Water-to-Air PAM4 Optical Camera Communication Using Long Short Term Memory Neural Network (LSTM-NN)

Yun-Han Chang[1], Shang-Yen Tsai[1], Ming-Chieh Tsai[1], Jia-Fu Li[1], Yin-He Jian[1], Chi-Wai Chow[1*] and Chien-Hung Yeh[2]

[1]Department of Photonics, College of Electrical and Computer Engineering, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan

[2]Department of Photonics, Feng Chia University, Seatwen, Taichung 40724, Taiwan

*cwchow@nycu.edu.tw

Abstract: We demonstrate a wide field-of-view (FOV) water-to-air transmission using rolling-shutter-based optical-camera-communication (OCC). Long-short-term-memory-neural-network (LSTM-NN) is utilized to mitigate the wavy water-surface induced link outage and to decode 4-level pulse-amplitude-modulation (PAM4) rolling-shutter pattern. © 2024 Author(s)

OCIS codes: (060.2605) Free-space optical communication; (060.4510) Optical communications

1. Introduction

Underwater activities are becoming more and more important. Nowadays, autonomous underwater vehicles (AUVs) are used for deep-sea surveillance, exploration, monitoring and rescue, etc. Traditionally, these AUVs utilize cable or acoustic wave for transmitting data. However; there are limitations, such as low bandwidth, limited AUV's mobility and coverage [1]. Recently, optical wireless communication (OWC) is considered as a promising candidate to relieve the highly congested radio-frequency (RF) communication for providing wireless communication [2]. Compared with the traditional underwater communication techniques using acoustic signal, RF, and cable; underwater wireless optical communication (UWOC) could provide distinct advantages including higher bandwidth and higher flexibility. Furthermore, UWOC could also be a promising candidate for the realization of direct connectivity in the water-air interface [1, 3]. The water-air optical wireless transmission is highly desirable to transmit real-time data from underwater AUV to unmanned aerial vehicle (UAV) or airplane, which then sends the data to ground stations using OWC or RF. Despite the aforementioned merits, UWOC and water-air optical communication have several challenges. Besides the absorption and scattering introduced by water, the water turbulence also generates power fluctuation at the receiver (Rx), producing scintillation effect. In the water-air optical communication, the wavy water surface will significantly deflect the optical signal reaching the Rx, producing communication link outage. To mitigate the deflection of optical signal; higher field-of-views (FOV) for both the transmitter (Tx) and Rx are highly desirable. Optical camera communication (OCC) is one realization of OWC using camera as Rx [4]. The advantages of OCC include the high availability of complementary-metal-oxide-semiconductor (CMOS) cameras in mobile-phones and UAVs; large size of photo-sensitive area; as well as the rolling shutter effect of CMOS camera allowing OCC data rate higher than camera frame rate. Table 1 summaries the recent rolling shutter camera based OCC systems. On-off keying (OOK) is usually used; and thresholding schemes, including polynomial-curve-fitting [4], extreme-value-averaging [5], beacon-jointed packet reconstruction [6] were utilized for decoding the rolling shutter pattern. Recently, artificial intelligence/machine learning (AI/ML) schemes were also utilized to enhance the OCC performances, including convolutional neural network (CNN) [7, 8], pixel-per-symbol labeling neural network (PPS-NN) [9].

In this work, we propose and preset the first demonstration up to the authors' knowledge a wide FOV water-to-air optical transmission using rolling shutter based OCC. Long-short-term-memory-neural-network (LSTM-NN) [10] is utilized to mitigate the wavy water-surface induced link outage and to decode 4-level pulse-amplitude-modulation (PAM4) rolling-shutter pattern. Experimental results show that the OCC system can support ±70^o, ±30^o, and ±30^o rotations around the *z*-, *y*- and *x*-directions, respectively when operated at 6 kbit/s.

Table I. Recent works Rolling shutter based OCC systems.

Ref.	Modulation & Decoding Scheme	Data Rate (kbit/s)	Distance (m)	Water-Air Interface
[4]	OOK (Poly. curve fitting)	3.1	0.35	--
[5]	OOK (Extreme value averaging)	7.68	0.3	--
[6]	OOK (Beacon jointed packet reconstruction)	10.32	0.2	--
[7]	OOK & RGB (2D-CNN)	47	0.4	--
[8]	OOK & RGB (Double-EQ)	111	0.4	--
[9]	PAM4 (PPS-NN)	14.4	2	--
This work	PAM4 (LSTM-NN)	7.2 (no ripple) / 6 (high ripple)	1(Air) + 0.1(Water)	Yes

2. Algorithm, Experiment and Results

Fig. 1(a) shows the scenario of the water-to-air optical communication from the AUV to UAV. As discussed before, the wavy water surface will significantly deflect the optical signal reaching the Rx, producing communication link outage. To mitigate the optical signal deflection; higher FOVs for both Tx and Rx are highly desirable. During the rolling shutter in CMOS camera, the camera Rx will integrate signal in a row-by-row manner similar to a scanning function. Hence, bright and dark fringes can be observed in the rolling shutter pattern representing LED "ON" and "OFF" [4, 5]. The rolling shutter effect allows the data rate of OCC system faster than the camera frame-rate. Since the next pixel-row will start to activate before the completion of the previous pixel-row during the rolling shutter operation, high inter-symbol interference (ISI) occurs when the LED is modulated at high data rate. This problem is even more severe in PAM4 OCC decoding.

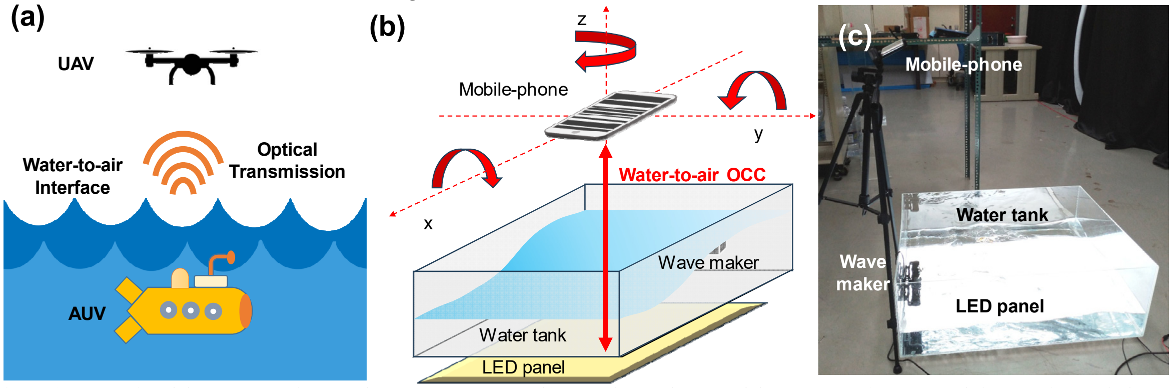

Fig. 1. (a) Scenario of the water-to-air OCC from AUV to UAV. (b) Experimental setup of the water-to-air OCC, and (b) experimental photo.

Figs. 1(b) and (c) show the proof-of-concept experiment and the photo of the proposed water-air optical transmission using rolling-shutter OCC. The PAM4 signal is generated by a 22 W LED light panel connected to an arbitrary waveform generator (AWG, Tektronix® AFG3252C). The dimensions of the LED panel and water tank are 0.58 m × 0.88 m; and 0.73 m × 0.88 m, respectively. The camera has the resolution of 1920 × 1440 pixels and 30 fps. The distance between the mobile-phone and water surface is 1 m, and the water depth is 0.1 m. A wave maker (Jebao® SCP-120) is used to produce the wavy water surface. The percentage increase in water ripple is define as $\{[(h_{peak} - h_{ave})/h_{peak}] - 1\}\times 100\%$, where h_{peak} and h_{ave} are the height of water peak and average water level, respectively. Here, the wave maker used to produce wavy water surface mechanically has only two operating modes; hence, we only emulate 3 scenarios: no water ripple, 9% and 12% increases of water ripple.

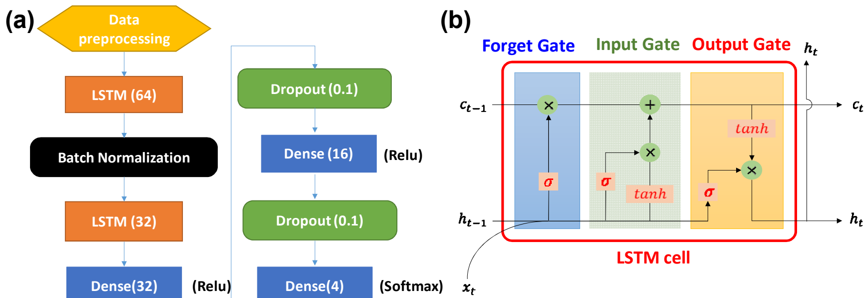

Fig. 2. Architectures of (a) LSTM-NN model used to decode the PAM4 OCC signal; (b) LSTM cell.

Fig. 2(a) shows the algorithm of the LSTM-NN model used to decode the PAM4 OCC signal and to mitigate the wavy water-surface induced link outage. Image frames from video are read-in and converted into grayscale values (0 to 255). Column matrix selection is performed to select the highest grayscale value in each pixel-row. In the data pre-processing unit, features of the input signal, including present symbol value, pre-/post-symbol relations and symbol average are extracted. After this, the data is sent to the LSTM layer. In this proposed LSTM-NN model, we have optimized the decoding performance and the complexity. The model has two LSTM layers with neuron numbers of 64 and 32 respectively. Batch normalization is utilized. The last three layers are the fully-connected (FC) dense layers, with neuron numbers of 32, 16 and 4 respectively. The dropouts are used to prevent overfitting. Relu is used at the 1st and 2nd FC dense layers; and Softmax is the activation function in the last layer. The loss function is sparse categorical cross entropy and the optimizer is Adam. Fig. 2(b) illustrates the LSTM cell used in the LSTM layer. C_{t-1}, C_t, x_t, σ, h_{t-1}, h_t, are the memory from the last LSTM cell, newly updated memory, current input, Sigmoid function, output of last LSTM cell, and current output respectively. Three internal gates, e.g. forget, input and output gates are used to decide whether the information is erased, stored and output based on the cell state after filtering.

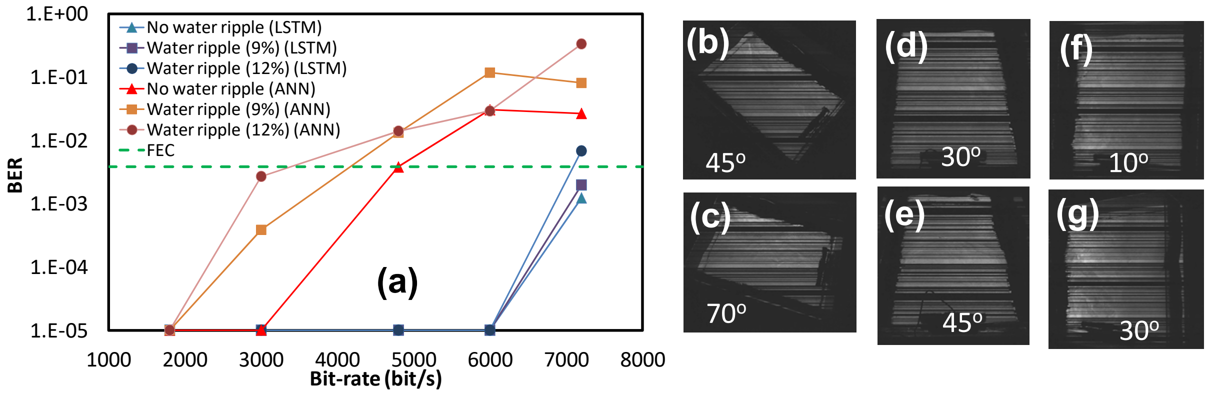

Fig. 3. Measured BERs of the water-to-air rolling-shutter-based OCC system at different water ripple conditions.

Fig. 3 shows the measured bit-error-rate (BER) of the water-to-air rolling-shutter-based OCC. The wave maker produces two kinds of water ripples: frequency 1.2 Hz (9% increase of water ripple); frequency 2 Hz (12% increase of water ripple). We also compare the BER performance with a conventional artificial neural network (ANN). The ANN is also optimized with an input, output, and 6 FC hidden layers. The neuron numbers in the FC hidden layers are 36, 64, 128, 64, 16, 4. Relu and Softmax activation functions are used for the first 5 FC layers and the last layer respectively. At no ripple case, we can observe that LSTM-NN can successfully decode PAM4 at 7.2 kbit/s, satisfying the pre-forward error correction (FEC) BER threshold (BER = 3.8×10^{-3}). However, the traditional ANN can only decode PAM4 at 4.8 kbit/s. At the 9% water ripple case, the proposed LSTM-NN can also operate at 7.2 kbit/s satisfying the FEC; however, the ANN can only operate at 3 kbit/s. We also evaluate the FOVs of the proposed system. As shown in Fig. 1(b), the mobile-phone is rotated in *z*-, *y*- and *x*-directions, respectively. Figs. 3(b)-(c) show the rolling shutter pattern rotated around the *z*-axis at 45° and 70°, respectively. Assume at zero rotation, the LED panel occupies *x* and *y* pixels in the horizontal and vertical directions respectively. When the mobile-phone is rotated, the occupied pixels of the LED panel in vertical direction of the screen decrease. When the *z*-rotation angle is > 70°, BER increases significantly. Figs. 3(d)-(e) show the rolling shutter pattern rotated around the *y*-axis at 30° and 45°, respectively. When *y*-axis rotation is > 30°, the LED panel becomes "trapezium" shaped and the occupied pixels in vertical direction decreases, increasing the BER. Figs. 3(f)-(g) show the rolling shutter pattern rotated around the *x*-axis at 10° and 30°. When x-axis rotation is > 30°, the optical reflection from the water tank side wall becomes severe, and interferes with the OCC signal. Experimental results show that the proposed OCC system can support ±70°, ±30°, and ±30° rotations around the *z*-, *y*- and *x*-directions, respectively, satisfying the FEC (i.e. BER = 3.8×10^{-3}) when operated at 6 kbit/s and decoded using LSTM-NN.

3. Conclusion

We demonstrated for the first time a wide FOV water-to-air rolling shutter based OCC system. LSTM-NN was utilized to mitigate the wavy water-to-are surface induced link outage. Experimental results showed that the proposed OCC system can support ±70°, ±30°, and ±30° rotations around the *z*-, *y*- and *x*-directions, respectively, satisfying the FEC (i.e. BER = 3.8×10^{-3}) when operated at 6 kbit/s and decoded using LSTM-NN.

Acknowledgment This work was supported by National Science and Technology Council, Taiwan (NSTC-112-2221-EA49-102-MY3, NSTC-110-2221-E-A49-057-MY3).

5. References

[1] L. K. Chen, et al, "Underwater and water-air optical wireless communication," J. Lightwave Technol, 40, 1440 (2022).
[2] C. W. Chow, et al, "Enabling techniques for optical wireless communication systems," OFC 2020, paper M2F.1.
[3] H. M. Oubei, et al, "Performance evaluation of underwater wireless optical communications links in the presence of different air bubble populations," IEEE Photon. J., 9, 7903009 (2017).
[4] C. Danakis, et al, "Using a CMOS camera sensor for visible light communication", IEEE Globecom Workshops, 2012, 1244.
[5] C. W. Chen, et al, "Efficient demodulation scheme for rolling-shutter-patterning of CMOS image sensor based visible light communications," Opt. Exp., 25, 24362 (2017).
[6] W. C. Wang, et al, "Beacon jointed packet reconstruction scheme for mobile-phone based visible light communications using rolling shutter," IEEE Photon. J., 9, 7907606 (2017).
[7] L. Liu, et al, "47-kbit/s RGB-LED-based optical camera communication based on 2D-CNN and XOR-based data loss compensation," Opt. Exp., 27, 33840 (2019).
[8] L. Liu, et al, "Beyond 100-kbit/s transmission over rolling shutter camera-based VLC enabled by color and spatial multiplexing," OFC 2020, Paper M1J.4.
[9] 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 (2021).
[10] S. Hochreiter, et al, "Long short-term memory," Neural Computation, 9, 1735 (1997).