Front. Optoelectron. 2019, 12(2): 190–196
https://doi.org/10.1007/s12200-018-0817-x

RESEARCH ARTICLE

A novel modulation format identification based on amplitude histogram space

Tianliang WANG, Xiaoying LIU (✉)

National Engineering Laboratory for Next Generation Internet Access System, School of Optical and Electronic Information (SOEI), Huazhong University of Science and Technology, Wuhan 430074, China

Abstract In this paper, we proposed a novel modulation format identification method for square *M*-quadrature amplitude modulation (*M*-QAM) signals which is based on amplitude histogram space of the incoming data after analog-to-digital conversion, chromatic dispersion compensation at the receiver. We demonstrated the identification of quadrature phase-shift keying (QPSK), 16-QAM, 64-QAM formats with an amplitude histogram space. Simulation results show that it achieve 100% identification accuracy when the incoming signal OSNR is 14 dB to identify the modulation format of QPSK, 16-QAM, and 64-QAM signals in digital coherent systems. The method has low complexity and small delay.

Keywords modulation format identification (MFI), amplitude histogram space, high-order modulation format, optical performance monitoring

1 Introduction

With the growth of network data, for example, the development of video on demand, online games, virtual reality, a higher rate of optical network is needed. At the same time, the old business should be retained. This requires that next optical network must be flexible, elastic and heterogeneous. Elastic optical networks (EON) have been researched to have more efficiency at bandwidth [1]. In the EONs, the receiver needs to identify different modulation formats, so that it can demodulate them by different algorithms.

In recent years, there are more and more modulation formats in wireless and optical communication systems [2,3]. The modulation format identification (MFI) in optical fiber communication systems has attracted more and more scholars' interest. One method is based on the peak to average power ratio (PAPR) of received signal [4–6]. It identifies the modulation format of the signal by collecting the distribution of the value of the power of the signal, but it only identify some existing fixed modulation formats. Another method is based on asynchronous delay tapped sampling (ADTS), two acquisition cards collect the same optics signal. The delay time is fixed [7–11]. The amplitude of the acquisition is P_i and Q_i, respectively. The distribution of statistical P_i and Q_i can be made into a two-dimensional amplitude histogram. The PQ two-dimensional histogram of the optical signals with different modulation formats and different bit rates is also different. So different modulation formats are identified by analyzing the differences of these histograms. But it requires two high-speed acquisition cards to sample the same optical signal asynchronously, so the cost is higher. The other method is to use deep neural network based on the asynchronous amplitude histograms of the received signal [12–15]. The neural network was trained with asynchronous amplitude histograms to identify the modulation formats. But it needs a lot of number of input data and the iterative algorithm of the neural network is too complex.

In this paper, we propose a novel MFI method for square *M*-quadrature amplitude modulation (*M*-QAM) optics signals based on amplitude histogram space which has lower time complexity. We demonstrate identification of 35 GBaud quadrature phase-shift keying (QPSK), 16-QAM, 64-QAM signals. We explain the principle of the proposed MFI in Section 2. In Section 3, we will show simulation setup. In Section 4, we will discuss about simulation results. In Section 5, it is the conclusion of our work.

2 Principle of the proposed MFI

This paper mainly discusses three common high-order modulation formats, which are QPSK, 16-QAM and 64-

Received February 11, 2018; accepted May 3, 2018

E-mail: liuxy@mail.hust.edu.cn

QAM. First the constellation of the three modulation formats is analyzed. The VPI TransmissionMaker™ simulation software is used to set up the transmission and the receiving laser's linewidth to 100 kHz. Single polarization signals are transmitted in the case of back to back. The constellation diagrams of the three modulation formats are obtained before equalization, as shown in Fig. 1. Due to the effect of laser linewidth, the points on the constellation diagram are not ideal to gather to a point, but a certain deflection. At the same time, because of the phase error, the constellation we obtained is not strictly along the X axis or Y axis symmetry, but rotated around the origin point (0, 0). The constellation deflection and rotation will not affect the amplitude of the signal. Therefore, the scheme has a high tolerance to frequency offset and phase error.

Combined with Figs. 1 and 2, even though the influence of laser linewidth and phase error, the distribution of amplitude histogram of high-order modulation format is basically unchanged. It can be seen from the figures that:

1) Amplitude histogram of QPSK signal has only a vertical line, that constant amplitude QPSK signal accords with the theoretical analysis.

2) There are three vertical lines at 16-QAM signal amplitude histogram. 16-QAM signal has three kinds of power value, which is consistent with our theoretical analysis.

3) 64-QAM signal amplitude histogram has nine peak vertical line. There are nine kinds of 64-QAM signal amplitude.

From the above analysis, it can conclude that the amplitude histograms of these three modulation formats are different. Therefore, the amplitude information of signals is used to distinguish these three modulation formats to achieve the purpose of recognition. Because the amplitude is not affected by frequency offset and phase error, it has better tolerance of laser linewidth and phase noise.

In high speed optical communication systems, coherent reception is usually used. Figure 3 is a block diagram of an algorithm for digital signal processing of a coherent receiver. As shown in Fig. 3, different modulation formats have different algorithm. For example, constant modulus algorithm (CMA) used by QPSK and 16-QAM are different. So, modulation format identification (MFI) must be completed before CMA equalization. The part of the blue tagging in Fig. 3 is exactly what this paper is going to study. The chromatic dispersion (CD) will cause the widening of the optical pulse and the intersymbol interference, which will further affect the peak value of the signal. At the same time, it will also change the amplitude histogram of the signal, resulting in the change

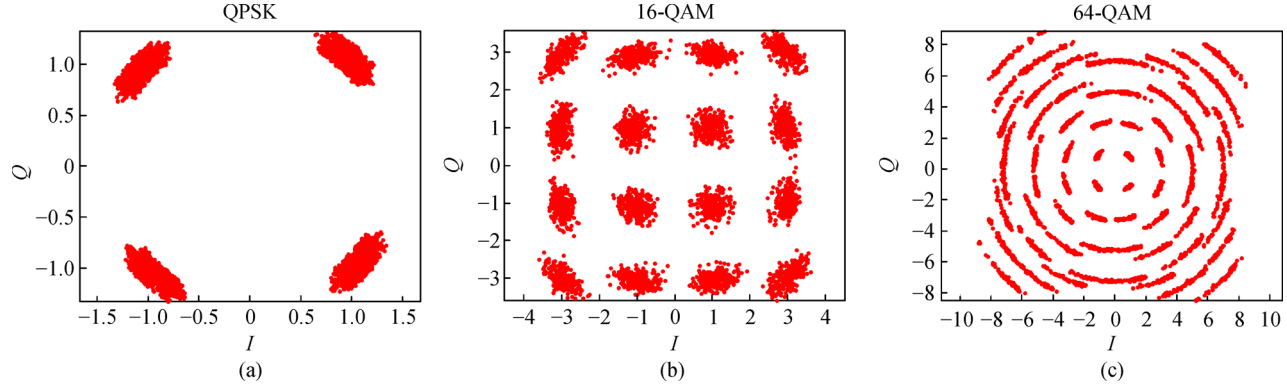


Fig. 1 Simulation constellation diagram of high-order modulation format (a) QPSK; (b) 16-QAM; (c) 64-QAM

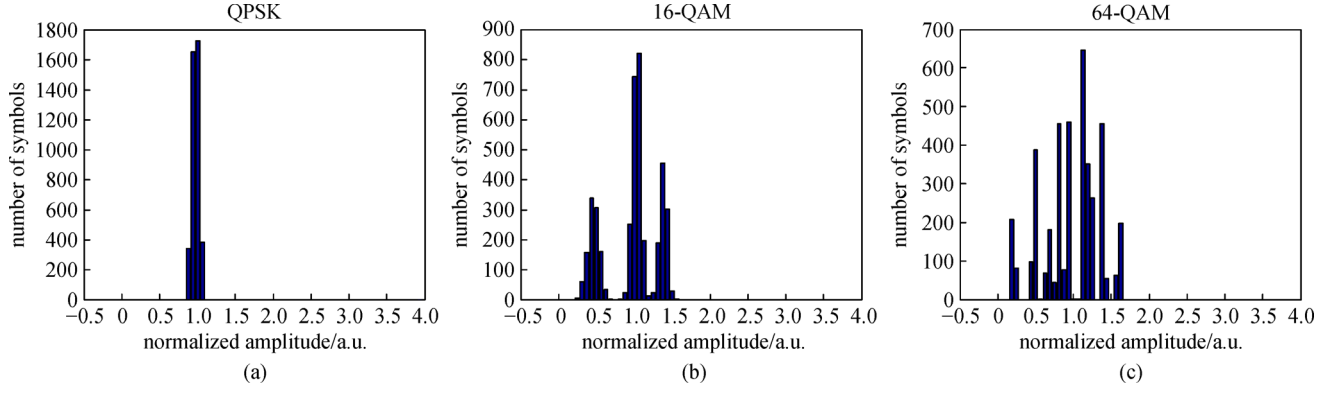


Fig. 2 Simulation amplitude histogram of high-order modulation format (a) QPSK; (b) 16-QAM; (c) 64-QAM

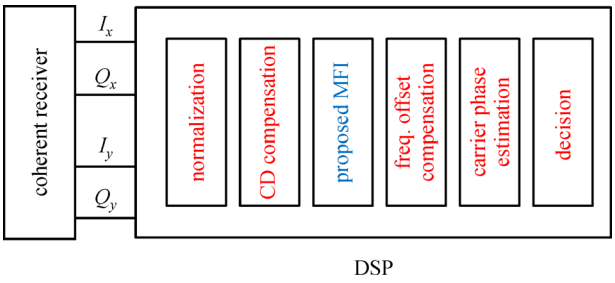


Fig. 3 Digital signal processing (DSP) flow with the proposed MFI

of the amplitude histogram of the three modulation formats. So, before MFI recognition, it is necessary to compensate chromatic dispersion. In this paper, the digital dispersion blind equalization algorithm is used to compensate the chromatic dispersion damage in the signal.

Here is our MFI scheme. First, four channels of digital signals I_x, Q_x, I_y and Q_y are acquired through coherent detection and polarization diversity reception. These four signals are combined as two orthogonal polarization signals of $I_x + \mathrm{j}Q_x$ and $I_y + \mathrm{j}Q_y$. We compute the modulus length of these two signals, and get the amplitude of these signals in each polarization direction. The statistical distribution of amplitude draws an amplitude histogram. Each histogram can be represented by a vector, denoted as $\boldsymbol{h}_1 = \{a_1,a_2,...,a_n\}$. Making the histogram vector of various modulation formats together to form a histogram matrix, denoted as $\boldsymbol{A}$.

$$[\boldsymbol{h}_1,\boldsymbol{h}_2,\boldsymbol{h}_3,...,\boldsymbol{h}_n] = \begin{bmatrix} a_1 & b_1 & c_1 & \cdots & t_1 \\ a_2 & b_2 & c_2 & \cdots & t_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_n & b_n & c_n & \cdots & t_n \end{bmatrix} = \boldsymbol{A}. \quad (1)$$

The covariance matrix $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ of $\boldsymbol{A}$ is equivalent to the singular value decomposition of $\boldsymbol{A}$, which is the representation of the difference of these amplitude histograms. We compute the eigenvalues of $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ as $\lambda_1,\lambda_2,...,\lambda_n$, and feature vectors as $v_1,v_2,...,v_n$. Then we use these eigenvectors as the basis vectors to form a pattern space V_n. Calculating the coordinates of amplitude histogram vector of known modulation format in the pattern space, denoted by $(x_1^i,x_2^i,...,x_n^i)$. We also calculate the coordinates of amplitude histogram vector of unknown optical signal received in the pattern space, denoted by $(u_1,u_2,...,u_n)$. Calculating the Euclidean distance of received optical signal amplitude histogram vector with known modulation format amplitude histogram vector in the pattern space, denoted as d_i.

$$d_i = \sqrt{\sum_{k=1}^{n} (u_k - x_k^i)^2}. \quad (2)$$

Finding the minimum value of $\{d_i\}$, assuming it is d_k. The Euclidean distance between k-th known modulation format and the received optical signal modulation format in the amplitude histogram space is closest, which means the k-th known modulation format is closest to the modulation format of received optical signal. Therefore, the modulation format of the received optical signal is identified as the format k. In this paper, there are three known modulation formats. They are QPSK, 16-QAM, and 64-QAM.

3 Simulation setup

The structure of the proposed transceiver is depicted in Fig. 4. At the transmitter-side, external cavity laser (ECL) generates a continuous laser with a center wavelength of 1550 nm, arbitrary waveform generator (AWG) produces QPSK, 16-QAM and 64-QAM electrical signals. The IQ modulator modulates the optical wave with these electrical signals. The optical channel consists of optical amplifiers and an 80 km standard single mode fiber (SSMF). Erbium-doped fiber amplifier (EDFA) amplifies received optical signals. The optical band-pass filter (OBPF) filters out the spontaneous emission spectrum of the amplifier as noise. At the receiver side, the optical signals are converted to the electrical signals by the integrated coherent receiver. After

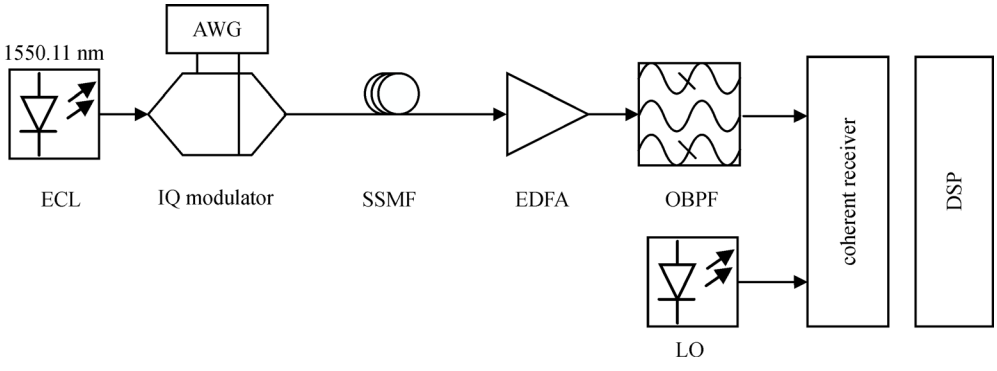


Fig. 4 Structure of the proposed transceiver. ECL: external cavity laser; EDFA: erbium-doped fiber amplifier; LO: local oscillator laser; SSMF: standard single mode fiber; OBPF: optical band-pass filter

the digitization using analog-to-digital converters (ADCs), DSP is used to recover and decode the signals. Because adaptive equalizer process and carrier phase estimation are dependent to the modulation format, MFI should be applied before these procedures. After orthonormalization, CD compensation must be processed due to the chromatic dispersion effects of optical fiber. In this simulation, the 1550 nm band is used to simulate the high-order modulation signal of single carrier after long distance transmission. As shown in Table 1, there are some simulation parameters of the optical communication system.

4 Simulation results and discussion

The VPI TransmissionMaker^{TM} software is used to build the optical communication system. The results of the simulation are obtained. Before MFI, the dispersion blind equalization algorithm is used to complete the chromatic dispersion compensation. This blind equalization algorithm is independent of the modulation formats. So it can be applied to *M*-QAM signals. Figure 5 shows the Euclidean distances of sampled signals with standard QPSK, 16-QAM and 64-QAM. In Fig. 5, the red line represents the Euclidean distance between sampling signals and the standard QPSK signal in the amplitude histogram space; the green line represents the Euclidean distance between sampling signals and the standard 16-QAM signal in the amplitude histogram space; the blue line represents the Euclidean distance between the sampling signal and standard 64-QAM in the amplitude histogram space. As shown in Fig. 5, the Euclidean distance between the sampled signal and the same type of standard signal is the minimum in the amplitude histogram space. Therefore, the modulation format of the signal is identified based on the minimum value of the Euclidean distance.

1) In Fig. 5(a), the 40 groups of random sampled QPSK signals, are average 61.5 standard space units deviation from the standard QPSK signal in the amplitude histogram space. The Euclidean distance between the sampled QPSK signals and the standard 16-QAM signals are average 516.3 standard space units. The Euclidean distance between the 40 groups of sampled QPSK signals and the standard 64-QAM format are average 565.5 standard space units. Since the sampled signals are nearest to the standard QPSK signal in the amplitude histogram space, the sampled signals are identified as QPSK signals. The average distance between the sampled signals and the standard 64-QAM is 9.2 times the average distance between them and the standard QPSK. The average distance between the sampled signals and the standard 16-QAM is 8.4 times the average distance between them and the standard QPSK signal. There is an average of 18.8 dB noise tolerance between QPSK and 16-QAM or 64-QAM.

2) In Fig. 5(b), the average Euclidean distance between the 40 groups of 16-QAM sampled signals and the standard QPSK signal in the amplitude histogram space is 510.1 standard space units. The average Euclidean distance between these sampled signals and the standard 16-QAM format is 59.5 standard space units. The average Euclidean distance between them and the standard 64-QAM format is 115.7 standard space units. Since the sampled signal is closest to the standard 16-QAM signal in the mode space, the signal is identified as a 16-QAM signal. This is in accordance with our simulation settings. The average distance from the standard QPSK signal is 8.6 times that of the standard 16-QAM. The average distance from the standard 64-QAM is 1.9 times the average distance from the standard 16-QAM. There is 6 dB noise tolerance between 16-QAM and 64-QAM.

3) In Fig. 5(c), the 40 groups of sampled signals are in the format of 64-QAM. The average distance between them and standard QPSK signal in amplitude histogram space is 572.6 standard space units. The average Euclidean distance between the sampled signals and standard 16-QAM signal is 124.2 standard space units. The average Euclidean distance between the sampled signals and standard 64-QAM format is 62.8 standard space units. Since the sampled signals of each group are the closest to the standard 64-QAM signal in the amplitude histogram space, these 40 groups of sampled signals are all identified as 64-QAM signals. This is the same as our simulation setting. The average distance from the standard QPSK signal is 9.1 times that of the standard 64-QAM. The average distance from the standard 16-QAM signal is 2 times that of the standard 64-QAM. Therefore, this again verifies that the noise tolerance of 16-QAM and 64-QAM is 6 dB.

In summary, we conclude that in the amplitude histogram space, the average noise tolerance between QPSK signal and 16-QAM signal is 18.6 dB. There is 19.2 dB noise tolerance between QPSK and 64-QAM. The noise tolerance between 16-QAM and 64-QAM is lowest, which is only 6 dB.

In this section, we will discuss the effect of optical signal to noise ratio (OSNR) on recognition results. The OSNR is adjusted by adding Gauss white noise to the original

Table 1 Simulation parameters of optical communication system

λ_c/nm	baud rate	sample rate	laser power/mW	laser linewidth/kHz	fiber loss/(dB·km^{-1})	CD/(ps·nm^{-1}·km^{-1})
1550	35e9	70e9	2	100	0.2±0.01	16±0.2

Note: λ_c, center wavelength

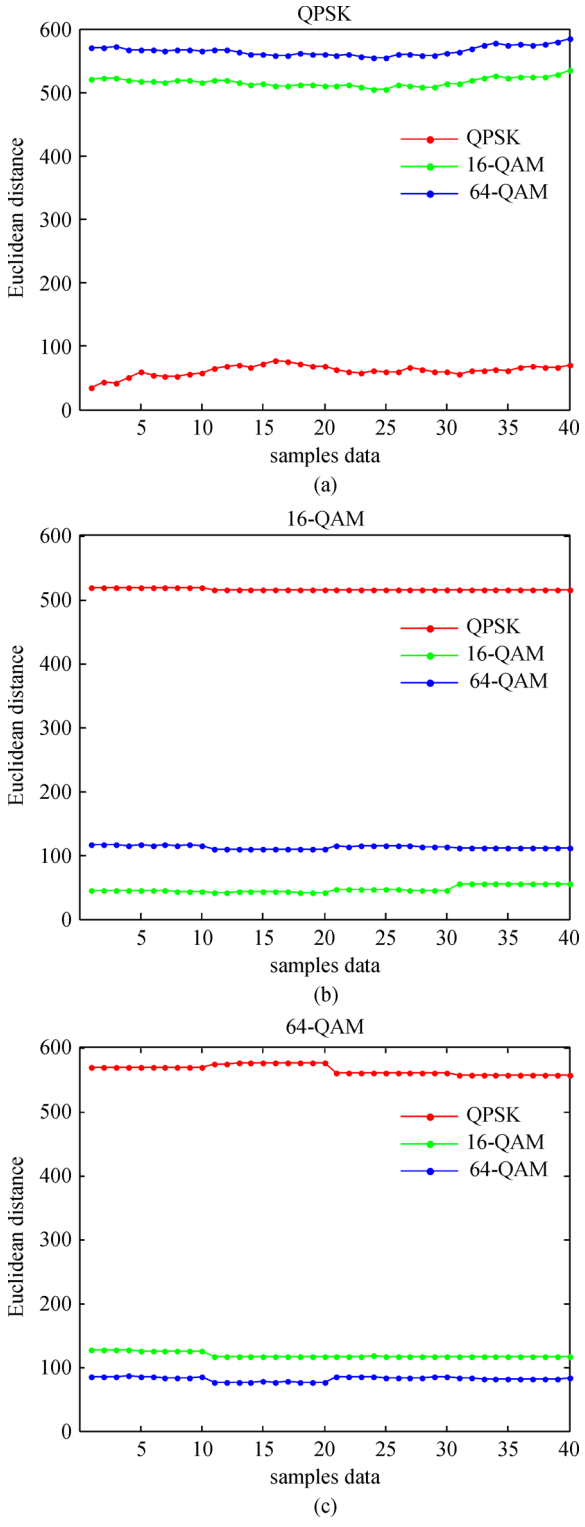


Fig. 5 Calculated Euclidean distance of the samples signals with standard (a) QPSK, (b) 16-QAM and (c) 64-QAM in amplitude histogram space

optical signal. The calculation of signal-to-noise ratio is shown in Eq. (3).

$$\mathrm{OSNR(dB)} = 10\log_{10}\left(\frac{P_\mathrm{s}}{P_\mathrm{n}}\right), \tag{3}$$

where P_s is the power of the signal and P_n is the power of the noise.

A group of samples are obtained at each 1 dB from 1 to 20 dB. And the QPSK, 16-QAM and 64-QAM signals are simulated respectively. The signal data sampled from the coherent receiver is converted to the amplitude histogram space after the normalization and the chromatic dispersion blind equalization. The Euclidean distances between the sampled signals and the standard QPSK, the standard 16-QAM, and the standard 64-QAM format in the amplitude histogram space are calculated. The relationship between the Euclidean distances and OSNR is concluded in Fig. 6.

Figure 6(a) shows the relationship between the Euclidean distances of the QPSK sampled signal with the standard modulation formats in the amplitude histogram space and its OSNR. When OSNR is greater than 12 dB, the distance between the sampled QPSK signal and the standard QPSK signal is less than that of the standard 16-QAM and the standard 64-QAM. And with the increase of OSNR, the distance between the sampled QPSK signal and the standard QPSK signal is getting smaller and smaller. The distance between the sampled signal and the standard 16-QAM format or 64-QAM format is increasing in the amplitude histogram space. This shows that the greater the OSNR, the higher the accuracy of recognition. And the minimum OSNR that can be correctly identified is 12 dB.

Figure 6(b) shows the relationship between the Euclidean distances of the sampled 16-QAM signal and the standard modulation formats and the OSNR in the amplitude histogram space. The OSNR of the arrow is the minimum OSNR that is correctly identified. When OSNR is greater than 14 dB, the distance between the sampled 16-QAM signal and the standard 16-QAM is the minimum of the three. And with the increase of OSNR, the distance between the sampled signal of 16-QAM and the standard 16-QAM format in the amplitude histogram space is always the minimum. This shows that the minimum OSNR tolerance for 16-QAM signals is 14 dB.

The relationship between the Euclidean distance and the OSNR of the sampled 64-QAM signal in the amplitude histogram space with the standard modulation format is shown in Fig. 6(c). The distance between the sampled 64-QAM format and the standard 64-QAM format in the amplitude histogram space is the minimum of the three when OSNR is greater than 1 dB. And with the increase of OSNR, the change of Euclidean distance is not large. This shows that in the case of only three modulation formats, the relationship between the recognition rate and the OSNR is not significant for the 64-QAM signal.

In conclusion, to correctly identify the sampled signal is which of the three types (QPSK, 16-QAM and 64-QAM), the minimum optical signal to noise ratio is 14 dB.

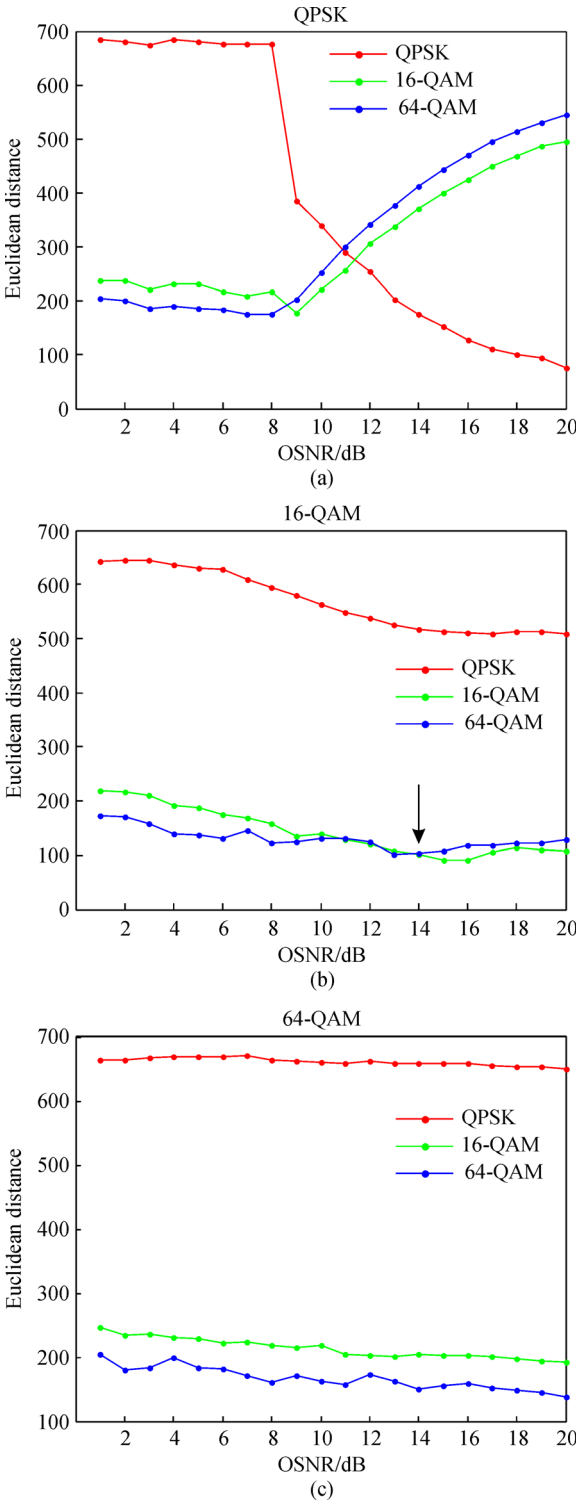


Fig. 6 Relationship between the Euclidean distance of the sampled signal with the standard modulation format and the OSNR. (a) QPSK; (b) 16-QAM; (c) 64-QAM

5 Conclusions

In this paper, we proposed a novel modulation format identification method for square *M*-QAM signals based on amplitude histogram space. We simulate this MFI method with VPI TransmissionMaker™ software in the case of single carrier. The identification of the formats of multi-carrier signals, such as orthogonal frequency division multiplexing (OFDM) has not been verified. Based on the minimum value of the Euclidean distances of the sampled signal with standard modulation formats in amplitude histogram space, it is possible to identify the modulation format of QPSK, 16-QAM, and 64-QAM signals when the OSNR is greater than 14 dB. No extra hardware is needed. Utilizing the digital signal processing, this method has low cost.

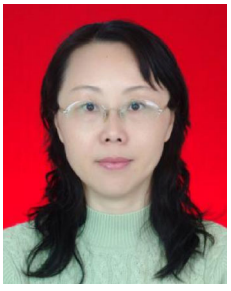
Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Nos. 61575071, and 61331010).

References

1. Paper C W. Cisco Global Cloud Index: Forecast and methodology, 2016–2021. San Jose: Cisco Systems Inc, 2018, 3–13
2. Hu S, Yang X, Yuan J. Indoor ultra-wideband channel modeling based on time-domain data. Communications Technology, 2013, 46 (3): 10–15
3. Shen D, Li P, Zheng J, Zhao M, Ma Q, Zhou W, Zhao Z. Research of ultra-wideband pulse amplitude modulation based on $LiNbO_3$ Mach-Zehnder modulator. Journal of Optoelectronics • Laser, 2013, 24(1): 79–86
4. Bilal S M, Bosco G, Dong Z, Lau A P, Lu C. Blind modulation format identification for digital coherent receivers. Optics Express, 2015, 23(20): 26769–26778
5. Liu J, Dong Z, Zhong K P, Lau A P T, Lu C, Lu Y. Modulation format identification based on received signal power distributions for digital coherent receivers. In: Proceedings of Optical Fiber Communication Conference. San Francisco: OSA Technical Digest, 2014, Th4D.3
6. Liu J, Zhong K, Dong Z, Guo C, Lau A P T, Lu C, Lu Y. Signal power distribution based modulation format identification for coherent optical receivers. Optical Fiber Technology, 2017, 36: 75–81
7. Guesmi L, Menif M. Method of joint bit rate/modulation format identification and optical performance monitoring using asynchronous delay-tap sampling for radio-over-fiber systems. Optical Engineering (Redondo Beach, Calif.), 2016, 55(8): 084108
8. Khan F N, Zhou Y, Sui Q, Lau A P T. Non-data-aided joint bit-rate and modulation format identification for next-generation heterogeneous optical networks. Optical Fiber Technology, 2014, 20(2): 68–74
9. Tan M C, Khan F N, Alarashi W H, Zhou Y, Tao Lau A P. Simultaneous optical performance monitoring and modulation format/bit-rate identification using principal component analysis. Journal of Optical Communications and Networking, 2014, 6(5): 441–448
10. Khan F N, Yu Y, Tan M C, Al-Arashi W H, Yu C, Lau A P, Lu C. Experimental demonstration of joint OSNR monitoring and

modulation format identification using asynchronous single channel sampling. Optics Express, 2015, 23(23): 30337–30346
11. Khan F N, Lu C, Lau A P T. Joint modulation format/bit-rate classification and signal-to-noise ratio estimation in multipath fading channels using deep machine learning. Electronics Letters, 2016, 52 (14): 1272–1274
12. Guesmi L, Hraghi A, Menif M. Automatic modulation format recognition for the next generation optical communication networks using artificial neural networks. In: Proceedings of SPIE-Physics and Simulation of Optoelectronic Devices XXIII. SPIE, 2015, 9357: 93571I
13. Khan F N, Zhong K, Al-Arashi W H, Yu C, Lu C, Lau A P T. Modulation format identification in coherent receivers using deep machine learning. IEEE Photonics Technology Letters, 2016, 28 (17): 1886–1889
14. Zhang S, Peng Y, Sui Q, Li J, Li Z. Modulation format identification in heterogeneous fiber-optic networks using artificial neural networks and genetic algorithms. Photonic Network Communications, 2016, 32(2): 246–252
15. Cui S, He S, Shang J, Ke C, Fu S, Liu D. Method to improve the performance of the optical modulation format identification system based on asynchronous amplitude histogram. Optical Fiber Technology, 2015, 23: 13–17

Tianliang Wang received the bachelor degree from Huazhong University of Science and Technology. He is currently pursuing the Master degree in National Engineering Laboratory for Next Generation Internet Access System of Huazhong University of Science and Technology. His present research work involves modulation format identification in high speed optical communication systems.



Xiaoying Liu is an Associate Professor in the School of Optical and Electronic Information at Huazhong University of Science and Technology, Wuhan, China. She was a postdoctoral researcher of College of Optoe-lectronics Science and Technology at Huazhong University of Science and Technology in 2003. Her current research includes advanced modulation formats, all-optical tunable filter, micro-ring resonators, optical fiber sensor and wireless sensor network.