
Article:
Zhang, Youwen, Li, Junhuan, Zakharov, Yuriy et al. (Accepted: 2019) Deep Learning based Single Carrier Communications over Time-Varying Underwater Acoustic Channel. IEEE Access. ISSN: 2169-3536 (In Press)

Deep Learning based Single Carrier Communications over Time-Varying Underwater Acoustic Channel

Youwen Zhang, Junxuan Li, Yuriy Zakharov, *Senior Member, IEEE*, Jianghui Li, *Member, IEEE*, Yingsong Li, *Member, IEEE*, Chuan Lin, Xiang Li

***Abstract*—In recent years, deep learning (DL) techniques have shown great potential in wireless communications. Unlike DL-based receivers for time-invariant or slow time-varying channels, we propose a new DL-based receiver for single carrier communication in time-varying underwater acoustic (UWA) channels. Without the off-line training, the proposed receiver alternately works with online training and test modes for accommodating the time variability of UWA channels. Simulation results show a better detection performance achieved by the proposed DL-based receiver and with a considerable reduction in training overhead compared to the traditional channel-estimate (CE) based decision feedback equalizer (DFE) in simulation scenarios with a measured sound speed profile. The proposed receiver has also been tested by using the data recorded in an experiment in the South China Sea at a communication range of 8 km. The performance of the receiver is evaluated for various training overheads and noise levels. Experimental results demonstrate that the proposed DL-based receiver can achieve error free transmission for all 288 burst packets with lower training overhead compared to the traditional receiver with a CE-based DFE.**

***Index Terms*—Channel equalization, deep learning, deep neural network, DFE, machine learning, single carrier communication, underwater acoustic network.**

I. Introduction

Underwater acoustic (UWA) channel features frequency-dependent limited bandwidth, long time-varying multipath spread and severe Doppler effect, which pose a great challenge for reliable and effective UWA communications [1]–[7] and networks [8]–[10], leading to relatively low data rates in a range between a few bits/s (bps) to several tens of kbits/s ($kbps$) and often unsatisfied performance [11]. Generally, single-carrier (SC) modulation schemes with time-domain equalization techniques enjoy high spectral efficiency and robust performance at the cost of a high receiver complexity [1], [4], [12]. Historically, in order to combat the inter-symbol interference (ISI) induced by the time-varying multipath spread, many channel equalization techniques have been thoroughly studied, g.e., linear equalizer (LE), decision feedback equalizer (DFE) etc. [13]–[16]. However, there is still a great room for the improvement in UWA communication systems.

In recent years, machine learning techniques have attracted attention in different fields. In particular, deep learning (DL) techniques feature great potential for solving nonparametric problems such as object detection and recognition, voice recognition, and object tracking [17]–[19], [21]. Although DL has been adopted for terrestrial radio wireless communication only recently [22]–[25], it has also been utilized in UWA communications [26]. In [22]–[24], DL techniques were proposed for joint channel estimation and symbol detection in OFDM systems. Simulation results demonstrate that deep neural network (DNN) has the ability to learn and analyze characteristics of wireless channels with nonlinear distortion and interference in addition to the frequency selectivity. In [25], learning assisted (LA) algorithms are proposed for estimation of time-varying channels. The DNN based channel estimators are utilized to track channel variations. Simulation results validate the effectiveness of the algorithms in online tracking the channel variations. In [26], inspired by the works in [23], [24], a DL-based UWA OFDM communication scheme is proposed and verified by simulation in a UWA channel with a measured sound speed profile (SSP). Despite the success of DNN in time-invariant or slow time-varying channels, DNN-based wireless communication over fast time-varying channels induced by severe Doppler effects has not been studied yet.

In this paper, we propose a DL-based receiver for UWA SC communications over time-varying channels. As compared to existing works, our main contributions are summarized below:

1) Different from the existing DL-based receivers with the offline training and online test modes, we propose a DL-based receiver with online training and test modes for accommodating time-varying UWA channels.
2) The performance of the proposed DL-based receiver is evaluated by a statistical channel simulator with a measured SSP. Numerical results show that the proposed receiver achieves a better detection performance than

Youwen Zhang and Xiang Li are with the Acoustic Science and Technology Laboratory and the College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin, Heilongjiang, 150001, China. Email: zhangyouwen@hrbeu.edu.cn and lixianghrbeu1982@163.com.

Junxuan Li is with the College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin, Heilongjiang, 150001, China. Email: lijunxuan@hrbeu.edu.cn.

Yuriy Zakharov is with the Department of Electronic Engineering, University of York, York, YO10 5DD, UK. Email: yury.zakharov@york.ac.uk.

Jianghui Li is with the Institute of Sound and Vibration Research, University of Southampton, U.K., e-mail: J.Li@soton.ac.uk.

Yingsong Li is with the College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China. Email: liyingsong@ieee.org.

Chuan Lin is with the Acoustics Science and Technology LaboratoryHarbin Engineering University, Heilongjiang, 150001, China. He is also with the Key Laboratory for Ubiquitous Network and Service Software of Liaoning province, School of Software, Dalian University of Technology, Dalian, 116024, China. Email: chuanlin1988@gmail.com.

the traditional CE-based decision-feedback equalization (DFE) receiver and with lower training overhead.

3) The performance of the proposed receiver has been tested in the South China Sea experiment, at a communication distance of 8 km. We show that the proposed DL-based receiver can achieve a substantial performance gain over the traditional CE-based DFE receiver with the second-order PLL. With less training overhead, the proposed receiver can achieve error-free transmission of 288 data packets at the signal-to-noise ratio (SNR) as low as SNR=5dB, while the traditional CE-based DFE receiver cannot achieve that even at as high SNR as SNR = 15 dB and with a significantly longer training sequence. At the best of our knowledge, this is the first time that a DL based receiver is validated using data from a sea trial instead of simulated data.

The rest of this paper is organized as follows. In Section II, the time-varying UWA SC communication system model is presented and the CE-based DFE receiver with second-order PLL is reviewed. Section III presents the proposed DL-based receiver for UWA communication over time-varying UWA channels. Simulation results are presented in Section IV. Section V presents results from the sea trial. Conclusions are drawn in Section VI.

Notation: Matrices and vectors are denoted by boldface uppercase and lowercase letters, respectively. $(\cdot)^{\dagger}$, $(\cdot)^{*}$ and $(\cdot)^{T}$ denote the Hermitian transposition, complex conjugate and transposition, respectively.

II. System Model for UWA SC Communications over Time-Varying Channel

A. Signaling Model

We consider a single-input single-output UWA SC communication system. Fig. 1 depicts the block diagram of the transmitter. The binary information bit vector $\mathbf{b}$ is split into groups of P bits, where P represents the number of bits per symbol, and each group is mapped to one of the 2^P-ary symbols of the alphabet $\mathcal{A} = \{\alpha_p\}_{p=1}^{P}$, where α_p is a complex number. The sequence of 2^P-ary symbols is multiplexed with a training symbol sequence of length N_t, producing the payload symbol vector $\mathbf{x} \triangleq \{x_n\}_{n=1}^{N_s}$. Symbols from the vector $\mathbf{x}$ pass through a square-root raised-cosine pulse-shaping filter with an impulse response $g(t)$ to produce the baseband signal $b(t)$. The complex baseband signal $b(t)$ is expressed as [27]

$$b(t) = \sum_{n=1}^{N_s} x_n g(t - nT_s), \tag{1}$$

where x_n is the transmitted symbol, $g(t)$ is a square-root raised cosine pulse-shaping filter with roll-off factor γ, T_s is the symbol interval. The signal $b(t)$ is then modulated onto a carrier of angular frequency ω_c to produce the transmitted signal $s(t)$ as [1], [27]

$$s(t) = Re\{b(t)\, e^{j\omega_c t}\}. \tag{2}$$

Preamble and postamble linear frequency modulation (LFM) waveforms are added before and after $s(t)$ for the purpose of the coarse frame synchronization and Doppler estimation.

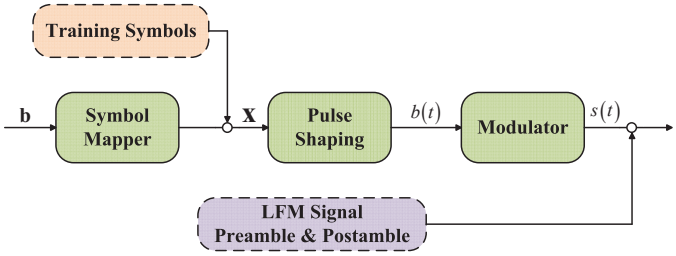


Fig. 1. Block diagram of the transmitter for the UWA SC communication system.

In this paper, we consider the narrowband signaling model, i.e., the Doppler effect can be represented as a carrier frequency offset without time scaling. We assume that the maximum channel delay (in symbol intervals) is L.

So the received baseband signal distorted by multipath spread and noise can be expressed as

$$y^b(t) = \sum_{l=0}^{L-1} \beta_l(t) b\big(t - \tau_l\big) e^{j\omega_c(-\tau_l)} + \eta(t), \tag{3}$$

where $\beta_l(t)$ is time-varying amplitude fading factor corresponding to the l-th path, τ_l is the delay associated with the l-th path and $\eta(t)$ is additive complex white Gaussian noise (AWGN) with zero mean and variance σ_η^2 at hydrophone, which is independent from $b(t)$.

B. Traditional CE-based DFE Receiver with PLL

Fig. 2 depicts the receiver with the traditional CE-based DFE and PLL. This type of DFE is widely used for combatting the inter-symbol interference (ISI) and phase distortion in UWA communication channels [15], [27].

On the receiver side, the passband signal $\mathbf{y}^p(n)$ is transformed into baseband signal $\mathbf{y}^b(n)$ by a demodulator, and the baseband signal $\mathbf{y}^b(n)$ is downsampled into $y(k)$, where k is the time index in the symbol interval T_s. Assume that a DFE consists of an L_f-length feedforward filter (FFF) with the tap vector $\mathbf{f}(k)$ and L_b-length feedback filter (FBF) with the tap vector $\mathbf{g}(k)$, and the equalizer delay is l [27], [28].

At time instant k, the transmitted symbol $x(k-l)$ estimated by DFE as [28]

$$\hat{x}(k-l) = \mathbf{y}^T(k)\mathbf{f}(k) + \hat{\mathbf{x}}^T(k)\mathbf{g}(k), \tag{4}$$

where $\mathbf{y}(k) = [y(k), y(k-1), \cdots, y(k-L_f+1)]^T$, and the already estimated symbol vector $\hat{\mathbf{x}}(k) = [\hat{x}(k-l-1), \hat{x}(k-l-2), \cdots, \hat{x}(k-l-L_b)]^T$. Fig. 3 depicts the structure of this type DFE in details. The equation (4) can be written in vector form as [27], [28]

$$\hat{x}(k-l) = \mathbf{m}^T(k)\mathbf{n}(k), \tag{5}$$

where

$$\mathbf{m}(k) = \begin{bmatrix} \mathbf{y}(k) \\ \hline -\hat{\mathbf{x}}(k) \end{bmatrix}, \mathbf{n}(k) = \begin{bmatrix} \mathbf{f}(k) \\ \hline \mathbf{g}(k) \end{bmatrix}. \tag{6}$$

With an adaptive channel estimator, we can estimate the equivalent time-varying baseband channel matrix $\hat{\mathbf{H}}(k)$ by

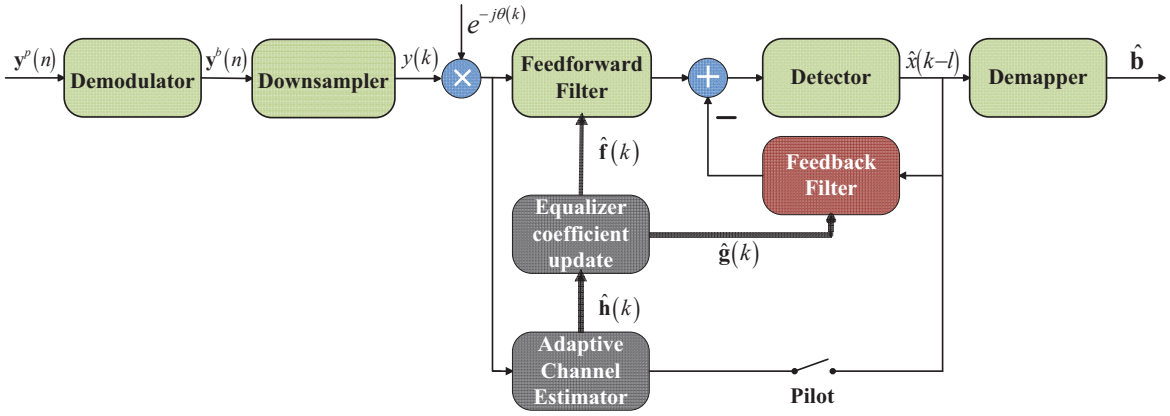


Fig. 2. Block diagram of the receiver structure for the traditional CE-based DFE.

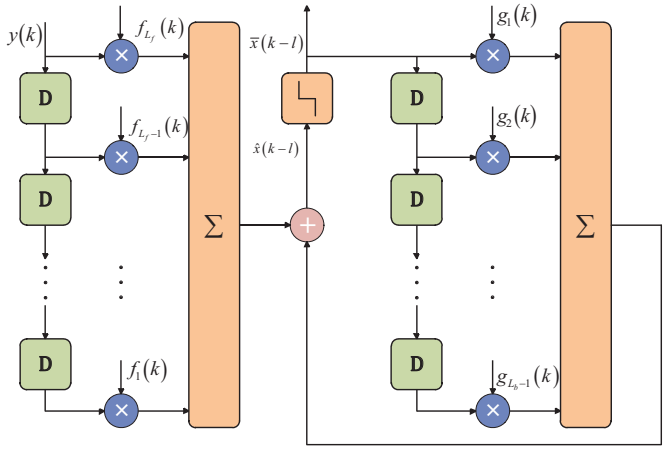


Fig. 3. Structure of a symbol-spaced DFE. $\bar{x}(k-l)$ denotes the hard decision of $\hat{x}(k-l)$

using the training symbols or the hard decisions of estimated data symbols [28]. Finally, the equalizer taps $\hat{\mathbf{f}}(k)$ and $\hat{\mathbf{g}}(k)$ are generated and updated by solving the equivalent mean square error (MSE) equation as follows [27]

$$\Gamma(k)\mathbf{n}(k) = \Psi(k), \tag{7}$$

where $\Gamma(k) = E\left\{\mathbf{m}(k)\mathbf{m}^H(k)\right\}$, $\Psi(k) = E\left\{\mathbf{m}(k)\hat{x}^*(k-l)\right\}$.

In order to compensate the phase distortion induced by the Doppler effect, the second-order PLL is incorporated into the CE-based DFE for joint carrier phase synchronization and equalization [29]. For easy understanding, we set the decision delay $l = 0$, so at time instant k, output of the FFF is given by

$$p(k) = \mathbf{f}^\dagger(k)\mathbf{y}(k)e^{-j\theta(k)}, \tag{8}$$

and the output of the FBF is written as

$$q(k) = \hat{\mathbf{g}}^\dagger(k)\hat{\mathbf{x}}(k), \tag{9}$$

where the time-varying phase can be tracked by the second-order PLL as follows [29]

$$\theta(k+1) = \theta(k) + K_{f_1}\Phi(k) + K_{f_2}\sum_{i=0}^{k}\Phi(i), \tag{10}$$

where K_{f_1} and K_{f_2} are proportional coefficient and integral coefficient of loop filter, respectively.

III. Proposed DL-based Receiver for UWA SC Communication over Time-Varying Channels

A. Review of DNN

DNN is an artificial neural network (ANN) with multiple hidden layers composed of many neurons [30]. Fig. 4 depicts the general structure of a DNN model with $Q > 3$ layers, where the layer 1 and layer Q are called as input layer and output layer, respectively, and the other layers are hidden layers. Data flow propagates from the input layer to the hidden layers, then the output layer.

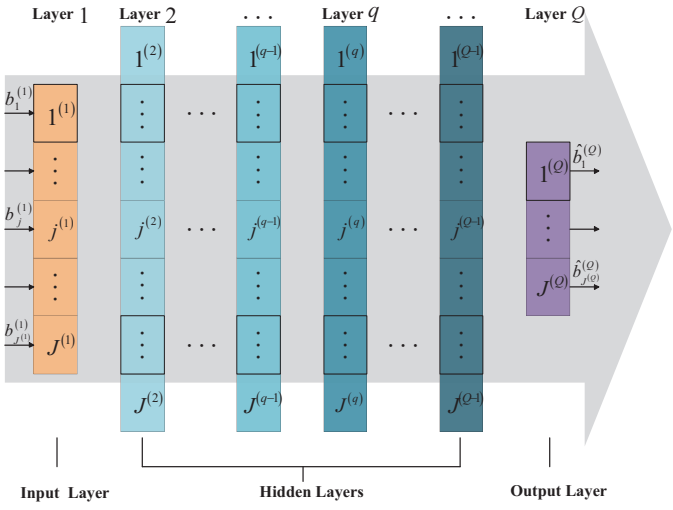


Fig. 4. General structure of a DNN model.

The input layer (i.e. layer 1) has $J^{(1)}$ variables in vector $\mathbf{b}^{(1)} = \left[b_1^{(1)}, \cdots, b_j^{(1)}, \cdots, b_{J^{(1)}}^{(1)}\right]^T$, then the j-th neuron's input of layer 2 is

$$a_j^{(2)} = \sum_{i=1}^{J^{(1)}} u_{ij}^{(1)} b_i^{(1)} + v_j^{(1)}, \quad j = 1, 2, \cdots, J^{(2)}, \tag{11}$$

where $u_{ij}^{(1)}$ is a weight between the i-th neuron of layer 1 and the j-th neuron of layer 2, $v_j^{(1)}$ is a bias of the j-th neuron in layer 2, $J^{(2)}$ is the number of neurons of layer 2. In each hidden layer, there is a nonlinear activation function $f(\cdot)$ which transforms the linear combinations of inputs to non-linear outputs. The j-th neuron's output of layer 2 is thus given by [30]

$$b_j^{(2)} = f^{(2)}\left(a_j^{(2)}\right). \tag{12}$$

Then $\mathbf{b}^{(2)} = \left[b_1^{(2)}, \cdots, b_j^{(2)}, \cdots, b_{J^{(2)}}^{(2)}\right]^T$ will be the next layer's input. Similarly, the j-th neuron's input $a_j^{(q)}$ and output $b_j^{(q)}$ of layer q are

$$a_j^{(q)} = \sum_{i=1}^{J^{(q-1)}} u_{ij}^{(q-1)} b_i^{(q-1)} + v_j^{(q-1)}, j = 1, 2, \cdots, J^{(q)}, \tag{13}$$

$$b_j^{(q)} = f^{(q)}\left(a_j^{(q)}\right), \tag{14}$$

where $J^{(q-1)}$ and $J^{(q)}$ are the number of neurons of layer $(q-1)$ and q, respectively.

The nonlinear activation functions may be the Sigmoid function $f_S(n) = \frac{1}{1+e^{-n}}$, or Rectified Linear Unit (ReLU) function $f_R(n) = \max(0, n)$ [30]. Hence, the final output of the DNN $\hat{\mathbf{b}}^{(Q)} = \left[\hat{b}_1^{(Q)}, \cdots, \hat{b}_j^{(Q)}, \cdots, \hat{b}_{J^{(Q)}}^{(Q)}\right]^T$ is a cascade nonlinear transformation of input $\mathbf{b}^{(1)} = \left[b_1^{(1)}, \cdots, b_j^{(1)}, \cdots, b_{J^{(1)}}^{(1)}\right]^T$, and can be expressed as

$$\begin{aligned}\hat{\mathbf{b}}^{(Q)} &= f\left(\mathbf{b}^{(1)}, \cdots, \mathbf{b}^{(q)}, \cdots, \mathbf{b}^{(Q-1)}; \mathbf{u}^{(1)}, \cdots, \mathbf{u}^{(q)}, \right. \\ &\quad \left. \cdots, \mathbf{u}^{(Q-1)}; \mathbf{v}^{(1)}, \cdots, \mathbf{v}^{(q)}, \cdots, \mathbf{v}^{(Q-1)}\right) \\ &= f^{(Q-1)}\left(f^{(Q-2)}\left(\cdots f^{(1)}\left(\mathbf{b}^{(1)}\right)\right)\right),\end{aligned} \tag{15}$$

where $f^{(q)}(\cdot)$ is the activation function adopted by layer q. Vectors $\mathbf{u}^{(q)}$ and $\mathbf{v}^{(q)}$ denote the weights and bias at layer q. Through the offline or online training process, the weights and bias can be optimized following a target function.

B. Proposed DL-based Receiver

In a time-invariant channel, a DL-based receiver usually comprises two stages: 1) offline training stage, and 2) online deployment/test stage [23]. However, this type of training and deployment mode is not suitable for the DL-based receiver over time-varying channels.

In order to accommodate the time variability of UWA channels, we propose a DL-based receiver with online training and test mode. As shown in Fig. 5(a), the proposed receiver is alternatively working at two modes: 1) online training mode, and 2) online test mode. As shown in Fig. 5(b), the whole payload is divided into D sub-blocks with N_s symbols in each. For the i-th sub-block, the first $\{N_p^i\}_{i=1}^D$ symbols are utilized as the training symbols and the remaining $N_d^i = N_s - N_p^i$ symbols are the data symbols. So the number of total training symbols is $N_p = \sum_{i=1}^D N_p^i$. The resulting training overhead is $\beta = N_p/N_s/D$.

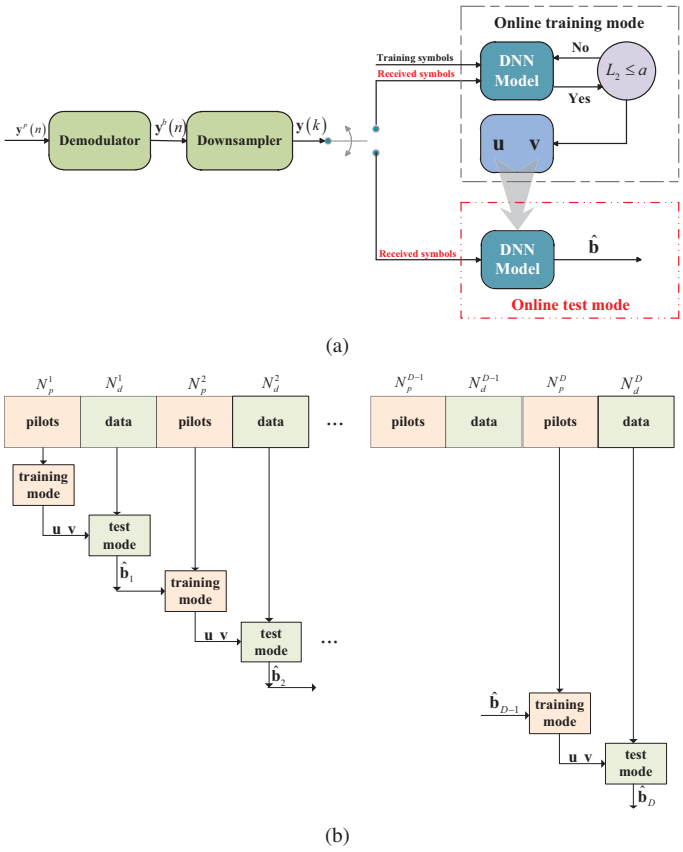


Fig. 5. Structure of the proposed DL-based receiver for time-varying UWA channels: (a) DL-based receiver, and (b) alternatively working between the training mode and test mode.

When the receiver is receiving known training symbols, the DL-based receiver switches to the online training mode. Given the received symbols $\mathbf{y}(k)$, the known training symbols $\mathbf{x}(k)$ and estimated symbols of previous sub-block, the DNN utilizes the Adam (Adaptive moment estimation) optimizer which is based on the stochastic gradient descent algorithm [30] to minimize the mean-square error of the loss function given by

$$L_2^i = \frac{1}{N_p^i} \sum_{k=0}^{N_p^i - 1} \left(\hat{\mathbf{b}}_i(k) - \mathbf{b}_i(k)\right)^2, i = 1, 2, \cdots, D. \tag{16}$$

The online training mode is stopped if L_2^i becomes lower than a predefined threshold a as shown in Fig. 5(a) or if $k = N_p - 1$. When the training mode ends, the DNN produces the weight set $\mathbf{u} \triangleq \left\{\mathbf{u}^{(q)}\right\}_{q=1}^{Q-1}$ and bias set $\mathbf{v} \triangleq \left\{\mathbf{v}^{(q)}\right\}_{q=1}^{Q-1}$, which are utilized in the online test mode.

The DL-based receiver switches into the online test mode after obtaining the weights vector $\mathbf{u}$ and bias vector $\mathbf{v}$. In the online test mode, we obtain the estimate $\{\hat{\mathbf{b}}^i\}_{i=1}^D$ of the transmitted symbols $\{\mathbf{b}^i\}_{i=1}^D$ by using the equation (15).

IV. Simulation Results

In this section, we evaluate the performance of the proposed DL-based receiver and compare it to the traditional DFE receiver with the recursive least squares (RLS) based channel estimator and embedded second-order PLL by using simulated time-varying channels.

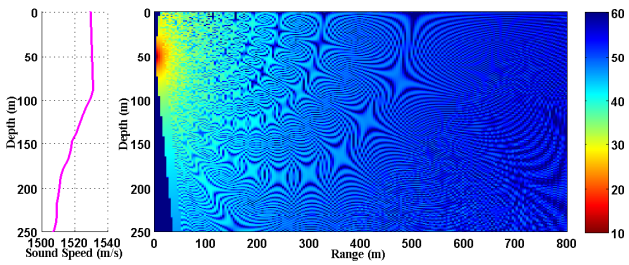


Fig. 6. Measured SSP and predicted TL with a source at 50 m depth. The transmission loss is shown in dB.

TABLE I
SIMULATION PARAMETERS SETUP

sea depth (H)	250 m
channel distance (d)	800 m
spreading factor (k)	1.7
speed of sound in bottom (c_b)	1574 m/s
density in bottom (ρ_b)	1.269 g/m^3
attenuation coefficient in bottom (α)	0.01875 dB/wavelength
minimum frequency (f_l)	8 kHz
carrier frequency (f_c)	12 kHz
bandwidth (B)	8 kHz
modulation type	BPSK
sample frequency (f_s)	48 kHz
symbol rate (R_s)	4 kHz

A. Simulation Environment

A statistical channel simulator [31] is used to generate the time-varying UWA channels for evaluating the performance of the proposed DL-based receiver and the traditional CE-based DFE receiver. The parameters for the time-varying channel simulator are listed in Table I. We use a SSP measured in a sea experiment and shown in Fig. 6. We also compute the transmission loss (TL) as shown in Fig. 6, the computation is done using the Bellhop acoustic toolbox [31]. The maximum TL is approximately 60 dB. In order to investigate how the surface and bottom affect the UWA channel characteristics, we construct three configurations of transducer-hydrophone pair listed in Table II.

Following the parameters and configurations shown in Table I and Table II, we obtain the time-varying UWA CIRs and corresponding channel scattering functions as shown in Fig. 7, Fig. 8, and Fig. 9. It can be seen that the simulated channels are time varying.

B. Training Scheme

For short packet bursts, we adopt the same training scheme for both the proposed DL-based receiver and the traditional CE-based DFE receiver as shown in Fig. 10. For each received packet with $N_s = 500$ symbols, the first N_p symbols are utilized as the training symbols and the remaining $N_d = N_s - N_p$ symbols are the data symbols. The resulting training overhead is $\beta = N_p/(N_p + N_d)$, and the corresponding data rate is $(1-\beta) \times R_s \ kbps$.

For the proposed DL-based receiver, the DNN has $Q = 4$ layers, in which $J^{(1)}, J^{(2)}, J^{(3)}$ and $J^{(4)}$ are set to 128, 96, 48, and 32, respectively. All layers utilize the sigmoid function as the active functions. The number of pilot symbols N_p is set to 64, then the resulting training overhead is $\beta = 12.8\%$.

TABLE II
DEPTH OF THE TRANSDUCER-HYDROPHONE PAIR

Configuration	Transmitter Depth (m)	Receiver Depth (m)
C1	50	50
C2	200	50
C3	200	200

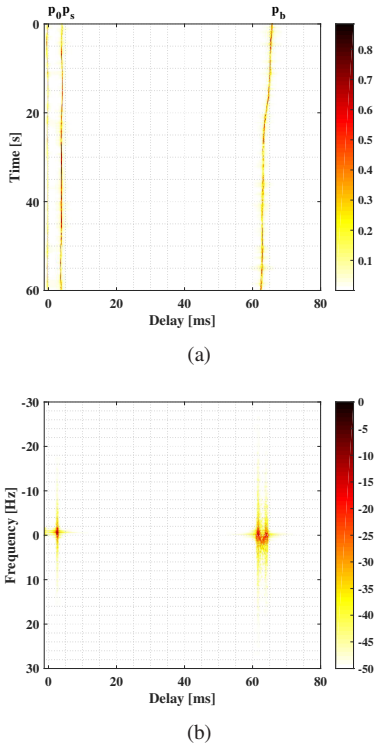
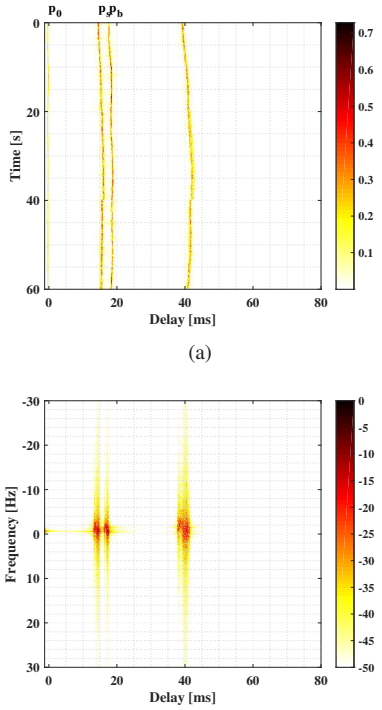


Fig. 7. Simulated time-varying channel characteristics under the configuration C1: (a) time-varying CIRs, and (b) corresponding channel scattering function.

For the traditional CE-based DFE receiver as shown in Fig. 2, we set N_p to 200 for ensuring that the traditional CE-based DFE receiver can get good performance. The length of the feed forward filter and feedback filter are set according to the CIRs shown in Fig. 7, Fig. 8, and Fig. 9, but the following parameters are common for all the three configurations: forgetting factor λ of the RLS adaptive algorithm is set to 0.995, proportional coefficient and integral coefficient of PLL, i.e. K_{f_1} and K_{f_2}, are set to 0.001 and 0.000001, respectively.

C. Test Results

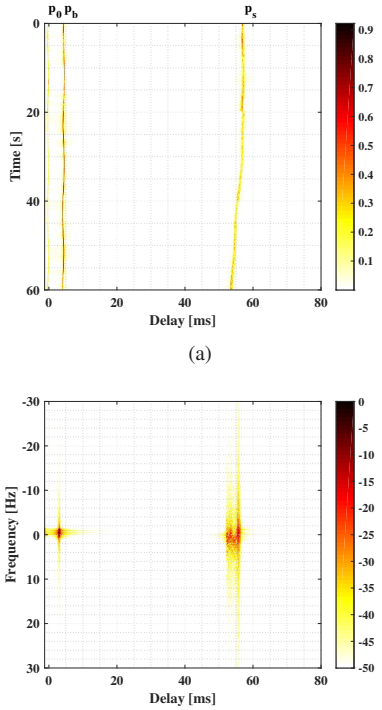
Following the above three simulation configurations, we compare the performance of the proposed DL-based receiver with that of the traditional CE-based DFE receiver in terms of bit error rate (BER). As shown in Fig. 11(a), Fig. 11(b), and Fig. 11(c), with lower training overhead, the proposed DL-based receiver consistently outperforms the traditional CE-based DFE receiver.



(a)

(b)

Fig. 8. Simulated time-varying channel characteristics under the configuration C2: (a) time-varying CIRs, and (b) corresponding channel scattering function.

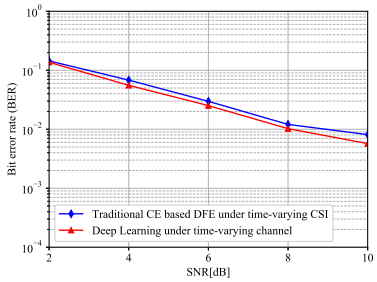


(a)

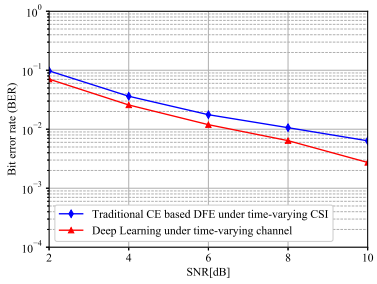
(b)

Fig. 9. Simulated time-varying channel characteristics under the configuration C3: (a) time-varying CIRs, and (b) corresponding channel scattering function.

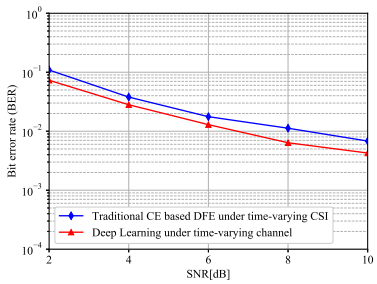
Fig. 10. Training scheme for the proposed DNN-based receiver and the traditional CE-based DFE receiver.



(a)



(b)



(c)

Fig. 11. BER performance of the proposed DL-based receiver and the traditional CE-based DFE receiver under the configuration: (a) C1, (b) C2, and (c) C3.

V. Experimental Results

A. Experimental Environment

The experiment was carried out in the South China Sea in November 2014. Fig. 12 depicts the layout of this experiment. The sea depth at the experimental site is about 99 m. One transducer was deployed to a depth of approximately 15 m from a ship. During the experiment, the ship was drifting on the sea surface. A receive vertical linear array of 48 hydrophones was moored with the first hydrophone at about 72 m below the surface, and other hydrophones evenly spaced by 0.25 m. The communication range was about 8 km at the start of the experiment. The system timers of transmitter and received array are synchronized by the GPS time before deployment.

B. Data Structure

For transmission, the input bits were encoded by a rate $R_c = 1/2$ convolutional coder with generator polynomial [171, 133] in octal format. The carrier frequency was $f_c = 3$

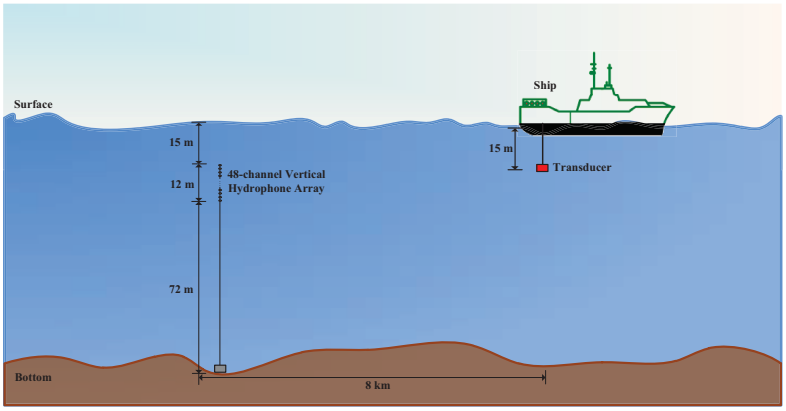


Fig. 12. Layout of the South China Sea experiment.

kHz and the symbol rate was 1 k symbols per second (ksps). The pulse shaping filter was a square-root raised cosine filter with a roll-off factor of 0.5 [27], leading to an occupied channel bandwidth of about 1.5 kHz. The sampling rate was 25 kHz at the receiver end. The structure of the transmitted data stream and relevant parameters are shown in Fig. 13. Preamble up-chirp and postamble down-chirp, Doppler-insensitive waveforms, were added before and after the data burst for coarse frame synchronization and estimation of an average Doppler shift over the whole data burst. Following the frame synchronization signal is one data packet (payload). Only data with BPSK modulation was used for performance evaluation for the proposed DL-based receiver. The payload is separated from up-chirp or down-chirp signal by the gap of the duration 150 ms for avoiding the inter-block interference. The length of each payload is 5500 symbols between two guard intervals. Each burst packet is transmitted every 6.1 s. The approximate SNR, which is estimated by using the signal part and silent part of the received signal, is in the range of 15 dB to 16 dB.

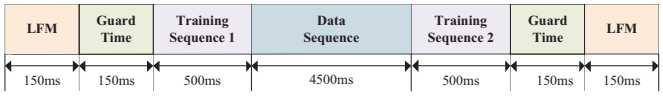


Fig. 13. Structure of the signal transmitted in the sea experiment.

In order to show characteristics of the UWA channel during the experiment, we estimate the CIRs by using the matched filter applied to the preamble and postamble chirp signals and the RLS algorithm applied to the data signals. It can be seen that the channel is fast time varying within a single burst. From Fig. 14, we can observe that the channel multipath spread is about $15 \sim 30$ ms, corresponding to a channel length of $15 \sim 30$ taps in terms of the symbol rate $R_s = 1$ ksps. The arrival paths fluctuate very rapidly and CIRs are clustering.

C. Training Scheme

In order to evaluate the performance of the proposed DL-based receiver and traditional CE-based DFE receiver, we use 6 transmitted bursts. For each burst, we have 48 received packets, so in total we have 288 received packets. We choose 5000 symbols including training sequence and data sequence in each packet depicted in Fig. 13 to test the performance of above receivers.

For the proposed DL-based receiver, the DNN has $Q = 5$ layers, in which $J^{(1)}, J^{(2)}, J^{(3)}, J^{(4)}$ and $J^{(5)}$ are set to 96, 48, 16, 8, and 1, respectively. All layers utilize the sigmoid function as the active functions. Training symbols are periodically inserted into the data to train the DNN. As shown in Fig. 15(a), the whole payload is divided into $D = 10$ sub-blocks with $N_s = 500$ symbols in each. For the i-th sub-block, the first N_p^i symbols are utilized as the training symbols and the remaining $N_d = N_s - N_p^i$ symbols are the data symbols. So the number of total training symbols is $N_p = \sum_{i=1}^{10} N_p^i$. The resulting training overhead is $\beta = N_p/500/10$.

For the traditional CE-based DFE receiver, as shown in Fig. 15(b), we follow the training scheme usually used in this receiver, then with $N_s = 5000$, the first N_p symbols are utilized as the training symbols and $N_d = N_s - N_p$ symbols are the data symbols. The length of the feed forward filter and feedback filter are set to 60 and 29 according to the channel characteristic analysis. Forgetting factor is 0.995, proportional coefficient and integral coefficient of the PLL loop filter are set to 0.001 and 0.000001, respectively.

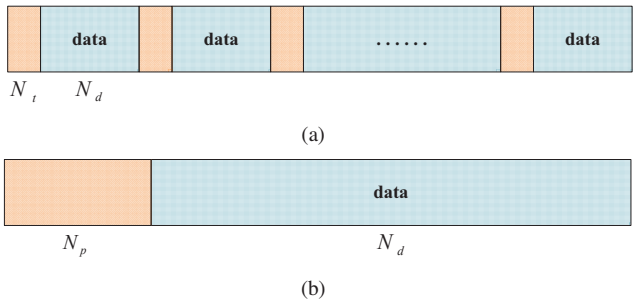
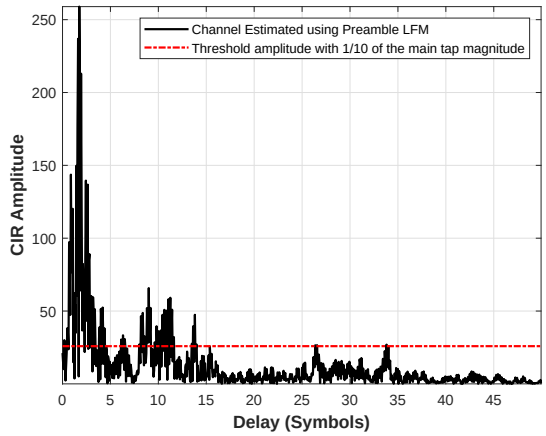
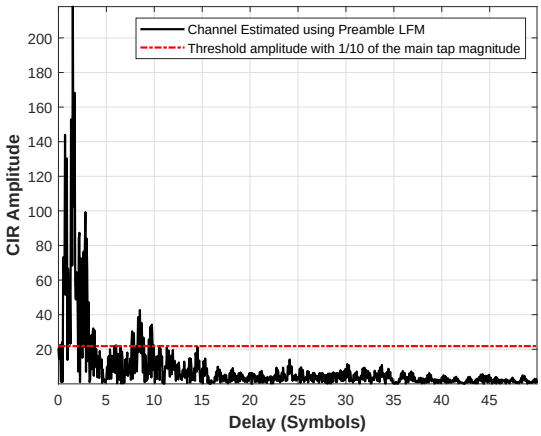


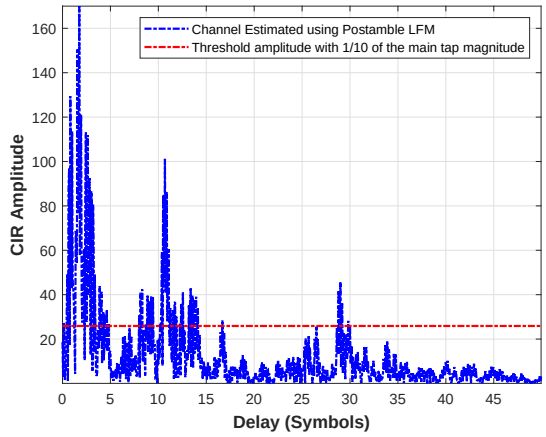
Fig. 15. Training schemes for (a) proposed DNN-based receiver; (b) traditional CE-based DFE receiver.



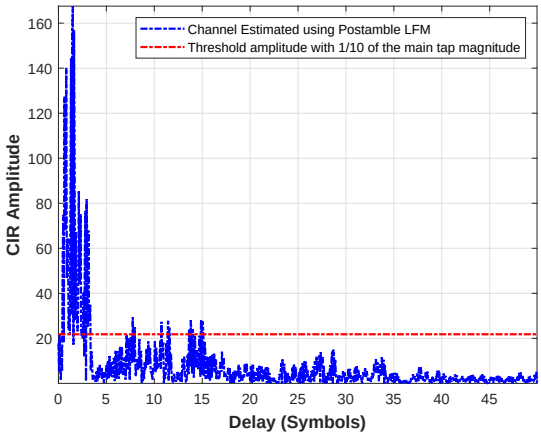
(a)



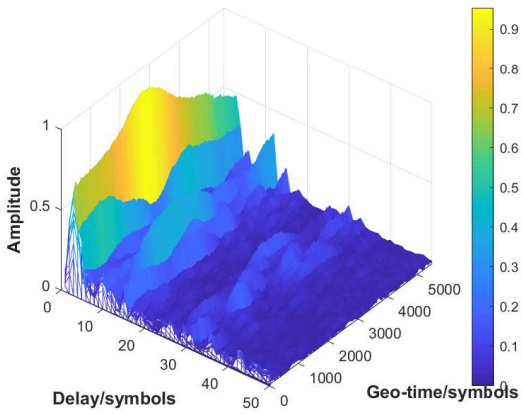
(b)



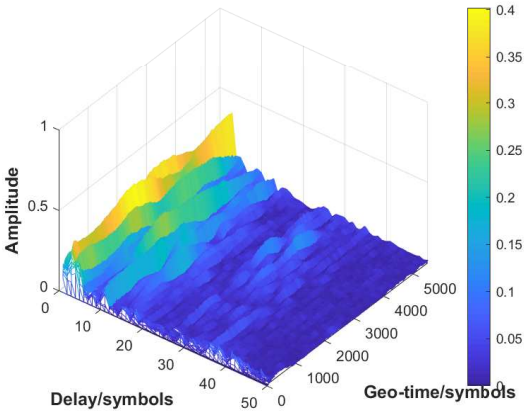
(c)



(d)



(e)



(f)

Fig. 14. Examples of the CIR estimated over one burst transmission. The CIRs measured between transducer and first hydrophone are shown in (a), (c), (e). The CIRs measured between transducer and last hydrophone are shown in (b), (d), (f). CIR is measured using: (a) and (b) the preamble up-chirp with the correlation method; (c) and (d) the postamble down-chirp with the correlation method; (e) and (f) data signals and the classical RLS algorithm with $\lambda = 0.995$.

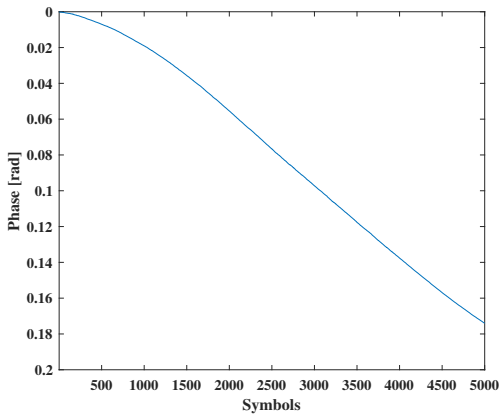


Fig. 16. Phase estimated by PLL at SNR $\approx$ 15 dB.

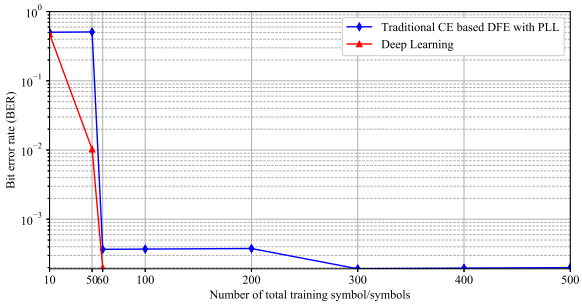


Fig. 17. BER versus the training overhead at SNR $\approx$ 15 dB.

D. Test Results

Fig. 16 shows the time-varying phase estimated by using the PLL for one received signal packet. Fig. 17 shows how the performance of receivers is effected by the training overhead. We observe that the total number of training symbols N_p significantly affects the performance of the traditional CE-based DFE receiver. It can be seen that, when N_p is less than 50, the traditional CE-based DFE receiver cannot converge. The performance of the proposed DL-based receiver can be improved with a few training symbols per subblock. When $N_p^i = 5$ for the i-th subblock, i.e. the total number of training symbols $N_p = 10 \times 5 = 50$, the BER can reach 10^{-2}. With $N_p^i = 6$, the proposed DL-based receiver can reach zero BER for all 288 received packets. The traditional CE-based DFE receiver needs at least 60 training symbols to converge to a BER below 3.7×10^{-4}.

For the traditional CE-based DFE receiver, the improvement in BER performance is small with the increase of training symbols. The error free transmission cannot be achieved even with a training overhead $\beta = 300/5000 = 6\%$, while the proposed DL-based receiver only needs 60 pilot symbols to achieve the error-free transmission for all 288 packets with a training overhead as low as $\beta = 6 \times 10/5000 = 1.2\%$.

Since the data were originally acquired in a relatively high SNR, we can evaluate the performance of the receivers over different noise levels by adding recorded noise into the received data. Fig. 18 and Fig. 19 show how the performance

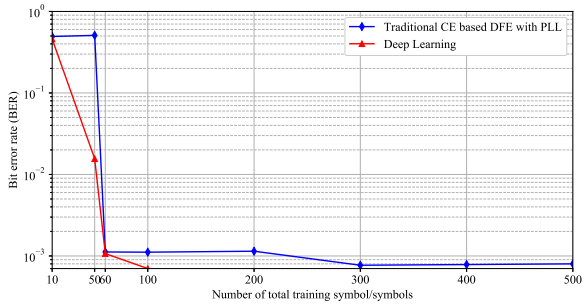


Fig. 18. BER versus the training overhead at SNR $\approx$ 10 dB.

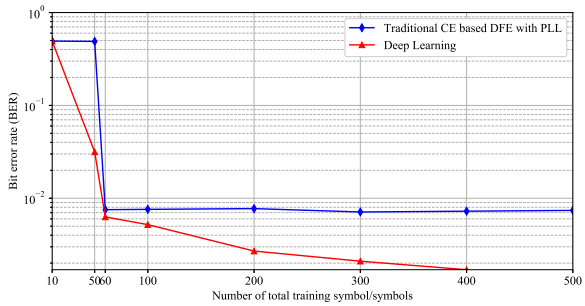


Fig. 19. BER versus the training overhead at SNR $\approx$ 5 dB.

of the receivers is effected by the lower SNR.

Fig. 18 shows the performance at an SNR of 10 dB. It can be seen that, when N_p is less than 50, the traditional CE-based DFE receiver cannot converge. With $N_p = 60$ BER can only reach BER=1.1×10^{-3}. With the proposed DL-based receiver, when the $N_p^i = 5$ for each subblock, i.e. the total training number is 50, all 288 packets can be received without errors.

Fig. 19 shows the receiver performance at SNR= 5 dB. It can be seen that, with the the lower SNR, the performance of the two receivers degrades. For the traditional CE-based DFE receiver an error floor is BER=7.5×10^{-3}. The proposed DL-based receiver, with $N_p = 400$, can reach zero BER with a training overhead $\beta = 400/5000 = 8\%$.

VI. Conclusions

In this paper, we have proposed a DL-based receiver for UWA SC communications over time-varying UWA channels. Unlike the DL-based receivers over time-invariant channel, the proposed receiver works with the online training stage and online test stage for accommodating the time variability of UWA channels. Simulation results show that the proposed receiver outperforms the traditional CE-based DFE receiver even if using a significantly shorter training sequence. The proposed receiver has also been tested using sea trial data recorder at a communication range of 8 km. The performance of the receiver is evaluated for various training overheads and SNRs. Experimental results demonstrate that the proposed DL-based receiver achieves error-free transmission at all SNR conditions with lower training overhead compared to the traditional CE-based DFE receiver, which cannot provide the error-free transmission.

ACKNOWLEDGMENT

The work of Y. Zhang was supported in part by the National Natural Science Foundation of China under Grant 61471138, Grant 61531012, and Grant 50909029, in part by the China Scholarship Council Funding, in part by the Program of International Science and Technology Cooperation under Grant 2013DFR20050, in part by the Defense Industrial Technology Development Program under Grant B2420132004, and in part by the Acoustic Science and Technology Laboratory in 2014. The work of Y. Zakharov is partly supported by the U.K. Engineering and Physical Sciences Research Council under Grant EP/P017975/1 and Grant EP/R003297/1. The work of C. Lin is partly Supported by the Fund of Acoustics Science and Technology Laboratory.

REFERENCES

[1] M. Stojanovic, J. Catipovic, and J. Proakis, "Phase-coherent digital communications for underwater acoustic channels," *IEEE Journal of Oceanic Engineering*, vol. 19, no. 1, pp. 100-111, Jan. 1994.

[2] T. H. Eggen, A. B. Baggeroer, and J. C. Preisig, "Communication over Doppler spread channels. Part I: Channel and receiver presentation," *IEEE J. Ocean. Eng.*, vol. 25, no. 1, pp. 62-71, Jan. 2000.

[3] T. H. Eggen, A. B. Baggeroer, and J. C. Preisig, "Communication over Doppler spread channels. Part II: Receiver characterization and practical results," *IEEE J. Ocean. Eng.*, vol. 26, no. 4, pp. 612-621, Oct. 2001.

[4] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 84-89, Jan. 2009.

[5] J. Li and Y. V. Zakharov, "Multibranch autocorrelation method for Doppler estimation in underwater acoustic channels," *IEEE J. Ocean. Eng.*, vol. 43, no. 4, pp. 1099-1113, Oct. 2018.

[6] J. Li and Y. V. Zakharov, "Efficient use of space-time clustering for underwater acoustic communications," *IEEE J. Ocean. Eng.*, vol. 43, no. 1, pp. 173-183, Jan. 2018.

[7] Y. Li, Z. Jiang, W. Shi, X. Han, and B. Chen, "Blocked Maximum Correntropy Criterion Algorithm for Cluster-Sparse System Identifications," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. PP, no. PP, pp. 1-1, Jan. 2019.

[8] G. Han, J. Jiang, N. Bao, L. Wan, and M. Guizani, "Routing protocols for underwater wireless sensor networks," *IEEE Communications Magazine*, vol. 53, no. 11, pp. 72-78, 2015.

[9] G. Han, J. Jiang, N. Sun, and L. Shu, "Secure communication for underwater acoustic sensor networks," *IEEE Communications Magazine*, vol. 53, no. 8, pp. 54-60, 2015.

[10] G. Han, S. Shen, H. Song, T. Yang, and W. Zhang, "A stratification-based data collection scheme in underwater acoustic sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10671-10682, 2018.

[11] Y. Zhang, Y. Zakharov, J. Li, "Soft-decision-driven sparse channel estimation and turbo equalization for MIMO underwater acoustic communications," *IEEE Access*, vol. 6, pp. 4955-4973, 2018.

[12] T. Arikan, T. Riedl, A. Singer and J. Younce, "Comparison of OFDM and single-carrier schemes for Doppler tolerant acoustic communications," in *OCEANS 2015 - Genova*, Genoa, Italy, May 2015, pp. 1-7.

[13] M. Johnson, L. Freitag, and M. Stojanovic, "Improved Doppler tracking and correction for underwater acoustic communications," in *Proceedings of the 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97)*, Munich, Germany, vol. 1, April 1997, pp. 575-578.

[14] M. Johnson, L. Freitag, and M. Stojanovic, "Efficient equalizer update algorithms for acoustic communication channels of varying complexity," in *Proc. IEEE Oceans'97 Conference*, Halifax, Nova Scotia, Canada, October 1997, pp. 580-585.

[15] M. Stojanovic, L. Freitag, and M. Johnson, "Channel-estimation-based adaptive equalization of underwater acoustic signals," in *Proc. MTS/IEEE OCEANS*, vol. 2, pp. 590-595, Sep. 1999.

[16] T. Chen, Y. V. Zakharov, C. Liu, "Low-complexity channel-estimate based adaptive linear equalizer," *IEEE Signal Processing Letters*, vol. 18, no. 7, pp. 427-430, 2011.

[17] R. M. Mehmood, R. Du, and H.-J. Lee, "Optimal feature selection and deep learning ensembles method for emotion recognition from human brain EEG sensors," *IEEE Access*, vol. 5, pp. 14797-14806, July 2017.

[18] W. Liu et al., "An ensemble deep learning method for vehicle type classification on visual traffic surveillance sensors," *IEEE Access*, vol. 5, pp. 24417-24425, Oct. 2017.

[19] X. L. Zhang and D. Wang, "A deep ensemble learning method for monaural speech separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 5, pp. 967-977, Mar. 2016.

[20] P.J. Werbos, "Computational intelligence for the smart grid-history, challenges, and opportunities," *IEEE Comput. Intell. Mag.*, vol. 6, no. 3, pp. 14-21, 2011.

[21] C.-W. Tsai, C.-F. Lai, M.-C. Chiang, L. T. Yang, et al., "Data mining for internet of things: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 77-97, 2014.

[22] T. Wang, C. K. Wen, H. Wang, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges," *China Communications*, vol. 14, no. 11, pp. 92-111, Nov. 2017.

[23] H. Ye, G. Y. Li, and B. H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114-117, Feb. 2018.

[24] H. He, C. K. Wen, S. Jin, and Y. Li, "Deep learning based channel estimation for beamspace mmWave massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 852-855, May 2018.

[25] X. Ma, H. Ye, and Y. Li, "Learning assisted estimation for time-varying channels," in *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, Lisbon, Portugal, pp. 1-5, Aug. 2018.

[26] Y. Zhang, J. Li, Y. Zakharov, D. Sun, and J. Li, "Underwater acoustic OFDM communications using deep learning," in *the 2nd Franco-Chinese Acoustic Conference (FCAC)*, 29-31 Oct. 2018, Le Mans, France, pp. 1-6.

[27] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw Hill, 2000.

[28] T. Chen, "Novel adaptive signal processing techniques for underwater acoustic communications," *University of York*, 2011.

[29] M. Stojanovic, J. Catipovic, and J. Proakis, "An algorithm for multichannel coherent digital communications over long range underwater acoustic telemetry channels," in *Proc. IEEE OCEANS'92 Conference*, Newport, RI, USA, October 1992, vol. 2, pp. 577-582.

[30] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Massachusetts, USA: MIT Press, 2016.

[31] P. Qarabaqi, M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 4, pp. 701-717, 2013.