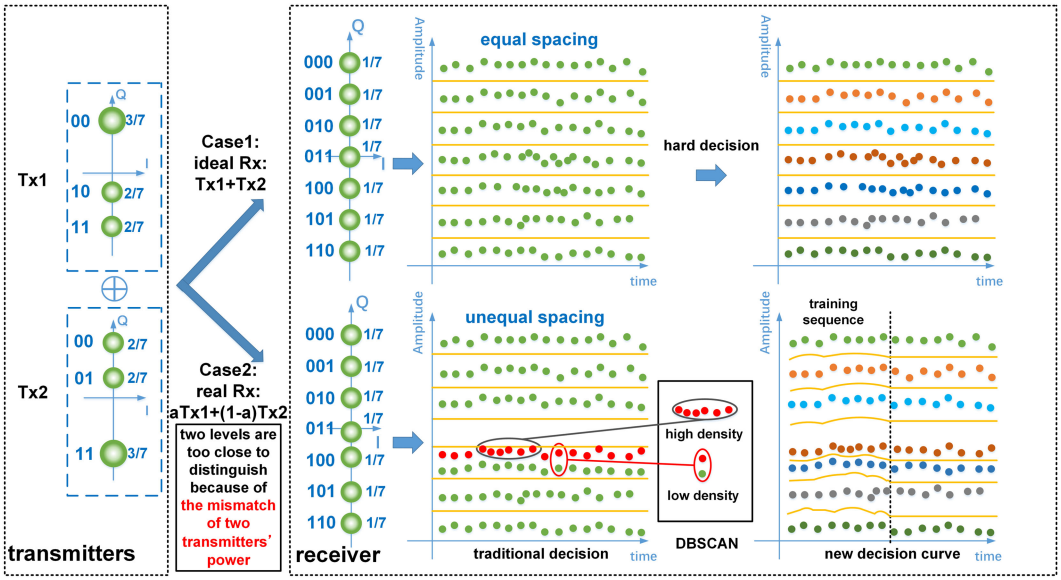
IEEE Photonics Journal
An IEEE Photonics Society Publication

Open Access

Enhanced Performance of PAM7 MISO Underwater VLC System Utilizing Machine Learning Algorithm Based on DBSCAN

Volume 11, Number 4, August 2019

Meng Shi
Yiheng Zhao
Weixiang Yu
Yuchong Chen
Nan Chi, *Member, IEEE*



DOI: 10.1109/JPHOT.2019.2928827

Enhanced Performance of PAM7 MISO Underwater VLC System Utilizing Machine Learning Algorithm Based on DBSCAN

Meng Shi, Yiheng Zhao, Weixiang Yu, Yuchong Chen, and Nan Chi, *Member, IEEE*

Key Laboratory for Information Science of Electromagnetic Waves (MoE), Shanghai Institute for Advanced Communication and Data Science, Fudan University, Shanghai 200433, China

DOI:10.1109/JPHOT.2019.2928827

Manuscript received May 13, 2019; revised July 6, 2019; accepted July 11, 2019. Date of publication July 15, 2019; date of current version July 29, 2019. This work was supported in part by National Key Research and Development Program of China under Grant 2017YFB0403603, and in part by the NSFC project under Grant 61571133. Corresponding author: Nan Chi (e-mail: nanchi@fudan.edu.cn).

Abstract: After visible light communication drawing increasing attention, underwater visible light communication (UVLC) has attracted more interest in the research community nowadays. As multiple input single output (MISO) is getting increasingly widely used to improve the transmission speed in UVLC system, the unbalance between multiple transmitters' power is still a common phenomenon, which leads to the unequal spacing between each adjacent level and damages the system performance. In this paper, we study and analyze the unbalance between the two transmitters. Compared to a traditional hard decision, a density-based spatial clustering of applications with noise (DBSCAN) of a machine learning method is employed to get the actual center of each cluster and distinguish each level of PAM7 signals. In this way, a new decision curve substitutes traditional standard straight line as a constellation discrimination method. The experimental results show that up to 1.22 Gb/s over 1.2 m underwater visible light transmission can be achieved by using DBSCAN for PAM7 MISO signals. The measured bit error rate is well under the hard decision-forward error correction threshold of 3.8×10^{-3}.

Index Terms: MISO, DBSCAN, underwater visible light communication.

1. Introduction

Visible light communication (VLC) has become an attracting and promising technology for high-speed indoor wireless communications over ten years [1]–[6]. The major issue to improve the data rate and transmission distance in the VLC system is how to overcome system bandwidth limitation and compensate multiple impairments effectively. To solve the above problems, new high spectral efficiency modulation formats and equalization techniques are investigated by lots of research institutes [7]–[15]. A high-speed VLC system can be achieved by employing orthogonal frequency division multiplexing (OFDM) [7], [8], discrete fourier transform spread orthogonal frequency division multiplexing (DFT-S OFDM) [9], [10], carrier-less amplitude and phase (CAP) modulation [11], [12], pulse amplitude modulation (PAM) [13], discrete multi-tone (DMT) modulation and so on. It shows the feasibility of these advanced modulation formats in the VLC system.

Nowadays, underwater visible light transmission is an interesting area need to be studied [16]–[24]. Besides the limitation exists in VLC, such as limited modulation bandwidth, non-linearity impairments of transceivers, the absorption, scattering and diffraction effect in UVLC would attenuate the transmitting signals and induce more non-linearity of the UVLC system. Researchers make an effort to increase both transmission speed and link distance by adopting advanced modulation formats and signal processing techniques. A data rate of 161 Mb/s system over 2 m distance, which adopted OFDM modulation format and blue light emitting diode (LED) was achieved by Xu *et al.* from Zhejiang University in 2016 [19]. In 2017, Tian *et al.* in Fudan University considered the complexity and selected on-off keying (OOK) modulation to realize a 200 Mb/s transmission over 5.4 m [20]. Kong *et al.* realized a system with arrayed transmitter/receiver and optical superimposition PAM4 signals and reached 12.288 Mb/s after transmitting through a 2 m underwater channel [21].

Considering to improve the communication speed with acceptable complexity, single input single output (SISO) is replaced by multiple input single output (MISO) technique [25]–[28]. We proposed an equiprobable pre-coding PAM7 modulation for non-linearity mitigation in 2×1 MISO UVLC system [28]. And a data rate of 918.75 Mb/s transmission was successfully achieved. However, we found that even the two LEDs we adopted in our previous experiment are under the same working point, there are difference exists which produce unbalance between transmitters. The unbalance between two LEDs will destroy the superimposed signals in PAM7 MISO system, cause symbol misjudgment and damage the performance. So the mismatch between two transmitters needs to be investigated.

Recently, advanced digital signal processing algorithm, nonlinear compensation and other algorithms based on machine learning are being investigated. Several VLC systems have employed machine learning to improve the performances [29]–[31]. S. Ma proposed three ML demodulators in VLC system based on convolutional neural network (CNN), the deep belief network (DBN), and adaptive boosting (AdaBoost), respectively. [30]. What's more, the DBSCAN algorithm is one of the most typical density-based clustering algorithms of machine learning [32]. DBSCAN can find actually center of clusters according to the density.

In this paper, we conduct a detailed theoretical and experimental analysis on the mismatch between the two transmitters'. At the same time, we use DBSCAN to process the signals at the receiving end, which can solve the wrong judgment problem caused by the mismatch between the two LEDs. To reduce the complexity of employing DBSCAN, a part of the data can be used as a training sequence to obtain an actual center of each level. Then a new decision level for remaining data is calculated. Finally, a 1.22 Gb/s PAM7 MISO system is successfully achieved over 1.2 meters of UVLC transmission under HD-FEC threshold of 3.8×10^{-3}.

The remainder of this paper is organized as follows. The theoretical analysis about the mismatch between two transmitters in PAM7 MISO system is presented first. Then simulation results, the experimental setup and experimental results in UVLC are showed. Final is the conclusion.

2. Principle

Firstly, the PAM7 signals generated by the direct superimposition of two special PAM3 signals which adopted the equiprobable pre-coding [25] method. The signals can be expressed as

$$\mathrm{Rx(PAM7)} = aTx_1(PAM3) + (1-a)Tx_2(PAM3) \tag{1}$$

$$Tx_1 = \{-3, 1, 3\}, Tx_2 = \{-3, -1, 3\} \tag{2}$$

Where a is the degree of mismatch between the two transmitters. When $a = 0.5$, the PAM7 signals are standard with equal probability.

Fig. 1 shows the original level, new level and deviation of each level at the receiver of PAM7 signals. In an ideal situation, the probability of seven levels in PAM7 signals is equal to 1/7. However, there are mismatch exists in practical systems. By calculating according formula (1), we can get a new expression of each level in PAM7 symbols. The new levels are $+6$, $-4a + 6$, $8a - 2$, $-12a + 6$, $8a - 6$, $-4a - 2$, -6, respectively. In traditional digital signal processing of PAM modulation, if one

Number	Original level	probability	New level	deviation
1	+6	1/7	+6	0
2	+4	1/7	-4a+6	|-4a+2|
3	+2	1/7	8a-2	|8a-4|
4	0	1/7	-12a+6	|-12a+6|
5	-2	1/7	8a-6	|8a-4|
6	-4	1/7	-4a-2	|-4a+2|
7	-6	1/7	-6	0

Fig. 1. Original level, new level and deviation of each level at the receiver of PAM7 signals.

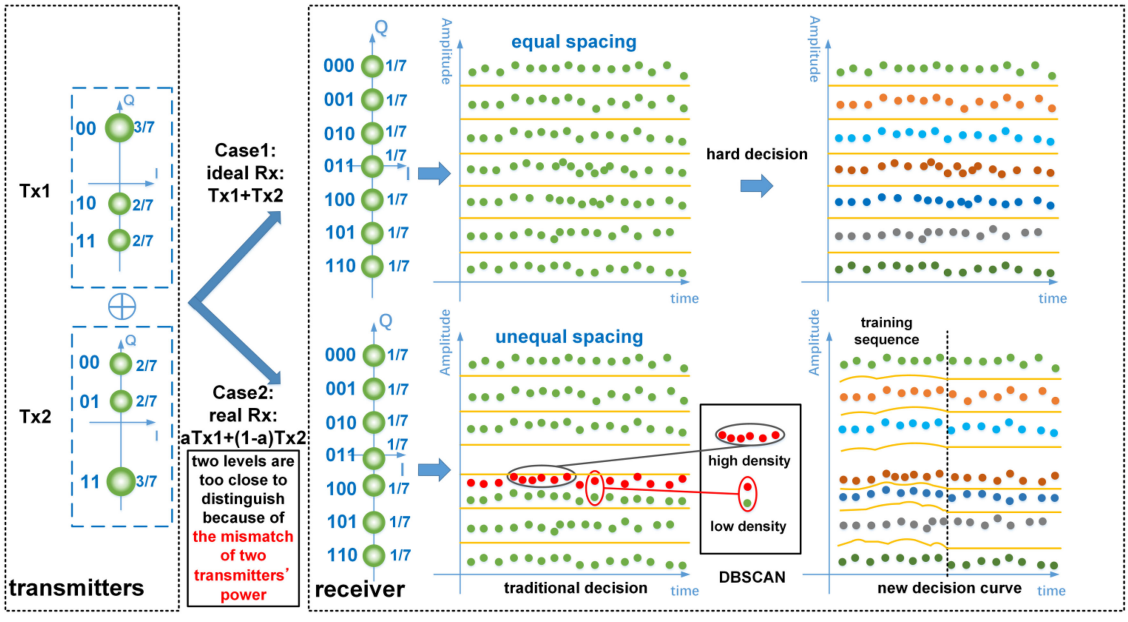


Fig. 2. Principle of PAM7 MISO system employing DBSCAN method.

level deviates from the standard level more than 0.5 amplitude, it would be misjudged to another level and cause error symbol.

Fig. 2 shows the principle of PAM7 MISO system employing DBSCAN method. At the transmitters, the PAM3 signals are an unequal probability. In case1 (ideal situation), after superimposition of two PAM3 signals, an equiprobable PAM7 signal can be generated. The PAM7 signals are with equal spacing. So it is easy to distinguish each level by adopting the traditional hard decision. However, in a real situation (case2), there always is a mismatch between the two transmitter's power. As a

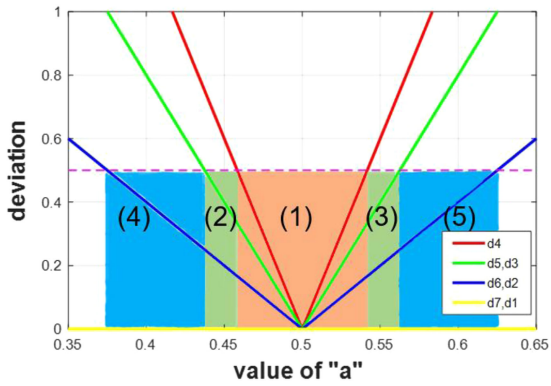


Fig. 3. Deviation of each level versus the value of a.

Number	area of a	misjudged level
(1)	[0.46,0.54]	/
(2)	[0.44,0.46]	d4
(3)	[0.54,0.56]	d4
(4)	[0.38,0.44]	d3,d4,d5
(5)	[0.56,0.62]	d3,d4,d5

Fig. 4. Different area of a and misjudged PAM level in each area.

result, the PAM7 signals may have unequal spacing. And the two levels are too close to distinguish each other accurately using the traditional hard decision as drawn in Fig. 1. What's important, the density is different. By applying DBSCAN which is a density clustering method, the signals with different density can be classified to different clusters. To reduce the complexity, we use a certain percentage of data as a training sequence to applying DBSCAN to find the actual center of each cluster. Then the average value of each level and new decision level are calculated in order. Finally, we use the new decision curve to re-judgment received PAM7 signals.

To see the change curve of deviation of each level, Fig. 3 is drawn. From the figure, we can find that d7, d1 which represents $+6$ and -6 level in PAM7 signals would not be misjudged all the time. The reason is that normalization adopted in data procession, and the two levels are always decided to $+6$ and -6. When the value of a is larger than 0.5, the deviation of d2, d3, d4, d5, d6 will increase with a. On the contrary, when the value of a is smaller than 0.5, the deviation of d2, d3, d4, d5, d6 will decrease with a. As we discussed before, when the deviation larger than 0.5 amplitude, the signals would be misjudged. So a pink dot line is added in Fig. 3. As a result, the area of a can be divided into five parts. Fig. 4 shows the detailed value of a and misjudged level in each part. For example, the detailed value of a of area(1) in Fig. 2 is [0.46, 0.54]. When the value of a is in [0.46, 0.54] in an ideal situation, no level will be misjudged. However, when a is in [0.44, 0.46] or [0.54, 0.56], d4 which represents the middle level of PAM7 signals, i.e., 0, would be misjudged because of the mismatch between two transmitters. What's more, when the degree of mismatch increase and a is in [0.38, 0.44] or [0.56, 0.62], the number of the misjudged levels will increase to 3. The

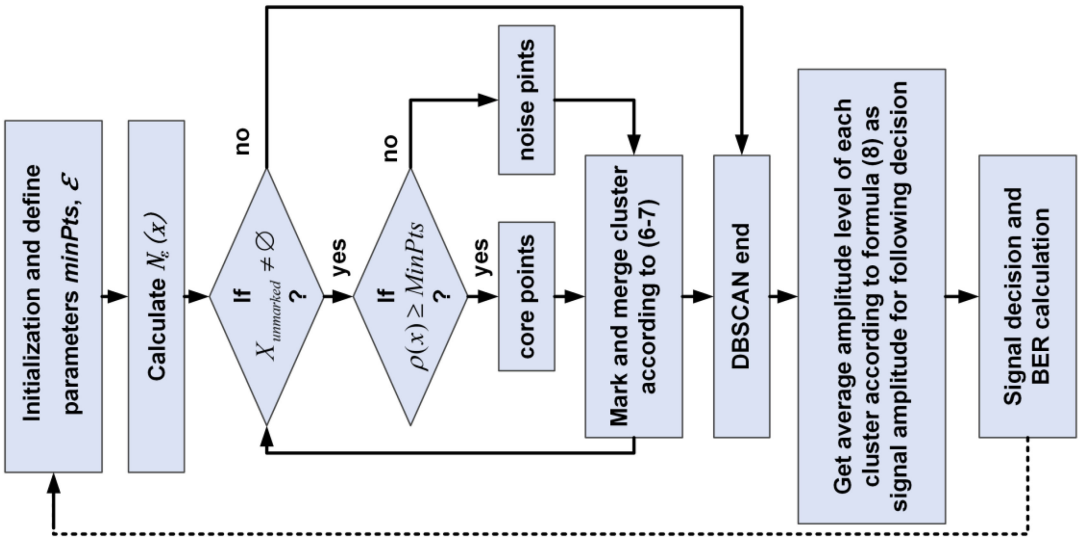


Fig. 5. Flow chart of DBSCAN utilized in PAM7 MISO UVLC system.

d3, d4, and d5 level will be misjudged. And in this case, it's hard to improve the BER performance because too many levels are stacked together. So in this paper, we only consider how to correct the misjudged levels when *a* is in [0.44, 0.54].

DBSCAN is adopted in this paper to process the received data. Fig. 5 shows the flow chart of DBSCAN utilized in our PAM7 MISO UVLC system. First, we initialize and define parameters. There are two important parameters in DBSCAN algorithm. *MinPts* represents the minimum number of points in a given area. ε is the minimum neighborhood radius of a given point. Then we calculate

$$N_\varepsilon(x) = \{y \in X : d(y, x) \leq \varepsilon\} \tag{3}$$

$X = \{x(1), x(2), \ldots, x(N)\}$ is the data set. $d(y, x)$ is the distance between y and x. Here, Euclidean distance is adopted. $N_\varepsilon(x)$ is the area whose radius is less than ε. Then each point will be marked in order until $X_{unmarked} = \emptyset$.

$$\rho(x) = \left|N_\varepsilon(x)\right| \tag{4}$$

$\rho(x)$ is the density of ε neighborhood. It can be understood that if there are more points in the ε neighborhood of x, the density is larger. For each point, it is needed to decide the identity according to formula (5).

$$x \in \begin{cases} \mathrm{X_{Core}}, & \rho(x) \geq MinPts \\ \mathrm{X_{Noise}}, & \rho(x) < MinPts \end{cases} \tag{5}$$

$\mathrm{X_{Core}}$ is the collection of all core points in X. It corresponds to a point inside the dense area. However, $\mathrm{X_{Noise}}$ is the collection of all noise points in X. It corresponds to a point in the sparse area. The noise point is the interference data in the data set, it does not belong to any cluster.

The goal of the DBSCAN algorithm is to divide the data set X into K clusters and noise point sets. We need to mark and merge clusters according to formula (6–7). A tag array m_i can be generated by:

$$m_i \in \begin{cases} k(k > 0), & x(i) \in cluster_k \\ -1, & x(i) \in X_{\mathrm{Noise}} \end{cases} \tag{6}$$

$$cluster_k = N_\varepsilon(x_i) \cup N_\varepsilon(x_j), \quad x_{i,j} \in X_{\mathrm{Core}} \tag{7}$$

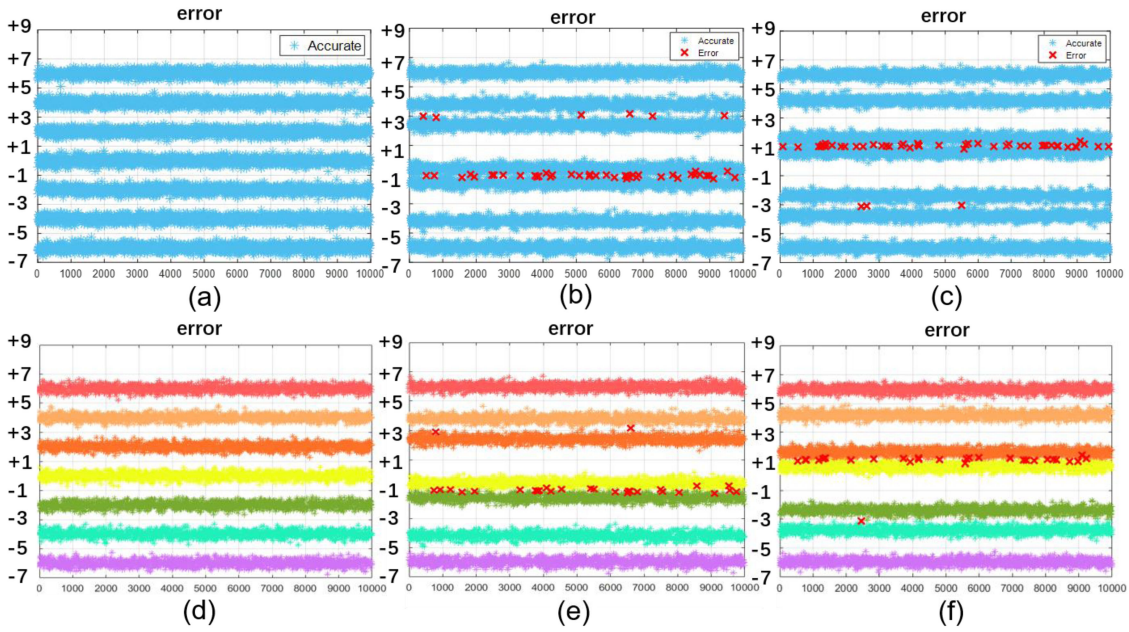


Fig. 6. PAM7 signals at receiver w/o DBSCAN in different area of a: (a) in area(1), (b) in area(2), (c) in area(3); w/ DBSCAN in different area of a: (d) in area(1), (e) in area(2), (f) in area(3).

x_i and x_j is the adjacent core point. A new $cluster_k$ will be merged by adjacent core points and its accessory nodes. Next, it is important to get the average amplitude level of each cluster according to formula (8) as the signal amplitude for the following decision.

$$Y_{out\_i} = \frac{1}{n}\overline{\sum_{i=1}^{n} x_i}, \quad x_i \in cluster_k \tag{8}$$

Where Y_{out_i} is the data out after employing DBSCAN. It is the mean value of each point in the same cluster. n is the number of points in $cluster_k$. Finally, perform signal decision and calculate BER for processed data.

Fig. 6 shows PAM7 signals at the receiver without or with DBSCAN in a different area of a. As we discussed before, Fig. 6(a) and (d) which is in the area(1) will have zero error because no level would be misjudged. Fig. 6(b) and (e) are in area(2). It can be easily found that level 0 and -2 are stacked closely which is hard to distinguish each other. The red cross is the point which is misjudged and the blue star is the point which is judged accurately. After adopting the DBSCAN algorithm, the number of misjudged points decrease which results in better BER performance. Fig. 6(c) and (f) are in the area(3) and level 0 and -2 are too close to distinguish each other correctly. It also can be improved by utilizing DBSCAN from the figure.

Based on the above analysis, we found DBSCAN can resist the mismatch between two transmitters. In order to know how the effect is in our previous experiment [25], DBSCAN is adopted. Fig. 7 shows PAM7 MISO experimental BER versus the value of a with or without DBSCAN. The value of a is calculated according to previous Vpp value of each transmitter in our previous experiment in [25]. From the figure, we can find that BER performance will be destroyed when the value of a is too large or too small which means a high degree of mismatch. What's important, the area of a under FEC threshold can be enlarged by employing DBSCAN. The error points will decrease too.

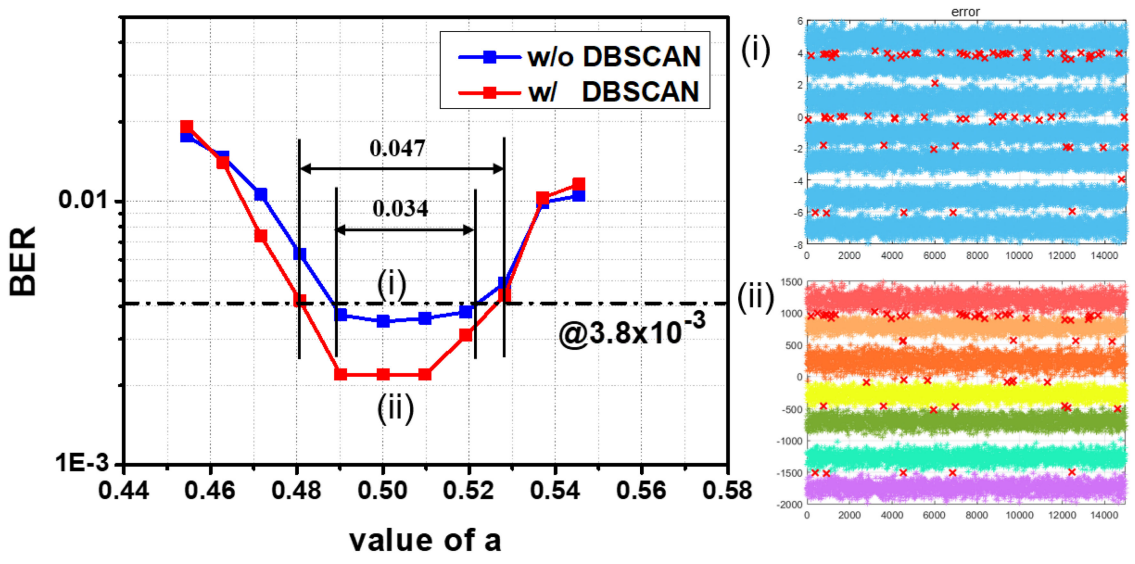


Fig. 7. PAM7 MISO experimental BER versus value of a with or without DBSCAN.

3. Numerical Simulation Analysis

To investigate the effect of mismatch between two transmitters, new simulation results and experimental results are presented followed. First of all, to quantify the mismatch between the two transmitters caused by the devices, we processed the transmitting signals in MATLAB, which meets the formula (9). And the two LEDs in transmitters were substituted by only one LED. In this way, the effect of two devices itself can be excluded partly, the system becomes simpler and the degree of mismatch between the two transmitters can be accurately quantified. Actually, the signals transmitted are generated by direct superposition according to formula (9). To get as close as possible to the actual situation, Gaussian white noise is added to the original transmitting signals according to formula (9). The signal to noise ratio (SNR) is expressed in formula (10).

$$\begin{aligned} &Tx\_new = a(Tx_1 + awgn) + (1-a)Tx_2, \quad a < 0.5 \\ &\text{or} \quad aTx_1 + (1-a)(Tx_2 + \text{awgn}), \quad a > 0.5 \end{aligned} \tag{9}$$

$$\text{SNR} = 30 + 20\log_{10}\left(\frac{a}{1-a}\right) \tag{10}$$

Where, *Tx_new* represents new transmitting signals we adopted, *awgn* is Gaussian white noise we added to the transmitter. From formula (9), the PAM3 signals in transmitter who has smaller amplitude will be added Gaussian white noise. In this way, the system performance will be worse and closer to the real situation in practice.

Fig. 8 shows the simulated BER versus the value of *a* with or without DBSCAN. It indicates that the DBSCAN can improve the BER performance when *a* is in a certain area. On the contrary, when the degree of mismatch is larger, DBSCAN will destroy system performance.

4. Experimental Setup and Results

The experimental setup of PAM7 UVLC system is shown in Fig. 9. At the transmitter, the original binary bit sequence is first mapped in MATLAB to PAM3 symbols according to the equiprobable pre-coding method presented in [25]. The PAM3 symbols are then encoded using PS-Manchester method for spectral shaping. In this way, PAM signals can be transmitted directly at baseband. Then signals are up-sampled by 4 times. Next, two PAM3 signals were added together according

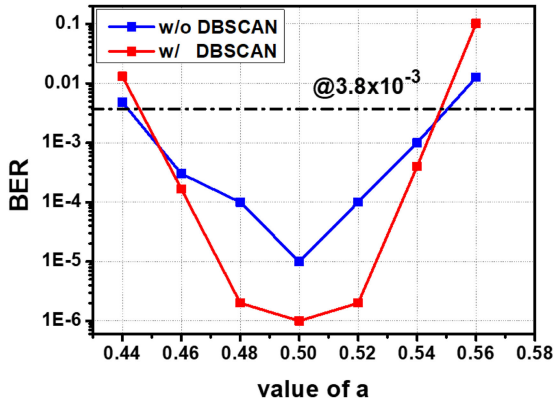


Fig. 8. PAM7 simulated BER versus the value of *a* with or without DBSCAN.

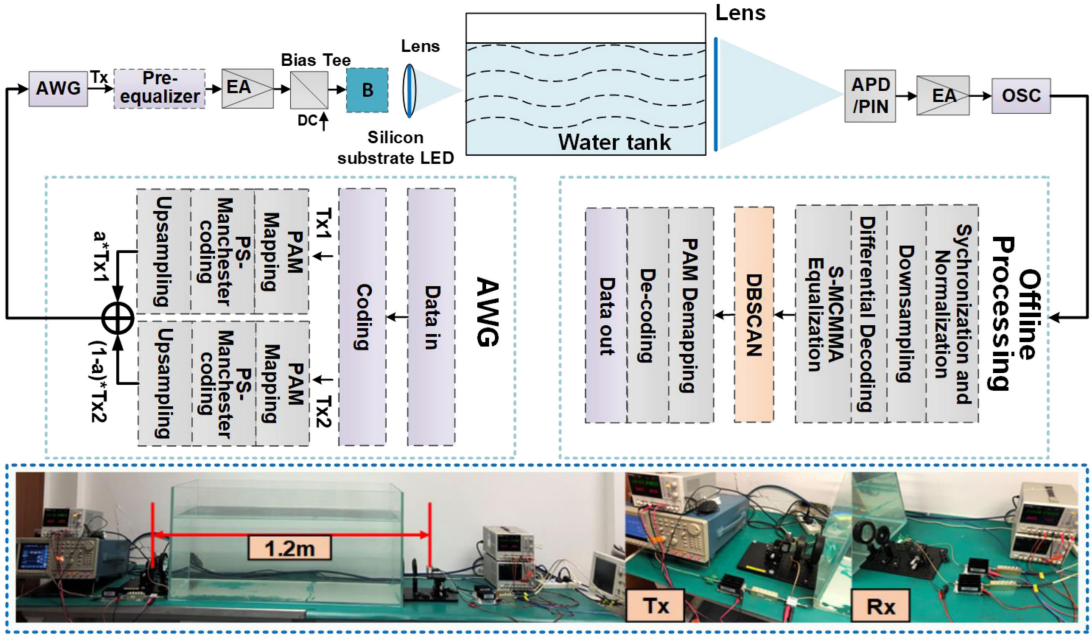


Fig. 9. Experimental setup of PAM7 UVLC system.

to formula (9) to form a new PAM7 symbol. In the experiment, electrical signals were generated by inputting data into the channel of an arbitrary waveform generator (AWG, Tektronix AWG710B). Then to compensate high-frequency attenuation of the channel, signals are passed through a self-designed bridge-T base pre-equalizer (hardware pre-equalization). Followed by an electrical amplifier (EA, 25 dB gain) and DC bias tee, the signals are coupled to the blue chip (457 nm) of RGBYC silicon substrate LED lamp (researched by Nan Chang University) [12] via an AD-DC coupler to drive the LED to emit light to generate an optical signal. The transmission distance of our experiment in underwater is 1.2 meters.

At the receiver, after focusing light by using the lens in front of the receiver, a photodiode (PIN, Hamamatsu 10784) is employed to detect the received light signals. It can convert transmitting optical signals to electrical signals. Then signals are amplified by EA. A digital storage oscilloscope (OSC, Agilent DSO54855A) is used to quantize the received signals for offline processing. In offline processing, signal synchronization and normalization, down-sampling are performed to get PAM7

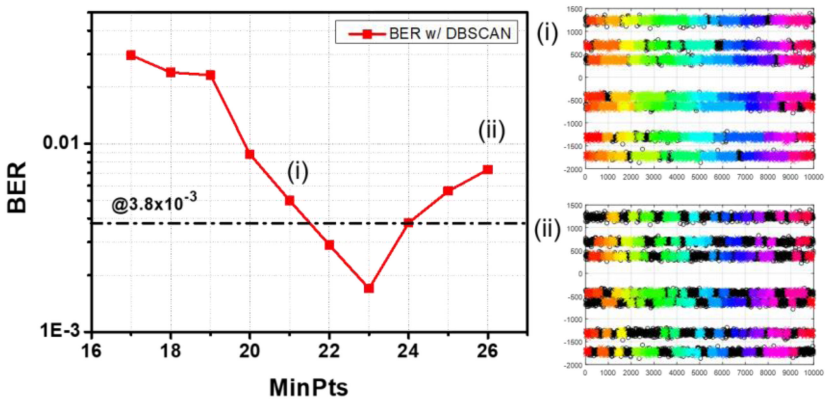


Fig. 10. Experimental BER at a = 0.46 versus different MinPts.

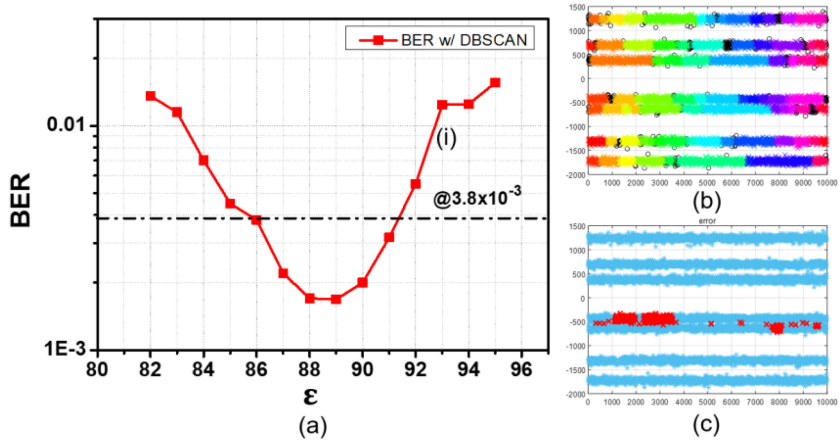


Fig. 11. Experimental (a) BER at a = 0.46 versus different ε, (b) clustering results at $\varepsilon = 93$, (c) decision results at $\varepsilon = 93$.

signals. Differential decoding is applied to eliminate common noise and improve performance. An adaptive post-equalizer based on scalar modified cascading multi-modulus algorithm (S-MCMMA) is utilized to compensate inter-symbol interference. Next, DBSCAN is added to optimize the worse system performance which is caused by the mismatch between two transmitters. The signals are de-mapped and decoded to obtain the original bit sequence. Finally, the system BER performance is calculated.

In the experiment, the voltage of bias-tee is 3.031 V, the current is 130 mA and the peak-to-peak voltage of signals is 0.7 V.

First of all, to know the effect of parameters of the DBSCAN algorithm, Fig. 10 and Fig. 11 are presented. Fig. 10 is the experimental BER at a = 0.46 versus different *MinPts*. As described in the DBSCAN algorithm, *MinPts* is the minimum points in the ε neighborhood area. In Fig. 10, it can be easily found that when the value of *MinPts* is too large or too small, the BER performance with DBSCAN is worse to transmit information. The optimal *MinPts* is 23 here. It can be reasonable that when *MinPts* is too small, DBSCAN will find lots of clusters, and the clustering performance is not good. However, when the *MinPts* is too large, many points may be judged as noise points as drawn in the inset (ii). So the BER performance is bad too. In inset figures, the black points represent noise points. And points with different color are in different clusters.

As described in the DBSCAN algorithm, ε is the minimum neighborhood radius of a given point. Fig. 11(a) shows the experimental BER at a = 0.46 versus different ε. In Fig. 11(a), the area of ε

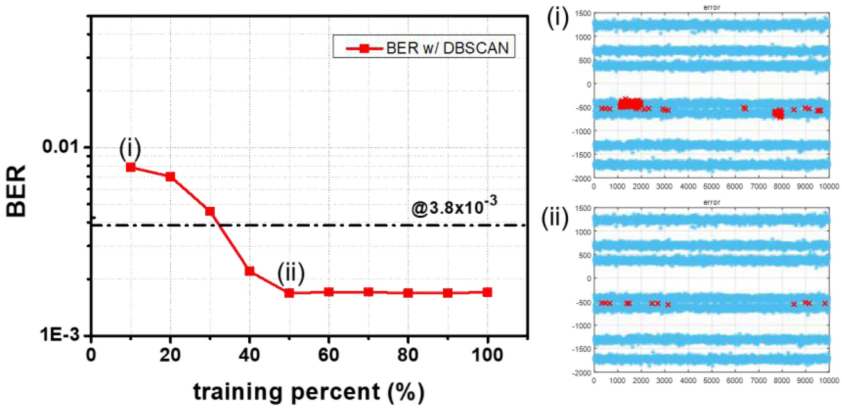


Fig. 12. Experimental BER results versus training percent.

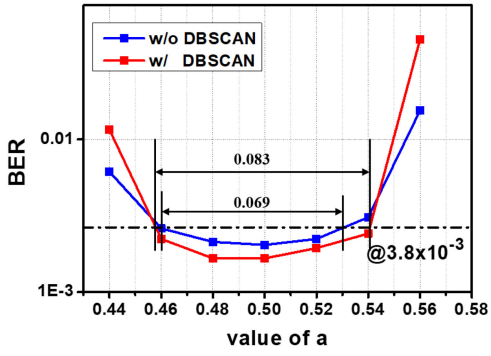


Fig. 13. BER versus the value of a with or without DBSCAN.

under FEC threshold is from 86 to 91. It can be found that the value of ε is too large or too small will produce worse BER performance too. That's because if a smaller ε is selected according to the denser clusters, the density of the nodes in the thinner cluster may be divided into multiple similar clusters. In contrast, if a larger ε is selected according to the thinner clusters, those clusters that are closer and denser will likely be merged into the same cluster, and the difference between them will be ignored. Both of these will lead to worse BER performance. As a result, there is a trade-off between BER performance and the value of ε. In this figure, the optimal value of ε is 88. To figure out how a high ε leads to a high BER, Fig. 11(b) and (c) are drawn. Fig. 11(b) is the clustering results and Fig. 11(c) is the decision results. It shows clearly that the high BER comes from the wrong decision and wrong clustering. Two adjacent levels are adjudged to the same cluster wrongly because of higher ε.

To find out the effect of the training sequence, we change the number of the training sequence. Fig. 12 shows the BER performance versus the percentage of the training sequence. The value of a is 0.45. The parameters of DBSCAN are set at $\varepsilon = 93$ and $MinPts = 23$. The figure shows clearly that when the training percent increases, the BER performance will be improved. And when training percent is larger than 50%, the BER curve tends to be smooth. That's mean that increasing training sequence would not improve BER performance. Considering the complexity of the system, 50% training sequence can be selected. From the inset figure (i), we can find that when 10% training sequence is used, there are lots of wrong judged points, resulting in a high BER.

The experimental BER versus the value of *a* with or without DBSCAN is shown in Fig. 13. From the figure, we can find that DBSCAN can enlarge the area of *a*. It is the same as our simulated results. When increasing the degree of mismatch, DBSCAN will degrade system performance.

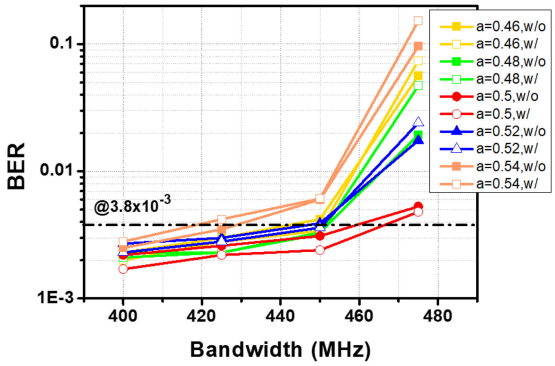


Fig. 14. BER versus bandwidth with or without DBSCAN.

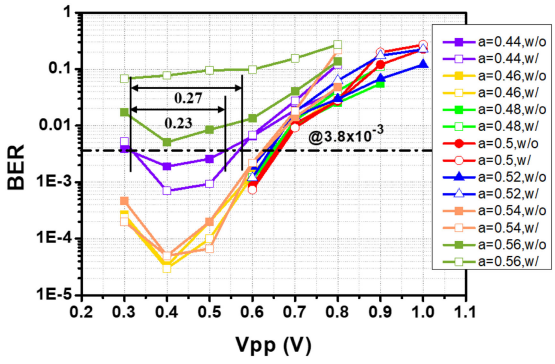


Fig. 15. BER versus Vpp with or without DBSCAN.

The reason is that the number of misjudged level is larger than 1. It is too hard to distinguish each level in received PAM signals. The area of a is increased 20.3% from 0.069 to 0.083 by employing DBSCAN. However, the experimental BER in Fig. 13 has more slightly improvement compared to simulation results in Fig. 8. Although Gaussian white noise is added to the simulation, the simulation result is close to the ideal situation, and the passing channel is not the actual channel. In actual experiments, the channel attenuation and actual noise will be more serious. So the improvement in BER in Fig. 13 is not as large as in the simulation result in Fig. 8.

To know how the BER performance under different *a* is when channel bandwidth increases. The result is presented in Fig. 14. First of all, the BER performance would be worse with bandwidth increases. In all cases, DBSCAN can increase the maximum bandwidth under the FEC threshold. When *a* is 0.5, the maximum bandwidth with DBSCAN is 466.5 MHz. The real bit rate can be calculated by 466.5*3(bit)*(7/8) $=$ 1.22 Gbit/s. It's worth mentioning that the performance at a $=$ 0.46 is not the same as when a $=$ 0.54 (The two cases are symmetric.). Because we add Gaussian white noise to a smaller PAM3 signal to simulates the actual situation of 2 $\times$ 1 MISO system. In real 2 $\times$ 1 MISO system, the transmitter with smaller amplitude is more sensitive to noise.

Fig. 15 shows the measured experimental BER versus peak-to-peak voltage (Vpp) with or without DBSCAN. First, the BER increases with Vpp value when Vpp is larger than 0.4 V. The reason is that high Vpp voltage means signals suffer high nonlinear and will be damaged seriously. It is unable to optimize system performance in that case by utilizing DBSCAN. Second, when the mismatch is not severe, DBSCAN makes system performance can tolerate higher Vpp range with DBSCAN under FEC threshold. For example, when a $=$ 0.44, the range of Vpp under FEC threshold is 0.23 without DBSCAN and 0.27 with DBSCAN. Although the improvement of 0.04 is not large, it is still useful. Because at this time, the deviation between the two transmitters is large enough.

Actually, it is possible to use more Txs to achieve PAM7 or higher level of PAM in theory. However, it will increase the device cost. And when the number of Tx increases, the mismatch among Txs will increase which leads to worse system performance. In that case, DBSCAN may be more necessary to improve the performance. Further explorations on MISO system with more Txs should be carried out in our future study.

5. Conclusions

In this paper, a detailed study of the unbalance between the two transmitters is presented. When the decision of each level in PAM7 signals which is judged by a conventional decision method has no more than one wrong level decision, DBSCAN can improve the system performance in a certain degree. The effect of minPts, ε, training percent, the degree of *a*, bandwidth and Vpp have been discussed in our experimental results. Finally, a transmitting rate of 1.22 Gbit/s over 1.2 m underwater VLC transmission is easily achieved by PAM7 MISO system with DBSCAN technique. From the experimental results, the area of *a* under FEC threshold is increased 20.3% from 0.069 to 0.083 by employing DBSCAN. DBSCAN can improve system performance when the unbalance exists between two transmitters. It shows that DBSCAN is a potential post-equalization technique in the future underwater VLC system.

References

[1] A. Nakajima *et al.*, "ShindaiSat: A visible light communication experimental micro-satellite," in *Proc. Int. Conf. Space Opt. Syst. Appl.*, Ajaccio, Corsica, France, Oct. 9–12, 2012, pp. 12–17.

[2] N. Chi, J. Shi, Y. Zhou, Y. Wang, J. Zhang, and X. Huang, "High speed LED based visible light communication for 5G wireless backhaul," in *Proc. IEEE Photon. Soc. Summer Topical Meeting Ser.*, Newport Beach, CA, USA, 2016, pp. 4–5.

[3] N. Chi, H. Haas, M. Kavehrad, TDC. Little, and X. Huang, "Visible light communications: Demand factors, benefits and opportunities," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 5–7, Apr. 2015.

[4] H. L. Minh *et al.*, "80 Mbit/s visible light communications using pre-equalized white LED," in *Proc. Eur. Conf. Opt. Commun.*, Brussels, Belgium, 2008, pp. 1–2.

[5] C. Kottke, J. Hilt, K. Habel, J. Vučić, and K.-D. Langer, "1.25 Gbit/s visible light WDM link based on DMT modulation of a single RGB LED luminary," in *Proc. Eur. Conf. Opt. Commun.*, Amsterdam, The Netherlands, 2012, paper We.3.B.4.

[6] G. Cossu, A. M. Khalid, P. Choudhury, R. Corsini, and E. Ciaramella, "3.4 Gbit/s visible optical wireless transmission based on RGB LED," *Opt. Express*, vol. 20, no. 26, pp. B501–B506, 2012.

[7] R. Deng, J. He, M. Chen, and Y. Zhou, "Experimental demonstration of a real-time gigabit OFDM-VLC system with a cost-efficient precoding scheme," *Opt. Commun.*, vol. 423, pp. 69–73, Sep. 2018.

[8] Y. Shao, Y. Hong, Z. Hu, and L. Chen, "Capacity maximization of OWC systems via joint precoding and probabilistic shaping," *IEEE Photon. Technol. Lett.*, vol. 31, no. 13, pp. 1013–1016, Jul. 2019.

[9] Q. Chen *et al.*, "Experimental research on adaptive 128/64QAM DFT-spread IFFT/FFT size efficient OFDM with a high SE in VLLC system," *IEEE Photon. J.*, vol. 9, no. 1, Feb. 2017, Art. no. 7900408.

[10] M. Shi, F. Wang, M. Zhang, Z. Wang, and N. Chi, "PAPR reduction of 2.0 Gbit/s DFT-S OFDM modulated visible light communication system based on interleaved sub-banding technique," in *Proc. IEEE Int. Conf. Commun. Workshops*, 2017, pp. 337–342.

[11] Y. Wang, X. Huang, L. Tao, J. Shi, and N. Chi, "4.5-Gb/s RGB-LED based WDM visible light communication system employing CAP modulation and RLS based adaptive equalization," *Opt. Express*, vol. 23, no. 10, pp. 13626–13633, 2015.

[12] Y. Wang, L. Tao, X. Huang, J. Shi, and N. Chi, "8-Gb/s RGBY LED-based WDM VLC system employing high-order CAP modulation and hybrid post equalizer," *IEEE Photon. J.*, vol. 7, no. 6, Dec. 2015, Art. no. 7904507.

[13] N. Chi, M. Zhang, Y. Zhou, and J. Zhao, "3.375-Gb/s RGB-LED based WDM visible light communication system employing PAM-8 modulation with phase shifted Manchester coding," *Opt. Express*, vol. 24, no. 19, pp. 21663–21673, 2016.

[14] H. Chun *et al.*, "LED based wavelength division multiplexed 10 Gb/s visible light communications," *J. Lightw. Technol.*, vol. 34, no. 13, pp. 3047–3052, Jul. 2016.

[15] X. Zhu, F. Wang, M. Shi, N. Chi, J. Liu, and F. Jiang, "10.72 Gbs visible light communication system based on single packaged color mixing LED utilizing QAM-DMT modulation and hybrid equalization," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, 2018, paper M3K.3.

[16] C. Wang, H. Y. Yu, and Y. Zhu, "A long distance underwater visible light communication system with single photon avalanche diode," *IEEE Photon. J.*, vol. 8, no. 5, Oct. 2017, Art. no. 7906311.

[17] M. Doniec, I. Vasilescu, M. Chitre, C. Detweiler, M. Hoffmann-Kuhnt, and D. Rus, "AquaOptical: A lightweight device for high-rate long-range underwater pointto-point communication," in *Proc. IEEE Oceans*, 2010, pp. 1–6.

[18] G. Cossu *et al.*, "Experimental demonstration of high speed underwater visible light communications," in *Proc. Int. Workshop Opt. Wireless Commun. Conf.*, 2014, pp. 11–15.

[19] J. Xu *et al.*, "OFDM-based broadband underwater wireless optical communication system using a compact blue LED," *Opt. Commun.*, vol. 369, pp. 100–105, 2016.
[20] P. Tian *et al.*, "High-speed underwater optical wireless communication using a blue GaN-based micro-LED," *Opt. Express*, vol. 25, no. 2, pp. 1193–1201, 2017.
[21] M. Kong *et al.*, "Underwater wireless optical communication using an arrayed transmitter/receiver and optical superimposition-based PAM-4 signal," *Opt. Express*, vol. 26, no. 3, pp. 3087–3097, 2018.
[22] F. W. Y. Liu, F. Jiang, and N. Chi, "High speed underwater visible light communication system based on LED employing maximum ratio combination with multi-PIN reception," *Opt. Commun.*, vol. 425, pp. 106–112, 2018.
[23] H. M Oubei *et al.*, "Light based underwater wireless communications," *Jpn. J. Appl. Phys.*, vol. 57, no. 8S2, 2018, Art. no. 08PA06.
[24] M. Elamassie, F. Miramirkhani, and M. Uysal, "Channel modeling and performance characterization of underwater visible light communications," in *Proc. IEEE Int. Conf. Commun. Workshops*, 2018, pp. 1–5.
[25] A. Mostafa and L. Lampe, "Physical-layer security for MISO visible light communication channels," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 9, pp. 1806–1818, Sep. 2015.
[26] Y. J. Zhu *et al.*, "An optimal power allocation for multi-LED phase-shifted based MISO VLC systems," *IEEE Photon. Technol. Lett.*, vol. 27, no. 22, pp. 2391–2394, Nov. 2015.
[27] Y. Y. Zhang *et al.*, "Energy-efficient space–time modulation for indoor MISO visible light communications," *Opt. Lett.*, vol. 41, no. 2, pp. 329–332, 2016.
[28] M. Shi *et al.*, "Equiprobable pre-coding PAM7 modulation for nonlinearity mitigation in underwater 2 $\times$ 1 MISO visible light communications," *J. Lightw. Technol.*, vol. 36, no. 22, pp. 5188–5195, Nov. 2018.
[29] Y.-C. Chuang, Z.-Q. Li, C.-W. Hsu, Y. Liu, and C.-Wai Chow, "Visible light communication and positioning using positioning cells and machine learning algorithms," *Opt. Express*, vol. 27, pp. 16377–16383, 2019.
[30] S. Ma *et al.*, "Signal demodulation with machine learning methods for physical layer visible light communications: Prototype platform, open dataset, and algorithms," *IEEE Access,* vol. 7, pp. 30588–30598, 2019.
[31] C. Hsu *et al.*, "Accurate indoor visible light positioning system utilizing machine learning technique with height tolerance," in *Proc. Opt. Fiber Commun. Conf.*, San Diego, CA, USA, 2018, paper M2K.2.
[32] E. Schubert *et al.*, "DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN," *ACM Trans. Database Syst.*, vol. 42, no. 3, pp. 1–21, 2017.