Seeing through Wave—Real-time Beam Tracking via a ResNet-based Model in Water-air OWC Systems

Anzi Xu, Yujie Di, Xiangyu Yue, Lian-Kuan Chen
Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China
Author e-mail address: {xa022, dy019, xyyue, lkchen}@ie.cuhk.edu.hk

Abstract: A ResNet-based model using one wave-distorted image input is demonstrated for real-time beam tracking. Packet loss rates reduce from 18% to 4% under a wave's ASCR of 0.614 rad/s, realizing a robust water-air OWC system. © 2024 The Author(s)

1. Introduction

In the emerging deep-sea mining and exploration, autonomous underwater vehicles (AUVs) are employed to gather seabed information that can be transferred back to the control site situated on land. Compared with acoustic and RF systems, the optical wireless link is promising for water-air communications due to its low attenuation in underwater and water-air interfaces [1]. However, the wavy sea surface presents a key challenge for water-air optical wireless communication (WA-OWC) systems, since lights will be deflected by the wavy surface when passing the water-air interface, causing beam wandering on the receiver plane. This leads to a high bit error rate (BER) and packet loss rate (PLR).

To curtail the aforementioned beam wander issue, a WA-OWC beam tracking scheme was proposed [2] using a photodiode array, but the tracking resolution was limited by the array's size. A camera can accurately capture the laser's position owing to its high resolution. However, when seeing through waves from underwater, the beam position and its background image are distorted by the uneven and dynamically changing water surface. It is challenging to restore the distorted image for accurate beam offset estimation since the wave-induced distortion is non-rigid and fast-changing. Computer vision methods such as mean pixel [3], lucky fusion [4], and refraction stereo [5] have been developed for the restoration of wave-distorted images to see through waves without distortion. Compared with these schemes, machine learning-based methods vastly enhance image restoration and complexity performance [6]. Given the capability of extracting spatial and time relationships, convolutional neural networks (CNN) and recurrent neural networks (RNN) have been applied to restore distorted images [7]. However, the scheme is time-consuming and too complex for a real-time tracking system.

In this paper, a neural-network-assisted image-based real-time tracking system is proposed to realize a robust WA-OWC system. An optimized residual neural network (ResNet) based model is employed to estimate a beam position using only one distorted image, and the process repeats periodically. The model achieves a position estimation error of 1.89 mm. It is shown experimentally that a 600-Mbit/s water-air transmission is realized with a 4% PLR via the ResNet-based tracking, whereas the transmission without tracking suffers an 18% PLR. This is the first realization of a real-time water-air tracking system using a neural network. The results also suggest that wave-distorted images contain wave profile information, which can be extracted by machine learning for beam tracking.

2. Principle

When the laser beam passes the water-air interface, it will be deflected by the wavy surface. The laser beam falls on a checkerboard at the receiver (Rx) side. At the transmitter (Tx) side, an underwater camera captures the Rx-side checkerboard, and the laser position is estimated by a computer with the captured image. Then, the estimated beam position is fed to a microcontroller unit (MCU). Via using a proportional-integral-derivative (PID) algorithm, the MCU calculates the required value for the micro-electromechanical system (MEMS) mirror's beam-steering angle to tilt the beam back to the Rx center. For short-distance tracking systems, images of the beam on the receiver plane, disregarding whether they are taken at the Tx or Rx sides, provide sufficient information for the estimation of beam position offset. However, for WA-OWC, Rx-side images captured from an underwater camera at the Tx side are severely distorted by waves, resulting in an erroneous estimation of beam position.

We propose a ResNet-based model to estimate the beam position using only one wave-distorted image. It is worth noting that the ResNet model proposed in [8] was for the feature extraction of high-resolution images. Therefore, it has numerous conventional layers, leading to a long processing latency. In our system, a ResNet-based model utilizes image distortion features and estimates the Rx beam position. As shown in Fig. 1(a), we optimize the model for both estimation accuracy and processing time reduction to realize a real-time tracking system as follows. (i) The input size is reduced to 10×10 pixels. (ii) The model only consists of four residual blocks. (iii) Rather than restoring the whole distorted image, one fully connected layer only outputs the estimated beam coordinates in the x and y directions. Eventually, the Rx-side beam position can be estimated within 10 ms, as compared to the 17-ms estimation time before model optimization.

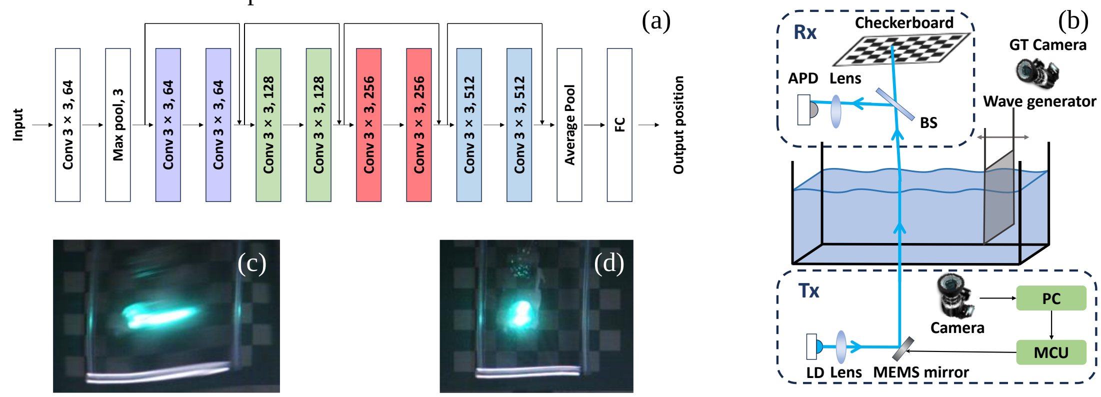

Fig. 1. (a) The structure of the ResNet-based model, (b) the experimental setup of the ResNet-based tracking system, and underwater camera captured images (c) with and (d) without wave.

3. Experimental setup and results

In Fig. 1(b), a ResNet-assist image-based tracking system is illustrated. A 60×30×40 cm (length×width×depth) tank is filled with 0.1-m-deep water, and the air distance is 1.05 m. The wave generator is a servo-motor-controlled paddle with tunable moving speeds and distances. The wave average slope changing rate (ASCR) is calculated as in [9] to indicate how volatile the wave slope (the most critical parameter) is changing. At the Rx, it contains a 20×20 cm checkerboard, a beam splitter (BS), and a 1-GHz avalanche photodiode (APD, Hamamatsu, C5658). The beam size at the receiver is 9 mm. A lens is placed in front of the APD to enlarge the Rx's field of view. At Tx, an OOK signal of different data rates is generated by an arbitrary waveform generator (AWG, Tektronix 7122 C), coupled with a 6.4-V bias via a bias-tee, and applied to the laser diode (LD, 510 nm, 20 mW). One hundred packets are collected under different conditions (ASCR and data rate), and each packet contains 10,000 symbols. The maximum angle of the MEMS mirror (Mirrorcle, A8L2.2-4600AL-TINY48.4-A/TP) is ~ ±5 degrees. The underwater and ground-truth (GT) cameras are both software-triggered high-speed cameras (HTSUA33GC/M, 790 fps, 640×480). The image is sent to our model that is implemented in a personal computer (AMD Ryze 7 5800H). Fig. 1(c) shows wave-induced image blurs and non-rigid distortion, compared with that in Fig. 1(d). To train the model, ground-truth beam positions are measured by the GT camera above water. The ground-truth beam positions are only used to calculate the mean square error (MSE) to train the model. The training and test sets are collected under the conditions of a wave ASCR of 0.498 rad/s, a 1.05-m air path, and a 0.1-m water path.

Fig. 2(a) and (b) show the estimation results of the trained ResNet-based model for different image frames (corresponding to different time instants). The fluctuation in the x direction is larger since the wave generator moves back and forth in the x direction. For the model, the MSE is 3.814, and the mean absolute error (MAE) is 1.6102. In Fig. 2(c), we compare the tracking performance of the proposed method and a method using direct beam position estimation without image restoration under different ASCRs. Standard deviations of the beam position are calculated to evaluate tracking performance. A larger standard deviation implies a larger fluctuation of the laser beam. With the ResNet-based tracking method, the standard deviation is 5.526 mm when ASCR is 0.614 rad/s, showing a 42% reduction of standard deviation when compared with the non-tracking case. For the non-image-restoration tracking case, the image is blurred and distorted when ASCR increases, resulting in degraded tracking performance. This can be attributed to the fact that the method without image restoration cannot distinguish beam offset due to (i) beam wandering and (ii) image distortion by wave.

Then, we investigate the average BER and PLR performance of 600-Mbit/s OOK signals for different ASCRs in Fig. 3(a). Average BER is calculated based on all 100 packets in each case. PLR is the ratio of the number of lost packets and the total number of packets. A packet is defined as lost if its BER is above the soft-decision forward error correction (SD-FEC) limit (2.0×10^{-2}). By controlling the speed of the wave generator, five ascending ASCRs are set, from 0.486 rad/s to 0.74 rad/s. With ResNet-based tracking, average BERs are below SD-FEC for cases with ASCR < 0.614 rad/s. While without tracking, average BERs are all above SD-FEC for ASCRs. In addition, the PLR ranges from 1% to 8% for different ASCRs with tracking, whereas it is 7% to 19% without tracking. Albeit the model is only trained on a dataset with an ASCR of 0.498 rad/s, the model can estimate the beam position and realize a tracking system with an ASCR ranging from 0.48 rad/s to 0.74 rad/s. This is plausible as different ASCRs have similar wave profiles, which can be extracted by the model. At high ASCR, the ResNet-based tracking system performance is reduced due to the limited MEMS's response time and the camera's exposure time. We also analyze system performance under different data rates as shown in Fig. 3(b). With tracking, the average BERs are all below SD-FEC, and PLRs are no more than 2% for data rates below 1 Gbit/s. The average BERs for non-tracking cases are all above SD-FEC and PLRs degrade to 6%–14% for data rates below 1 Gbit/s.

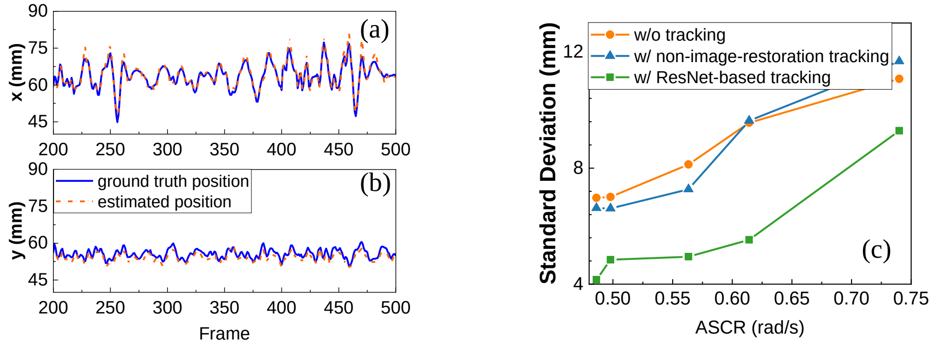

Fig. 2. Beam position estimation using trained ResNet-based model of (a) x- and (b) y-direction, and (c) standard deviation of beam positions versus wave average slope changing rate (ASCR).

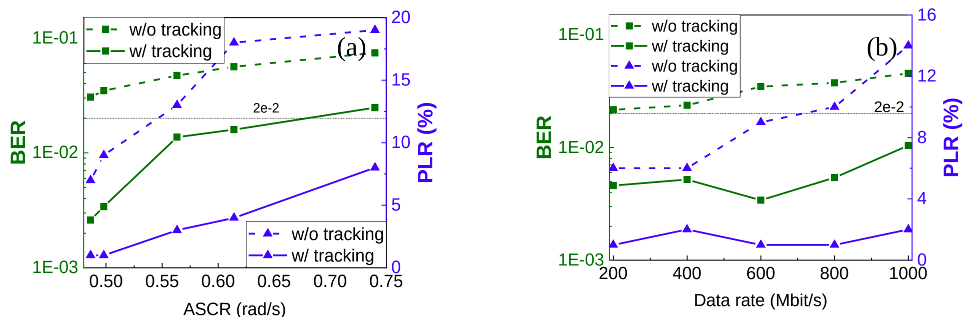

Fig. 3. BER and PLR performance without tracking and with ResNet-based tracking versus (a) ASCR with a 600-Mbit/s OOK signal and (b) data rate with a wave ASCR of 0.498 rad/s.

4. Conclusion

In this paper, a ResNet-assisted image-based tracking system is proposed and experimentally verified. The performance of the tracking system is investigated under different ASCRs and data rates compared with the non-tracking system. Experimental results show that a maximum throughput of 980 Mbit/s can be achieved under an ASCR of 0.498 rad/s with the tracking system. The PLR is reduced to 4% under an ASCR of 0.614 rad/s, as compared to 18% of the non-tracking system. By refining and training a ResNet-based model, a water-air OWC system with real-time beam tracking is achieved using only an underwater Tx-side camera. The proposed model can be further improved to predict future beam positions using the wave-distorted image, which may further enhance tracking performance. This work is supported in part by HKSAR GRF 14204921 and 14219322.

5. References

[1] L. -K. Chen, Y. Shao, and Y. Di, "Underwater and Water-Air Optical Wireless Communication," in *Journal of Lightwave Technology*, vol. 40, no. 5, pp. 1440-1452, 2022.
[2] Y. Di, Y. Shao, and L. -K. Chen, "Mitigation of Wave-induced Packet Loss for Water-air Optical Wireless Communication by A Tracking System," in *Proc. of OFC*, pp. 1-3, 2021.
[3] Iosif M Levin, Victor V Savchenko, and Vladimir Ju Osadchy, "Correction of an image distorted by a wavy water surface: laboratory experiment" in *Applied Optics*, vol. 47, pp. 6650–6655, 2008.
[4] David L. Fried, "Probability of getting a lucky short-exposure image through turbulence*," JOSA A, vol. 68, pp. 1651-1658, 1978.
[5] N. J. W. Morris and K. N. Kutulakos, "Dynamic refraction stereo," in *Proc. of ICCV*, Vol. 2, pp. 1573-1580, 2005.
[6] N. Li, S. Thapa, C. Whyte, A. Reed, S. Jayasuriya, and J. Ye, "Unsupervised Non-Rigid Image Distortion Removal via Grid Deformation," in *Proc of ICCV*, pp. 2502-2512, 2021.
[7] Thapa, Simron, N. Li, and J. Ye., "Dynamic fluid surface reconstruction using deep neural network," in *Proc of CVPR*, pp. 21-30, 2020.
[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc of CVPR*, pp. 770-778, 2016.
[9] Y. Di, Y. Shao, and L. -K. Chen, "Real-Time Wave Mitigation for Water-Air OWC Systems via Beam Tracking," in *IEEE Photonics Technology Letters*, vol. 34, no. 1, pp. 47-50, 2022.