed using short-time Fourier transform (STFT). Three popular time-frequency analysis techniques, i.e., STFT, analytic wavelet transform (AWT) and S transform (ST), have been examined. AWT and ST are often applied in engine signal analyses. In particular, an AWT expression in terms of the quality factor *Q* and an analytical relationship between ST and AWT have been derived. The time-frequency resolution of a Gaussian function windowed STFT was studied via numerical simulation. Based on the simulation, the empirical limits for the lowest distinguishable frequency as well as the time and frequency resolutions were determined. These can provide insights for window width selection, spectrogram interpretation and artifact identification. Gaussian function windowed STFTs were applied to some cylinder head vibration signals. The spectrograms of the same signals from ST and AWT were also determined for comparison. The results indicate that the uniform resolution feature of STFT is not necessarily a disadvantage for time-frequency analysis of vibration signals when the engine is in stationary state because it can more accurately localize the frequency components excited by transient excitations without much loss of time resolution.

Key words: Short-time Fourier transform (STFT), Gaussian window, Time-frequency resolution limits, Internal combustion (IC) engine, Vibration signals, Analytic wavelet transform (AWT), S transform (ST)

doi:10.1631/jzus.A1000384 **Document code:** A **CLC number:** TB53; TP206+.1

1 Introduction

Joint time-frequency analysis is a versatile method by which a signal in time-frequency domain can be localized. It has been used in condition monitoring, noise source localization and fault diagnosis in internal combustion (IC) engines. Among all the techniques used in joint time-frequency analysis, the continuous wavelet transform (CWT) using complex Morlet wavelets, also called the analytic wavelet transform (AWT), is probably the most popular (Hao and Han, 2004; Peng and Chu, 2004; Zhu and Kim, 2005; Wu and Chen, 2006; Jing and Hao, 2009). The S transform (ST) has recently also become popular (Wang *et al*., 2003; Hao *et al*., 2008). In contrast, the earliest developed technique, the short-time Fourier transform (STFT), has not yet received much attention probably because of its uniform resolution feature, which has been considered to be its major shortcoming compared to the other two methods.

The advantages and disadvantages of a specific signal processing technique are contingent on the signals processed. A technique perfect for a specific signal type is quite often not applicable to another type. In this study, we aimed to find out if AWT and ST are better than STFT when applied to time-frequency analysis of cylinder head vibration signals under steady engine speeds.

For simplicity, only the Gaussian function was used as the STFT window function in this study. Thus, the term "STFT" in this paper refers to STFT with Gaussian window, unless noted otherwise.

* Project (No. 2011BAE22B05) supported by the National Key Technologies Supporting Program of China during the 12th Five-Year Plan Period

2 Gaussian function windowed STFT, ST, AWT, and their relationship

The Gaussian function windowed STFT, ST and AWT are all closely related to the Gaussian function. Both AWT and ST can be expressed in the form of STFT. Therefore, the resolutions of all three techniques are limited by the uncertainty principle and the spectrograms they generate are all related to the window width.

2.1 Gaussian function

The Gaussian function is the most localized function among all the functions in the time-frequency domain, i.e., the product of its time duration and frequency bandwidth is the smallest among all functions. The energy normalized form (Qian, 2002) is given by

$$g(t)=\sqrt[4]{\frac{\alpha}{\pi}}\mathrm{e}^{-\frac{\alpha}{2}t^2}, \tag{1}$$

where t is time and α is a constant. Both the time and frequency centers of $g(t)$ are 0. Its time and frequency domain standard deviations σ_t and σ_f are determined by Eqs. (2) and (3) below, respectively. σ_t and σ_f are usually employed to measure the time duration and frequency bandwidth, respectively, of a signal. In the following, we will refer to σ_t as the "window width".

$$\sigma_\mathrm{t}^2=\frac{1}{2\alpha}, \tag{2}$$

$$\sigma_\mathrm{f}^2=\frac{\alpha}{8\pi^2}. \tag{3}$$

From Eqs. (2) and (3), we can obtain

$$\sigma_\mathrm{f}=\frac{1}{4\pi\sigma_\mathrm{t}}. \tag{4}$$

A distinctive feature of the Gaussian function is that the function value is less than 1% of its peak value when $|t|>3\sqrt{2}\sigma_\mathrm{t}$. The discrete form of the Gaussian function $g(t)$ can be written as

$$g^i(t)=\sqrt[4]{\frac{1}{2\pi\sigma_\mathrm{t}^2}}\mathrm{e}^{-\frac{t_i^2}{4\sigma_\mathrm{t}^2}}=\sqrt[4]{\frac{1}{2\pi\sigma_\mathrm{t}^2}}\mathrm{e}^{-\frac{i_t^2}{4\sigma_\mathrm{t}^2\cdot f_\mathrm{s}^2}}$$

$$=\sqrt{f_\mathrm{s}}\sqrt[4]{\frac{1}{2\pi\sigma_\mathrm{t}^2 f_\mathrm{s}^2}}\mathrm{e}^{-\frac{i_t^2}{4\sigma_\mathrm{t}^2\cdot f_\mathrm{s}^2}}, \tag{5}$$

where $i_\mathrm{t}=-N_\mathrm{t}, -N_\mathrm{t}+1,\ldots, N_\mathrm{t}-1$; f_s represents the sampling frequency; $2N_\mathrm{t}$ denotes the number of the points in the time domain; and i_t is the index of the discrete window series.

Discrete Gaussian windows for several specific parameters are shown in Fig. 1. The most significant advantage of Gaussian windows over the commonly used Cosine windows such as Hanning, Hamming or Blackman, is that they have an intuitive and physically meaningful parameter σ_t that is independent of the number of points (Fig. 1a). In particular, if the number of the points is large enough, the time-frequency resolution of a Gaussian function can be fixed well by keeping σ_t constant while increasing the number of points. In this way, the Fourier transform (FT) lines can be increased, which helps improve the smoothness of the spectrogram without sacrificing the resolution.

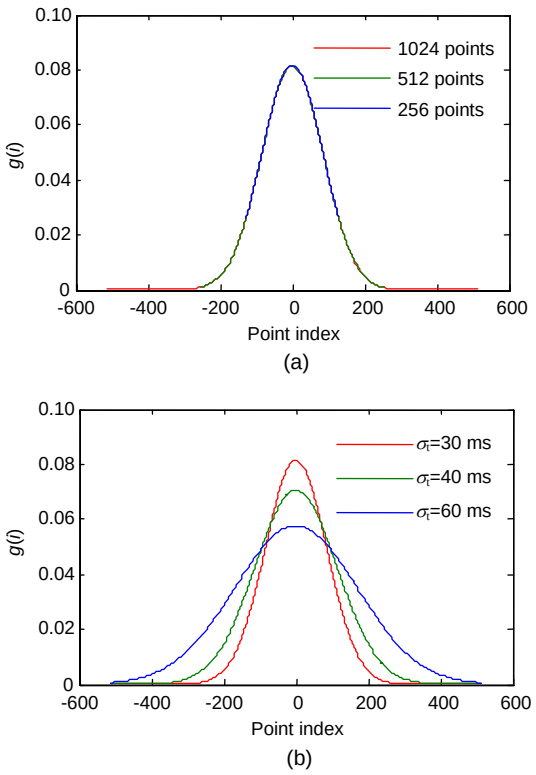

Fig. 1 Plots of discrete Gaussian windows of different parameters
(a) f_s=2 kHz, σ_t=30 ms; (b) f_s=2 kHz, 1024 points

2.2 Gaussian function windowed STFT

The Gaussian function windowed STFT of a signal $s(t)$ can be written as

$$\begin{aligned}\mathrm{STFT}_{s,g}(\tau, f) &= \int_{-\infty}^{+\infty} s(t) g^*(t-\tau) \mathrm{e}^{-j2\pi ft} \mathrm{d}t \\ &= \int_{-\infty}^{+\infty} s(t) g(t-\tau) \mathrm{e}^{-j2\pi ft} \mathrm{d}t,\end{aligned} \tag{6}$$

where * denotes the complex conjugate operator. The second expression is correct because $g(t)$ is real.

2.3 AWT

AWT is a special form of CWT (Zhu and Kim, 2005), whose mother wavelet is a complex Morlet wavelet, also known as the Gabor wavelet. Because all the wavelets in the Gabor wavelet family are analytic, that is, they have no negative frequency component, this specific CWT is also referred to as the AWT.

2.3.1 Gabor wavelet family in terms of the quality factor Q

The energy normalized mother Gabor wavelet is defined by

$$\psi_{0,g}(t) = g(t)\mathrm{e}^{j2\pi f_0 t} = \sqrt[4]{\frac{\alpha_0}{\pi}} \mathrm{e}^{-\frac{\alpha_0}{2}t^2 + j2\pi f_0 t}, \tag{7}$$

where f_0 and α_0 are the center frequency and a constant related to the time duration and frequency bandwidth of the mother wavelet, respectively. A family of Gabor wavelets can be obtained by translating and dilating the mother wavelet. The quality factor Q of each wavelet, which is defined as the ratio of σ_f, a half of the frequency bandwidth (i.e., frequency domain standard deviation), to f, the center frequency, is always the same

$$Q = \frac{\sigma_\mathrm{f}}{f} = \frac{1}{4\pi\sigma_\mathrm{t} f}. \tag{8}$$

The Gabor wavelet with center frequency, f, and time translation, τ, can be written as

$$\begin{aligned}\psi_{f,\tau}(t) &= g_{f,0}(t-\tau)\mathrm{e}^{j2\pi f(t-\tau)} \\ &= \sqrt[4]{8\pi Q^2 f^2}\,\mathrm{e}^{-4\pi^2 Q^2 f^2 (t-\tau)^2} \mathrm{e}^{j2\pi f(t-\tau)},\end{aligned} \tag{9}$$

where

$$g_{f,0}(t) = \sqrt[4]{8\pi Q^2 f^2}\,\mathrm{e}^{-4\pi^2 Q^2 f^2 t^2} \tag{10}$$

is the envelope of $\psi_{f,0}(t)$. It is also a normalized Gaussian function. Clearly, the whole Gabor wavelet family can be described by a single parameter Q. The frequency bandwidth increases and the time width of a Gabor wavelet with a given center frequency decreases with Q.

2.3.2 AWT in terms of quality factor Q

The AWT of a signal $s(t)$ can be rewritten as a function of Q

$$\begin{aligned}\mathrm{AWT}_s(\tau, f) &= \int_{-\infty}^{+\infty} s(t)\psi_{f,\tau}^*(t)\mathrm{d}t \\ &= \int_{-\infty}^{+\infty} s(t)\sqrt[4]{8\pi Q^2 f^2}\,\mathrm{e}^{-4\pi^2 Q^2 f^2 (t-\tau)^2} \mathrm{e}^{-j2\pi f(t-\tau)}\mathrm{d}t.\end{aligned} \tag{11}$$

2.3.3 Amplitude normalized AWT (AWT^m)

The wavelets of the AWT defined by Eq. (11) have the same energy at different frequencies/scales. The AWT transform of a signal, by Eq. (11), that is composed of two single tones of the same amplitude but with different frequencies shows a large magnitude at the low frequency and small amplitude at the high frequency. More precisely, the magnitude is inversely proportional to the square root of the frequency. However, in some real-world applications, such as IC engine signals, the amplitude of the signal components also need to be measured (Yang and Hao, 2006). The amplitude normalized AWT is given in Eq. (12):

$$\mathrm{AWT}_s^\mathrm{m}(\tau, f) = \sqrt{\frac{f}{f_\mathrm{s}}}\mathrm{AWT}_s(\tau, f). \tag{12}$$

2.4 ST

2.4.1 ST

The window function in STFT does not change at each point in time-frequency domain. In contrast, an essential feature of ST (Stockwell, 1996; Wang, 2010) is that the window width of window functions $w_\mathrm{S}(t,f)$ is inversely proportional to the frequency

$$w_S(t, f) = \frac{|f|}{\sqrt{2\pi}} e^{-\frac{f^2}{2}t^2}. \quad (13)$$

Although $w_S(t,f)$ is a Gaussian function, it is not energy normalized. The window function of ST at frequency f (Eq. (13)) can be related to $g_{f,0}(t)\big|_{Q=1/(2\sqrt{2}\pi)}$ (Eq. (10)), which is the envelope of a member of center frequency f in the Gabor wavelet family with $Q=1/(2\sqrt{2}\,\pi)$, by

$$w_S(t, f) = \sqrt[4]{\frac{f^2}{4\pi}}\, g_{f,0}(t)\big|_{Q=1/(2\sqrt{2}\pi)}. \quad (14)$$

Consequently, $s(t)$'s ST can be expressed as

$$\begin{aligned}
&\mathrm{ST}_s(\tau, f) = \mathrm{STFT}_{s,w_S}(\tau, f)\\
&= \int_{-\infty}^{+\infty} s(t) w_S^*(t-\tau, f) e^{-j2\pi f t} dt\\
&= \sqrt[4]{\frac{f^2}{4\pi}} e^{-j2\pi f\tau} \int_{-\infty}^{+\infty} s(t)\, g_{f,0}(t-\tau)\big|_{Q=1/(2\sqrt{2}\pi)} e^{-j2\pi f(t-\tau)} dt\\
&= \frac{1}{\sqrt[4]{4\pi}} e^{-j2\pi f\tau} \sqrt{f}\, \mathrm{AWT}_s(\tau, f)\big|_{Q=1/(2\sqrt{2}\pi)}\\
&= \frac{\sqrt{f_s}}{\sqrt[4]{4\pi}} e^{-j2\pi f\tau}\, \mathrm{AWT}_s^{\mathrm{m}}(\tau, f)\big|_{Q=1/(2\sqrt{2}\pi)}, \quad (15)
\end{aligned}$$

which indicates that ST is no more than an amplitude and phase modulated version of a specific $\mathrm{AWT}^{\mathrm{m}}$ with $Q=1/(2\sqrt{2}\pi)$.

2.4.2 Modified S transform (ST^{m})

A modified ST can be defined as

$$\begin{aligned}
\mathrm{ST}_s^{\mathrm{m}}(\tau, f) &= \sqrt[4]{\frac{4\pi}{f_s^2}} \mathrm{ST}_s(\tau, f)\\
&= e^{-j2\pi f\tau}\, \mathrm{AWT}_s^{\mathrm{m}}(\tau, f)\big|_{Q=1/(2\sqrt{2}\pi)}.
\end{aligned} \quad (16)$$

In this way, the numerical implementation of $\mathrm{ST}^{\mathrm{m}}(\tau, f)$ and $\mathrm{AWT}^{\mathrm{m}}(\tau, f)$ can be unified. Moreover, $\mathrm{ST}^{\mathrm{m}}(\tau, f)$ can be obtained by a simple phase shift of $\mathrm{AWT}_s^{\mathrm{m}}(\tau, f)\big|_{Q=1/(2\sqrt{2}\pi)}$. In addition, the magnitudes of $\mathrm{ST}^{\mathrm{m}}(\tau, f)$ and $\mathrm{AWT}_s^{\mathrm{m}}(\tau, f)\big|_{Q=1/(2\sqrt{2}\pi)}$ are always identical.

3 Simulation tests of STFT

3.1 The lowest distinguishable frequency

It is possible that only a few periods of a single tone signal are inside the time window due to its limited time width. So, an STFT spectrogram may not be able to show a single tone clearly. A synthetic signal Sig1 was designed to help determine the lowest distinguishable frequency of STFT:

$$\mathrm{Sig1}(t) = 7\sin(2\pi f_0 t),$$

where f_0 is a constant. The STFTs with a given σ_t were performed on Sig1 with different f_0 to find out the critical value $f_{\sigma_t}^{\min}$ for f_0, at which the STFT spectrogram can show the signal clearly. A series of simulations showed that $f_{\sigma_t}^{\min}$ is about $2.5\sigma_f$. Two STFT spectrograms with σ_t=66.5 ms are shown in Fig. 2. L indicates the length of the Gaussian window, which is also the fast Fourier transform (FFT) frame size because no zero padding is applied; TS indicates the time step of the window, in number of sample points, in STFT implementation. The equivalent overlap will be $(L-\mathrm{TS})/L\times 100\%$.

In Fig. 2a, $2.5\sigma_f$ is 3 Hz, which is smaller than 7 Hz. Clearly, the time window was sufficient to show the frequency of the signal. However, as the frequency of Sig1 is reduced to 3 Hz and equal to $2.5\sigma_f$ (Fig. 2b), the corresponding STFT spectrogram shows artifacts at low frequencies. It became more and more difficult to distinguish spectrograms as the frequency of Sig1 diminished.

3.2 Time-frequency resolution limits

IC engine vibration signals have two essential features. First, there are different frequency response components during the occurrence of excitation events. Second, some excitations are very close in the time domain, even simultaneous. These may cause response components of the same frequency. The STFT can be used to identify these excitation events if these response components can be differentiated from each other via STFT spectrograms (Vulli *et al*., 2009). Therefore, a study of the ability of STFT to distinguish between small frequencies spaced response components at the same instant and components of the same frequency but of small time interval, is highly desirable.

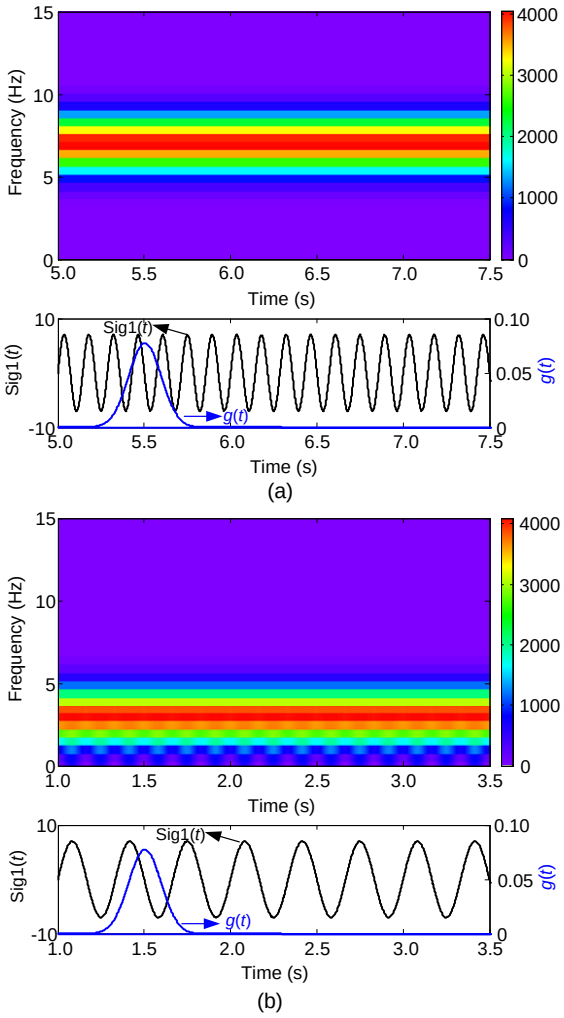

Fig. 2 STFT spectrograms and time waveforms of Sig1
(a) f_0=7 Hz, σ_t=66.5 ms, $2.5\sigma_f$=2.5/(4πσ_t) =3 Hz; (b) f_0=3 Hz, σ_t=66.5 ms, $2.5\sigma_f$=2.5/(4πσ_t)=3 Hz. f_s=1000 Hz, L=2048, TS=2. The color represents square of the STFT magnitude in linear mode

The spectrograms from CWT, ST and STFT are all approximate descriptions or decompositions of the original signal. The time width σ_t and frequency bandwidth σ_f of a Gaussian window, which are indirect and qualitative descriptions of the time-frequency resolution of STFT, are inversely proportional to each other. The better is the time resolution, the worse is the frequency resolution and vice versa. For a given window width σ_t, what is the minimum time/frequency spacing between the two components that can be clearly differentiated using the STFT spectrogram? To help answer this question and to quantitatively determine the STFT's time-frequency resolution limits, we defined a synthetic signal Sig2 as follows:

$$\begin{aligned}\mathrm{Sig2}(t) &= \mathrm{S1}(t)+\mathrm{S2}(t)+\mathrm{S3}(t)\\ &\quad+\mathrm{S4}(t)+\mathrm{S5}(t)+\mathrm{S6}(t),\end{aligned}$$

where

$$\mathrm{S1}(t)=\sin(2\pi(200t+5t^2)),\ t\in[0,10],$$

$$\mathrm{S2}(t)=\sin(2\pi(300t-5t^2)),\ t\in[0,10],$$

$$\mathrm{S3}(t)=\sin(2\pi\times150t)+\sin(2\pi\times700t),$$
$$t\in[4.05,4.1]\cup[4.35,4.4]\cup[4.65,4.7],$$

$$\mathrm{S4}(t)=\sin(2\pi\times700t)+\sin(2\pi\times740t),$$
$$t\in[5.05,5.1]\cup[5.15,5.2]\cup[5.25,5.3],$$

$$\mathrm{S5}(t)=\sin(2\pi\times150t),$$
$$t\in[5.05,5.1]\cup[5.13,5.18]\cup[5.21,5.26],$$

$$\mathrm{S6}(t)=8,\ t\in[6,6].$$

Sig2 contains short-time components occurring simultaneously and with close frequencies, such as the two short-time single tone components in S4, which have a frequency of 700 and 740 Hz, respectively. Sig2 also contains short-time components of the same frequency which are very close in the time domain, such as the components in S3 with a time interval 250 ms, S4 with a time interval 50 ms, and S5 with a time interval 30 ms. Similarly, Sig2 also has long-time evolving components such as S1 and S2, and a spike component such as S6. The frequency range of the components in Sig2 is broad.

The intuitive interpretation of the theoretical components of Sig2 can be inferred from Fig. 3. It is probably the best STFT spectrogram among many of different σ_t, whose window width σ_t is 10 ms. Fig. 3 shows clearly almost all the theoretical components of Sig2, except the portions between 3.5–6.5 s of S1 and S2. Sig2 was analyzed by STFT with different σ_t values. Three typical spectrograms are shown in Figs. 3–5 for comparison.

It can be inferred from Fig. 3 that:

1. During the period of 3.5–6.5 s, the frequency difference between S1 and S2 at the same time instant is less than 1/(3σ_t)=33.3 Hz.

2. During the period from 3.5–6.5 s, as the frequency difference between S1 and S2 at the same time instant becomes smaller, it is harder to distinguish the two components S1 and S2. In the time domain, this phenomenon can be explained by the

fact that, as the frequencies of S1 and S2 become closer, the modulation frequency of the amplitude modulated wave formed by the superposition of S1 and S2 becomes smaller. The carrier frequency is the average of the instantaneous frequencies of S1 and S2 while the modulation frequency is one half of the difference between these two frequencies. Accordingly, the time distance between the nodes and antinodes becomes larger in time domain. This can be seen clearly from the detailed waveform in Fig. 5. The dark and bright bands along the vertical direction in the spectrogram correspond to the nodes and antinodes in the time waveform, respectively.

3. Fig. 3 clearly shows the two short-time single tones in S4. In particular, the frequency difference between the single tones is 40 Hz, being larger than $1/(3\sigma_t)$=33.3 Hz.

4. Although the spike occurring at 6 s has a large value of 8, its corresponding spectrogram values in Fig. 3 are much smaller than those of S1 and S2, which have an amplitude of 1.

Fig. 4 shows the result of σ_t=5 ms. The following observations can be made:

1. Artifacts between S1 and S2 are observed from 2–8 s, where the frequency difference between S1 and S2 at the same time instant is not larger than 60 Hz, being smaller than $1/(3\sigma_t)$=66.7 Hz.

2. In S4, a vertical artifact appears between the

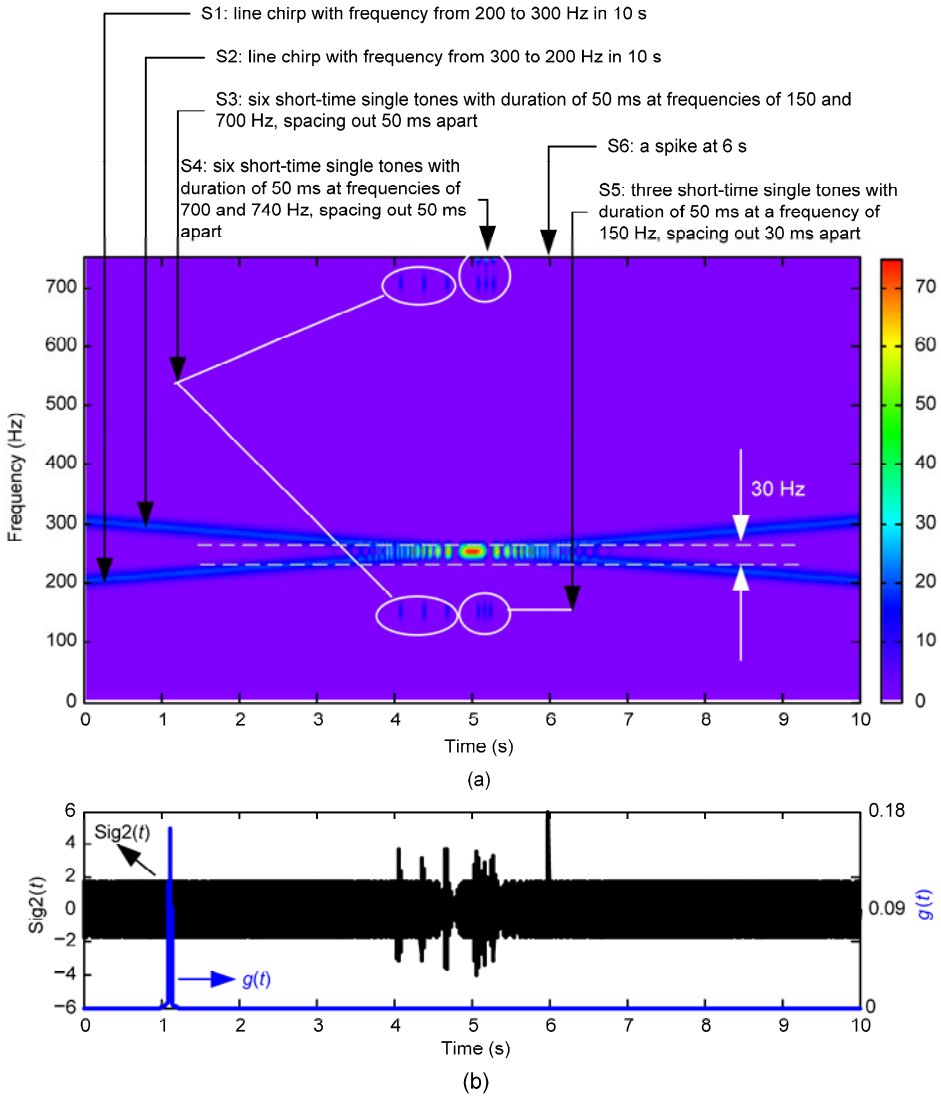

Fig. 3 STFT spectrogram with σ_t=10 ms (a) and time waveforms (b) of Sig2

f_s=1500 Hz, L=1024, TS=8. $3\sigma_t$=30 ms, $1/(3\sigma_t)$=33.3 Hz, $2.5\sigma_f$=$2.5/(4\pi\sigma_t)$=19.9 Hz. The color represents square of the STFT magnitude in linear mode

short-time single tones of frequencies 700 and 740 Hz.

Note that the two regions mentioned above in Fig. 4 can be easily misinterpreted as a single component with a larger bandwidth.

A few points can be made from Fig. 5 in which σ_t is 20 ms:

1. The artifact region between S1 and S2 is reduced to 4.2–5.8 s. In this interval, the S1 and S2 frequency difference at the same time point is not larger than 16 Hz.

2. Artifacts can be observed in both S4 and S5. In particular, the artifact in S5 is more pronounced than that in S4. The time spacings between components of the same frequency in S5 and S4 are 30 and 50 ms, respectively.

3. The spike component S6 cannot be identified.

We can conclude from Figs. 3–5 that for an STFT spectrogram with a window width of σ_t:

1. To be able to differentiate two components in the STFT spectrogram, the low limit of the frequency spacing between these two components at the same time instant is $\Delta f_{\sigma_t}^{\min} \approx 1/(3\sigma_t)$.

2. To be able to identify two components of the same frequency in the STFT spectrogram (components which do not overlap in the time domain), the low limit of their time spacing is $\Delta t_{\sigma_t}^{\min} \approx 3\sigma_t$.

4 Discussion

1. The ability to differentiate the components in a signal from STFT spectrograms is critically dependent on the selection of window width. Specifically, to be able to clearly differentiate the components with a slowly varying frequency, the stationary components, and the short-time single tones in a signal from the STFT spectrograms, the lowest distinguishable component frequency should be about $2.5\sigma_t$, while the time and frequency resolution limits should be around $3\sigma_t$ and $1/(3\sigma_t)$, respectively.

2. These three limits have to be met simultaneously. For instance, as shown in Fig. 4, although the frequencies of both S1 and S2 are larger than $2.5\sigma_t$, it

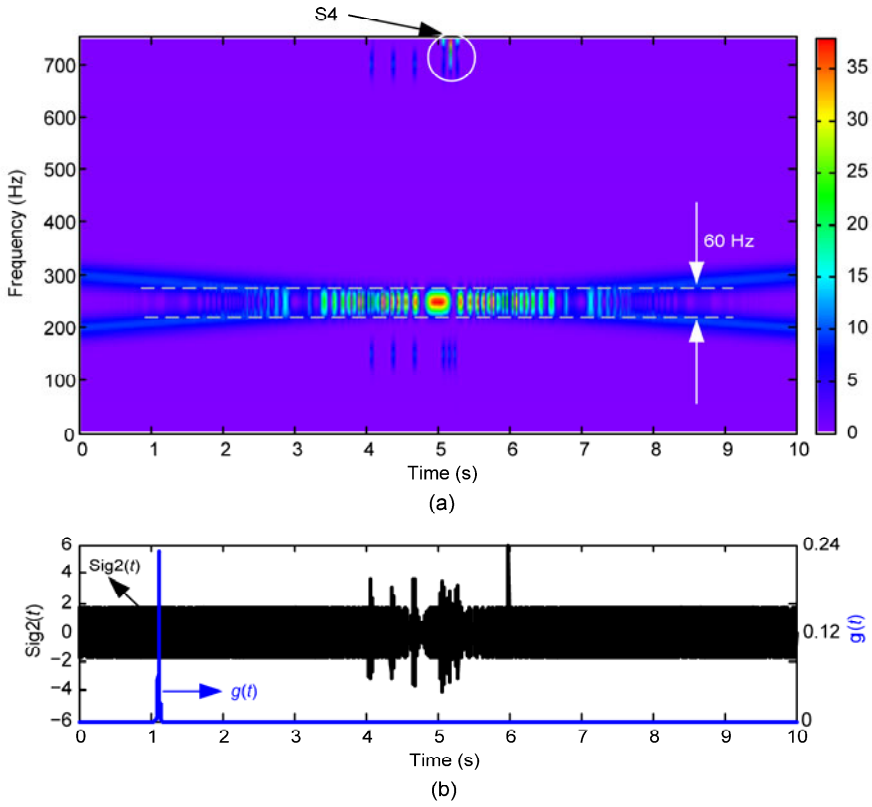

Fig. 4 STFT spectrogram with σ_t=5 ms (a) and time waveforms (b) of Sig2

f_s=1500 Hz, L=1024, TS=8. $3\sigma_t$=15 ms, $1/(3\sigma_t)$=66.7 Hz, $2.5\sigma_t$=$2.5/(4\pi\sigma_t)$=39.8 Hz. The color represents square of the STFT magnitude in linear mode

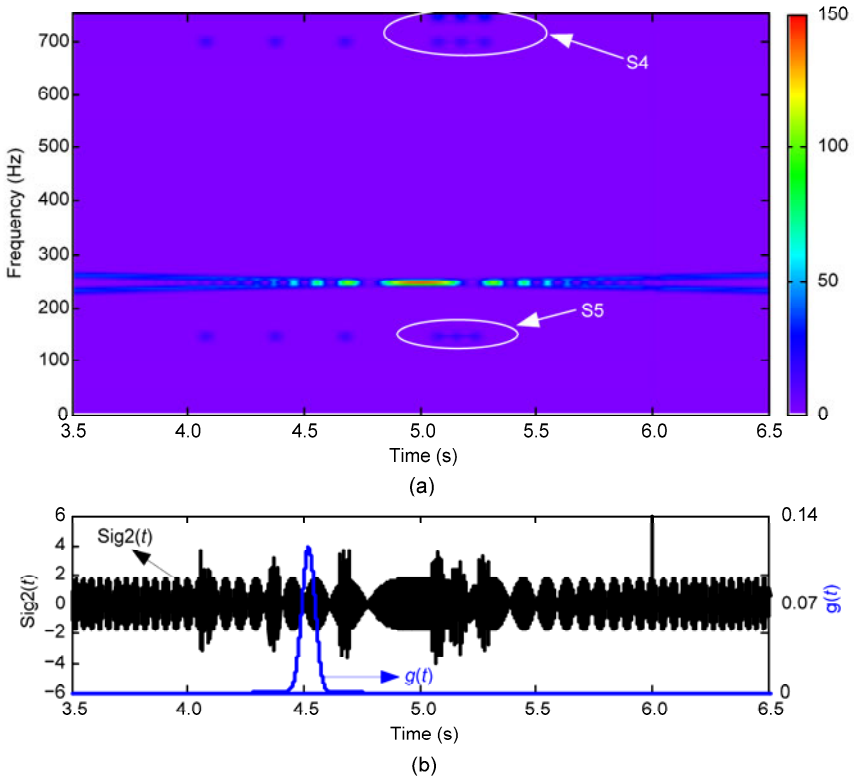

Fig. 5 STFT spectrogram with σ_t=20 ms (a) and time waveforms (b) of Sig2

f_s=1500 Hz, L=1024, TS=8. $3\sigma_t$=60 ms, $1/(3\sigma_t)$=16.7 Hz, $2.5\sigma_f=2.5/(4\pi\sigma_t)$=9.9 Hz. The color represents square of the STFT magnitude in linear mode

is still impossible to clearly differentiate them in the period from 2–8 s, for the frequency resolution limit $1/(3\sigma_t)$ is not satisfied.

3. By varying σ_t, the STFT spectrograms can also be used to identify the short-period components and the transient spike components that are mixed with pseudo-stationary components as in Sig2. However, those small-energy and short-period components may be hard to differentiate in the spectrogram due to their small energy. By comparing the results shown in Figs. 3–5, we can conclude that small window widths are good for highlighting short-period small-energy components. This is particularly true for a spike.

4. It is usually difficult to find a suitable window width that can resolve all the components in a signal. Discrimination depends upon the detailed composition of the signal. In practical applications, compromises are usually needed to realize a specific concern. Although the three limits proposed in this section are based on simulation results using harmonic waves and seem to be empirical, they can be used to interpret the spectrograms correctly and to identify the artifacts.

5 Characterizing the excitations in an IC engine under stationary state and engine body vibration signals

The excitations arousing vibration in an IC engine in a stationary state can be categorized into two types: long-term continuous and transient.

Continuous excitations are smooth. Their frequencies are very low and are usually an integral or fractional multiples of the engine's turning frequency, i.e., harmonics. Continuous excitations include mainly the reciprocating, centrifugal inertial force/moment, and overturning moment. They usually cause the stationary rigid body vibration of the entire engine, which is also continuous.

Intermittent and transient excitations may result

from fuel injection, cylinder pressure fluctuation due to combustion, impacts between matching parts of clearance, and gas pressure pulsation during the intake and exhaust strokes. Among all the impacts between parts, slapping of the cylinder bore by the piston and impacts on the valve train are prominent. Both are transient, and consequently, the corresponding responses are damped and intermittent. The constituent frequencies of these responses involve the frequency spectrum of excitation and the frequency response function (FRF) between the excitation point and the response point. Because the natural frequencies of the engine structural components, such as cylinder head and cylinder block, are usually very high, so are their responses to transient events. The response frequencies do not change a lot at a steady engine speed.

Therefore, the vibration responses to the excitations in an IC engine at a steady engine speed are the superposition of cyclic and continuous waves as well as the short-time damping waves corresponding to a variety of transient excitations. These signals can be analyzed well using Gaussian function windowed STFT.

6 Applications of Gaussian function windowed STFT

6.1 Signal collection

Vibration signals were collected at an idle speed of 800 r/min in a 4-cylinder 4-stroke common-rail diesel engine, whose firing order was 1-3-4-2. Two single-axial accelerometers were located at positions A1 and A4 on the side of the cylinder head (Fig. 6). They were aligned with the central axis of cylinders 1 and 4, respectively. Lateral vibration signals that were perpendicular to the longitudinal axis of the crankshaft were measured.

The signals collected were called A01 and A04 according to their respective locations. A photoelectric tachometer was used to indicate the top dead center (TDC) position as the event reference, which generated one pulse per crankshaft revolution. All three signals were measured simultaneously. The sampling frequency was 12.76 kHz and the lowpass cutting frequency was 4 kHz.

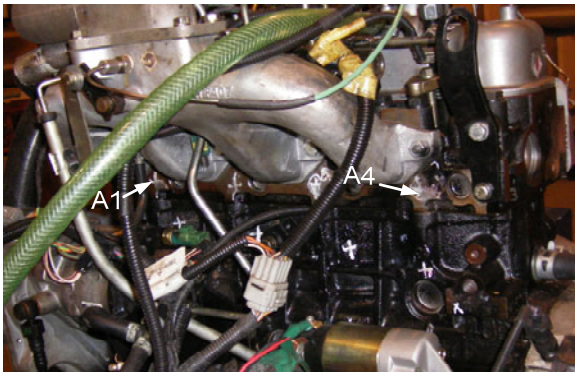

Fig. 6 Test engine and locations of accelerometers

6.2 Time-frequency analysis of the signals

6.2.1 Initial window width σ_t

As mentioned above, an STFT spectrogram depends greatly on the window width. The window width selection here was guided by the conclusions drawn in Section 3. We take a specific signal as an example: a major source of mechanical noise in an IC engine is the impacts between the parts as the valves are opening or closing. A schematic valve diagram for the test engine is shown in Fig. 7. It can also be viewed as the sequence of major excitation events on the cylinder head. The exhaust valves are open regularly at a 26° crankshaft angle (CA) (5.417 ms at 800 r/min) after each power stroke TDC. The intake valves are closed regularly at 55° CA (11.5 ms at 800 r/min) after each power stroke TDC. Each such event may excite vibrations of components at the same frequency. As shown in Section 3, an STFT spectrogram with a window width σ_t has a time resolution of about $3\sigma_t$. Accordingly, to be able to differentiate between the vibration responses due to combustion and those due to exhaust valve closing (providing that there is sufficient corresponding energy in the vibration signal), and 5.417/3=1.81 ms should be selected as the initial value of σ_t. In Section 3, the time resolution limit is referred to as the time spacing between the end of one signal component and the start of another, while here 5.417 ms is actually the time interval between the start points of the excitation events. We use σ_t=5.417/3=1.81 ms as an initial value of σ_t because impulse responses on typical engine structures last for a few milliseconds (Vulli *et al*., 2009).

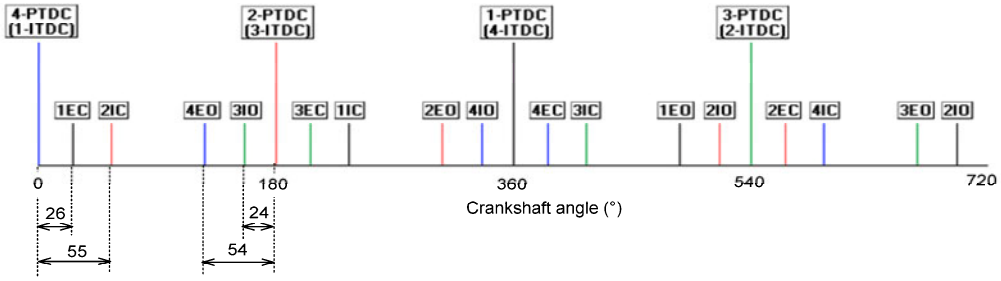

Fig. 7 Valve diagram of the test engine

PTDC: TDC of power stroke; ITDC: TDC of intake stroke; EO: exhaust valve opening; EC; exhaust valve closing; IO: intake valve opening; IC: intake valve closing. The number before the symbols denotes the cylinder number

The STFT spectrograms with σ_t=1.81 ms of the measured signals are shown in Figs. 8 and 9.

6.2.2 Discussion

The vibration signals in Figs. 8 and 9 are comprised of continuous and transient, intermittent responses. In the idle state, both the combustion and valve train mechanical excitations are small and, as a result, the measured response signals to the vibrations are also small. The FT of the two signals revealed that their energies were almost completely concentrated at the frequencies of 53 and 80 Hz, being the 4th and 6th order turning frequencies, respectively, at 800 r/min. These are the predominant excitation frequencies of the stationary reciprocating inertia force and overturning moment in this 4-cylinder 4-stroke engine. However, the transient responses to the transient excites were hard to identify from FT. The lowest distinguishable frequency of a single tone from an STFT spectrogram with a window width of 1.81 ms is 110 Hz according to Section 3. Therefore, to identify the continuous and low-frequency components as well as the intermittent high-frequency response components, a different window width should be used. In the following part of this section, the focus will be on the responses to transient excitations, which are believed to be the major noise sources.

A careful examination of Figs. 8 and 9 indicates that the STFT spectrograms with σ_t=1.81 ms are not bad despite having some regions that cannot be identified clearly.

In Figs. 8 and 9, the crankshaft angle positions at which the tachometer indicated the occurrence of a TDC are represented by the dashed line L1, while the instants of exhaust and intake valve closing are indicated by the dashed lines L2 and L3, respectively. The dashed line L1 is spaced at 26° CA from L2 and 55° CA from L3. The dashed lines L4 and L1 are separated from each other by one cycle—720° CA. The differences in vibration responses to the excitations at different crankshaft angles are obvious. The variations over the combustion process can also be revealed by comparing the spectrograms at L1 and L4. The following observations can be made:

1. In Fig. 8, the responses to exhaust valve closing are hard to identify from A01's spectrogram. This may be due to the low response energy. However, in Fig. 9, the weak components near the frequencies of 900 Hz and 3.5 kHz at L2 are very likely caused by exhaust valve closing.

2. In Fig. 8, the components of the frequencies 2.84 and 3.07 kHz at L3 in A01 are probably the responses to the intake valve closing, whereas they are hard to distinguish from A04 in Figs. 4 and 5.

3. In Fig. 9, the four components at L1 around the frequencies 1.4, 1.9, 2.3 and 2.8 kHz are very likely due to combustion excitations as well as piston slapping. The part highlighted by a smaller circle in Fig. 9 looks like a component but its frequency band exceeds the frequency resolution limit 184 Hz (the window width σ_t=1.81 ms). It is reasonable to conclude that this pseudo-component may be composed of several real components with close frequencies. This can be verified by doing STFT with increased window widths. The STFT spectrogram with σ_t=4 ms of this pseudo-component is shown in Fig. 9 highlighted by a bigger circle, in which four components are clearly evident.

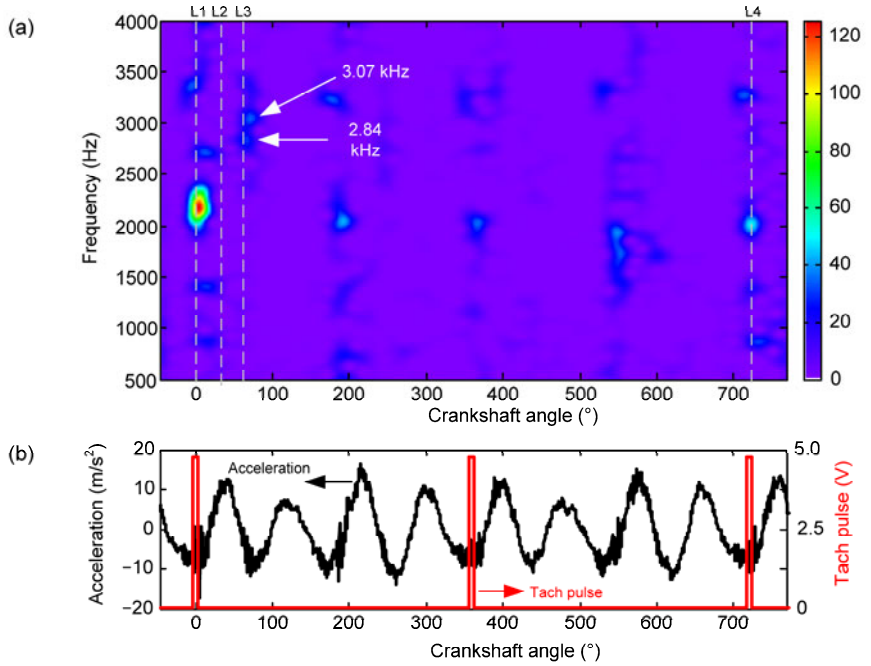

Fig. 8 A01's STFT spectrogram with σ_t=1.81 ms (a) and time waveforms (b)

f_s=12.76 kHz, L=4096, TS=4. $1/(3\sigma_t)$=184 Hz, $2.5\sigma_t=2.5/(4\pi\sigma_t)$=110 Hz. The color represents square of the STFT magnitude in linear mode

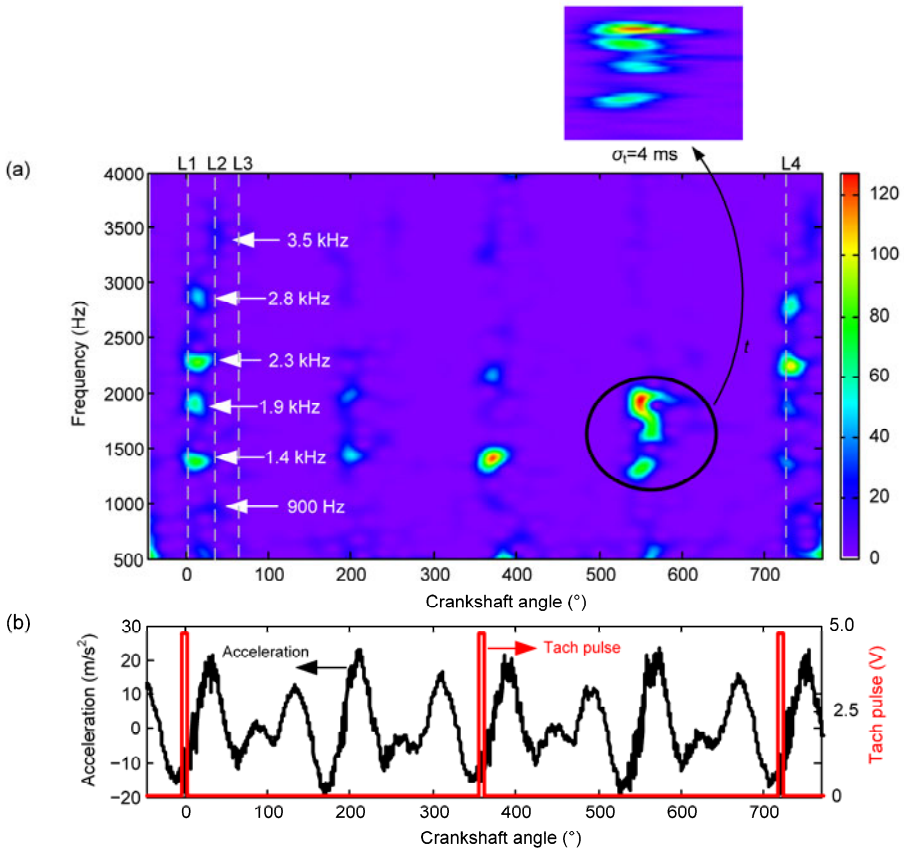

Fig. 9 A04's STFT spectrogram with σ_t=1.81 ms (a) and time waveforms (b)

f_s=12.76 kHz, L=4096, TS=4. $1/(3\sigma_t)$=184 Hz, $2.5\sigma_t=2.5/(4\pi\sigma_t)$=110 Hz. The color represents square of the STFT magnitude in linear mode

The A01's ST^m and AWT^m spectrograms with Q=0.025 are shown in Figs. 10 and 11, respectively. The ST/ST^m and AWT^m/AWT employ different window widths at different frequencies, a wider window at lower frequencies and a narrower window at higher frequencies.

Specifically, ST/ST^m always uses a window width of 0.707 ms at a frequency of 1 kHz (frequency resolution=471 Hz) and a window width of 0.202 ms at 3.5 kHz (frequency resolution=1.65 kHz). Accordingly, it cannot differentiate the components of frequency difference 471 Hz in the frequency range from 1 to 3.5 kHz. The response frequencies cannot be identified accurately from Fig. 10. In particular, without knowledge of the effects of window width on the spectrogram, the circled region in Fig. 10 could be easily interpreted as a single component of a wide frequency band from 1.8 to 4 kHz.

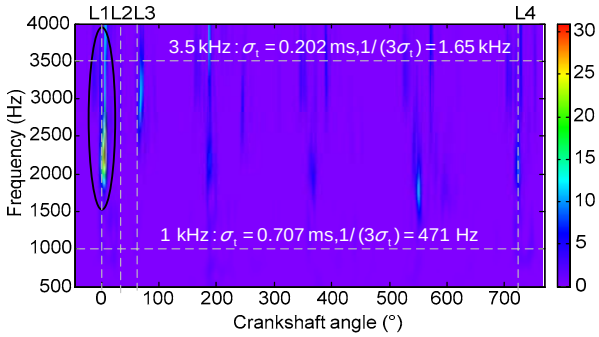

Fig. 10 A01's ST^m spectrogram

The color represents square of the ST^m magnitude in linear mode

The AWT^m spectrogram of A01' with Q=0.025 is shown in Fig. 11. The window width at 1 kHz is 3.183 ms (the frequency resolution limit=105 Hz), and the window width at 3.5 kHz is 0.909 ms (the frequency resolution limit=366.5 Hz). The frequency resolution limits at each frequency in Fig. 11 are smaller than those in Fig. 10, while the frequency resolution limits at the frequencies larger than 1.76 kHz (where σ_t=1.81 ms) in Fig. 11 are larger than those in Fig. 8. Therefore, the two components of 2.84 and 3.07 kHz at L3 cannot be identified from Fig. 11, whereas they can be clearly seen in Fig. 8. Although many Q have been tried, it is hard to find an appropriate value due to the multi-resolution feature of AWT (Fig. 12).

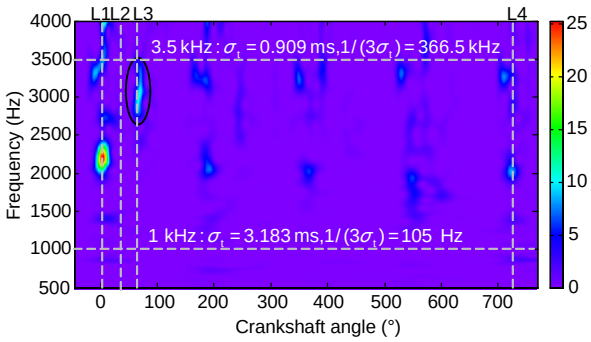

Fig. 11 A01's AWT^m spectrogram with Q=0.025

The color represents square of the AWT^m magnitude in linear mode

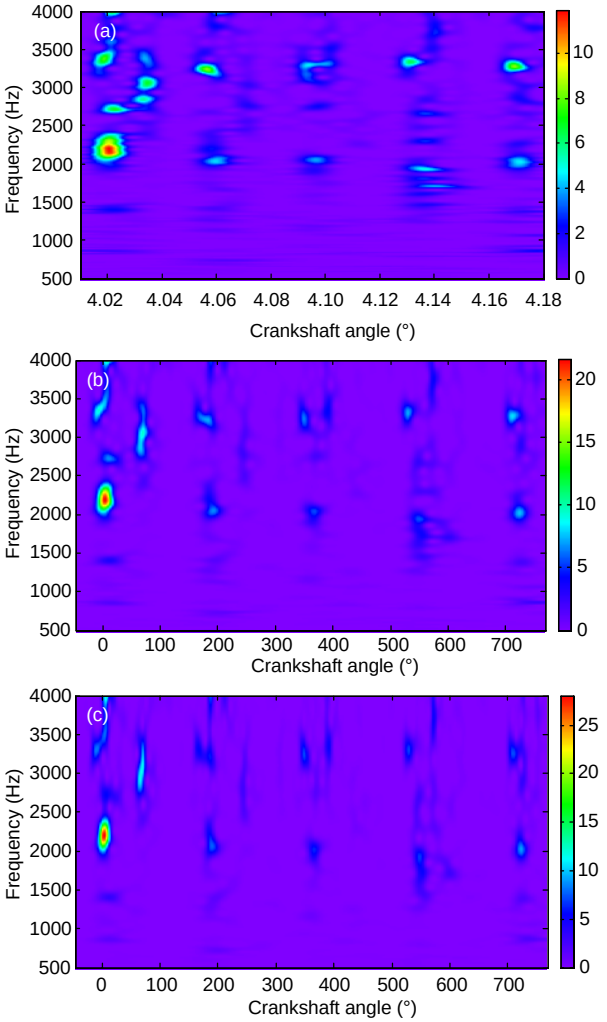

Fig. 12 A01's AWT^m spectrogram with (a) Q=0.01, (b) Q=0.02 and (c) Q=0.03

The color represents square of the AWT^m magnitude in linear mode

The results presented in Figs. 8–12 indicate that the uniform resolution is not necessarily a disadvantage for the joined time-frequency analysis of stationary IC engine vibration signals.

7 Conclusions

1. The validity and applications of the lowest distinguishable frequency limit $2.5\sigma_f$, the time resolution limit $3\sigma_t$, and the frequency resolution limit $1/(3\sigma_t)$ for an STFT spectrogram with Gaussian window were confirmed.

2. The vibration signals in an IC engine at a steady speed responding to transient excitations consist of continuous and intermittent parts. The continuous part is usually at low frequencies, while the intermittent part usually has high frequencies. A single window width by which both parts can be clearly distinguished may not exist. This can be circumvented by performing STFT using a few different window widths.

3. The spectrograms of three time-frequency transforms, STFT, ST, or AWT, are closely related to the window width (for STFT) or the quality factor (for ST or AWT). This needs to be borne in mind when interpreting the spectrogram to avoid mistaken or unconvincing interpretations. In addition, ST has no adjustable parameters, while the AWT spectrogram varies as a function of Q.

4. In a vibration signal from an IC engine at a steady speed, the appearance of signal components with an abrupt frequency change is very rare. Therefore, STFT is applicable. However, this does not mean that rational judgments are no longer needed. The resolution of STFT is limited, and artifacts always exist.

5. For cylinder head vibration signals, STFT spectrograms with a reasonable window width exhibit a better frequency localizing ability in time-frequency domain than either ST or AWT without greatly sacrificing time resolution. STFT may be useful in several applications including fault diagnosis, vibroacoustic source identification, and transfer path analysis.

References

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

Hao, Z., Xu, H., Zheng, G., Jing, G., 2008. Study on the Time-Frequency Characteristics of Engine Induction Noise in Acceleration Based on S Transform. Congress on Image and Signal Processing, Sanya, China. IEEE, Piscataway, USA, p.242-246.

Jing, G., Hao, Z., 2009. A new technique based on traditional wavelet transform used in NVH application of internal combustion engine. *Mechanical Systems and Signal Processing*, **23**(3):979-985. [doi:10.1016/j.ymssp.2008.08.009]

Peng, Z.K., Chu, F.L., 2004. Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. *Mechanical Systems and Signal Processing*, **18**(2):199-221. [doi:10.1016/S0888-3270(03)00075-X]

Qian, S., 2002. Introduction to Time-Frequency and Wavelet Transforms. Prentice Hall PTR, Upper Saddle River, NJ, p.41-42.

Stockwell, R.G., 1996. Localization of the complex spectrum: the S transform. *IEEE Transactions on Signal Processing*, **44**(4):998-1001. [doi:10.1109/78.492555]

Vulli, S., Dunne, J.F., Potenza, R., Richardson, D., King, P., 2009. Time-frequency analysis of single-point engine-block vibration measurements for multiple excitation-event identification. *Journal of Sound and Vibration*, **321**(3-5):1129-1143. [doi:10.1016/j.jsv.2008.10.011]

Wang, C.D., Zhang, Y.Y., Xia, Y., 2003. Fault diagnosis for diesel valve train based on S transform. *Transactions of CSICE*, **21**(04):271-275 (in Chinese).

Wang, Y., 2010. The Tutorial: S Transform. Graduate Institute of Communication Engineering, National Taiwan University, Taipei.

Wu, J., Chen, J., 2006. Continuous wavelet transform technique for fault signal diagnosis of internal combustion engines. *NDT & E International*, **39**(4):304-311. [doi:10.1016/j.ndteint.2005.09.002]

Yang, J., Hao, Z., 2006. Using A-weighting CWT to identify internal combustion engine sound source. *Journal of Zhejiang University (Engineering Science)*, **40**(07):1174-1177 (in Chinese).

Zhu, X.D., Kim, J., 2005. Application of the Analytic Wavelet Transform for Time-Frequency Analysis of Impulsive Sound Signals. SAE Noise and Vibration Conference and Exhibition, Grand Traverse, MI, USA. [doi:10.4271/2005-01-2391]