100Gbps IM/DD Transmission over 25km SSMF using 20G-class DML and PIN Enabled by Machine Learning

Peixuan Li, Lilin Yi*, Lei Xue, Weisheng Hu

State Key Lab of Advanced Optical Communication Systems and Networks, Shanghai Jiao Tong University, Shanghai 200240, China,
*lilinyi@sjtu.edu.cn

Abstract: We experimentally demonstrate 100Gb/s/λ IM/DD transmission over 25km SSMF using 20G-class DML and PIN enabled by convolutional neural network based equalization technique. Transmission performance in C-band and O-band using PAM-8 and PAM-16 are compared.

OCIS codes: (060.2330) Fiber optics communications; (060.2360) Fiber optics links and subsystems; (060.4510) Optical communications

1. Introduction

In these years, fast development of Internet services such as cloud computing and video streaming is driving up the demand for high speed optical inter-datacenter interconnection, which requires high-speed low-cost intensity modulation and direct detection (IM/DD) transmission system. To meet this requirement, advanced modulation formats like pulse amplitude modulation (PAM), carrier-less amplitude and phase modulation (CAP) and discrete multi-tone (DMT) are proposed to efficiently use limited bandwidth. Due to its simple nature, PAM has gained popularity among the IM/DD transmission community. For high-speed transmission with low-bandwidth optical devices, PAM-8 and PAM-16 seems to be a reasonable choice, which have been realized in various previous work [1-6]. To truly meet the requirement of low cost, instead of using high-bandwidth external modulation like external modulated laser (EML) or Mach Zehnder modulator (MZM), low-cost low-bandwidth directly modulated laser (DML) is highly desired. However, due to the severe inter-symbol interference (ISI) caused by the limited bandwidth, strong chirp and nonlinearity of the optical devices, traditional equalization techniques are not satisfactory in this condition, such as feed-forward equalization (FFE) and volterra nonlinear equalization (VNE). Machine learning has been popular these years and shown its powerful capability in the domain of computer vision and machine translation. Some related previous work has tried feed-forward fully-connected neural network (FCN) [5,6] for inter-datacenter interconnection, which is a machine learning algorithm but providing limited performance. As a powerful tool of machine learning in computer vision and pattern recognition, convolutional neural network (CNN) helps AlphaGo defeat various professional Go players [7] and also shows good performance on optical performance monitoring [8].

In this paper, we experimentally demonstrate 100Gbps PAM-8/PAM-16 IM/DD transmission over 25km standard single-mode fiber (SSMF) using 20G-class DML and P-I-N photodiode (PIN) in C-band & O-band with CNN based post-equalization. The total 3-dB end-to-end bandwidth are 13.3 GHz for C-band devices and 16.8 GHz for O-band. The proposed CNN based post-equalizer has an architecture similar to the most classic hand-written digit recognition CNN, LeNet5 [9] and shows its equalization capability upon 100Gbps PAM-8/PAM-16 signal.

2. Principle of CNN based post-equalizer

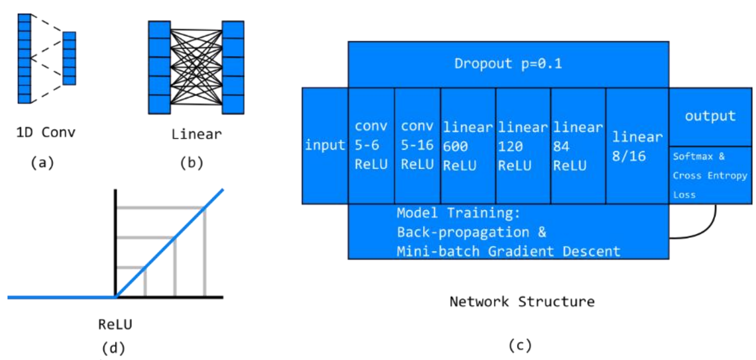

Fig. 1 Structure of (a) 1D convolutional layer, (b) linear layer and (c) proposed CNN-based equalizer. (d) Graph of rectified linear unit (ReLU)

CNN has proven its success in pattern recognition due to its local feature extraction capability, which motivates us from the similarity between images and received sequences. After all, in both domains we want to recover some pattern from the noisy, distorted signal. The most important building block of CNN is the convolutional layer, which can be designed to be one-dimensional (1D), 2D and 3D but the underlying principle are same. In a convolutional layer, there are many output channels, each of which has a sliding window doing cross-correlation with the input array. Figure 1(a) shows the structure of a channel of a 1D convolutional layer. Another important building block is the linear layer, in which input is multiplied by a weight matrix to form the output, as shown in Fig. 1 (b). In the proposed CNN-based equalizer, we setup two consecutive 1D convolutional layers with a sliding window of five and output channel of 6 and 16. The input of the first convolutional layer is 101. After the convolutional layers, the signal is unrolled to be a column vector and fed into four consecutive linear layers with number of output be 600, 128, 84 and 8 or 16. The 8 or 16 output units denotes the levels of PAM-8 and PAM-16, with the highest to be decision symbol. In the middle of convolutional and linear layers, the activation layer of rectified linear unit is applied to provide nonlinearity while avoiding the training problem of gradient vanishing, which is shown in Fig. 1(d). The whole network structure can be found in Fig. 1(c).

$$CrossEntropy = -\sum_{i=0}^{n} target_i \times log(softmax(output_i)), Softmax(x) = \frac{e^x}{\sum_{i=0}^{n} e^{x^i}} \tag{1}$$

The loss function of the network is the cross-entropy loss, which is shown in Eq. (1). The cross-entropy loss is computed combining a special activation function called Softmax, which converts the output unit to its probability approximation. In fact, the concept of cross entropy comes from information theory. With the minimization of cross entropy, the probability distribution of output units come close to the target distribution, thus the network output gradually approaching the desired output and the error rate becomes lower and lower while training. The network is trained with back-propagation algorithm and mini-batch gradient descent. Using back-propagation algorithm, the gradient of the whole network is gained supporting the gradient descent optimization rule. To improve training efficiency, mini-batch gradient descent is applied with batch size of 512, meaning simultaneously 512 samples are sent for training, taking advantage of the parallel computing capability of an Nvidia video card. To avoid over-fitting, dropout mechanism is applied to the network while training, which randomly drops the hidden units out from training by setting them to zero. The probability of dropout is set to 0.1 after tuning. The whole dataset contains ten independent $2^{15}-1$ pseudorandom-binary-sequences (PRBS15) with six of them forming the train dataset of around 200k symbols, and two for cross validation and test dataset contain each. The train dataset is used to train the model, while the model is tuned upon the train and cross validation dataset. The final BER result is calculated based upon the test dataset in order to keep the generalization of the model. Once the training is completed, the optimized parameters of CNN can be stored and only forward calculation is required for real-time error counting.

3. Experimental setup & results

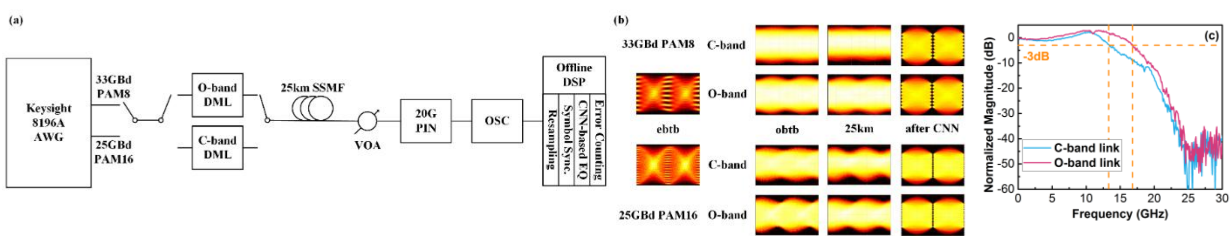

Fig. 2 (a) Experimental setup (b) Eye diagram (c) Frequency Response of C-band & O-band links

Figure 2(a) shows the experimental setup of the 100 Gbps IM/DD transmission system. A Keysight 8196A arbitrary waveform generator (AWG) is used to generate 33 GBd PAM-8 and 25 GBd PAM-16 signal at transmitter side. The signal is then modulated to optical domain with C-band / O-band DML. After 25km SSMF transmission, the optical signal is attenuated by a variable optical attenuator (VOA) to control the optical power into a 20G-class PIN. After PIN, the signal is digitized and sampled by a LeCroy digital sampling oscilloscope (DSO) with 30-GHz bandwidth and 80-GSa/s sample rate. The end-to-end 3-dB channel bandwidth is 13.3 GHz for the C-band link and 16.8 GHz for O-band link, as shown in Fig. 2(c). The sampled signal is then offline-processed using MATLAB and Python. The offline digital signal processing (DSP) flow chart can also be found in Fig. 2(a). First the digitized signal is resampled to one sample per symbol, which enables real-time implementation with 33 GS/s analog-digital converter (ADC). Then the transmitted PRBS-15 sequence is extracted and fed into T-spaced CNN-based equalizer. After the training and test process described in previous section, errors are counted and bit error rate is calculated upon the output of the equalizer.

The eye diagrams of 33 GBd PAM-8 can be found in the upper half of Fig. 2(b), and 25 GBd PAM-16 in the lower half. The electronic back-to-back (BtB) signals are quite good, but after electrical-optical (E-O) and optical-electrical

(O-E) conversion, the eyes are completely closed due to the severe ISI caused by the limited channel bandwidth and the nonlinearity sensitivity of PAM-8 and PAM-16 signals. After the CNN-based equalization performing both equalization and demapping, the eyes are open again.

We compare the BER performance of different modulation formats and wavebands for 25km SSMF and optical BtB transmission, as shown in Fig. (3). With the help of proposed CNN-based equalization, all the transmission configuration achieves BER under 7% hard-decision forward-error-correction (HD-FEC) limit of 3.8e-3. The BER performance is similar for both 25km and optical BtB transmission, with receiver sensitivity ranging from -6.5 dBm to -5.5 dBm. As clearly shown in Fig. 3, the result of O-band transmission outperforms C-band due to the higher channel bandwidth of O-band link. The dispersion in C-band nearly does not bring penalty due to the strong linear equalization capability of CNN, which implies longer transmission distance can be supported. What's more, even with higher baud rate, the BER result of 33 GBd PAM-8 is better than 25 GBd PAM-16, which might come from the higher optical devices nonlinearity and noise sensitivity of the PAM-16 signal.

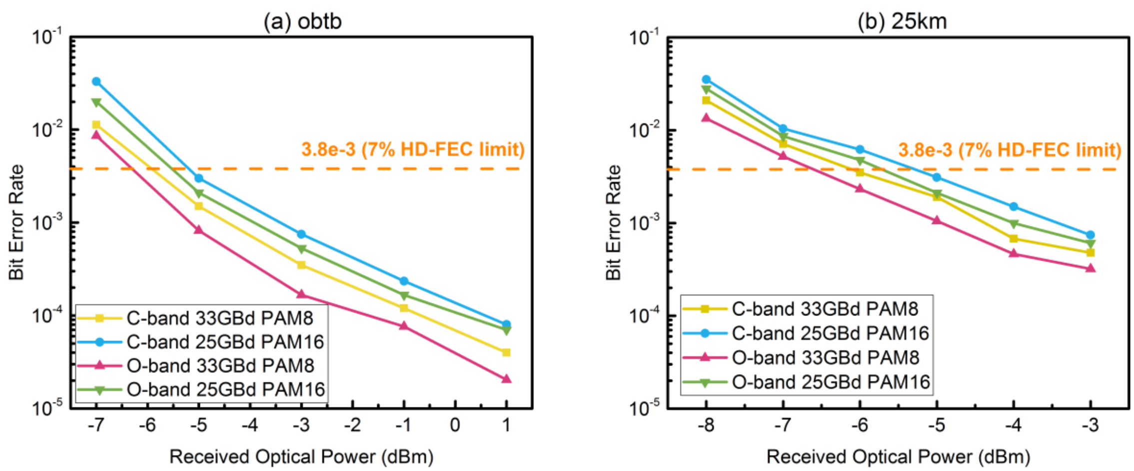

Fig. 3 BER performance comparison of different modulation formats and wavebands for (a) optical back-to-back and (b) 25km-SSMF transmission.

4. Conclusions

We experimentally demonstrate a 100Gbps IM/DD data link over 25km SSMF comparing PAM-8 / PAM-16 and C-band / O-band with the help of proposed CNN-based equalization. To achieve the goal of low cost, DML and PIN are used in the experiment as the E-O and O-E conversion optics with 13.3 GHz end-to-end 3 dB channel bandwidth for C-band link and 16.8 GHz for O-band link. Receiver sensitivity ranging from -6.5 dBm to -5.5 dBm is achieved. By comparing different modulation formats and wavebands, we found 33 GBd PAM-8 outperforms the 25 GBd PAM-16 due to its resistance to device nonlinearity and system noise and CNN shows strong equalization capability on dispersion and bandwidth, which could be worthy of further research.

5. References

[1] J. Zhang, J. Yu, and H. Chien, "EML-based IM/DD 400G (4x112.5-Gbit/s) PAM-4 over 80 km SSMF Based on Linear Pre-Equalization and Nonlinear LUT Pre-Distortion for Inter-DCI Applications," in *Proc. OFC*, paper W4I.4, 2017.

[2] Y. Gao, J. C. Cartledge, S. S. H. Yam, A. Rezania and Y. Matsui, "112 Gb/s PAM-4 Using a Directly Modulated Laser with Linear Pre-Compensation and Nonlinear Post-Compensation," in *Proc. ECOC*, paper M.2.C.2, 2016.

[3] N. Kaneda, J. Lee and Y. K. Chen, "Nonlinear equalizer for 112-Gb/s SSB-PAM4 in 80-km dispersion uncompensated link," in *Proc. OFC*, paper Tu2D.5, 2017.

[4] S. van der Heide , A. Albores-Mejia , F. Gomez-Agis , B. Docter and C. Okonkwo, "112-Gbit/s Single Side-Band PAM-4 Transmission over Inter-DCI distances Without DCF Enabled by Low-complexity DSP," in *Proc. ECOC*, paper P2.SC5.8, 2017

[5] S. Gaiarin, X. Pang, O. Ozolins, R. T. Jones, E. P. da Silva, R. Schatz, U. Westergren, S. Popov, G. Jacobsen, and D. Zibar, "High Speed PAM-8 Optical Interconnects with Digital Equalization based on Neural Network," in *Proc. ACP*, paper AS1C.1, 2016

[6] J. Estaran, R. Rios-Mueller, M. A. Mestre, F. Jorge, H. Mardoyan, A. Konczykowska, J.-Y. Dupuy and S. Bigo, "Artificial Neural Networks for Linear and Non-Linear Impairment Mitigation in High-Baudrate IM/DD Systems," in *Proc. ECOC*, paper M.2.B.2, 2016.

[7] David Silver et al., "Mastering the Game of o with Deep Neural Networks and Tree Search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[8] D. Wang and M. Zhang, "Convolutional Neural Network-based Deep Learning for Intelligent OSNR Estimation on Eye Diagrams" in *Proc. ECOC*, paper P1.SC3.42, 2017.

[9] Yann Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," in *Proc. IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.