Multi-Carrier Signal Recognition Method Based on Multi-Feature Input and Hybrid Training Neural Network

Shanshan Li 1, Yi Cui 1, Qi Zhang 1,*, Zhipei Li 2,*, Ran Gao 2, Feng Tian 1, Qinghua Tian 1, Bingchun Liu 3, Jinkun Jiang 1, Yongjun Wang 1 and Xiangjun Xin 1

Abstract: In order to achieve automatic identification of modulation formats in orthogonal frequency division multiplexing (OFDM) subcarrier signals, a recognition method based on multiple feature inputs and Hybrid Training Neural Network (HTNN) is proposed, in which an HTNN structure is designed to obtain high-order statistical correlation features and constellations of OFDM subcarriers. The recognition performance of the proposed method in free space channel transmission and atmospheric time-varying channel transmission is studied by simulation. Simulation results show that the overall identification accuracy of the recognition model based on the proposed method exceeded 93.37% in the wide Signal-to-Noise Ratio (SNR) range of the free space channel. With an SNR higher than 7.5 dB, identification accuracy performance of the learning model culminated, achieving 100% identification accuracy of OFDM subcarrier signals. Under weak turbulent atmospheric and time-varying channel conditions, the overall identification accuracy curve tended to increase as SNR increased and stabilized at more than 95%. Compared with the two comparison methods, the proposed method reduced the sensitivity to channel variations and improved the convergence speed on the basis of the guaranteed identification accuracy, and enabled reliable identification of OFDM subcarrier signals in a wide SNR range.

Keywords: modulation format identification; OFDM; machine learning

1. Introduction

Due to the increase in space engineering projects and tasks, the complexity of modulation formats in satellite laser communication systems is increasing in order to meet the needs of different users and services and make full use of channel capacity [1–7]. To expand transmission capacity, the Optical Frequency Division Multiplexing (OFDM) technique has attracted attention, and the OFDM system was found to be highly resistant to multipath interference and spectrally selective fading as well as having good BER performance [8–15]. To meet the requirements of high-speed data transmission in satellite laser communication systems, the OFDM technique can improve the transmission rate and reliability of the system. In order to ensure the security and accuracy of information transmission, the receiver needs to obtain the modulation format of the laser signal to demodulate the received signal, and it has become a key procedure in the study of automatic modulation format identification of OFDM subcarrier signals.

In 1969, Weaver et al. published the first paper on automatic modulation recognition, in which a pattern recognition technique was used to achieve fast automatic recognition.
of the modulation format of high-frequency radio signals, not only in cases where the modulation format was unknown, but also to recognize new signals that had never been encountered before [16]. In the 1990s, Nandi and Azzouz first published a monograph on an automatic modulation recognition technique exploring two algorithms that enable the recognition of analog and digital modulation formats: decision theory and an artificial neural network [17]. Compared with the decision theory classifier, the feature-based classifier achieves the performance close to the former at the cost of less complexity, and has more advantages in application. Based on neural network algorithms, instantaneous information features, transform domain features, cumulative features, and signal constellation features, researchers proposed a variety of signal modulation format recognition methods to improve recognition performance [18–27]. However, the aforementioned methods of OFDM signal identification mostly focus on the differentiation between OFDM signals and single-carrier signals, whereas there are few studies related to the recognition of different modulation formats of OFDM subcarrier signals in atmospheric channels.

In this paper, based on studying the signal characteristics of multi-carrier laser signals in time-varying channels, a multi-carrier signal recognition method based on multi-feature input and HTNN is proposed. The method exploits the features of high-order statistics and OFDM subcarrier constellations. The features of higher-order statistics can suppress the effect of Gaussian noise, which helps to improve the robustness of the classifier over a wide range of SNR [28], and the features of constellations have better discriminability under atmospheric weak turbulence channel conditions. As the double-input characteristics of the proposed multi-input hybrid training neural network, the HTNN is trained to obtain the high-order correlation features of multi-feature inputs and get the model to realize automatic modulation format recognition among OFDM subcarriers. Through the analysis of the performance under the two conditions of free-space channel transmission and atmospheric time-varying channel transmission, simulation results verified that the proposed method could reduce the sensitivity to channel changes on the basis of ensuring the identification accuracy, improve the convergence speed, and realize the reliability recognition of the OFDM subcarrier signal under a wide range of SNRs.

2. Principle

2.1. OFDM Signal Model

OFDM, as a special frequency division multiplexing method, transmits serial high-speed data by converting the data into blocks of data and then using Fourier transform techniques to encode the data on separate orthogonal subcarriers in the frequency domain. Transmitted OFDM signal in the time domain can be represented as:

$$s(t) = \sum_{k=1}^{n} (a_k \cos(kw_0 t) - b_k \sin(kw_0 t))$$

(1)

where $w_0 = 2\pi\Delta f$. $\Delta f$ is the guard interval, $n$ is the number of sub-carriers, and in frequency domain it is expressed as:

$$s(t) = \sum_{k=1}^{n} S_k e^{-jkw_0 t}$$

(2)

2.2. High-Order Statistics Feature

In statistics, for statistical analysis of random variables or stochastic processes with Gaussian distribution, first-order and second-order statistics such as mean, variance and the autocorrelation function are usually used, whereas for random signals that do not satisfy Gaussian distribution, the information carried in the signal cannot be described by first-order and second-order statistics, and the characteristics of the signal can be characterized only by third-order or fourth-order statistics, and this method is called the higher-order statistical method, in which the higher-order statistics include higher-order
moments, higher-order cumulants, higher-order cumulant spectra, and other forms. In addition, higher-order statistics can be used as a mathematical analysis tool to identify non-minimum systems or reconstruct non-minimum phase signals, extract signal nonlinear information, detect cyclic smoothness in characterized signals, etc.

A random signal $x$ without a Gaussian distribution, whose probability density function is $f(x)$, and its first characteristic function are expressed as:

$$
\phi(\omega) = E\{e^{j\omega x}\} = \int_{-\infty}^{\infty} f(x)e^{j\omega x}dx
$$

(3)

Since the probability density function is non-zero, the maximum value is obtained at the origin in Equation (3), and the $k$-derivative is:

$$
\phi^{(k)}(\omega) = \frac{d^{(k)}\phi(\omega)}{d\omega^k} = j^kE\{x^ke^{j\omega k}\}
$$

(4)

When $\omega = 0$, the $k$ moment of $x$ can be expressed as:

$$
m_k = E\{x^k\} = -j^k\frac{d^{(k)}\phi(\omega)}{d\omega^k}|_{\omega=0} = -j^k\phi^k
$$

(5)

The Taylor expansion of Equation (5) can be obtained as:

$$
c_{kx} = -j^k\frac{d^{(k)}\ln\phi(\omega)}{d\omega^k}|_{\omega=0} = -j^k\psi^k
$$

(6)

where $\psi(\omega) = \ln\phi(\omega)$. It is called the cumulant of order $k$ of the random signal $x$, also known as the cumulant generating function. The third-order cumulant of imaginary parts of OFDM signals with different modulation formats is shown in Figure 1.

![Figure 1. Comparison of third-order cumulant of imaginary parts of OFDM signals.](image)

2.3. Constellation Diagram Feature

The constellation diagram of the modulation signal reflects the amplitude and phase information of the signal and can in principle be used to classify the modulation format. Considering the four modulation formats of BPSK, QPSK, 8PSK, and 16QAM, Gaussian
white noise was added to simulate the noise in the channel and the SNR was set to 20 dB. Channel conditions with different turbulence intensifications were simulated by setting different light intensity variances. Finally, the constellations comparison of OFDM signals with different modulation formats transmitted through the channel is shown in Figure 2.

2.4. Multi-Feature Input and Hybrid Training Neural Network

To meet the requirements of high-speed data transmission in satellite laser communication systems, the OFDM technique can improve the transmission rate and reliability of the system. In order to correctly receive the information transmitted by OFDM subcarrier signals in atmospheric time-varying channels, it is necessary to design a signal identification method without prior knowledge to improve the reliability and convergence of identification. The essence of deep learning is to simulate the human brain neural networks using a large amount of training data to improve the accuracy of the classification or prediction. Considering that the higher-order statistical features of OFDM signals can suppress the effect of Gaussian noise, it is beneficial to improve the robustness of the classifier in a large SNR range. Figure 3 is the structure diagram of the HTNN. High-order statistics features and constellation features of OFDM subcarrier signals are taken as the dual-input features of the network, and the HTNN is trained to independently mine high-order correlation features and obtain a learning model to realize automatic modulation format recognition among OFDM subcarrier signals.
3. Simulation Setup and Performance Analysis

3.1. Simulation Setup

The simulation process is as follows.

Step 1: Determine the input parameters, including the number of OFDM carriers, the number of symbols, IFFT length, cyclic prefix length, cyclic suffix length, plus the ascending cosine window coefficient, modulation formats, the number of OFDM signals, and the division ratio of the training and test data; the fixed simulation parameters were set as shown in Table 1. Randomly generated [01] data were transmitted, and Gray codes were used to encode the transmitted data. The modulation formats were BPSK, QPSK, 8PSK, and 16QAM.

Table 1. Parameters in the simulation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of sub-carriers</td>
<td>200</td>
</tr>
<tr>
<td>Number of symbols</td>
<td>100</td>
</tr>
<tr>
<td>IFFT length</td>
<td>512</td>
</tr>
<tr>
<td>Cyclic prefix length</td>
<td>128</td>
</tr>
<tr>
<td>Cyclic suffix length</td>
<td>20</td>
</tr>
<tr>
<td>Ascending cosine window coefficient</td>
<td>1.5/32</td>
</tr>
<tr>
<td>Modulation formats</td>
<td>BPSK\QPSK\8PSK\16QAM</td>
</tr>
<tr>
<td>Division ratio of training and test data</td>
<td>3:1</td>
</tr>
</tbody>
</table>

Step 2: Change input parameters. The input parameters were varied in order to compare the two identifications under time-varying channel transmission conditions in free-space and atmospheric channel transmission effects. The light intensity variation range was set to [0, 0.1, 0.2, 0.5, 0.8] and the SNR range was [0, 2.5, 5, 7.5, 10, 12.5, 15, 17.5, 20], where 0 represents the free space channel.

Step 3: OFDM signal acquisition. Based on a log-normal light intensity distribution model to simulate turbulent channels, OFDM signals with different light intensity variances and different SNRs were obtained. In the case of a single optical intensity variant and signal-to-noise ratio, 10,000 signals were generated for each modulation format, for a total of 40,000 OFDM signals.

Step 4: Obtain the constellation feature set and the third-order cumulant feature set. Based on the OFDM signal set, the constellation mapping was carried out and the third-order cumulant value was calculated to respectively form the constellation feature set and the third-order cumulant feature set.
Step 5: The training set and test set were separated, and the training set and test set were randomly selected according to the ratio of 3:1, with sample sizes of 30,000 and 10,000, respectively.

Step 6: Hybrid training neural network training. The constellation feature set and the third-order cumulative feature set from the training set were fed into the hybrid training neural network for training to obtain the model.

Step 7: Performance measurement of the recognition model. The recognition effect of the recognition model constructed in step 6 was tested on the test set. The accuracy value of the classifier (i.e., recognition rate) was mainly considered as an indicator of recognition effectiveness.

Step 8: Check whether the light intensity difference and SNR were traversed. If yes, compare and analyze the recognition effect under different channel conditions and SNR, and the simulation ends. If not, return to step 3.

In order to analyze the recognition performance of the proposed method in free-space channels, the simulation of OFDM subcarrier signal modulation format recognition in free-space channel transmission conditions was completed by using MATLAB and Python co-simulation according to the above steps. The signal generation, constellation feature set acquisition and third-order cumulative feature set calculation parts were completed based on MATLAB, and the model learning part was completed based on Python.

3.2. Performance Analysis
3.2.1. Identification Performance under Different Neural Network Model Parameters

The hybrid training neural network model parameters are shown in Table 2. Cross entropy was adopted as the loss function, and the Adam optimizer was selected by the optimizer with 10 iterations. For the constellation feature set and the third-order cumulative feature set in the input training set, 64 feature samples were selected for training at a time. In order to compare the influence of network model parameters on the model recognition effect, the learning rate was set to change between [0.01, 0.001, 0.0001], and the change curve of the loss function and identification accuracy in the training of mixed training neural network under different learning rates were obtained, as shown in Figures 4 and 5.

Table 2. HTNN parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss function</td>
<td>Cross-entropy</td>
</tr>
<tr>
<td>Optimizer</td>
<td>Adam optimizer</td>
</tr>
<tr>
<td>Epochs</td>
<td>10</td>
</tr>
<tr>
<td>Batch</td>
<td>64</td>
</tr>
</tbody>
</table>

![Figure 4. Loss function curve of the HTNN under different learning rates vs. epochs.](image)
As can be seen from Figure 5, the loss function curves at different learning rates show that the loss function showed a decreasing trend as the amount of training increased. The loss function curves under different learning rates varied greatly, which indicates that the designed hybrid training neural network is sensitive to the change of learning rate. When the learning rate was $1 \times 10^{-2}$, the loss function of the network decreased the fastest, but there were several fluctuations in the subsequent training, indicating that the convergence of the model was not stable. When the learning rate was $1 \times 10^{-3}$, the loss function of the network decreased the fastest and was stable, and fluctuated slightly during the training process, which indicates that the model converged well at this time. When the learning rate was $1 \times 10^{-3}$, the loss function of the network decreased the slowest and fluctuated.

The identification accuracy curves varied greatly at different learning rates, which validated the sensitivity of the hybrid training neural network to changes in the learning rate. When the learning rate was $1 \times 10^{-2}$, the identification accuracy of the network increased the fastest, but fluctuated several times during the subsequent training, indicating that the convergence of the model was not stable at this time. When the learning rate was $1 \times 10^{-3}$, the identification accuracy of the network increased the fastest and was stable, and fluctuated less during the training process, which indicates that the convergence of the model was good at this time. When the learning rate was $1 \times 10^{-3}$, the identification accuracy of the network improved the slowest and fluctuated.

The above results show that the loss function and identification accuracy of the designed hybrid training neural network are more sensitive to the change of the model parameter learning rate. In order to obtain the best performance in the subsequent simulation, the learning rate was set as $1 \times 10^{-3}$ according to the above simulation results.

3.2.2. Recognition Performance using Different Methods

In order to verify the feasibility of the proposed method to realize OFDM signal recognition, we compared the proposed method with two other methods, i.e., third-order cumulant feature with a MLP-based recognition method, and constellation feature with a CNN-based recognition method. In a free-space channel, identification accuracy performance of all three methods under different SNR conditions was compared.

As shown in Figure 6, the identification accuracy curves of OFDM signal recognition using the three methods had the same trend of variation, increasing with the increase of SNR. The overall identification accuracy of the recognition model trained by a constellation feature and CNN exceeded 79.4% in all SNR ranges, whereas it was poor at low SNRs and converged to 100% when the SNR was greater than 7.5 dB. The recognition model trained by a third-order cumulant feature and MLP had poor recognition performance in all SNR ranges, with the overall identification accuracy fluctuating around 73%, which indicates that the characteristic parameters of high-order statistics of OFDM signals are not affected by the SNR. The overall identification accuracy of the proposed method was more than
93.37% in all SNR ranges, and converged to 100% when the SNR was greater than 7.5 dB. The above results show that the proposed method is insensitive to the variation of the SNR and achieves high accuracy identification of OFDM signals.

As can be seen from Figure 7, under the condition of free space channel transmission, as the SNR increased, the identification accuracy of the proposed method for OFDM signals under different modulation formats increased, and finally converged to 100%. On the whole, the identification accuracy of the learning model based on the proposed method for the four modulation formats was more than 98.63% in all SNR ranges, and the accuracy of 16QAM OFDM signals was more than 99.9%. The above results show that the proposed method has good recognition performance for OFDM signals with different modulation formats in free-space channels and is minimally affected by the variation of the SNR, outperforming the model proposed in [29] in terms of SNR tolerance.

3.2.3. Identification Performance in Turbulence Channel

This section shows the simulation results of the modulation format identification process of single-carrier signals under atmospheric time-varying channel transmission conditions. The variances of light intensity representing the atmospheric turbulence intensity was set as [0.1, 0.2, 0.5, 0.8]. With the increase of the non-uniformity of the channel, the fluctuation of the light field intensity gradually increased. The identification accuracy curves against the SNR of the learning model based on the proposed method at different turbulence intensities are shown in Figure 8.

Figure 6. Identification accuracy comparison of different methods against SNR.

Figure 7. Identification accuracy curves of four modulation formats in free-space channels.
The overall identification accuracy showed an increasing trend as the SNR increased under the weakly turbulent atmospheric time-varying channel conditions, and the overall identification accuracy of the learning model based on the proposed method reached more than 73.5% under all turbulence intensities and SNR ranges. The overall identification accuracy curve plateaued after a rapid increase. With the gradual increase of the light intensity variance, the effect of turbulence gradually increased and the overall performance of the classifier slightly decreased, and the stable identification accuracy value gradually decreased. However, numerically, the converged identification accuracy of the proposed method reached more than 95% under weak turbulence conditions. The results showed that the proposed method could achieve fast convergence of OFDM multi-carrier signal recognition under weak turbulence conditions of atmospheric time-varying channels with guaranteed identification accuracy.

An identification accuracy histogram of different modulation formats of OFDM signals under different turbulence intensities against the SNR is shown in Figure 9. As can be seen from the diagram, with the increase of SNR under all four turbulence intensities, identification accuracy for all four modulation formats increased. In particular, the identification accuracy of 16QAM OFDM signals was stable at 100%. The results above show that the proposed method had the best recognition performance on 16QAM OFDM signals and was not affected by the variation of the SNR and turbulence intensity. On the other hand, influenced by the SNR and turbulence intensity, the recognition accuracy of the model based on the proposed method deteriorated for OFDM signals in BPSK, QPSK, and 8PSK formats to varying degrees. Under the conditions of a low SNR and strong turbulence intensity, the identification accuracy decreased greatly. With the increase of the SNR and turbulence, the identification accuracy gradually decreased and finally stabilized at more than 90%.

![Identification accuracy against SNR under different turbulence intensities.](image-url)
4. Conclusions

Based on the study of the characteristics of OFDM subcarrier laser signals in atmospheric time-varying channels, a multi-carrier signals method based on a multi-feature input and hybrid training neural network was proposed. The recognition performance of the multi-carrier signal recognition method based on multi-feature input and an HTNN was analyzed. The simulation results show that when the learning rate of the HTNN is $10^{-3}$, the classification accuracy performance is most stable and best, and the training process fluctuates less and converges better. The overall classification accuracy of the learning model based on the proposed method in all turbulence strengths and SNR of the time-varying channel under weak turbulence conditions reaches more than 73.5%, and converged at more than 95% with an increasing SNR. The simulation results demonstrate the feasibility of the proposed method for OFDM multi-carrier modulated signal recognition in atmospheric time-varying channels and its robustness to turbulence intensity variations and SNR.

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

Funding: This research was funded by National Key Research and Development Project of China, under Grant 2019YFB1803701, and in part by the National Natural Science Foundation of China under Grants 61835002, 61727817, 62021005, 61935005.

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

References


8. Ohtsuki, T. Multiple-subcarrier modulation in optical wireless communications. IEEE Commun. Mag. 2003, 41, 74–79. [CrossRef]


