

Article

Automatic Modulation Classification of Digital Communication Signals Using SVM Based on Hybrid Features, Cyclostationary, and Information Entropy

Yangjie Wei * , Shiliang Fang * and Xiaoyan Wang

Key Laboratory of Underwater Acoustic Signal Processing of Ministry of Education, Southeast University, Nanjing 210096, China

* Correspondence: 230169359@seu.edu.cn (Y.W.); slfang@seu.edu.cn (S.F.)

Received: 13 July 2019; Accepted: 26 July 2019; Published: 30 July 2019



Abstract: Since digital communication signals are widely used in radio and underwater acoustic systems, the modulation classification of these signals has become increasingly significant in various military and civilian applications. However, due to the adverse channel transmission characteristics and low signal to noise ratio (SNR), the modulation classification of communication signals is extremely challenging. In this paper, a novel method for automatic modulation classification of digital communication signals using a support vector machine (SVM) based on hybrid features, cyclostationary, and information entropy is proposed. In this proposed method, by combining the theory of the cyclostationary and entropy, based on the existing signal features, we propose three other new features to assist the classification of digital communication signals, which are the maximum value of the normalized cyclic spectrum when the cyclic frequency is not zero, the Shannon entropy of the cyclic spectrum, and Renyi entropy of the cyclic spectrum respectively. Because these new features do not require any prior information and have a strong anti-noise ability, they are very suitable for the identification of communication signals. Finally, a one against one SVM is designed as a classifier. Simulation results show that the proposed method outperforms the existing methods in terms of classification performance and noise tolerance.

Keywords: modulation classification; digital communication signals; cyclostationary; information entropy; SVM

1. Introduction

Automatic modulation classification (AMC) of digital communication signals has now become an established research area [1]. It plays an important role in many applications. Some of these applications are for civilian purposes such as signal confirmation and spectrum management. The others are for military purposes such as surveillance, electronic warfare, and threat analysis. Therefore, if the types of the enemy signals are recognized, it will be of great significance for us to analyze and interfere with the enemy information.

Many methods for the modulation recognition of the communication signals have been published in recent years. In general, these methods can be divided into two categories: one is based on the decision-theoretic framework and the other is based on the statistical pattern recognition. The decision-theoretic approach is made by maximizing the probability of a certain modulation being sent given the received signal. The maximum likelihood algorithm is the most popular algorithm used in this approach. In the pattern recognition approach, the decision is made based on a set of features extracted from the intercepted signal, which is widely used in practical engineering. Extracting features from the intercepted signal is often followed by a pattern recognizer that determines the signal modulation. The following is an overview of some of these modulation recognition algorithms.

In [2], Kim et al. develop a modulation recognition method of MPSK signals based on the decision-theoretic approach. This method is less robust and lots of prior information such as carrier frequency, initial phase, and symbol rate are all assumed to be available to the classifier. Nandi et al. [3] follow the statistical pattern recognition approach and use the instantaneous features to discriminate the communication signals. These communication signals include 2ASK, 4ASK, 2PSK, QPSK, 2FSK, 4FSK. The classifier is a tree classifier and through simulations, they demonstrate that this recognizer performed well when the SNR is greater than 15 dB. The success recognition rate is $>88\%$ for these signals at the SNR of 15 dB. This approach is easy to implement and does not need any prior information, thus making it widely used. However, the features extracted by this method is sensitive to the noise and interference, and the method also needs to set a fixed decision threshold, which is often selected empirically. The shortcomings mentioned above have seriously affected the recognition performance of this method. In [4], three layered deep neural networks have been employed for the classification of BPSK, QPSK, 8PSK, 16QAM, and 64QAM with 21 statistical features. The method can achieve $>90\%$ classification accuracy when the SNR is greater than 10 dB. In [5], Afan Ali et al. develop a method for automatic modulation classification using the deep learning architecture in a combination of the In-phase and Quadrature constellation points of the received signal as the training example. The recognition rate of BPSK, 4QAM, 16QAM, 64QAM by this method is $>90\%$ when the SNR is greater than 5 dB. In [6], Weihua Jiang et al provide a modulation recognition method of non-cooperation underwater acoustic communication signals using principal component analysis and an artificial neural network. The recognition rate of the BPSK, QPSK, MFSK by this method is $>91\%$ when the SNR is greater than 5 dB. Although this method has good performance, it cannot distinguish between 2FSK and 4FSK. In [7], the modulation of the communication signals is recognized by the wavelet transform. The percentage of correct identification for PSK signals is $>80\%$ when $SNR > 6$ dB, and the percentage of correct identification for FSK signals is $>80\%$ when $SNR > 12$ dB. Although this method is a good tool to identify PSK and FSK, how to choose the appropriate wavelet function is a difficult problem to solve. Although many methods of modulation recognition have been proposed in the past [1–15], as far as we know, the methods to identify BPSK, QPSK, 2FSK, 4FSK and MSK signals have been rarely proposed so far. Therefore, this paper will focus on the identification of these signals.

In recent years, information entropy has been widely used in signal recognition [16,17]. Entropy is used to measure the uncertainty of signal distribution and represents the complexity degree of the signal; therefore, information entropy provides a theoretical basis for signal characterization description [18]. In [17], the Renyi Entropy of the Wigner-Ville distribution (WVD) and the continuous wavelet transform (CWT), and the singular spectrum entropy are extracted to identify the 2FSK, BPSK, 16QAM, 32QAM, and MSK signals. The average correct recognition rate of all signals is $>90\%$ when $SNR > 5$ dB. Although the method performs well at low SNRs, it is sensitive to the parameters of the WVD and CWT. Because of underlying periodicities due to various periodic signal processing operations such as sampling, scanning, modulating, multiplexing, and coding, or due to periodicity in the physical phenomenon that gives rise to the time series, many signals can be modeled as cyclostationary signals such as communication signals, radar signals, sonar signals and so on [19–21]. The cyclostationary of a signal is usually reflected by the spectral correlation function, also known as the cyclic spectrum. According to [20,21], we can see that the power spectra of different communication signals may be the same, but the cyclic spectra are sometimes significantly different. Moreover, since the noise does not have the characteristics of the cyclostationary, the cyclic spectrum has good anti-noise performance. Based on the advantages mentioned above, the cyclic spectrum is very suitable for identifying communication signals.

In this paper, a novel method for automatic modulation classification of digital communication signals, using SVM based on hybrid features, cyclostationary, and information entropy, is proposed. In this proposed method, by combining the theory of the cyclostationary and entropy, based on the existing signal features, we propose three other new features to assist the classification of communication signals, which are the maximum value of the normalized cyclic spectrum when

the cyclic frequency is not zero, the Shannon entropy of the cyclic spectrum, and Renyi entropy of the cyclic spectrum respectively. Since these new features do not require any prior information and have a strong anti-noise ability, they are very suitable for the identification of communication signals. Finally, a one against one SVM is used as a classifier. Simulation results show that the proposed method in this paper performs well in the low SNR condition.

The rest of the paper is arranged as follows: In Section 2, the mathematical model of the communication signals to be identified is given. In Section 3, the proposed features for signal classification in this paper are described in detail. The proposed SVM classifier is given in Section 4. The simulation results are displayed in Section 5. Finally, conclusions are addressed in Section 6.

2. Signal Model

In this work, the mathematical model of the signals to be recognized is expressed as

$$y(n) = x(n) + v(n), \quad (1)$$

where $n = 0, 1, \dots, N-1$, N represents the signal length. $y(n)$, $x(n)$ and $v(n)$ are respective the transmitted modulation signal, the intercepted signal, and the noise sample at discrete time n . The transmitted signal $\{x(n), n = 0, 1, \dots, N-1\}$ is drawn from an unknown constellation set Ψ which in turn belongs to a set of possible modulation formats $\{\Psi_1, \Psi_2, \dots, \Psi_K\}$. The modulation classification problem refers to the determination of the constellation set Ψ to which the transmitted signal belong based on the intercepted signal $\{y(n), n = 0, 1, \dots, N-1\}$. In this paper, we have considered the following digital communication signals for classification: BPSK, EPSK, 2FSK, 4FSK and MSK.

3. Feature Extraction

3.1. Instantaneous Features

In [3], instantaneous features, which contain hidden modulation information in a single domain, were demonstrated to be suitable for signal classification. According to the considered digital communication signals in this paper, the following instantaneous features are selected.

- (1) σ_{ap} : the standard deviation of the absolute value of the nonlinear component of the instantaneous phase in the non-weak segments of the intercepted signal:

$$\sigma_{ap} = \sqrt{\frac{1}{C} \left(\sum_{A_n(n) > a_t} \phi_{NL}^2(n) \right) - \left(\frac{1}{C} \sum_{A_n(n) > a_t} |\phi_{NL}(n)| \right)^2}, \quad (2)$$

where $A(n)$ denotes the instantaneous amplitude and $\phi(n)$ denotes the instantaneous phase of the intercepted signal both at time instants $t = n/f_s$. $A_n(n) = A(n)/m_a$, m_a is the average value of the instantaneous amplitude over one frame, that is

$$m_a = \frac{1}{N} \sum_{n=1}^N A(n), \quad (3)$$

$\phi_{NL}(n)$ is the value of the non-linear component of the instantaneous phase at time instants $t = n/f_s$, a_t is a threshold for $A(n)$ below which the estimation of the instantaneous phase is very sensitive to the noise, and C is the number of samples in $\{\phi_{NL}(n)\}$ for which $A_n(n) > a_t$. σ_{ap} is mainly used to distinguish the MPSK signals and it also can differentiate the modulation schemes of MFSK signals to some extent.

- (2) σ_{af} : standard deviation of the absolute value of the normalized-centered instantaneous frequency over non-weak segments of the intercepted signal:

$$\sigma_{af} = \sqrt{\frac{1}{C} \left(\sum_{A_n(n) > a_t} f_N^2(n) \right) - \left(\frac{1}{C} \sum_{A_n(n) > a_t} |f_N(n)| \right)^2}, \tag{4}$$

where, $f_N(n) = f_c(n) / r_s$, $f_c(n) = f(n) - m_f$, $m_f = (1/N) \sum_{i=1}^N f(i)$. σ_{af} can differentiate between the modulation types without frequency information and the FSK modulation types and also between 2FSK and 4FSK.

Figures 1 and 2 show the relationship between the features σ_{ap} and σ_{af} of different modulation signals with the SNRs. In this simulation, the sampling frequency $f_s = 10$ KHz, the signal length $N = 4096$, and the noise $v(n)$ is the white Gaussian noise. According to Figures 1 and 2, it is clear that although feature σ_{ap} and feature σ_{af} of different signals are different, and the difference is not obvious when the SNR is low. Therefore, we need to extract other new features to assist the identification of the modulation types of these signals in low SNR environments, and these features are introduced in the next section.

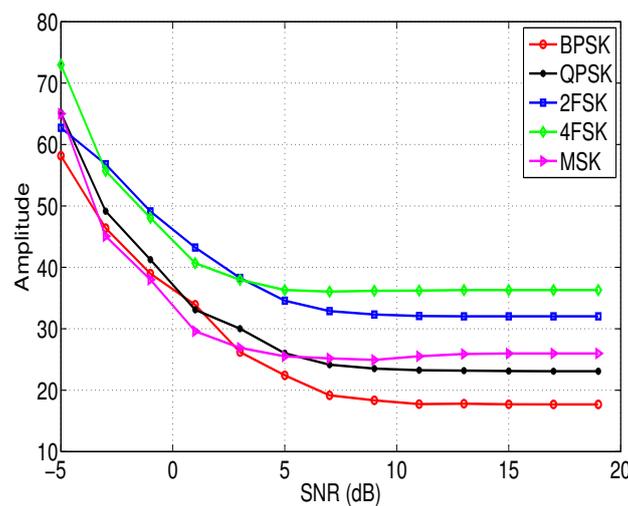


Figure 1. The relationship between the feature σ_{ap} of different modulation signals and the SNRs.

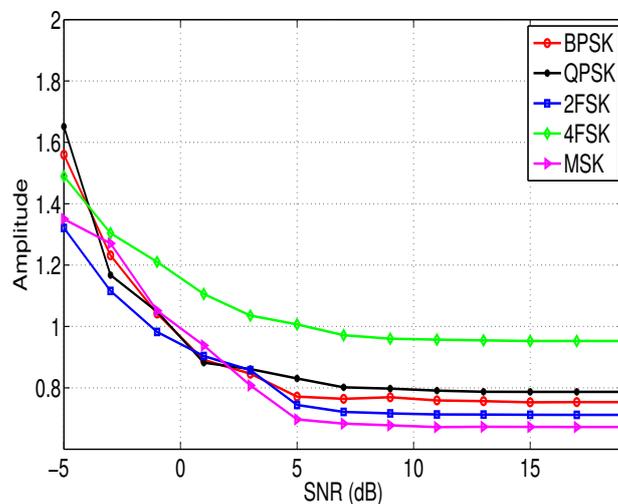


Figure 2. The relationship between the feature σ_{af} of different modulation signals and the SNRs.

3.2. Features Based on Cyclostationary and Information Entropy

3.2.1. The Cyclostationary of Communication Signals

A discrete time process $x(n)$ is cyclostationary, if the discrete time cyclic autocorrelation function

$$R_x^\alpha(k) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N [x(n+k)e^{-i\pi\alpha(n+k)}] [x(n)e^{i\pi\alpha n}]^*, \quad (5)$$

exists and is not identically zero when $\alpha \neq 0$. A particularly convenient characterization of cyclostationary is the cyclic spectrum,

$$S_x^\alpha(f) = \sum_{k=-\infty}^{\infty} R_x^\alpha(k) e^{-i2\pi fk}, \quad (6)$$

which is the Fourier series transform of the $R_x^\alpha(k)$. In (5) and (6), $i = \sqrt{-1}$, $*$ denotes the conjugation operation, α is called the cyclic frequency and f is the spectral frequency. Moreover, it is obvious that the cyclic autocorrelation function and the cyclic spectrum will reduce to the conventional autocorrelation function and the power spectral density when the cyclic frequency $\alpha = 0$.

The cyclic spectrum can be calculated directly through the double limit of the time smoothed cyclic periodogram,

$$S_x^\alpha(f) = \lim_{T \rightarrow \infty} \lim_{\Delta t \rightarrow \infty} S_{x_T}^\alpha(n, f)_{\Delta t}, \quad (7)$$

where

$$S_{x_T}^\alpha(n, f)_{\Delta t} = \langle X_T(n, f + \alpha/2) X_T^*(n, f - \alpha/2) \rangle_{\Delta t}, \quad (8)$$

$X_T(n, f)$ is called the complex demodulate of $x(n)$ and is the complex envelop of a narrow-band band-pass component of $x(n)$ centered at f with bandwidth $\Delta f \approx 1/T$,

$$X_T(n, f) = \sum_{m=n-T/2}^{n+T/2} x(m) e^{-i2\pi fm}, \quad (9)$$

where Δf and Δt are called the frequency and time resolutions of the estimation. The time smoothed cyclic periodogram is used to estimate the cyclic spectrum point by point.

When $x(n)$ is the MPSK signal, the mathematical model of $x(n)$ can be described as

$$x(n) = A \cos(2\pi f_c n / f_s + \phi(n)), \quad (10)$$

and

$$\phi(n) = \sum_{m=-\infty}^{\infty} \theta_m q(n - mT_0), \quad (11)$$

where f_c denotes the carrier frequency, A denotes the signal amplitude, $\theta_m \in \{2k\pi/M, k = 0, 1, \dots, M-1\}$ is phase of the transmitted MPSK symbol, T_0 denotes the symbol period. Here, $q(n)$ is a rectangle pulse,

$$q(n) = \begin{cases} 1, & |n| \leq T_0 f_s / 2 \\ 0, & |n| > T_0 f_s / 2 \end{cases}, \quad (12)$$

and therefore

$$Q(f) = \frac{\sin(\pi f T_0)}{\pi f}. \quad (13)$$

When $M = 2$, $x(n)$ is a BPSK signal, then based on (6) the cyclic spectrum of the BPSK signal can be obtained

$$S_x^\alpha(f) = \begin{cases} \frac{1}{4T_0} [Q(f - f_c + \frac{\alpha}{2}) Q^*(f - f_c - \frac{\alpha}{2}) \\ + Q(f + f_c + \frac{\alpha}{2}) Q^*(f + f_c - \frac{\alpha}{2})] \\ , \alpha = p/T_0 \\ \frac{1}{4T_0} Q(f - f_c + \frac{\alpha}{2}) Q^*(f + f_c - \frac{\alpha}{2}) \\ , \alpha = 2f_c + p/T_0 \\ \frac{1}{4T_0} Q(f + f_c + \frac{\alpha}{2}) Q^*(f - f_c - \frac{\alpha}{2}) \\ , \alpha = -2f_c + p/T_0 \\ 0, else \end{cases} , \quad (14)$$

For all integers p . Similarly, when $M \geq 4$, the cyclic spectrum of the MPSK signal can be written as

$$S_x^\alpha(f) = \begin{cases} \frac{1}{4T_0} [Q(f - f_c + \frac{\alpha}{2}) Q^*(f - f_c - \frac{\alpha}{2}) \\ + Q(f + f_c + \frac{\alpha}{2}) Q^*(f + f_c - \frac{\alpha}{2})] \\ , \alpha = p/T_0 \\ 0, else \end{cases} . \quad (15)$$

According to (14) and (15), we can see that the cyclic spectra of the BPSK signals have large values, when $\alpha = p/T_0$ and $\alpha = \pm 2f_c + p/T_0$. However, the cyclic spectra of the MPSK ($M \geq 4$) signals only have nonzero values at $\alpha = p/T_0$.

Similarly, the type of FSK signals can be expressed as

$$x(n) = \cos \left[2\pi f_c n / f_s + \sum_{r=-\infty}^{\infty} \sum_{m=1}^M \delta_m(r) [2\pi f_m (n/f_s - rT_0) + \theta_m(r)] q(n - rT_0) \right], \quad (16)$$

where the M frequencies $\{f_c + f_m, m = 1, 2, \dots, M\}$ are keyed randomly. When the phase sequences θ_m are constant

$$\theta_r(n) = \phi_m, \quad (17)$$

then the FSK is called clock phase coherent FSK and the MSK signal belong to this type of signal. Simultaneously, (16) can be written as

$$x(n) = \sum_{r=1}^{\infty} \sum_{m=1}^M \delta_m(r) q_m(n - rT_0), \quad (18)$$

where

$$q_m(n) = \cos [2\pi (f_c + f_m) n / f_s + \phi_m], \quad (19)$$

and therefore

$$Q_m(f) = \frac{\sin [\pi (f - f_c - f_m) T_0] e^{j\phi_m}}{2\pi (f - f_c - f_m)} + \frac{\sin [\pi (f + f_c + f_m)] e^{-j\phi_m}}{2\pi (f + f_c + f_m)}, \quad (20)$$

then based on (6) the cyclic spectrum of (18) can be given by

$$S_x^\alpha(f) = \frac{1}{MT_0} \sum_{m=1}^M Q_m(f + \alpha/2) Q_m^*(f - \alpha/2) - \frac{1}{M^2 T_0} \left[\sum_{m=1}^M Q_m(f + \alpha/2) \right] \left[\sum_{n=1}^M Q_n(f - \alpha/2) \right]^* \cdot \left[1 - \frac{1}{T_0} \sum_{n=-\infty}^{\infty} \delta(f + \alpha/2 - n/T_0) \right], \alpha = p/T_0 \tag{21}$$

Letting $f'_m = f_c + f_m$, it can be shown that $|S_x^\alpha(f)|$ has its maximum values at $f = \pm f'_m$, and if $f'_m T_0$ are integers, there are additional peaks at $\alpha = \pm f'_m$ and $f = 0$. There are also secondary maxima, down by the factor $M - 1$ from the primary maximum, at $\pm \alpha = f'_m \pm f'_n$ and $\pm f = (f'_m \mp f'_n)/2$.

When the phase sequence $\{\theta_m(r), m = 1, 2, \dots, M\} = \theta_n$ is independent and identically distributed with uniform fraction of time distribution over $(-\pi, \pi]$, $x(n)$ in (16) is called phase incoherent FSK and the 2FSK/4FSK signal belong to this type of signal, and (16) can be re-written as

$$x(n) = \sum_{r=-\infty}^{\infty} \sum_{m=1}^M a_r(n) q(n - rT_0), \tag{22}$$

and

$$a_r(n) = \cos [2\pi (f_c + f(r)) n/f_s + \theta_r], \tag{23}$$

where

$$f(r) = \sum_{m=1}^M \delta_m(r) f_m, \tag{24}$$

For a purely stationary $f(r)$ with discrete M -ary fraction of time distribution $\{P_m\}_1^M$, the cyclic spectrum of (22) can be expressed as

$$S_x^\alpha(f) = \frac{1}{4T_0} \sum_{m=1}^M P_m \left[Q\left(f + f'_m + \frac{\alpha}{2}\right) Q^*\left(f + f'_m - \frac{\alpha}{2}\right) + Q\left(f - f'_m + \frac{\alpha}{2}\right) Q^*\left(f - f'_m - \frac{\alpha}{2}\right) \right], \alpha = p/T_0 \tag{25}$$

Comparing (21) with (25), we can see that there are no impulses in (18), and there are no peaks at $\alpha = \pm 2f'_m$ for phase incoherent FSK.

Figure 3 shows the cyclic spectra of different communication signals under different SNR environments. The noise is the additive white Gaussian noise (AWGN). From Figure 3, it is clear that the cyclic spectra ($\alpha > 0$) of these communication signals are not only distinct but also have strong noise immunity, which means the cyclic spectrum is a very good tool for identifying these signals. In this simulation, $f_s = 10$ KHz, $N = 4096$.

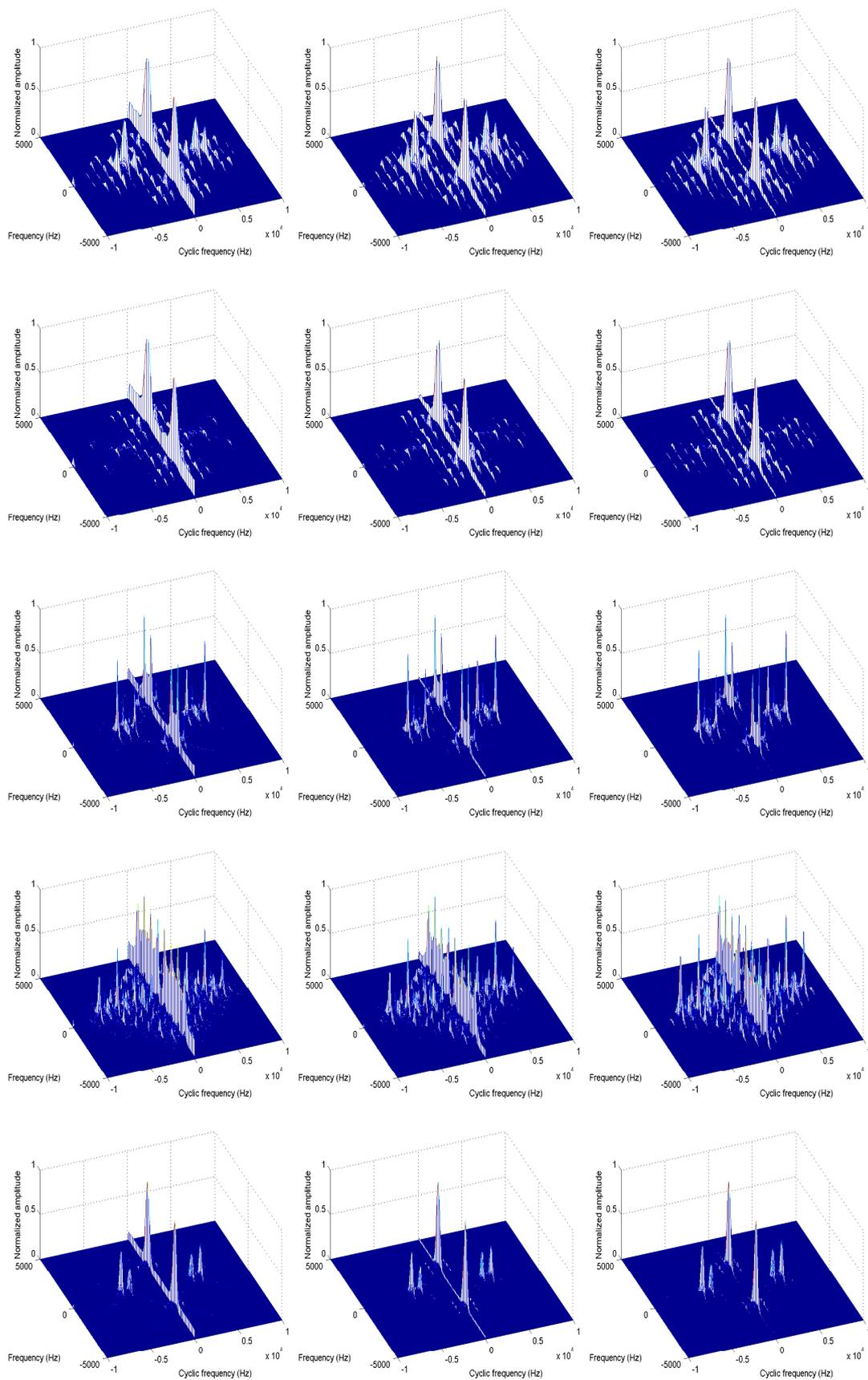


Figure 3. The cyclic spectra of five different modulation signals under different SNRs, the figures from top to bottom show the cyclic spectra of BPSK, QPSK, 2FSK, 4FSK, MSK. The figures in the left, middle and right columns correspond to 0 dB, 5 dB, and 10 dB respectively.

3.2.2. The Feature of the Cyclic Spectrum

From Figure 4 and the theory mentioned in Section 3.2.1, we can obtain that the maximum H_c of the normalized cyclic spectrum is a good characteristic to distinguish the communication signals, H_c is defined as:

$$H_c = \max \left\{ \left| S_x^{\alpha > 0}(f) \right| \right\}. \tag{26}$$

where $\max \{ \}$ denotes taking the maximum value, $| |$ represents taking the amplitude. The relationship between the feature H_c of different modulation signals and the SNR is shown in Figure 4. The signal propagation channel is AWGN. It can be seen from Figure 4 that the feature H_c of different modulation signals varies significantly, which means H_c is a good feature to distinguish them.

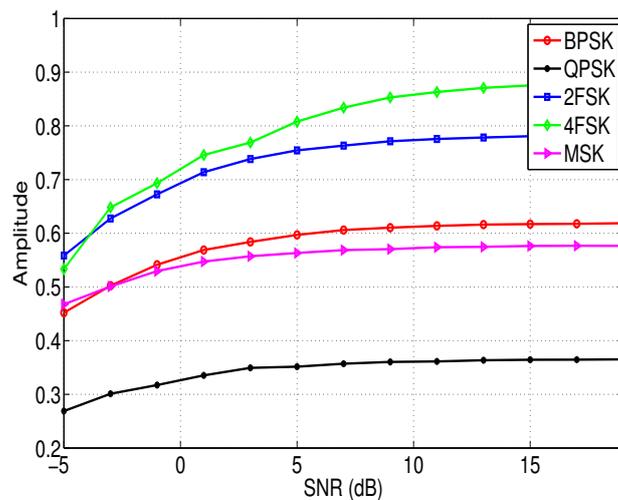


Figure 4. The relationship between the feature H_c of different modulation signals and the SNRs.

3.2.3. The Information Entropy Features of the Cyclic Spectrum

The information entropy was first proposed by Shannon, which was used to measure the uncertainty of signal distribution and represents the complexity degree of the signal. Therefore, information entropy provides a theoretical basis for the signal characterization description. Presently, entropy is applied many subjects [22–28]. Because of the symmetry of the cyclic spectrum, that is

$$|S_x^\alpha(f)| = |S_x^{-\alpha}(f)|. \tag{27}$$

Then according to [17], the Shannon entropy of the cyclic spectrum can be defined as

$$H_s = - \sum_f \sum_{\alpha > 0} P_x^\alpha(f) \log_2(P_x^\alpha(f)), \tag{28}$$

where

$$P_x^\alpha(f) = \frac{|S_x^\alpha(f)|}{\sum_f \sum_\alpha |S_x^\alpha(f)|}, \alpha > 0. \tag{29}$$

From (28), the entropy H_s has several important properties:

- (I) Symmetry: when the order of each component $P_x^\alpha(f)$ changed, the H_s will not be changed, which means the entropy is only related to the whole statistical properties of the data set. According to this property, we can obtain that the entropy H_s is robust to the signal modulation parameters such as carrier frequency, code rate, etc.

(II) Non-negative property: the entropy H_s is a non-negative value, that is

$$H_s \geq 0. \tag{30}$$

(III) Extreme property: when each component in data set existed in equal probability, the entropy H_s will get its maximum value. that is

$$H_s \leq \log_2 \Omega, \tag{31}$$

where Ω represents the number of the $P_x^\alpha(f)$.

Similarly, according to [17], the two-dimensional Renyi entropy of the cyclic spectrum can be defined as follows:

$$H_\beta = \frac{1}{1-\beta} \log_2 \sum_f \sum_{\alpha>0} (P_x^\alpha(f))^\beta, \tag{32}$$

where β is the order of the Renyi entropy of the cyclic spectrum, and $\beta \geq 0, \beta \neq 1$. Compared with Shannon entropy, the Renyi entropy can better reflect the difference between two different distributions [29].

The relationship between the information entropy H_s and H_β of different modulation signals and the SNR is shown in Figures 5 and 6. In these simulations, the sampling frequency $f_s = 10$ KHz, the signal length $N = 4096$, and without loss of generality, the order β is set to 5. The noise $v(n)$ is the white Gaussian noise. From Figure 5, it is clear that the entropy H_s is a good feature to distinguish 2FSK, 4FSK, and MSK in a low SNR environment. Similarly, the entropy H_β is a good feature to distinguish between BPSK and QPSK, 4FSK and 2FSK, 4FSK and MSK when the SNR is low.

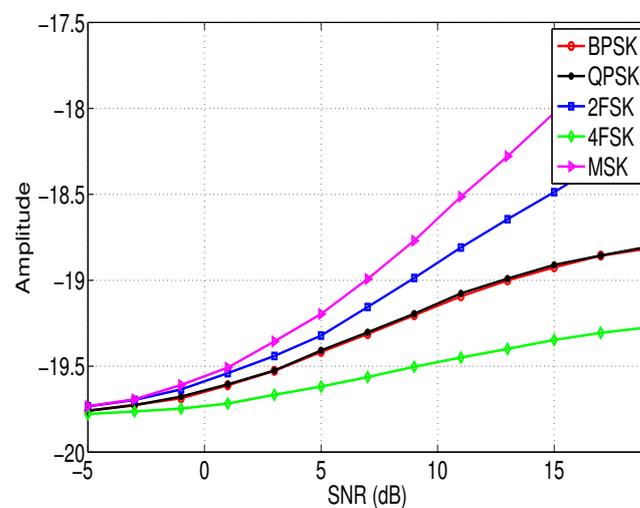


Figure 5. The relationship between the entropy feature H_s of different modulation signals and the SNRs.

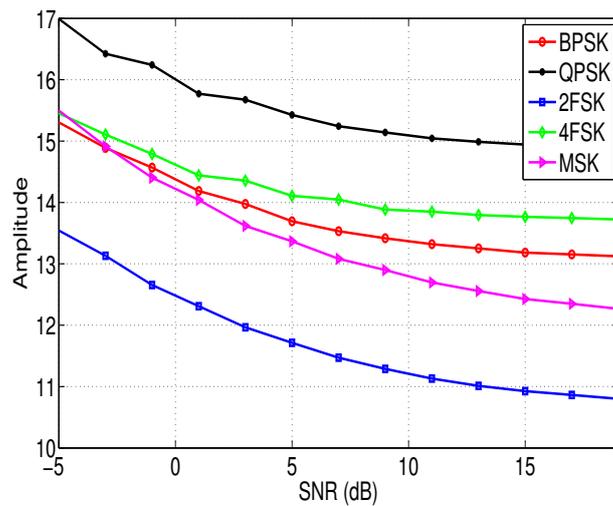


Figure 6. The relationship between the entropy feature H_{β} of different modulation signals and the SNRs.

4. The Proposed SVM Classifier

The traditional artificial neural networks (ANNs) often encounter problems such as overfitting and local minimization. Meanwhile, the large amount of sample data needed for full training of an ANN cannot be guaranteed in practical applications [30–32]. The SVM based on the structural risk minimization criterion cannot only minimize the classification error but also improve the generalization ability and has outstanding small sample learning ability. Therefore, based on the mentioned above, this paper will use the SVM to design the classifier to automatically identify the types of the modulation signals.

Given a training set of instance-label pairs $(x_i, y_i), i = 1, 2, \dots, l$ where $x_i \in R^n$ is the input vector and $y_i \in \{1, -1\}^l$ represents two classes label. Then the mathematical model for the two classes of SVM classifiers can be defined as follows:

$$\begin{aligned} \min_{w, b, \xi} & \frac{1}{2} \|w\|^2 + D \sum_{i=1}^l \xi_i \\ \text{s.t.} & y_i (w^T \Phi(x_i) + b) \geq 1 - \xi_i \end{aligned} \tag{33}$$

where $i = 1, 2, \dots, l, w$ is the vector of weight coefficient, $\xi_i \geq 0$ is the slack variable for the errors, $D > 0$ is the penalty parameter of the error term, a larger D corresponding to assigning a higher penalty to errors.

Each x_i is then mapped to a $\Phi(x_i)$ in the kernel-induced feature space, which is related to the kernel function

$$K(x_i, x_j) = \Phi(x_i)^T \Phi(x_j). \tag{34}$$

Then the standard SVM tries to find a hyperplane $w^T \Phi(x) + b$ that has a large margin and small training error. The kernel function has many types, such as linear function, polynomial function, radial basis function (RBF), and sigmoid function. The expressions of these functions are given as follows [33]:

(I) The linear kernel function:

$$K(x_i, x_j) = \gamma x_i^T x_j. \tag{35}$$

(II) The polynomial kernel function:

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d. \tag{36}$$

(III) The RBF kernel function:

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}. \quad (37)$$

(IV) The sigmoid kernel function:

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r). \quad (38)$$

where γ is the reciprocal of the number of signal types to be classified. Obviously, in this paper $\gamma = 1/5$. The effect of these kernel functions on the classification performance of SVM is discussed in detail in the next section.

SVM was originally only used for two types of classification problems, in order to achieve multi-classification problems a multiclass SVM comprising ten two-class sub-SVMs is designed. The number of sub-SVMs is $U(U - 1)/2$, where U is the number of the signal types. Figure 7 shows the classification procedure structure of the multiclass SVM proposed in this paper.

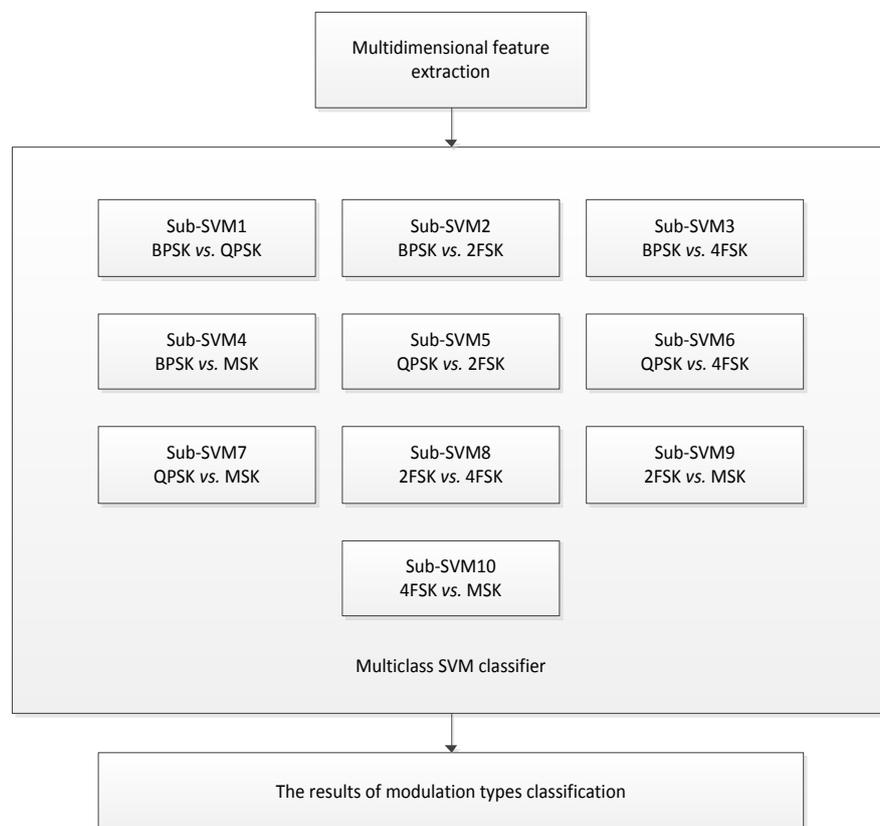


Figure 7. The structure of multiclass SVM classifier based on ten sub-SVMs using the one-versus-one algorithm.

5. Simulation Results

This section shows the simulation results of the proposed method for the classification of the considered digital modulation signals $\{BPSK, QPSK, 2FSK, 4FSK, MSK\}$, and the feature set adopted in these tests is $\{\sigma_{ap}, \sigma_{af}, H_c, H_s, H_\beta\}$. The sampling frequency $f_s = 10$ KHz, and the signal length $N = 4096$. The noise is the additive white Gaussian noise and was added according to SNRs $\{-5$ dB, -4 dB, \dots , 20 dB $\}$. Each modulation type has 2000 realizations and half of the realizations with SNRs of -5 dB, 0 dB, 5 dB, 15 dB, and 20 dB are used for training. Simulations results have been given

in figures and tables, and we use the accuracy metric to test the recognition performance. Furthermore, we have given some confusion matrixes for particular experiments that are considerable.

5.1. Classification in AWGN Channel

Tables 1–3 show the confusion matrixes over the AWGN channel of the proposed methods under different SNRs. The kernel function used here is the RBF function. From these tables, we can obtain that the overall accuracy of the proposed method for different modulation signals can reach 85.92%, when the $SNR = 0$ dB, and the overall accuracy will be greater than 99% when the $SNR \geq 6$ dB.

Table 1. The confusion matrix over AWGN channel, $SNR = 0$ dB.

Actual Modulation Type	Predicted Modulation Type				
	BPSK	QPSK	2FSK	4FSK	MSK
BPSK	912	0	0	4	128
QPSK	0	954	0	0	0
2FSK	2	0	856	130	76
4FSK	0	0	4	786	8
MSK	86	46	140	80	788
Overall Accuracy	85.92%				

Table 2. The confusion matrix over AWGN channel, $SNR = 6$ dB.

Actual Modulation Type	Predicted Modulation Type				
	BPSK	QPSK	2FSK	4FSK	MSK
BPSK	1000	0	0	0	15
QPSK	0	998	0	0	0
2FSK	0	2	1000	4	7
4FSK	0	0	0	996	0
MSK	0	0	0	0	978
Overall Accuracy	99.44%				

Table 3. The confusion matrix over AWGN channel, $SNR = 12$ dB.

Actual Modulation Type	Predicted Modulation Type				
	BPSK	QPSK	2FSK	4FSK	MSK
BPSK	1000	0	0	0	3
QPSK	0	1000	0	0	0
2FSK	0	0	1000	2	0
4FSK	0	0	0	998	0
MSK	0	0	0	0	997
Overall Accuracy	99.9%				

To evaluate the performance of different kernel functions for multiclass SVM. Table 4 shows the overall accuracy of the proposed method when using different kernel functions, and the $SNR = 6$ dB. According to Table 4, it is obvious that for the method proposed in this paper, the RBF function has the best performance, and the Sigmoid function has the worst performance. Therefore, the RBF function is recommended for the kernel function of the SVM classifier designed in this paper.

In practical applications, the complexity of the algorithm is an important consideration. Then, in order to measure the computational complexity of the proposed method in this paper, the recognition time of each sub-SVM is shown in Table 5. The simulation is implemented on a computer with a CPU of Intel Core 2.6 GHz i5-3230M and 4-Gb RAM, under the 64-bit Windows 7 system (Microsoft, Redmond, WA, USA). The multiclass SVM is accomplished via MATLAB2011b (MathWorks, Natick,

MA, USA). In practice, it is easy to find DSP with similar performance, such as TMS320C6678 and so on. Since the proposed SVM classifier in this paper uses a parallel structure, the time spent by the SVM classifier is equal to the maximum time spent by one of the sub-SVMs. From Table 5, we can obtain that the maximum time of the sub-SVMs is 35.708 μ s, which is acceptable in practical applications.

To show the superiority of the method proposed in this paper, the performance of the proposed method is investigated by making comparisons with the existing methods in [7,17]. Figure 8 shows the overall accuracy of different methods under different SNRs. The test uses 1000 Monte Carlo experiments. According to Figure 8, we can obtain that when the $SNR < 5$ dB, the recognition performance of the method proposed in this paper is better than that of the methods in [7,17], and when the $SNR \geq 5$ dB, the recognition performance of the method proposed in this paper is comparable to that of the method in [17].

Table 4. Comparison of different kernel functions for overall accuracy over AWGN channel, SNR = 6 dB.

Kernel Function	Overall Accuracy
Linear function	98.65%
Polynomial function	99.32%
RBF	99.44%
Sigmoid	89.6%

Table 5. Each sub-SVM recognition time in the real-time recognition system.

Sub-SVM	Time (μ s)
BPSK vs. QPSK	9.4673
BPSK vs. 2FSK	1.4247
BPSK vs. 4FSK	4.9047
BPSK vs. MSK	35.708
QPSK vs. 2FSK	3.05
QPSK vs. 4FSK	1.0533
QPSK vs. MSK	1.6783
2FSK vs. 4FSK	2.2073
2FSK vs. MSK	2.494
4FSK vs. MSK	1.0521
Maximum time	35.708

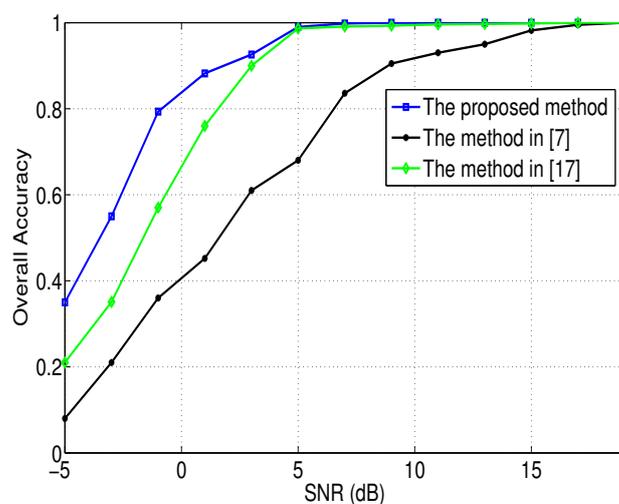


Figure 8. The overall accuracy of different methods under different SNRs.

5.2. Classification in Fading Channel

In practical environments, the propagation of signals is often affected by the channels. Table 6 shows the performance of the proposed method when the channel is the Rayleigh channel, and the SNR = 6 dB. It is assumed that there are two channels of multipath signals, the delay of multipath signals is 0.005 s and 0.01 s respectively, and the frequency deviation of multipath signals is 5 Hz and 10 Hz respectively. From Table 6 we can see that the overall accuracy is 99.22% on this occasion, which is comparable to that shown in Table 2. This is because since the multipath effect will affect the amplitude of the cyclic spectrum but not the shape of the cyclic spectrum, the multipath effect has little influence on the entropy characteristics proposed in this paper, so at this point, the performance of the method presented in this paper will not be seriously affected.

Table 6. The confusion matrix over Rayleigh channel, SNR = 6 dB.

Actual Modulation Type	Predicted Modulation Type				
	BPSK	QPSK	2FSK	4FSK	MSK
BPSK	975	6	0	0	4
QPSK	17	994	0	0	0
2FSK	0	0	999	2	1
4FSK	0	0	0	998	0
MSK	8	0	1	0	995
Overall Accuracy	99.22%				

6. Conclusions

Since digital communication signals are widely used in various military and civilian applications, it is important to improve the recognition rate of digital communication signals. In this paper, a novel signal classification method using SVM based on hybrid features, cyclostationary, and information entropy is proposed. The method combines the theory of the cyclostationary and entropy and uses a one against one SVM as a classifier. Simulation results show that the proposed method has a good recognition performance for the signals considered in this paper when in low SNR environments and fading channels.

Author Contributions: The work presented in this paper has been carried out in collaboration with all authors. Y.W. designed and implemented the improved method. S.F. and X.W. discussed the results with Y.W. and refined the overall manuscript.

Funding: This work was supported in part by National Natural Science Foundation of China under Grant no.11604048, and in part by the National Natural Science Foundation of China under Grant no.11704069, and in part by the Fundamental Research Funds for the Central Universities 2242016K30013, and in part by the Fundamental Research Funds for the Central Universities 2242019K30021, and in part by the National Natural Science Foundation of China under Grant no.11574048, and in part by the National Natural Science Foundation of China under Grant no.11674057.

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

References

1. Dobre, O.A.; Abdi, A.; Bar-Ness, Y.; Su, W. Survey of automatic modulation classification techniques: Classical approaches and new trends. *IET Commun.* **2007**, *1*, 137–156. [[CrossRef](#)]
2. Kim, K.; Polydoros, A. Digital modulation classification: The BPSK and QPSK case. In Proceedings of the MILCOM 88, 21st Century Military Communications—What’s Possible? Conference Record. Military Communications Conference, San Diego, CA, USA, 23–26 October 1998; Volume 10, pp. 431–436.
3. Nandi, A.K.; Azzouz, E.E. Algorithms for automatic modulation recognition of communication signals. *IEEE Trans. Commun.* **1998**, *334*, 431–436. [[CrossRef](#)]

4. Kim, B.; Kim, J.; Chae, H.; Yoon, D.; Choi, J.W. Deep neural network-based automatic modulation classification technique. In Proceedings of the 2016 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, South Korea, 19–21 October 2016; pp. 579–582.
5. Ali, A.; Yangyu, F.; Liu, S. Automatic modulation classification of digital modulation signals with stacked autoencoders. *Digit. Signal Process.* **2017**, *71*, 108–116. [[CrossRef](#)]
6. Jiang, W.H.; Tong, F.; Dong, Y.Z.; Zhang, G.Q. Modulation recognition of non-cooperation underwater acoustic communication signals using principal component analysis. *Appl. Acoust.* **2018**, *138*, 209–215. [[CrossRef](#)]
7. Ho, K.C.; Prokopiw, W.; Chan, Y.T. Modulation identification of digital signals by the wavelet transform. *IEE Proc. Radar Sonar Navig.* **2000**, *147*, 169–176. [[CrossRef](#)]
8. Polydoros, A.; Kim, K. On the detection and classification of quadrature digital modulations in broad-band noise. *IEEE Trans. Commun.* **1990**, *38*, 1199–1211. [[CrossRef](#)]
9. Jallon, P.; Chevreuil, A.; Loubaton, P. Separation of digital communication mixtures with the CMA: Case of unknown rates. *Digit. Signal Process.* **2010**, *90*, 2633–2647. [[CrossRef](#)]
10. Soliman, S.S.; Hsue, S.-Z. Signal classification using statistical moments. *IEEE Trans. Commun.* **1992**, *40*, 908–916. [[CrossRef](#)]
11. Oner, M. On the classification of binary space shift keying modulation. *IEEE Trans. Commun. Lett.* **2018**, *22*, 1584–1587. [[CrossRef](#)]
12. Swami, A.; Sadler, B.M. Hierarchical digital modulation classification using cumulants. *IEEE Trans. Commun.* **2000**, *48*, 416–429. [[CrossRef](#)]
13. Chan, Y.T.; Gadbois, L.G. Identification of the modulation type of a signal. *Signal Process.* **1989**, *16*, 149–154. [[CrossRef](#)]
14. Aslam, M.W.; Zhu, Z.; Nandi, A.K. Automatic modulation classification using combination of genetic programming and KNN. *IEEE Trans. Wirel. Commun.* **2012**, *11*, 2742–2750.
15. Abu-Romoh, M.; Aboutaleb, A.; Rezki, Z. Automatic modulation classification using moments and likelihood maximization. *IEEE Commun. Lett.* **2018**, *22*, 938–941. [[CrossRef](#)]
16. Li, J.; Ying, Y. Radar Signal Recognition Algorithm Based on Entropy Theory. In Proceedings of the 2014 2nd International Conference on Systems and Informatics (ICSAI 2014), Shanghai, China, 15–17 November 2014; Volume 147, pp. 718–723.
17. Wang, H.; Guo, L.; Dou, Z.; Lin, Y. A New Method of Cognitive Signal Recognition Based on Hybrid Information Entropy and D-S Evidence Theory. *Mob. Netw. Appl.* **2018**, *23*, 677–685. [[CrossRef](#)]
18. Liu, S.; Lu, M.; Liu, G.; Pan, Z. A Novel Distance Metric: Generalized Relative Entropy. *Entropy* **2017**, *19*, 269. [[CrossRef](#)]
19. Gardner, W.A. Measurement of Spectral Correlation. *IEEE Trans. Acoust. Speech Signal Process.* **1986**, *34*, 1111–1123. [[CrossRef](#)]
20. Gardner, W.A. Spectral Correlation of Modulated Signals: Part I-Analog Modulation. *IEEE Trans. Commun.* **1987**, *35*, 584–594. [[CrossRef](#)]
21. Gardner, W.A.; Brown, W.; Chen, C.-K. Spectral Correlation of Modulated Signals: Part 11-Digital Modulation. *IEEE Trans. Commun.* **1987**, *35*, 595–601. [[CrossRef](#)]
22. Pawar, S.U.; Doherty, J.F. Modulation Recognition in Continuous Phase Modulation Using Approximate Entropy. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 843–852. [[CrossRef](#)]
23. Li, J.; Guo, J. A New Feature Extraction Algorithm Based on Entropy Cloud Characteristics of Communication Signals. *Math. Probl. Eng.* **2015**, *2015*, 1–8. [[CrossRef](#)]
24. He, Z.-Y.; Cai, Y.-M.; Qian, Q.-Q. A Study of Wavelet Entropy Theory and its Application in Electric Power System Fault Detection. *Proc. CSEE* **2005**, *5*, 40–45.
25. Jiang, B.; Dong, X.; Shi, S. Application of approximate entropy to cross-country fault detection in distribution networks. *Power Syst. Prot. Control* **2015**, *7*, 15–21.
26. Hosseini, S.A.; Naghibi-Sistani, M.B. Emotion recognition method using entropy analysis of EEG signals. *Int. J. Image Graph. Signal Process.* **2011**, *5*, 30–36. [[CrossRef](#)]
27. Li, Y.; Chen, X.; Yu, J.; Yang, X.; Yang, H. The data-driven optimization method and its application in feature extraction of ship-radiated noise with sample entropy. *Energies* **2019**, *12*, 359. [[CrossRef](#)]

28. Li, Y.; Wang, L.; Li, X.; Yang, X. A novel linear spectrum frequency feature extraction technique for warship radio noise based on complete ensemble empirical mode decomposition with adaptive noise, duffing chaotic oscillator, and weighted-permutation entropy. *Entropy* **2019**, *21*, 507. [[CrossRef](#)]
29. Zyczkowski, K. Renyi extrapolation of Shannon entropy. *Physics* **2013**, *10*, 297–310. [[CrossRef](#)]
30. Burges, C.J.C. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *10*, 121–167. [[CrossRef](#)]
31. Shang, X.; Veldhuis, R.N.J. Grip-Pattern Recognition in Smart Gun Based on Likelihood-Ratio Classifier and Support Vector Machine. In Proceedings of the ICISP'08 Proceedings of the 3rd International Conference on Image and Signal Processing, Cherbourg-Octeville, France, 1–3 July 2008; pp. 289–295.
32. Burges, C.J.C. Support vector networks. *Mach. Learn.* **1995**, *20*, 144–152.
33. Wei, J.; Huang, Z.; Su, S.; Zuo, Z. Using Multidimensional ADTPE and SVM for Optical Modulation Real-Time Recognition. *Entropy* **2016**, *18*, 30. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).