2019 3rd International Conference on Mechanical and Electronics Engineering (ICMEE 2019)

Research on Feature Extraction Method of Engine Acoustic Signal

Kai Chen[a], Yuxiang Wang[b]

Army Artillery Air Defense Academy, Hefei 230031, China.

[a]348211143@qq.com, [b]479259070@qq.com

Keywords: Engine, Acoustic signal characteristics, extraction methods

Abstract: For the most critical feature extraction problem in vehicle engine abnormal sound fault diagnosis, this paper selects a certain type of transport vehicle as the research object, and collects the audio signal of the engine in normal operation and lack of cylinder operation as a sample. Firstly, the original signal is preprocessed, including pre-emphasis, Sub-frame, plus windows. Then the short-time energy of the frame signal is obtained in the time domain. Finally, the power spectrum is calculated in the frequency domain, which is passed through the Mel filter bank, and the MFCC is obtained by DCT. The research results show that the time and frequency domain characteristic parameters obtained by this method can reflect the operating state information of the engine, and can effectively distinguish whether there is a lack of cylinder operation. It is suitable as the feature vector for vehicle engine fault diagnosis and prediction. This will lay a foundation for future model training and matching recognition.

1. Introduction

Since the Second World War, the pace of motorization and mechanization has accelerated rapidly, and the number and types of military vehicles have also increased dramatically. Many special vehicles with different functions and uses have appeared. Among the powerful weapon equipment such as military vehicles, the engine is one of its most important equipment. However, due to its complicated structure and poor working conditions, the failure rate is high. Failure to diagnose and perform maintenance guarantees in the event of a war will have serious consequences. Therefore, efficient and simple vehicle engine fault diagnosis technology is an urgent need in the field of weapon equipment maintenance support. The most critical part of fault diagnosis is the feature extraction of fault information. [1].

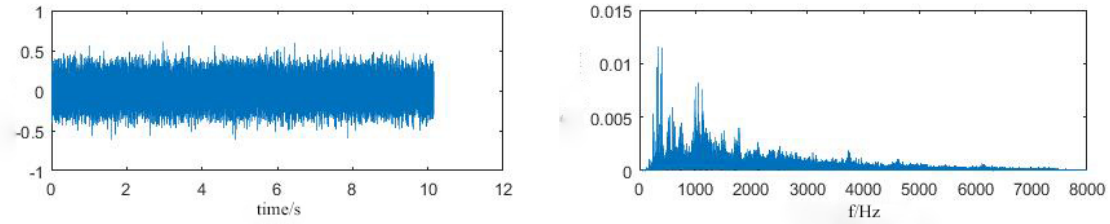
The near-field acoustic signal of the engine is rich in periodic information and pulse components, so it usually appears in such signals when it fails. Feature extraction for non-stationary signals such as engine running sounds is generally divided into two steps, first using non-stationary signal processing methods such as short-time Fourier transform (STFT), empirical mode decomposition (EMD), wavelet analysis, etc. Signal time-frequency decomposition, and then traditional methods including time domain statistical analysis, spectrum analysis, power spectrum analysis, cepstrum analysis, etc. for fault feature extraction [2].

In this paper, the time domain and frequency domain analysis methods are used to extract the parameters reflecting the fault characteristics in the engine acoustic signal, which provides the basis for fault diagnosis and prediction. The short-time energy of the acoustic signal is selected as the characteristic parameter in the time domain, and the abnormal noise signal of the engine is processed by the method of the frequency spectrum analysis in the frequency domain. The MFCC mainly reflects the static characteristics of the sound [3]. The inside of the engine is a mechanical system that reciprocates cyclically. Regular noise is generated during operation, static characteristics can be identified, and MFCC parameters have good recognition performance and noise immunity. Therefore, the MFCC simulating the human ear hearing mechanism is introduced into the fault diagnosis field of the engine abnormal sound signal as a characteristic parameter of the fault information.

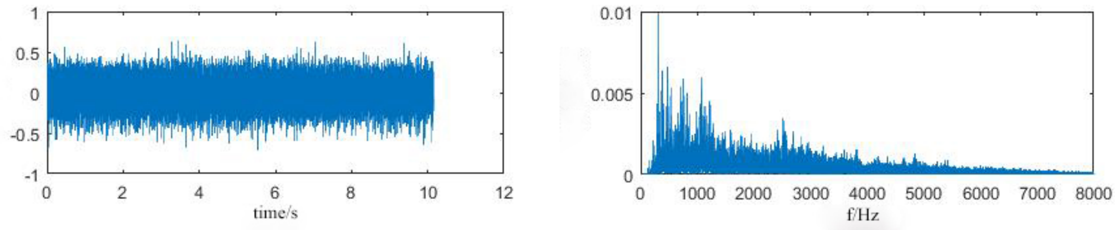
2. Acoustic signal acquisition and characterization

In this paper, the sound of a certain type of transport vehicle engine is selected as the sample. The vehicle is equipped with a six-cylinder and four-stroke diesel engine, which respectively collect the idle speed sound signal under normal engine condition and the idle speed sound signal when the two cylinders are not working. The SONY ICD-SX2000 stereo digital recording stick is used to collect the sound signal of the car engine. The audio format is .WAV, the sampling rate is 16 KHz, and the sampling time is 30 seconds. Intercept the middle segment for about 10 seconds as a sample analysis to avoid noise interference at the beginning and end. The engine is a mechanical system with cyclic reciprocating motion. The waveform of the acoustic signal also appears as a periodic cycle. Intercepting the signal analysis in the middle part can also reduce the amount of calculation and improve the feature extraction speed.

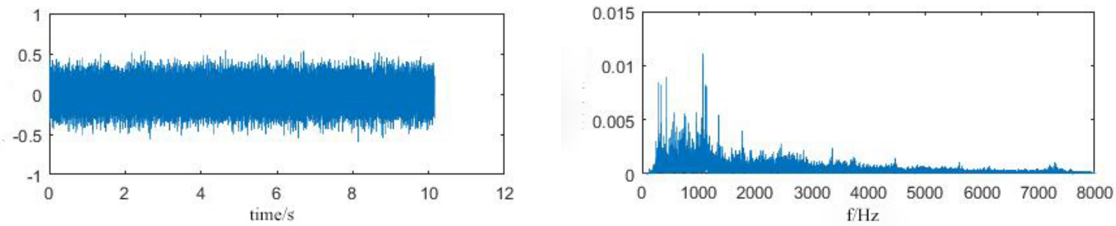
The waveform of the acoustic signal in the time domain is difficult to observe the signal characteristics. Therefore, it is generally converted to the frequency domain by fast Fourier transform, and the signal characteristics are analyzed by observing the amplitude-frequency distribution map. The time-frequency domain analysis of the collected audio signal, the time domain waveform of the normal operation sound signal of the engine and its spectrum diagram are shown in Fig. 1. In the figure, a, b and c are three vehicles of the same type. When the engine is running normally, the time domain waveform of the acoustic signal is stable, and the frequency is mainly concentrated in the low frequency region, and the highest is about 8 KHz. The engine fault is set artificially, so that the two cylinders are not working when the engine is running, and the time domain waveform and spectrum map corresponding to the running acoustic signal are shown in Fig. 2. In contrast, the amplitude range of the acoustic signal in the time domain of the acoustic signal is larger and more unstable. Although the spectrum map is different but not obvious, further frequency domain analysis is needed to extract the feature parameters.



(a)

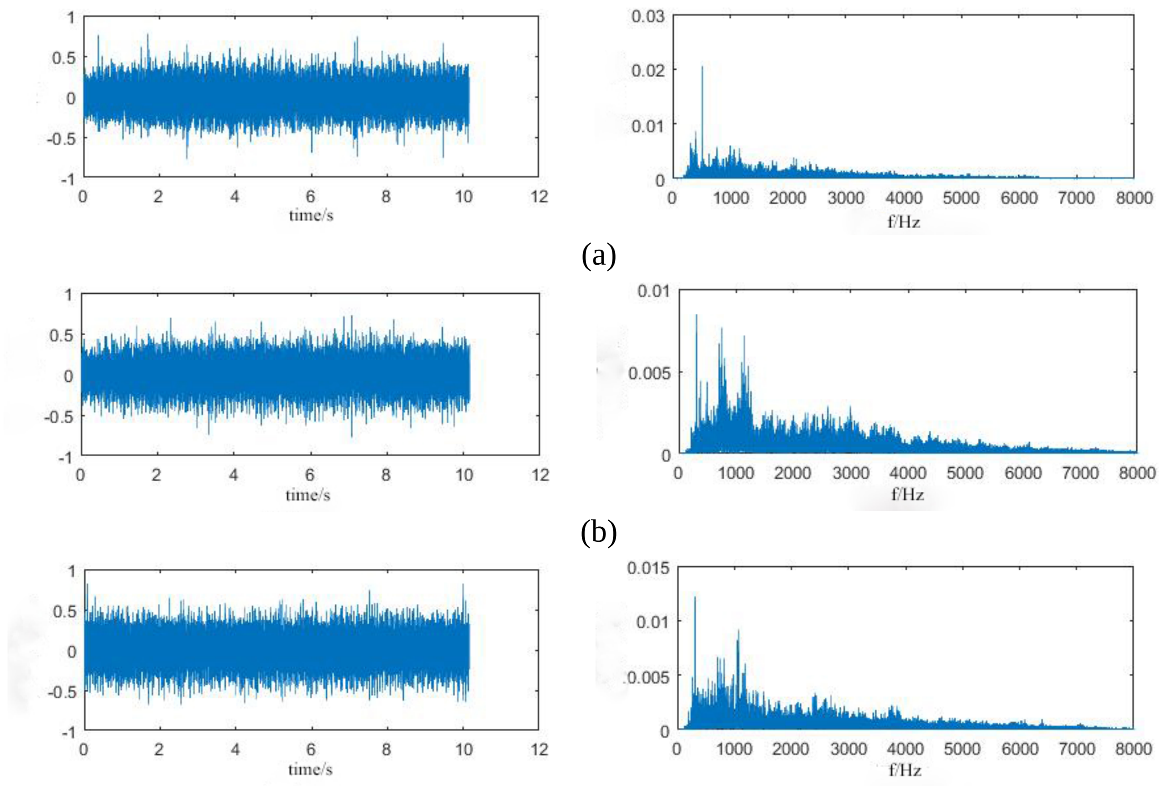


(b)



(c)

Figure 1. Audio signal waveform and spectrum of engine idle running.



(c)

Figure 2. Audio signal waveform and spectrum of engine idle running without cylinder.

3. feature parameter extraction

The main purpose of feature extraction is to simplify recognition by summing up a large amount of audio data without losing the acoustic properties of the original audio. The sound signal of the automobile engine contains rich sound source information, which can reflect it's abnormal or fault state. The audio characteristic parameter extraction obtains a feature that can describe the engine running state from the original audio signal, and lays down the pattern matching and fault identification. Basis. Before the feature extraction, the original acoustic signal needs to be preprocessed for analysis.

3.1 Preprocessing

Assuming the input audio signal is x (n), the preprocessing process is as follows.

(1) Pre-emphasis. Under normal circumstances, the energy of the sound signal will exponentially decay with the increase of its frequency, so that the intensity of the low frequency signal is greater than the high frequency signal, affecting the subsequent analysis and processing. The pre-emphasis process essentially passes the sound signal through a high-pass filter to amplify the high-frequency signal, making the spectrum of the signal flatter throughout the frequency band to facilitate feature extraction. The pre-emphasis process is usually implemented with a first-order digital filter with 6dB/octave, as shown in equation (1):

$$H(Z) = 1 - \mu z^{-1} \tag{1}$$

Where μ is a constant, generally 0.97.

(2) Framing and windowing. Since the engine belongs to the cyclic motion device, the physical characteristic parameters and spectral characteristics of the running noise for a short period of time can be considered to remain basically the same, or slow, that is, the short-term stationary characteristics of the sound signal [4]. Using this feature, the audio signals are overlapped and framed, and a small segment is intercepted for analysis. Take 512 samples of data as one frame, and the

corresponding time length is 32ms. In order to avoid signal discontinuity, there is a certain overlap between every two adjacent frames, and 256 points of data are overlapped between frames. In order to reduce the discontinuity of the signals at both ends of the frame caused by the Gibbs phenomenon, windowing processing is required, and a Hamming window is used to multiply each frame signal. The Hamming window is as follows:

$$W(n,a) = (1-a) - a \times cos\frac{2\pi n}{N-1}\ ,\ 0 \le n \le N-1 \tag{2}$$

Where a is a constant, usually 0.46.N is the frame length.

3.2 Short-time energy

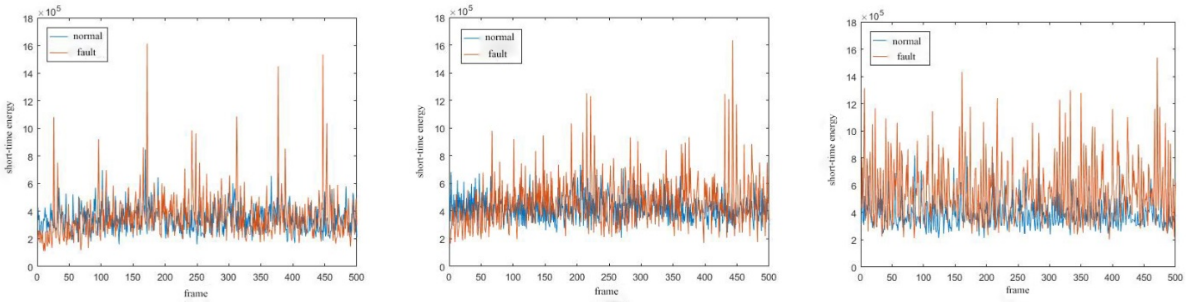
The signal characteristics of each frame of the audio signal after preprocessing are basically unchanged, and can be regarded as a whole for analysis and processing. The total energy of one frame signal is the short-term energy of the signal [5]. The short-time energy of the audio signal represents the level of the volume. When the engine fails, there is often a lot of noise. By comparing the short-time energy of the acoustic signal during normal operation with the short-time energy of the abnormal signal when the fault occurs, the engine state can be judged. The analysis of the short-time energy of the signal can reflect the change of the energy of the acoustic signal in different states in the time domain, and can effectively reduce the influence of ambient noise, and has intuitiveness and good anti-noise.

The pre-processed audio signal is$x(\tau)$, and the short-time energy is calculated as follows:

$$E(n) = \sum_{\tau=0}^{N-1} |x(\tau)|^2 \tag{3}$$

$E(n)$ Is the short-term energy of a single frame, and N is the frame length. In this paper, 512 data points are taken as one frame.

The short-time energy of the sound signal in the normal operation and the cylinder-less operation state of the engine is calculated separately, and the relationship between the short-time energy and the number of frames is made.



(a) (b) (c)

Figure 3. Vehicle engine running audio signal short-time energy.

As shown in FIG. 3, the short-term energy of the audio signal of the vehicle engine is partially manifested as the occurrence of multiple sharp points in some frame signals when the cylinder is running, which is 2~3 times higher than the short-term energy of the signal when the cylinder is running. On the whole, it is observed that the short-time energy of the signal when the cylinder is missing is higher than that of the signal when the cylinder is missing. The samples collected in this experiment all meet this conclusion after analysis, so the short-term energy of the audio signal of vehicle engine running can be taken as one of the characteristics reflecting the fault information.

3.3 MFCC

MFCC is the most commonly used characteristic parameter in speech recognition and speaker recognition, and the initial method of abnormal sound signal diagnosis is to use the sound heard by the human ear to judge the fault, and experienced professionals can run through the engine. The sound judges the state of the machine. Considering the different sensitivity of the human auditory system in response to different frequency signals, the MFCC first maps the linear spectrum to the Mel nonlinear spectrum based on the human ear frequency perception, and then converts it to the cepstrum to extract the cepstrum parameters. The MFCC parameters reflect the characteristics of the audio short-term amplitude spectrum, and have good recognition performance and anti-noise ability, so they are widely used in speech recognition. MFCC can reflect the difference between human voice signal in time and frequency domain. If the engine of the vehicle is regarded as "vocal cord" and the exhaust tract is regarded as "channel", the vehicle noise signal can be regarded as a kind of "vehicle voice"[6]. Therefore, the MFCC method in the field of speech signal recognition can be transplanted into vehicle noise signal analysis for feature extraction. The extraction process of the MFCC parameters is as follows.

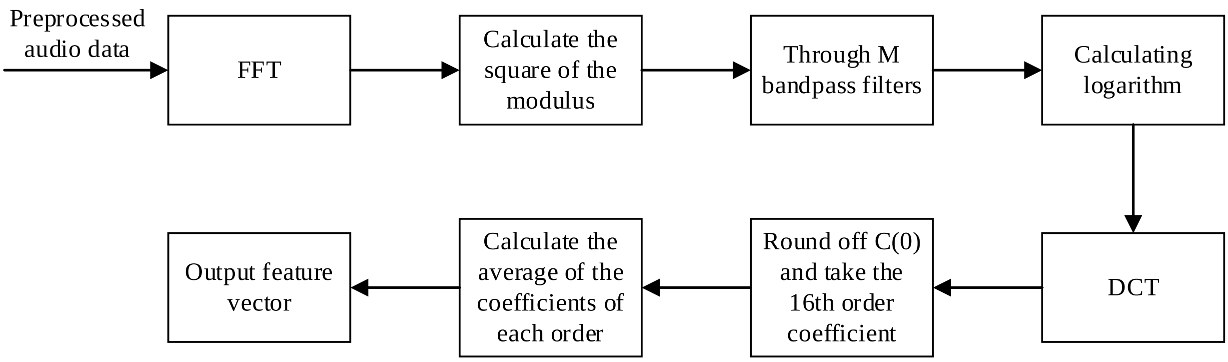


Figure 4. Vehicle engine running audio signal short-time energy.

The specific calculation process is as follows:

(1) FFT

Since the characteristics of the signal are not easily seen in the time domain, the pre-processed analysis frame signal is usually subjected to FFT to obtain the spectrum of each frame, and the spectral modulo square is calculated to obtain the discrete power spectrum of the signal.

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi k/N}, 0 \leq k \leq N \tag{4}$$

$$P(k) = |X(k)|^2 \tag{5}$$

Where N represents the number of points of the FFT.

(2) Mel filter bank

The power spectrum signal is input to the Meyer filter bank, and the lowest frequency is defined as zero, and the highest frequency is half of the signal sampling frequency. The Mel filter bank divides the signal frequency domain into a series of triangular filter sequences to simulate the logarithmic relationship and masking effect similar to the human ear perception [7]. In the spectrum, the number of low-band filters is large, and the high-band filter the number is small; its central frequency domain is linearly distributed in the Mel frequency domain. The conversion relationship between the actual frequency and the Mel frequency is:

$$M(f) = 1125ln(1 + f/700) \tag{6}$$

$$f(M) = 700(e^{\frac{M}{1125}} - 1) \tag{7}$$

In this experiment, 40 Mel band-pass filters are used to filter the power spectrum signal. The transfer function of the band-pass filter can be expressed as:

$$H_m(k) = \begin{cases} 0 & , k < f(m-1) \\ \frac{2(k-f(m-1))}{(f(m+1)-f(m-1))(f(m)-f(m-1))}, & f(m-1) \leq k \leq f(m) \\ \frac{2(f(m+1)-k)}{(f(m+1)-f(m-1))(f(m)-f(m-1))}, & f(m) \leq k \leq f(m+1) \\ 0 & , k > f(m+1) \end{cases} \tag{8}$$

$$\sum_{m=0}^{M-1} H_m(k) = 1 \quad (9)$$

The logarithmic energy of the power spectrum $P(k)$ output through the Mel filter bank is as follows:

$$S(m) = ln(\sum_{k=0}^{N-1} P(k)H_m(k)) , 0 \leq m \leq M \quad (10)$$

(3) DCT

The MFCC coefficients are obtained by performing DCT on the vector of the filter output.

$$C(n) = \sum_{m=0}^{N-1} S(m) \cos\left(\frac{\pi n(m-0.5)}{M}\right), n = 0,1,2, \dots, L \quad (11)$$

In the above formula, $C(n)$ is the nth MFCC coefficient, $S(m)$ is the filter output logarithmic energy Mel spectrum, and L is the MFCC coefficient order, where L is 16. Then, each stage can obtain 17-order MFCC, wherein the 0-order coefficient obtained at n=0 reflects the spectral energy, and its energy is large and represents the DC component, which is generally not used as a characteristic parameter, so the i-th frame signal The 16 MFCC feature values can form a feature vector K_{0I}:

$$K_{0I} = [k_{01}, k_{02}, \dots, k_{015}, k_{016}]$$

A schematic diagram of the MFCC characteristic values of the normal and missing cylinder running audio signals of the vehicle engine is shown in FIG. It can be seen from the figure that the MFCC characteristic value difference of the audio signal in the two operating states of the engine is mainly concentrated between the third to eighth-order coefficients, and the sound characteristics can distinguish the sounds of normal operation and lack of cylinder operation. Therefore, the MFCC coefficient of the engine acoustic signal can be used as one of the characteristic parameters of the final fault diagnosis and prediction.

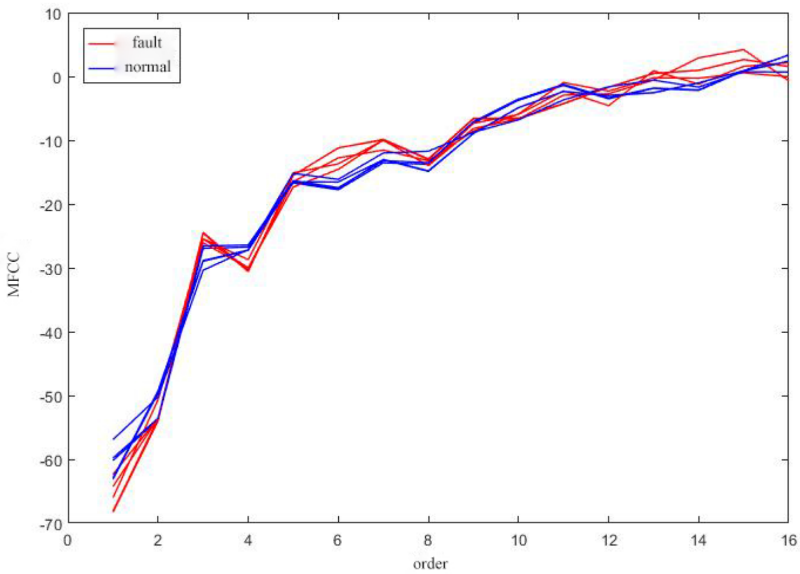


Figure 5. MFCC characteristic value of vehicle engine running sound signal.

4. Conclusion

After the pre-processing of the vehicle engine audio signal, the short-term energy characteristics obtained in the time domain range and the MFCC coefficients extracted in the frequency domain range can be used as characteristic parameters reflecting the engine operating state, and used to distinguish whether the engine is out of cylinder operation. The two types of feature parameters are independent of each other, and a single feature can be combined into a hybrid feature, thereby improving the accuracy of the recognition match. A method for feature extraction is provided for analyzing other fault types, which lays a foundation for further identifying and matching different faults, and finally achieves the purpose of fault diagnosis and prediction.

References

[1] Liu M Q, Bing-Bing L I, Guo W L. Discussion on power spectrum and estimation of random signal analysis course [J]. Journal of Electrical & Electronic Education, 2016.

[2] Fault diagnosis of stamping process based on empirical mode decomposition and learning vector quantization [J]. A.M. Bassiuny,Xiaoli Li,R. Du. International Journal of Machine Tools and Manufacture. 2007 (15)

[3] Combining dynamic texture and structural features for speaker identification. Zhao G, Huang X, Gizatdinova Y,et al. ACM Workshop on Multimedia in Forensics,Security and Intelligence . 2010

[4] Time-frequency analysis for efficient fault diagnosis and failure prognosis for interior permanent-magnet AC motors. Strangas, Elias G, Aviyente, Selin, Zaidi, Syed Sajjad H. IEEE Transactions on Industrial Electronics. 2008

[5] Feature Extraction and Recognition for Rolling Element Bearing Fault Utilizing Short-Time Fourier Transform and Non-negative Matrix Factorization [J]. GAO Huizhong, LIANG Lin, CHEN Xiaoguang, XU Guanghua. Chinese Journal of Mechanical Engineering. 2015(01)

[6] Combining vocal source and MFCC features for enhanced speaker recognition performance using GMMs. HOSSEINZADEH D, KRISHNAM S. IEEE9th Workshop on Multimedia Signal Processing. 2007

[7] Hossan M A, Memon S, Gregory M A. A novel approach for MFCC feature extraction [C]// Signal Processing and Communication Systems (ICSPCS), 2010 4th International Conference on. IEEE, 2011.