Article

Signal Pattern Recognition Based on Fractal Features and Machine Learning

Chang-Ting Shi

College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China; shichangting@hrbeu.edu.cn

check for updates

Abstract: As a typical pattern recognition method, communication signal modulation involves many complicated factors. Fractal theory can be used for signal modulation feature extraction and recognition because of its good ability to express complex information. In this paper, we conduct a systematic research study by using the fractal dimension as the feature of modulation signals. Box fractal dimension, Katz fractal dimension, Higuchi fractal dimension, Petrosian fractal dimension, and Sevcik fractal dimension are extracted from eight different modulation signals for signal pattern recognition. Meanwhile, the anti-noise function, box-diagram, and running time are used to evaluate the noise robustness, separability, and computational complexity of five different fractal features. Finally, Bback-Propagation (BP) neural network, grey relation analysis, random forest, and K-nearest neighbor are proposed to classify the different modulation signals based on these fractal features. The confusion matrices and recognition results are provided in the experimental section. They indicate that random forest had a better recognition performance, which could reach 96% in 10 dB.

Keywords: pattern recognition; fractal dimension; feature evaluation; random forest classifier

1. Introduction

As a classic pattern recognition problem, the modulation recognition of communication signals has many important research prospects and application value. In the military field, it is a precondition to carry out communication reconnaissance and jamming. Once the signal modulation of the enemy's communication system is clarified, the enemy's signals can be demodulated and communication information can be obtained. In the civil filed, signal modulation recognition can be used to signal confirmation, interference identification, spectrum management, and spectrum monitoring. Therefore, a secure and reliable feature extraction method is needed to effectively recognize the different signal modulations [1] in a complex environment.

Modulation pattern recognition generally uses statistical pattern recognition methods and decision-making methods. At present, most of the applications are statistical pattern recognition methods, which can be divided into two parts: feature extraction and classifier design. Good features are a key factor to improve the classification accuracy. In recent decades, researchers have proposed many methods for extracting feature parameters such as instantaneous features [2], fractal features [3], signal constellation reconstruction [4], etc. As another key factor of signal pattern recognition, classifiers directly determine the classification results. Machine learning classifiers are widely used in signal classification, and include the decision tree classifier [5], K-nearest neighbor (KNN) Classifier [6], support vector machine (SVM) classifier [7], neural network classifier [8], etc. In addition, confusion matrix, recognition rate, and running time are widely used as evaluation methods. Reference [9] proposed a novel method called the A-test for evaluating the structural risk of a classifier both quantitatively and graphically. The structural risk of a classification method was defined as the

instability of the method with the new test data. In the A-test method, the classification error percentage of the different values of z is calculated by using the balanced Z-fold verification method.

In this paper, we focus mostly on fractal feature extraction. Different methods are proposed for estimating the fractal dimension. However, considering that fractal features are limited to a few fractal types (e.g., box dimension), the application of fractal methods for the purpose of extracting discriminative features has not been deeply investigated. The main goal of the present paper is thus to help fill this research gap. Therefore, many kinds of fractal dimensions (i.e., box fractal dimension, Katz fractal dimension, Higuchi fractal dimension, Petrosian fractal dimension, and Sevcik fractal dimension) are extracted in order to provide a comprehensive description.

In the classifier section, Back-Propagation (BP) neural network, grey relation analysis, random forest, and K-nearest neighbor are used to classify the modulated signals. BP neural networks have powerful pattern recognition capabilities, providing better robustness and potential fault tolerance. Therefore, it is easy to obtain a higher recognition rate. A kind of elastic back propagation neural network is used as classifier in Reference [8]. In Reference [10], an improved bee colony algorithm is applied to digital modulation recognition by optimizing a BP neural network. Grey correlation can be used to analyze and describe the attributes and characteristics of the model, which is suitable for mathematical statistical analysis. In Reference [11], the basic grey relation model is proposed. Random forest has high prediction accuracy, and good tolerance to outliers and noise. It is also not prone to overfitting. In References [12,13], random forest is used for dimensionality reduction in low signal-to-noise ratio (SNR). In Reference [14], the information entropy is used as the modulation feature, and random forest is used as the classifier. A high recognition rate is obtained in low SNR. The K-nearest neighbor algorithm has the advantages of simple calculation and high recognition efficiency [6]. A KNN–SVM joint classifier is proposed in Reference [7], which can reduce the sensitivity to the choice of kernel function parameters and ease the difficulty of parameter selection.

Figure 1 shows the process of signal modulation recognition in this paper. Firstly, five fractal dimensions are extracted for the eight different kinds of digital modulated signals. Then, the noise robustness, data distribution, and computational complexity are evaluated. After that, four different kinds of classifiers are introduced. Finally, the classification performance is given through the confusion matrix and recognition rate.
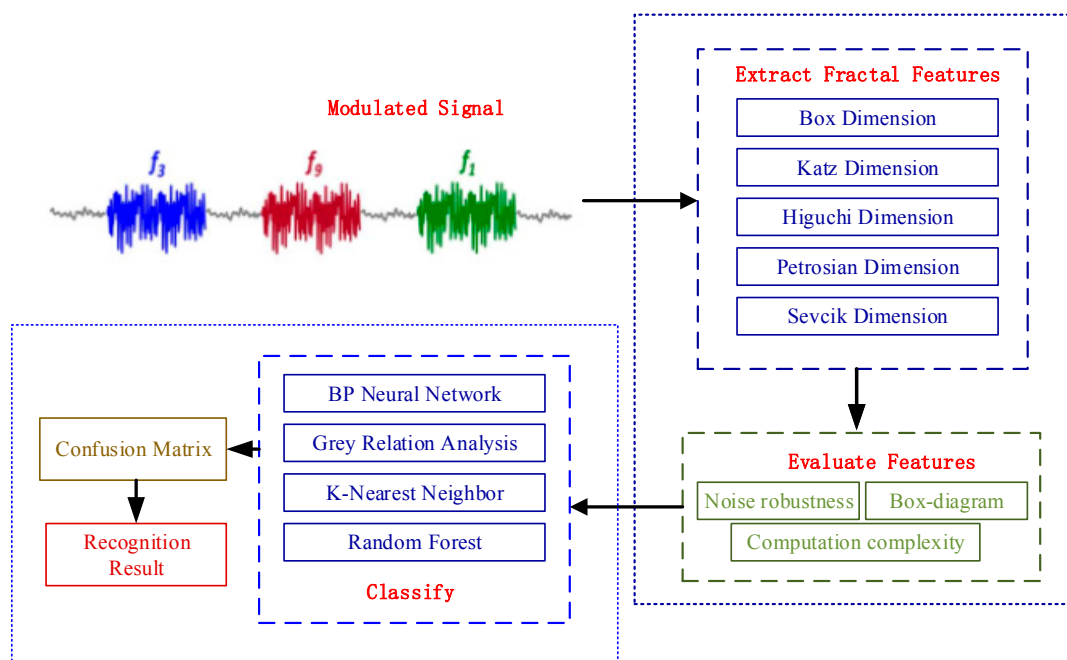

Figure 1. Signal pattern recognition process in this paper.

2. Related Work

In this section, we provide a description of some works related to fractal feature extraction. Fractal-based pattern recognition has been paid increasing attention, because it can effectively deal with feature information in a complex electromagnetic environment. One of the first studies of fractal theory [15,16] was conducted by B.B. Mandelbrot in 1975, as a branch of modern mathematical theory. The fractal dimension is the main parameter in fractal theory. It quantitatively describes the complexity of the fractal set. As a kind of time series, the communication modulation signal can be effectively described by the fractal dimension. There are many definitions of and measurement methods for fractal dimension [17–20].

The common one-dimensional fractal dimensions are box dimension, Hausdorff–Besicovitch dimension, Sevcik fractal dimension, Katz dimension, Higuchi dimension, and Petrosian dimension. Among them, the box dimension is easy to calculate, and a box with different side lengths is used to describe the change of the signal waveform. Smaller side lengths of the box lead to a longer calculation time, but the recognition rate of the signal will increase. In References [21,22], the box dimension was used to characterize the signal characteristics. The Hausdorff–Besicovitch dimension [23,24] was the basic fractal dimension. Hausdorff–Besicovitch and box dimensions were considered to minimize the antenna size, holding a high radiation efficiency in Reference [25]. This confirmed the importance of Hausdorff–Besicovitch Dimension, but also indicated that it has a high computational complexity, making it difficult to realize in practice. The Sevcik fractal dimension was proposed to calculate the fractal dimension of waveforms in Reference [26]. This method could be used to quickly estimate the Hausdorff–Besicovitch dimension of waveforms and measure the complexity and randomness of waveforms. Petrosian proposed a fast method to estimate the fractal dimension of a finite sequence [27], which converts the data to binary sequence before estimating the fractal dimension from time series. Tricot compared the estimation accuracy delivered by Katz's and Higuchi's methods on four synthetic fractal waveforms. The results indicated that Katz's method invariably under-estimated the true dimension, but Higuchi's method was more accurate in estimating the fractal dimensions.

In addition, in Reference [28], the definition and mathematical analysis of the Weierstrass–Mandelbrot fractal function were described. The application of fractal functions in engineering was discussed in Reference [29], including fractal modeling and calculation. In Reference [30], the effect of multiple fractals on a digital communication signal was studied under different noise distributions. The simulation results showed that multiple fractals could effectively extract the feature of FSK (Frequency Shift Keying) signals under different noise distributions. Additionally, the use of multiple fractals was more effective in depicting subtle features [31–33]. However, the problem of large complexity in multifractal calculation could not be avoided [34,35].

Therefore, how to select the proper fractal dimension algorithm according to the signal complexity, and then accurately classify the different communication modulation signals are the key problems in this paper. Therefore, we conducted a systematic empirical study on how to choose fractal dimension features for classifying communication modulation signals.

3. Fractal Dimension

3.1. Fractal Box Dimension

Self-similar dimensions are difficult to apply to objects which are not strictly self-similar, and the box dimension can be used to overcome this problem. The idea is to apply continuous hypercube mesh coverage to the curve. The result is a value similar to the Hausdorff dimension, which is another standard method for calculating the fractal dimension. In the metric space (X, d), A is one of the M

nonempty emergency clusters of X. For a box with a side length of ε, the minimum number $N(A, \varepsilon)$ of boxes required to cover A is shown as follows:

$$N(A, \varepsilon) = \left\{ M : A \subset \sum_{i=1}^{M} N(x_i, \varepsilon) \right\}, \tag{1}$$

where $x_1, x_2, \cdots x_M$ are different points of X. When ε approaches 0, the box dimension is shown as follows:

$$D_b = \lim_{\varepsilon \to 0} \frac{InN(A, \varepsilon)}{In(1/\varepsilon)}. \tag{2}$$

As indicated above, the box dimension merely represents the geometric dimension of the signal, but does not reflect the density distribution in the planar space.

3.2. Katz Fractal Dimension

For a signal (x_i, y_i) of length N, the Katz fractal dimension is shown as follows:

$$D = \frac{\log(N)}{\log(N) + \log(\frac{d}{L})}, \tag{3}$$

where L is the length of the waveform, and d is the maximum distance between the initial point (x_1, y_1) to the other points. They can be defined as follows:

$$L = \sum_{i=0}^{N-2} \sqrt{(y_{i+1} - y_i)^2 + (x_{i+1} - x_i)^2}, \tag{4}$$

$$d = \max\left(\sqrt{(x_i - x_1)^2 - (y_i - y_1)^2} \right). \tag{5}$$

3.3. Higuchi Fractal Dimension

Like the box counting, Higuchi's method is iterative in nature. However, it is especially useful to handle waveforms as objects.

Supposing the time sequence is $x(1), x(2), \cdots, x(N)$, the reconstructed time sequence is shown as follows:

$$x_m^k = \left\{ x(m), x(m+k), x(m+2k), \cdots, x\left(m + \left[\frac{N-m}{k} \right] \cdot k \right) \right\}, \tag{6}$$

where k represents the interval between two adjacent time series, $m = 1, 2, \cdots, k$ represents the initial time value of the sequence, symbol $[x]$ represents the integer part of x. For every x_m^k, the average length is shown as follows:

$$L_m(k) = \frac{\sum_{i=1}^{[(N-m)/k]} |x(m+ik) - x(m+(i-1) \cdot k)| \cdot (N-1)}{[(N-m)/k] \cdot k}. \tag{7}$$

Therefore, the total average length of the discrete time signal sequence is shown as follows:

$$L(k) = \sum_{m=1}^{k} L_m(k). \tag{8}$$

By the logarithmic transformation, we can get the following result:

$$\ln(L(k)) \propto D \cdot \ln\left(\frac{1}{k}\right). \tag{9}$$

The slope D is the Higuchi fractal dimension.

3.4. Petrosian Fractal Dimension

Firstly, the time series is $x(1), x(2), \cdots, x(N)$. Suppose the waveform signal consists of a series of points $\{y_1, y_2, \cdots, y_N\}$, binarize the sequence first, and set the binary matrix z_i, then:

$$z_i = \begin{cases} 1, & x_i > mean(y) \\ -1, & x_i \leq mean(y) \end{cases}, i = 1, 2, \cdots, N. \tag{10}$$

Then, find the total number of adjacent symbol changes in the sequence:

$$N_\Delta = \sum_{i=1}^{N-2} \left| \frac{z_{i+1} - z_i}{2} \right|. \tag{11}$$

The Petrosian fractal dimension is shown as follows:

$$D = \frac{\log_{10}^N}{\log_{10}^N + \log_{10}\left(\frac{N}{N + 0.4N_\Delta}\right)}. \tag{12}$$

3.5. Sevcik Fractal Dimension

Similarly, supposing the signal consists of a series of points (x_i, y_i), the length of signal is N. Normalize the signal first:

$$x_i^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, y_i^* = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \tag{13}$$

$x_{\min}, y_{\min}$ are the minimum values among x_i, y_i. $x_{\max}, y_{\max}$ are the maximum values among x_i, y_i. Then, the Sevcik fractal dimension D can be calculated as follows:

$$D = 1 + \frac{\ln(L) + \ln(2)}{\ln[2 \times (N-1)]}, \tag{14}$$

where L is the length of waveform, shown as follows:

$$L = \sum_{i=0}^{N-2} \sqrt{\left(y_{i+1}^* - y_i^*\right)^2 + \left(x_{i+1}^* - x_i^*\right)^2}. \tag{15}$$

4. Classifier Algorithm

4.1. BP Neural Network

Neurons are the basic units of artificial neural networks. The neurons' structure is just like in human cells. To date, hundreds of artificial neural network models have been proposed. BP neural networks have been widely used in pattern recognition, classification research, data regression, and more. The neurons' weight updating mainly uses the error back propagation learning algorithm. The sample is placed at the input, and the results of the BP neural network are compared to the original correct output. The error is extended from back to front for updating neurons' weight. The learning process is shown as follows:

Algorithm: BP Neural Network Classifier

Input: training dataset, testing dataset
Output: classification result (label out)
1. Network initialization
2. Calculate the output of the network
3. Adjust the neurons' weight in the network
4. Achieve the minimum value of the objective function
5. Output the classification result

4.2. Grey Relational Analysis

The mathematical theory of grey relation analysis (GRA) originates from space theory, which is based on clusters. The calculation of grey relation degree reveals the relationship between two different discrete sequences. According to the relationship between the training samples, we can divide them into several groups, and then analyze the uncertain relationship between the reference sequences and the training samples. Finally, the category type with the greatest relationship to the reference sequences can be found.

The main steps of GRA include determining the comparison sequences, defining the reference sequences, and calculating the grey relationship degree. The calculation process is shown as follows:

Algorithm: Grey Relational Analysis

Input: training dataset, testing dataset
Output: classification result (label_out)
1. Determining the comparative sequence $x_k(k = 1, 2, \ldots, K)$
2. For $i = 1$ to K
      Calculate the correlation degree γ_{0i}
   End
3. $\gamma_{\max} = \max(\gamma_{0i})$
4. Label out = find $(\gamma_{0i} == \gamma_{\max})$
5. End

4.3. K-Nearest Neighbor

K-nearest neighbor is a target classification method based on the closest training instance in feature space. The function can only be locally approximated, and all calculations are deferred to the classification. The KNN algorithm is divided into two phases: the training phase and the classification phase. The training examples are vectors in the multidimensional feature space. Each vector has a class label. The training phase of the algorithm only stores the feature vector (reference vector library) and class label of the training samples. In the classification phase, K is a user-defined constant, and an unlabeled vector (a query or test point) is classified by assigning the label which is most frequent among the K training samples nearest to that query point. In other words, the KNN method compares the query point or an input feature vector with a library of reference vectors and the query point is identified to be of the class of library feature vector to which it has the closest distance. The calculation process is shown as follows:

Algorithm: KNN Classifier

Input: training dataset, testing dataset

Output: classification result (label out)

1. Calculate the weight of characteristic: w_{ij}

2. Calculate the vector space model of the training sample and the sample to be tested:

$d_i = (w_{i1}, w_{i2}, w_{i3}, \cdots, w_{in})$

3. The distance of samples: $Sim(A, B) = \dfrac{\sum\limits_{j=1}^{n} w_{ij} w_{kj}}{\sqrt{\sum\limits_{j=1}^{n} w_{ij}{}^2} \sqrt{\sum\limits_{j=1}^{n} w_{kj}{}^2}}$

4. For $i = 1$ to K

    Calculate the weight of characteristic item: $wei(x, C_t) = \sum sim(x, d_i) y(d_i, C_t)$

5. $wei_{\max} = \max(wei)$

6. Label out = find ($wei == wei_{\max}$)

7. End

4.4. Random Forest

The final discriminant result of random forest is the voting result of all decision trees. However, this voting method does not consider the difference between a strong classifier and a weak classifier. Once the number of decision trees with wrong results is greater than the number of decision trees with correct classification results, the recognition result will become worse. Therefore, we do not consider voting in the output, but a probability value is assigned to each category in the form of probability, which is the basic probability assignment in evidence theory. The recognition results of random forest subclassifiers placed in different positions are synthesized, and the category with the highest fusion probability is selected as the final category, which can greatly improve the recognition rate. The processing flow is shown in Figure 2.

The calculation process is as follows:

Algorithm: Random Forest Classifier

Input: training dataset, testing dataset

Output: classification result (label out)

1. Training J random forest classifiers

2. For $i = 1$ to J

        Calculate each probability of every classifier p_{ij}: $p_{ij} = \dfrac{N_{ij}}{ntree}$

    End

3. Use Dempster–Shaferevidence theory to calculate each result probability of every classifier Pi

4. $P_{\max} = \max(P_i)$

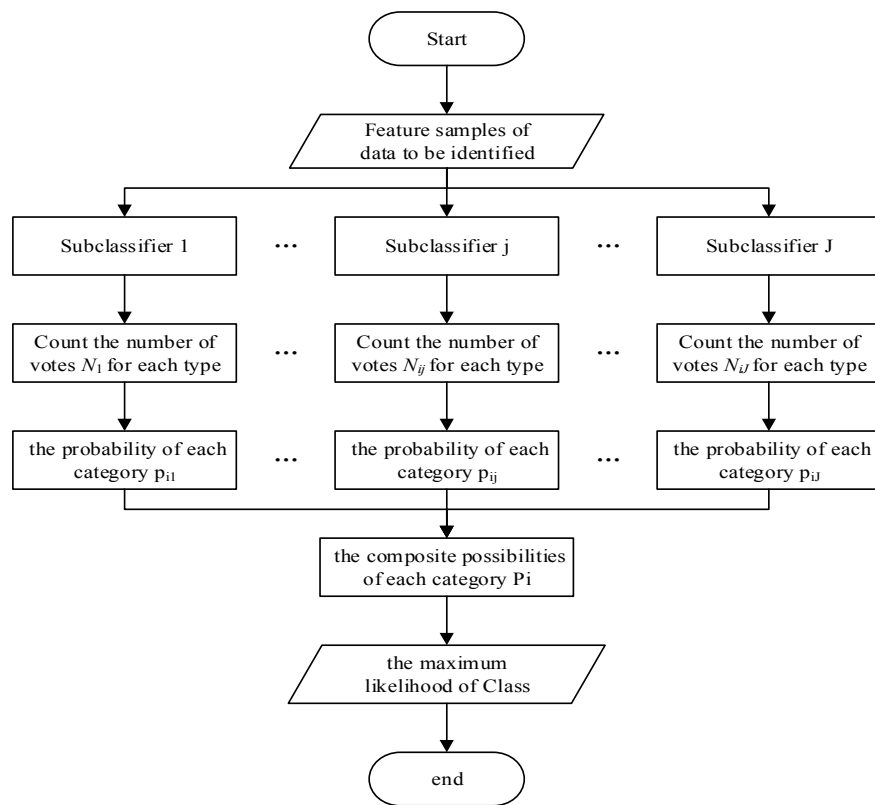5. Label out = find ($P_i == P_{\max}$)

6. End

Figure 2. Random forest structure.

5. Simulation Results

5.1. Simulation Analysis of Fractal Features

Multiple Amplitude Shift Keying (MASK, M = 2, 4), Multiple Frequency Shift Keying (MFSK, M = 2, 4, 8), Multiple Phase Shift Keying (BPSK), 16QAM (Quadrature Amplitude Modulation), and 32QAM were used to conduct the simulation. Signal correlation parameters were carrier frequency f_c =4 MHz $f_c = 4$MHz, sampling frequency $f_s = 4 \times f_c$, the initial frequency of MFSK, f_1 =4 MHz, and frequency deviation Δf = 1 MHz. Signal length N_s = 2048, digital signal symbol rate R_s =1000 Sps. Random codes forming a baseband signal were used. The SNR range was −5 to 10 dB. The simulation results of the calculated signals are shown in Figure 3.

Through Figure 3, we can directly see that even for eight digitally modulated signals, the difference in features between the signals were more obvious. Meanwhile, in the whole SNR range, the feature distribution was slightly fluctuating. That is, the fractal feature was not sensitive to the noise environment. Under each fractal feature, the data distribution interval was quite different, and the number of signal features of the Katz dimension jumped the most among the eight signal features. The box dimension distinguished the eight signals better, and the Katz dimension and Sevcik dimension distinguished 4ASK, 16QAM, and 32QAM better. The Petrosian dimension distinguished the 2FSK and 16QAM signals better. The Higuchi fractal dimension characteristic signal was the worst. The aliasing between signals was serious.

One of the main difficulties in automatic modulation pattern recognition is the problem of most recognition methods having poor recognition in a large dynamic SNR environment. In practical applications, the SNR changes rapidly during the working process and it is difficult to maintain stability. Therefore, high requirements are placed on the adaptability of large dynamic SNR (i.e., noise robustness). For the characteristics of the automatic recognition algorithm based on feature extraction and machine learning, the noise robustness of the recognizer is mainly reflected in the selection of

features. From Figure 3 we can see that for various modulated signals, each fractal feature under different SNR showed low noise sensitivity. When evaluating the noise robustness of each feature, an algorithm to analyze the overall anti-noise performance was used as follows.
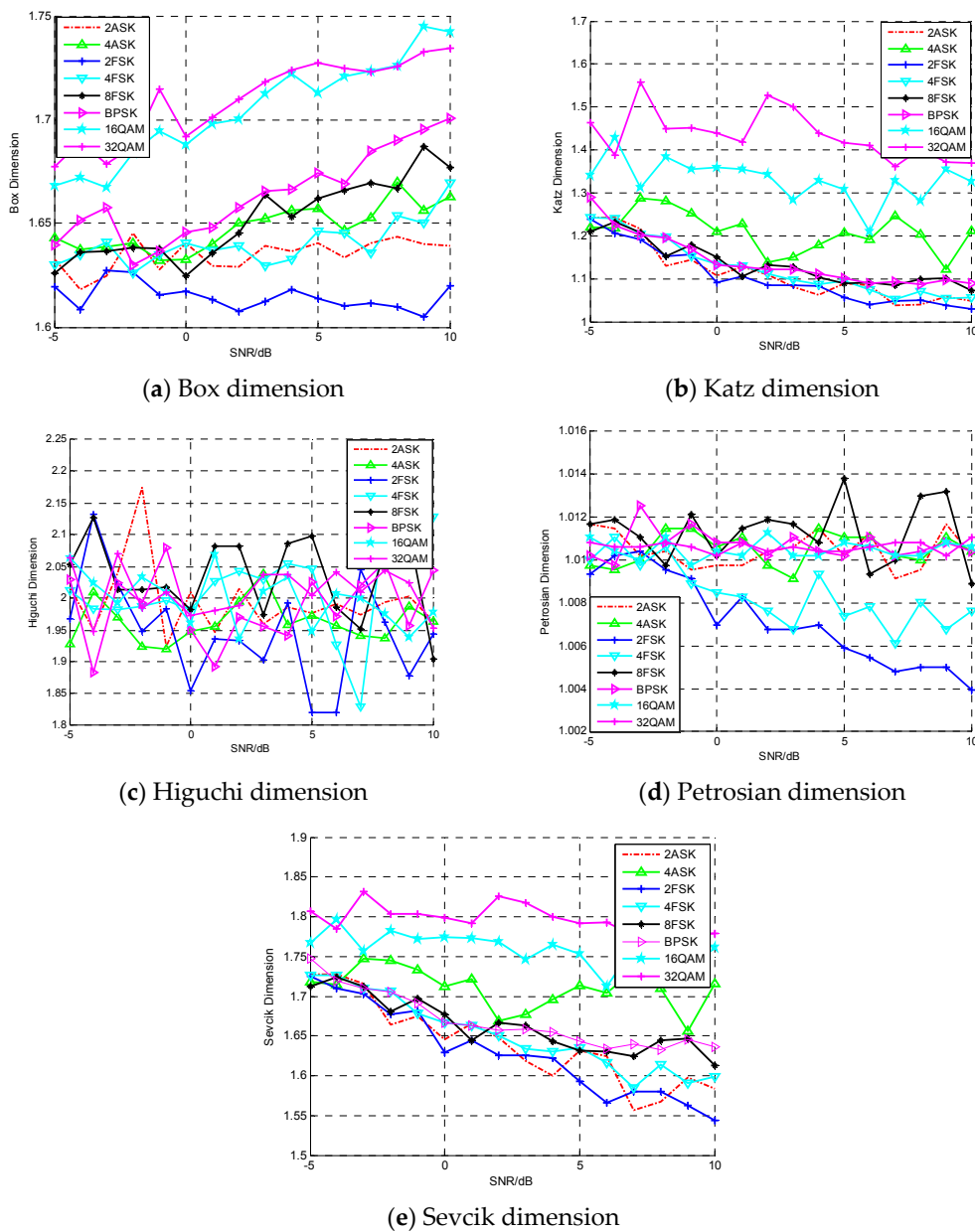

(**a**) Box dimension

(**b**) Katz dimension

(**c**) Higuchi dimension

(**d**) Petrosian dimension

(**e**) Sevcik dimension

Figure 3. Fractal dimension of different communication modulation signals.

According to Figure 3, the differences among the eight different signals were obvious. Meanwhile, in the whole SNR range, the feature distribution was slightly fluctuating, which means that the fractal feature was not sensitive to the noise environment. Comparing the five different fractal features, the data distribution was also quite different. The Katz dimension differed the most among the five different features. The box dimension could distinguish the eight signals effectively. Katz and Sevcik dimensions could better distinguish 4ASK, 16QAM, and 32QAM. Petrosian dimension could better distinguish the 2FSK and 16QAM signals. The Higuchi fractal dimension was the worst. The aliasing among signals was serious.

One of the main difficulties in automatic modulation recognition is the problem of poor recognition rate in a large dynamic SNR environment. In practical applications, the SNR changes rapidly with the environment and it is difficult to maintain stability. Therefore, high requirements are placed on the adaptability of large dynamic SNR. For automatic modulation recognition based on feature extraction and machine learning, the noise robustness of the classifier mainly depends on the selection of features. From Figure 3, we can see that the fractal feature under different SNRs showed low noise sensitivity. Therefore, an algorithm for analyzing the overall anti-noise performance is provided as follows.

Calculating a feature parameter C for the *i*th modulation method, the zero-center normalized statistical variance $Var(C_i)$ of the corresponding parameter values at different SNR is shown as follows:

$$Var(C_i) = \frac{1}{K}\sum_{k=1}^{K}[C_i(k) - E(C_i(k))]^2. \tag{16}$$

K is the number of samples for each feature. $C_i(k)$ is the zero-center normalized statistical value of parameter value C, which corresponds to the feature value $c_i(k)$ for the *i*th modulation under different SNRs. $C_i(k)$ can be calculated by the following equation:

$$C_i(k) = \frac{c_i(k)}{\frac{1}{K}\sum_{k=1}^{K}c_i(k)} - 1, k = 0, 1, 2, \ldots K \tag{17}$$

Then, the sum of the zero-center normalized statistical variance $Var(C_i)$ is calculated with n different modulation modes as $SVar(C)$:

$$SVar(C) = \sum_{i=1}^{n}Var(C_i) \tag{18}$$

Table 1 shows the anti-noise evaluation results of the five different fractal features. The $SVar(C)$ of the five different fractal features are all within $\times 10^{-3}$, which reflects the superior noise immunity of fractal features. Therefore, fractal features have an advantage in overcoming the problem of poor robustness in traditional automatic modulation recognition.

Table 1. Robustness evaluation (SVar, zero-center normalized statistical variance).

Type	Box	Katz	Higuchi	Petrosian	Sevcik
SVar/$\times 10^{-3}$	2.5	1.88	9	0.03	9.3

Furthermore, in order to show the data distribution of the five different features more clearly, the box diagram distribution of these features is given at SNR = 0 dB. A box diagram is a statistical method to reflect the discrete data distribution. The box diagram consists of five data points: minimum, lower quartile, median, upper quartile, and maximum. The lower quartile, median, and upper quartile make up a "box with compartments". An extension line is established between the quartiles and the maximum value. There is always a wide variety of "dirty data" in real data, which are also called "outliers". Outliers are picked up individually. Therefore, the outliers of a dataset can be recognized intuitively, and the data dispersion and bias can be judged. The box diagram distribution is shown in Figure 4.

According to Figure 4, the data distribution of the Sevcik dimension had the biggest difference, which means it had the best ability to distinguish the eight different kinds of communication modulation signals. The data distribution of the box dimension and Katz dimension were also ideal, where the intra-class distribution of the Katz dimension was the most concentrated and the aggregation degree was the best. The data ranges of Higuchi dimension and Petrosian dimension were small, and the data distribution between different signals was similar.
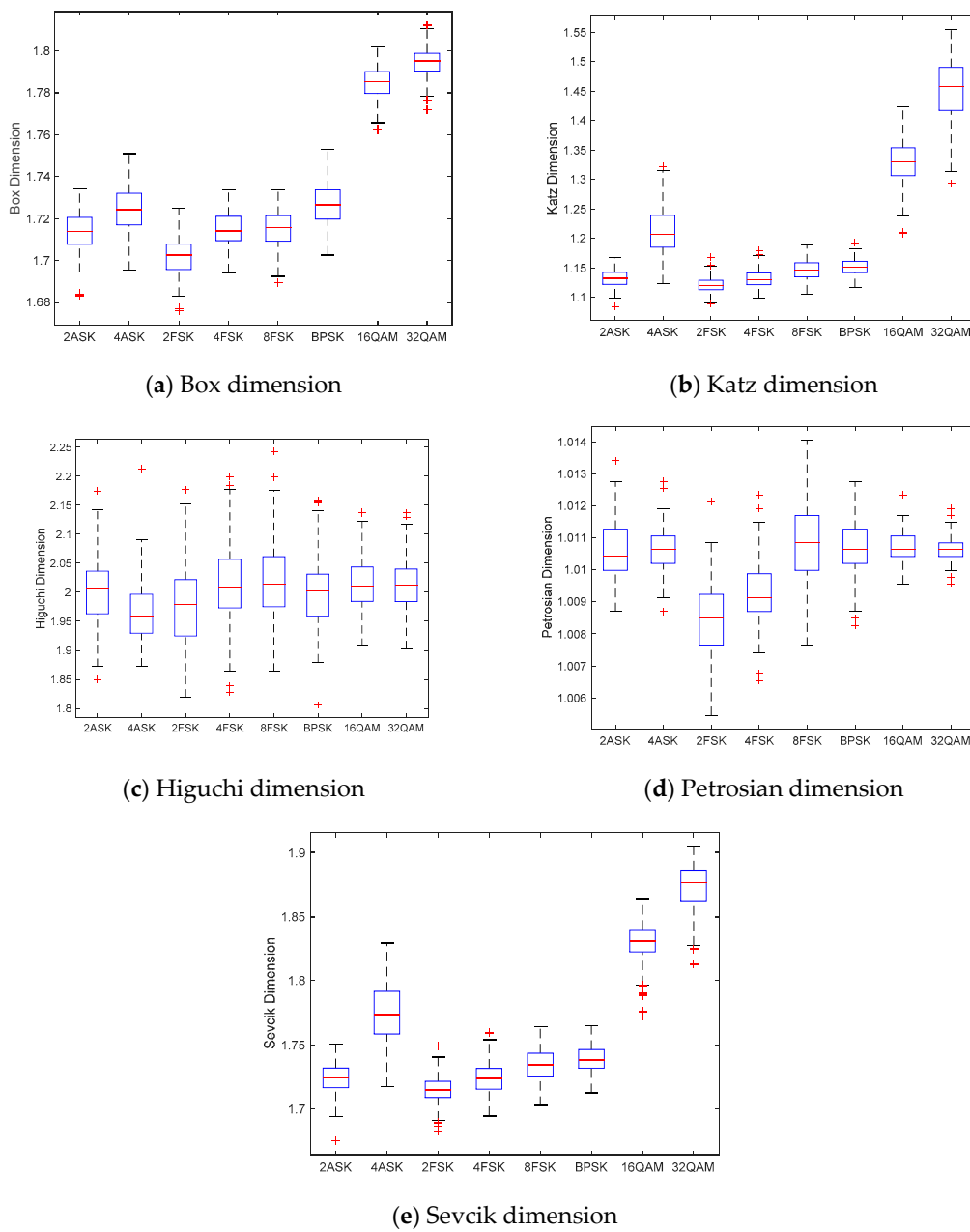
(**a**) Box dimension

(**b**) Katz dimension

(**c**) Higuchi dimension

(**d**) Petrosian dimension

(**e**) Sevcik dimension

Figure 4. Characteristic box diagram evaluation of five different fractal dimensions.

Computational complexity is also very important for evaluating features. Therefore, the running times of the five fractal features were calculated with a DELL Inspiron 14 laptop. The CPU was an Intel(R) Core(TM) i7-4510U, and the memory was 8 GB. The results are shown in Table 2.

Table 2. Characteristic complexity evaluation.

Type	Box	Katz	Higuchi	Petrosian	Sevcik
Runtime/s	1.92	1.66	82.199	2.83	1.26

The Table 2 indicates that the running time of the Sevcik dimension was the shortest, meaning that the Sevcik dimension had the lowest computational complexity. The running time of the Higuchi dimension was more than 30 times longer than the others. Therefore, the computational complexity of the Higuchi dimension was much higher than that of the others.

5.2. Classification Results

Four classifiers were used: BP neural network, grey relation analysis, random forest, and K-nearest neighbor. The training dataset contained 25,600 samples. From −5 dB to 10 dB, there were 200 samples per signal under each SNR. The testing dataset contained 25,600 samples. From −5 dB to 10 dB, there were 200 samples per signal under each SNR. The parameters of the BP neural network were as follows: the number of hidden layers was 120. The training time was 500. The training accuracy was 0.04, and the learning rate was 0.1. The number of trees grown of the random forest was 300.

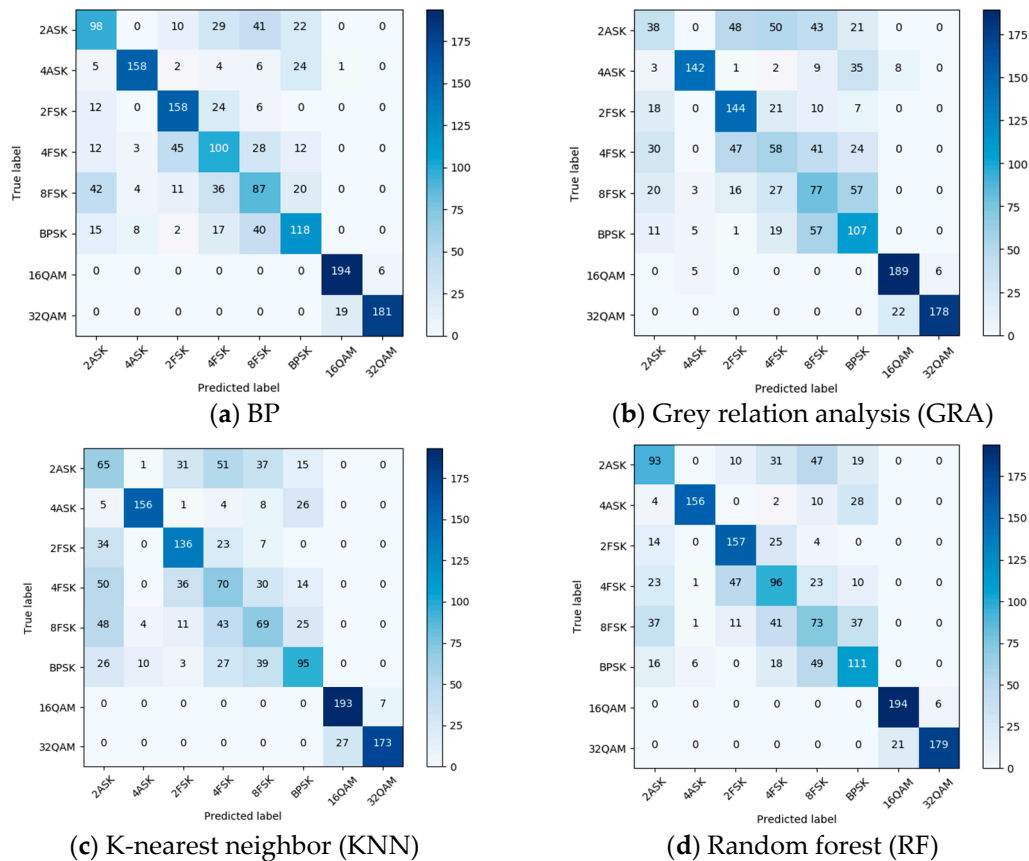The confusion matrices of the eight signals under 0 dB are given in Figure 5.

(a) BP

(b) Grey relation analysis (GRA)

(c) K-nearest neighbor (KNN)

(d) Random forest (RF)

Figure 5. Confusion matrices at 0 dB signal-to-noise ratio (SNR).

The simulation results indicated that the BP neural network and random forest classifiers performed better at SNR = 0 dB. For eight different kinds of communication modulation signals, the 4ASK, 2FSK, 16QAM, and 32QAM had better recognition rates than the others. The recognition rate of 8FSK was the worst. The classification errors of 8FSK were the largest among the eight different signals.

Similarly, in order to analyze the recognition effect under high SNR, the confusion matrices of four classifiers under 10 dB are given in Figure 6.

As shown in Figure 6, among the eight different signals, 8FSK was again the worst recognition rate. At 10 dB, the recognition rate of the random forest classifier was the best. The BP neural network had a great ability to distinguish most signals, but the recognition rate of 8FSK was not very good, the number of correct classifications was 0. The recognition rates of the grey correlation and K-nearest neighbor classifiers were poor, but the K-nearest neighbor classifier was better than the grey correlation classifier. On the whole, compared with 0 dB, the recognition rate was significantly improved at 10 dB SNR.
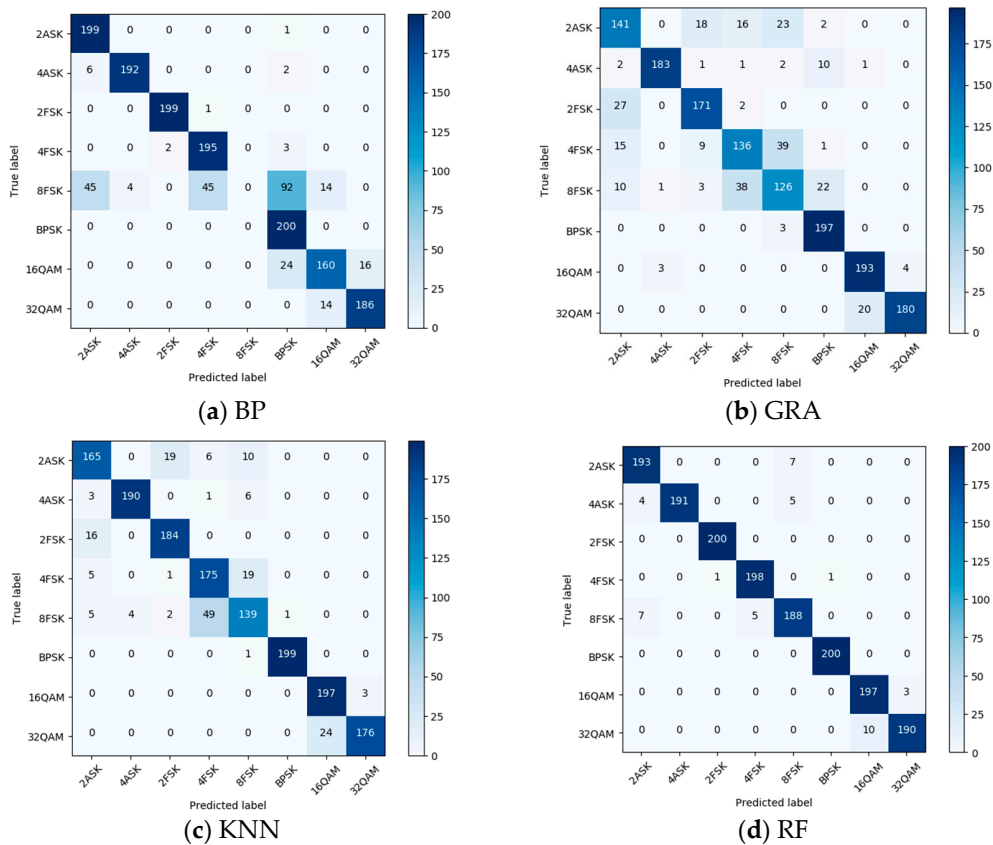
(**a**) BP　　　　　　　　　　　　　(**b**) GRA

(**c**) KNN　　　　　　　　　　　　　(**d**) RF

Figure 6. Confusion matrices at 10 dB SNR.

Figure 7 shows the overall recognition rate curves under different SNRs. It indicates that the recognition rate of the four different classifiers was about 45% at SNR = −5 dB. With increasing SNR, the recognition rate increased continuously. It was stable at SNR = 8 dB. The BP neural network, grey relational, and K-nearest neighbor classifiers stabilized at 85%, but the random forest classifier stabilized at 96%. In addition, the recognition rate of the random forest classifier was better than the others in the whole SNR range. Therefore, the recognition performance of the random forest classifier was optimal.
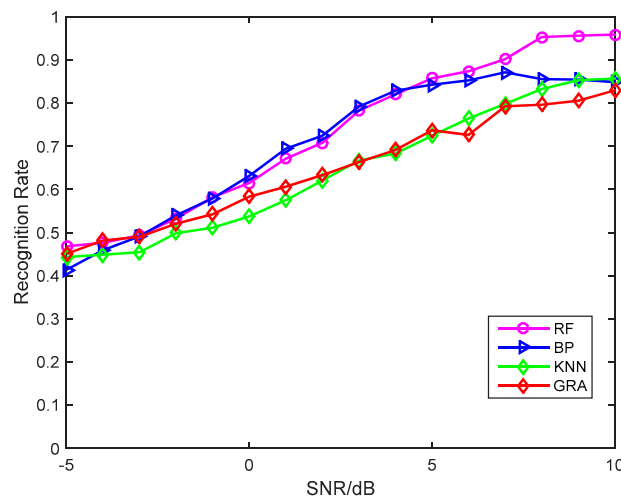

Figure 7. Recognition rate curve under different SNRs.

6. Conclusions

In this paper, a systematic empirical study of fractal theory is used to extract the fractal features of eight different communication modulation signals. Five kinds of fractal features are used, including box fractal dimension, Katz fractal dimension, Higuchi fractal dimension, Petrosian fractal dimension, and Sevcik fractal dimension. Additionally, the evaluation methods of five fractal features are proposed. The noise robustness is analyzed by an anti-noise function. The data distribution is calculated using box diagrams, and the computational complexity is evaluated by the running time. Finally, BP neural network, grey relation analysis, random forest, and K-nearest neighbor classifiers are used to classify the communication modulation signals. The experimental results showed that the recognition rate of random forest could reach 96% at SNR = 10 dB. In addition, the recognition rate of the random forest classifier was superior to the others in the whole SNR range. Therefore, the classification performance of the random forest classifier was optimal.

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

References

1. Liu, K.; Zhang, X.; Chen, Y.Q. Extraction of Coal and Gangue Geometric Features with Multifractal Detrending Fluctuation Analysis. *Appl. Sci.* **2018**, *8*, 463. [CrossRef]
2. Dahap, B.I.; Liao, H.S. Advanced algorithm for automatic modulation recognition for analogue & digital signals. In Proceedings of the International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering, Khartoum, Sudan, 7–9 September 2015; pp. 32–36.
3. Shen, Y.; Liu, X.; Yuan, X. Fractal Dimension of Irregular Region of Interest Application to Corn Phenology Characterization. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 1402–1412. [CrossRef]
4. Zhang, J.; Wang, X.; Yang, X. A method of constellation blind detection for spectrum efficiency enhancement. In Proceedings of the International Conference on Advanced Communication Technology, Pyeongchang, Korea, 31 January–3 February 2016.
5. Wei, D.; Wei, J. A MapReduce Implementation of C4.5 Decision Tree Algorithm. *Int. J. Database Theory Appl.* **2014**, *7*, 49–60.
6. Jiang, L.B.; Liu, Z.Z.; Wen, D.G. A New Recognition Algorithm of Digital Modulation Signals. *Microelectron. Comput.* **2013**, *44*, 112–123.
7. Mehmood, R.M.; Lee, H.J. Emotion classification of EEG brain signal using SVM and KNN. In Proceedings of the IEEE International Conference on Multimedia & Expo Workshops, Turin, Italy, 29 June–3 July 2015; pp. 1–5.
8. Jia, Y.U.; Chen, Y. Digital modulation recognition method based on BP neural network. *Transducer Microsyst. Technol.* **2012**, *5*, 7.
9. Gharehbaghi, A.; Lindén, M. A Deep Machine Learning Method for Classifying Cyclic Time Series of Biological Signals Using Time-Growing Neural Network. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**. [CrossRef] [PubMed]
10. Shi, X.; Liu, Y. Modified Artificial Bee Colony Algorithm Optimizing BP Neural Network and Its Application in the Digital Modulation Recognition. *J. Jiangnan Univ.* **2015**, *4*, 4.
11. Deng, J.L. Introduction to Grey system theory. *J. Grey Syst.* **1989**, *1*, 1–24.
12. Wang, X.; Wang, J.K.; Liu, Z.G.; Wang, B.; Hu, X. Spectrum Sensing for Cognitive Networks Based on Dimensionality Reduction and Random Forest. *Int. J. Signal Process. Image Process.* **2014**, *7*, 443–452. [CrossRef]
13. Wang, X.; Gao, Z.; Fang, Y.; Yuan, S.; Zhao, H.; Gong, W.; Qiu, M.; Liu, Q. A Signal Modulation Type Recognition Method Based on Kernel PCA and Random Forest in Cognitive Network. In Proceedings of the International Conference on Intelligent Computing, Taiyuan, China, 3–6 August 2014; pp. 522–528.
14. Zhang, Z.; Li, Y.; Zhu, X.; Lin, Y. A Method for Modulation Recognition Based on Entropy Features and Random Forest. In Proceedings of the IEEE International Conference on Software Quality, Reliability and Security Companion, Prague, Czech Republic, 25–29 July 2017.

15. Zhang, Y.D.; Chen, X.Q.; Zhan, T.M.; Jiao, Z.Q.; Sun, Y.; Chen, Z.M.; Yao, Y.; Fang, L.T.; Lv, Y.D.; Wang, S.H. Fractal Dimension Estimation for Developing Pathological Brain Detection System Based on Minkowski-Bouligand Method. *IEEE Access* **2016**, *4*, 5937–5947. [CrossRef]

16. Sanchez, J.P.; Alegria, O.C.; Rodriguez, M.V.; Abeyro, J.A.; Almaraz, J.R.; Gonzalez, A.D. Detection of ULF Geoma-gnetic Anomalies Associated to Seismic Activity Using EMD Method and Fractal Dimension Theory. *IEEE Latin Am. Trans.* **2017**, *15*, 197–205. [CrossRef]

17. Zhao, N.; Yu, F.R.; Sun, H.; Li, M. Adaptive Power Allocation Schemes for Spectrum Sharing in Interference-Alignment-Based Cognitive Radio Networks. *IEEE Trans. Veh. Technol.* **2016**, *65*, 3700–3714. [CrossRef]

18. Lin, Y.; Zhu, X.; Zheng, Z.; Dou, Z.; Zhou, R. The individual identification method of wireless device based on dimensionality reduction and machine learning. *J. Supercomput.* **2017**, *74*, 1–18. [CrossRef]

19. Shi, C.; Dou, Z.; Lin, Y.; Li, W. Dynamic threshold-setting for RF-powered cognitive radio networks in non-Gaussian noise. *Phys. Commun.* **2018**, *27*, 99–105. [CrossRef]

20. Yang, Z.; Deng, J.; Nallanathan, A. Moving Target Recognition Based on Transfer Learning and Three-Dimensional Over-Complete Dictionary. *IEEE Sens. J.* **2016**, *16*, 5671–5678. [CrossRef]

21. Prieto, M.D.; Espinosa, A.G.; Ruiz, J.R.; Urresty, J.C.; Ortega, J.A. Feature Extraction of Demag-netization Faults in Permanent-Magnet Synchronous Motors Based on Box-Counting Fractal Dimension. *IEEE Trans. Ind. Electron.* **2011**, *58*, 1594–1605. [CrossRef]

22. Zhou, J.T.; Xu, X.; Pan, S.J.; Tsang, I.W.; Qing, Z.; Goh, R.S.M. Transfer Hashing with Privileged Information. *arXiv*, **2016**, arXiv:1605.04034.

23. Vanthana, P.S.; Muthukumar, A. Iris authentication using Gray Level Co-occurrence Matrix and Hausdorff Dimension. In Proceedings of the International Conference on Computer Communication and Informatics, Coimbatore, India, 8–10 January 2015.

24. Gui, Y. Hausdorff Dimension Spectrum of Self-Affine Carpets Indexed by Nonlinear Fibre-Coding. In Proceedings of the International Workshop on Chaos-Fractals Theories and Applications, Liaoning, China, 6–8 November 2009; pp. 382–386.

25. Guariglia, E. Entropy and Fractal Antennas. *Entropy* **2016**, *18*, 84. [CrossRef]

26. Sevcik, C. A Procedure to Estimate the Fractal Dimension of Waveforms. *Complex. Int.* **1998**, *5*, 1–19.

27. Petrosian, A. Kolmogorov Complexity of Finite Sequences and Recognition of Different Preictal EEG Patterns. In Proceedings of the Computer-Based Medical Systems, Lubbock, TX, USA, 9–10 June 1995; pp. 212–217.

28. Berry, M.V.; Lewis, Z.V.; Nye, J.F. On the Weierstrass-Mandelbrot fractal function. *Proc. R. Soc. Lond.* **1980**, *370*, 459–484. [CrossRef]

29. Guariglia, E. Spectral Analysis of the Weierstrass-Mandelbrot Function. In Proceedings of the 2nd International Multidisciplinary Conference on Computer and Energy Science, Split, Croatia, 12–14 July 2017.

30. Wang, H.; Li, J.; Guo, L.; Dou, Z.; Lin, Y.; Zhou, R. Fractal Complexity-based Feature Extra-ction Algorithm of Communication Signals. *Fractals* **2017**, *25*, 1740008. [CrossRef]

31. Liu, S.; Pan, Z.; Cheng, X. A Novel Fast Fractal Image Compression Method based on Distance Clustering in High Dimensional Sphere Surface. *Fractals* **2017**, *25*, 1740004. [CrossRef]

32. Lin, Y.; Wang, C.; Wang, J.; Dou, Z. A Novel Dynamic Spectrum Access Framework Based on Reinforcement Learning for Cognitive Radio Sensor Networks. *Sensors* **2016**, *16*, 1675. [CrossRef] [PubMed]

33. Liu, S.; Pan, Z.; Fu, W.; Cheng, X. Fractal generation method based on asymptote family of generalized Mandelbrot set and its application. *J. Nonlinear Sci. Appl.* **2017**, *10*, 1148–1161. [CrossRef]

34. Liu, S. Special issue on advanced fractal computing theorem and application. *Fractals* **2017**, *25*, 1740007.

35. Liu, S.; Zhang, Z.; Qi, L.; Ma, M. A Fractal Image Encoding Method based on Statistical Loss used in Agricultural Image Compression. *Multimed. Tools Appl.* **2016**, *75*, 15525–15536. [CrossRef]