Long Short-Term Memory Neural Network to Enhance the Data Rate and Performance for Rolling Shutter Camera Based Visible Light Communication (VLC)

Ching-Wei Peng[1,2], Deng-Cheng Tsai[1,2], Yun-Shen Lin[1,2], Chi-Wai Chow[1,2*], Yang Liu[3] and Chien-Hung Yeh[4]
[1]Department of Photonics, College of Electrical and Computer Engineering, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan
[2]Department of Photonics, College of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan
[3]Philips Electronics Ltd., N.T., Hong Kong
[4]Department of Photonics, Feng Chia University, Seatwen, Taichung 40724, Taiwan
*cwchow@faculty.nctu.edu.tw

Abstract: We propose and demonstrate using Long-Short-Term-Memory neural-network (LSTM-NN) to mitigate inter-symbol-interference (ISI) in 4-level pulse-amplitude-modulation (PAM4) camera based visible-light-communication (VLC) system. Data-rate of 14.4-kbit/s with 3-m free-space transmission is achieved. © 2022 The Author(s)
OCIS codes: (060.2605) Free-space optical communication; (060.4510) Optical communications

1. Introduction

Optical wireless communication (OWC) and visible light communication (VLC) have evolved significantly recently and is now employed in many applications, such as internet-of-things (IOT), free-space terrestrial communications, undersea communications, etc. The camera or image sensor based OWC, which is also known as optical camera communication (OCC) [1, 2] has attracted particular interest recently. OCC has the advantage that it does not need any hardware modification on the receiver (Rx). Mobile phones, vehicle cameras and surveillance cameras can be employed for the OCC systems. To increase the OCC data rate while using low frame rate cameras (i.e. 30 or 60 frame-per-second, fps), rolling shutter mode in the complementary-metal-oxide-semiconductor (CMOS) camera can be used [1, 2]. In the rolling shutter mode, CMOS camera does not receive light at the same time. Instead, different pixel-rows in the camera are activated sequentially depending on the row-by-row exposure time. When the optical Tx is modulated faster than the frame rate of the camera but slower than the row-by-row exposure time, dark and bright fringes representing optical signal "OFF" and "ON" will be seen in each frame. Although the rolling shutter mode can increase the OCC data rate higher than the frame rate, the row-by-row exposure delay will create decoding challenge. When the optical Tx is operated at high modulation speed, only a few pixel-rows in the CMOS image sensor can be used to stand for a logic bit (i.e. low pixel-row per bit). Hence, high inter-symbol interference (ISI) will be observed. Besides using low pixel-row per bit, multi-level modulation can also be used to future increase the OCC data rate. However, it is very challenging to identify the multi-level in the rolling shutter pattern due to the uneven exposure, which is known as "blooming" effect [2]. Different rolling shutter decoding techniques have been proposed, such as using adaptive Extreme-Value-Averaging (EVA) [3]; however, the transmission distance is limited to 30 cm. Recently, a 2-D convolutional neural network (CNN) was reported, and 111 kbit/s with 0.4 m transmission was obtained utilizing red-green-blue (RGB) light emitting diodes (LEDs) [4]. This means a high bit-rate distance product of 14.8 kbit/s • m per color was demonstrated [4]. Z-score averaging neural network (Z-NN) was reported to enhance the rolling shutter decoding [5]. Inspired by this idea, an OCC system using 4-level pulse amplitude modulation (PAM4) supporting a bit-rate distance product of 28.8 kbit/s • m per color was reported [6].

In this work, we propose and demonstrate the Long Short-Term Memory neural network (LSTM-NN) for mitigating the transmission impairments of PAM4 rolling shutter camera based VLC (i.e. OCC) system. The LSTM-NN is a special type of recurrent neural network (RNN) that has memory cells for temporal processing [7]. Unlike the typical artificial neural network (ANN), LSTM-NN has feedback and memory operations. We discover that the proposed LSTM-NN rolling shutter decoding scheme is very effective to decode the multi-level pattern with very low pixel-row per bit. Experiment results show that a record 43.2 kbit/s • m per color can be achieved for the first time.

2. LSTM-NN Algorithm and Experiment

Fig. 1(a) shows the experiment of using LSTM-NN rolling shutter pattern decoding PAM4 OCC system. The Tx is a LED light panel (Li-Cheng Corp.®) with output power of 21.6 W and dimensions of 88 cm × 58 cm. The OCC signal is detected by a CMOS camera with resolution of 1920 × 1080 pixels after 1-3 m free-space transmission. The PAM4 signal has a dark 8-bit header and a variable bit-length payload. Fig. 1(b) illustrates the proposed decoding algorithm for the PAM4 rolling shutter pattern. First, the image is read-in, then it is converted to grayscale values from 0 to 255, representing dark and bright fringes respectively. The grayscale values in each pixel-row are arranged in accessing order; and a vertical column matrix with the highest grayscale value is chosen to produce the grayscale waveform pattern. After this, the pattern is sent to data pre-processing module (i.e. orange boxes in Fig. 1(b)). Padding payloads to the same length and re-sampling by the FIR filter are performed to make sure all the input patterns has the same time dimension. The LSTM-NN is modified from the RNN that possesses memory cells for handling long-term temporal dependency; hence, the features of present symbol value, symbol relationship and symbol average should be extracted acting as the input features to the LSTM-NN as shown in Fig. 1(b).

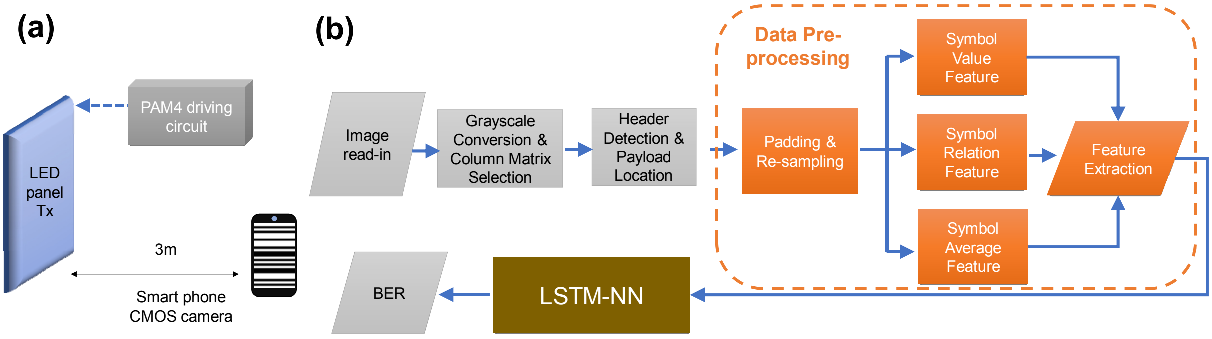

Fig. 1. (a) Experiment of using LSTM-NN for signal decoding in OCC system. (b) Proposed decoding algorithm of the PAM4 rolling shutter pattern.

Figs. 2(a) and (b) show the proposed LSTM-NN in flow diagram and simplified topology. It has an input layer with 172, 106, 56 features depending on the distances (i.e. 1, 2 and 3 m) respectively, 4 hidden layers, and an output layer. The first 2 hidden layers are the LSTM layers with LSTM cell. There are 128 and 64 neurons in this two hidden layers respectively. The other 3 layers are the fully-connected (FC)/ fully-connected (FC)/Softmax. 64 and 16 neurons are used in the two FC layers. The last layer for classification. The signal at the output classification layer is divided into four probabilities of 00, 01, 10, 11 in the PAM4 signal via the Softmax layer. After the LSTM-NN operation, bit-error ratio (BER) measurement based on Tx and Rx bit-by-bit comparison is performed. Fig. 2(c) shows the architecture of the "Symbol Average Feature" process in the data preprocessing module. It consists of an n^{th} order delay on the basis of the number of samples per symbol. The delay time T between taps is a sampling period. The outputs of the delay sequence subtracting the input signal is performed in order to obtain the difference characteristic between two sequences. The features, including averaging of one symbol to accelerate convergence and enhance the model accuracy, are converted in the parallel-to-serial (P/S) block. Fig. 2(d) shows the architecture of the LSTM cell used in the LSTM-NN. It contains different nonlinear activation functions and point-wise multiplication processes. C_{t-1}, C_t, x_t, σ, h_{t-1},h_t, are the memory from last LSTM unit, newly updated memory, current input, Sigmoid operation, output of last LSTM unit and current output respectively. The LSTM cell contains internal gates, which can store or erase data of the cell state.

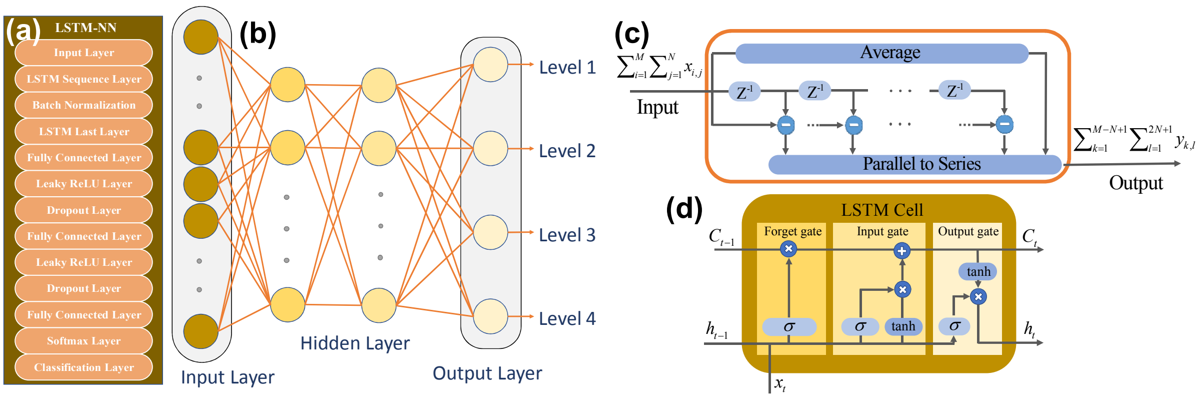

Fig. 2. Proposed LSTM-NN in (a) flow diagram and (b) simplified topology. Architectures of (c) "Symbol Average Feature" process in the data preprocessing module and (d) LSTM cell.

3. Results and Discussion

Figs. 3 (a)-(c) show the original electrical PAM4 pattern for the Tx, decoding patterns by the traditional ANN and the LSTM-NN respectively. The data rate of 6 kbit/s is used. At around 60^{th} pixel-row, bit error can be observed when the OCC decoding is based on ANN. When the decoding is based on the proposed LSTM-NN, the correct data can be retrieved as indicated in Fig. 3(c). Figs. 4(a) and (b) show the measured BER curves at different free-space transmission distances when using the ANN and the proposed LSTM-NN respectively. The OCC transmission distances are from 1 to 3 m. The measured illuminance are ~220 lux and ~100 lux at the distances of 2 and 3 m respectively away from the LED panel. We can observe that the ANN can decode OCC signal at 6 kbit/s with 2 m transmission, fulfilling the pre-forward error correction (pre-FEC BER = 3.8×10^{-3}) threshold. At 3 m transmission distance, the data rate should be reduced to 3.6 kbit/s. When utilizing the proposed LSTM-NN, BER fulfilling the pre-FEC threshold can be observed even at a high data rate of 14.4 kbit/s with 3 m transmission as presented in Fig. 4(b). The experimental results illustrate that the ANN cannot efficiently decode the multi-level rolling shutter pattern with high ISI, while the proposed LSTM-NN can decode the multi-level pattern even with high ISI. We also evaluate the lowest achievable pixel-row per bit using different rolling shutter decoding schemes in Fig. 4(c). In ANN decoding scheme, the lowest pixel-row per bit is 7. In the LSTM-NN decoding, a lower pixel-row per bit of 5 is needed. Hence, the LSTM-NN decoding can support a higher OCC data rate.

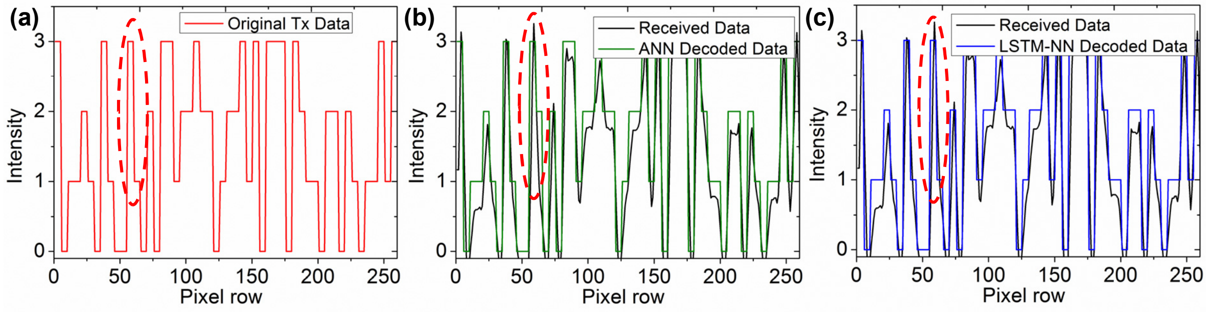

Fig. 3. PAM4 rolling shutter pattern (a) driving the Tx and decoded patterns using (b) ANN and (c) proposed LSTM-NN.

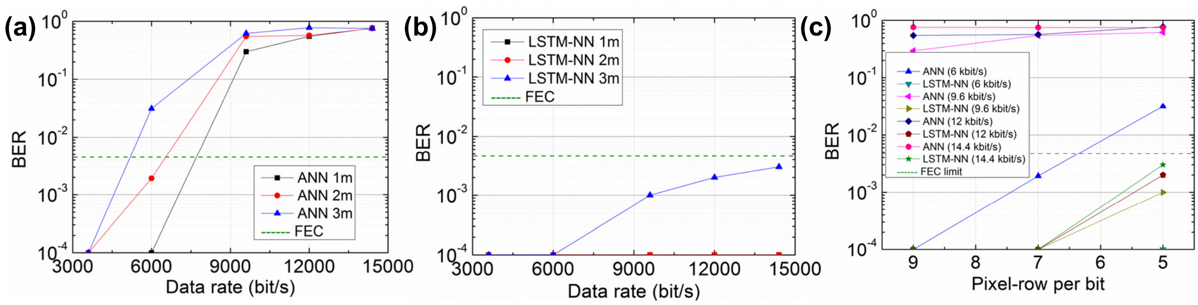

Fig. 4. Measured BERs at different transmissions using (a) ANN and (b) the proposed LSTM-NN. (c) BER of different schemes at different pixel-row per bit.

4. Conclusion

We proposed and demonstrated the LSTM-NN for mitigating the transmission impairments of PAM4 rolling shutter OCC system. The LSTM-NN is a special type of RNN having memory cells for temporal processing. The proposed LSTM-NN rolling shutter decoding scheme is very effective to decode the multi-level rolling shutter pattern with very low pixel-row per bit. Experiment results show that a record 43.2 kbit/s • m per color is achieved.

Acknowledgment This work was supported by the Ministry of Science and Technology, Taiwan, MOST-110-2221-E-A49-057-MY3, MOST-109-2221-E-009-155-MY3 and Ministry of Education (MOE) in Taiwan.

5. References

[1] C. Danakis, et al, "Using a CMOS camera sensor for visible light communication", Proc. IEEE Globecom Workshops 2012, pp. 1244-1248.
[2] C. W. Chow, et al, “Visible light communication using mobile-phone camera with data rate higher than frame rate,” Opt. Exp. **23**, 26080-26085 (2015).
[3] C. W. Chen, et al, “Efficient demodulation scheme for rolling-shutter-patterning of CMOS image sensor based visible light communications,” Opt. Exp. **25**, 24362-24367 (2017).
[4] L. Liu, et al, "Beyond 100-kbit/s transmission over rolling shutter camera-based VLC enabled by color and spatial multiplexing," Proc. OFC 2020, Paper M1J.4.
[5] Y. S. Lin, et al., "Z-score averaging neural network and background content removal for high performance rolling shutter based optical camera communication (OCC)," Proc. OFC, 2021, Paper F1A.4.
[6] Y. S. Lin, et al, "PAM4 rolling-shutter demodulation using a pixel-per-symbol labeling neural network for optical camera communications," Opt. Exp. **29**, 31680-31688 (2021).
[7] S. Hochreiter et al, "Long short-term memory," Neural Computation **9**, 1735–1780 (1997).