

Article

Emitter Identification of Digital Modulation Transmitter Based on Nonlinearity and Modulation Distortion of Power Amplifier

Yue Chen, Xiang Chen and Yingke Lei *

School of Electronic Countermeasures, National University of Defense Technology, Hefei 230000, China; chenyue@nudt.edu.cn (Y.C.); xiaxiang83@126.com (X.C.)

* Correspondence: leiyingke@163.com

Abstract: Specific transmitter identification (SEI) is a technology that uses a received signal to identify to which individual radiation source the transmitted signal belongs. It can complete the identification of the signal transmitter in a non-cooperative scenario. Therefore, there are broad application prospects in the field of wireless-communication-network security, spectral resource management, and military battlefield-target communication countermeasures. This article demodulates and reconstructs a digital modulation signal to obtain a signal without modulator distortion and power-amplifier nonlinearity. Comparing the reconstructed signal with the actual received signal, the coefficient representation of the nonlinearity of the power amplifier and the distortion of the modulator can be obtained, and these coefficients can be used as the fingerprint characteristics of different transmitters through a convolutional neural network (CNN) to complete the identification of specific transmitters. The existing SEI strategy for changing the modulation parameters of a test signal is to mix part of the test signal with the training signal so that the classifier can learn the signal of which the modulation parameter was changed. This method is still data-oriented and cannot process signals for which the classifier has not been trained. It has certain limitations in practical applications. We compared the fingerprint features extracted by the method in this study with the fingerprint features extracted by the bispectral method. When SNR < 20 dB, the recognition accuracy of the bispectral method dropped rapidly. The method in this paper still achieved 86% recognition accuracy when SNR = 0 dB. When the carrier frequency of the test signal was changed, the bispectral feature failed, and the proposed method could still achieve a recognition accuracy of about 70%. When changing the test-signal baud rate, the proposed method could still achieve a classification accuracy rate of more than 70% for four different individual radiation sources when SNR = 0 dB.



Citation: Chen, Y.; Chen, X.; Lei, Y. Emitter Identification of Digital Modulation Transmitter Based on Nonlinearity and Modulation Distortion of Power Amplifier. *Sensors* **2021**, *21*, 4362. <https://doi.org/10.3390/s21134362>

Academic Editor: Xianbin Wang

Received: 21 May 2021

Accepted: 18 June 2021

Published: 25 June 2021

Keywords: specific emitter identification; power-amplifier nonlinearity; modulator distortion; convolutional neural network

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



Copyright: © 2021 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/>).

1. Introduction

SEI is a technology that uses signal-processing methods to extract essential difference information of different transmitters from the received signal, thereby distinguishing each communication transmitter [1]. Since the beginning of the 21st century, the development of sensor networks and mobile communication technologies has also enabled the development of the information society from the electronic age to the Internet of Things era and has promoted the birth of countless communication radiation-source devices. In the civil field, the security of wireless communication systems is receiving exponential attention. According to the characteristics of the physical layer, SEI can determine whether the received signal comes from an authorized transmitter without relying on the user key. This can undoubtedly reduce security incidents caused by key theft. Therefore, it is extensively studied. In the military field, the use of received radio signals to determine from which radio station a signal comes is conducive to grasping the information initiative on the battlefield and supporting communication confrontation.

The key to SEI is finding fingerprint features that represent the difference of specific transmitters. Fingerprint features should be universal, unique, stable, and measurable in a given target set. All individual radiation sources have versatility in fingerprint features. Uniqueness shows that the fingerprint features of individuals on different radiation sources are different. The fingerprint features that need to be extracted for stability have certain antinoise interference capabilities, and the measurability shows that the individual identification of specific radiation sources can be realized in engineering. The fingerprint characteristics of a radiation source are mainly divided into transient and steady-state characteristics. An extracted feature from the signal when the emitter is started is transient, and the extracted feature from a signal that the emitter transmits stable information is steady-state. Lin Y. analyzed transient and steady-state features. He used the power spectral density method and the fractional Fourier transform method to analyze transient signals, and the line integral bispectral method to analyze steady-state signals. Results showed that the SEI method using transient signals had a higher recognition rate than that of the SEI method using steady-state signals under a high signal-to-noise ratio (SNR). However, transient features are susceptible to noise. When the SNR is low, the recognition rate rapidly drops [2] because the transient signal does not contain the interference of sending the message, and it has rich transmitter characteristics. The radio-frequency fingerprint vector extracted from a transient signal can characterize the transmitter device well, but its stability is poor, it is susceptible to noise interference, and it is difficult to obtain a sufficient transient signal in non-cooperative scenarios. This limits its applications in engineering. Due to the presence of modulation information, some features of the steady-state signal are overwhelmed, so the recognition rate of this method is low, but when the SNR is large enough, the features that are not overwhelmed by the message still have very good individual representation performance. Steady-state signals are easy to implement, and it is easy to obtain a sufficient number of steady-state signals. Therefore, steady-state characteristics have always been the main research object of scholars.

For steady-state features, statistical features in the time domain need to first be extracted from a received signal. Reising, Williams, and others extracted the instantaneous amplitude, instantaneous phase, the mean value of instantaneous frequency, variance, kurtosis, and other statistical parameters for the transient and pilot parts of different types of signals (such as GSM, WiMax, and ZigBee signals) [3–6]. Regarding the characteristics of the individual radiation source, experiments verified that the use of statistical time-domain parameter characteristics can identify individual radiation sources well. S. Deng proposed an algorithm of linear skewness and linear kurtosis based on skewness and kurtosis. For limited samples, linear moment estimation is more robust and accurate than other estimation methods are, and even better than maximum-likelihood estimation. Therefore, linear skewness and linear kurtosis are not sensitive to outliers, so non-Gaussian high-precision measurement can be achieved [7]. Research on radiation-source identification methods based on statistical time-domain characteristics is at a relatively early stage, but these methods are susceptible to noise; in this case, they are not enough to analyze non-Gaussian and nonstationary signals. Some subtle features are easily overwhelmed by noise or extracted signal features only reflect noise features. Statistical feature extraction based on high-order moments and high-order spectral parameters can provide richer amplitude and phase information, and this has strong anti-Gaussian noise capabilities, but the amount of calculation is large, so the integrated bispectral feature is often used to characterize individual radiation-source differences. Integrated one-dimensional spectral characteristics can be obtained by integrating the two-dimensional spectrum. According to different integration paths, this can be divided into the radial integrated bispectrum (RIB), axial integrated bispectrum (AIB), square integrated bispectrum (SIB), and circular integral bispectrum (CIB) [1,8–11].

Weak signals such as parasitic modulation and parasitic elements generated by the nonlinearity of radiation-source transmitter equipment are difficult to accurately analyze by time- and frequency-domain analysis methods. Researchers also applied feature extraction

in the signal-transformation domain to the individual identification of communication radiation sources. This type of method attempts to observe or count the characteristics of signal parameters in other domains through various signal conversions, to distinguish between different individual radiation sources. Methods such as wavelet analysis, time-frequency analysis, empirical-mode decomposition (EMD) transformation, and intrinsic time-scale decomposition (ITD) transformation were applied to analyze the slightly different information of individual radiation sources, achieving a good individual recognition effect [1,12]. He. B. decomposed and preprocessed a received signal from EMD, ITD, or variational-mode decomposition (VMD), and then extracted the skewness and kurtosis of the decomposed signal to form a feature vector that characterized the individual signal. The ITD-based method obtained the highest recognition rate in the SEI problem, followed by EMD, and lastly the VMD decomposition method [13]. Jie proposed an SEI recognition method based on a 3D Hilbert energy spectrum and multiscale segmentation; the time-frequency energy spectrum was derived via Hilbert–Huang transform, which could be defined as a complicated curved surface in the three-dimensional space, namely, the 3D Hilbert energy spectrum. Then, via fractal theory, four features were extracted to compose the feature vector under multiscale segmentation. Lastly, the communication of the identification of 13 individual emitters was achieved by utilizing a support vector machine (SVM). When SNR > 26 dB, this method exceeded 90% correct recognition rate based on three simulation datasets [14].

The third type of radiation-feature extraction algorithm is based on the nonlinearity of transmitter hardware. Various analog components inside the communication radiation-source transmitter, such as digital to analog (DAC), power amplifier (PA), modulators, and filters exhibit a certain nonlinearity during operation [15–17]. Aiming at the nonlinear error introduced by DAC [18], Polak used a random Brownian bridge process to model the nonlinear behavior of the radiator transmitter's DAC device [19]. Zhang and Liu used a memoryless polynomial model to describe the power-amplifier nonlinearity of different individual radiation sources [20–22], which simplified the solution of the nonlinear power-amplifier model, but insufficiently describing the memory effect of the power amplifier. To characterize the nonlinear behavior of a power amplifier with a weak memory effect [23], Liu presented a radio-frequency front-end nonlinearity estimator that performed SEI based on the knowledge of a training sequence. The algorithm provided robust identification by first using alternative degrees of nonlinearities associated with symbol amplitudes for initial estimation, and then iteratively estimating the channel coefficients and distorted transmit symbols to overcome the intersymbol interference effect [24]. Jian C. considered that a low-amplitude signal is less affected by the PA, the pulse of the channel response of the low-amplitude symbol could be estimated, and the nonlinearity of the PA could be extracted, reducing the influence of noise iteratively and through larger amplitude symbols. In other words, this method is only applicable to signals with different symbol amplitudes. Simulation results showed that this method could achieve a recognition rate of 90% under the condition of 18 dB [25].

The essential reason for the difference between radiation-source signals is that various device parameters that constitute a radiation source have a certain error range. These subtle parameter differences lead to differences in the amplitude, frequency, skewness, kurtosis, and double-order spectrum of the signal. We can distinguish a single radiation source from the surface phenomenon of the signal, but extracted features cannot characterize the difference in the internal device parameters of the radiation source. Therefore, in this paper, through the mathematical modeling of the distortion of the internal components of the radiation source and the nonlinearity of the power amplifier, parameters characterizing the distortion and nonlinearity are extracted from the received RF signal as the feature vector of different radiation sources to distinguish different radiation sources. First, the received signal was demodulated to obtain the transmitted message bit sequence, and then some basic parameters were obtained (such as carrier frequency, baud rate, modulation mode, and phase) through analysis of the signal. We could use the obtained parameters

and message bit sequence to recover the signal sent by the signal sender through the signal-processing tool. The simulated signal did not pass through the transmitter. We could compare the ‘pure signal’ that we recovered without any nonlinear distortion with the actual received signal. In this way, we know what kind of nonlinear distortion the message bits experience through the transmitter. The entire process is shown in Figure 1. This method extracts the characteristics of the nonlinear parameters of the transmitter and does not change with the change in the transmitted message signal. Through processing any signal, the characteristic parameters of the hardware difference of the transmitter can be obtained, and the signal sent from the transmitter can be individually identified without being affected by the test signal carrier frequency, modulation mode, and modulation rate.

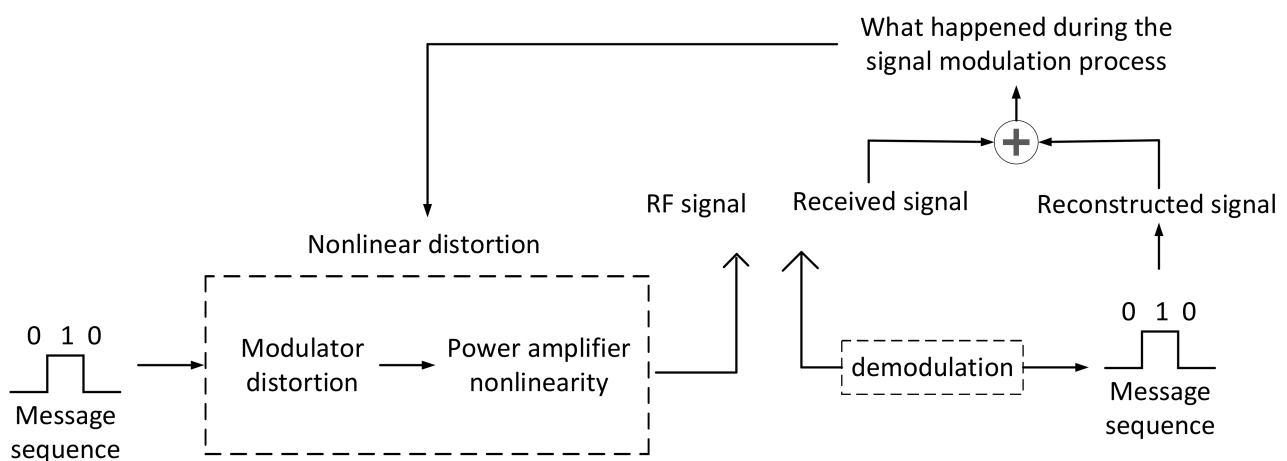


Figure 1. Feature-extraction flowchart.

The rest of this article is arranged as follows: Section 2 first gives the nonlinear model of the power amplifier, establishes the modulator distortion models of PAM and QPSK signals, and presents the feature-extraction formulas under these two modulation modes. It also briefly introduces the traditional feature-extraction integral bispectral algorithm. Section 3 verifies the recognition performance of this method via simulation, and presents the recognition performance of pulse amplitude modulation (PAM) signal and quadrature phase-shift keying (QPSK) signals under the condition of variable carrier frequency and modulation rate. Section 4 presents the results and discussion of the experiment.

2. Modulator Distortion and Nonlinear Power Amplifier (PA) Model Establishment

2.1. Nonlinear PA Model

PA is an important part of various wireless transmitters. In the pre-stage circuit of the transmitter, the power of the RF signal generated by the modulation oscillator circuit is very small, and it needs to pass through the power amplifier to obtain sufficient RF power before it can be fed, radiating towards the antenna. When the power amplifier works in the linear-amplification region, its efficiency is relatively low, but when the power amplifier enters the saturated working region, it produces serious nonlinear distortion, which expands the signal spectrum, interferes with adjacent channels, and produces intermodulation interference [1,23].

The mathematical modeling methods of power amplifiers are mainly divided into memory and nonmemory models. When the input signal is narrow-band, a memoryless model is often used to describe the nonlinear behavior of the power amplifier, for example, the polynomial, Saleh, and RAPP models. When the input signal is wide-band, the nonlinear behavior of the power amplifier is often modeled by a memory model, such as a Volterra series model, a polynomial model with memory, a delayed neural network model, a Hammerstein model, and a Wiener model.

In practical applications, even if the input is a narrow-band power-amplifier signal, there is generally a weak memory effect. Therefore, we use the linear memory (Hammerstein) model to describe the nonlinear distortion of the power amplifier. The Hammerstein models are shown in Figure 2. By introducing intermediate functions, the input–output relationship of the model can be more concise. The content shown in Figure 2 can be expressed by the following formula:

$$a(t) = \sum_{i=1}^M \lambda_i (x(n))^i \quad (1)$$

$$y(t) = \sum_{i=1}^D \alpha_i y(n-i) + a(t) \quad (2)$$

where M is the order of the nonlinear model, λ_i is the model coefficient, D is the order of the memory model, and the corresponding model coefficient is α_i .

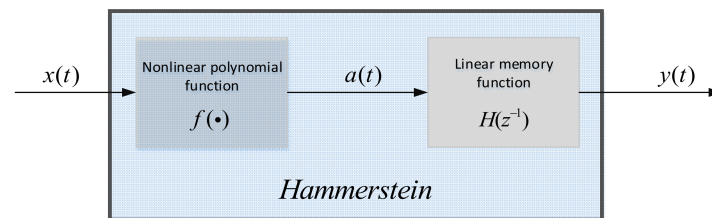


Figure 2. Hammerstein model.

2.2. PAM Modulation-Distortion Model

In digital PAM, the signal waveform can be expressed as:

$$S_m(n) = \text{Re} \left[A_m g(n) e^{j2\pi f_c n} \right] = A_m g(n) \cos(2\pi f_c n) \quad (3)$$

where A_m is the amplitude of the signal, $g(n)$ is the shaped pulse, and f_c is the carrier frequency. In digital amplitude modulation, the distortion of the transmitter includes the following. (a) The deviation of the A_m amplitude. When there are multiple pulse amplitudes, different radiation sources produce the same amplitude due to the allowable error of the digital-to-analog converter (DAC) and other devices. There are subtle differences. (b) The nonlinearity of the power amplifier. The radio-frequency signal is amplified by the PA, which produces nonlinear distortion. The amplitude offset produced by ΔA is represented, and we used the linear memory Hammerstein model to describe the nonlinearity of the power amplifier:

$$x(n) = \text{Re} \left\{ \xi(n) e^{j\omega n} \right\} = \frac{1}{2} \left[\xi(n) e^{j\omega n} + \xi^*(n) e^{-j\omega n} \right] \quad (4)$$

where $\xi(n) = (A_m + \Delta A)g(n)$ is the equivalent baseband complex envelope signal.

$$a(n) = \sum_{i=1}^M \lambda_i (x(n))^i = \sum_{i=1}^M \lambda_i \sum_{k=0}^i \frac{1}{2^i} C_i^k (\xi(n))^k (\xi^*(n))^{i-k} e^{j\omega n(2k-i)} \quad (5)$$

There is a frequency multiplication component in $a(n)$, and the actual signal, except for the center frequency passband, is filtered out by the filter; therefore, $2k - i = \pm 1$. Formula (5) shows that i is the only term with retained odd numbers. Therefore, the equivalent low-pass of the above formula can be written as:

$$a'(n) = \sum_{i=1}^{(M-1)/2} \frac{1}{2^{2i}} C_{2i+1}^{i+1} \lambda_{2i+1} (\xi(n))^{i+1} (\xi^*(n))^i \quad (6)$$

Combined with the memory-term formula, the equivalent output of the power amplifier of the Hammerstein model is:

$$y(n) = \sum_{i=1}^D \alpha_i y(n-i) + \xi(n) \sum_{i=0}^{(M-1)/2} \frac{1}{2^{2i}} C_{2i+1}^{i+1} \lambda_{2i+1} |\xi(n)|^{2i} \quad (7)$$

The above formula can be written as:

$$y(n) = \mathbf{u}(n) \mathbf{h} \quad (8)$$

where $\mathbf{u}(n)$ is only related to $y(n-i)$ and $\xi(n)$, and $\mathbf{h}$ is only related to model parameters, not related to power amplifier input:

$$\begin{aligned} \mathbf{u}(n) &= [\xi(n), \xi(n)|\xi(n)|^2, \dots, \xi(n)|\xi(n)|^{M-1}, y(n-1), \dots, y(n-D)] \\ \mathbf{h} &= [\mu_0, \dots, \mu_{(M-1)/2}, \alpha_1, \dots, \alpha_D]^T \\ \mu_i &= \frac{1}{2^{2i}} C_{2i+1}^{i+1} \lambda_{2i+1} \quad i = (0, \dots, (M-1)/2) \end{aligned}$$

At the receiving end, $\mathbf{u}(n)$ in Formula (8) is only related to the current input $\xi(n)$ of the power amplifier and output $y(n-i)$ of the power amplifier sometime before, where $y(n-i)$ can be observed, and $\xi(n)$ can be obtained by demodulating and reconstructing the signal. Therefore, $\mathbf{u}(n)$ is observable. Assuming the channel is additive white Gaussian noise (AWGN), Formula (8) can be written as:

$$y(n) = \mathbf{u}(n) \mathbf{h} + v_n \quad (9)$$

where v_n is zero-mean Gaussian white noise. For observation time steps N , the above formula is written in matrix form:

$$\mathbf{y} = \mathbf{U} \mathbf{h} + \mathbf{V} \quad (10)$$

where

$$\begin{aligned} \mathbf{y} &= [y(n), \dots, y(n-N)]^T, \mathbf{U} = \begin{bmatrix} \mathbf{u}(n) \\ \vdots \\ \mathbf{u}(n-N) \end{bmatrix} \\ \mathbf{u}(n-j) &= [\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{(M-1)/2}, y(n-j-1), \dots, y(n-j-D)] \quad (j = 0, \dots, N) \\ \varepsilon_i &= \xi(n-j) |\xi(n-j)|^{2i} \quad (i = 0, \dots, (M-1)/2) \end{aligned}$$

Formula (10) can be solved by the least-squares method; then,

$$\mathbf{h} = \mathbf{U}^H (\mathbf{U} \mathbf{U}^H)^{-1} \mathbf{y} \quad (11)$$

Considering the impact of amplitude deviation,

$\xi(n) = (A_m + \Delta A)g(n) = \xi(n) + \Delta A g(n)$, where $g(n)$ is a rectangular pulse, ΔA is a constant, and $\Delta A g(n)$ can be considered as a constant $A = \Delta A g(n)$. The final effect is that there is a certain deviation between the amplitude of the actual signal and the amplitude of the reconstructed signal.

$$\mathbf{y} - A = \mathbf{U} \mathbf{h} + \mathbf{V} \quad (12)$$

We traverse A when A obtains optimal solution A' , $A' = \underset{A}{\operatorname{argmin}} v(\mathbf{h})$, where $v(\mathbf{h})$ represents the variance of solution feature $\mathbf{h}$. After searching and solving A' , Formula (11) is substituted to obtain an individual characteristic vector $\mathbf{h}$ of the radiation source. The final feature vector is $\mathbf{e} = [A', \mathbf{h}]$.

2.3. Quadrature Modulation Distortion Mode

The flow chart of quadrature-phase modulation is shown in Figure 3. The fingerprint characteristics of the radiation source usually come from modules composed of analog devices, so the analog circuit modules after the DAC module shown in the figure may generate the radiation-source fingerprint. The error factor of the in-phase/quadrature (I/Q) modulator includes the four following parts.

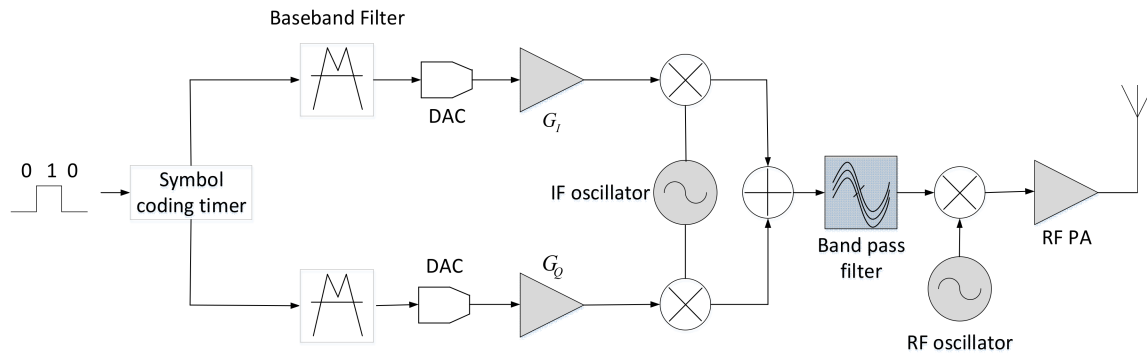


Figure 3. Phase-modulation flowchart.

- I/Q gain imbalance

This is mainly caused by the difference in the gain characteristics of the components on the I/Q two channels, which are represented as G_I and G_Q in Figure 4a; this leads to the phase point to change within the quadrilateral of the ‘abcd’ siege, as shown in Figure 4a.

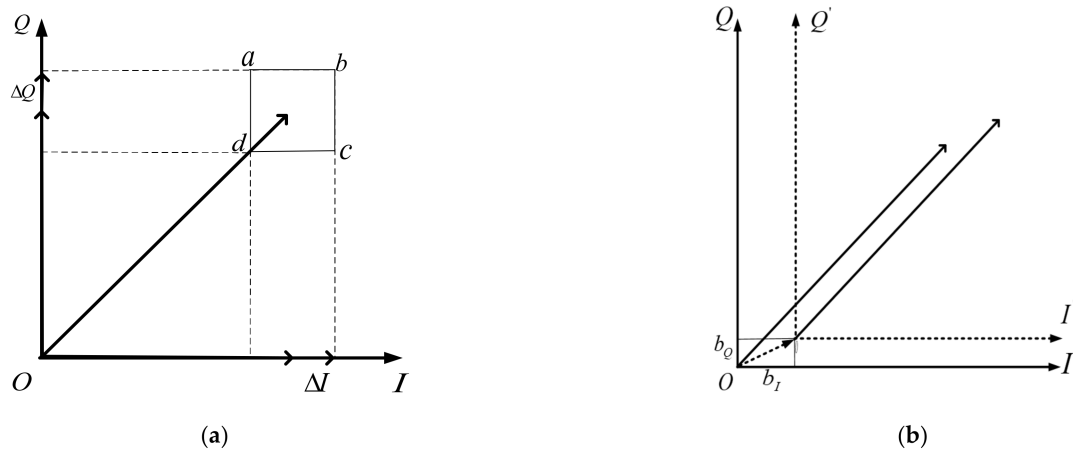


Figure 4. Phase deviation caused by modulator distortion. (a) Offset of phase center point caused by I/Q gain imbalance; (b) offset of origin of phase coordinate caused by DC offset.

- I/Q DC offset and carrier leakage

This is mainly caused by the DC offset generated by the amplifier on the I/Q path and the carrier leakage generated by the mixer; as shown in Figure 4, the DC offset causes the origin of the phase coordinate to drift and cause phase errors.

- I/Q delay mismatch

This is mainly due to the difference in the delay characteristics of the analog components (amplifiers, mixers) on the I/Q channel.

- I/Q quadrature error

When the local oscillator generates two orthogonal local carriers due to the characteristic error of the components, the phase difference is not 90° , that is, the quadrature error. These errors distort the constellation diagram on the demodulation constellation diagram, and the error characteristics of different individual transmitters are often different.

For the I/Q quadrature modulator, by gain imbalance $G_m = G_I / G_Q$, the DC offset of the I and Q channels is denoted by b_I and b_Q , respectively. The delay mismatch is τ , and the quadrature error is denoted as ε . Then, I channel signal $I(n)$ and Q channel signal $Q(n)$ can be expressed as:

$$I(n) = G_m \cos \left(2\pi \sum_i h_i d_i g(n - iT_d) \right) + b_I \quad (13)$$

$$Q(n) = \sin \left(2\pi \sum_i h_i d_i g(n - iT_d - \tau) \right) + b_Q \quad (14)$$

where h is the modulation coefficient, d_i is the symbol sequence generated by the symbol encoder, and $g(n)$ is the instantaneous pulse. The signal obtained by quadrature modulation is:

$$x(n) = I(n) \cos(\omega_c n + \varepsilon/2) + Q(n) \sin(\omega_c n - \varepsilon/2) \quad (15)$$

Substituting Formulas (13) and (14) into Formula (15) produces:

$$x(n) = x_m(n) e^{j2\pi f_0 n} = e^{j2\pi f_0 n} (u_m \xi(n) + v_m \xi^*(n) + b_I + b_Q) \quad (16)$$

where f_0 is the residual frequency offset, $\xi(n) = e^{j2\pi \sum_i h_i d_i g(n - iT_d)}$; let $B = b_I + b_Q$ be the DC offset, and

$$\begin{aligned} u_m &= \frac{1}{2}(G_m + 1) \cos(\varepsilon/2) + \frac{1}{2}j(G_m - 1) \sin(\varepsilon/2) \\ v_m &= \frac{1}{2}(G_m - 1) \cos(\varepsilon/2) + \frac{1}{2}j(G_m + 1) \sin(\varepsilon/2) \end{aligned} \quad (17)$$

Combined with the Hammerstein model of the power amplifier, the effect of bias is not considered, $x(n) = u_m \xi(n) + v_m \xi^*(n)$, the equivalent discrete form of the output signal of the power amplifier is:

$$\begin{aligned} y(n) &= \sum_{i=1}^D \alpha_i y(n-i) + \sum_{i=1}^M \lambda_i (x(n))^i \\ &= \sum_{i=1}^D \alpha_i y(n-i) + \sum_{i=1}^{(M-1)/2} \frac{1}{2^i} C_{2i+1}^{i+1} \lambda_{2i+1} \sum_{j=-i}^{i+1} \left(\sum_{k=0}^i C_{i+1}^{i+k} C_i^k \chi_{j+k,k}^{i+1-j-k,i-k} \right) \kappa_{i+j,i+1-j}^n \end{aligned} \quad (18)$$

where D and M are the AR order and polynomial order in the Hammerstein model, $\alpha = [\alpha_1, \dots, \alpha_D]$; $\lambda = [\lambda_1, \dots, \lambda_M]$ are the AR coefficient and polynomial coefficient, respectively, and, κ and χ are defined as follows:

$$\kappa_{a,b}^n = (\xi(n))^a (\xi(n)^*)^b, \chi_{a,b}^{c,d} = (u_m)^a (u_m^*)^b (v_m)^c (v_m^*)^d$$

Formula (18) can also be written as:

$$q(n) = \mathbf{u}(n) \mathbf{h} + \mathbf{v} \quad (19)$$

where

$$\begin{aligned} u_n &= [\rho_0^n, \dots, \rho_{M-1/2}^n, y_n'] \in \mathbb{R}^{1 \times ((M+1)(M+3)/4 + D)} \\ h &= [\eta_0, \eta_1, \dots, \eta_{M-1/2}, \alpha]^T \in \mathbb{R}^{1 \times ((M+1)(M+3)/4 + D)} \\ \rho_i^n &= [\kappa_{0,2i+1}^n, \dots, \kappa_{j+i,i+1-j}^n, \dots, \kappa_{2i+1,0}^n] \quad (i = 0, \dots, M-1/2) \\ y_n' &= [y(n-1), \dots, y(n-D)] \\ \eta_i &= \frac{1}{2^{2i}} C_{2i+1}^{i+1} \lambda_{2i+1} \left[\chi_{0,i}^{i+1,0}, \dots, \sum_{k=0}^i C_{i+1}^{i+k} C_i^k \chi_{j+k,k}^{i+1-j-k,i-k}, \dots, \chi_{i+1,0}^{0,i} \right] \\ &\quad (i = 0, \dots, M-1/2, j = -i, \dots, i+1) \end{aligned}$$

For observation period N , the above formula can be written in a concise matrix form:

$$\mathbf{y} = \mathbf{U}\mathbf{h} + \mathbf{v} \quad (20)$$

where

$$\mathbf{y} = [y(n), \dots, y(n-N)]^T, \mathbf{U} = \begin{bmatrix} \mathbf{u}_n \\ \vdots \\ \mathbf{u}_{n-N} \end{bmatrix}$$

Considering the impact of DC bias, $x(t) = u_m \xi(t) + v_m \xi^*(t) + B$. Bias B is a real number, and the final effect is that a certain deviation exists between the amplitude of the signal that did not pass through the power amplifier, and the amplitude of the reconstructed signal. We traverse B when B obtains optimal solution $B' = \underset{B}{\operatorname{argmin}} v(\mathbf{h})$, where $v(\mathbf{h})$ represents the variance of solution feature $\mathbf{h}$. After searching and solving B' , Formula (20) is substituted to obtain an individual characteristic vector $\mathbf{h}$ of the radiation source. The final feature vector is $\mathbf{e} = [B', \mathbf{h}]$.

2.4. Integral Bispectrum

High-order spectral analysis is widely used in signal processing. In theory, a high-order spectrum can completely suppress any Gaussian noise and symmetrically distributed non-Gaussian noise. It can retain signal amplitude and phase information and is independent of time. Therefore, high-order spectral analysis is the current mainstream feature extraction method [11,26,27]. The third-order spectrum is the simplest high-order spectrum, also known as the bispectrum. Though the processing method is relatively simple, it can describe the nonlinear characteristics of the signal. For the digital zero-IF I/Q signal $r(t)$ of the communication radiation source, its bispectrum can be defined as:

$$\mathbf{B}(\omega_1, \omega_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} c_{3r}(\tau_1, \tau_2) e^{[-j(\omega_1 \tau_1 + \omega_2 \tau_2)]} d\tau_1 d\tau_2 \quad (21)$$

where $c_{3r}(\tau_1, \tau_2) = E\{r^*(t)r(t+\tau_1)r(t+\tau_2)\}$ is the third-order cumulant of $r(t)$.

Compared with the power spectrum, it can provide phase information and is widely used. Integrating the bispectrum is the best way to reduce the dimensionality of two-dimensional bispectral features. According to different integration paths, it can be divided into RIB, AIB, SIB, and CIB. The integration path of each integral bispectrum is illustrated in Figure 5b.

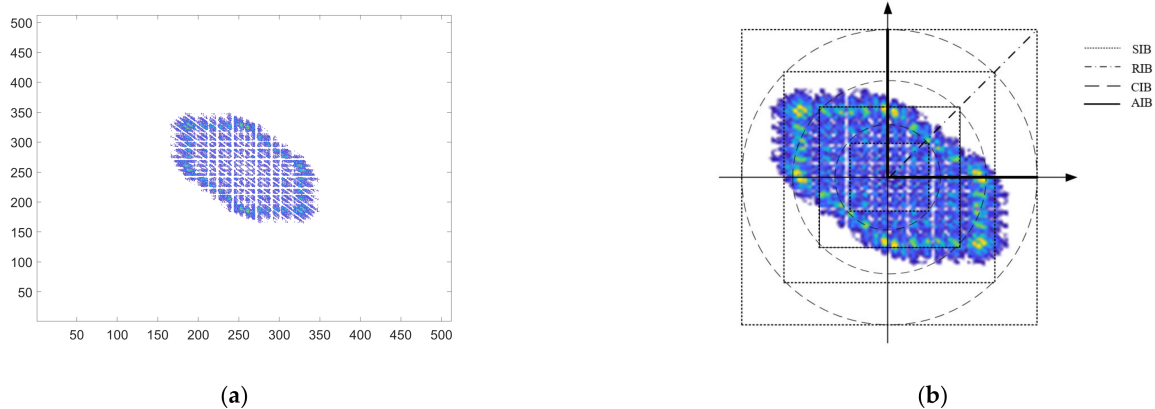


Figure 5. (a) Result of bispectral analysis of signal by Emitter 1 under signal-to-noise ratio of 15 dB. (b) Four paths for integrating two-dimensional bispectral features.

3. Simulation and Experiment Results

3.1. Simulation Verification of PAM-Modulated Radiation Source

To verify the identification performance of the method on the ASK signal, the amplitude deviation of the transmitter and the nonlinearity of the power amplifier were simulated through a simulation signal. The first selection of λ_i and α was determined by our demodulation and reconstruction of the real signal, and by fitting. Then, the subtle differences on the fitted data were randomly changed to become our four simulated radiation-source power-amplifier parameters, and amplitude error ΔA was added on this basis.

The simulation signal constructed four radiation-source PAM signals with different parameters, five times oversampling; the radiation-source simulation parameter setting table is shown in Table 1. Figure 6 shows that polynomial order M of the nonlinear model of each radiation-source power amplifier and order D of the memory model was set to 11 and 2, respectively [1]. The number of signal samples of each radiation source was 1000, which were proportionally divided into training and test sets. By adding Gaussian white noise, the signal-to-noise ratio varied from 0 to 35 dB.

Table 1. ASK-modulated radiation-emitter parameter setting.

Parameter	Emitter 1	Emitter 2	Emitter 3	Emitter 4
λ_1	1	1	1	1
λ_3	−42.38	−40.29	−44.57	−88.62
λ_5	2686.28	2409.38	2028.72	2242.64
λ_7	16,675	16,420	15,892	17,024
λ_9	−1652.42	−1589.88	−1892.46	−1686.27
λ_{11}	−5648.64	−6278.96	−7224.84	−5946.56
α_1	1.9122	1.9108	1.9048	1.9143
α_2	−0.9794	−0.9986	−0.9042	−0.9342
ΔA	0.1	0.15	0.08	0.1

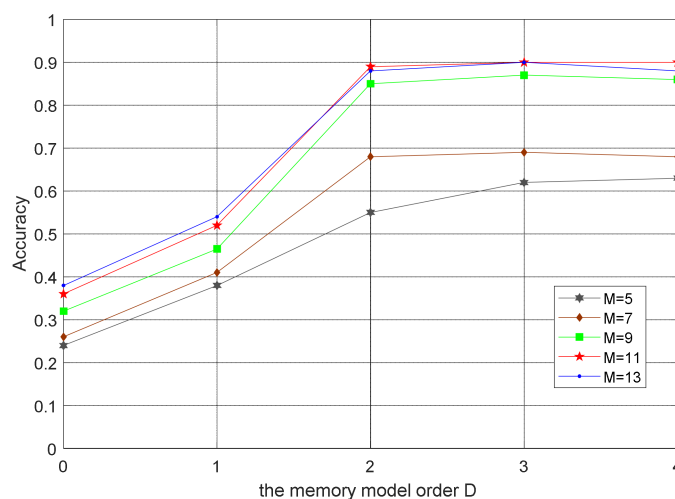


Figure 6. Influence of different model structures on accuracy.

We used software to simulate four simple PAM signals concerning the above nonlinear coefficients of the power amplifier. The simulated signal carrier was $f_c = 1$ MHz, the information rate was 100 bit/s, and the sampling rate was five times the carrier frequency. The simulated signal is shown in Figure 7.

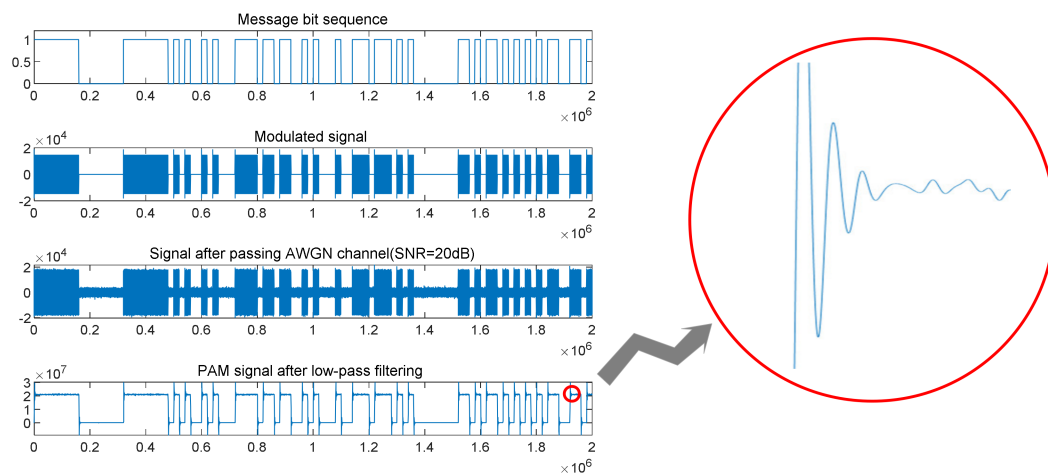


Figure 7. Signal simulation flowchart: entire signal process from original message bits to demodulation after passing through the channel; partial signal enlargement after low-pass filtering. Jitter reflects nonlinear characteristics of the power amplifier in the signal-amplification process.

3.1.1. Emitter Separability Analysis

The feature vector $\mathbf{e}$ of the radiation source was extracted using Formulas (12) and (20). We chose commonly used classifier algorithms by scholars, namely, convolutional neural network (CNN) and support vector machine (SVM). The distortion and nonlinear parameters were extracted and compared with the two classifiers.

Since extracted features $\mathbf{e}$ were only nine dimensions, we used three convolutional layers, two fully connected layers, and softmax layers to form our neural-network structure; the ‘padding’ method of the CNN was the same, the size of the convolutional kernel was 2×2 , the stride was 1, pooling size was 2×2 , and the stride was 1.

The SVM algorithm was originally designed for binary classification problems. When dealing with multiclass problems, it is necessary to construct a suitable multiclass classifier. At present, there are two main methods for constructing SVM multiclass classifiers: first, the direct method, which directly modifies the objective function, combines the parameter solving of multiple classification surfaces into an optimization problem and realizes multiclass classification by solving the optimization problem once. This method seems simple, but its computational complexity is relatively high, and it is relatively difficult to implement. It is only suitable for small problems.

Second, the indirect method, which mainly combines two classifiers to achieve the construction of multiple classifiers. The method is to design an SVM between any two types of samples, so $k(k-1)/2$ SVMs need to be designed for k types of samples. When classifying an unknown sample, the category with the most votes is the category of the unknown sample. To classify four different radiation sources, we needed to design 6 SVM classifiers. The randomly selected training samples for each category were 800, and the test samples were the remaining 200. The kernel function of the SVM classifier was ‘rbf’, which is used to implement nonlinear classification.

Under different signal-to-noise ratio conditions, the recognition accuracy of the two classification methods is shown in Figure 8.

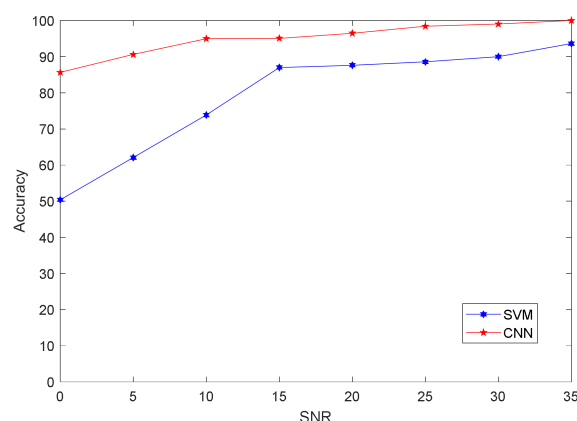


Figure 8. Accuracy of ASK signal features under different signal-to-noise ratio conditions through CNN and SVM classifiers.

As illustrated in Figure 8, the CNN classifier algorithm was generally better than the SVM classifier algorithm, and the CNN algorithm could more strongly distinguish with a low signal-to-noise ratio. When the signal-to-noise ratio was greater than 15 dB, the classification accuracy rates of the two classification algorithms, SVM and CNN, exceeded 80%. Figure 9a shows the three-dimensional visualization of the 1-, 7-, and 9-dimensional vectors of the individual characteristics of the four radiation sources extracted by the algorithm when the signal-to-noise ratio was 15 dB. Figure 9b is the confusion matrix given by CNN in this signal-to-noise ratio. Figure 9a shows that radiation sources 4 and 1 were classified incorrectly. Generally speaking, the feature vector obtained by this method could discriminate well.

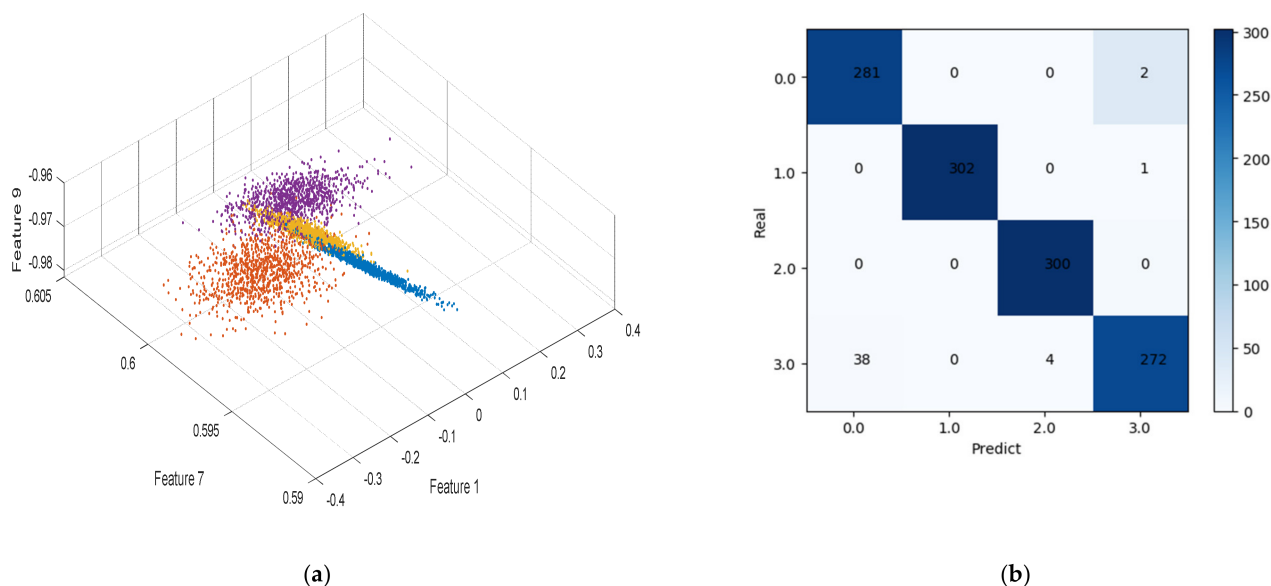


Figure 9. With a SNR ratio of 15 dB: (a) three-dimensional visualization map that composed 1-, 7-, and 9-dimensional data of feature vector e ; (b) confusion matrix corresponding to CNN classification result.

3.1.2. Comparison Experiment of Different Feature-Extraction Algorithms

In [7,10], the authors used the skewness and kurtosis algorithm to extract fingerprint features of a radiation source hidden in the signal. The authors in [10] used EMD, ITD, and VMD to decompose the signal, and then extracted the skewness and kurtosis information of each component as the fingerprint feature of the radiation source. In [26], the authors

optimized skewness and kurtosis, and proposed a method for identifying individual characteristics of radiation sources based on linear skewness and linear kurtosis, which achieved a 100% distinction of three transmitters. The authors in [24] used the combination of an integral SIB bispectrum and a CNN to achieve a good recognition effect on five simulation data.

Figure 10 shows the recognition-accuracy results of the proposed method and the integrated bispectrum through the CNN under different signal-to-noise ratio conditions. The figure shows that, when the SNR was less than 20 dB, the recognition rate of the proposed method was higher than that of the bispectral method. When the SNR was greater than 20 dB, the two methods had similar levels of recognition accuracy. This also shows that the bispectral method could achieve better recognition performance under a high signal-to-noise ratio.

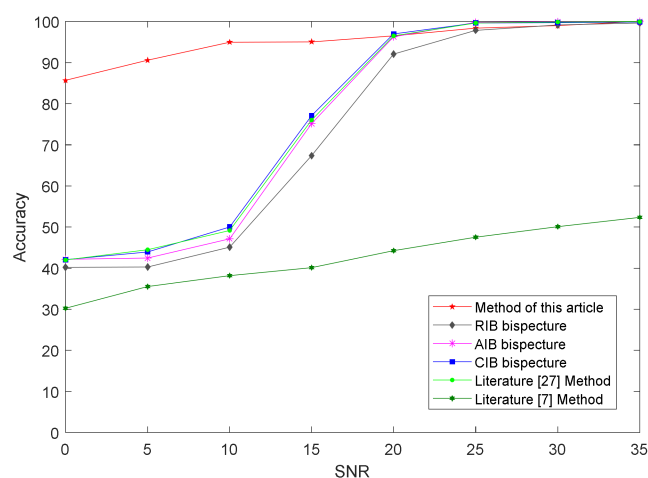


Figure 10. Recognition results of integral bispectrum algorithms, the algorithm in [27], and the proposed algorithm under different signal-to-noise ratio conditions.

3.1.3. Changing Carrier Frequency

At present, individual radiation-identification methods use training-set signals to train the classifier, and then use signals with the same modulation parameters as the training set as the test set for testing. The actual radiation source can change the modulation parameters of the transmitted signal, such as changing the carrier frequency, modulation rate, and modulation mode of the transmitted signal. In this way, the traditional radiation-source identification method is no longer applicable because the classifier is not trained to change the modulation parameters of the signal. Changes in these modulation parameters make some existing research methods no longer have good robustness.

The method in this paper extracts the distortion parameters of the inherent hardware of the transmitter, which does not change with a change in the carrier frequency, rate, and modulation mode of the transmitted signal. Therefore, it can recognize variable modulation parameters well. To verify the robustness of the proposed method under different parameter settings, we changed the carrier frequency of the test data. The carrier frequency of the training set was $f_c = 1$ MHz, and the carrier frequency of the test set was set to $f_c = 1, 1.02, 1.04$, and 1.06 MHz. Figure 11 shows the relationship between recognition accuracy and SNR after changing the carrier frequency of the test set using the proposed method and the dual-spectrum method. After the two methods changed the test set frequency, the recognition accuracy was reduced.

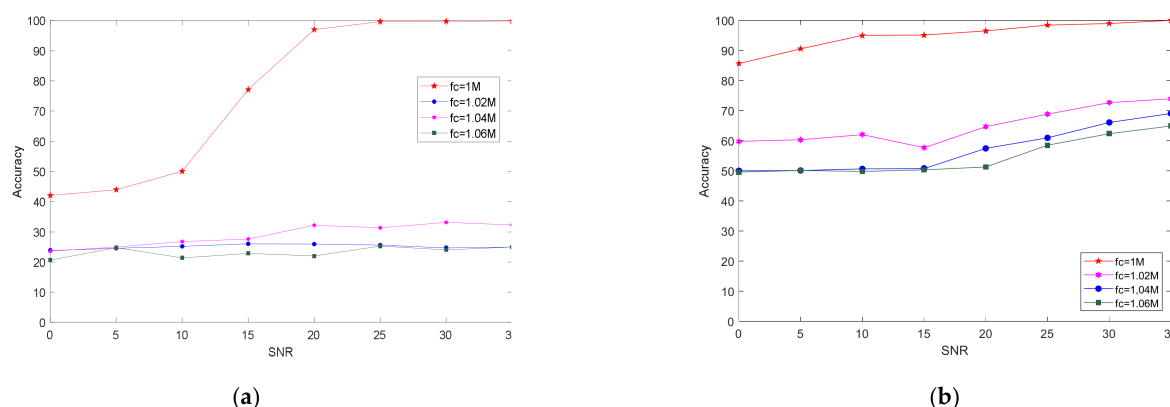


Figure 11. Experimental results when using different carrier-frequency data and (a) bispectral method or (b) proposed method as the test set.

The specific reasons are analyzed as follows: for the integrated bispectral algorithm, the bispectral feature extracted information such as the frequency and phase of the signal. When using a test signal that was different from the carrier frequency of the training set, due to the test set signal frequency and the training set signal frequency, the difference was large. Although the signals were emitted by the same transmitter, their bispectral characteristics were quite different, which failed to achieve the purpose of identifying individual radiation sources. For the algorithm in this paper, changing different carrier frequencies, when we demodulated and reconstruct the signal, due to the error in the estimation of the carrier frequency and the estimation of the symbol rate, the reconstructed message signal had a certain error, which led to a decline in recognition accuracy.

Figure 12 shows that, in the classification experiment of the four individual radiation sources, the SIB algorithm had a recognition accuracy of about 25% after changing the carrier frequency of the test set. The bispectral method appeared to not face the test signal with new modulation parameters to be able to classify individuals. The proposed method still had 50% accuracy when the signal-to-noise ratio was 0 dB. When the signal-to-noise ratio was high enough, recognition accuracy could exceed 70%. Figure 12 shows that, when the frequency of the test set was farther away than $f_c = 1\text{ MHz}$, recognition accuracy showed a downward trend.

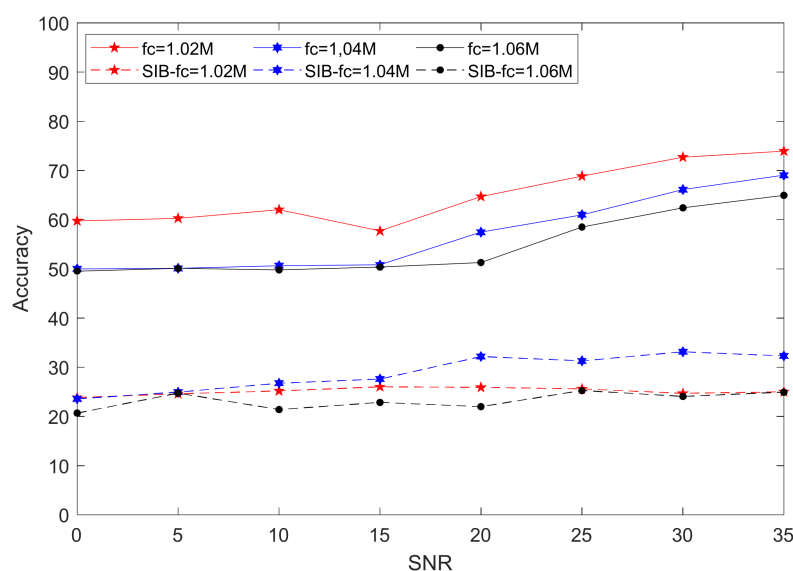


Figure 12. Relationship between recognition accuracy and SNR of the proposed method and bispectral method on test-set data of different carrier frequencies.

3.1.4. Change in Information Rate

To verify the robustness of the proposed method under different parameter settings, we changed the information rate of the test data. The information rate of the data we use for training was 100 bit/s, and the information rate of the test data was $R = 60, 80, 100, 120$, and 140 bit/s. Experiment results are shown in Figure 13 below.

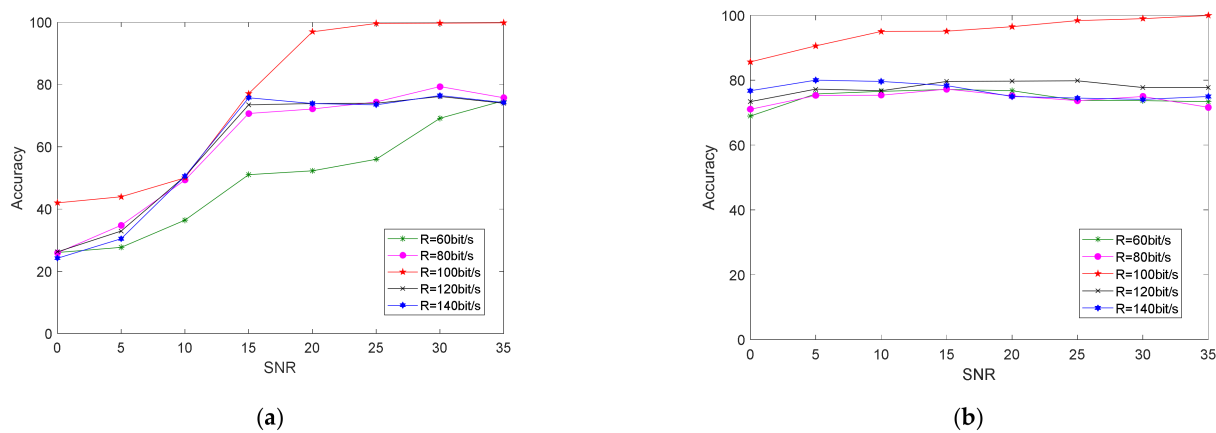


Figure 13. After changing the data information rate of the test set, the classification accuracy results using (a) bispectral method and (b) proposed method under different signal-to-noise ratio conditions.

As shown in Figure 14, after changing the message rate of the test set, the recognition accuracy of the two methods decreased. When the signal-to-noise ratio was greater than 15 dB, the recognition accuracy of the two algorithms was maintained at 70–80%.

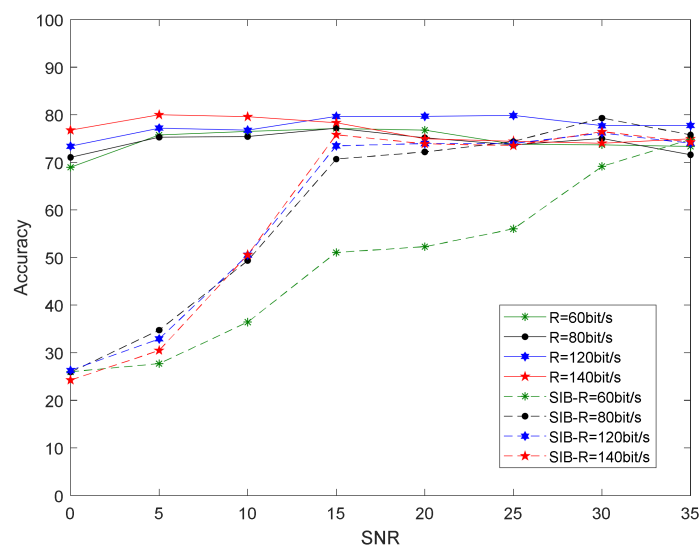


Figure 14. Classification-accuracy results of the proposed method and bispectral method under different SNR conditions after changing test-set information rate.

When changing the message rate, because the bispectral feature extracted the information of the signal frequency and phase, it had certain robustness to the change in message rate. Both algorithms had a certain drop in accuracy because, when the message rate changed, the length of the baseband shaping signal also changed. This means that the proportion of unstable points caused by the distortion of the modulator and the jitter caused by the nonlinearity of the power amplifier in the entire piece of data changed, as shown in Figure 15 (α_1 and α_2). The proposed method needs to use all data points to

calculate the final feature vector e . Therefore, changing the message rate of the test set produced a certain error in the results.

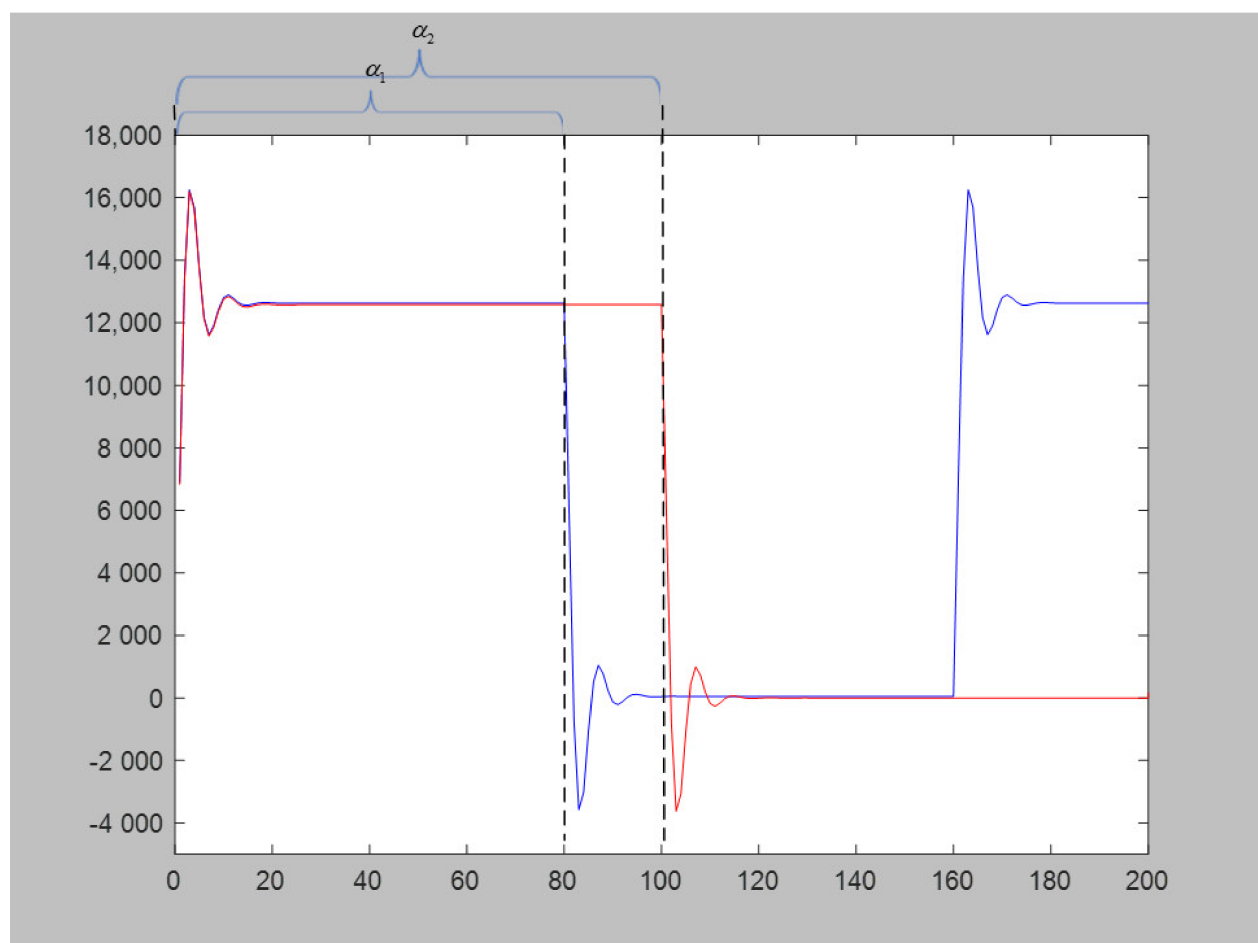


Figure 15. Baseband signal after passing through the power amplifier.

3.2. Simulation and Verification of QPSK Modulated Radiation Source

To verify the ability of this method to identify QPSK signals, we simulated four QPSK signals, and the nonlinear parameters of the power amplifier were set as in the previous experiment. On this basis, unique gain imbalance G_m , DC offset, carrier leakage B , and quadrature error ε of the QPSK signal were added. Parameter settings are shown in Table 2.

Table 2. QPSK-modulated radiation-source parameter setting.

Parameter	Emitter 1	Emitter 1	Emitter 3	Emitter 4
G_m	0.9	1	1.2	1.3
B	0.1	0	0.12	0.05
ε	1°	0°	1°	0°

By setting the value of the parameter, we could obtain a model with only power nonlinearity, a model with only modulator distortion, and a model with both power amplifier nonlinearity and modulator distortion. Experiment results obtained are shown in Figure 16.

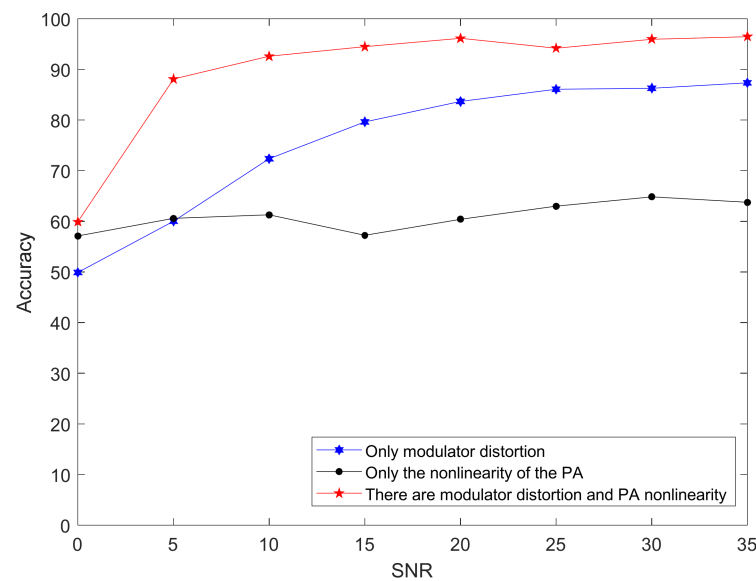


Figure 16. Classification-accuracy rate of three different models through CNN classifier under different signal-to-noise ratio conditions.

The Figure 17 shows that, when considering both the modulator distortion and the nonlinearity of the power amplifier, the model had the highest recognition rate under any SNR condition. When only considering the power amplifier, the accuracy rate did not change much with the increase in signal-to-noise ratio, indicating that the power amplifier did not have much influence on the signal because the QPSK signal was a constant envelope signal, and the phase of the signal was modulated. There was little correlation with the amplitude of the signal.

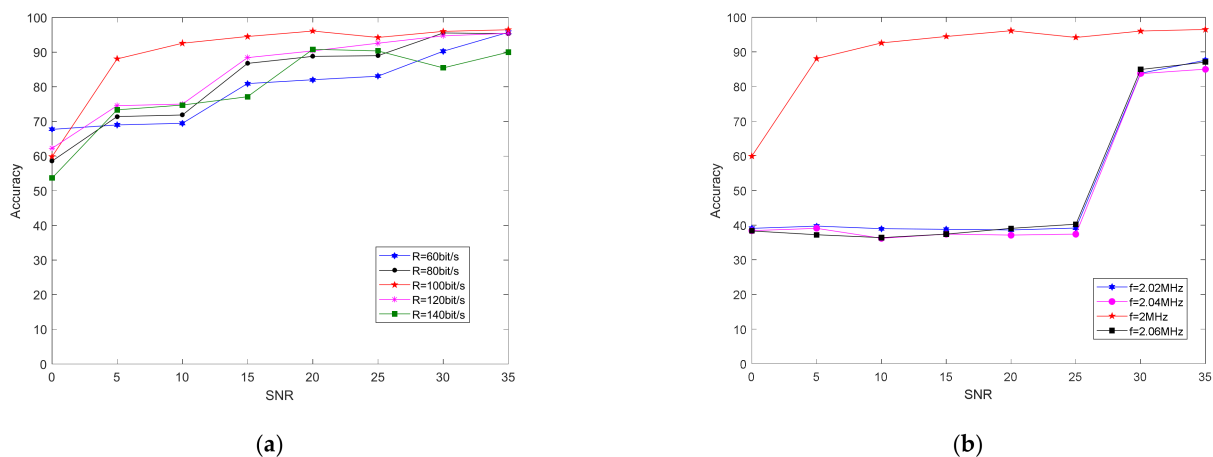


Figure 17. Classification-accuracy results of the proposed method under different signal-to-noise ratio conditions after changing (a) the data information rate and (b) the carrier frequency of the test set.

4. Discussion

By demodulating the digital modulation signal, the initial information sequence was obtained; then, by reconstructing the message signal, a signal without any modulation distortion and power amplifier nonlinearity were obtained. By comparing the deconstructed signal with the actual received RF signal through the feature-extraction algorithm, the parameter representation of the difference between the two could be obtained. The parameter that characterized the difference between the two was the feature vector that we needed, which can represent the difference between individual radiation sources. The proposed

method puts forward the reason for the difference in a radiation-source signal rather than in the appearance. Through the extraction and identification of distortion parameters, the effective identification of individual radiation sources is achieved.

The traditional SEI idea is to use training-set signals to train the classifier, and then use the signal with the same modulation parameters as the training set as the testing signal. When the test- and training-set signals come from the same transmitter but have different modulation parameters, changes in these modulation parameters make some existing research methods no longer have good robustness. The proposed method extracts the distortion parameters of the inherent hardware of the transmitter, and the distortion parameters do not change with changes in carrier frequency and the rate of the transmitted signal. Therefore, it can identify variable modulation parameters well.

The characteristics of the transmitter extracted by the proposed method can represent the characteristics of the transmitter hardware. The characteristics do not change with the change in signal-modulation parameters (such as modulation mode, modulation rate, and carrier frequency) transmitted by the transmitter. The intelligent identification of a specific transmitter can then be realized.

5. Conclusions

This paper proposed a method for individual radiation-source identification based on modulator distortion and power-amplifier nonlinearity. This method can complete the extraction of transmitter hardware characteristics. Signals with different modulation rates and carrier frequencies from the same transmitter have robust recognition capabilities. It was experimentally verified that, when the signal-to-noise ratio was lower than 20 dB, the proposed feature method was better than the bispectral feature-extraction method. When the signal-to-noise ratio was 0 dB, the recognition ability of four different transmitters using the proposed feature-extraction method exceeded 85%. When the carrier frequency was changed, the signal characteristics extracted by the bispectral method could not distinguish individual transmitters. The proposed method could still achieve 70% accuracy in the identification of four different PAM transmitters when the signal-to-noise ratio was high enough. The identification of four different QPSK transmitters could achieve a correct rate of more than 80% when the signal-to-noise ratio was sufficiently high.

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

Funding: This research was funded by National Natural Science Foundation of China, grant number 62071479.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Jia, Y. The Research on Communication Emitters Identification Technology. Ph.D. Thesis, University of Electronic Science and Technology of China, Cheng Du, China, 2017.
2. Lin, Y.; Jia, J.; Wang, S.; Ge, B.; Mao, S. Wireless Device Identification Based on Radio Frequency Fingerprint Features. In Proceedings of the ICC 2020—2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020; pp. 1–6.
3. Reising, D.R.; Temple, M.A.; Mendenhall, M.J. Improved wireless security for GMSK-based devices using RF fingerprinting. *Int. J. Electron. Secur. Digit. Forensics* **2010**, *3*, 41–59. [[CrossRef](#)]

4. Williams, M.D.; Munns, S.A.; Temple, M.A.; Mendenhall, M.J. RF-DNA Fingerprinting for Airport WiMax Communications Security. In Proceedings of the International Conference on Network & System Security IEEE, Melbourne, Victoria, Australia, 1–3 September 2010.
5. Dubendorfer, C.; Ramsey, B.; Temple, M. ZigBee Device Verification for Securing Industrial Control and Building Automation Systems. In Proceedings of the International Conference on Critical Infrastructure Protection, Washington, DC, USA, 18–20 March 2013; Springer: Berlin/Heidelberg, Germany, 2013.
6. Dubendorfer, C.K.; Ramsey, B.W.; Temple, M.A. An RF-DNA verification process for ZigBee networks. In Proceedings of the Military Communications Conference IEEE, Orlando, FL, USA, 29 October–1 November 2012.
7. Deng, S.; Huang, Z.; Wang, X. A novel specific emitter identification method based on radio frequency fingerprints. In Proceedings of the 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCI), Beijing, China, 8–11 September 2017; pp. 368–371. [\[CrossRef\]](#)
8. Carroll, T.L. Phase space method for identification of driven nonlinear systems. *Chaos Interdiscip. J. Nonlinear Sci.* **2009**, *19*, 348. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Chandran, V.; Carswell, B.; Boashash, B.; Elgar, S. Pattern recognition using invariants defined from higher order spectra. *IEEE Trans. Image Process.* **1997**, *6*, 703–712. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Tugnait, J.K. Detection of non-Gaussian signals using integrated polyspectrum. *IEEE Trans. Signal Process.* **1994**, *42*, 3137–3149. [\[CrossRef\]](#)
11. Zhang, X.-D.; Shi, Y.; Bao, Z. A new feature vector using selected bispectra for signal classification with application in radar target recognition. *IEEE Trans. Signal Process.* **2001**, *49*, 1875–1885. [\[CrossRef\]](#)
12. Qian, Y.; Qi, J.; Kuai, X.; Han, G.; Sun, H.; Hong, S. Specific Emitter Identification Based on Multi-Level Sparse Representation in Automatic Identification System. *IEEE Trans. Inf. Forensics Secur.* **2021**, *16*, 2872–2884. [\[CrossRef\]](#)
13. He, B.; Wang, F. Cooperative Specific Emitter Identification via Multiple Distorted Receivers. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 3791–3806. [\[CrossRef\]](#)
14. Han, J.; Zhang, T. Novel Method for SEI Based on 3D-Hilbert Energy Spectrum and Multi-scale Segmentation Features. In Proceedings of the 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 25–26 May 2017; pp. 351–355. [\[CrossRef\]](#)
15. Wambacq, P.; Sansen, W. Distortion Analysis of Analog Integrated Circuits. *Int. Ser. Eng. Comput. Sci.* **1998**, *3*, 335–345. [\[CrossRef\]](#)
16. Shi, M.; Huang, Y.; Wang, G. Carrier Leakage Estimation Method for Cross-Receiver Specific Emitter Identification. *IEEE Access* **2021**, *9*, 26301–26312. [\[CrossRef\]](#)
17. Ozbek, O. Improving DAC Integral Nonlinearity (INL) through Gain Correction. Cypress Semiconductor Corp. 2011; pp. 1–6. Available online: <https://www.cypress.com/file/109966/download> (accessed on 15 June 2021).
18. Polak, A.C.; Dolatshahi, S.; Goeckel, D.L. Identifying Wireless Users via Transmitter Imperfections. *IEEE J. Sel. Areas Commun.* **2011**, *29*, 1469–1479. [\[CrossRef\]](#)
19. Zhang, J.; Wang, F.; Dobre, O.; Zhong, Z. Specific Emitter Identification via Hilbert–Huang Transform in Single-Hop and Relaying Scenarios. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 1192–1205. [\[CrossRef\]](#)
20. Liu, M.-W.; Doherty, J.F. Nonlinearity estimation for Specific Emitter Identification in multipath environment. In Proceedings of the 2009 IEEE Sarnoff Symposium, Princeton, NJ, USA, 30 March–1 April 2009; pp. 1–5. [\[CrossRef\]](#)
21. Shi, F.; Chen, Z.; Cheng, X. Behavior Modeling and Individual Recognition of Sonar Transmitter for Secure Communication in UASNs. *IEEE Access* **2019**, *8*, 2447–2454. [\[CrossRef\]](#)
22. Gilabert, P.; Montoro, G.; Bertran, E. On the Wiener and Hammerstein Models for Power Amplifier Predistortion. In Proceedings of the 2005 Asia-Pacific Microwave Conference, Suzhou, China, 4–7 December 2005.
23. Liu, M.W.; Doherty, J.F. Nonlinearity Estimation for Specific Emitter Identification in Multipath Channels. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 1076–1085.
24. Jian, C.; Duan, T. Nonlinear feature extraction of power amplifiers in multipath. In Proceedings of the 2015 4th International Conference on Computer Science and Network Technology (ICCSNT) IEEE, Harbin, China, 19–20 December 2015.
25. Huang, G.; Yuan, Y.; Wang, X.; Huang, Z. Specific Emitter Identification Based on Nonlinear Dynamical Characteristics. *Can. J. Electr. Comput. Eng.* **2016**, *39*, 34–41. [\[CrossRef\]](#)
26. Yao, Y.Y.; Yu, L.; Chen, Y. Specific Emitter Identification Based on Square Integral Bispectrum Features. In Proceedings of the IEEE 20th International Conference on Communication Technology, Nanning, China, 28–31 October 2020.
27. Ji, H.; Wang, T.; Xiong, W. A method for specific emitter identification based on surrounding-line bispectrum and convolutional neural network. In Proceedings of the 2020 IEEE 3rd International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), Shenyang, China, 20–22 November 2020.