

# Research on Signal Extraction and Classification for Ship Sound Signal Recognition

Shuai Fang<sup>1</sup>, Jianhui Cui<sup>1</sup>, Ling Yang<sup>1</sup>, Fanbin Meng<sup>2</sup>, Huawei Xie<sup>2</sup>, Chunyan Hou<sup>2</sup> and Bin Li<sup>2</sup>

Received: 20 June 2023 / Accepted: 03 November 2023

© Harbin Engineering University and Springer-Verlag GmbH Germany, part of Springer Nature 2024

## Abstract

The movements and intentions of other ships can be determined by gathering and examining ship sound signals. The extraction and analysis of ship sound signals fundamentally support the autonomous navigation of intelligent ships. Mel scale frequency cepstral coefficient (MFCC) feature parameters are improved and optimized to form NewMFCC by introducing second-order difference and wavelet packet decomposition transformation methods in this paper. Transforming sound signals into a feature vector that fully describes the dynamic characteristics of ship sound signals and the high- and low-frequency information solves the problem of the inability to transport ordinary sound signals directly as signals for training in machine learning models. Radial basis function kernels are used to conduct support vector machine classifier simulation experiments. Five types of sound signals, namely, one type of ship sound signals and four types of interference sound signals, are categorized and identified as classification targets to verify the feasibility of the classification of ship sound signals and interference signals. The proposed method improves classification accuracy by approximately 15%.

**Keywords** Ship signal identification; Signal extraction; Automatic classification; Intelligent ships; Support vector machine

## 1 Introduction

Recent years have witnessed the accelerated digitalization and intelligent transformation of the shipbuilding industry owing to the widespread application of advanced technologies such as artificial intelligence and the Internet of Things (Lu et al., 2023). Ships must emit sound whistles when entering or leaving ports or when navigating in foggy conditions. The sound whistles act as a mode of communication between vessels and ports and as a warning signal.

Ship sound signals are special sound signals emitted by ships to indicate their navigational intentions or to call the attention of other vessels. Although autonomous ships substantially reduce accidents caused by human factors, uncertainty remains in their navigation safety (Wróbel et al., 2017). If intelligent ships can autonomously recognize ship sound signals and interact with traditional ships, they can achieve fully autonomous ship navigation.

A sound signal is complex and cannot be directly used as input sent to a machine learning model for training and classification even after preprocessing and denoising; otherwise, the model will have high computational difficulty and poor performance (Bin et al., 2009). To solve this problem, features must be extracted from various types of sounds, and the sound signal must be converted into a simple but more distinctive representation of the actual signal.

The extraction of MFCC feature parameters, building on the characteristics of MFCCs, is proposed in this paper. The extensive research on using MFCC for sound signal processing includes numerous studies, such as heartbeat recognition and Chinese language recognition (Deng et al., 2020; Yan et al., 2011; Tomchuk and Ieee, 2018). However, the traditional MFCC method has limitations, which are outlined below:

1) Limited dynamic range: traditional MFCC computation entails taking logarithms of the magnitude spectrum, which can restrict the dynamic range of the resulting features and might lead to the loss of fine-grained informa-

## Article Highlights

- Dynamic time warping and wavelet packet decomposition methods are introduced to improve the Mel scale frequency cepstral coefficient.
- The temporal and spectral characteristics of sound are combined to construct a 42-dimensional sound feature to enhance sound recognition accuracy.
- A ship siren recognition and classification method based on a support vector machine is proposed, and its classification performance is verified. The experimental verification obtained well results.
- This provides a theoretical reference for intelligent ships to autonomously recognize ship sound signals, interact with traditional ships, and achieve fully autonomous navigation.

✉ Jianhui Cui  
jianhuicui@email.tjut.edu.cn

<sup>1</sup> Maritime College, Tianjin University of Technology, Tianjin 300384, China

<sup>2</sup> Tianjin Navigation Instrument Research Institute, Tianjin 300131, China

tion, particularly in scenarios with a wide range of signal amplitudes.

2) Fixed window length: MFCC calculation involves fixed-length analysis windows, which may not satisfactorily capture the temporal dynamics of sound signals with varying time scales. This limitation could affect the ability to capture rapid changes in sound characteristics.

3) Lack of frequency resolution: traditional MFCCs use a linear frequency scale, which may not be ideal for capturing the nonlinear frequency perception of the human auditory system and could limit the representation of frequency-related information.

To address these shortcomings, predecessors have attempted numerous improvements in various application areas. Zheng and Zhang found that the frequency band energy can improve the performance of MFCC (Zheng et al., 2001). Zhang applied bark wavelet to MFCC and used it for preprocessing before Fast Fourier Transform (FFT) (Zhang et al., 2006). Shi proposed smoothing MFCC, which aims to improve the MFCC algorithm based on smoothing short-term spectral amplitude envelope (Shi and Wang, 2011). Bhattacharjee and KshirodSarmah investigated why the performance of a baseline gaussian mixed model-mel-scale frequency cepstral coefficient (GMM–MFCC)-based language identification system improves remarkably when prosodic features are considered additional features with MFCC (Bhattacharjee et al., 2013). Lv proposed combining MFCC features with song channel features using an auditory model. They extracted three features reflecting the properties of the vocal tract: the all poles of vocal tract system (LPCAS), the spectrum plopsestrum of the vocal tract system (plpsestrum), and perceptual linear predictive (PLP) of the vocal tract system cepstrum. In an experiment, they mixed these three channel features with MFCC (Lv et al., 2020). Studies on similar improvements are many, but their application in ship sound signal recognition still cannot address the traditional shortcomings (Maurya et al., 2017; Mahesha et al., 2017; Amelia et al., 2019; Bhattacharjee et al., 2013; Xie et al., 2021; Wang and Hu, 2018). To alleviate these drawbacks, the extraction steps of MFCC (Jiang et al., 2017) and feature parameters based on the characteristics of MFCCs in ship sound signal extraction are presented in this paper. Furthermore, the MFCC feature parameters are improved and optimized by introducing second-order difference and wavelet packet decomposition transformations. These enhancements result in the development of NewMFCC, which effectively represents the dynamic characteristics and high-/low-frequency information of ship sound signals. Additionally, audio features are constructed from the temporal and spectral perspectives. Seven statistical measures, namely, maximum, minimum, mean, standard deviation, root mean square, skewness, and kurtosis, are computed for each feature. The corresponding feature vectors generated are then normalized. Moreover, a control group experiment is conducted to com-

pare NewMFCC with MFCC. The proposed approach improves the ability to extract dynamic and static and high- and low-frequency information from sound signals, enabling a more comprehensive representation of the information content in sound signals and laying a foundation for subsequent processing tasks.

In ship sound signal classification, several machine learning algorithms that can be used in the field of sound recognition are comprehensively compared. The support vector machine (SVM) (Cortes and Vapnik, 1995) algorithm is selected due to its advantages of requiring fewer training samples, satisfactory generalization performance, and high prediction accuracy (Zhang and Ma, 2012; Tuncer and Aydemir, 2020). The feasibility of using SVM with a radial basis function (RBF) kernel to classify ship sound signals and interference signals is verified through simulation experiments. Experimental designs are conducted using five types of sound signals, namely, one type of ship sound signals and four types of interference signals, as classification targets. Then, 600 audio segments are selected for training based on the sound characteristics of each category. The system is trained to distinguish ship sound signals automatically. Finally, the experimental results are obtained and compared. The results confirm the feasibility of the ship sound signal recognition system based on SVMs for classifying ship sound signals and interference signals, and classification accuracy improves by approximately 15%.

## 2 Ship sound information extraction and analysis

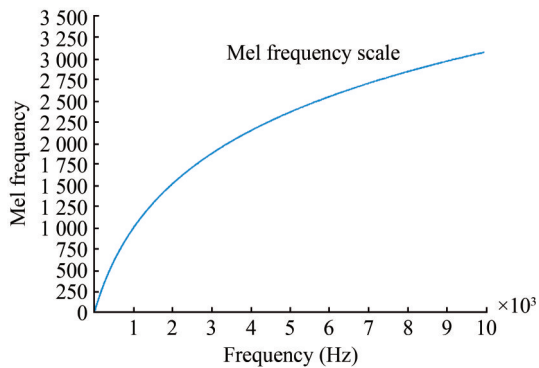
### 2.1 MFCC overview

The human auditory system is a special nonlinear system possessing different auditory sensitivities for various frequencies of sound waves (Geoffrey and Collie, 2004). According to the latest automatic recognition technology, a large gap remains between its recognition ability and that of the human auditory system (Wang et al., 2023). If the frequency perceived by the human ear and the actual frequency of the sound wave are below 1 000 Hz, the relationship between them is linear. When both are above 1 000 Hz, the relationship between them is logarithmic (Yang and Chen, 2020). Currently, the most commonly used technique in sound recognition is MFCC, which can fully simulate the human auditory mechanism (Lin et al., 2006). The Mel scale describes the nonlinear frequency characteristics of the human ear, and its relationship can be approximated by Equation (1).

$$\text{Mel}(f) = 2595 \times \lg\left(1 + \frac{f}{700}\right) \quad (1)$$

where  $f$  represents the frequency received by the human ear.

Figure 1 exhibits that when the frequency is small, the Mel scale changes rapidly with the frequency; by contrast, when the frequency is large, the Mel scale increases slowly, and the slope of the curve is small. This outcome indicates the human ear is sensitive to low-frequency tones but not to high-frequency tones.



**Figure 1** The relationship curve between human ear receiving frequency and Mel scale

Because low-frequency sounds are transmitted further in the basilar membrane of the cochlea than high-frequency sounds, low-frequency sounds are more likely to mask high-frequency sounds, while masking low-frequency sounds with high-frequency sounds is more difficult. The critical bandwidth of masking in the low-frequency range is smaller than that in the high-frequency range, which inspired the distribution of Mel filter banks (Chu et al., 2022). This behavior corresponds to the objective rule that the higher the frequency is, the duller the human ear becomes (Nassim and Abderahmane, 2017).

## 2.2 MFCC feature parameter extraction involving several steps

MFCC considers the characteristics of human hearing. First, MFCC maps the linear spectrum to the Mel nonlinear spectrum based on auditory perception. Then, MFCC converts the Mel nonlinear spectrum into cepstrum. This feature neither relies on the nature of the signal nor makes any assumption or restriction on the input sound signal. Compared with other models, the MFCC model utilizes the research achievements of the auditory model and exhibits better robustness, noise resistance, and recognition performance.

The specific steps of MFCC feature parameter extraction are described below (Wang and Zhao, 2020).

### 1) Preprocessing

Sound signal  $x(n)$  undergoes preprocessing, namely, pre-emphasis, framing, and Hamming windowing  $x_i(m)$ , to obtain (Xie et al., 2010), where  $i$  denotes the  $i$ th frame after framing.

### 2) FFT

Sound signals must be transformed from the time domain

to the energy distribution in the frequency domain to facilitate observation. Different features of sounds can be represented by dissimilar energy distributions. After the signal is divided into frames and multiplied by the Hamming window, each frame must undergo another FFT to obtain the energy distribution of the entire spectrum, as shown in Equation (2).

$$X(i, k) = \text{FFT}[x_i(m)] \quad (2)$$

### 3) Computing spectral line energy

Equation (3) shows that for the spectrum of each frame of the sound signal obtained after FFT, the spectral energy is obtained by taking the square of its absolute value.

$$E(i, k) = |X(i, k)|^2 \quad (3)$$

### 4) Translation: calculating the energy passed through Mel filters

The resolution of high-frequency components in the Mel filter banks is lower, whereas that of low-frequency components is higher. Therefore, only certain frequency components are allowed to pass through, and the amplitude of high-frequency information is attenuated. The energy of each frame's power spectrum is passed through the Mel filter banks, and the energy in the Mel filter banks is quantified. In the frequency domain, this process is equivalent to multiplying and adding the power spectrum  $E(i, k)$  of each frame with the frequency response  $H_m(k)$  of the Mel filter banks. Equation (4) is shown below.

$$S(i, m) = \sum_{k=0}^{N-1} E(i, k) H_m(k), \quad 0 \leq m \leq M \quad (4)$$

where  $i$  is the  $i$ th frame of the signal, and  $k$  is the  $k$ th spectral line in the frequency domain.

### 5) Logarithmic operation

After the sound signal is transformed by the spectrogram FFT, convolution becomes multiplication. Logarithmic operations can turn it into addition; that is, FFT and logarithm operations turn the convolution signal into an additive signal.

### 6) Computation of discrete cosine transform (DCT) cepstrum

Different Mel filters have intersections, so they are correlated. To improve recognition accuracy, the correlation between signals in different dimensions can be removed by using DCT, which maps the signal to a lower-dimensional space. Equation (5) is shown below.

$$\text{mfcc}(i, n) = \sqrt{\frac{2}{M}} \sum_{m=0}^{M-1} \log[S(i, m)] \cos\left[\frac{\pi n(2m-1)}{2M}\right] \quad (5)$$

where  $m$  is the  $m$ th filter ( $M$  filters in total), and  $n$  is the spectral line after DCT.

After the random selection of the audio section following the above steps, MFCC can be obtained. The first several dimensions of MFCC possess greater discriminative power

in distinguishing sounds, and increasing the number of dimensions does not necessarily improve the recognition results. To reduce unnecessary computational complexity, the first 12 coefficients are taken as the 12-dimensional MFCC feature parameters.

The MFCC feature parameter extraction is described in detail in Figure 2.

### 3 MFCC feature parameter optimization

#### 3.1 $\Delta$ MFCC based on second-order difference

MFCC is a satisfactory representation of the characteristics of sound, but it only describes the energy spectrum of a single frame of the sound signal, which is a static feature. The extraction of MFCC feature parameters segments the sound signal into frames of 10–30 millisecond to meet the requirement of short-term stationarity. However, ship acoustic signals are nonstationary and require improvement of the feature parameters to adopt dynamic characteristics.

Performing first-order differentiation on MFCC, which is the difference between two adjacent frames, can reflect the relationship between consecutive frames of the audio signal. Second-order differentiation is the relationship between the first-order differentiations of the preceding and subsequent frames, reflecting the dynamic relationship between three adjacent frames of the audio signal. Differential parameters can be calculated using Equation (6).

$$d_t = \begin{cases} C_{t+1} - C_t, & t < K \\ \frac{\sum_{k=1}^K K(C_{t+k} - C_{t-k})}{\sqrt{2 \sum_{k=1}^K K^2}}, & \text{others} \\ C_t - C_{t-1}, & t > Q - K \end{cases} \quad (6)$$

where  $d_t$  is the  $t$ th first-order differentiation,  $C_t$  is the  $t$ th cepstral coefficient,  $Q$  is the order of cepstral coefficients,

and  $K$  is the time difference of the first-order differentiation, which can be 1 or 2. The result of the above equation can be further used to compute the second-order differential MFCC feature parameters, which are denoted as  $\Delta$ MFCC (Chomorlig et al., 2012).

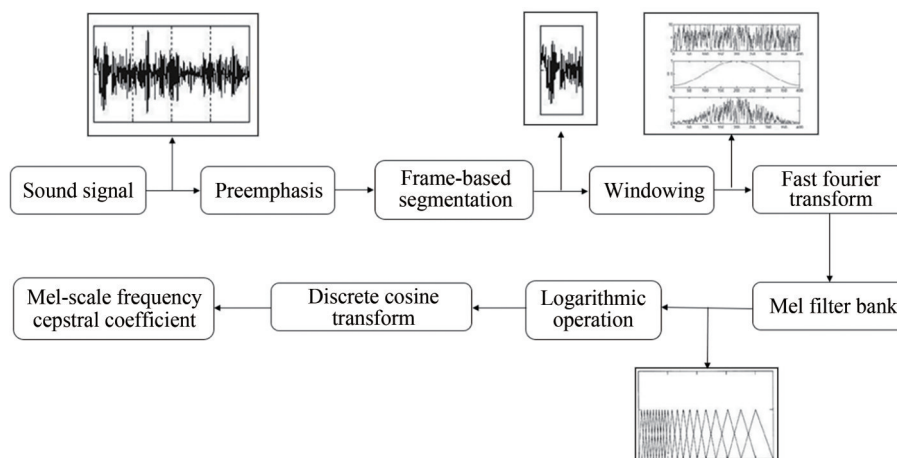
$\Delta$ MFCC can capture the continuous dynamic changes of sound features. Thus, combining the feature parameters of second-order differential  $\Delta$ MFCC can address the disadvantage that MFCC is unable to represent dynamic characteristics.

#### 3.2 WPMFCC based on wavelet packet decomposition transformation

In the previous section, the signal analysis method used in extracting MFCC feature parameters is Fourier transform, but the window function added to each frame of the audio signal is the same. After the signal is transformed by Fourier transform, it contains all frequency band information, which can neither reflect the temporal information in the time domain nor describe the local characteristics of the audio signal in the time and frequency domains. In practical engineering applications, given nonstationary ship sound signals, the extraction of MFCC feature parameters can be improved further.

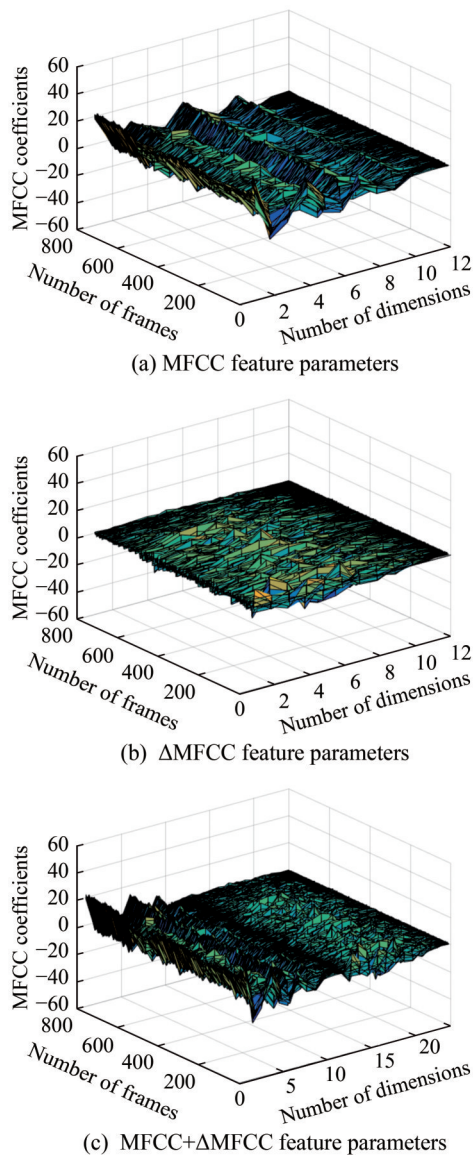
Wavelet transform is a time–frequency localization analysis method developed to decompose nonstationary signals. Wavelet transform uses fixed-sized window functions that can change in shape in time and frequency in the extraction. Wavelet transform can perform time–frequency localization analysis on the signal, but it can only further decompose its low-frequency part; the high-frequency part, that is, the detail part of the signal, is no longer decomposed; consequently, the information in the low-frequency region gradually decreases as the frequency increases, and the frequency domain of the signal cannot be uniformly distributed (Yu et al., 2019).

Ship sound signal recognition is performed in complex water environments where noise signals must be considered.



**Figure 2** MFCC feature parameter extraction



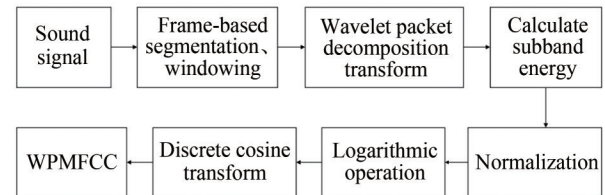


**Figure 3** MFCC, ΔMFCC, and MFCC+ΔMFCC feature parameter 3D plot

Hence, more mature wavelet packet decomposition methods can be introduced based on wavelet transform to improve the frequency resolution of the signal and the performance of feature parameters in complex noise environments while extracting the feature parameters.

The wavelet packet decomposition transform method uses high- and low-pass filters to screen sound signals separately. The wavelet packet decomposition transform method can simultaneously decompose low- and high-frequency components and achieve the frequency-domain uniformity of sound signals. MFCC feature parameters based on wavelet packet decomposition transform are called WPMFCC (Srivastava et al., 2012). The specific extraction is shown in Figure 4.

In this paper, the wavelet packet decomposition is performed according to the Mel scale, which not only simu-



**Figure 4** Flowchart of WPMFCC based on wavelet packet decomposition

lates the human auditory mechanism but also avoids excessive decomposition layers that generate  $2^n$  nodes and considerably increase computational complexity. Thus, the training efficiency of the model is influenced. The frequency  $f(j)$  of the  $j$ th sub-band of the wavelet packet decomposition after filtering by the Mel filter bank can be expressed as Equation (7).

$$\text{Mel}(f) = 2595 \times \lg \left( 1 + \frac{f(j)}{700} \right) \quad (7)$$

The energy of each frequency band is given by Equation (8).

$$W_j = \sum_i^{N_j} [W(j, i)]^2 \quad (8)$$

where  $i = 1, 2, \dots, N_j$  is the total number of wavelet coefficients in the  $j$ th sub-band. The normalization of the formula derives Equation (9).

$$S_j = \frac{W_j}{N_j} \quad (9)$$

After normalization, the energy of each sub-band is logarithmically transformed and subjected to DCT to obtain WPMFCC, as shown in Equation (10).

$$\text{WPMFCC}(n) = \sum_{j=1}^M \lg S_j \cos \left( \frac{\pi n (2j - 1)}{2M} \right) \quad (10)$$

where  $n$  is the dimension of WPMFCC based on wavelet packet decomposition.

### 3.3 NewMFCC feature parameters

To represent the characteristics of sound fully and achieve excellent classification and recognition accuracy, an improved method for extracting MFCC feature parameters is presented in this paper. After extracting the 12-dimensional static features of the sound signal using MFCC, they are combined with the dynamic features of ΔMFCC obtained by second-order difference and the WPMFCC feature parameters, including high and low frequencies, improved by wavelet packet decomposition. A 36-dimensional feature parameter called NewMFCC is formed.

The NewMFCC feature parameters consider the dynamic and static information of the audio signal at the same time, which compensates for the shortcomings of traditional MFCC. Moreover, the NewMFCC feature parameters regard the high- and low-frequency information of the signal, which can fully represent the information of the audio signal.

## 4 Sound temporal and spectral analysis and feature extraction

### 4.1 Temporal analysis

Temporal analysis describes the relationship between a mathematical function or physical signal and time, which is more intuitive and visual. This paper assumes signal  $x(n)$  is divided into frames as  $y_i(n)$ , where  $L$  is the frame length, and  $fn$  is the total number of frames after division. Several types of temporal characteristics of sound are classified and introduced in the following section (Fan, 2017).

#### 1) Short-term energy

Short-term energy reflects the intensity of the audio signal at various times. Energy cannot be calculated as a whole, and the energy of each frame must be determined frame by frame. The short-term energy of the  $i$ th frame of the audio signal can be expressed as Equation (11).

$$E(i) = \sum_{n=0}^{N-1} y_i^2(n), \quad 1 \leq i \leq fn \quad (11)$$

#### 2) Short-term average amplitude

Short-term average magnitude is the energy level of the sound signal in a frame. Using short-term energy and short-term average magnitude facilitates distinguishing voiced segments from unvoiced segments, as shown in Equation (12).

$$M(i) = \sum_{n=0}^{N-1} |y_i(n)|, \quad 1 \leq i \leq fn \quad (12)$$

#### 3) Short-time zero crossing rate

Short-time zero crossing rate (ZCR) is the number of times a speech signal waveform crosses the horizontal axis (zero level), that is, the number of times the sample values change their sign. Short-time average ZCR can be used to identify speech signals from background noise and to determine the start and end positions of silent segments and speech segments, as defined in Equation (13).

$$Z(i) = \frac{1}{2} \sum_{n=0}^{N-1} |\text{sgn}[y_i(n)] - \text{sgn}[y_i(n-1)]|, \quad 1 \leq i \leq fn \quad (13)$$

#### 4) Short-term autocorrelation function

The autocorrelation function can measure the similarity of the time waveform of the signal itself. The short-term

autocorrelation function is used in signal processing because audio signals possess nonstationary characteristics. The short-term autocorrelation function is obtained by calculating the autocorrelation of a small segment of signal sample points near the current time point using a short-time window, as shown in Equation (14).

$$R_i(k) = \sum_{n=0}^{L-k-1} y_i(n)y_i(n+k) \quad (14)$$

where  $k$  is the delay.

### 4.2 Spectral analysis

The most important perceptual characteristics of audio signals are reflected in the power spectrum, and phase changes play a small role. Thus, the frequency domain analysis of audio signals is particularly important (Zhong and Cai, 2019). The frequency domain analysis of audio signals can greatly clarify certain features of the signal that cannot be represented in the temporal domain.

#### 1) Fourier transform

Fourier transform decomposes the signal into a combination of different frequency components, which relates the time-domain characteristics of the signal with its frequency-domain properties. Fourier transform plays a very crucial role in signal processing and is suitable for analyzing periodic, transient, or stationary random signals. The relationship between the time-domain and frequency-domain characteristics of Fourier transform is “global” and cannot reflect local features.

Ship sound signals and sea interference signals are non-stationary, and Fourier analysis on the entire sound signal loses time information. Thus, short-time Fourier transform (STFT) is used to examine time–frequency localization. STFT entails performing Fourier transform on short frames of the sound signal rather than the entire signal. This approach reflects the characteristics of the frequency spectrum of the sound signal over time. STFT is expressed as Equation (15).

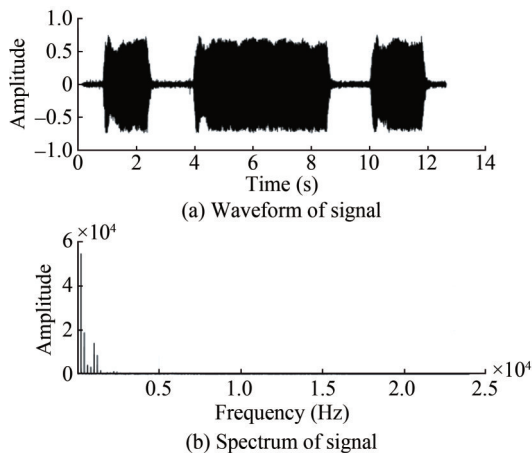
$$\text{STFT}(t, f) = \int_{-\infty}^{\infty} x(\tau)h(\tau - t)e^{-j2\pi f\tau}d\tau \quad (15)$$

#### 2) Spectrum

Any signal can be decomposed into a DC component (i.e., a constant) and a sum of several (infinite) sinusoidal signals through Fourier transformation. Each sinusoidal component has a frequency and an amplitude. The amplitude of each sinusoidal component is plotted against its corresponding frequency, where frequency is the horizontal axis and amplitude is the vertical axis. This plot, called the frequency spectrum, visually represents the spectrum, sound, or other signal, which is a key characteristic of the audio signal.

#### 3) Power spectrum

The power spectrum density function, based on the finite and STFT of the signal, can obtain the short-time power



**Figure 5** Waveform and corresponding spectrum of sound signal

spectrum by transforming the signal power in the frequency band. The short-time power spectrum is the square of the amplitude of the STFT. The power spectrum contains some amplitude information of the frequency spectrum but discards the phase information.

#### 4) Frequency spectrum centroid

The spectral centroid is a crucial feature of the energy and frequency distribution of the sound signal. The spectral centroid is the frequency at which a certain portion of the total spectral energy is centered, weighted by the energy at each frequency band. The spectral centroid is measured in Hz and represents the center of gravity of the frequency components. The spectral centroid indicates where the energy of the signal is concentrated in the frequency domain (Hu et al., 2021), as shown in Equation (16).

$$SC = \frac{\sum_{i=1}^N i \times |X_i|}{\sum_{i=1}^N |X_i|} \quad (16)$$

where  $X_i$  is the spectral energy at frequency  $i$  obtained by the STFT of the signal.

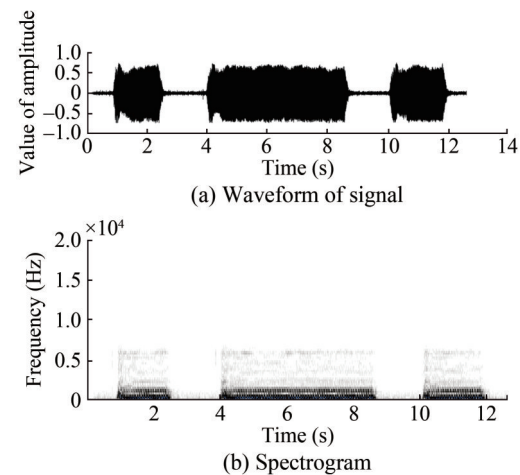
#### 5) Spectrogram

A spectrogram visually represents the frequency characteristics of the sound signal as it varies over time. The intensity of a frequency component at a certain time is represented by the intensity of the corresponding point's gray-scale or color tone. A spectrogram combines the characteristics of a frequency spectrum and a time-domain waveform; it provides a clear view of how the spectrum of a sound varies over time (Yan et al., 2013). The spectrogram is also known as a dynamic spectrum and has important practical value in sound analysis. The vertical axis of the spectrogram corresponds to the frequency, the horizontal axis corresponds to the time, and the grayscale or color represents energy.

### 4.3 Sound signal feature extraction

Various types of sounds, including noisy ship signals,

dolphin whistle sounds, island bird chirping, human communication sounds, and land vehicle horns, can be distinguished by extracting the characteristic values of sound. Combined with the time-domain and frequency-domain characteristics of the sound signal, the short-time energy, short-time average amplitude, short-time ZCR, short-time autocorrelation function, power spectrum, frequency spectrum centroid, and NewMFCC feature parameters of the preprocessed and denoised sound signal are calculated to construct 42-dimensional sound feature values, as shown in Table 1.



**Figure 6** Waveform and corresponding spectrogram of a sound signal

**Table 1** 42-dimensional feature values type display

Serial number	Feature value type	Serial number	Feature value type
1	Short-term energy	6	Frequency spectrum centroid
2	Short-term average amplitude	7–18	MFCC 12-dimensional
3	Short-time zero crossing rate	19–30	ΔMFCC 12-dimensional
4	Short-term auto-correlation function	31–42	WPMFCC 12-dimensional
5	Power spectrum		

For each feature value, the maximum value, minimum value, mean value, standard deviation, root mean square, skewness, and kurtosis are computed, and a set of 294-dimensional sound signal feature vectors based on MFCCs is obtained.

The pattern recognition of sound signals requires not only the feature extraction of the signals but also the design of classifiers for training after the feature extraction. The feature data must be preprocessed to avoid the influence of some vectors with large numerical values on the accuracy of the recognition results and to reduce the computational complexity because different dimensions of feature values have the same effect on the recognition results. Data normalization is an important step in data preprocessing, where the max-min normalization method is used to make the 294-dimensional

data fall within the range of (0, 1). Equation (17) is shown below.

$$x = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (17)$$

To compare the improvement effect of NewMFCC feature parameters with traditional MFCC feature parameters, a control group is established in this experiment, which takes the first 18 dimensions of features in the table. Consistent with the above steps, the feature vector set is constructed and normalized to obtain 126-dimensional data.

## 5 Ship sound signal classification study

### 5.1 Sound classification and recognition algorithm analysis

The classification of sound signals involves pattern recognition for different sound signals, the individual comparison of the test samples with the reference of known samples, and the selection of the best matching pattern as the recognition result. The current research on speech recognition is more mature than that on nonlinguistic sound signal recognition, so inspiration is taken from speech recognition technology. Commonly used models for speech recognition include hidden Markov models, artificial neural networks, random forest algorithms, and SVMs.

Compared with other machine learning algorithms, SVM possesses the advantages of requiring fewer training samples (Tian et al., 2007), satisfactory generalization performance, and high prediction accuracy. Thus, the SVM algorithm is used for sound signal classification and recognition in the case of limited experimental data.

### 5.2 SVM theoretical analysis

SVM is a linear classifier whose maximum margin is defined in the feature space (Bryant and Garber, 1999). SVM can be formalized as a solution to a convex quadratic programming problem. At present, SVM is primarily used to solve problems of nonlinearity, small sample size, and high dimensionality of data.

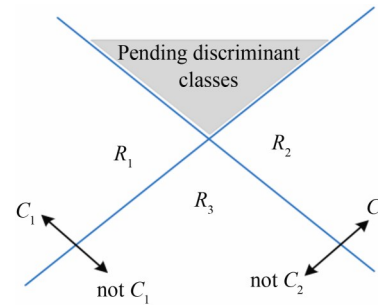
RBF kernel can more effectively solve nonlinear problems compared with linear kernel function, has fewer parameters than polynomial kernel function, reduces computational complexity, and has low complexity. Thus, the RBF kernel is used for classifying ship whistle features in this paper, as expressed in Equation (18).

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right) \quad (18)$$

The SVM algorithm was initially designed to solve binary classification problems, but this experiment requires the recognition of multiple types of sound signals. Two methods for multiclass classification using SVM are currently used: “one-versus-rest” and “one-versus-one” (Tang et al., 2018).

#### 1) One-versus-rest

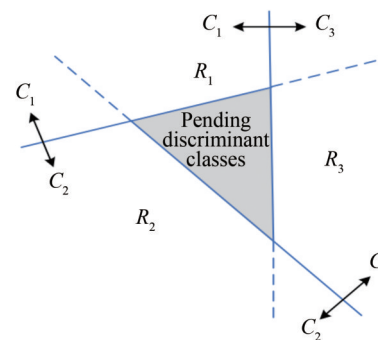
The one-versus-rest method transforms multiple classes into two classes. During training, one group of samples is classified as one class, and the remaining samples are classified as another class. For example, assuming  $n$  classes in a training set,  $n$  SVM classifiers are created. During classification, if a sample belongs to multiple classes simultaneously, as shown in the shaded area in Figure 7, the classification discriminant function values are determined separately for each class. The class corresponding to the maximum classification function value is selected as the class of the sample. The advantage of the “one-versus-rest” method is that it requires fewer SVM classifiers, but its training speed is substantially slow due to the increased sample data.



**Figure 7** “One-versus-rest” classification diagram

#### 2) One-versus-one

The one-versus-one method designs an SVM classifier between every two categories when training the sample dataset and then uses a voting method to determine its class, as shown in the shaded area in Figure 8. The category with the most votes is the class of the sample. For a sample dataset with  $n$  categories,  $(n(n-1))/2$  SVM classifiers must be constructed.



**Figure 8** “One-versus-one” classification diagram

The advantage of the “one-versus-one” method is that it possesses satisfactory training accuracy. However, given



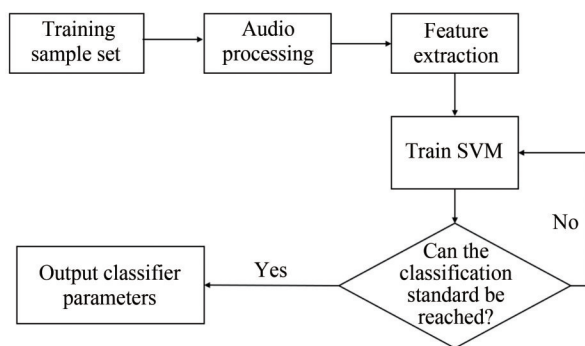
many classes of sample data, numerous classifiers are needed, the corresponding training time is long, and the complexity is high.

### 5.3 Voice signal classification experiment based on SVM

#### 5.3.1 Experimental procedure

Data processing and programming are based on MATLAB R2016a. This experiment uses a one-to-one strategy to solve the multiclassification problem, which is simply a voting algorithm. When the number of votes is equal, the sum of the absolute values of probabilities can be used as the basis for the final classification output. The specific approach is to set an SVM between any two categories of audio samples. Five categories, namely, ship sound signals, dolphin whistles, island bird songs, human communication sounds, and land vehicle horns, and 10 SVM classifiers must be trained. The goal is to recognize and parse the audio signal perceived as ship sound signals.

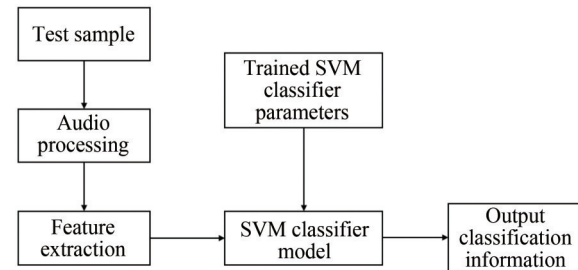
The experiment is divided into two parts. The first part is the establishment and training of the classifier model. The labeled audio is taken as the input and processed. Then, the 42-dimensional feature parameters of the sound signal are extracted, the feature vector set is constructed, the feature vectors are normalized, the SVM model is trained, and various parameters are output. If the classification effect does not meet the expected value, new parameters are used to retrain the model, and the process is repeated until the classification standard is met and the parameters are output. The training structure flowchart is shown in Figure 9.



**Figure 9** Training structure of the SVM classifier model

The second part is recognition, which consists of preprocessing and noise reduction of the test audio signal, feature extraction, input of the normalized data set into the trained model, and output of the recognized category result. The recognition is shown in Figure 10.

Two parameters influence the fitting degree of the SVM classification model: penalty parameter  $C$  and kernel function coefficient. Penalty parameter  $C$  is the tolerance for errors, selected from  $10^t$  ( $t=[-4,4]$ ), between 0.000 1 and 10 000.



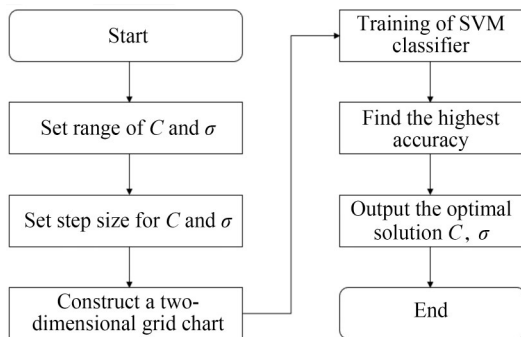
**Figure 10** Structure framework of SVM classifier recognition

If  $C$  is extremely large, the model excessively punishes the error examples, which may cause overfitting; that is, the model learns the input data very thoroughly and cannot correctly categorize it during testing. If  $C$  is extremely small, the margin between the samples and the hyperplane is exceedingly large; that is, the model cannot recognize the data well, making distinguishing the samples challenging for the classification model. Exceedingly large or extremely small values of  $C$  influence the generalization ability of the model.

In theoretical analysis, RBF kernel can more effectively solve nonlinear problems compared with linear kernel function. Moreover, the RBF kernel requires fewer parameters than the polynomial kernel function, which results in reduced computational complexity and lower complexity of the model. Therefore, in this paper, the RBF kernel is adopted for the classification of ship sound signal features. However, if the kernel function coefficient  $\sigma$  is not selected properly, the classification performance may be worse than that of the linear kernel function. A larger  $\sigma$  value results in fewer support vectors, whereas a smaller  $\sigma$  value results in more support vectors. The number of support vectors influences the speed of training and prediction. In practical applications, the penalty parameter  $C$  and the coefficient  $\sigma$  of the RBF kernel need to be adjusted and optimized continuously to improve classification accuracy and achieve optimal performance for the model (Ma and Zheng, 2015).

Grid search is a commonly used algorithm for SVMs. Grid search sets the range and step size of the parameters  $C$  and  $\sigma$  according to the samples, respectively. Given  $M$  values for  $C$  and  $N$  values for  $\sigma$ ,  $C$  and  $\sigma$  can form an  $M \times N$  grid, and each point on the grid represents a possible combination of parameters (Du et al., 2021). The best parameters are selected based on the highest average accuracy on the training set by traversing each node on the grid, training the SVM classifier, and validating the accuracy of the prediction, as shown in Figure 11.

Assessing the performance of a model is essential. In this experiment, cross-validation is used to evaluate the performance of the model. This cross-validation involves randomly dividing the sample data into two parts: the training set and the validation set. First, the classifier is trained using the training set, the accuracy of the model and optimization of the parameters are evaluated using the validation set, and the final classification accuracy is recorded. Then, the



**Figure 11** Flowchart of grid search algorithm

sample is reshuffled, and a new training set and test set are selected to continue training the data and testing the model. The optimal classification model is determined by repeatedly cross-validating and measuring the quality of the model.

### 5.3.2 Experimental evaluation metrics

#### 1) Confusion matrix

The confusion matrix assesses the model results and is a part of model evaluation, which can be used to judge the quality of a classifier (Xu and Liu, 2021). The confusion matrix characterizes the relationship between the true attributes of the sample data and the types of classification predictions in a matrix form.

#### 2) Classification accuracy

The proportion of correctly classified samples is calculated by Equation (19).

$$\text{Accuracy} = \frac{D}{N} \times 100\% \quad (19)$$

where  $D$  is the number of correctly classified audio segments, and  $N$  is the total number of sample audios.

## 5.4 Experimental result analysis

All audio data used in the experimental investigation of the ship signal recognition and analysis system are provided by maritime personnel. These authentic sound recordings are captured by crew members aboard vessels using recording equipment. The sound recordings encompass two sound formats: mp3 and advanced audio coding (AAC). The signals are initially sampled at a frequency of 48 kHz, which is later down-sampled to 8 000 Hz. The sample precision for the recordings is 16 bits. The cumulative duration of the acoustic signals is 2 hours. After processing using Adobe Audition audio processing software, approximately 600 individual independent sound segments are obtained. The duration of each segment varies between 1 second and 2 minutes. The sound signals are preliminarily annotated, resulting in five major categories: ship signals with noise, dolphin whistles, island bird calls, human communication signals,

and land vehicle horn sounds. The ship signal dataset encompasses various horn sounds recorded during anchoring, maritime navigation, or near-port activities. The ship signal dataset features complete, clear sound signals. A total of 600 audio segments are collected in the experiment, and the feature vector set of 294 dimensions based on NewMFCC feature parameters and the feature vector set of 126 dimensions of the control group are used as inputs. Each time, 3/4 of the processed data, that is, 450 audio segments, are taken as the training set of the model and passed to the SVM-TRAIN function. The LIBSVM tool is used to perform a grid search for the optimal parameters to obtain the trained model. The 150 remaining data segments are used as the test set, and combined with the trained model, they are passed to the support vector machine predict function. The confusion matrix and classification accuracy are outputs to cross-validate the performance of the classifier and to construct the best classifier for ship whistle recognition.

When the feature parameter selected is NewMFCC, after continuous training of the model, the optimal classification model is obtained when the penalty parameter  $C$  is 3.943 7, and the kernel function coefficient  $\sigma$  is 1.257 3. After recognizing 150 audio segments in the test set, 47 ship signals are correctly identified, and 4 ship signals are categorized as interference signals. The classifier achieves an accuracy of 92.16%.

Table 2 presents a statistical summary of the classification results for specific categories, which can facilitate comprehending the relationship between the five types of sound signals and the recognition accuracy.

**Table 2** Classification result summary

Category	Identified	Unidentified	Accuracy (%)
Ship sound signals	47	4	92.16
Island bird songs	18	1	94.74
Human communication sounds	35	3	92.11
Land vehicle horns	20	3	82.96
Dolphin whistles	17	2	84.47
Total	137	13	91.33

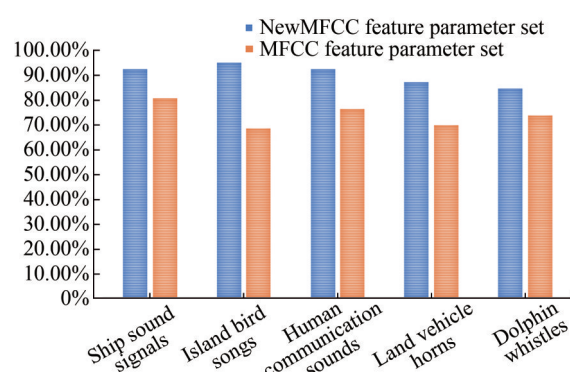
The experimental results reveal that 137 audio segments are classified correctly, and the accuracy of the classifier is 91.33%. This high recognition rate sufficiently demonstrates the feasibility of the sound signal classification model based on SVMs. The control group experiment follows the above experimental steps. The recognition accuracy of each category in the test set is measured, as shown in Table 3.

Based on the data in Table 3, the MFCC feature values of the audio signal are extracted, combined with the time-frequency characteristics of the sound, and recognized by the SVM classification model. The classification accuracy is 75.33%. The comparison of the recognition accuracy of the two groups of experiments is shown in Figure 12. The classification accuracy of each category based on the NewMFCC feature parameter set is remarkably higher

than that based on the MFCC feature parameter set alone. The experiment confirms that the NewMFCC feature parameters obtained by fusing second-order differential and wavelet packet decomposition transformation are favorable to improving classification accuracy.

**Table 3** Classification result statistics for the control group

Category	Identified	Unidentified	Accuracy (%)
Ship sound signals	41	10	80.39
Island bird songs	13	6	68.42
Human communication sounds	29	9	76.32
Land vehicle horns	16	7	69.57
Dolphin whistles	14	5	84.47
Total	113	37	75.33



**Figure 12** Experimental result comparison chart

## 6 Conclusion

The extraction of MFCC feature parameters, building on the MFCC's characteristics, is proposed in this paper. However, the traditional MFCC method has limitations. To mitigate these drawbacks, NewMFCC is introduced by incorporating second-order differences and wavelet packet decomposition, and feature extraction is enhanced to capture dynamic and high-/low-frequency information better. This improvement addresses the shortcomings of traditional MFCC and provides a more comprehensive representation of sound signals. A 42-dimensional audio feature is then constructed using the fused improved NewMFCC from the perspective of the temporal and spectral characteristics of sound. For each feature dimension, seven statistical quantities are calculated, namely, maximum value, minimum value, average value, standard deviation, root mean square, skewness, and kurtosis, to form a 294-dimensional vector set. This set is then normalized to prepare for ship whistle classification and recognition. Then, approximately 600 individual independent sound segments are obtained. SVM theory is used to design the classifier. The comparison of experimental results exhibits that the classification accuracy of

the classifier reaches 91.33%, which represents an improvement of approximately 15%. The results successfully verify the feasibility of the ship whistle recognition system based on SVM in classifying ship whistles and interference signals.

The utilization of signal extraction and classification techniques for ship sound signal recognition holds remarkable potential across various domains and possesses intrinsic societal value. This methodology not only facilitates the exchange of information among vessels, mitigating the risk of collisions and enhancing navigational safety but also offers invaluable data for ship traffic management. This method aids the relevant authorities in comprehending waterway utilization, traffic dynamics, and vessel distribution, hence optimizing traffic oversight and augmenting port operational efficacy.

The applications of the proposed method extend beyond navigation and encompass the monitoring of marine ecological conditions, piscine migrations, and oceanic acoustics. The proposed method contributes to the enhancement of marine ecosystem preservation and encapsulates the broader concept of maritime environmental safeguarding.

**Competing interest** The authors have no competing interests to declare that are relevant to the content of this article.

## References

- Amelia F, Gunawan D (2019) DWT-MFCC method for speaker recognition system with noise. 7th International Conference on Smart Computing & Communications (ICSCC), Miri, Malaysia, 310-314. DOI: 10.1109/ICSCC.2019.8843660
- Bhattacharjee U, Kshirod S (2013) Language identification system using MFCC and prosodic features. International Conference on Intelligent Systems and Signal Processing (ISSP), 194-197. DOI: 10.1109/Issp.2013.6526901
- Bin Hamzah HI, Bin Abdullah A, Candrawati R (2009) Biologically-inspired abstraction model to analyze sound signal. IEEE Student Conference on Research and Development, Serdang, Malaysia, 180-183. DOI: 10.1109/Scored.2009.5443168
- Bryant ML, Garber FD (1999) SVM classifier applied to the mstar public data set. Proc. SPIE 3721, Algorithms for Synthetic Aperture Radar Imagery VI. <https://doi.org/10.1117/12.357652>
- Chomorig, Zhang Z, Xiang YL (2012) Research of feature extraction in mongolian speech based on an improved algorithm of MFCC parameter. 2nd International Conference on Advanced Engineering Materials and Technology (AEMT), Zhuhai, 833-837. DOI: 10.4028/www.scientific.net/AMR.542-543.833
- Chu XT, Wang HP, Yang HT, Lin NH (2022) Speaker recognition based on convolutional neural network. Police Technology (1): 46-50
- Cortes C, Vapnik V (1995) Support-vector networks. Machine Learning 20(3): 273-297. DOI: 10.1023/A:1022627411411
- Deng MQ, Meng TT, Cao JW, Wang SM, Zhang J, Fan HJ (2020) Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. Neural Networks 130: 22-32. DOI: 10.1016/J.Neunet.2020.06.015
- Du C, Shao JH, Yang W, Et Al (2021) Indoor visible light positioning based on support vector machine optimized by grid search method.

- Laser Journal 42(3): 104-109. DOI: 10.14016/J.Cnki.Jgzz.2021.03.104
- Fan P (2017) A complex sound recognition method in noisy environments. Hefei University of Technology (10): 18-21. DOI: 10.7666/D.Y3235272
- Geoffrey L, Collie A (2004) Comparison of novices and experts in the identification of sonar signals. *Speech Communication* 43(4): 297-310. DOI: 10.1016/J.Specom.2004.03.003
- Hu L, Hu XJ, Huang ZH, Xu L, Hu K, Zhang JM (2021) Mura defect detection based on effective background reconstruction and contrast enhancement. *Journal of Liquid Crystals and Displays* 36(10): 1395-1402. DOI: 10.37188/CJLCD.2021-0177
- Lin W, Yang L, Xu B (2006) Speaker recognition based on modified MFCC parameters of chinese mandarin whispered speech. *Journal of Nanjing University (Natural Sciences)* 42(1): 54-62. DOI: CNKI: SUN:NJZ.0.2006-01-007
- Lu MJ, Dong SJ, Tang M, Wang CR, Cao L, Yan XP (2023) Research on the development of china's marine transportation equipment industry. *Chinese Engineering Science* (3): 53-61. DOI: 10.15302/J-Sscae-2023.03.006
- Lv DJ, Zhang Y, Fu QJ, Xu HF, Liu J, Zi JL, Huang X (2020) Birdsong recognition based on MFCC combined with vocal tract properties. 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Harbin, 1523-1526. DOI: 10.1109/Icmce51767.2020.00334
- Ma YJ, Zheng XY (2015) Research and application based on improved support vector machine. *Industrial Control Computer* 28(12): 25-26+28. DOI: CNKI:SUN:GYKJ.0.2015-12-012
- Mahesha P, Vinod DS (2017) LP-Hillbert transform based MFCC for effective discrimination of stuttering dysfluencies. 2nd IEEE International Conference on Wireless Communications, Signal Processing and Networking (Wispnet), Chennai, 2561-2565. DOI: 10.1109/Wispnet.2017.8300225
- Maurya A, Kumar D, Agarwal RK (2017) Speaker recognition for hindi speech signal using MFCC-GMM approach. 6th International Conference on Smart Computing and Communications (ICSCC), Kurukshetra, 880-887. DOI: 10.1016/J.Procs.2017.12.112
- Nassim A, Abderrahmane A (2017) Boosting scores fusion approach using front-end diversity and adaboost algorithm, for speaker verification. *Computers and Electrical Engineering* 62: 648-662. DOI: 10.1016/J.Compeleceng.2017.03.022
- Jiang N, Qiu M, Dai W (2017) SROC: A speaker recognition with data decision level fusion method in cloud environment. *J Sign Process Syst* 86: 123-133. DOI: 10.1007/S11265-015-1100-7
- Shi YB, Wang L (2011) Improved MFCC algorithm in speaker recognition system. *International Conference on Graphic and Image Processing (ICGIP 2011)*, Cairo, Egypt, 828567. <https://doi.org/10.1117/12.913462>
- Srivastava S, Bhardwaj S, Bhandari A, Gupta K, Bahl H, Gupta JRP (2012) Wavelet packet based mel frequency cepstral features for text independent speaker identification. 1st International Symposium on Intelligent Informatics (ISI12), Chennai, India, 237-247. [https://doi.org/10.1007/978-3-642-32063-7\\_26](https://doi.org/10.1007/978-3-642-32063-7_26)
- Tang L, Tian YJ, Pardalos Panos M (2018) A novel perspective on multiclass classification: Regular simplex support vector machine. *Information Sciences* 480: 324-338. <https://doi.org/10.1016/j.ins.2018.12.026>
- Tian J, Xue SH, Huang HN, Zhang CH (2007) Classification of underwater still objects based on multi-field features and SVM. *J Mar. Sc. Appl.* 6: 36-40. DOI: 10.1007/S11804-007-6042-4
- Tomchuk KK (2018) Spectral masking in MFCC calculation for noisy speech. *Wave Electronics and Its Application In Information and Telecommunication Systems (WECONF)*, St Petersburg, Russia, 1-4. DOI: 10.1109/WECONF.2018.8604460
- Tuncer T, Aydemir E (2020) An automated local binary pattern ship identification method by using sound. *Acta Infologica* 4(1): 57-63. DOI: 10.26650/Acta.762809
- Wang LH, Zhao ZH (2020) Marine mechanical noise monitoring system based on MFCC-SVM. *Automation and Instrumentation* 35(12): 54-58. DOI: 10.19557/J.Cnki.1001-9944.2020.12.012
- Wang S, Ding N, Li N, Zhang JJ, Zong CQ (2023) Language cognition and language computation: language understanding by human and machine. *Chinese Science: Information Science* 52(10): 1748-1774. DOI: 10.48550/Arxiv.2301.04788
- Wang Y, Hu WP (2018) Speech emotion recognition based on improved MFCC. In *Proceedings of the 2nd International Conference on Computer Science and Application Engineering (CSAE'18)*, New York, Article 88: 1-7. DOI: 10.1145/3207677.3278037
- Wróbel K, Montewka J, Kujala P (2017) Towards the assessment of potential impact of unmanned vessels on maritime transportation safety. *Reliability Engineering & System Safety* 165: 155-169. DOI: 10.1016/J.Res.2017.03.029
- Xie GD, Yang JC, Yi Q, Han ZF (2010) Research on battlefield target sound detection technology. *Control Engineering* 17(S1): 41-44. DOI: 10.14107/J.Cnki.Kzgc.2010.S1.025
- Xie SS, Xu HF, Liu J, Zhang Y, Lv DJ (2021) Research on bird songs recognition based on MFCC-HMM. *International Conference on Computer, Control and Robotics (ICCCR)*, Shanghai, 262-266. DOI: 10.1109/Icccr49711.2021.9349284
- Xu H, Liu YL (2021) Unmanned aerial vehicle sound recognition algorithm based on deep learning. *Computer Science* 48(7): 225-232. DOI: 10.11896/Jsikx.200500091
- Yan CF, Chun HZ, Yuan LD, Zhi YS (2013) Coarse frequency offsetestimation of TD-LTE based on spectrum centroid. *Advanced Materials Research* 756-759: 3602-3606. DOI: 10.4028/www.scientific.Net/AMR.756-759.3602
- Yan Q, Zhou ZJ, Li S (2011) Chinese accents identification with modified MFCC. *International Conference on Instrumentation, Measurement, Circuits and Systems (ICIMCS 2011)*, Hong Kong, 659-666. [https://doi.org/10.1007/978-3-642-27334-6\\_77](https://doi.org/10.1007/978-3-642-27334-6_77)
- Yang Y, Chen X (2020) Experimental design of speaker recognition based on neural networks. *Laboratory Research and Exploration* 39(9): 38-41+50. DOI: 10.3969/J.Issn.1006-7167.2020.09.008
- Yu Y, Xu JX, Zhang MC, Shi H (2019) Comparative analysis of wavelet analysis and wavelet packet analysis in bearing fault diagnosis. *Coal Mine Machinery* 40(12): 170-173. DOI: CNKI: SUN:MKJX.0.2019-12-055
- Zhang C, Ma Y (2012) Ensemble machine learning: methods and applications. *Springer Science & Business Media*, 10-329. DOI: 10.1007/9781441993267
- Zhang XY, Bai J, Liang WZ (2006) The speech recognition system based on bark wavelet MFCC. 8th International Conference on Signal Processing, Guilin, 833-835. DOI: 10.1109/ICOSP.2006.345539
- Zheng F, Zhang GL, Song ZJ (2001) Comparison of different implementations of MFCC. *Journal of Computer Science and Technology* 16(6): 582-589. DOI: 10.1007/BF02943243
- Zhong M, Cai W (2019) Feature fusion-based recognition of marine mammal sounds. *Electronic Science and Technology* 32(5): 32-37. DOI: 10.16180/J.Cnki.Issn1007-7820.2019.05.007



# Journal Of Marine Science And Application

## Copyright

Submission of a manuscript implies: that the work described has not been published before (except in the form of an abstract or as part of the published lecture, review, or thesis); that it is not under consideration for publication elsewhere; that its publication has been approved by all co-authors, if any, as well as - tacitly or explicitly - by the responsible authorities at the institution where the work was carried out. The author warrants that his/her contribution is original and that he/she has full power to make this grant. The author signs for and accepts responsibility for releasing this material on behalf of any and all co-authors. Transfer of copyright to Harbin Engineering University and Springer becomes effective if and when the article is accepted for publication. After submission of the Copyright Transfer Statement signed by the corresponding author, changes of authorship or in the order of the authors listed will not be accepted by Harbin Engineering University and Springer. The copyright covers the exclusive right (for U.S. government employees: to the extent transferable) to reproduce and distribute the article, including reprints, translations, photographic reproductions, microform, electronic form (offline, online) or other reproductions of similar nature.

An author may self-archive an author-created version of his/her article on his/her own website. He/she may also deposit this version on his/her institution's and funder's (funder designated) repository at the funder's request or as a result of a legal obligation, including his/her final version, provided it is not made publicly available until after 12 months of official publication. He/she may not use the publisher's PDF version which is posted on [www.springerlink.com](http://www.springerlink.com) for the purpose of self-archiving or deposit. Furthermore, the author may only post his/her version provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The original publication is available at [www.springerlink.com](http://www.springerlink.com)".

All articles published in this journal are protected by copyright, which covers the exclusive rights to reproduce and distribute the article (e.g. as offprints), as well as all translation rights. No material published in this journal may be reproduced photographically or stored on microfilm, in electronic data bases, video disks, etc., without first obtaining written permission from the publishers. The use of general descriptive names, trade names, trademarks, etc., in this publication, even if not specifically identified, does not imply that these names are not protected by the relevant laws and regulations.

While the advice and information in this journal is believed to be true and accurate at the date of its going to press, neither the authors, the editors, nor the publishers can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Special regulations for photocopies in the USA: Photocopies may be made for personal or in-house use beyond the limitations stipulated under Section 107 or 108 of U.S. Copyright Law, provided a fee is paid. All fees should be paid to the Copyright Clearance Center, Inc., 222 Rosewood Drive Danvers, MA 01923, USA, Tel.: +1-978-7508400, Fax: +1-978-6468600, <http://www.copyright.com>, stating the ISSN of the journal, the volume, and the first and last page numbers of each article copied. The copyright owner's consent does not include copying for general distribution, promotion, new works, or resale. In these cases, specific written permission must first be obtained from the publishers.

## Subscription information

**ISSN print edition:** 1671-9433

**ISSN electronic edition:** 1993-5048

Volume 23 (4 issues) will appear in 2024

## Subscription rates

For information on subscription rates please contact: Customer Service

**North and South America:** [journals-ny@springer.com](mailto:journals-ny@springer.com)

**Outside North and South America:** [subscriptions@springer.com](mailto:subscriptions@springer.com)

## Orders and inquiries:

### North and South America

Springer New York, Inc.

Journal Fulfillment

P.O. Box 2485

Secaucus, NJ 07096 USA

Tel: 1-800-SPRINGER or 1-201-348-4033

Fax: 1-201-348-4505

Email: [journals-ny@springer.com](mailto:journals-ny@springer.com)

### Outside North and South America

Springer Distribution Center

Customer Service Journals

Haberstr. 7

69126 Heidelberg, Germany

Tel: +49-6221-345-0, Fax: +49-6221-345-4229

Email: [subscription@springer.com](mailto:subscription@springer.com)

**Cancellations** must be received by September 30 to take effect at the end of the same year.

**Changes of address:** Allow for six weeks for all changes to become effective. All communications should include both old and new addresses (with postal codes) and should be accompanied by a mailing label from a recent issue. According to § 4 Sect. 3 of the German Postal Services Data Protection Regulations, if a subscriber's address changes, the German Federal Post Office can inform the publisher of the new address even if the subscriber has not submitted a formal application for mail to be forwarded. Subscribers not in agreement with this procedure may send a written complaint to Customer Service Journals with 14 days of publication of this issue.

**Microform editions** are available from: ProQuest. Further information available at <http://www.il.proquest.com/uni>

## Electronic edition

An electronic version is available at [springerlink.com](http://springerlink.com).

## Production

Printed in People's Republic of China

Springer-Verlag is part of Springer Nature Singapore Pte Ltd