



Article

Modulation Mode Recognition Method of Non-Cooperative Underwater Acoustic Communication Signal Based on Spectral Peak Feature Extraction and Random Forest

Tao Fang^{1,2,3}, Qian Wang^{1,2,3}, Lanyue Zhang^{1,2,3} and Songzuo Liu^{1,2,3,*}

- ¹ Acoustic Science and Technology Laboratory, Harbin Engineering University, Harbin 150001, China; fangtao@hrbeu.edu.cn (T.F.); wq15663765839@163.com (Q.W.); zhanglanyue@hrbeu.edu.cn (L.Z.)
- ² Key Laboratory of Marine Information Acquisition and Security, Harbin Engineering University, Ministry of Industry and Information Technology, Harbin 150001, China
- ³ College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin 150001, China
- * Correspondence: liusongzuo@hrbeu.edu.cn

Abstract: The modulation mode recognition of non-cooperative underwater acoustic (UWA) communication signal faces great challenges due to the influence of the UWA channel and the demand for efficient recognition. This work proposes a recognition method for UWA orthogonal frequency division multiplexing (OFDM), binary frequency shift keying (2FSK), four-frequency shift keying (4FSK), and eight-frequency shift keying (8FSK) by using spectral peak feature extraction combined with random forest (RF). First, a new spectral peak feature extraction method is proposed. In this method, pre-processing, waveform optimization, and feature extraction are used to ensure that the extracted feature maintains high robustness in the UWA channel. Then, we designed an RF classifier that can meet the demand for high-efficiency recognition and good performance. Finally, simulation and experimental results verified the feasibility of the recognition method.



Citation: Fang, T.; Wang, Q.; Zhang, L.; Liu, S. Modulation Mode Recognition Method of Non-Cooperative Underwater Acoustic Communication Signal Based on Spectral Peak Feature Extraction and Random Forest. *Remote Sens.* **2022**, *14*, 1603. <https://doi.org/10.3390/rs14071603>

Academic Editor: Jaroslaw Tegowski

Received: 11 February 2022

Accepted: 23 March 2022

Published: 26 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: underwater acoustic communication; modulation mode recognition; feature extraction; random forest

1. Introduction

With the increasing number of underwater acoustic sensor networks (UWSNs) (Figure 1) deployed in underwater environments to perform certain tasks, such as remote control, submarine assistance, intelligent monitoring, and marine data acquisition [1,2], the modulation mode recognition technology of non-cooperative underwater acoustic (UWA) communication signals [3,4] has gradually become a research hotspot. The traditional modulation mode recognition method of a non-cooperative communication signal can be divided into feature extraction and classifier design.

The feature extraction methods in radio communication were presented in [5–8]. The features of the high-order cumulant [5] and wavelet transform [6] show good performance in the Gaussian channel, but poor performance in the multipath channel. The calculation process of the cyclic spectral feature [7] is particularly complex, and a priori information such as the carrier frequency is required before obtaining the instantaneous information feature [8], so they cannot be directly applied in non-cooperative conditions. The feature extraction methods in UWA communication were presented in [9–12]. Reference [9] used high-order cumulant and continuous wavelet transform features to recognize multiple modulation modes, but the modeled channel was a Gaussian channel. Reference [10] extracted the spectral peak feature in the autocorrelation waveform to recognize orthogonal frequency division multiplexing (OFDM) based on the character of the cyclic prefix (CP). However, the extracted spectral peak feature will become unstable under a low signal-to-noise ratio (SNR). Reference [11] extracted the feature in the frequency domain waveform to recognize multiple modulation modes. However, the extracted feature is only effective

in the ideal UWA channel. Reference [12] extracted the mean value and the variance of the spectral envelope as the features to recognize M-ary frequency shift keying (MFSK) and the square spectrum envelope as the feature to recognize binary phase shift keying (BPSK). However, this method does not consider the influence of the UWA multipath channel, which is characterized by frequency-selective fading in the frequency domain.

Scholars have designed numerous classifiers, such as decision tree [13], support vector machine (SVM) [14], and k-nearest neighbor (KNN) [15]. The decision tree classifier contains the recognition order and multiple threshold settings, making it difficult to achieve optimal performance. KNN has high complexity in processing a large amount of data. The SVM classifier is mostly used in binary classification problems. The classifier is composed of multiple SVM classifiers, and in the multi-classification problem, the performance will be reduced and the complexity will increase. Deep learning has also become a popular method for modulation mode recognition [16]. However, deep learning requires a substantial amount of training data, which may be limited in application considering the great difference in the channel environment in various sea areas.

On the basis of the above references, this work optimized the autocorrelation and frequency domain waveforms to obtain a high robustness spectral peak and directly extracted multiple local maximums as the spectral peak features from optimized waveforms to form a feature set for training the classifier. A low-complexity random forest (RF) classifier was designed to meet the demand of high-efficiency recognition in the underwater environment. Finally, an effective recognition of orthogonal frequency division multiplexing (OFDM), binary frequency shift keying (2FSK), four-frequency shift keying (4FSK), and eight-frequency shift keying (8FSK) was realized. The proposed method can effectively reduce the influence of a low SNR and multipath in the UWA channel. In addition, the complexity of the classifier is low. The structure of this paper is as follows: The Section 2 of this paper mainly describes the system model. In the Section 3, we propose the extraction method of the spectral peak feature. In the Section 4, we mainly describe the design of the RF classifier. In the Section 5, we analyze the simulation data. In the Section 6, we analyze the experimental data. In the Section 7, we summarize the full paper.

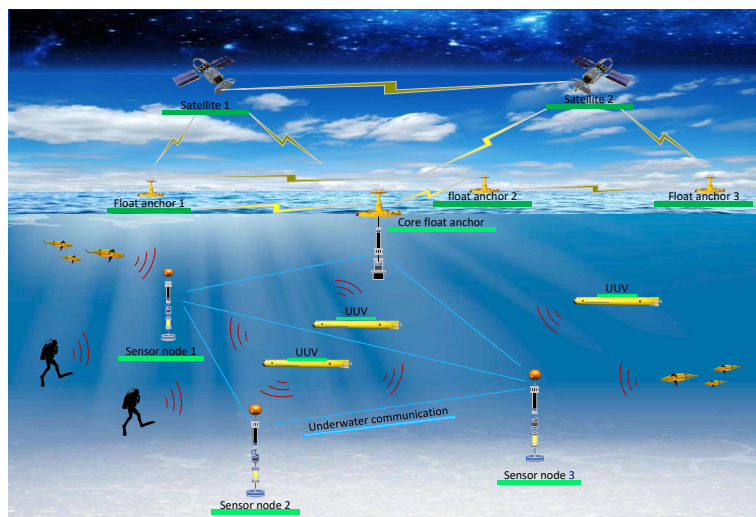


Figure 1. UWSNs.

2. System Model

Assuming that the transmitted signal is $s(n)$, the received signal can be expressed as follows:

$$x(n) = s(n) * h(n) + w(n), \quad (1)$$

where $h(n)$ is the channel impulse response (CIR), $w(n)$ is the channel noise, and $*$ represents the convolution operation. In this work, we assumed that the modulation mode of

the received signal is one of OFDM, 2FSK, 4FSK, or 8FSK. This study aimed to recognize the modulation mode of the received signal.

3. Proposed Spectral Peak Feature Extraction Method

In this section, we aim to optimize the autocorrelation and frequency domain waveforms, so that the optimized waveforms only have a large local maximum at the spectral peak, while the local maximums in other areas are small. The optimized autocorrelation and frequency domain waveforms of different modulation modes have obvious differences in the amplitude of local maximums. Accordingly, these local maximums can be directly extracted to realize the recognition of modulation modes. The process of the spectral peak feature extraction method is shown in Figure 2, which includes three steps. The first step is to pre-process the signal by the autocorrelation and Fourier transform. We can obtain the autocorrelation and frequency-domain waveforms containing spectral peaks. The second step is to optimize the autocorrelation and frequency-domain waveforms, including the absolute value acquisition, spectral peak enhancement, moving average filtering, and Gaussian fitting. The third step is to find local maximums. Each step is described in detail below.

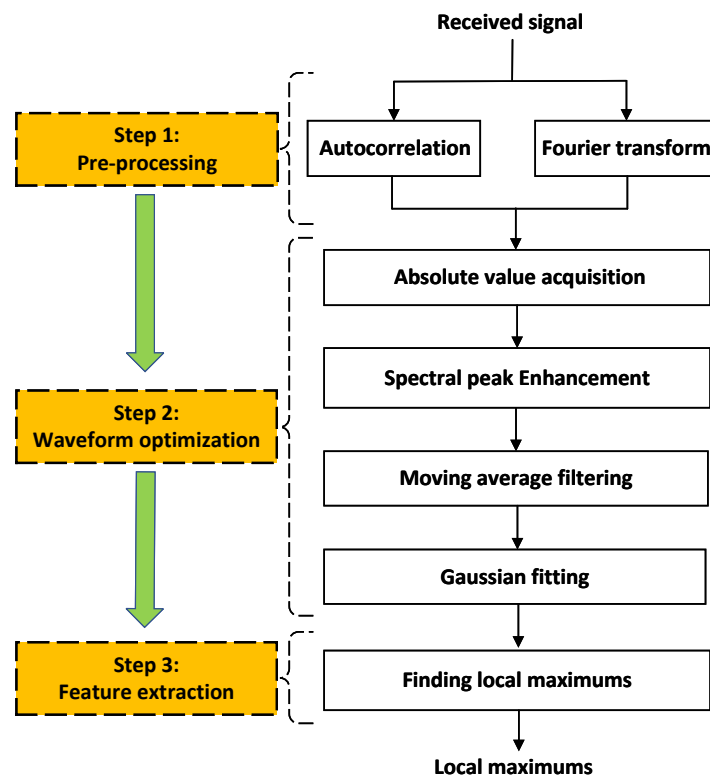


Figure 2. The process of the spectral peak feature extraction method.

3.1. Pre-Processing

Assuming that the length of the received signal $x(n)$ is N , the autocorrelation process of $x(n)$ can be expressed as follows:

$$R_{xx}(k) = \sum_{n=0}^{N-1} x(n+k) \cdot x(n), k = 0, 1, \dots, N-1, \quad (2)$$

where $R_{xx}(k)$ is the autocorrelation output and \cdot represents the multiplication operation. UWA OFDM uses the CP to resist the interference of the multipath channel, and the CP is the copy of partial OFDM symbols. Accordingly, the autocorrelation operation in OFDM will form another correlation peak other than the main peak. We removed the main peak

according to the empirical value. This notion means that we only extracted the spectral peak feature in the areas outside the main peak.

The frequency domain waveform represents the frequency distribution of the signal. The N -point discrete Fourier transform (DFT) of $x(n)$ can be expressed as follows:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{N}kn}, k = 0, 1, \dots, N-1, \quad (3)$$

where $X(k)$ is the frequency-domain output. According to the characteristics of the DFT, we only took the first $N/2$ points for processing. 2FSK, 4FSK, and 8FSK have significantly different frequency-domain waveforms. 2FSK has two spectral peaks; 4FSK has four spectral peaks; 8FSK has eight spectral peaks; obvious spectral lines will appear in the spectral peaks at each modulation frequency.

3.2. Waveform Optimization

3.2.1. Absolute Value Acquisition

We use $A(q)$ to represent $R_{xx}(k)$ and $X(k)$ for the convenience of formula derivation. We assumed that the length of $A(q)$ is Q . We took the absolute value of $A(q)$ to obtain the better spectral peaks from the autocorrelation and frequency-domain waveforms. The output waveform after the absolute value operation can be expressed as follows:

$$B(q) = |A(q)|, \quad (4)$$

where $|\cdot|$ represents the absolute value operation.

3.2.2. Spectral Peak Enhancement

We propose a new spectral peak enhancement method by using the energy difference between two sliding windows. We established two sliding windows with different lengths centered on the sampling point q of $B(q)$ and obtained the gain coefficient $c(q)$ through calculating the energy difference between the two sliding windows. $c(q)$ can be expressed as follows:

$$c(q) = E_{s_{win}} - (E_{l_{win}} - E_{s_{win}}), \quad (5)$$

where:

$$E_{s_{win}} = \sum_{i=q-s_{win}}^{q+s_{win}} |B(i)|^2, \quad (6)$$

$$E_{l_{win}} = \sum_{i=q-l_{win}}^{q+l_{win}} |B(i)|^2, \quad (7)$$

where s_{win} corresponds to a short sliding window, with the window length being $2s_{win} + 1$, and l_{win} corresponds to a long sliding window, with the window length being $2l_{win} + 1$. A relationship of $l_{win} = 2s_{win}$ can be observed. $E_{s_{win}}$ is the energy of the short sliding window, and $E_{l_{win}}$ is the energy of the long sliding window. In addition, we made $c(q) = 0$ directly in the special case of $c(q) < 0$, $q - s_{win} \leq 0$, $q - l_{win} \leq 0$, $q + s_{win} > Q$ or $q + l_{win} > Q$. The output waveform after spectral peak enhancement can be expressed as follows:

$$C(q) = B(q) \cdot c(q). \quad (8)$$

Therefore, the waveform at the spectral peak is amplified, and the those in other areas are suppressed. The length of the sliding window should include the spectral peak as much as possible, but it should not be considerably long. 2FSK, 4FSK, and 8FSK have multiple spectral peaks. A particularly long sliding window will lead to multiple spectral peaks in the window. In addition, if the length of the sliding window is considerably short, then the ability to resist interference will also be weakened.

3.2.3. Moving Average Filter

We performed moving average filtering on waveform $C(q)$. This step can reduce the random noise and obtain the envelope of the spectral peak. The output waveform after moving average filtering can be expressed as follows:

$$H(q) = \frac{1}{2U+1} \sum_{i=q-U}^{q+U} C(i), \quad (9)$$

where the length of the moving average filter is $2U+1$. A large number of pseudo envelopes will form if the length of the moving average filter is considerably short. An amount of useless information will be generated if the length of the moving average filter is long. We set the length of the moving average filter to be consistent with the length of the short sliding window to effectively obtain the envelope of the spectral peak waveform [17,18]. In addition, we made $H(q) = C(q)$ directly in the special case of $q-U \leq 0$ or $q+U > Q$.

3.2.4. Gaussian Fitting

We performed Gaussian fitting on waveform $D(q)$. This step can further smoothen $D(q)$ [19], which is conducive to the subsequent finding of local maximums. The Gaussian window function can be expressed as follows:

$$G(q) = e^{-\frac{1}{2} \left(\alpha \frac{q}{(L-1)/2} \right)^2} = e^{-q^2/2\sigma^2}, \quad (10)$$

where the length of the Gaussian window is L , $-(L-1)/2 \leq q \leq (L-1)/2$, and α is inversely proportional to the standard deviation $\sigma = (L-1)/2\alpha$ of the Gaussian window function. We set the length of the Gaussian window to be consistent with the length of the short sliding window and the length of the moving average filter to effectively fit the waveform. The output waveform after Gaussian fitting can be expressed as follows:

$$V(q) = \frac{1}{L} \sum_{i=q-(L-1)/2}^{q+(L-1)/2} H(i) \cdot G(i). \quad (11)$$

In addition, we made $V(q) = D(q)$ directly in the special case of $q - (L-1)/2 \leq 0$ or $q + (L-1)/2 > Q$. Finally, we normalized $V(q)$ to obtain an optimized waveform.

We took OFDM, 2FSK, 4FSK, and 8FSK as examples to illustrate how the proposed waveform optimization method works for clarity. The channel used in the simulation was the same as the channel in Section 5. The bandwidth of each modulation mode was 4 kHz, and the center frequency was 12 kHz. An optimal performance was difficult to achieve by setting the same parameters because these modulation modes have different spectral peak structures. Accordingly, we set the length of the short sliding window of OFDM as 200 sampling points, the length of the short sliding window of 2FSK as the frequency point length corresponding to the 600 Hz bandwidth, the length of the short sliding window of 4FSK as the frequency point length corresponding to the 400 Hz bandwidth, and the length of the short sliding window of 8FSK as the frequency point length corresponding to the 200 Hz bandwidth according to the empirical value. Figures 3–6 show the process of OFDM, 2FSK, 4FSK, and 8FSK waveform optimization. After the absolute value of the waveform is multiplied by the gain coefficient, the spectral peak is enhanced. Subsequently, the spectral peak envelope is formed after moving average filtering. Then, the envelope is further smoothed by Gauss fitting. At this time, only a local maximum exists at the spectral peak, while the local maximum is small or even tends to zero in other areas. Therefore, we can extract these local maximums directly in the optimized waveform as the spectral peak feature to realize the modulation mode recognition.

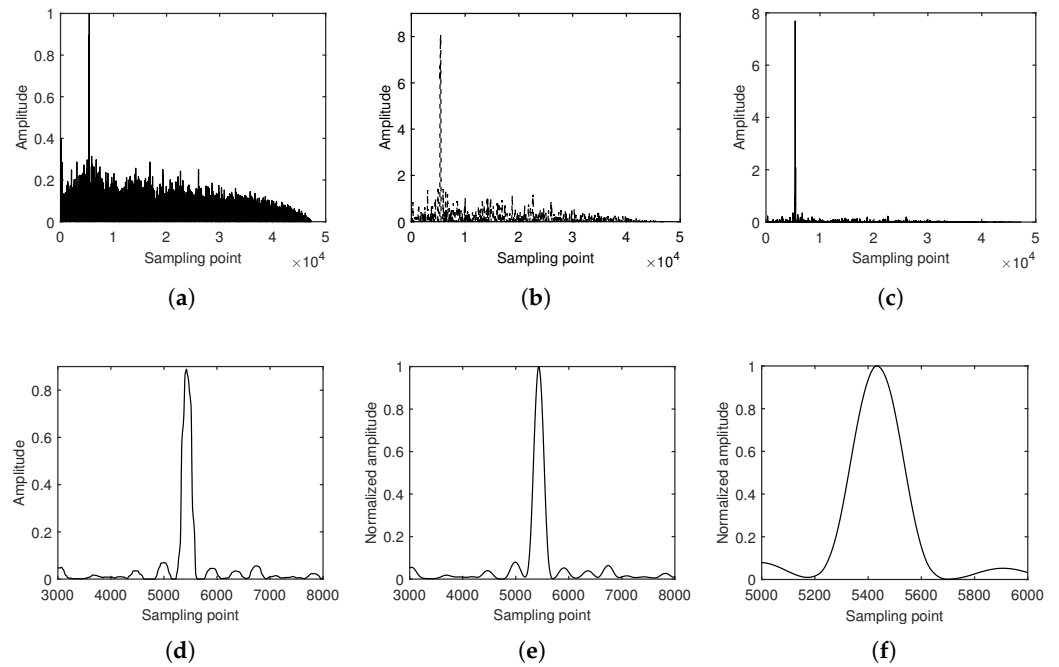


Figure 3. Waveform optimization of OFDM. (a) Waveform after absolute value acquisition, (b) gain coefficient, (c) waveform after spectral peak enhancement, (d) partial enlargement of the waveform after moving average filtering, (e) partial enlargement of the waveform after Gaussian fitting, and (f) partial enlargement of (e).

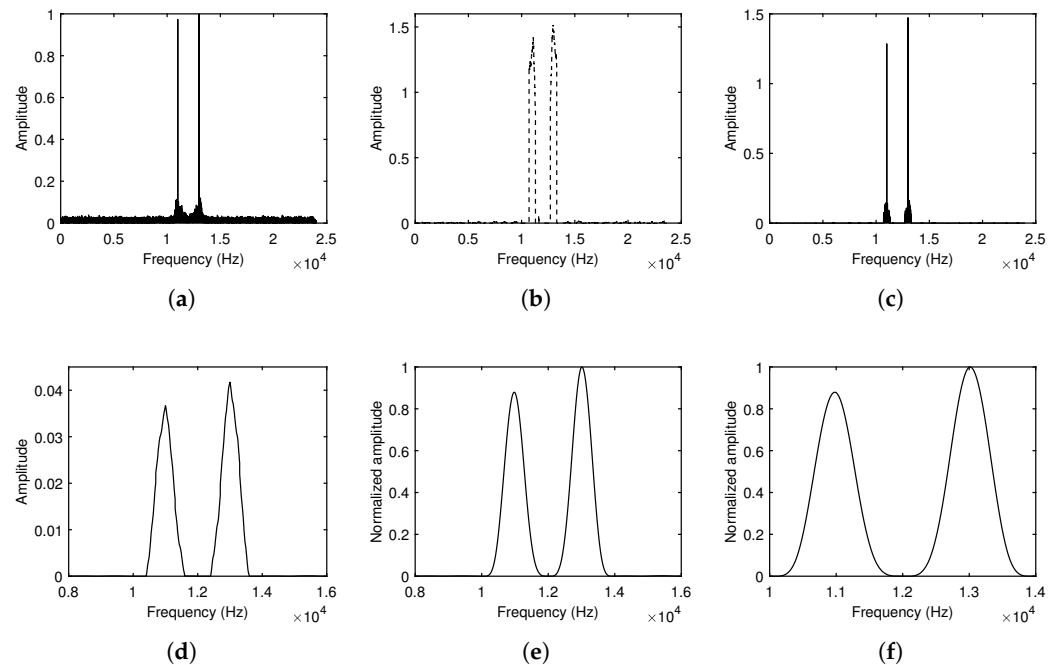


Figure 4. Waveform optimization of 2FSK. (a) Waveform after absolute value acquisition, (b) gain coefficient, (c) waveform after spectral peak enhancement, (d) partial enlargement of the waveform after moving average filtering, (e) partial enlargement of the waveform after Gaussian fitting, and (f) partial enlargement of (e).

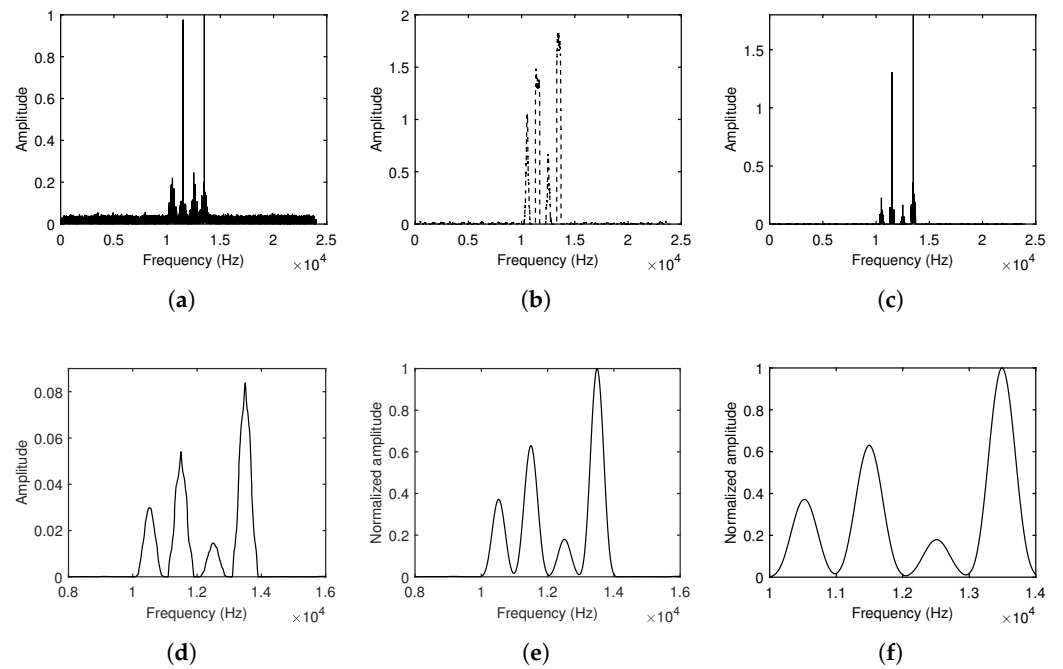


Figure 5. Waveform optimization of 4FSK. (a) Waveform after absolute value acquisition, (b) gain coefficient, (c) waveform after spectral peak enhancement, (d) partial enlargement of the waveform after moving average filtering, (e) partial enlargement of the waveform after Gaussian fitting, and (f) partial enlargement of (e).

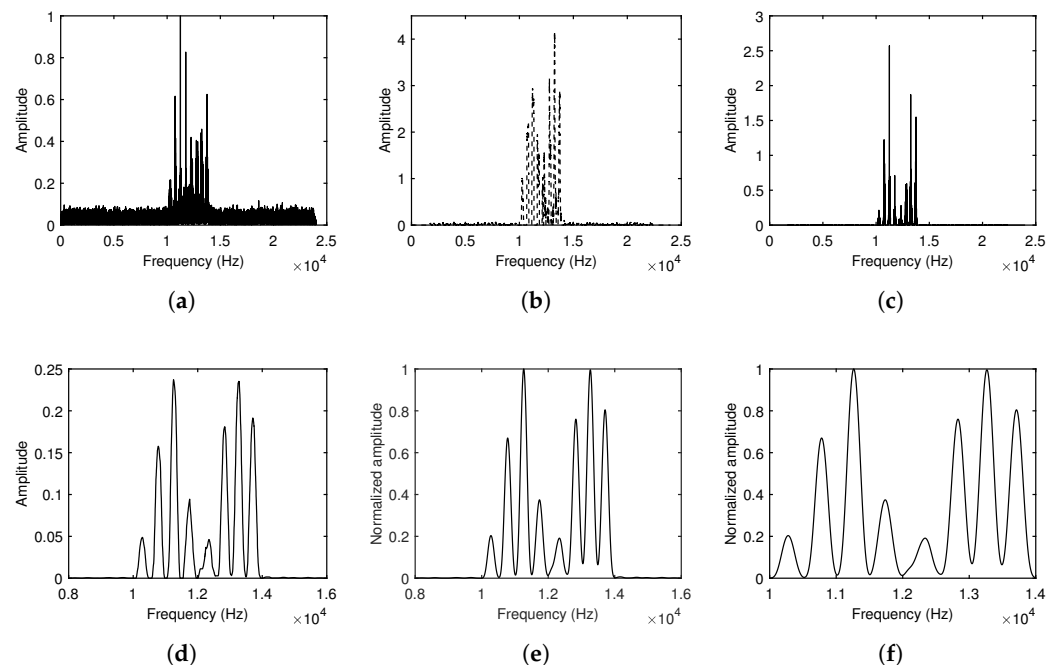


Figure 6. Waveform optimization of 8FSK. (a) Waveform after absolute value acquisition, (b) gain coefficient, (c) waveform after spectral peak enhancement, (d) partial enlargement of the waveform after moving average filtering, (e) partial enlargement of the waveform after Gaussian fitting, and (f) partial enlargement of (e).

3.3. Feature Extraction

First, we extracted all the local maximums of the optimized waveform. Then, we sorted the local maximums according to the amplitude of the local maximum. Finally,

we extracted the corresponding number of local maximums according to the modulation mode. In the above simulation, we extracted the second local maximum because it is relatively small when the modulation mode is OFDM. We extracted the second and third local maximums because the former is greater than zero, while the latter tends to zero when the modulation mode is 2FSK. We extracted the local maximums from the second to the fifth because the local maximums from the second to the fourth are all greater than zero, while the fifth local maximum tends to zero when the modulation mode is 4FSK. We extracted the local maximums from the second to the ninth because the local maximums from the second to the eighth are all greater than zero, while the ninth local maximum tends to zero when the modulation mode is 8FSK. However, we did not know the modulation mode of the received signal in the actual processing. Accordingly, we extracted the spectral peak features according to the following process. First, we assumed the modulation mode of the received signal. Then, we extracted the corresponding local maximums by setting the parameters in the waveform optimization method according to the modulation mode. Hence, a total of 15 local maximums were extracted. Finally, we used these extracted local maximums to form a feature set.

4. RF Classifier

4.1. Classifier Principle

The RF classifier is composed of multiple decision trees. Assuming that the number of features in the feature set is F , each decision tree node of the RF will randomly select K features from the feature set, and an optimal feature is selected from the K features for division. We used the CART decision tree [20] to illustrate how the optimal feature is selected. CART decision tree uses the Gini index to select the optimal feature. We assumed that a feature set called D contains M modulation modes, and the purity of D can be measured by the Gini value. The Gini value can be expressed as follows:

$$\text{Gini}(D) = 1 - \sum_{m=1}^M p_m p_m, \quad (12)$$

where $p_m (m = 1, 2, \dots, M)$ represents the proportion of modulation mode m . Equation (12) illustrates that the smaller $\text{Gini}(D)$ is, the higher the purity of D will be. D will be divided into K branches with K features randomly selected by each decision tree node of the RF. Any feature is assumed to be the k th feature, which divides D into D_1^k and D_2^k . The Gini index of the k th feature can be expressed as follows:

$$\text{Gini_index}(D, k) = \frac{\langle D_1^k \rangle}{\langle D \rangle} \text{Gini}(D_1^k) + \frac{\langle D_2^k \rangle}{\langle D \rangle} \text{Gini}(D_2^k), \quad (13)$$

where $\langle \cdot \rangle$ represents the number of samples in the feature set. Finally, the CART decision tree selects the feature that minimizes the Gini index for division. The training process of the RF classifier is shown in Figure 7. Assuming that the number of decision trees of the RF is Q , feature sets $D_1 \sim D_Q$ are generated by the bootstrap sampling method [20], and each decision tree is trained by the corresponding feature set. Figure 8 demonstrates the decision process of the RF classifier. We summarize the recognition results generated by each decision tree and screened out the modulation mode that frequently appeared in the results.

4.2. Complexity Analysis

The RF classifier has less computational complexity, compared with common SVM and KNN classifiers. The computation of linear SVM is proportional to the product of the number of support vectors and the dimension of the feature set [14] when linear SVM realizes the modulation mode recognition of a single signal for binary classification problems. The computation of linear SVM will increase accordingly for multi-classification

problems. In order to realize the modulation mode recognition of a single signal, by using the KNN classifier, we needed to calculate the distance between the signal and all training data to obtain its nearest neighbor, of which the computation was the highest among the three classifiers. Assuming that the types of modulation modes to be classified are P , in order to realize the modulation mode recognition of a single signal, the RF classifier needs $Q(P - 1)$ comparators at most in the decision process and $Q - 1$ adders and $P - 1$ comparators in the final voting process when there are Q decision trees. In addition, no complex multiplication operation exists in the classification process of the RF classifier. We can see from the comparison that the complexity of the classification process by using the RF classifier was relatively lower. In practical applications, the modulation mode recognition of non-cooperative UWA communication signals requires real-time monitoring. Therefore, using the RF classifier can meet the needs of efficient recognition.

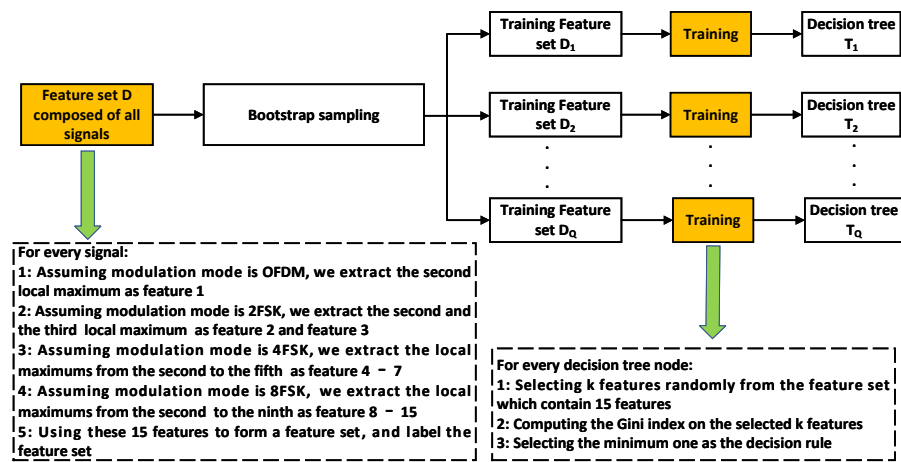


Figure 7. The training process of the RF classifier.

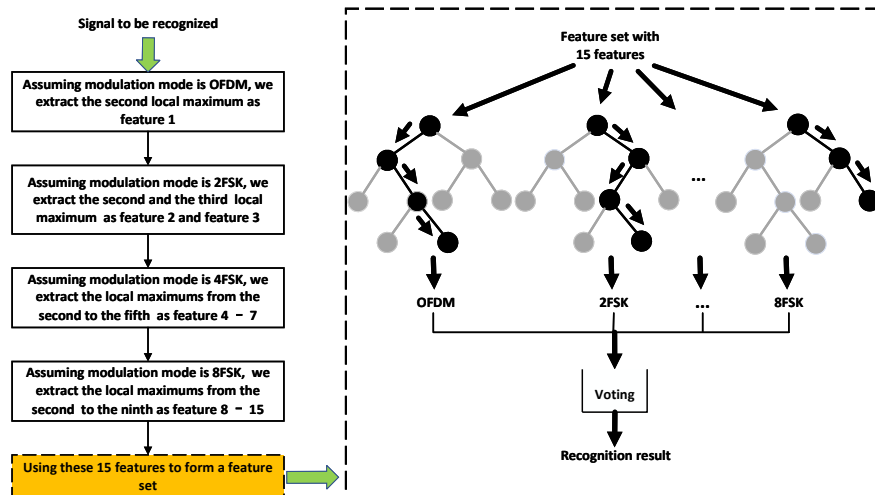


Figure 8. The decision process of the RF classifier.

5. Results

5.1. Simulation Data Analysis

5.1.1. Parameter Setting

The sampling rate of the simulation system was 48 kHz. The UWA channel in the simulation was cited from Reference [21]. Assuming the path number of channels was 10, the delay difference of adjacent paths followed an exponential distribution with an average of 2 ms, and the average multipath delay spread was 20 ms. In order to satisfy the

time-varying characteristics of the UWA channel, the multipath channel was generated randomly when the communication signal was generated. We can see from Figure 9 that when 100 communication signals were generated, 100 channels were randomly generated. Because the number of signals with different modulation modes under each SNR was 150, and we had 41 SNRs and 4 modulation modes, and the total number of signals with different modulation modes under all SNR was 24,600. The parameter setting of the simulated signals is shown in Table 1.

Table 1. Parameter setting of the simulated signals.

| Modulation Mode | Center Frequency | Bandwidth | CP Length | SNR Range | Number of Signals Under Each SNR |
|-----------------|------------------|-----------|-----------|---------------------------------|----------------------------------|
| OFDM | 12 kHz | 4 kHz | 20 ms | −20 dB–20 dB, 1 dB increment | 50 |
| | | | 30 ms | | 50 |
| | | | 40 ms | | 50 |
| 2FSK | 12 kHz | 3 kHz | - | −20 dB–20 dB, 1 dB increment | 50 |
| | | 4 kHz | | | 50 |
| | | 5 kHz | | | 50 |
| 4FSK | 12 kHz | 3 kHz | - | −20 dB–20 dB, 1 dB increment | 50 |
| | | 4 kHz | | | 50 |
| | | 5 kHz | | | 50 |
| 8FSK | 12 kHz | 3 kHz | - | −20 dB–20 dB, 1 dB increment | 50 |
| | | 4 kHz | | | 50 |
| | | 5 kHz | | | 50 |

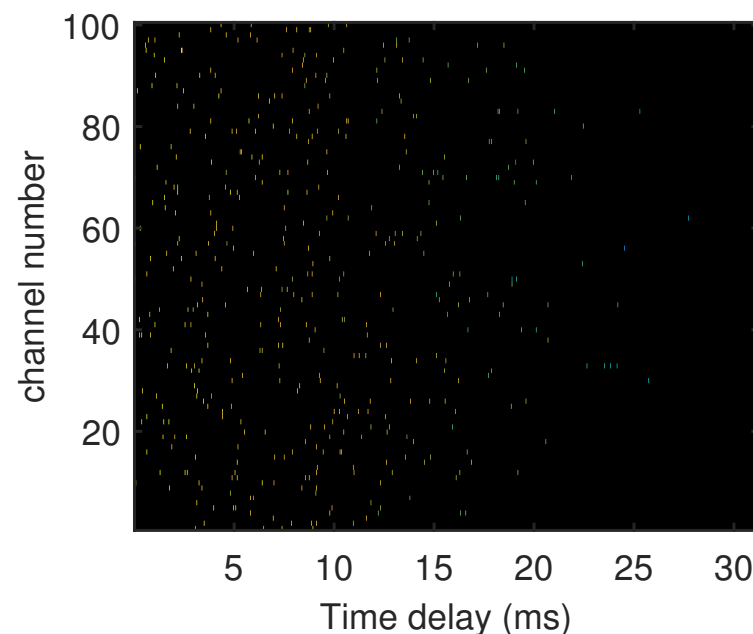


Figure 9. CIR of 100 channels.

5.1.2. Robustness Analysis of Spectral Peak Features

Figure 10 shows that these 15 local maximums changed with the increase of the SNR under the four modulation modes of OFDM, 2FSK, 4FSK, and 8FSK, and each local maximum in the figure was obtained by averaging 150 signals. We can find from Figure 10

that these extracted local maximums can remain stable with the increase of the SNR and have a different distribution.

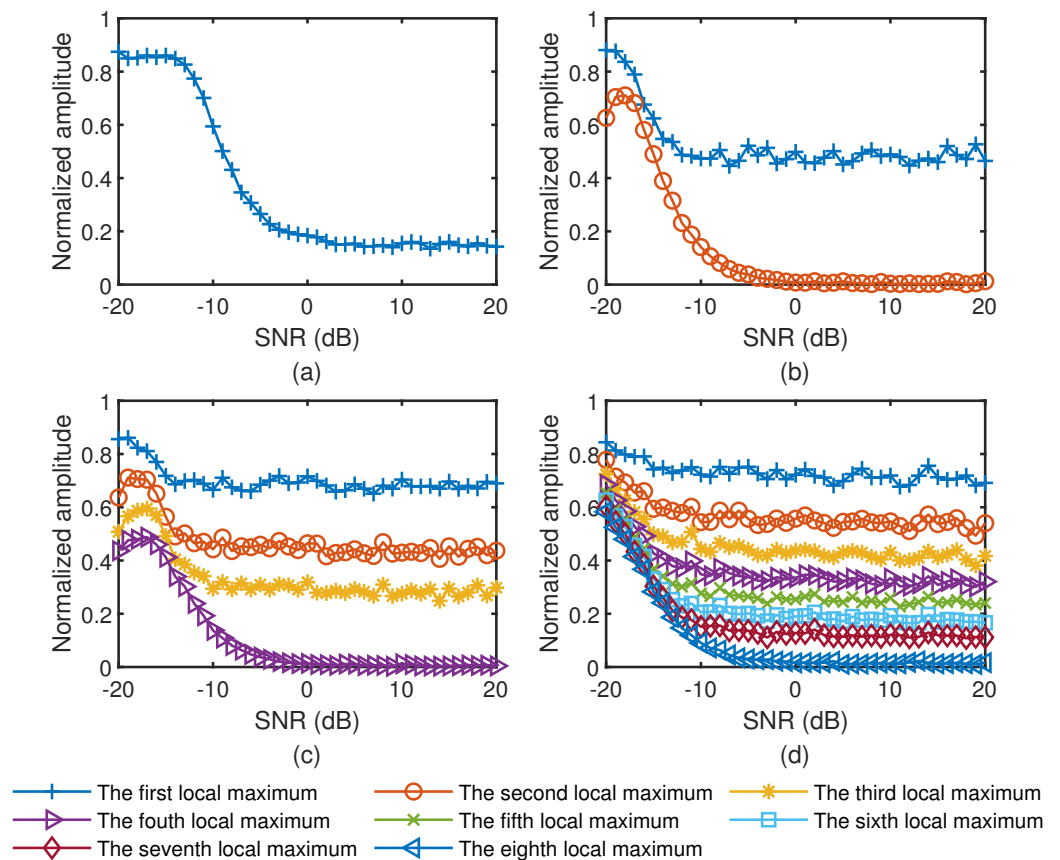


Figure 10. The variation of the extracted local maximum with the increase of the SNR. (a) OFDM, (b) 2FSK, (c) 4FSK, and (d) 8FSK.

We show the distribution of these 15 local maximums under different modulation modes by using boxplots [22] (Figure 11). Firstly, we found out the upper and lower edges, the median, and two quartiles of the local maximums in the feature set. Then, we connected two quartiles to form the box. Finally, we connected the upper and lower edges with the box. What we need to pay attention to is that the red line in the box is the median, and feature values outside the upper and lower edges are outliers. We can find from Figure 10 that these local maximums' distribution tended to be stable when $\text{SNR} \geq -10$ dB. Therefore, we only used the simulation data when $\text{SNR} \geq -10$ dB in Table 1 in the following simulation. We can find only the second local maximum of OFDM was small, while the second local maximum of other modulation modes was large from Figure 11a. We can find that the second and the third local maximum of 2FSK had different distributions compared to the other modulation modes and that the third local maximum tended to zero from Figure 11b,c. We can find that the second to the fifth local maximum of 4FSK had different distributions compared to the other modulation modes and that the fifth local maximum tended to zero from Figure 11d–g. We can find that the second to the ninth local maximum of 8FSK had different distributions compared to the other modulation modes and that the ninth local maximum tended to zero from Figure 11h–o. We can find that these extracted local maximums had different distributions under different modulation modes from the simulation results, as shown in Figures 10 and 11.

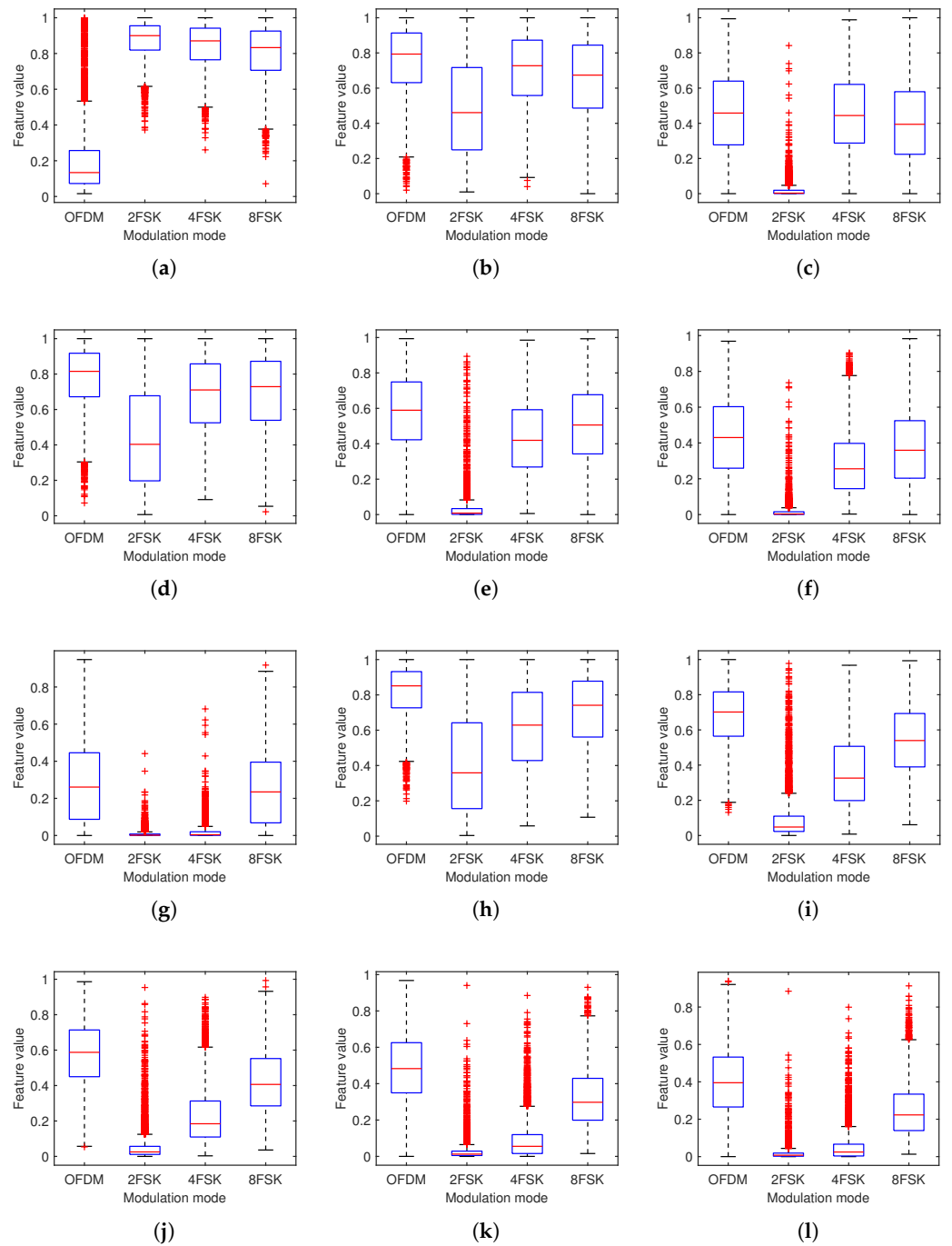


Figure 11. Cont.

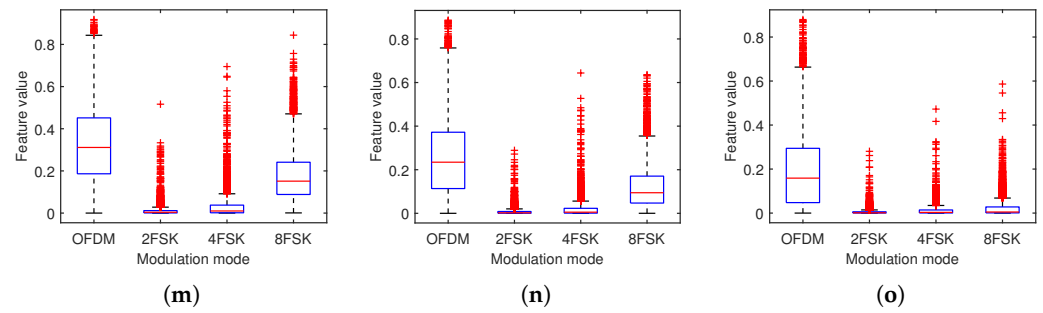


Figure 11. Distribution of 15 extracted features under different modulation modes. (a) The second local maximum of OFDM, (b) the second local maximum of 2FSK, (c) the third local maximum of 2FSK, (d) the second local maximum of 4FSK, (e) the third local maximum of 4FSK, (f) the fourth local maximum of 4FSK, (g) the fifth local maximum of 4FSK, (h) the second local maximum of 8FSK, (i) the third local maximum of 8FSK, (j) the fourth local maximum of 8FSK, (k) the fifth local maximum of 8FSK, (l) the sixth local maximum of 8FSK, (m) the seventh local maximum of 8FSK, (n) the eighth local maximum of 8FSK, and (o) the ninth local maximum of 8FSK.

5.1.3. Parameter Determination of the Classifier

We determined the parameters of the RF classifier based on the recognition rate. The sample in the simulation was 18,600 groups of signals when $\text{SNR} \geq -10$ dB in Table 1. The recognition rate was calculated by tenfold cross-validation. The classification process of the RF classifier needs to determine two parameters: the number of decision trees Q and the number of randomly selected features K .

Firstly, we determined parameter Q by setting $K = \log_2 F$ [20]. Figure 12a shows the influence of parameter Q on the recognition rate after 50 Monte Carlo simulations. We can find that the recognition rate tended to be stable with the increase of Q and that the highest recognition rate was about 95% when Q was higher than 60 from Figure 12a. Figure 12b shows the influence of parameter K on the recognition rate after 50 Monte Carlo simulations. We can find from Figure 12b that the recognition rate increased first and then decreased with the increase of K and that the highest recognition rate was 95.4% when $K = 8$. Finally, we determined $Q = 60$ and $K = 8$ according to the simulation results in Figure 12.

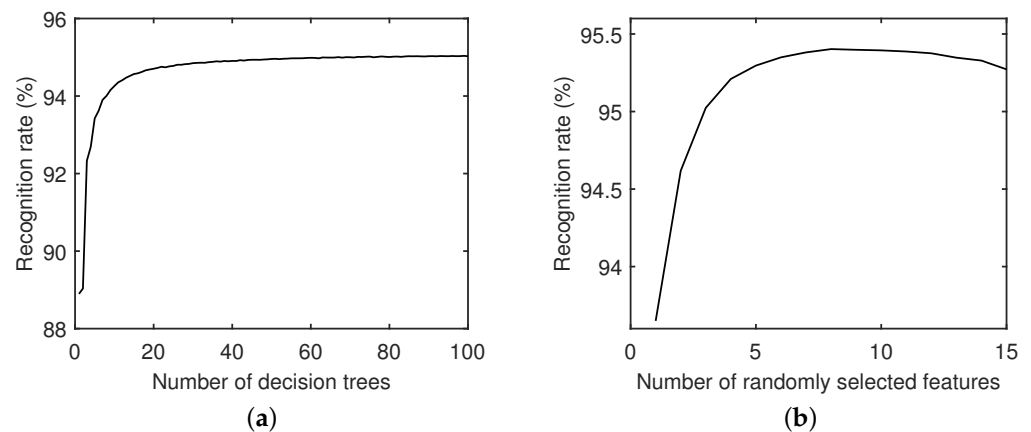


Figure 12. Parameter determination of the classifier. (a) Number of decision trees and (b) number of randomly selected features.

5.1.4. Performance Analysis of The Proposed Method

We analyzed the recognition performance of these extracted features combined with the SVM, KNN, and RF classifiers. The SVM, KNN, and RF classifiers were trained on

18,600 groups of signals with $\text{SNR} \geq -10$ dB in Table 1. The validation data were newly generated, and there were 1500 signals of different modulation modes under each SNR. The multipath channel was generated randomly when the communication signal was generated. The parameter setting of the validation set is shown in Table 2.

Table 2. Parameter setting of the validation set.

| Modulation Mode | Center Frequency | Bandwidth | CP Length | SNR Range | Number of Signals Under Each SNR |
|-----------------|------------------|-----------|-----------|---------------------------------|----------------------------------|
| OFDM | 12 kHz | 4 kHz | 20 ms | −20 dB–20 dB, 1 dB increment | 500 |
| | | | 30 ms | | 500 |
| | | | 40 ms | | 500 |
| 2FSK | 12 kHz | 3 kHz | - | −20 dB–20 dB, 1 dB increment | 500 |
| | | 4 kHz | | | 500 |
| | | 5 kHz | | | 500 |
| 4FSK | 12 kHz | 3 kHz | - | −20 dB–20 dB, 1 dB increment | 500 |
| | | 4 kHz | | | 500 |
| | | 5 kHz | | | 500 |
| 8FSK | 12 kHz | 3 kHz | - | −20 dB–20 dB, 1 dB increment | 500 |
| | | 4 kHz | | | 500 |
| | | 5 kHz | | | 500 |

The simulation results are shown in Figure 13, and we can find that RF classifier performed best from the simulation results. The poor performance of the KNN classifier was mainly due to the problem of insufficient performance when processing high-dimensional data [23], and the performance of the linear SVM classifier will decline when the variance of the feature set is large. Therefore, we focused on the performance of the RF classifier here. We found that the recognition rate of OFDM reached 90% when the SNR was higher than −6 dB and that the recognition rate of OFDM was stable at 98% when the SNR was higher than 5 dB from Figure 13a. We found that the recognition rate of 2FSK reached 93.4% when the SNR was higher than −9 dB and that the recognition rate of 2FSK was stable at 98% when the SNR was higher than −6 dB from Figure 13b. We found that the recognition rate of 4FSK reached 92.27% when the SNR was higher than −8 dB and that the recognition rate of 4FSK was stable at 96.5% when the SNR was higher than −3 dB from Figure 13c. We found that the recognition rate of 8FSK reached 92.47% when the SNR was higher than −6 dB and that the recognition rate of 8FSK was stable at 95% when the SNR was higher than 3 dB from Figure 13d. This was because the higher the recognition rate, the better the recognition performance of the classifier was. Therefore, we can know that the recognition performance of 8FSK was the worst and that of 2FSK was the best. Because the UWA channel in the simulation was randomly generated, the simulation results verified the effectiveness of the proposed recognition method. The simulation results also showed that the proposed spectral peak feature extraction method can reduce the influence of the UWA channel effectively.

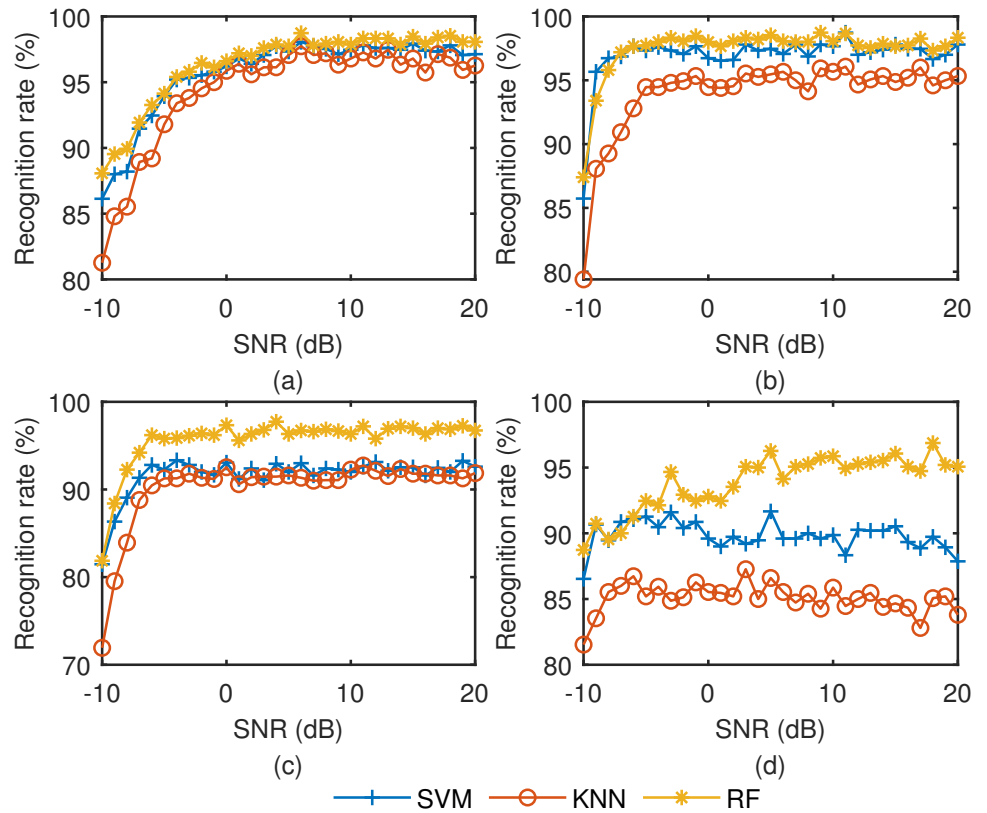


Figure 13. Comparison of the recognition rate under different classifiers. (a) OFDM, (b) 2FSK, (c) 4FSK, and (d) 8FSK.

5.2. Experimental Data Analysis

The experimental data were acquired in Bohai Sea. The depth of the transducer was about 10 m, the depth of the hydrophone about 22 m, and the distance between transducer and hydrophone about 1.5 km. The experimental diagram is shown in Figure 14a. The data collection equipment used in the experiment is shown in Figure 14b. The emission source level was about 170 dB, and the SNR of the received signal was about 12 dB in band. The CIR measured during the experiment is shown in Figure 14c. We can find from Figure 14c that there were obvious multipaths in the channel. The SVM, KNN, and RF classifiers were trained on 18,600 groups of data with SNR ≥ −10 dB in Table 1. The sampling frequency of the system was 100 kHz. The parameter setting of the experimental signals is shown in Table 3.

Table 3. Parameter setting of the experimental signals.

| Modulation Mode | Center Frequency | Bandwidth | CP Length | SNR Range | Number of Signals Under Each SNR |
|-----------------|------------------|-----------|-----------|-----------|----------------------------------|
| OFDM | 12 kHz | 4 kHz | 32 ms | 12 dB | 240 |
| 2FSK | 12 kHz | 4 kHz | - | 12 dB | 240 |
| 4FSK | 12 kHz | 4 kHz | - | 12 dB | 240 |
| 8FSK | 12 kHz | 4 kHz | - | 12 dB | 240 |

Table 4 shows the comparison of the recognition rates of the SVM, KNN, and RF classifiers, and we can see the classification performance of the RF classifier was the best among them from Table 4. The recognition rate of OFDM was 91.25%, that of 2FSK

100%, that of 4FSK 97.08%, and that of 8FSK 96.25%. We can find from the experimental results that the RF classifier trained on the simulation data realized a high recognition rate classification result on the experimental data, which fully proves the effectiveness of the proposed recognition method.

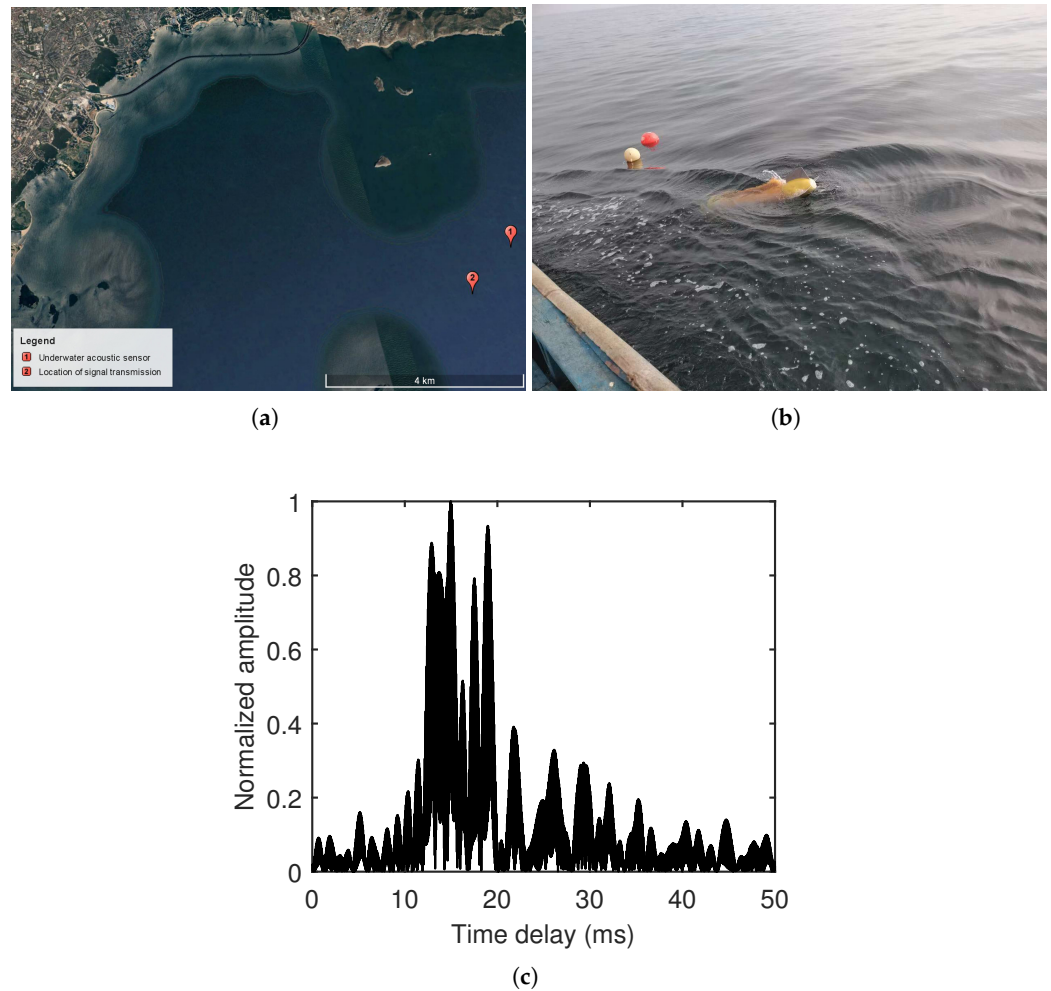


Figure 14. Experimental environment. (a) Experimental diagram, (b) data acquisition equipment, and (c) measured CIR.

Table 4. Classification results of the experimental data.

| Predicted \ Actual | OFDM | 2FSK | 4FSK | 8FSK |
|--------------------|---|---|---|---|
| OFDM | SVM: 217(90.42%) KNN: 208(86.67%) RF: 219(91.25%) | SVM: 1 KNN: 1 RF: 3 | SVM: 6 KNN: 7 RF: 4 | SVM: 16 KNN: 24 RF: 14 |
| 2FSK | SVM: 0 KNN: 0 RF: 0 | SVM: 240(100%) KNN: 231(96.25%) RF: 240(100%) | SVM: 0 KNN: 7 RF: 0 | SVM: 0 KNN: 2 RF: 0 |
| 4FSK | SVM: 0 KNN: 0 RF: 0 | SVM: 11 KNN: 13 RF: 3 | SVM: 224(93.33%) KNN: 222(92.50%) RF: 233(97.08%) | SVM: 5 KNN: 5 RF: 4 |
| 8FSK | SVM: 0 KNN: 5 RF: 2 | SVM: 4 KNN: 11 RF: 0 | SVM: 34 KNN: 58 RF: 7 | SVM: 202(84.17%) KNN: 166(69.17%) RF: 231(96.25%) |

6. Discussion

6.1. Significance of the Proposed Method

The spectral peak extraction method proposed in this work provides a new idea for UWA communication signal feature extraction. After preprocessing and waveform optimization, the spectral peak features can maintain high robustness in the UWA channel. The feature with high robustness can greatly improve the recognition rate of UWA communication signals.

The proposed spectral peak feature extraction method can be further applied to the estimation of the carrier frequency of the UWA FSK signal. The multipath of the UWA channel will lead to frequency-selective fading. Hence, the carrier frequency of the FSK signal is difficult to estimate based on the traditional method. The position of the frequency point corresponding to the spectral peak is the estimated frequency in the proposed method.

The designed RF classifier can greatly ensure the efficiency and performance of recognition. In practical applications, the recognition performance of the RF classifier will be higher with the addition of more experimental data to train the RF classifier.

The application scenario of the proposed method is to realize the recognition of non-cooperative UWA communication signals in sea areas. After realizing reliable recognition, parameter estimation and blind demodulation can be further achieved. Therefore, the proposed method can be used as the basis for subsequent research.

6.2. Limitations of the Proposed Method

The setting of the empirical values in the spectral peak feature extraction method proposed in this work is not applicable to FSK signals with all bandwidths. If the bandwidth of FSK signals is small or large, then the performance of the spectral peak extraction method may be reduced. In the practical application of the proposed method, the bandwidth of the signal should be estimated first. Then, the corresponding empirical value should be set according to the bandwidth to achieve the optimal performance of the proposed method.

7. Conclusions

This work presented an efficient recognition method for different modulation modes, including OFDM, 2FSK, 4FSK, and 8FSK, in the UWA channel. First, the autocorrelation and frequency-domain waveforms of the communication signal are easily affected by the UWA channel, resulting in the spectral peak feature instability. To solve this problem, we proposed a new spectral peak feature extraction method to obtain a high robustness feature, which contains pre-processing, waveform optimization, and feature extraction. Then, we designed an RF classifier that can meet the needs of efficient recognition. Finally, we realized effective recognition by combining the extracted features with the RF classifier. We verified the proposed method through simulation and experimental data. The simulation results showed that the recognition rate of OFDM reached 98% when the SNR was higher than 5 dB, that of 2FSK reached 98% when the SNR was higher than -6 dB, that of 4FSK reached 96.5% when the SNR was higher than -3 dB, and that of 8FSK reached 95% when the SNR was higher than 3 dB. Although the RF classifier was trained by the simulation data, the recognition rate of all modulation modes can still reach more than 90% according to the experimental results.

Author Contributions: Conceptualization, T.F.; methodology, T.F. and S.L.; software, T.F.; validation, T.F. and L.Z. and Q.W.; formal analysis, Q.W.; investigation, L.Z.; resources, T.F.; data curation, T.F.; writing—original draft preparation, T.F.; writing—review and editing, S.L.; visualization, Q.W.; supervision, T.F.; project administration, T.F.; funding acquisition, S.L. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported in part by the National Natural Science Foundation of China under Grant Nos. 11974090 and 11774074 and in part by the Ph.D. Student Research and Innovation Fund of the Fundamental Research Funds for the Central Universities under Grant No. 3072021GIP0505.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this paper are available after contacting the corresponding author.

Acknowledgments: The authors would like to thank the anonymous Reviewers for their careful reading and valuable comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Petrioli, C.; Petrocchia, R.; Potter, J.R.; Spaccini, D. The SUNSET framework for simulation, emulation and at-sea testing of underwater wireless sensor networks. *Ad Hoc Netw.* **2015**, *34*, 224–238. [[CrossRef](#)]
2. Toso, G.; Masiero, R.; Casari, P. Field Experiments for Dynamic Source Routing: S2C EvoLogics modems run the SUN protocol using the DESERT Underwater libraries. In Proceedings of the MTS/IEEE Oceans Conference, Hampton Roads, VA, USA, 14–19 October 2012; pp. 1–10.
3. Fang, T.; Liu, S.Z.; Ma, L.; Zhang, L.Y. Subcarrier modulation identification of underwater acoustic OFDM based on block expectation maximization and likelihood. *Appl. Acoust.* **2021**, *173*, 107654. [[CrossRef](#)]
4. Fang, T.; Liu, S.Z.; Wu, X.B.; Yan, H.L. Non-cooperative MPSK Modulation Identification in SIMO Underwater Acoustic Multipath Channel. In Proceedings of the 2021 OES China Ocean Acoustics (COA), Harbin, China, 14–17 July 2021; pp. 1–6.
5. Ren, H.; Yu, J.L.; Wang, Z.X.; Chen, J. Modulation format recognition in visible light communications based on higher order statistics. In Proceedings of the 2017 Conference on Lasers and Electro-Optics Pacific Rim, Sands Expo and Convention Centre, Singapore, 31 July–4 August 2017; pp. 1–2.
6. Wang, L.X.; Guo, S.T.; Jia, C.J. Modulation format recognition in visible light communications based on higher order statistics. In Proceedings of the 2016 7th IEEE International Conference on Software Engineering and Service Science, Beijing, China, 26–28 August 2016; pp. 627–630.
7. Li, S.J.; Wang, Y.W. Method of modulation recognition of typical communication satellite signals based on cyclostationary. In Proceedings of the 2013 ICME International Conference on Complex Medical Engineering, Beijing, China, 25–28 May 2013; pp. 268–273.
8. Wei, Y.J.; Fang, S.L.; Wang, L.Y. Automatic Modulation Classification of Digital Communication Signals Using SVM Based on Hybrid Features, Cyclostationary, and Information Entropy. *Entropy* **2019**, *21*, 745. [[CrossRef](#)] [[PubMed](#)]
9. Chen, E.; Yan, J.Q.; Sun, H.X. Research on MPSK Modulation Classification of Underwater Acoustic Communication Signals. In Proceedings of the 2016 IEEE/OES China Ocean Acoustics (COA), Harbin, China, 9–11 January 2016; pp. 1–5.
10. Jiang, W.H.; Chen, D.S.; Wu, Y.Q. Modulation recognition of underwater acoustic OFDM signals based on the correlation property of the cyclic prefix. *Appl. Acoust.* **2016**, *35*, 42–49.
11. Jiang, N.; Wang, B. Underwater Communication Signals' Modulation Recognition Based on Sparse Autoencoding Network. *J. Signal Process.* **2019**, *35*, 103–114.
12. Jiang, W.H.; Tong, F.; Wang, B. Modulation recognition of non-cooperation underwater acoustic communication signals using principal component analysis. *Appl. Acoust.* **2016**, *37*, 1670–1676. [[CrossRef](#)]
13. Wang, M.C.; Liu, J.; Zhang, J.W. Modulation format identification based on phase statistics in Stokes space. *Opt. Commun.* **2020**, *480*, 1–7. [[CrossRef](#)]
14. Burges, C.J.C. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [[CrossRef](#)]
15. Zhu, Z.C.; Aslam, M.W.; Nandi, A.K. Augmented Genetic Programming for automatic digital modulation classification. In Proceedings of the 2010 IEEE International Workshop on Machine Learning for Signal Processing, Kittila, Finland, 29 August–1 September 2010; pp. 1–6.
16. Yao, X.H.; Yang, H.H.; Li, Y.Q. Modulation Recognition of Underwater Acoustic Communication Signals based on Convolutional Neural Networks. *Unmanned Syst. Technol.* **2018**, *1*, 68–74.
17. Rakshit, M.; Das, S. An efficient wavelet-based automated R-peaks detection method using Hilbert transform. *Biocybern. Biomed. Eng.* **2017**, *37*, 566–577. [[CrossRef](#)]
18. Yuan, L.; Yang, Y.; Hernandez, A. Novel Adaptive Peak Detection Method for Track Circuits Based on Encoded Transmissions. *IEEE Sens. J.* **2018**, *18*, 6224–6234. [[CrossRef](#)]
19. Chen, Y.; Yang, K.; Liu, H.L. Self-Adaptive Multi-Peak Detection Algorithm for FBG Sensing Signal. *IEEE Sens. J.* **2016**, *16*, 2658–2665. [[CrossRef](#)]
20. Zhou, Z.H. *Machine Learning*; Tsinghua University Press: Beijing, China, 2016; pp. 73–95, 178–180.
21. Berger, C.R.; Zhou, S.L.; Preisig, J.C. Sparse Channel Estimation for Multicarrier Underwater Acoustic Communication: From Subspace Methods to Compressed Sensing. *IEEE Trans. Signal Process.* **2010**, *58*, 1708–1721. [[CrossRef](#)]

-
22. Fang, L.; Pascal, D.; Olivier, D. Acoustic Resonance Detection Using Statistical Methods of Voltage Envelope Characterization in Metal Halide Lamps. *IEEE Trans. Ind. Appl.* **2017**, *53*, 5988–5996.
 23. Friedman, J.H.; Baskett, F; Shustek L.J. An Algorithm for Finding Nearest Neighbors. *IEEE Trans. Comput.* **1975**, *C-24*, 1000–1006. [[CrossRef](#)]