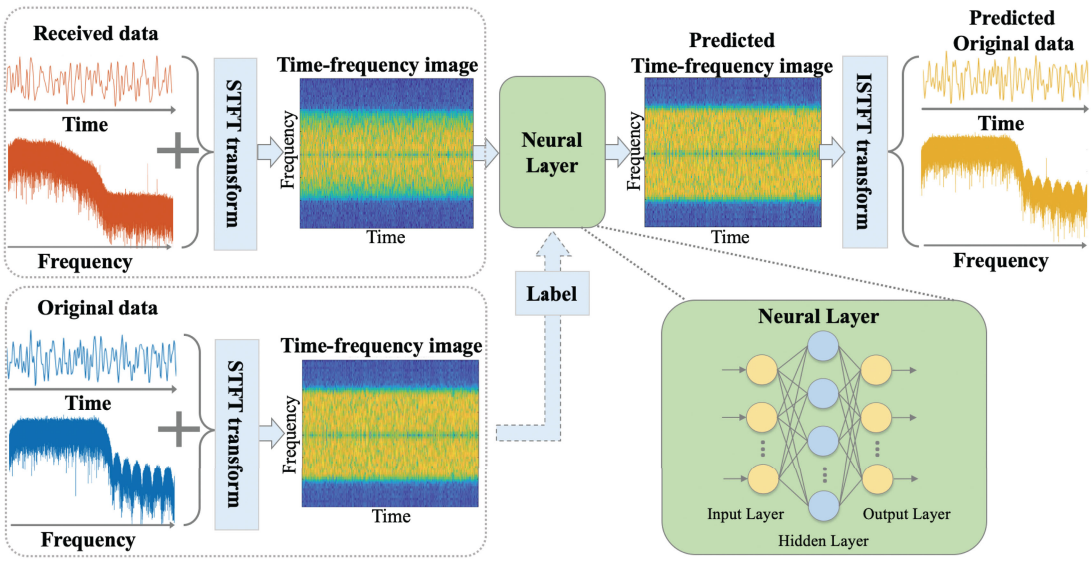
IEEE Photonics Journal
An IEEE Photonics Society Publication

Open Access

Nonlinear Resilient Learning Method Based on Joint Time-Frequency Image Analysis in Underwater Visible Light Communication

Volume 12, Number 2, April 2020

Hui Chen
Yiheng Zhao
Fangchen Hu
Nan Chi



DOI: 10.1109/JPHOT.2020.2981516

Nonlinear Resilient Learning Method Based on Joint Time-Frequency Image Analysis in Underwater Visible Light Communication

Hui Chen, Yiheng Zhao, Fangchen Hu, and Nan Chi

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

DOI:10.1109/JPHOT.2020.2981516

Manuscript received March 11, 2020; accepted March 14, 2020. Date of publication March 17, 2020; date of current version April 13, 2020. This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFB0403603, and in part by the NSFC project under Grant 61925104, Fudan University-CIOMP Joint Fund. Corresponding author: Nan Chi (e-mail: nanchi@fudan.edu.cn).

Abstract: In this paper, we propose a novel nonlinear resilient learning post equalizer named TFDNet in UVLC system. Unlike the traditional deep neural network (DNN) based post equalizers which merely consider the time domain, the proposed TFDNet exploits time-frequency image analysis which considers the time and frequency domains simultaneously and transforms the signal into 2D time-frequency image, which is further learned by neural network. Experimental results demonstrate that TFDNet outperforms Volterra and DNN based methods for compensating nonlinear distortions through a 1.2 m underwater channel using 64 quadrature amplitude modulation-carrierless amplitude modulation (64QAM-CAP). Even under severe nonlinear distortions where Volterra and DNN cannot work, TFDNet retains valid bit error rate (BER) below the 7% forward error correction (FEC) limit of 3.8×10^{-3}. The performance of TFDNet verifies the effectiveness of time-frequency image analysis which has been applied to tackle nonlinear distortions in UVLC system for the first time.

Index Terms: Underwater visible light communication, post equalizer, time-frequency image analysis, nonlinear distortions, neural network.

1. Introduction

Underwater visible light communication (UVLC) has drawn growing attention from all over the world in recent years. It is reported that the blue-green wavelengths (450∼550 nm) of the visible light spectra locate in the low absorption window of seawater [1], which established the theoretical foundation of UVLC system. Since then, lots of solid research works have been achieved by researchers to accelerate the development of UVLC [2]–[4]. Along with the upcoming era of IoT (Internet of Things), the massive data produced by the close-knit interaction of underwater things such as underwater vehicles, human activities, sensor networks, etc., would demand more advanced underwater communication technologies. Comparing with the widely applied underwater acoustic communication and radio frequency communication, UVLC has the unparalleled advantages of

higher speed, lower latency and more confidentiality [5], which has been considered as a promising solution for future underwater communication [6].

Light emitting diodes (LEDs) are widely employed as transmitters in UVLC system with their superiority of low cost, high energy efficiency and wide divergence. [7] reported a world-breaking record of 15.17 Gbps bit-loading-DMT underwater transmission using the special manufactured LED chips, which reveals the great potentials of LED-based underwater high speed communication. However, linear and nonlinear components in UVLC system would distort the signal and degrade system performance. Linear distortions which may come from optical multipath and sampling time offset can be compensated by using classical linear equalization methods such as recursive least square (RLS) [8] and least mean square (LMS) [9]. On the contrary, the factors that cause nonlinear distortions include the light source, the detector, and related circuits, etc. are more difficult to compensate due to high complexities.

To minimize the negative impact of nonlinear distortions in UVLC system, nonlinear post equalization methods have been proposed to tackle the problem. Traditional post equalizers such as decision feedback equalization (DFE) [10] and the widely used Volterra series [11] have limited ability to compensate nonlinear distortions, which are insufficient to further increase the performance of UVLC system. Fortunately, with the rapid development of deep learning, it has been proved that deep neural network (DNN) has strong ability to learn and approximate complex functions when basic mathematics are difficult to analyze or cannot be clearly described [12], which makes it a competitive candidate for a more effective post equalizer. DNN is also called multi-layer perception (MLP), which usually consists one input layer, multiple hidden layers and one output layer. Zhou *et al.* compared different post equalizers and found DNN outperforms other methods in VLC system [13]. Zhao *et al.* proposed a Gaussian-kernel aided DNN based post equalization algorithm [14] and further proposed a dual-branch DNN based post equalizer to improve the system performance [15]. It is obvious that these methods in common are based on DNN model and process the signal in time domain. However, we found in experiment that there is a noticeable disparity between the frequency spectra of the original signal and the signal which is compensated by DNN in time domain. Generally, if the compensated signal learned by DNN is indistinguishable from the original signal, then they are expected to have similar characteristics in frequency domain, yet, which is not tally with the fact in experiment. It is suggested that DNN is inadequate to compensate signal merely in time domain. Therefore, more domain information should be exploited to aid DNN for better nonlinear distortions compensation.

Time-Frequency (TF) analysis is a powerful tool for signal detection and estimation [16]. TF analysis can reveal comprehensive information of non-stationary signals due to the capability of analyzing a signal in the time and frequency domains simultaneously, which can be modeled as a 2D image for further analysis [17]. Thus, TF analysis has flourished in various applications such as audio speech enhancement [18], however, which has not been explored in UVLC as far as we know. Inspired by the fact that TF analysis can transform one-dimension audio speech into two-dimension time-frequency signal for further enhancement, we exploit the advantages of TF image analysis in this paper and propose a novel joint **T**ime-**F**requency post-equalizer based on **D**eep **N**eural N**et**work (TFDNet) and image analysis to compensate the nonlinear distortions in UVLC system. Unlike those post equalizers in the literature which are based on DNN only process the signal in time domain, we resort to short time Fourier transformation (STFT) which is one of the widely used TF analysis tool to combine the time and frequency domains at the same time, and expand the signal from one-dimension (time domain) to two-dimensions (time-frequency domain). Then the 2D time-frequency image is fed into the DNN, which learns the mapping relations of signal both in time domain and frequency domain. The addition of frequency domain could help DNN learn the complementary characteristics of signal better than merely in time domain. Experimental results show that the proposed TFDNet could resist severe nonlinear distortions and achieve the bit error rate (BER) below the 7% forward error correction limit of 3.8×10^{-3} even when Volterra and DNN based post equalizers fail to work, which verifies the effectiveness of the proposed method for UVLC system. To the best of our knowledge, this is the first time that TF image analysis has been used in UVLC system, which could possibly provide new inspirations of introducing more

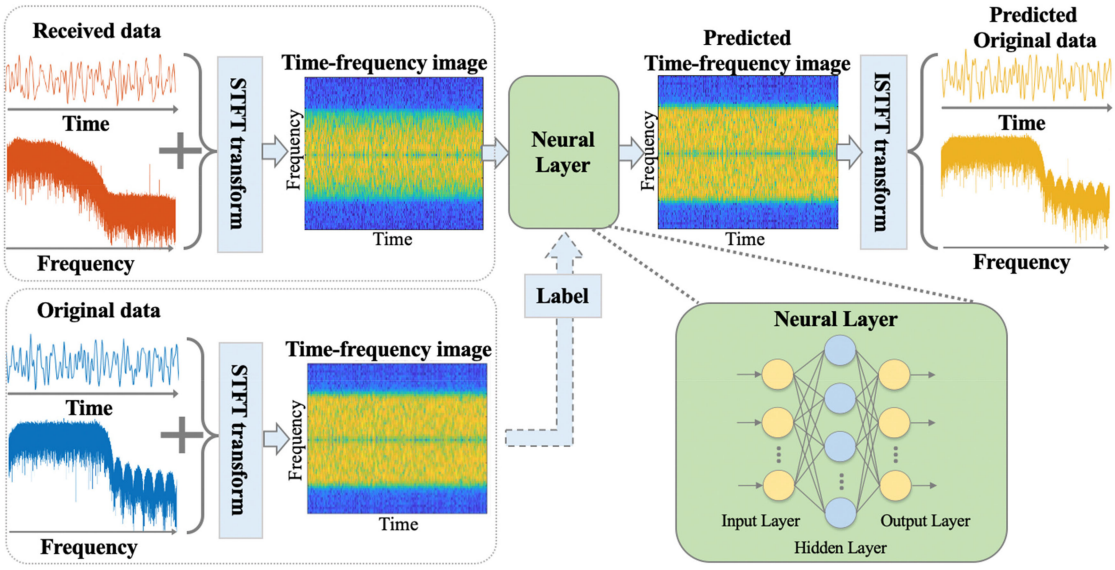


Fig. 1. Schematic diagram of the proposed TFDNet.

deep learning based image processing and analysis methods to tackle the problem of nonlinear distortions in UVLC system.

2. Principle

Given a real-value transmitted signal $x(n)$, the received signal $y(n)$ of UVLC system is modeled as:

$$y(n) = h(x(n)) + v(n) \tag{1}$$

where $h(\cdot)$ denotes the linear and nonlinear distortions and $v(n)$ denotes the additive noise.

The purpose of post equalizers is try to obtain the compensated signal $\hat{x}(n)$ from the distorted signal $y(n)$ and approximate the original signal $x(n)$ as closely as possible. The widely used DNN based post equalizers usually take several time series data from $y(n)$ as the input data of DNN model, which is mainly in consideration of the interference of adjacent symbols for the middle symbol. After the learning process of several hidden layers, DNN outputs the mapped middle symbol to form the estimated $\hat{x}(n)$. Based on DNN model, the proposed TFDNet joints time domain and frequency domain to learn more auxiliary information, which is schematically illustrated in Fig. 1.

The overall pipeline of TFDNet is as following: The received data with distortions $y(n)$ and its frequency spectrum are transformed into time-frequency domain by using STFT transformation, which is expressed as $Y = STFT(y(n))$. The STFT matrix Y could be considered as a 2D image. Then, the image is fed into the multi-layer neural network for the learning process. The original data $x(n)$ and its frequency spectrum are also transformed into time-frequency image which could provide the label information expressed as $X = STFT(x(n))$. After several learning iterations, the predicted time-frequency image $\hat{Y} = L(Y, \Theta)$ is obtained from the output of the well-trained network L with weighted paramters Θ. Finally, the inverse short time Fourier transform (ISTFT) is used to transform the time-frequency image and reconstruct the predicted original data defined as $\hat{x}(n) = ISTFT(\hat{Y})$, which is the goal of the proposed method. The details of the proposed TFDNet is described below.

Firstly, we utilize STFT [19] which is a widely used TF analysis with a time window function to comprehend the signal in this paper. The STFT transformation of the received time series signal $y(n)$ with total length N is calculated by sliding an analysis window of length M over the signal

and then the windowed data is calculated by discrete Fourier transform (DFT). The number of DFT points is D. Then, the window hops over the signal at intervals of R samples repeatedly until the end of the signal. Hamming window is used as the window function. It is noteworthy that in STFT most window functions taper off at the edges to avoid spectral ringing, the windowed segments ought to be overlap-added to compensate the signal attenuation at the window edges with a nonzero overlap length L. Finally, the DFT of each windowed segment is added to the STFT matrix that contains the magnitude and phase for each point in time and frequency, which is expressed as:

$$Y(f) = [Y_1(f), Y_2(f), \ldots, Y_c(f)] \tag{2}$$

The kth element of the STFT matrix is expressed as:

$$Y_k(f) = \sum_{n=-\infty}^{\infty} y(n) g(n - kR) e^{-j2\pi fn} \tag{3}$$

Where $g(n)$ is the window function with length M, R is hop size between adjacent DFTs, $R = M - L$.

The STFT matrix could be considered as a 2D image for the visualization of the time-frequency domain of the signal as shown in Fig. 1. Since the elements in STFT matrix have complex values, we concatenate the real values and the imaginary values in one column of the STFT matrix in sequence to form the corresponding column in the image. Thus, the height of the image is equal to the number of rows in the STFT matrix which is two times of the number of DFT points D, the width of the image equals to the number of columns in the STFT matrix. The height and width of the image are expressed as:

$$height = 2D \; width = \left\lfloor \frac{N - L}{M - L} \right\rfloor \tag{4}$$

where the $\lfloor \cdot \rfloor$ denotes the floor function.

After we obtian the 2D image, we feed the image into the neural network. The neural netwok has multi-nodes input layer, multi-nodes output layer and one hidden layer with the activation function of Rectified linear unit (ReLU). We use every column of the image as the input of the neural network. After propagated though the hidden layer, the network outputs the learned data which have the same dimension as the input data. The parameters optimization strategy follows the Adam algorithm [20] and the loss function of the neural network is mean square error (MSE), which can be given as follows:

$$L(\Theta) = \frac{1}{C} \sum_{c=1}^{C} \left\| \hat{Y}_c - X_c \right\|_2^2 \tag{5}$$

Where C is total number of training data.

Once the training process of neural network is completed, the predicted time-frequency image $\hat{Y}$ is then transformed by ISTFT. It is worth reminding that to ensure successful reconstruction of the original signal when using ISTFT [19], the analysis window must satisfy the COLA constraint [19]. The desired $\hat{x}(n)$ is reconstructed by taking the IFFT of each column of $\hat{Y}$ and overlap-adding the inverted signals. The ISTFT process is calculated as follows:

$$\begin{aligned} \hat{x}(n) &= \int_{-1/2}^{1/2} \sum_{k=-\infty}^{\infty} \hat{Y}_k(f) e^{j2\pi fn} df \\ &= \sum_{k=-\infty}^{\infty} \int_{-1/2}^{1/2} \hat{Y}_k(f) e^{j2\pi fn} df \\ &= \sum_{k=-\infty}^{\infty} \hat{x}_k(n) \end{aligned} \tag{6}$$

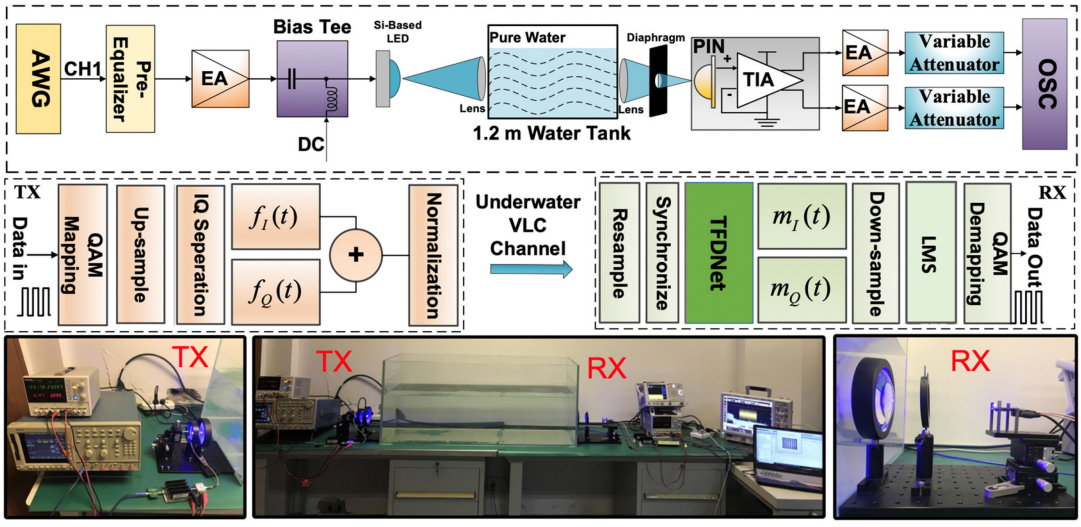


Fig. 2. Experimental setup of UVLC system.

3. Experimental Setup

The experimental setup of UVLC system is shown in Fig. 2. At the transmitter, the original bit sequence is firstly mapped into 64QAM complex symbols. After 4 times of up-sampling, the generated 64QAM signal is then separated by IQ separation for CAP modulation. After normalization, the signal is ready to be transmitted. An AWG (Tektronix AWGG710) is utilized to convert the digital signal into analogy signal. Then the signal passes through a bridged-T amplitude hardware pre-equalizer [21] which could extend the bandwidth of the transmitter, and is amplified by an electrical amplifier (EA, Mini-Circuits, 25 dB gain). Finally, the signal is emitted by silicon substrate blue LED through a 1.2 water tank. It is noticed that static water is used as the transmission channel of LED light due to the existing experimental condition limitation. Other disturbances of underwater channel such as turbulence, scattering, diffusion, etc. would be considered in future experiments. The light source, the receiver, electrical amplifier, etc. are mainly responsible for the nonlinearity of this UVLC system [22]–[24].

At the receiver, a commercial PIN photodiode is used to detect and convert the received optical signal into electrical signal with a differential amplifier circuit. The lens before PIN is used to focus light. Then the two differential output signals are recorded by a digital storage oscilloscope for offline processing. In the offline processing, Volterra, DNN and TFDNet are used as a post equalizer to mainly compensate nonlinear distortions, respectively. The LMS equalization is used to further compensate the linear distortions to improve the overall performance of the system [22], [23]. Then the matching filter is used for CAP demodulation. Finally, the binary data is recovered from QAM demodulation.

4. Results and Analysis

There are several parameters in the proposed TFDNet should be considered carefully, which mainly include the window size with length M of STFT, the number of DFT points D and the node numbers in neural layers. Therefore, in order to choose the appropriate parameters which are suitable for the system, BER versus different parameter values are measured in the UVLC system as shown in Fig. 3. The lower the BER is, the better performance the system is. Window size determines the time-frequency resolution in STFT transform. Whether the window size is too large or too small, the time-frequency resolution would be degraded unequally and decrease the system performance. The choice of the window size mainly depends on the level of inter-symbol-interference (ISI) of

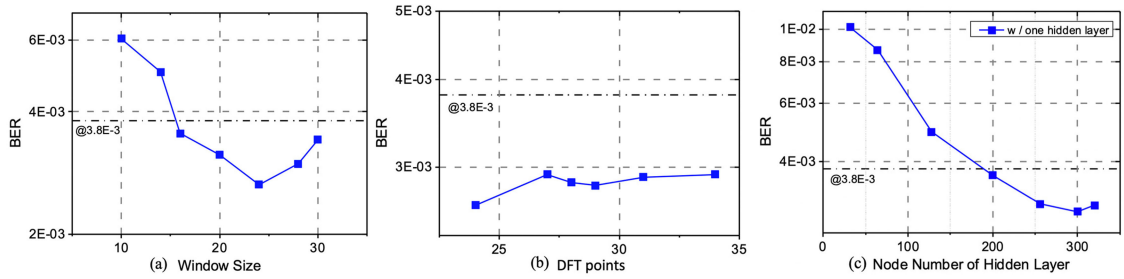


Fig. 3. BER using TFDNet versus different values of parameters (a) window size, (b) DFT points, and (c) nodes of hidden layer.

the system. The window size increases as the ISI increase and could be larger than ISI to gain good equalization effect in principle. However, if the window size is too large, there would be a risk of overfitting. It can be observed from Fig. 3(a) that the BER tends to decrease along with the increasing of window size. When the window size reaches at 24, the lowest BER is achieved. Then with the further increasing of window size, the BER tends to increase which means larger window size is not beneficial for the system performance. Thus, we choose 24 as the window size in the process of STFT transform. The number of DFT points determines the rows of the obtained STFT matrix. Generally, the number of DFT points is no less than the length of window size [19]. As shown in Fig. 3(b), the BER shows no significant fluctuation as the number of DFT points becomes larger. This is because when the window size is fixed at 24, the DFT transform is based on the windowed data no matter what the number of DFT points is. Larger DFT points do not mean it would generate more valuable information. Thus, we choose 24 as the number of DFT points which is the same as the window size. To satisfy the COLA constraint, the overlap length is set at 20. Moreover, the real and imaginary elements of STFT matrix are concatenated to form the time-frequency image as explained above. Therefore, the height of the image is two times of the number of DFT points, which equals to 48. As for the neural network in the proposed TFDNet, we use densely connected neural layers for the learning process. The number of nodes in the input layer is the same as the height of the image. Only one hidden layer is used in the network as to verify the feasible learnability of the proposed construction. BER versus different node number of hidden layer is illustrated in Fig. 3(c). It is shown that the system performance improves with the increasing node number of hidden layer. It can be inferred that with more nodes in hidden layer, the network would performance better. It remains relatively stable when the number is higher than 256. Thus, the node number of the hidden layer is fixed at 256. It is worth noting that more hidden layers could be added in the proposed method to form a deeper neural network. We use only one hidden layer because more hidden layers could bring more complexities. More importantly, the following experiments show the proposed method with only one hidden layer already outperforms DNN with multiple hidden layer, which verifies the effectiveness of the proposed method. The output layer has the same number of nodes as the input layer, which makes the dimensions of the predicted image consistent with the original image.

The optimization investigation of the hyperparameters of DNN in time domain follows the investigation process in [23]. We found in experiment that when the node number of the input layer, first hidden layer, second hidden layer and output layer is set at 33, 128, 64 and 1 respectively, DNN achieves best performance. Thus, DNN with these parameters is tested in the following experiments.

The comparison of BER performance versus different Vpp with different post equalizers is presented in Fig. 4. We compare the proposed TFDNet with the traditional used Volterra and DNN which processes data in time domain. These experiments are conducted at a fixed bias current (150 mA) with different Vpp under two data rate. In Fig. 4(a), we can see that Volterra has shown limited compensation ability, it has a valid operating Vpp range between 0.6 V to 0.8 V which is

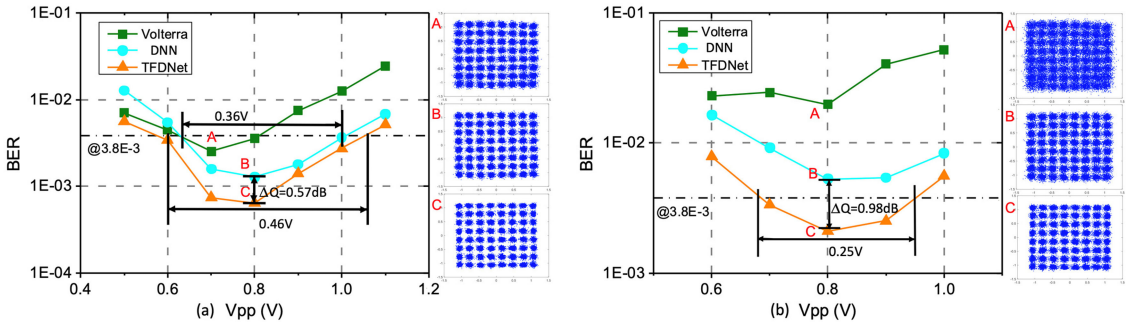


Fig. 4. BER versus different Vpp at data rate of (a) 2.7 Gbps and (b) 2.85 Gbps.

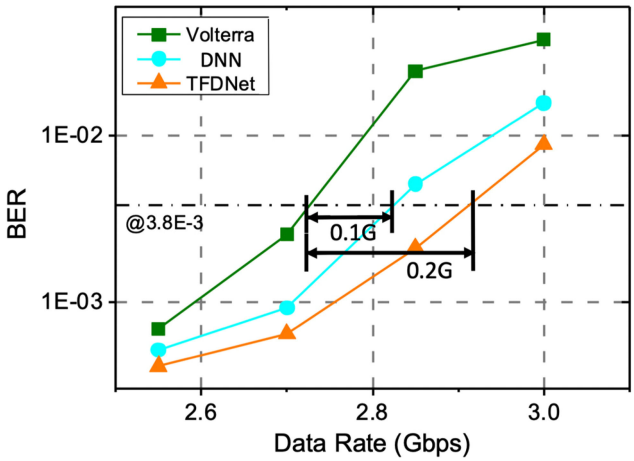


Fig. 5. BER versus different data rate.

below the error threshold of 3.8×10^{-3} and 0.7 Vpp is the optimal Vpp for Volterra. DNN performs better than Volterra by extending the valid Vpp scope from 0.8 Vpp to 1.0 Vpp. It is observed that the BER gap between Volterra and DNN becomes larger along with the increasing of Vpp which indicates higher nonlinear distortions and this is consistent with the conclusion in [13]. TFDNet by using the additional information provided by frequency domain outperforms other equalizers overall and further extends the operating Vpp scope to around 1.1 V. It shows that TFDNet has a flexible operating Vpp range of 0.46 V, which is larger than 0.36 V of DNN. Besides, 0.8 Vpp is the optimal Vpp both for TFDNet and DNN and TFDNet has a higher Q factor than DNN by 0.57 dB. The constellations of point A, B and C which locate in the lowest BER of Volterra (0.7 Vpp), DNN (0.8 Vpp) and TFDNet (0.8 Vpp) in Fig. 4(a) show that TFDNet has a more ordered and clear constellation, which indicates the enhanced performance by the proposed TFDNet.

It is noteworthy that only TFDNet achieves BER under the error threshold in Fig. 4(b). Different from Fig. 4(a), we increase the data rate in Fig. 4(b). It is observed that neither Volterra nor DNN has valid operating Vpp. This is mainly because with the increasing of data rate, more nonlinear distortions are brought into the UVLC system, which goes beyond the compensation ability of Volterra and DNN. TFDNet increases the Q factor by 0.98 dB and has a valid operating Vpp range of 0.25 V which is below the error threshold. It can be inferred that TFDNet has strong compensation ability under severe nonlinear distortions. Thus, TFDNet can be a competitive equalizer in cases where Volterra and DNN have frustrating performance.

We also evaluate the BER performance of different equalizers with different data rate as shown in Fig. 5. In order to illustrate the best performance of these equalizers, the optimal Vpp is set for

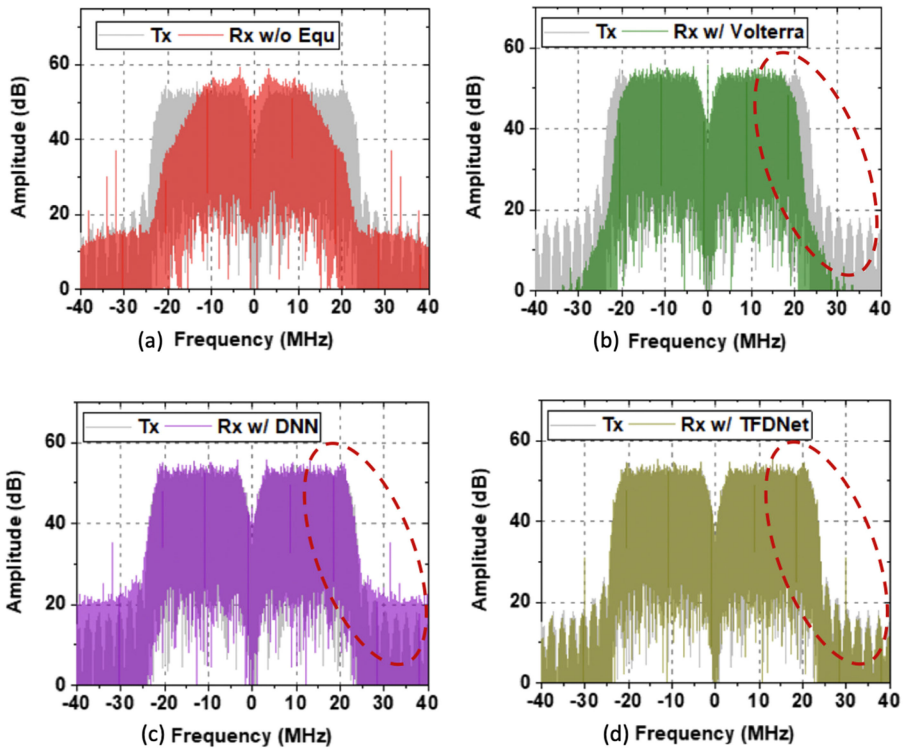


Fig. 6. Frequency spectra of (a) without equalzier, (b) with Volterra, (c) with DNN, and (d) with TFDNet.

Volterra, DNN and TFDNet as 0.7 Vpp, 0.8 Vpp and 0.8 Vpp, respectively. As the increasing of data rate, the BER is also rising. Despite this, TFDNet has lower BER than other equalizers at different data rate. The maximum data rate of TFDNet under the error threshold exceeds 0.1G more than DNN and 0.2 G more than Volterra, which verifies the superior compensation effect of TFDNet.

To Find out why TFDNet outperforms DNN and Volterra equalizers, Fig. 6 may give some hints of the remarkable performance of TFDNet. In Fig. 6, frequency spectra of the original transmitted signal and the received signal with different equalizers are presented. It can be observed that the transmitted signal is severe distorted at the receiver in Fig. 6(a). The received signal compensated by Volterra in Fig. 6(b) has large disparity both in the inner-band part and out-of-band part comparing with the original signal, which may explain the inadequate performance of Volterra. The received signal compensated by DNN in Fig. 6(c) is similar with the inner band of the original signal, however, the out-of-band signal is mistaken by DNN. Fig. 6(d) shows the received signal compensated by TFDNet and it is noticed that the characteristics of both inner-band and out-of-band parts are reconstructed similarly. Table 1 shows the quantitative results of mean absolute error (MAE) of the spectra disparity between the original signal and the received signal compensated by these equalizers. TFDNet has the lowest MAE comparing with others. It can be inferred that the joint of time domain and frequency domain provides TFDNet more insights into the intact features of the signal, which facilitate the learning ability of TFDNet.

It should be pointed out that the complexity of TFDNet is higher than Volterra and DNN, which is mainly caused by the extra computation of STFT and its inverse. The extra computation of TFDNet leads to more time overhead than Volterra and DNN. We conducted the experiments in Intel(R) Core(TM) i7-8700K CPU at 3.7 GHz. The processing time for Volterra, DNN and the proposed method is 0.486 s, 0.463 s, and 0.861 s, respectively. The processing time of TFDNet is higher than Volterra by 77.2% and DNN by 86%. Fortunately, some studies have focused on the complexity reduction of STFT [25], which could simplify the computational process significantly.

TABLE 1
MAE Between the Original Signal and the Received Signal Compensated by Different Equalizers

Equalizers	MAE
Volterra	6.6328
DNN	5.3018
TFDNet	1.7919

Besides, with the rapid developing of deep learning, the complexity of the network in TFDNet would be further reduced by the optimization methods such as pruning [26]. Most importantly, the construction of TFDNet may provide some inspiration about exploiting more image processing and analysis methods which are suitable to tackle problems in UVLC system.

5. Conclusions

In this paper, for the first time we propose to utilize time-frequency image analysis and transform the signal into 2D image to compensate nonlinear distortions in UVLC system. By combining both time and frequency domains simultaneously, the proposed TFDNet could learn more auxiliary information about the signal than DNN which learns signal merely in time domain. The proposed method could be valid for wireless optical communication in a wider perspective. UVLC is one typical representation of wireless optical communication. The experiments in this work illustrate the validation of TFDNet under higher nonlinear conditions, which verifies that the proposed method is qualified for tackling nonlinear distortions in UVLC system. We will consider to explore other 2D signal processing methods such as CNN (Convolutional Neural Networks) or GAN (Generative Adversarial Networks) to further enhance the performance of UVLC system in future work.

References

[1] S. Q. Duntley, "Light in the sea," *JOSA*, vol. 53, no. 2, pp. 214–233, 1963.
[2] N. Chi, H. Haas, M. Kavehrad, T. D. Little, and X.-L. Huang, "Visible light communications: Demand factors, benefits and opportunities [guest editorial]," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 5–7, Apr. 2015.
[3] N. Chi, Y. Zhou, S. Liang, F. Wang, J. Li, and Y. Wang, "Enabling technologies for high-speed visible light communication employing cap modulation," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 510–518, Jan. 2018.
[4] C. Fei *et al.*, "High-speed underwater wireless optical communications: From a perspective of advanced modulation formats," *Chin. Opt. Lett.*, vol. 17, no. 10, 2019, Art. no. 100012.
[5] A. Jovicic, J. Li, and T. Richardson, "Visible light communication: Opportunities, challenges and the path to market," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 26–32, Dec. 2013.
[6] C. Shen *et al.*, "20-meter underwater wireless optical communication link with 1.5 gbps data rate," *Opt. Express*, vol. 24, no. 22, pp. 25502–25509, 2016.
[7] Y. Zhou *et al.*, "Common-anode led on a si substrate for beyond 15 gbit/s underwater visible light communication," *Photon. Res.*, vol. 7, no. 9, pp. 1019–1029, 2019.
[8] Y. Wang, L. Tao, X. Huang, J. Shi, and N. Chi, "8-gb/s rgby led-base d wdm vlc system employing high-order cap modulation and hybrid post equalizer," *IEEE Photon. J.*, vol. 7, no. 6, Dec. 2015, Art. no. 7904507.
[9] Y. Wang, X. Huang, J. Zhang, Y. Wang, and N. Chi, "Enhanced performance of visible light communication employing 512-qam n-sc-fde and dd-lms," *Opt. Express*, vol. 22, no. 13, pp. 15328–15334, 2014.
[10] F.-M. Wu, C.-T. Lin, C.-C. Wei, C.-W. Chen, Z.-Y. Chen, and H.-T. Huang, "3.22-gb/s wdm visible light communication of a single rgb led employing carrier-less amplitude and phase modulation," in *Proc. IEEE Opt. Fiber Commun. Conf. Expo. Nat. Fiber Opt. Engineers Conf.*, 2013, pp. 1–3.
[11] G. Stepniak, J. Siuzdak, and P. Zwierko, "Compensation of a vlc phosphorescent white led nonlinearity by means of volterra dfe," *IEEE Photon. Technol. Lett.*, vol. 25, no. 16, pp. 1597–1600, Aug. 2013.
[12] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, 2016.

[13] Y. Zhou *et al.*, "Comparison of nonlinear equalizers for high-speed visible light communication utilizing silicon substrate phosphorescent white led," *Opt. Express*, vol. 28, no. 2, pp. 2302–2316, 2020.
[14] Y. Zhao, P. Zou, M. Shi, and N. Chi, "Nonlinear predistortion scheme based on gaussian kernel-aided deep neural networks channel estimator for visible light communication system," *Opt. Eng.*, vol. 58, no. 11, Art. no. 116108, 2019.
[15] Y. Zhao, P. Zou, and N. Chi, "3.2 gbps underwater visible light communication system utilizing dual-branch multi-layer perceptron based post-equalizer," *Opt. Commun.*, vol. 460, 2020, Art. no. 125197.
[16] S. Chikkerur, A. N. Cartwright, and V. Govindaraju, "Fingerprint enhancement using stft analysis," *Pattern Recognit.*, vol. 40, no. 1, pp. 198–211, 2007.
[17] K. Grochenig, *Foundations of Time-Frequency Analysis*. Cambridge, MA, USA: Birkhauser, 2001.
[18] T. Afouras, J. S. Chung, and A. Zisserman, "The conversation: Deep audio-visual speech enhancement," in *Proc. Conf. Int. Speech Commun. Assoc.*, 2018.
[19] D. Griffin and J. Lim, "Signal estimation from modified short-time fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 2, pp. 236–243, Apr. 1984.
[20] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, 2015.
[21] X. Huang *et al.*, "2.0-gb/s visible light link based on adaptive bit allocation ofdm of a single phosphorescent white led," *IEEE Photon. J.*, vol. 7, no. 5, Oct. 2015, Art. no. 7904008.
[22] N. Chi and M. Shi, "Advanced modulation formats for underwater visible light communications," *Chin. Opt. Lett.*, vol. 16, no. 12, 2018, Art. no. 120603.
[23] N. Chi, Y. Zhao, M. Shi, P. Zou, and X. Lu, "Gaussian kernel-aided deep neural network equalizer utilized in underwater pam8 visible light communication system," *Opt. Express*, vol. 26, no. 20, pp. 26700–26712, 2018.
[24] C. Fei, J. Zhang, G. Zhang, Y. Wu, X. Hong, and S. He, "Demonstration of 15-M 7.33-Gb/s 450-nm underwater wireless optical discrete multitone transmission using post nonlinear equalization," *J. Lightw. Technol.*, vol. 36, no. 3, pp. 728–734, Feb. 2017.
[25] B. Kim, S.-H. Kong, and S. Kim, "Low computational enhancement of stft-based parameter estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 8, pp. 1610–1619, Dec. 2015.
[26] P. Molchanov, A. Mallya, S. Tyree, I. Frosio, and J. Kautz, "Importance estimation for neural network pruning," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2019, pp. 11264–11272.