

Article

A Comparative Study of Multiscale Sample Entropy and Hierarchical Entropy and Its Application in Feature Extraction for Ship-Radiated Noise

Weijia Li ¹, Xiaohong Shen ^{1,2,*} and Yaan Li ^{1,*}

¹ School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China

² Key Laboratory of Ocean Acoustics and Sensing (Northwestern Polytechnical University), Ministry of Industry of Information Technology, Xi'an 710072, China

* Correspondence: xhshen@nwpu.edu.cn (X.S.); liyaan@nwpu.edu.cn (Y.L.); Tel.: +86-29-8849-5817 (Y.L.)

Received: 2 July 2019; Accepted: 13 August 2019; Published: 14 August 2019



Abstract: The presence of marine ambient noise makes it difficult to extract effective features from ship-radiated noise. Traditional feature extraction methods based on the Fourier transform or wavelets are limited in such a complex ocean environment. Recently, entropy-based methods have been proven to have many advantages compared with traditional methods. In this paper, we propose a novel feature extraction method for ship-radiated noise based on hierarchical entropy (HE). Compared with the traditional entropy, namely multiscale sample entropy (MSE), which only considers information carried in the lower frequency components, HE takes into account both lower and higher frequency components of signals. We illustrate the different properties of HE and MSE by testing them on simulation signals. The results show that HE has better performance than MSE, especially when the difference in signals is mainly focused on higher frequency components. Furthermore, experiments on real-world data of five types of ship-radiated noise are conducted. A probabilistic neural network is employed to evaluate the performance of the obtained features. Results show that HE has a higher classification accuracy for the five types of ship-radiated noise compared with MSE. This indicates that the HE-based feature extraction method could be used to identify ships in the field of underwater acoustic signal processing.

Keywords: underwater signal processing; feature extraction; multiscale sample entropy (MSE); hierarchical entropy (HE); ship-radiated noise

1. Introduction

Identification and classification of marine vehicles are important in the field of underwater signal processing, as they are of great value in the military and marine economy [1–4]. An important aspect of the ship classification problem is to extract effective features from received signals. Features extracted from a signal are the representation of part of the signal's characteristics. Insufficient characteristic reflection will lead to low accuracy of classification. Therefore, there is a great need for the development of feature extraction methods in the field of underwater signal processing.

The traditional feature extraction method is based on the frequency domain. There are many studies devoted to extracting the spectral characteristics of signals, such as the analysis of the power spectral density of signals [5]. However, studies show that traditional methods have shortcomings and limitations in practical applications. For example, the traditional spectrum-based method is based on the assumption of the linearity of the signals, which means the features extracted using this method will miss the signal's nonlinear characteristics [6]. In this paper, we use entropy as a feature extraction method, which is based on the time domain and quantifies the complexity of the signal as a feature.

Entropy, as a feature extraction method measuring the system's complexity in the time domain, has been maturely applied to fault diagnosis and pathological signal detection [7–11]. Pincus proposed the concept of approximate entropy (AE) based on the theory of Shannon entropy in 1991 [12]. However, the AE has self-matching terms in the calculation, which leads to bias of the result. This kind of bias results in two disadvantages in the calculation of AE. One is that the computation of AE is overly dependent on the length of data, and the other is the lack of correlation between the AE result and the signal complexity. Thus, in 2000, Richman and Moorman proposed an improvement of the AE, which is the sample entropy (SE) [13]. It solved the consistency problem in AE, and in the subsequent research, the fast sample entropy was proposed, which simplified the SE's calculation [14]. However, in pathological research, a single scale cannot illustrate the whole information carried in signals. To distinguish different kinds of pathological signals and calculate the complexity of interested signals more accurately, multiscale sample entropy (MSE) based on the coarse-graining process [15–17] and hierarchical entropy (HE) based on hierarchical decomposition [18,19] have been proposed. HE, as a method improving MSE, is capable of roller bearing fault diagnosis [20]. Compared with the MSE, which only considers the lower frequency components of signals in the calculation, HE that retains both the lower frequency and higher frequency components of signals can better recognize different pathological signals in practical applications.

The calculation of entropy on a single scale only takes into account the temporal information in the signal. Although it has many advantages such as simple calculation, sometimes it cannot reflect the complexity differences between different signals accurately. Many previous research works applied the coarse-graining process to entropy. This improvement can describe the complexity of signals at different scales. For example, Li proposed a method of extracting the features of ship-radiated noise combined with variational mode decomposition (VMD) and multiscale permutation entropy (MPE) in 2017 [21]. Yang combined VMD with fluctuation-based dispersion entropy [22]. Chen proposed a new method based on permutation entropy and coarse-graining [23]. Shashidhar applied MSE to weak signal detection problems [24]. All of the above studies have proven that entropy based on multiple scales has certain applicability in feature extraction of underwater acoustic signals. However, they did not consider the high-frequency components in the signal. In other words, much useful information may be missed at a high frequency. Meanwhile, research showed that the lower frequency components of ambient noise have increased over the past few decades [25]. This makes it more difficult to deal with the detection and feature extraction of ship-radiated noise. If the lower and higher frequency components of the signal can be separately analyzed when extracting the features, we may get more complete information contained in the signal.

In this paper, HE is used as a novel feature extraction method for ship-radiated noise. It has great advantages compared with methods such as MSE, preserving both the low-frequency and high-frequency components of the signal while performing multi-scale decomposition and calculating the complexity of the signals of interest. Hence, HE describes the signal characteristics more accurately. Several sets of simulation signals were used to compare the difference between HE and MSE in identifying different types of signals, verifying that the HE has good recognition ability, especially for signals with similar low-frequency components and different high-frequency components. For those signals with different low frequency components and similar high frequency components, since HE also considers the low-frequency components of the signal, the actual results are comparable to those of MSE. At the end of this paper, five different types of ship-radiated noises are presented, using SE, MSE, and HE for feature extraction, respectively. In order to compare the performance of the features extracted by different methods more clearly, we will pass the different features through a probabilistic neural network and criticize the performance of different features through the accuracy of classification.

The rest of this paper organized as follows: Section 2 introduces the concept of SE, MSE, and HE. In Section 3, the proposed method is applied to the simulated signal to show the properties of HE and

MSE. In Section 4, five types of ship-radiated noise are given to reflect the difference between the two feature extraction methods. Finally, Section 5 is the conclusion.

2. Basic Theory

2.1. Sample Entropy

Sample entropy quantifies a system’s degree of regularity by calculating the negative natural logarithm of conditional probability. It was developed by Richman and Moorman in 2000. Compared with approximate entropy, sample entropy eliminates the bias caused by self-matching. Meanwhile, it also reduces the computational time. Given a time series $\{x(i) : 1 \leq i \leq N\}$, N is the length of the original time series. $\{x(i) : 1 \leq i \leq N\}$ can be reconstructed into a set of sequences as follows: $X(i) = [x(i), x(i + 1), \dots, x(i + m - 1)] : 1 \leq i \leq N - m + 1$, where m is the embedding dimension. According to m and $N + m - 1$ sequences, which were obtained above, the distance $d[X(i), X(j)]$ between any two vectors can be defined, abbreviated as $D^m(i)$:

$$D^m(i) = d[X(i), X(j)] = \max[|x(i+k) - x(j+k)|] : 0 \leq k \leq m - 1; 1 \leq i, j \leq N - m + 1; i \neq j. \quad (1)$$

Since the time series $\{x(i) : 1 \leq i \leq N\}$ has already been given, the standard deviation (SD) of the time series can be readily obtained. Set $r = 0.1 SD \sim 0.25 SD$ to be the threshold, with the distance $d[X(i), X(j)]$. The formula of $B_i^m(r)$ is given by:

$$B_i^m(r) = \frac{1}{N - m} \{\text{the number of } d[X(i), X(j)] \leq r\}. \quad (2)$$

Equation (2) computes the probability of the distance between $X(i)$ and the remaining sequences within the threshold r . Moreover, the average of $B_i^m(r)$ can be obtained by Equation (3):

$$B^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} B_i^m(r). \quad (3)$$

Increasing the embedding dimension m to $m + 1$, then $B_i^{m+1}(r)$ can be analogously obtained by repeating the previous steps. Finally, the sample entropy ($SampEn(m, r, N)$) is given by the following equation:

$$SampEn(m, r, N) = -\ln[B^{m+1}(r)/B^m(r)]. \quad (4)$$

In order to better understand the calculation process of sample entropy, we briefly describe it through Figure 1.

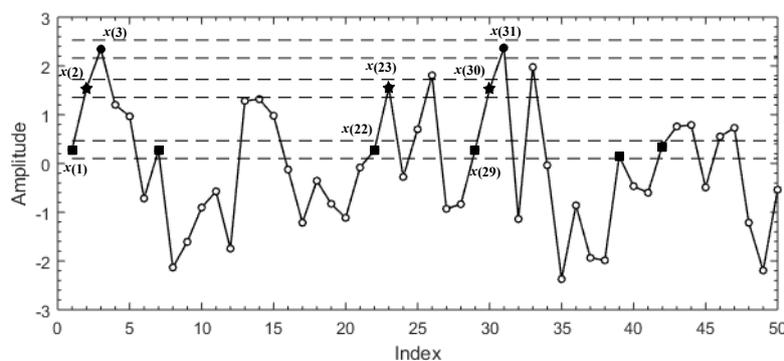


Figure 1. Stimulation signal.

A time series $\{x(i) : 1 \leq i \leq 50\}$ is given to illustrate the process for calculating $SampEn(m, r, N)$. We specify $m = 2$ and $r = 0.15 SD$. The horizontal dashed lines around $x(1)$, $x(2)$, and $x(3)$ represent

$x(1) \pm r$, $x(2) \pm r$, and $x(3) \pm r$, respectively. If the absolute difference between any two points is less than r , these two points match each other; also, it can be viewed as “indistinguishable”. In Figure 1, all of the points that match $x(1)$, $x(2)$, and $x(3)$ are represented with same symbol, respectively. Let $\{x(1), x(2)\}$ and $\{x(1), x(2), x(3)\}$ be a template sequence with two points and three points, respectively. Throughout $\{x(i) : 1 \leq i \leq 50\}$, there are two sequences $\{x(22), x(23)\}$ and $\{x(29), x(30)\}$ that match the template sequence $\{x(1), x(2)\}$. As for template sequence $\{x(1), x(2), x(3)\}$, there is only one sequence $\{x(29), x(30), x(31)\}$ that matches it. Count the number of the sequences that match $\{x(1), x(2)\}$ and $\{x(1), x(2), x(3)\}$. Repeat the previous steps for the next two-point sequence $\{x(2), x(3)\}$ and three-point sequence $\{x(2), x(3), x(4)\}$. Sum the number of sequences that match two-point and three-point sequence $\{x(2), x(3)\}$ and $\{x(2), x(3), x(4)\}$. Add them to the previous values that we already obtained. Repeat the same work mentioned above until all other possible template sequences ($\{x(1), x(2), x(3)\}, \dots, \{x(48), x(49), x(50)\}$) are considered. The ratio between the sum of two-point template matches and the sum of three-point template matches can be obtained. Therefore, $SampEn(m, r, N)$ is the natural logarithm of this ratio.

The value of $SampEn(m, r, N)$ is related to the parameters m and r . Therefore, the choices of these two parameters are also very important. According to Chen’s research [26], m is set to be one or two, and $r = 0.1 SD \sim 0.25 SD$ under most circumstances.

2.2. Multiscale Sample Entropy

Although SE has many advantages, in some circumstances, it cannot reflect the complexity differences between different signals accurately. The structure of signals generated from complex systems exhibits multiple temporal scale characteristics in the actual ocean environment. SE, as a single-scale-based method, does not account for the interrelationship between entropy and multiple scales. In order to overcome this shortage, Costa et al. developed the concept of multiscale sample entropy [15]. MSE can be viewed as SE with a coarse-graining process for the time series [27]. The coarse-graining process is based on averaging the samples inside moving, but non-overlapping windows. For a given time series $\{x(i) : 1 \leq i \leq N\}$, the coarse-graining process is denoted as:

$$y^{(n)} = \frac{1}{n} \sum_{j=1}^n x(ni - n + j) : 1 \leq i \leq N_n, \tag{5}$$

where N is the length of the time series and $N_n = \lfloor \frac{N}{n} \rfloor$ stands for the largest integer no greater than $\frac{N}{n}$. Hence, MSE at scale n is obtained by calculating the sample entropy of $y^{(n)}$. The MSE focuses on lower frequency components of a time series. However, it ignores the information contained in the higher frequency components of the signal. This problem leads to the development of hierarchical entropy.

2.3. Hierarchical Entropy

Hierarchical entropy (HE) is an algorithm quantifies the “complexity” of a time series based on SE and hierarchical decomposition. Unlike MSE, hierarchical decomposition takes both higher and lower frequency components of a time series into consideration [18]. Specifically, for a given time series, $x = \{x(i) : 1 \leq i \leq 2^n\}$. The definition of two operators Q_0 and Q_1 is as follows:

$$Q_0(x) = \left(\frac{x(2i-1) + x(2i)}{2} : 1 \leq i \leq 2^{n-1} \right), \tag{6}$$

$$Q_1(x) = \left(\frac{x(2i-1) - x(2i)}{2} : 1 \leq i \leq 2^{n-1} \right), \tag{7}$$

$Q_0(x)$ and $Q_1(x)$ are respectively the lower and higher frequency component of time series x , and their scale is two and their length 2^{n-1} . As a matter of fact, x can be reconstructed from $Q_0(x)$ and $Q_1(x)$.

$$x = Q_0(x)_j + Q_1(x)_j, Q_0(x)_j - Q_1(x)_j : 1 \leq j \leq 2^{n-1}. \tag{8}$$

$Q_0(x)_j$ and $Q_1(x)_j$ stand for the j^{th} value in $Q_0(x)$ and $Q_1(x)$, respectively. Thus, $Q_0(x)$ and $Q_1(x)$ constitute the two-scale hierarchical decomposition of the time series x .

After we obtain $Q_0(x)$ and $Q_1(x)$, each of them can also be decomposed by Q_0 and Q_1 . Consequently, we can get the hierarchical decomposition of the time series X at a scale of three. A tree graph can clearly show the relationship between each hierarchical component of the time series X in Figure 2.

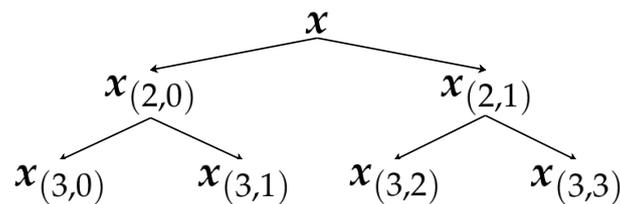


Figure 2. Hierarchical decomposition of the signal with three scales.

After the hierarchical decomposition, several sub-signals $x_{(n,e)}$ can be obtained, where n represents the scale and e stands for the e^{th} sub-signal at scale n . Calculate the SE for each sub-signal, and the HE result of X is obtained. It is important to choose the appropriate scales in different circumstances. On the one hand, high scales usually lead to computational redundancy. On the other hand, low scales may have insufficient accuracy in $SampEn(m, r, N)$'s computation.

3. Simulation Analysis of Different Signals Based on Hierarchical Entropy and Multiscale Sample Entropy

In this section, MSE and HE are compared using different simulation signals in order to illustrate their different characteristics. Before the simulation analysis, there are some previous steps that need to be done. In this paper, all the SE calculation's parameters are the same, which is $m = 2$, $r = 0.15 SD$, and the length of the data is at least 512 points for every SE calculation. In this part, the content is divided into the following subsections. First, we prove that the parameters chosen when calculating SE are appropriate. Second, three different orders of AR signals with different complexity are used to prove that HE is an effective measure of complexity. Third, different simulation signals are constructed, and their results of HE and MSE are compared. The results show that MSE pays more attention to the low-frequency components of the signal, and HE not only retains the information of the low-frequency components of the signal, but also retains the information of the high-frequency components of the signal. Finally, considering the noise interference in practical applications, this paper compares the robustness of the two methods to noise.

3.1. Parameter Selection for Sample Entropy

Both HE and MSE are based on SE. When we calculate the SE for a signal, it is important to choose the appropriate m and r . Since our main purpose is using entropy as a feature extraction method for ship-radiated noise, the simulation signals in this subsection are set as follows:

$$\begin{cases} S_1(n) = \sin(2\pi * 50n) + N(n), \\ S_2(n) = \sin(2\pi * 13n) + N(n), \end{cases} \quad (9)$$

In Equation (9), $S_1(n)$ and $S_2(n)$ are two sinusoidal signals mixed with Gaussian white noise. We use the sinusoidal signal in order to simulate the periodic signal produced by the ship engine or propeller. Meanwhile, Gaussian white noise is used to simulate the ambient noise. Since the composition of the ship-radiated noise is very complex, including ambient noise, cavitation noise, and signals produced by propellers and the engine, we simplify the model of ship-radiated noise as Equation (9). The signal-to-noise ratio (SNR) is set to be 5 dB, $m = 2$, and $r = 0.15 SD$. To demonstrate the impact of different data lengths on the calculation results, we calculated 60 sets of SE results with different

lengths of the two signals, each with 30 results. The data length increased from 150 equal intervals to 3150. The result is shown in Figure 3.

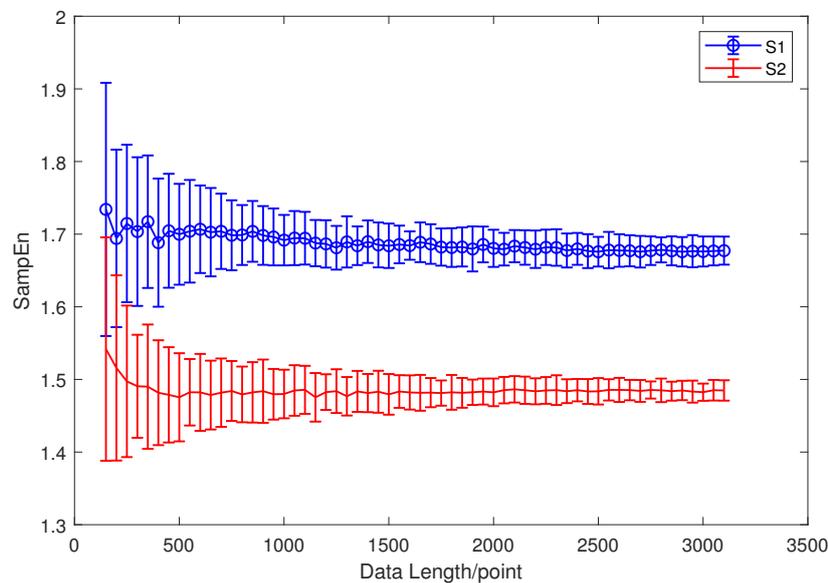


Figure 3. $S_1(n)$ and $S_2(n)$'s SE results with different lengths of data.

In Figure 3, as the length of the calculated sample entropy data increases, the results of the calculation become gradually stable. When the data length is too short, the SE results are too unstable to distinguish the sinusoidal signals of two different frequencies very well. Although the result becomes more stable as the data length increases, due to the consideration of the amount of calculation, when calculating the sample entropy in the paper, the data length is unified to 512. When we calculate HE in this paper, since the data length is 8192 points, we decompose the signal into a scale of five and guarantee that the SE's calculation that is contained in HE is at least 512 points.

After selecting the appropriate data length, the same simulated signals in Equation (9) are used to choose the value of m and r . The length of signal is set to be 512 points when calculating SE. The result is displayed in Figures 4 and 5.

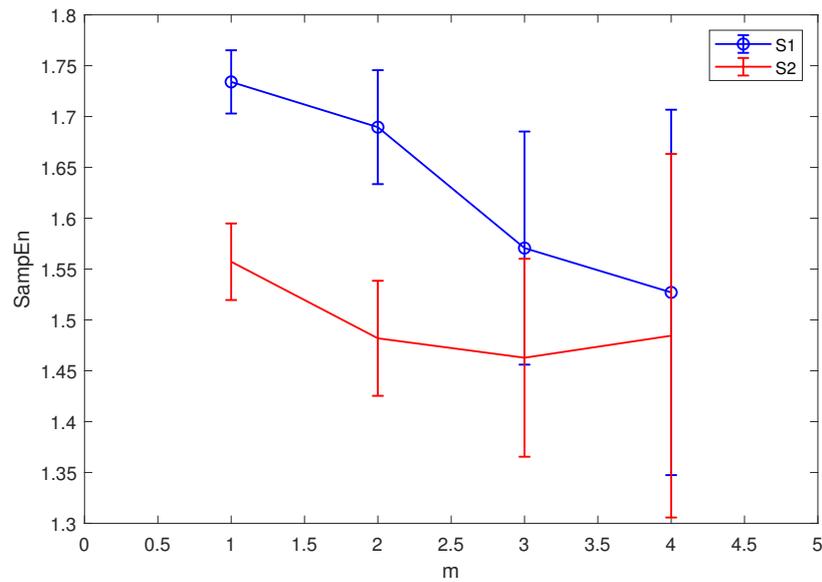


Figure 4. $S_1(n)$'s and $S_2(n)$'s SE results with different m .

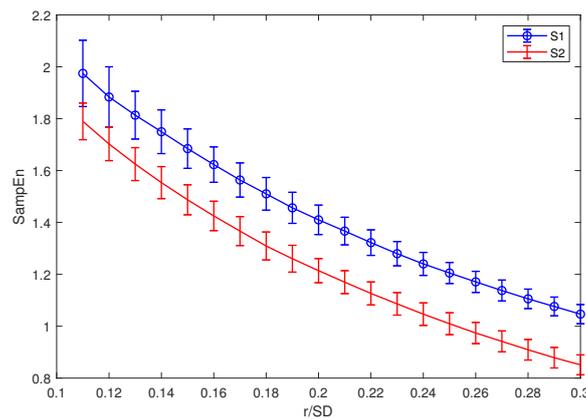


Figure 5. $S_1(n)$'s and $S_2(n)$'s SE results with different r .

From Figures 4 and 5, the result of SE is too close to distinguish two signals when $m = 3$, and it becomes unstable when m is larger than three, so we set $m = 2$ in this paper. As for r , the value of r has little effect on the stability of the results, so we set $r = 0.15 * SD$. The same parameters are discussed using the real ship-radiated noise employed in this paper [28], further verifying the conclusion in this section. The results are demonstrated in Figure 6. For some certain types of ship-radiated noise, SE cannot distinguish them very well according to Figure 6. This is why we need to introduce HE as a new feature extraction method to help us distinguish different signals.

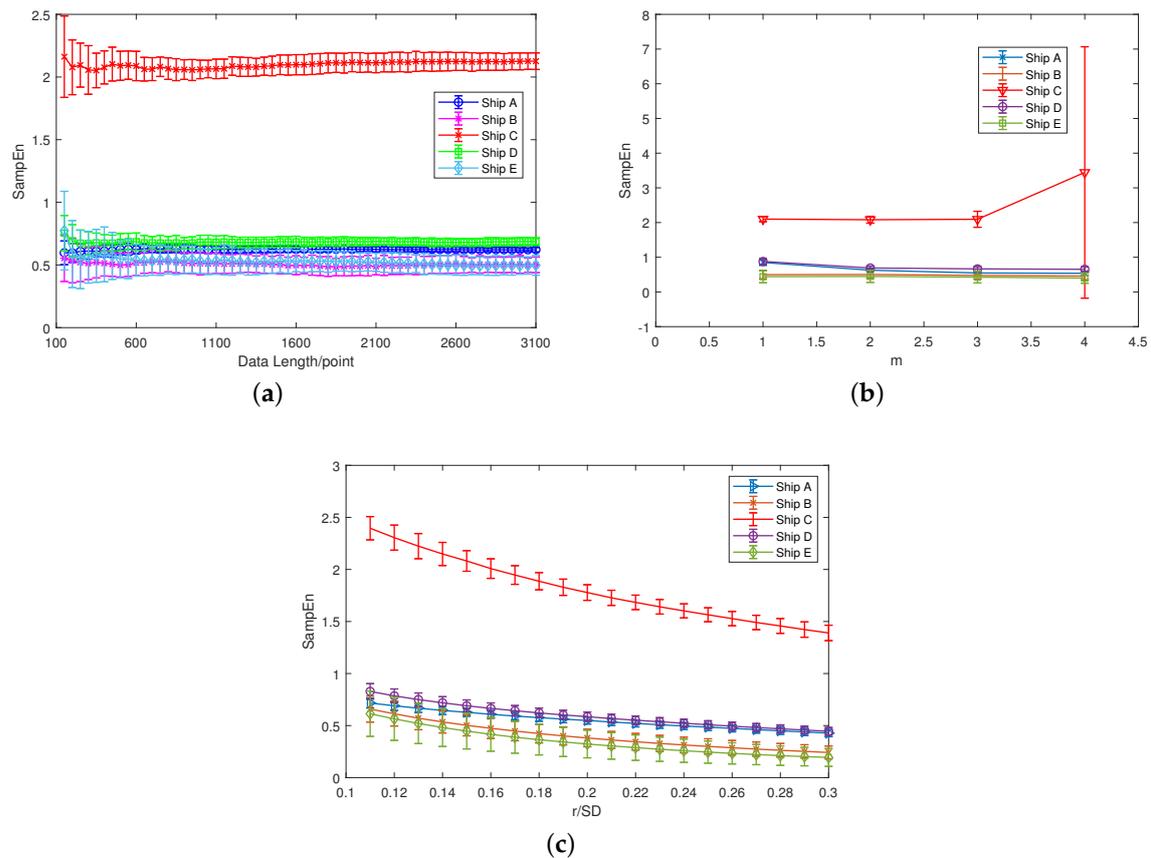


Figure 6. SE results for five types of ship-radiated noise with different parameters. (a) SE results with different data length. (b) SE results with different m . (c) SE results with different r .

3.2. Hierarchical Entropy Analysis for the AR Process

Three autoregressive processes (AR) with different orders will be given to demonstrate that HE is an effective method for measuring the complexity of different signals. The AR time series are given by:

$$AR_p(t) = \sum_{i=1}^p \alpha_i AR(t-i) + n(t), \tag{10}$$

where $n(t)$ is the Gaussian white noise with a standard normal distribution. The length of each AR process is 2^{13} . p indicates the order of the AR process, and α_i is the correlation coefficients. The value of α_i in each AR process is given in Table 1 according to [29].

Table 1. The correlation coefficients for generating AR processes.

	α_1	α_2	α_3	α_4	α_5	α_6	α_7
AR(1)	0.5	-	-	-	-	-	-
AR(4)	0.5	0.25	0.125	0.0625	-	-	-
AR(7)	0.5	0.25	0.125	0.0625	0.0313	0.0156	0.0078

The HE results of three AR time series are illustrated in Figure 7; $HE(n, e)$ stands for the e^{th} component of hierarchical entropy at scale n , and this abbreviation is used throughout this paper.

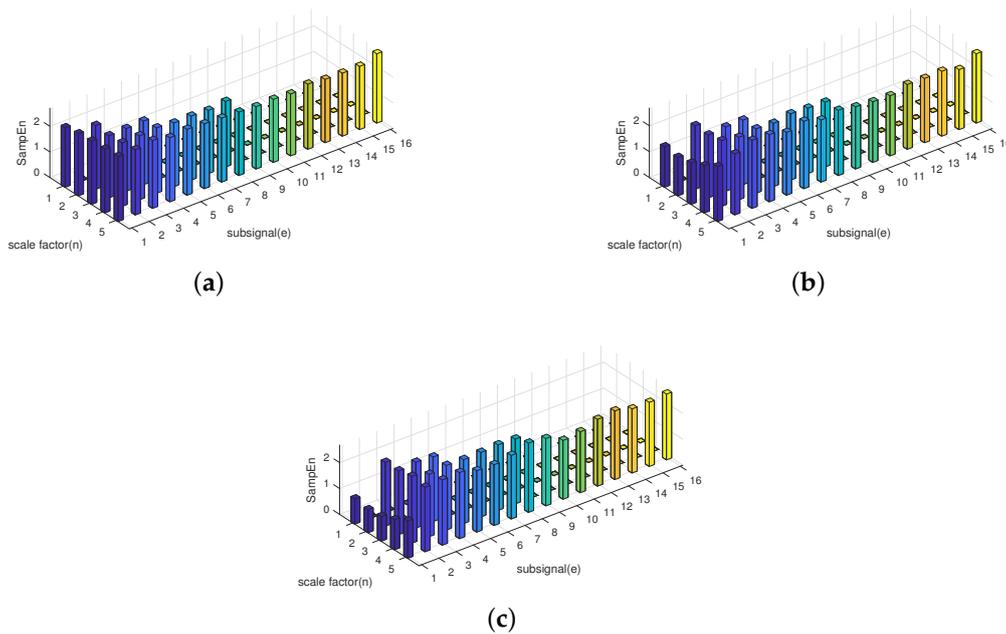


Figure 7. Hierarchical entropy results of AR(1), AR(4), and AR(7). (a) HE results for AR(1). (b) HE results for AR(4). (c) HE results for AR(7).

The AR process specifies that the output value is linearly dependent on its own previous and random terms. The dependence of the output value on the previous terms increases as the order p increases. Furthermore, as the order p increases, the correlation of the signal increases accordingly, making the model more predictable [23,29]. That is, the complexity of $AR(p + 1)$ is lower than that of $AR(p)$. Based on this idea, the value of HE should be negatively correlated to order p . Figure 7 depicts that the sample entropy of lower frequency components decreases while the order p of the AR time series increases. Hence, HE can be confirmed as an effective method for measuring the complexity of different time series.

3.3. Properties for Multiscale Sample Entropy

In this section, a set of simulation signals is employed to demonstrate the properties for MSE, which is focused on the lower frequency components of the signal. This property leads to the result that MSE performs well in distinguishing the signals with different low-frequency components. In order to highlight these properties of MSE, a set of signals is given as follows:

$$f_1(n) = \begin{cases} \sin(2\pi * 5n) : 1 \leq n \leq (2^{13} - 2^{10}), \\ \sin(2\pi * 60n) : (2^{13} - 2^{10}) + 1 \leq n \leq 2^{13}. \end{cases} \tag{11}$$

$$f_2(n) = \begin{cases} \sin(2\pi * 15n) : 1 \leq n \leq (2^{13} - 2^{10}), \\ \sin(2\pi * 60n) : (2^{13} - 2^{10}) + 1 \leq n \leq 2^{13}. \end{cases} \tag{12}$$

The lower frequency components of $f_1(n)$ and $f_2(n)$ are different, while the high-frequency components are the same. The waveform of $f_1(n)$ and $f_2(n)$ is shown in Figure 8. According to the theory of MSE, MSE should be able to distinguish between the two signals very well since the difference between the two signals is mainly in the lower frequency components. Figure 9 is the MSE result for $f_1(n)$ and $f_2(n)$ from a scale of 1–15.

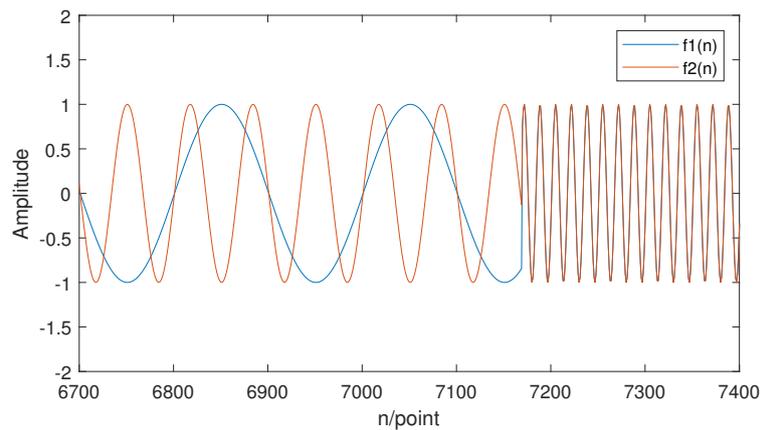


Figure 8. The waveform of $f_1(n)$ and $f_2(n)$.

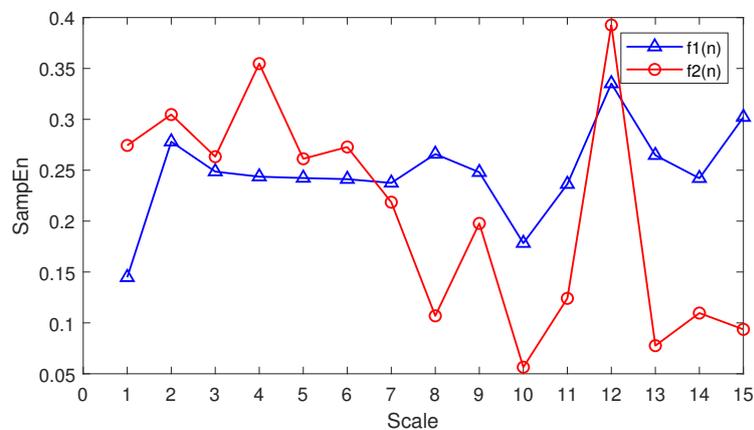


Figure 9. The MSE result for $f_1(n)$ and $f_2(n)$ at a scale of 1 ~ 15.

$f_1(n)$ and $f_2(n)$ can be distinguished by MSE easily since the two signals' MSE have a great difference when the scale is greater than eight. Therefore, MSE performs well when distinguishing signals with different low-frequency components.

3.4. Properties for Hierarchical Entropy

According to the basic theory of hierarchical entropy, it takes into account higher frequency components of the signal when calculating, while sample entropy and multiscale sample entropy do not. Consequently, hierarchical entropy performs better when measuring the complexity of those signals whose information is stored in both lower and higher frequency components. In order to illustrate this characteristic, a set of synthetic signals are given as follows:

$$f_3(n) = \begin{cases} \sin(2\pi * 5n) : 1 \leq n \leq (2^{13} - 2^{10}), \\ \sin(2\pi * 60n) : (2^{13} - 2^{10}) + 1 \leq n \leq 2^{13}. \end{cases} \quad (13)$$

$$f_4(n) = \begin{cases} \sin(2\pi * 5n) : 1 \leq n \leq (2^{13} - 2^{10}), \\ \sin(2\pi * 50n) : (2^{13} - 2^{10}) + 1 \leq n \leq 2^{13}. \end{cases} \quad (14)$$

$f_3(n)$ and $f_4(n)$ are signals that contain both higher and lower frequency components. Part of the waveform of $f_3(n)$ and $f_4(n)$ is shown in Figure 10.

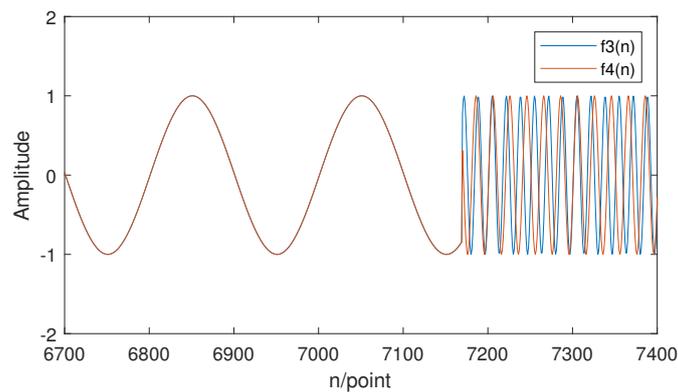


Figure 10. The waveform of $f_3(n)$ and $f_4(n)$.

It is obvious that the information stored in the lower frequency components is the same, while the information stored in the higher frequency components is different. Based on the theory of sample entropy and multiscale sample entropy, only the lower frequency part is considered, which will lead to lower accuracy in distinguishing different signals while using SE or MSE. However, HE still measures the complexity of $f_3(n)$ and $f_4(n)$ very well since it considers the information stored in the higher frequency component. The HE results of two signals is displayed in Figure 11. The numerical result of SE, MSE, and HE is also shown in Table 2.

Table 2. Different entropy’s results of $f_3(n)$ and $f_4(n)$.

	SE	MSE(2)	MSE(4)	HE(5,9)	HE(5,13)
$f_3(n)$	1.1447	0.2769	0.2419	0.2320	0.1533
$f_4(n)$	1.1442	0.2862	0.2460	0.3102	0.2645
Absolute Difference	0.0005	0.0093	0.0041	0.0782	0.1112

Before the interpretation of the results, first, some abbreviations are explained. $MSE(i)$ stands for the multiscale entropy of signals at scale i , and $HE(n, e)$ stands for the e^{th} component of the hierarchical entropy at scale n . These abbreviation are also used in the rest of this paper. According to the results displayed in Figure 11 and Table 2. The histogram at a scale of one is the sample entropy of the signal, $HE(i, 0)$ is equivalent to $MSE(2^{i-1})$. Based on this equivalence relationship between MSE and HE, the HE results of $f_3(n)$ and $f_4(n)$ illustrated in Figure 11 also include part of the results of MSE. From Figure 11c, the HE results of the low frequency components of the two signals are not much different, but in some of the high-frequency components, the two signals can be successfully distinguished. That is to say, MSE cannot distinguish between signals that differ only in high-frequency components. Hence, HE has a better performance than SE or MSE in distinguish different frequency signals, especially when the information of the signal is mainly stored in higher frequency components.

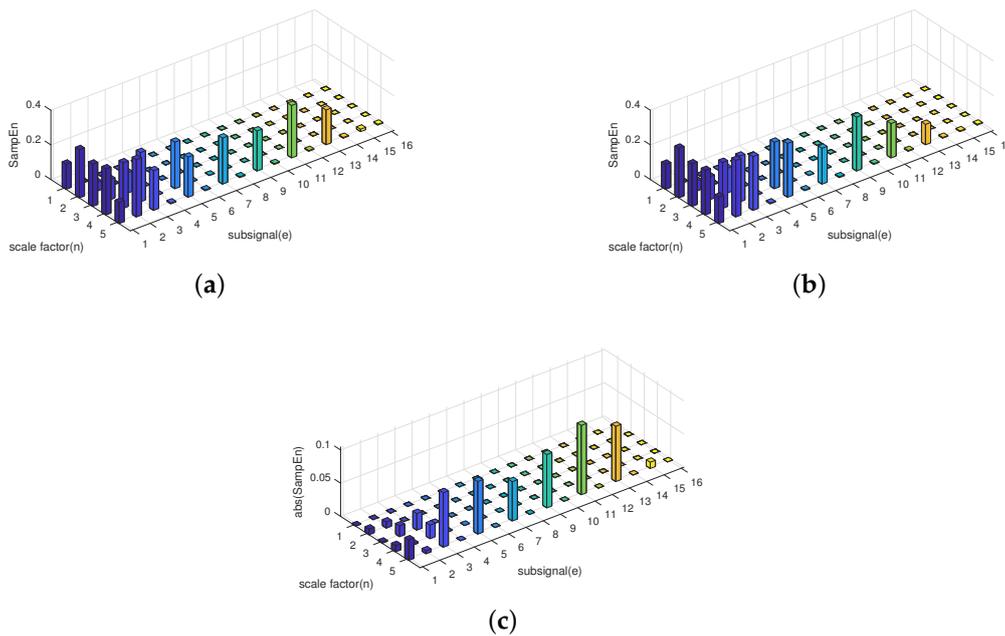


Figure 11. Hierarchical entropy results of $f_3(n)$ and $f_4(n)$. (a) HE results for $f_3(n)$. (b) HE results for $f_4(n)$. (c) HE's absolute difference.

4. Feature Extraction of Ship-Radiated Noise Based on Hierarchical Entropy

4.1. Feature Extraction Method Based on HE

The main steps of the feature extraction method based on HE are shown in Figure 12.

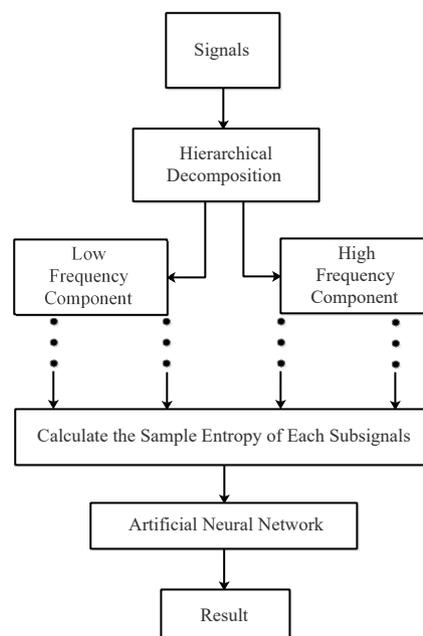


Figure 12. The flowchart of HE the feature extraction method.

- Step 1: Five types of ship-radiated noise are given in this paper; choose the appropriate hierarchical decomposition order to guarantee that the length of sub-signal is longer than 512.
- Step 2: By doing the hierarchical decomposition n times, 2^n sub-signals can be obtained, representing the lower and higher frequency components of the original signal, respectively.

- Step 3: Calculate the sample entropy for each sub-signal. Get the HE result.
- Step 4: Flatten the HE matrix into a vector. Pass the vector through an artificial neural network.
- Step 5: Get the classification results.

4.2. Feature Extraction of Ship-Radiated Noise Based on HE

In this section, five types of ship-radiated noise were employed for the feature extraction (the ship-radiated noise of Ships D and E can be obtained from <https://www.nps.gov/glba/learn/nature/soundclips.htm>). The sampling frequency of Ships A, B, and C was 52.7 kHz. As for Ships D and E, the sampling frequency was 44.1 kHz. Ship A was a cruise ship. The vessel was less than 50 m away from the hydrophone. The hydrophone depth was 4.8 m. Ship B was an ocean liner. The vessel was less than 50 m away from the hydrophone. The hydrophone depth was 5.8 m. Ship C was a motorboat. The distance between the vessel and the hydrophone changed from 50 m–100 m during the recording of the data approximately.

The hydrophone depth was 5.8 m. Further information for Ships A, B, and C can be found in [30]. Ships D and E were downloaded from a public website [31]. We chose a part of each signal and divided them into 100 segments separately. The length of each segment was 8192 sample points, namely 0.18 s of real-world data for Ships D and E and 0.15 s of real-world data for Ships A, B, and C. We can obtain 100 results for each type of ship-radiated noise by calculating the HE and MSE for every segment. The number of hierarchical decompositions was set as five. The waveform of five types of ship-radiated noise is demonstrated in Figure 13. Figure 14 gives the power spectrum density analysis results of the five types of signals.

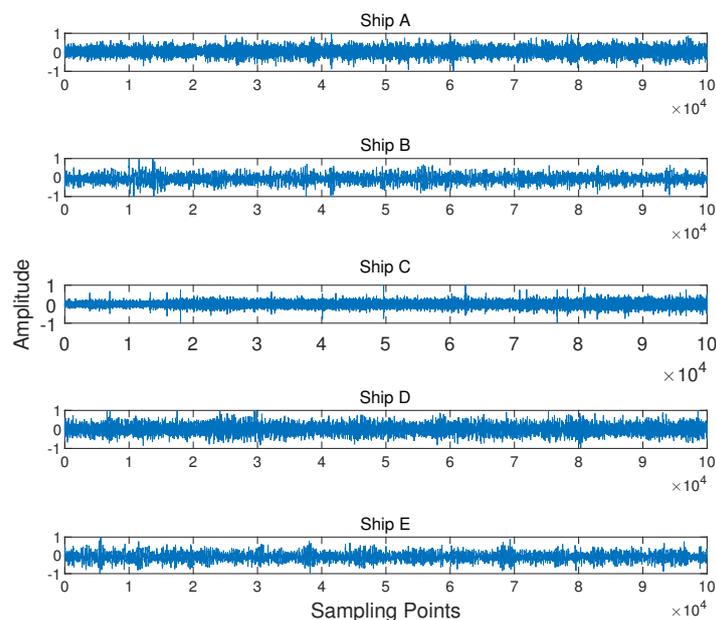


Figure 13. The waveform of the five types of ship-radiated noise.

Much useful information can be obtained from the power spectrum density analysis results of the five types of ship-radiated noise in Figure 14. The narrow-band spectral lines existing in Figure 14b,c make it easy to distinguish Ship B and Ship C. As for the rest of the types of ship, which are Ships A, D, and E in Figure 14a,d,e, few spectral lines can be found for us to distinguish different types of ship. Especially for Ships D and E, the fact that there was no evident distinction existing in their broadband spectral envelopes made it difficult for us to distinguish these two types of ships accurately. Therefore, classifying these five different ships using the spectrum as a feature is difficult.

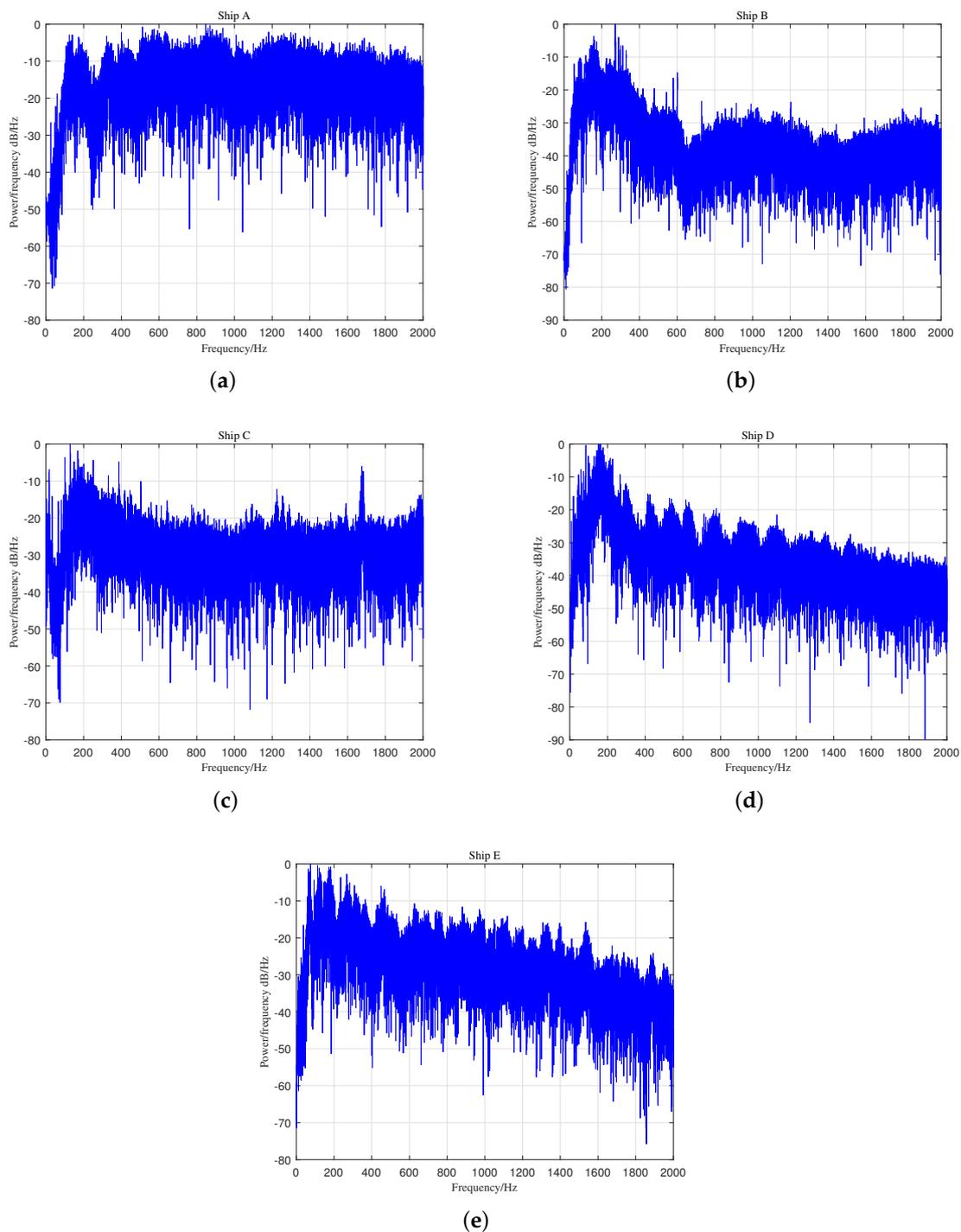


Figure 14. The power spectrum density analysis results of the five types of ship-radiated noise. (a) Ship A. (b) Ship B. (c) Ship C. (d) Ship D. (e) Ship E.

The HE results of the five types of ship-radiated noise are illustrated in Figure 15. In order to compare the performance when HE and MSE both calculate the same data length for their sub-signals, Figure 16 shows the MSE result of the five types of ships from a scale of 1–16. Guarantee that when calculating the HE at a scale five, the length of the sub-signal was 512 points, the same as MSE at a scale of 16. Since it is difficult to see the differences between the five types of ship-radiated noise through Figure 15, part of the HE results are also shown numerically through Table 3. HE(n) represents the HE result at scale n .

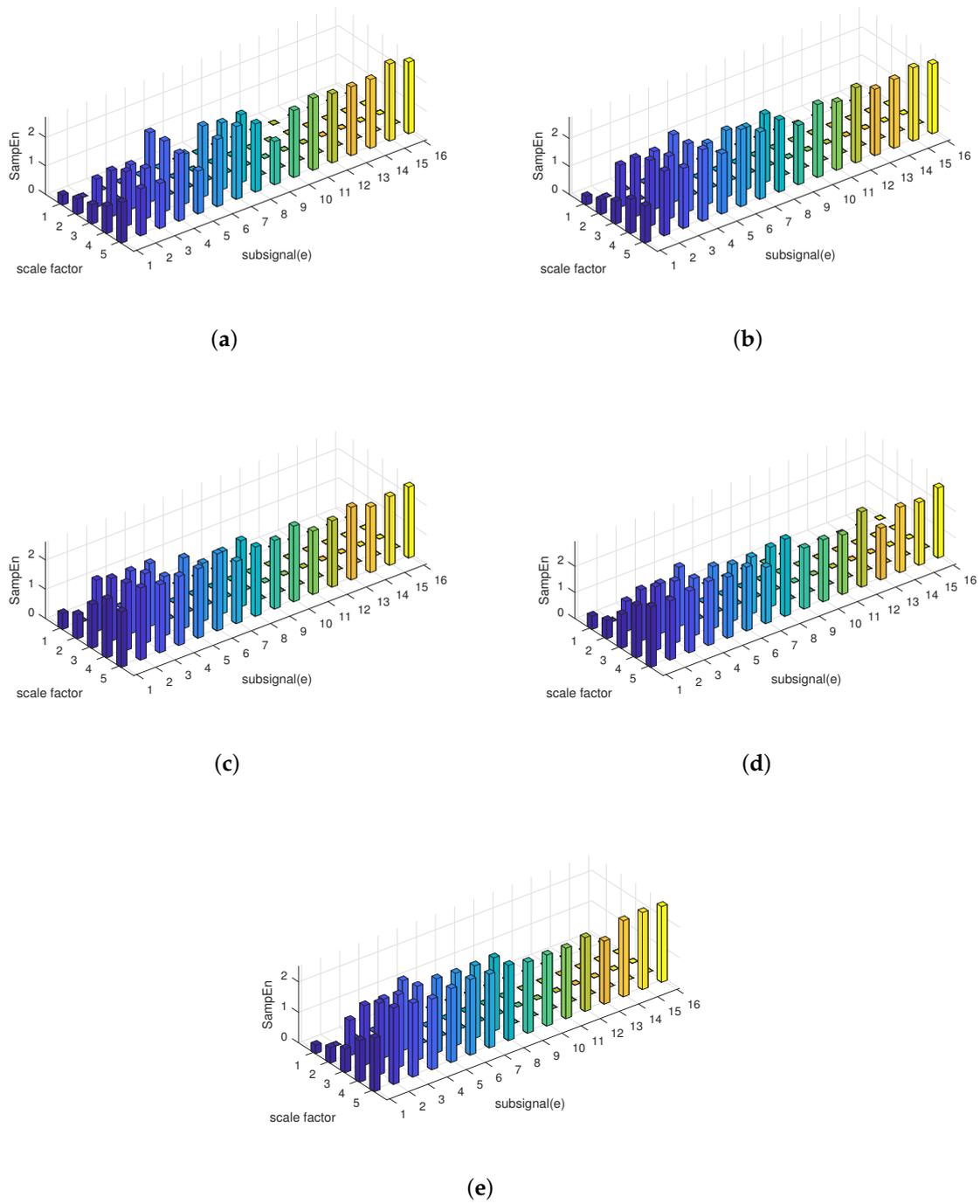


Figure 15. The HE results for the five types of ship-radiated noise. (a) Ship A. (b) Ship B. (c) Ship C. (d) Ship D. (e) Ship E.

According to the MSE result demonstrated in Figure 16, we can see that SE can only distinguish Ship C from other types of ship. Throughout the MSE result from a scale of 1–16, the entropy differences between Ships A and D and Ships B and E remained small.

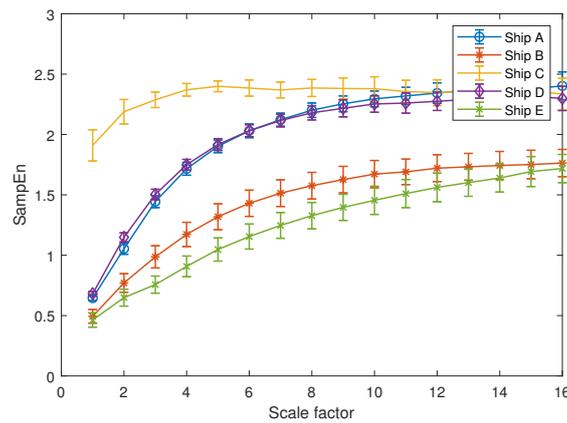


Figure 16. MSE results of the five types of ship-radiated noise.

Table 3. Part of HE results for five types of ship-radiated noise.

Ship Type	SE	MSE(2)	MSE(4)	MSE(8)	HE(3,3)	HE(4,7)	HE(5,3)	HE(5,13)
Ship A	0.64	1.04	1.72	2.21	2.25	2.19	2.08	2.17
Ship B	0.41	0.83	1.21	1.55	2.41	2.49	2.35	2.45
Ship C	1.92	2.13	2.23	2.37	2.36	2.45	2.41	2.51
Ship D	0.66	1.07	1.65	2.10	2.39	2.36	2.29	2.38
Ship E	0.42	0.75	1.06	1.53	2.37	2.47	2.72	2.61

To evaluate the performance of the above-mentioned feature extraction methods quantitatively, the results of two methods were separately classified and identified by a probabilistic neural network. Since the MSE’s results for the five types of ships were vectors of length 16, we fed the probabilistic neural network with these vectors to get the classification results. As for HE, we flattened the HE’s results from matrices into vectors of a length of 31, then fed the PNN with these vectors to get the classification results. The classification results are demonstrated in Tables 4–6. The training set for each type of ship was 70, and the test set was 30.

Table 4. Probabilistic neural network classification results of SE.

Type	Recognized as					Sensitivity	Specificity
	A	B	C	D	E		
A	28	0	0	2	0	93.3%	90%
B	3	27	0	0	0	90%	75%
C	0	0	30	0	0	100%	100%
D	9	0	0	21	0	70%	96.7%
E	0	30	0	0	0	0%	100%
Accuracy					70.7%		

Table 5. Probabilistic neural network classification results of MSE(1)~(16).

Type	Recognized as					Sensitivity	Specificity
	A	B	C	D	E		
A	21	0	0	9	0	70%	96.7%
B	0	25	0	0	5	83.3%	95.8%
C	0	0	30	0	0	100%	100%
D	4	0	0	26	0	86.7%	92.5%
E	0	5	0	0	25	83.3%	95.8%
Accuracy					84.7%		

Table 6. Probabilistic neural network classification results of HE(1)~(5).

Type	Recognized as					Sensitivity	Specificity
	A	B	C	D	E		
A	25	0	0	5	0	83.3%	99.2%
B	0	27	0	0	3	90%	100%
C	0	0	30	0	0	100%	100%
D	1	0	0	29	0	96.7%	95.8%
E	0	0	0	0	30	100%	97.5%
Accuracy						94%	

Before assessing the performance of the PNN, the definitions of “sensitivity” and “specificity” are given as follows:

$$\begin{cases} \text{Sensitivity} = \frac{TP}{TP+FN}, \\ \text{Specificity} = \frac{TN}{TN+FP}, \end{cases} \quad (15)$$

where TP, TN, FP, and FN are the abbreviations for “true positive”, “true negative”, “false positive”, and “false negative”, respectively. It is important to note that “accuracy” calculates the overall classification accuracy of neural networks, which is also the average of “sensitivity”.

From Tables 4–6, it is obvious that HE was able to classify five types of ships very well. Even for those types of ships that SE and MSE could not classify, their sensitivities in HE’s result were very high. The accuracy of HE increased 9.3% compared with MSE and 23.3% compared with SE. In order to eliminate the impact of sampling frequency, we reduced the sampling frequency of Ships A, B, and C from 52.7 kHz to 44.1 kHz, calculated the HE results for five types of ships, and passed the results through PNN. The classification result is demonstrated in Table 7. Through the table, we can see that the classification accuracy was 96%, very close to the accuracy of not reducing the sampling frequency.

Table 7. Probabilistic neural network classification results of HE(1)~(5) after reducing the sampling frequency.

Type	Recognized as					Sensitivity	Specificity
	A	B	C	D	E		
A	27	0	0	3	0	90%	98.3%
B	0	29	0	0	1	96.7%	100%
C	0	0	30	0	0	100%	100%
D	2	0	0	28	0	93.3%	97.5%
E	0	0	0	0	30	100%	99.1%
Accuracy						96%	

Moreover, we mixed five types of ship-radiated noise with Gaussian white noise. The SNR was set to be 5 dB, and the classification results are illustrated in Tables 8 and 9.

Table 8. (Noise) Probabilistic neural network classification results of HE(1)~(5).

Type	Recognized as					Sensitivity	Specificity
	A	B	C	D	E		
A	26	0	1	3	0	86.7%	97.5%
B	0	24	0	5	1	80%	97.5%
C	0	0	30	0	0	100%	99.2%
D	3	2	0	25	0	83.3%	92.5%
E	0	30	0	0	0	93.3%	99.2%
Accuracy						88.7%	

Table 9. (Noise) Probabilistic neural network classification results of MSE(1)~(16).

type	Recognized as					Sensitivity	Specificity
	A	B	C	D	E		
A	15	1	3	11	0	50%	90%
B	2	23	0	0	5	76.7%	93.3%
C	2	0	28	0	0	93.3%	95.8%
D	8	1	2	17	2	56.7%	89.2%
E	0	6	0	2	22	73.3%	94.2%
Accuracy						70%	

According to the results shown in Tables 8 and 9, as the noise mixed into the ship-radiated noise, both HE and MSE were affected. However, even though the accuracy of both methods decreased, HE's accuracy remained higher compared with MSE. The accuracy of HE decreased by 5.3% with added noise, while the accuracy of MSE decreased by 14.7% under the same conditions. Furthermore, even when the ship-radiated noise was mixed with noise, HE could still distinguish Ship C very well.

5. Conclusions

A new method was proposed for feature extraction of ship-radiated noise based on hierarchical entropy in this paper. The simulation analysis indicated that HE had better performance compared with MSE when the differences between signals were mainly focused on their high-frequency components. Applying two feature extraction methods to ship-radiated noise could help distinguish some signals that were not very different in the frequency domain. Moreover, in order to compare the performance of HE and MSE, we passed the extracted features through a neural network, and the classification results showed that the classification accuracy of HE was higher than MSE. In summary, since HE considered more information, as a new feature extraction method in the field of underwater acoustic signal processing, HE can better distinguish different signals in most circumstances than traditional entropy-based methods such as MSE.

Author Contributions: W.L., X.S., and Y.L. conceived and designed the research, W.L. analyzed the data and wrote the manuscript. X.S. and Y.L. revised the manuscript. All authors have read and approved the final manuscript.

Funding: This work was supported in part by Grant No. 11874302, No. 11574250, and No. 61571365 from NSFC. The research was also funded by the National Key R&D Program of China (2016YFC1400200).

Acknowledgments: This work was partially supported by the National Natural Science Foundation of China (No. 61571365, No. 11874302, No. 11574250) and the National Key R&D Program of China (2016YFC1400200).

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

Abbreviations

The following abbreviations are used in this manuscript:

AE	Approximate entropy
SE	Sample entropy
MSE	Multiscale sample entropy
HE	Hierarchical entropy
VMD	Variational mode decomposition
MPE	multiscale permutation entropy
EEMD	Ensemble empirical mode decomposition
SD	Standard deviation
SNR	Signal-to-noise ratio

References

1. Urick, R.J. *Principles of Underwater Sound for Engineers*; Tata McGraw-Hill Education: New York, NY, USA, 1967.
2. Rajagopal, R.; Sankaranarayanan, B.; Rao, P.R. Target classification in a passive sonar-an expert system approach. In Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Albuquerque, NM, USA, 3–6 April 1990; pp. 2911–2914.
3. Margarit, G.; Tabasco, A. Ship classification in single-pol SAR images based on fuzzy logic. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3129–3138. [[CrossRef](#)]
4. Li, Y.; Wang, L.; Li, X.; Yang, X. A novel linear spectrum frequency feature extraction technique for warship radio noise based on complete ensemble empirical mode decomposition with adaptive noise, duffing chaotic oscillator, and weighted-permutation entropy. *Entropy* **2019**, *21*, 507. [[CrossRef](#)]
5. Deaett, M. Signature modeling for acoustic trainer synthesis. *IEEE J. Ocean. Eng.* **1987**, *12*, 143–147. [[CrossRef](#)]
6. Bao, F.; Li, C.; Wang, X.; Wang, Q.; Du, S. Ship classification using nonlinear features of radiated sound: An approach based on empirical mode decomposition. *J. Acoust. Soc. Am.* **2010**, *128*, 206–214. [[CrossRef](#)] [[PubMed](#)]
7. Hornero, R.; Álvarez, D.; Abásolo, D.; del Campo, F.; Zamarron, C. Utility of approximate entropy from overnight pulse oximetry data in the diagnosis of the obstructive sleep apnea syndrome. *IEEE Trans. Biomed. Eng.* **2007**, *54*, 107–113. [[CrossRef](#)] [[PubMed](#)]
8. Howedi, A.; Lotfi, A.; Pourabdollah, A. Exploring Entropy Measurements to Identify Multi-Occupancy in Activities of Daily Living. *Entropy* **2019**, *21*, 416. [[CrossRef](#)]
9. Lake, D.E.; Richman, J.S.; Griffin, M.P.; Moorman, J.R. Sample entropy analysis of neonatal heart rate variability. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **2002**, *283*, R789–R797. [[CrossRef](#)]
10. Nicolaou, N.; Georgiou, J. The use of permutation entropy to characterize sleep electroencephalograms. *Clin. EEG Neurosci.* **2011**, *42*, 24–28. [[CrossRef](#)]
11. Li, Y.; Chen, X.; Yu, J.; Yang, X.; Yang, H. The Data-Driven Optimization Method and Its Application in Feature Extraction of Ship-Radiated Noise with Sample Entropy. *Energies* **2019**, *12*, 359. [[CrossRef](#)]
12. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [[CrossRef](#)]
13. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.-Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [[CrossRef](#)] [[PubMed](#)]
14. Pan, Y.H.; Wang, Y.H.; Liang, S.F.; Lee, K.T. Fast computation of sample entropy and approximate entropy in biomedicine. *Comput. Methods Programs Biomed.* **2011**, *104*, 382–396. [[CrossRef](#)] [[PubMed](#)]
15. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* **2002**, *89*, 068102. [[CrossRef](#)]
16. Costa, M.; Peng, C.K.; Goldberger, A.L.; Hausdorff, J.M. Multiscale entropy analysis of human gait dynamics. *Phys. A: Stat. Mech. Its Appl.* **2003**, *330*, 53–60. [[CrossRef](#)]
17. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of biological signals. *Phys. Rev. E* **2005**, *71*, 021906. [[CrossRef](#)]
18. Jiang, Y.; Peng, C.K.; Xu, Y. Hierarchical entropy analysis for biological signals. *J. Comput. Appl. Math.* **2011**, *236*, 728–742. [[CrossRef](#)]
19. Romero, R.; Monticelli, A. A hierarchical decomposition approach for transmission network expansion planning. *IEEE Trans. Power Syst.* **1994**, *9*, 373–380. [[CrossRef](#)]
20. Zhu, K.; Song, X.; Xue, D. A roller bearing fault diagnosis method based on hierarchical entropy and support vector machine with particle swarm optimization algorithm. *Measurement* **2014**, *47*, 669–675. [[CrossRef](#)]
21. Li, Y.; Li, Y.; Chen, X.; Yu, J. A novel feature extraction method for ship-radiated noise based on variational mode decomposition and multi-scale permutation entropy. *Entropy* **2017**, *19*, 342.
22. Yang, H.; Zhao, K.; Li, G. A New Ship-Radiated Noise Feature Extraction Technique Based on Variational Mode Decomposition and Fluctuation-Based Dispersion Entropy. *Entropy* **2019**, *21*, 235. [[CrossRef](#)]
23. Chen, Z.; Li, Y.; Liang, H.; Yu, J. Improved Permutation Entropy for Measuring Complexity of Time Series under Noisy Condition. *Complexity* **2019**, 2019. [[CrossRef](#)]

24. Siddagangaiah, S.; Li, Y.; Guo, X.; Chen, X.; Zhang, Q.; Yang, K.; Yang, Y. A complexity-based approach for the detection of weak signals in ocean ambient noise. *Entropy* **2016**, *18*, 101. [[CrossRef](#)]
25. Cato, D.H. A perspective on 30 years of progress in ambient noise: Source mechanisms and the characteristics of the sound field. *AIP Conf. Proc.* **2012**, *1495*, 242–260.
26. Chen, X.; Yin, C.; He, W. Feature extraction of gearbox vibration signals based on EEMD and sample entropy. In Proceedings of the 2013 10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Shenyang, China, 23–25 July 2013; pp. 811–815.
27. Yamamoto, Y.; Hughson, R.L. Coarse-graining spectral analysis: New method for studying heart rate variability. *J. Appl. Physiol.* **1991**, *71*, 1143–1150. [[CrossRef](#)] [[PubMed](#)]
28. Li, Y.; Li, Y.; Chen, Z.; Chen, X. Feature extraction of ship-radiated noise based on permutation entropy of the intrinsic mode function with the highest energy. *Entropy* **2016**, *18*, 393. [[CrossRef](#)]
29. Chanwimalueang, T.; Mandic, D. Cosine similarity entropy: Self-correlation-based complexity analysis of dynamical systems. *Entropy* **2017**, *19*, 652. [[CrossRef](#)]
30. Santos-Domínguez, D.; Torres-Guijarro, S.; Cardenal-López, A.; Pena-Gimenez, A. ShipsEar: An underwater vessel noise database. *Appl. Acoust.* **2016**, *113*, 64–69. [[CrossRef](#)]
31. Can, G.; Akbaş, C.E.; Cetin, A.E. Recognition of vessel acoustic signatures using non-linear teager energy based features. In Proceedings of the 2016 International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM), Reggio Calabria, Italy, 27–28 October 2016; pp. 1–5.



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).