Received November 27, 2020, accepted December 12, 2020, date of publication December 18, 2020, date of current version December 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3045706

Weighted TSVR Based Nonlinear Channel Frequency Response Estimation for MIMO-OFDM System

SHUO LI[1,2], WEIMIN LEI[1], (Member, IEEE), WEI ZHANG[1], XIAOYU WANG[2], AND LIDONG WANG[2]

[1]School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China

[2]School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan 114051, China

Corresponding author: Xiaoyu Wang (xiaoyuwxy@126.com)

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1702000, in part by the Fundamental Research Funds for the Central Universities of China under Grant N2016014 and Grant N2024005-1, and in part by the National Natural Science Foundation of China under Grant 61671141.

ABSTRACT A channel frequency response estimation method based on weighted twin support vector regression (TSVR) for pilot-aided multi-input multi-output (MIMO) orthogonal frequency division multiplexing (OFDM) system is proposed in this work. Nonlinearity of channel in wireless communication system is considered. The channel is fading in time domain produced by Doppler effect and in frequency domain by propagation multipath. An improved TSVR-weighted TSVR is adopted to estimate channel parameters in MIMO-OFDM system. The weights obtained by wavelet transform method are used to improve the regression performance of TSVR. The characteristic of the proposed algorithm is that different training samples are given weights calculated according to the distance from the samples to the mean values filtered by wavelet transform method. Due to the regression characteristics of TSVR, the channel frequency response estimation algorithm proposed in this work has good estimation performance and anti noise ability. Experimental results show that compared with the classical pilot aided channel estimation method, the proposed algorithm has better performance in estimating mean square error.

INDEX TERMS Channel estimation, wavelet transform, MIMO, OFDM, TSVR.

I. INTRODUCTION

High data rate wireless communication is a research hotspot in the emerging wireless local area networks, home audio/visual network and the fifth generation (5G) wireless cellular systems. However, due to the limitation of peak to average power ratio (PAPR) and signal-to-noise ratio (SNR) of the actual receiver, it is usually not feasible to use traditional high data rate wireless links in Non-line-of-sight (NLOS) networks. Additionally, there is a serious range penalty to be paid for high bandwidth systems. Multi-input multi-output (MIMO) wireless constitutes a technological breakthrough that will allow Gb/s speeds in NLOS wireless networks. The use of MIMO systems can improve the performance due to array gain, diversity gain, spatial multiplexing gain, and interference reduction. Orthogonal frequency division multiplexing (OFDM) is a kind of multi-carrier technique. It has good anti multipath fading ability and can support multi-user access. In OFDM system, the channel is divided into orthogonal subchannels, and high-speed serial data is converted into low-speed parallel data for transmission. Orthogonal signals can be separated by coherent technique at the receiver, which can reduce the inter channel interference (ISI). The signal bandwidth of each subchannel is less than the coherent bandwidth of the channel, hence, it can be seen as flat fading on each subchannel and the ISI can be eliminated. Due to the advantages of the MIMO and OFDM technologies, the MIMO-OFDM is adopted by the 4G and 5G wireless communication systems [1].

At receiving end of wireless communication, the performance of coherent demodulation is better than that of non-coherent demodulation under the same signal-to-noise ratio because of taking more channel information [2]. Therefore, channel estimation is very important to the performance of

The associate editor coordinating the review of this manuscript and approving it for publication was Eyuphan Bulut.

high quality information transmission. However, due to the frequency selectivity caused by multipath propagation and the time selectivity caused by Doppler effect, channel estimation becomes a very difficult task. OFDM system is sensitive to inter carrier interference (ICI), which is caused by mobility, and the carrier frequency drift and the increase of mobile speed can further worsen the transmission performance [3], [4]. To obtain effective communication in the time and frequency selective cases (doubly selective), a well-designed channel estimator is an important prerequisite.

Many channel estimation methods were given in the existing literature, such as Least square (LS) estimation method, linear minimum mean square error (LMMSE) and its improved algorithm [5]–[7]. For the doubly selective channel, the above estimation methods can reduce the transmission efficiency due to the large pilot ratio. In order to solve this problem, serval algorithms were proposed. Basis expansion model (BEM) for doubly selective channel was established and algorithms for channel estimation were proposed in [12]–[14]. Zhang *et al.* [18] designed adaptive weighted estimators for OFDM systems. Jan [15] used iterative processing method to estimate the doubly selective channel. Support vector regression (SVR) method was also adopted to estimate the parameters of doubly selective channel [16], [17].

In the process of signal propagation, the channel shows a certain degree of nonlinearity caused by some reasons like saturation of components, dispersion of optical fiber and coupling between devices [19]–[21], especially in the case of doubly selective channel. The nonlinearity can degrade the estimation performance if it is estimated under a linear assumption [17]. Therefore, some nonlinear methods were used for nonlinear channel estimation. A nonlinear Kalman filter-based high-speed channel estimation algorithm for OFDM systems was proposed in [22]. Yang *et al.* [23] proposed a channel estimator based on deep learning method. While the advantages of SVR are that it is suitable for regression of nonlinear systems and the number of training samples is small [24], [25]. The kernel trick can be adopted to extend the linear SVR method to the field of nonlinear regression according to the theory of Vapnik-Chervonenkis (VC) [26]. It is a powerful tool for system regression [27], and can be used for wireless channel estimation. A MIMO channel estimation method was proposed in [28], while the channel of which was assumed to be flat in frequency domain. In [16] and [17], basic SVR based OFDM channel estimators were proposed for some application scenarios. Later, many improved SVR algorithms were proposed. The twin SVR (TSVR) was proposed in [29], ν-SVR was introduced in [30], an asymmetric ν -TSVR method for asymmetric noised data regression was proposed in [27], and Y. Shao *et al.* proposed an improved algorithm of TSVR by introducing the regularization criterion in [31]. These algorithms mentioned above have been improved to a certain extent in terms of computational complexity and/or performance.

In most of the literature, the training data share the same weights, which are not conducive to noise suppression for the communication signal polluted by noise, and the performance will be reduced. In order to solve this problem, a wavelet transform based weighted TSVR (WTWTSVR) was proposed in our work [32]. In the propose algorithm, the weights calculated by data variance are given to the training data and the regression performance can be improved.

Inspired by this idea, an improved nonlinear channel frequency response estimation in MIMO-OFDM systems based on WTWTSVR is proposed in this paper. The contributions of this work are described as: the novel improved TSVR based channel frequency response estimator is proposed for nonlinear parameter estimation in MIMO-OFDM systems. The proposed algorithm outperforms traditional algorithms such as linear interpolation, LMMSE, BEM based estimation method and classic TSVR based algorithm. Moreover, wavelet transform is used to preprocess the sample data. The weights of the samples are obtained by calculating the distance between the sample points and the mean value obtained by wavelet transform. Then they are assigned to the sample points in objective functions to make use of the prior information of data and reduce the impact of noise and outliers. Due to the nature of wavelet filter, this method has advantages in processing time series signal such as communication signal.

The outline of this paper is as follows: Section II dwells on the MIMO-OFDM system model and Section III introduces space time coding. WTWTSVR based channel frequency response estimation algorithm is proposed in Section IV. Simulation results are demonstrated in Section V to exhibit the performance of the proposed estimator. Finally, Section VI summarizes the paper and prospects the future work.

II. MIMO-OFDM SYSTEM MODEL

Consider a MIMO-OFDM system with M_T transmit antennas and M_R receive antennas (Figure 1). The sequence to be transmitted $X(k)$, which get from QPSK or QAM constellation is sent to a space-time encoder. Then the output sequence $X_i(k)$ $(i = 1, 2, \ldots, M_T)$ is parsed into blocks of N symbols and transformed into a time-domain sequence using an N-point inverse discrete Fourier transform (IDFT). To avoid inter-block interference (IBI), a cyclic prefix (CP) of length L_C equal to or larger than the channel order L is inserted at the head of each block. Denote $x_i(n)$ as the time-domain signal of the i-th transmit antenna, which can be described as:

$$x_i(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_i(k) e^{j2\pi nk/N} \tag{1}$$

where $n = -L_C, \ldots, N-1$, $k = 0, \ldots, N-1$. Then $x_i(n)$ can be passed over the doubly selective wireless channel. At the receiver, the cyclic prefix will be removed, and the received signal vector of M_R receive antennas in time domain

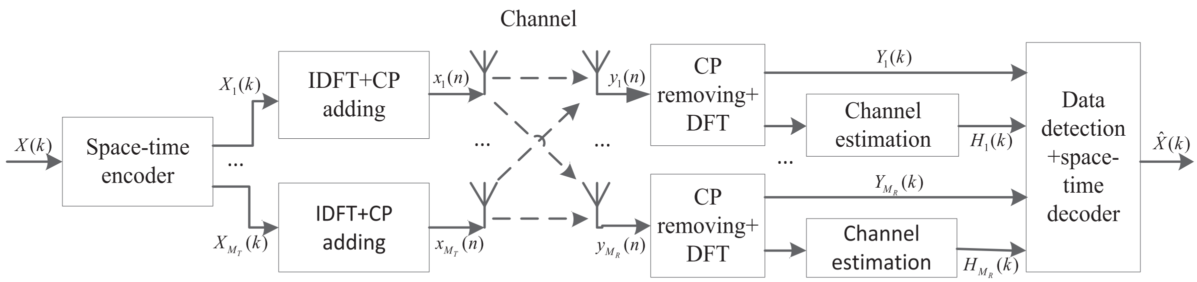


FIGURE 1. MIMO-OFDM system model.

$\mathbf{y}(n) \in \mathbb{C}^{M_R}$ can be expressed as

$$\mathbf{y}(n) = \sum_{i=1}^{M_T} \sum_{l=0}^{L-1} \mathbf{h}_i(n,l)x_i(n-l) + \mathbf{v}(n) \tag{2}$$

where $\mathbf{v}(n) = [v_1(n), v_2(n), \ldots, v_{M_R}(n)]^T$ is additive white Gaussian noise (AWGN). The expectation and variance value of $v_i(n)$ are 0 and σ_n^2, respectively, and $\mathbf{v}(n)$ is independent, i.e. $E(\mathbf{v}^T(m_1)\mathbf{v}(m_2)) = 0, \forall m_1 \neq m_2$. $\mathbf{h}_i(n,l) = [h_{i1}(n,l), h_{i2}(n,l), \ldots, h_{iM_R}(n,l)]^T$ is the channel response vector from the ith transmit antenna. Denoting $\mathbf{Y}(k) \in \mathbb{C}^{M_R}$ as the frequency response vector of the received sequence vector after removing CP at the kth subcarrier, we have

$$\mathbf{Y}(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{y}(n)e^{-j2\pi nk/N} \tag{3}$$

Substituting (2) in (3), it can be obtained

$$\begin{aligned}\mathbf{Y}(k) &= \sum_{i=1}^{M_T}\sum_{r=0}^{N-1} X_i(r) \sum_{l=0}^{L-1} \mathbf{H}_i(r-k,l)e^{-j2\pi lr/N} \\ &\quad + \mathbf{V}(k) &(4)\\ &= \sum_{i=1}^{M_T} X_i(k) \sum_{l=0}^{L-1} \mathbf{H}_i(0,l)e^{-j2\pi kl/N} + \mathbf{I}(k) \\ &\quad + \mathbf{V}(k) &(5)\\ &= \sum_{i=1}^{M_T} X_i(k)\mathbf{H}_i(k) + \mathbf{I}(k) + \mathbf{V}(k) &(6)\end{aligned}$$

where

$$\mathbf{H}_i(m,l) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{h}_i(n,l)e^{-j2\pi mn/N} \tag{7}$$

and

$$\mathbf{I}(k) = \sum_{i=1}^{M_T} \sum_{\substack{r=0,\\ r\neq k}}^{N-1} X_i(r) \sum_{l=0}^{L-1} \mathbf{H}_i(r-k,l)e^{-j2\pi lr/N}. \tag{8}$$

$\mathbf{V}(k) \in \mathbb{C}^{M_R}$ is the frequency response of noise $\mathbf{v}(n)$ at subcarrier k, and $\mathbf{I}(k) \in \mathbb{C}^{M_R}$ is ICI induced by subcarriers around the subcarrier k. At the receiver, the doubly fading channel can be equalized and the transmitted symbols can be recovered based on channel information.

III. SPACE–TIME CODING TRANSMISSION

In order to obtain the diversity gain, the Alamouti space–time coding scheme can be adopted for data transmission. The simplest MIMO-OFDM system is with two transmit antennas and one receive antenna, which can be expressed by the following matrix.

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_e & -X_o^* \\ X_o & X_e^* \end{bmatrix} \tag{9}$$

where $(\cdot)^*$ is the complex conjugate operator, X_1 and X_2 represent the two transmitted signals by antennas 1 and 2, respectively and X_o and X_e denote the odd and even sequences of the signal to be transmitted X, respectively. Two different signals X_1 and X_2 can be transmitted from antennas 1 and 2, respectively. We assume that the channel response remains unchanged over consecutive symbol periods. The signals received by the receive antenna over consecutive symbol periods in the frequency domain can be expressed as

$$\begin{bmatrix} Y_e(k) \\ Y_o(k) \end{bmatrix} = \begin{bmatrix} X_e(k) & X_o(k) \\ -X_o^*(k) & X_e^*(k) \end{bmatrix} \cdot \begin{bmatrix} H_1(k) \\ H_2(k) \end{bmatrix} + \begin{bmatrix} V_e(k) \\ V_o(k) \end{bmatrix} \tag{10}$$

where $X_{e/o}(k)$, $Y_{e/o}(k)$, and $V_{e/o}(k)$ denote the transmitted, received signal and the noise at the k-th subcarrier of the even/odd sequences, and $H_{1/2}(k)$ denotes the channel response of the 1st/2nd transmit antenna to the receive antenna at the k-th subcarrier.

IV. MIMO-OFDM SYSTEM CHANNEL ESTIMATION

Due to the fading in time domain and frequency domain, wireless communication channels are nonlinear, while the classical channel estimation is based on the linear assumption. TSVR based estimation algorithm is adopted in this work for its nonlinear properties. In this section, a novel TSVR, weighted TSVR based on wavelet transform is used for estimating the MIMO-OFDM channel frequency response.

A. CHANNEL ESTIMATION AT PILOT SUBCARRIERS

In this work, the pilots are inserted in time domain and frequency domain uniformly. Phrases 'pilot symbol' and 'data symbol' denote OFDM symbol inserted into the pilot and that without pilot, respectively. 'Pilot subcarrier' and 'data subcarrier' denote the subcarrier being pilot in a pilot symbol and the subcarrier being data, respectively. In order to realize channel estimation in space-time coding scheme, pilot symbols appear in pairs in time domain. The position set of pilot symbols in time domain can be expressed as $\{n\Delta t, n\Delta t+1 | n = 0, \ldots, N_t/2-1\}$, where the even number $\Delta t \geq 2$ is the pilot symbol interval and N_t is the pilot symbols number. The position set of pilot subcarriers can be described as $\{m\Delta f | m = 0, \ldots, N_f - 1\}$, N_f and Δf are the pilot number in one OFDM symbol and the pilot frequency interval, respectively. Figure 2 demonstrates the pilot insertion scheme. The channel frequency response of pilot subcarrier positions can be estimated by inverting (10), i.e

$$\begin{bmatrix} \hat{H}_{1p} \\ \hat{H}_{2p} \end{bmatrix} = \begin{bmatrix} X_{ep} & X_{op} \\ -X_{op}^{*} & X_{ep}^{*} \end{bmatrix}^{-1} \cdot \begin{bmatrix} Y_{ep} \\ Y_{op} \end{bmatrix} \tag{11}$$

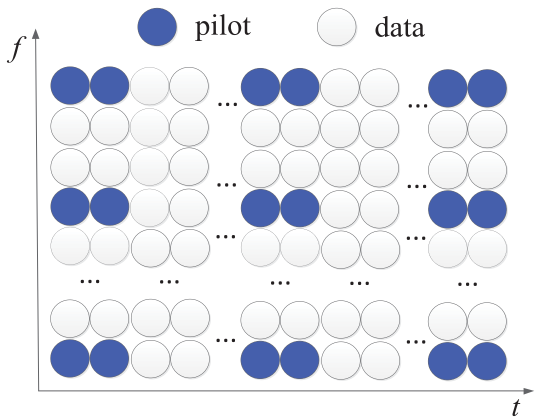


FIGURE 2. The pilot insertion scheme.

where $\hat{*}$ represents the estimation of $*$, and the subscribe p means the pilot, such as $\hat{H}_{1p} = \hat{H}_1(n\Delta t, m\Delta f)$ is the estimated frequency response at pilot positions $(n\Delta t, m\Delta f)$. All the estimated channel response $\hat{H}_{1p}$ $\left(\hat{H}_{2p}\right)$ at pilot positions of the channel from antenna 1 (antenna 2) to the receive antenna can be collected to form a matrix $\hat{\mathbf{H}}_{1P/2P}$.

Then the frequency response will be computed, and the estimated value of all subcarriers can be expressed as

$$\tilde{\mathbf{H}}_{1/2} = f(\hat{\mathbf{H}}_{1P/2P}) \tag{12}$$

B. CHANNEL ESTIMATION ALGORITHM

Define a training set $\mathbf{S} = \{(\tau_1, r_1), (\tau_2, r_2), \ldots, (\tau_m, r_m)\}$, where the position of pilot sample points in time domain and frequency domain is used as training input: $\tau_k \in \mathbb{R}^2$ and the impulse response of pilot sample points is used as training output: $r_k \in \mathbb{R}, k = 1, 2, \ldots, m$, which constitute a training input-output pair (τ_k, r_k), m is the samples number. The output vector can be expressed as $\mathbf{R} = (r_1, r_2, \ldots, r_m)^T \in \mathbb{R}^m$ and the training input as $\mathbf{T} = (\tau_1, \tau_2, \ldots, \tau_m)^T \in \mathbb{R}^{m\times 2}$. Define $\mathbf{e}$ and $\mathbf{I}$ as ones column vector and identity matrix, respectively.

In the wireless environment with relative mobility and multipath propagation, the channel shows selectivity in both time domain and frequency domain. Nonlinearity is the basic characteristic of doubly selective channel. The estimation of nonlinear channel using linear method will lead to large error. In this paper, we adopt the WTWTSVR to estimate the channel response in MIMO-OFDM system, since TSVR is superior in solving nonlinear regression or pattern recognition problems. Similar to classical TSVR, the WTWTSVR is composed by down-bound $f_1(\tau)$ and up-bound $f_2(\tau)$, and the regressor can be calculated by

$$f(\tau) = \frac{1}{2}(f_1(\tau) + f_2(\tau)). \tag{13}$$

The regressor (13) is a linear expression, it can be changed to nonlinear form by adopting kernel tricks [26], that is mapping the training samples into a higher-dimensional space: down and up bounds of the regressor are $f_1(\tau) = K(\tau, \mathbf{T}^T)\mathbf{g}_1 + \delta_1$ and $f_2(\tau) = K(\tau, \mathbf{T}^T)\mathbf{g}_2 + \delta_2$, respectively, where K is a kernel function, $\mathbf{g}_1, \mathbf{g}_2 \in \mathbb{R}^m$ and $\delta_1, \delta_2 \in \mathbb{R}$ are parameters to be estimated. The average of $f_1(\tau)$ and $f_2(\tau)$ is the final regression result, which shares the same expression as the linear one (13). The optimization problems can be expressed as (15) and (16) where $v_1, v_2, c_1, c_2, c_3, c_4 > 0$ are pre-selected parameters, $\varepsilon_1, \varepsilon_2$ are insensitive constants, ξ_1, ξ_2 are slack vectors reflecting whether the training samples locate in the ε tube or not [29]. $\mathbf{d} \in \mathbb{R}^m$ and $\mathbf{D} = diag(\mathbf{d}) \in \mathbb{R}^{m\times m}$ are weights, which will be calculated by wavelet filtering algorithm [32]. The weighting vector $\mathbf{d}\left(= [d_1, d_2, \ldots, d_m]^T\right)$ is computed by:

$$\mathbf{d} = Ae^{(-\left|\mathbf{R}-\hat{\mathbf{R}}\right|^2/\sigma^2)}, \tag{14}$$

where A and σ are amplitude and standard deviation of the Gaussian function, respectively. $\hat{\mathbf{R}}(= [\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_m]^T)$ represents the filtered value vector of output $\mathbf{R}(= [r_1, r_2, \ldots, r_m]^T)$. In the proposed algorithm, the $\hat{\mathbf{R}}$ can be calculated by the wavelet filter described by [32].

$$\begin{aligned} \min_{\mathbf{g}_1,\delta_1,\xi_1,\varepsilon_1} \; & \frac{1}{2}(\mathbf{R} - (K(\mathbf{T}, \mathbf{T}^T)\mathbf{g}_1 + \mathbf{e}\delta_1))^T \\ & \times \mathbf{D}(\mathbf{R} - (K(\mathbf{T}, \mathbf{T}^T)\mathbf{g}_1 + \mathbf{e}\delta_1)) + \frac{1}{2}c_1(\mathbf{g}_1^T\mathbf{g}_1 + \delta_1^2) \\ & + c_2(v_1\varepsilon_1 + \frac{1}{m}\mathbf{d}^T\xi_1) \\ \text{s.t. } & \mathbf{R} - (K(\mathbf{T}, \mathbf{T}^T)\mathbf{g}_1 + \mathbf{e}\delta_1) \geqslant -\varepsilon_1\mathbf{e} - \xi_1 \\ & \xi_1 \geqslant 0\mathbf{e} \; \varepsilon_1 \geqslant 0, \end{aligned} \tag{15}$$

$$\begin{aligned} \min_{\mathbf{g}_2,\delta_2,\xi_2,\varepsilon_2} \; & \frac{1}{2}(\mathbf{R} - (K(\mathbf{T}, \mathbf{T}^T)\mathbf{g}_2 + \mathbf{e}\delta_2))^T \\ & \times \mathbf{D}(\mathbf{R} - (K(\mathbf{T}, \mathbf{T}^T)\mathbf{g}_2 + \mathbf{e}\delta_2)) + \frac{1}{2}c_3(\mathbf{g}_2^T\mathbf{g}_2 + \delta_2^2) \\ & + c_4(v_2\varepsilon_2 + \frac{1}{m}\mathbf{d}^T\xi_2) \\ \text{s.t. } & (K(\mathbf{T}, \mathbf{T}^T)\mathbf{g}_2 + \mathbf{e}\delta_2) - \mathbf{R} \geqslant -\varepsilon_2\mathbf{e} - \xi_2 \\ & \xi_2 \geqslant 0\mathbf{e} \; \varepsilon_2 \geqslant 0. \end{aligned} \tag{16}$$

The first term in (15)/(16) is the sum of weighted squared errors from training points to the down/up bound function.

The regularization term (the second term) can smooth the regression function $f_{1/2}(t)$, and the third term is adopted to narrow the ε tube. The objective functions in (15) and (16) are proposed based on the principle of structural risk minimization. The structural minimization principle is helpful to weaken the problem of overfitting and improve the quality of normalization [29]. For a more detailed explanation of the objective functions, please refer to reference [32].

It is difficult to solve the optimization problems of (15) and (16) directly. However, we can map them into the dual problems by adopting Lagrangian multipliers. Then, (15) can be transformed as (17),

$$\begin{aligned} &L(\mathbf{g}_1, \delta_1, \xi_1, \varepsilon_1, \alpha, \beta, \gamma) \\ &= \frac{1}{2}(\mathbf{R}-(K(\mathbf{T},\mathbf{T}^T)\mathbf{g}_1+\mathbf{e}\delta_1))^T D(\mathbf{R}-(K(\mathbf{T},\mathbf{T}^T)\mathbf{g}_1+\mathbf{e}\delta_1)) \\ &\quad +\frac{1}{2}c_1(\mathbf{g}_1^T\mathbf{g}_1+\delta_1^2)+c_2(v_1\varepsilon_1+\frac{1}{m}\mathbf{d}^T\xi_1) \\ &\quad -\alpha^T(\mathbf{R}-(K(\mathbf{T},\mathbf{T}^T)\mathbf{g}_1+\mathbf{e}\delta_1)+\varepsilon_1\mathbf{e}+\xi_1)-\beta^T\xi_1-\gamma\varepsilon_1, \end{aligned} \tag{17}$$

where $\alpha = (\alpha_1, \ldots, \alpha_m)^T$, $\beta = (\beta_1, \ldots, \beta_m)^T$, and $\gamma \geqslant 0$ are Lagrangian multipliers. By using Karush–Kuhn–Tucker (KKT) conditions, we can get the dual problem as (18),

$$\begin{aligned} \min\ & \frac{1}{2}\alpha^T\Xi(\Xi^T\mathbf{D}\Xi+c_1\mathbf{I})^{-1}\Xi^T\alpha \\ &\quad -\mathbf{R}^T\mathbf{D}\Xi(\Xi^T\mathbf{D}\Xi+c_1\mathbf{I})^{-1}\Xi^T\alpha+\mathbf{R}^T\alpha \\ s.t.\ & 0\mathbf{e} \leqslant \alpha \leqslant \frac{c_2}{m}\mathbf{d}, \\ & \mathbf{e}^T\alpha \leqslant c_2 v_1. \end{aligned} \tag{18}$$

where

$$\Xi = [\,K(\mathbf{T},\mathbf{T}^T)\ \ \mathbf{e}\,] \tag{19}$$

By computing the dual QPP problem (18), it can be obtained

$$\mathbf{w}_1 = [\,\mathbf{g}_1^T\ \ \delta_1\,]^T = (\Xi^T\mathbf{D}\Xi+c_1\mathbf{I})^{-1}\Xi^T(\mathbf{D}\mathbf{R}-\alpha). \tag{20}$$

Similarly, the dual problem of (16) can be expressed as (21), and it is not difficult to get

$$\begin{aligned} \min\ & \frac{1}{2}\lambda^T\Xi(\Xi^T\mathbf{D}\Xi+c_3\mathbf{I})^{-1}\Xi^T\lambda \\ &\quad +\mathbf{R}^T\mathbf{D}\Xi(\Xi^T\mathbf{D}\Xi+c_3\mathbf{I})^{-1}\Xi^T\lambda-\mathbf{R}^T\lambda \\ s.t.\ & 0\mathbf{e} \leqslant \lambda \leqslant \frac{c_4}{m}\mathbf{d}, \\ & \mathbf{e}^T\lambda \leqslant c_4 v_2. \end{aligned} \tag{21}$$

$$\mathbf{w}_2 = [\,\mathbf{g}_2^T\ \ \delta_2\,]^T = (\Xi^T\mathbf{D}\Xi+c_3\mathbf{I})^{-1}\Xi^T(\mathbf{D}\mathbf{R}+\lambda). \tag{22}$$

C. SUMMARY OF THE ALGORITHM (TAKING 2 TRANSMIT ANTENNAS AND 1 RECEIVE ANTENNA AS AN EXAMPLE)

The channel frequency response estimation can be carried out according to the following procedure.

Input: Transmitted pilot value matrix $\mathbf{X}_P$ and pilot position $\mathbf{T} = [n\Delta t, m\Delta f]$, $n = 0, \ldots, N_t - 1$, $m = 0, \ldots, N_f - 1$; received pilot value $\mathbf{Y}_P$; the selected constants c_1, c_2, c_3, c_4, v_1 and v_2, in (15) and (16); A, σ in (14).

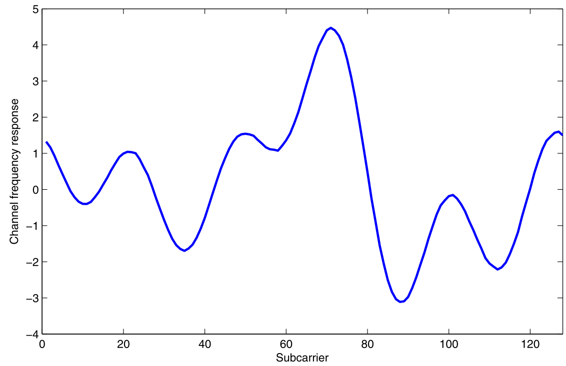


FIGURE 3. Channel response under v = 100km/h, L = 5.

Output: The estimated channel frequency response of all subcarriers $\tilde{\mathbf{H}}_{1/2}$.

Process:

1. Use (11) to compute channel frequency response at $\mathbf{T} = [n\Delta t, m\Delta f]$: $\hat{\mathbf{H}}_{1P}$ and $\hat{\mathbf{H}}_{2P}$.
2. Let $\mathbf{R} = \text{real}(\hat{\mathbf{H}}_{1P})$, use wavelet transform to filter $\mathbf{R}$ and get $\hat{\mathbf{R}}$. Use (14) to compute $\mathbf{d}$.
3. In (18) and (21), $\Xi = [\,K(\mathbf{T},\mathbf{T}^T)\ \mathbf{e}\,]$. By Solving (18) and (21), we can get α and λ.
4. Compute $\mathbf{w}_1$ and $\mathbf{w}_2$ by (20) and (22), respectively.
5. Calculate $h_{1\text{real}}(\tau) = \frac{1}{2}K(\tau,\mathbf{T}^T)(\mathbf{g}_1+\mathbf{g}_2)^T+\frac{1}{2}(\delta_1+\delta_2)$.
6. Let $\mathbf{R} = \text{imag}(\hat{\mathbf{H}}_{1P})$, $\mathbf{R} = \text{real}(\hat{\mathbf{H}}_{2P})$ and $\mathbf{R} = \text{imag}(\hat{\mathbf{H}}_{2P})$, respectively, and repeat 2-5, the $h_{1\text{imag}}(\tau)$, $h_{2\text{real}}(\tau)$ and $h_{2\text{imag}}(\tau)$ can be obtained.
7. Frequency response can be estimated as $(\tilde{\mathbf{H}})_{1\mathbf{t}/2\mathbf{t}} = h_{1/2\text{real}}(\mathbf{t}) + jh_{1/2\text{imag}}(\mathbf{t})$, where $(\tilde{\mathbf{H}})_{1\mathbf{t}/2\mathbf{t}}$ is the $\mathbf{t} = (n, m)$ element of $\tilde{\mathbf{H}}_{1/2}$.

V. SIMULATION RESULTS

In this section, the simulation experiments are presented to demonstrate the performance of the proposed weighted TSVR based nonlinear channel frequency response estimation algorithm for MIMO-OFDM system. The proposed algorithm, linear interpolation, TSVR in [29], BEM based estimation, LMMSE estimation and perfect estimation are presented for comparison. Computations are carried out on MATLAB R2014a.

At the receiving end, any one of the M_R receiving antennas needs to estimate M_T channel responses independently. Therefore, we generally choose 2I1O-OFDM as the test example. Consider a 2I1O-OFDM system with doubly fading channel.Channel taps are assumed to be independent and identically distributed (i.i.d.). The channel response is correlate in time, which can be expressed as $E[h(n_1, l_1)h^*(n_2, l_2)] = \sigma^2 J_0(2\pi f_{\max} T_s(n_1 - n_2))\delta(l_1 - l_2)$, where $E(\cdot)$ denotes taking expected value, $(\cdot)^*$ means conjugate, n_i and l_i $(i = 1, 2)$ are time index and channel path index, respectively. J_0 is Bessel function with the first kind zeroth-order, T_s and σ^2 are sampling interval and variance of the channel impulse response, respectively [33].

TABLE 1. Simulation parameters.

Data modulation	16 QAM
Carrier frequency	$f_c = 2.15GHz$
Mobile speed	$v = 100/300$ Km/h
Number of carriers	$N = 128$
Number of multipath	$L = 5/10$
SVR parameters	$c_1 = c_3 = 0.1$ $c_2 = c_4 = 0.1$ $\nu_1 = \nu_2 = 10$

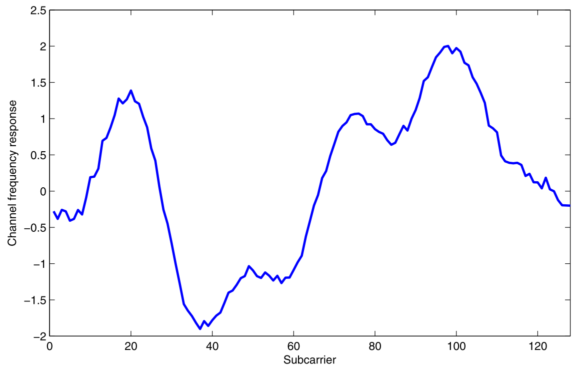


FIGURE 4. Channel response under v = 300km/h, L = 5.

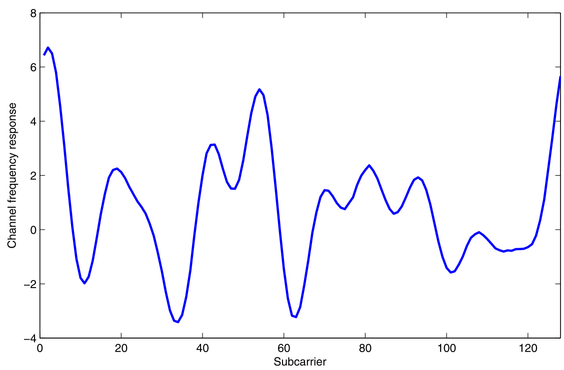


FIGURE 5. Channel response under v = 100km/h, L = 10.

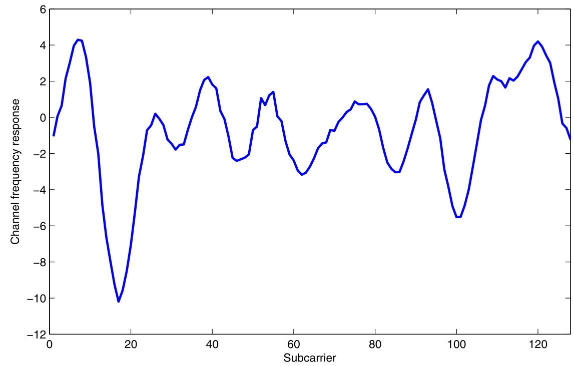


FIGURE 6. Channel response under v = 300km/h, L = 10.

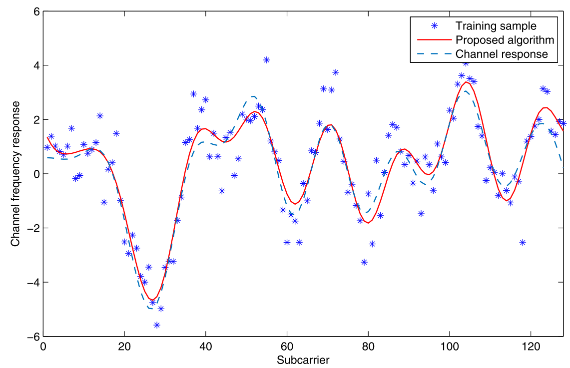


FIGURE 7. Regression of WTWTSVR to training samples of channel response.

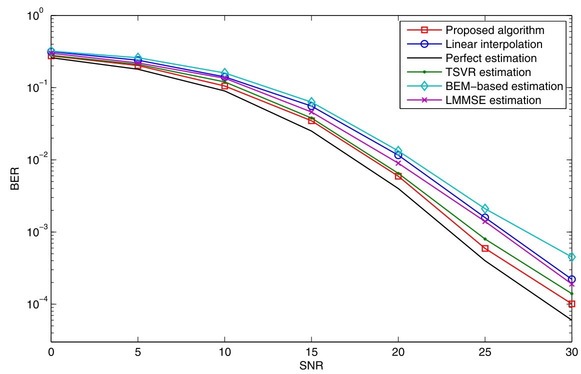


FIGURE 8. Bit error rate at moving speed v = 100km/h, Δt = 4, Δf = 4.

In the simulation, pilots are inserted both in time and frequency domain. Gaussian function is selected as the nonlinear mapping kernel in (19) for WTWTSVR based algorithm and TSVR [29] base estimator:

$$K(\mathbf{x}_1^T, \mathbf{x}_2^T) = exp(-\|\mathbf{x}_1 - \mathbf{x}_2\|^2/\eta), \tag{23}$$

where η is the width or the variance of the Gaussian function. Constants in (18) and (21) are determined by searching from $\{10^k|k = -2, -1, \ldots, 3\}$. The simulation parameters are listed in Table 1.

Two criteria, sum squared error (SSE) and bit error rate (BER) are selected to estimate the performance of algorithms. The criteria are specified as: SSE=$\sum_{i=1}^{n_t} |y_i - \hat{y}_i|^2$, BER=$n_e/n_T$, where $\hat{y}_i$ is the predicted value of the testing sample y_i, n_e and n_T are error data number and total data number in binary, respectively. Define SNR=10log(σ_x^2/σ_v^2), where $\sigma_x^2 = E(|x(k)|^2)$, σ_v^2 is AWGN variance.

Number of transmission multipath and the relative moving speed can affect the fading depth of the channel. Figure 3 - Figure 6 demonstrate the channel response in one OFDM symbol under moving speed being 100km/h, 300km/h and number of multipath being 5 and 10. It can be observed that large moving speed affects the channel response of adjacent subcarriers, resulting in ICI and more multipath makes the channel fading deeper. Therefore, higher speed and more paths will bring more challenges to channel estimation.

TABLE 2. SSE of estimation methods.

SNR(dB)	0	5	10	15	20	25	30
			v=100km/h, $\Delta t = 4$, $\Delta f = 4$				
WTWTSVR	0.2584	0.1884	0.0998	0.0314	0.0066	0.0003	0.0001
TSVR	0.2838	0.2055	0.1212	0.0496	0.0118	0.0006	0.0002
LMMSE	0.2863	0.2112	0.1288	0.0501	0.0125	0.0007	0.0002
Linear	0.2917	0.2257	0.1338	0.0516	0.0131	0.0009	0.0002
BEM	0.2918	0.2258	0.1342	0.0520	0.0142	0.0011	0.0003
			v=300km/h, $\Delta t = 4$, $\Delta f = 4$				
WTWTSVR	0.2604	0.1906	0.1041	0.0308	0.0058	0.0014	0.0002
TSVR	0.2833	0.2110	0.1229	0.0458	0.0090	0.0025	0.0004
LMMSE	0.2885	0.2193	0.1279	0.0482	0.0108	0.0029	0.0004
Linear	0.2938	0.2251	0.1380	0.0502	0.0120	0.0033	0.0005
BEM	0.2982	0.2296	0.1411	0.0516	0.0131	0.0038	0.0006
			v=100km/h, $\Delta t = 8$, $\Delta f = 8$				
WTWTSVR	0.2593	0.2026	0.1197	0.0487	0.0104	0.0016	0.0003
TSVR	0.2621	0.2101	0.1215	0.0510	0.0115	0.0022	0.0004
LMMSE	0.2653	0.2122	0.1243	0.0531	0.0121	0.0028	0.0004
Linear	0.2676	0.2140	0.1283	0.0555	0.0128	0.0032	0.0005
BEM	0.2679	0.2144	0.1290	0.0562	0.0134	0.0039	0.0006
			v=300km/h, $\Delta t = 8$, $\Delta f = 8$				
WTWTSVR	0.3609	0.2985	0.1138	0.0436	0.0122	0.0022	0.0007
TSVR	0.3667	0.3022	0.2199	0.0486	0.0149	0.0028	0.0011
LMMSE	0.3685	0.3042	0.2200	0.0491	0.0152	0.0033	0.0012
Linear	0.3694	0.3067	0.2208	0.0506	0.0157	0.0038	0.0013
BEM	0.3696	0.3066	0.2211	0.0519	0.0161	0.0041	0.0014

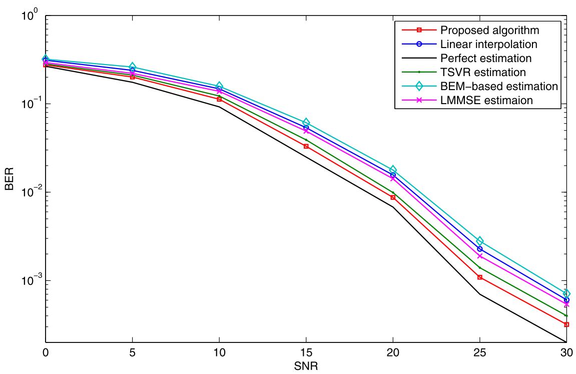


FIGURE 9. Bit error rate at moving speed v = 300km/h, Δt = 4, Δf = 4.

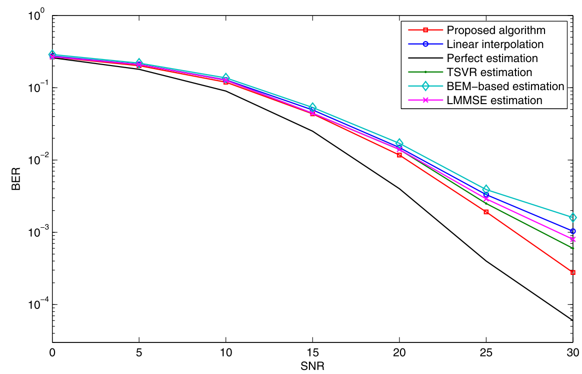


FIGURE 10. Bit error rate at moving speed v = 100km/h, Δt = 8, Δf = 8.

Figure 7 illustrates the effectiveness of WTWTSVR for regression of nonlinear channels. In this test, $SNR = 5$dB, path number $L = 10$, mobile speed $v = 100$km/h. The star points are noised channel response for training WTWTSVR. We can see that the estimate of the proposed method (dashed line) can fit the channel response (the solid line) well.

Figures 8 and 9 show the BER performance of the WTWTSVR based algorithm at mobile speed 100km/h and 300km/h with pilot insertion interval $\Delta t = 4$ in time domain and $\Delta f = 4$ in frequency domain, respectively. For comparison, curves of TSVR-based, linear interpolation, LMMSE, BEM based methods and perfect estimation are also demonstrated. Figures 10 and 11 show those of $\Delta t = 8$ and $\Delta f = 8$. The channel frequency response SSE of methods is shown in Table 2. The results are averaged from 50 runs.

It can be observed from Figures 8 - 11 that the improvement in BER performance becomes significant as SNR increases. While the improvement is negligible when SNR is less than 10dB due to the main interference effect of additive noise at low SNR. The increase of moving speed can also degrade the BER performance. This is because faster relative movement speed leads to greater ICI. The BER curves of the proposed method and the TSVR based algorithm have similar trends, and their performance is better than that of traditional methods, which shows that SVR has the regression

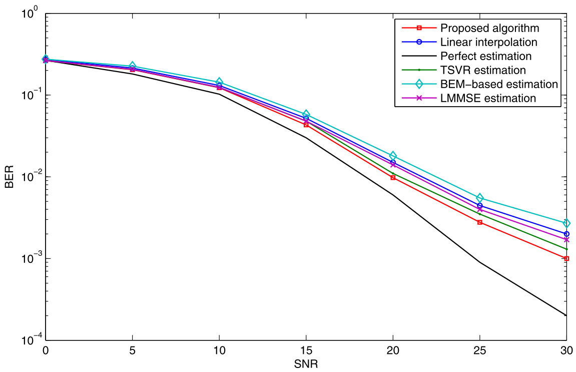


FIGURE 11. Bit error rate at moving speed v=300km/h, $\Delta t = 8$, $\Delta f = 8$.

advantage of nonlinear relationship. Comparing Figures 8, 9 with 10, 11, it can be seen that the increase of time and frequency interval degrades the regression performance, which is due to the decrease of information content and the increase of interpolation interval. As shown in Figures 8- 11, the proposed algorithm outperforms the TSVR based algorithm due to the function of weights obtained by wavelet transform preprocessing.

The results in Table 2 also show the same conclusion from the perspective of SSE, and the effectiveness of the proposed algorithm is further verified.

VI. CONCLUSION

A weighted TSVR based MIMO-OFDM channel frequency response estimator is proposed in this paper. The proposed estimation method is based on pilots inserted uniformly in time and frequency domain, which are known to transmitter and receiver. At the transmitter, the pilots are inserted into the data sequence, and then transmitted through the antenna after space-time coding and OFDM processing. At the receiver, after OFDM demodulation and space-time decoding, the channel frequency response of pilot symbol can be calculated by LS algorithm. Then the channel response of data symbol is interpolated by the WTWTSVR based estimator. In the proposed algorithm, training samples on the pilot position are given different weights according to the distance from the mean position based on wavelet transform.

The WTWTSVR method is suitable for dealing with time series signal denoising such as the channel estimation problems, which is determined by the property of wavelet filter. The effectiveness of the proposed method in the estimation of nonlinear wireless communication channel is confirmed by computational experiments. However, the computational complexity of SVR increases significantly with the increase of training samples. Therefore, how to reduce the computational complexity under the premise of ensuring the accuracy of estimation is our next research goal.

REFERENCES

[1] J. Lu, X. Chen, S. Liu, and P. Fan, "Location-aware ICI reduction in MIMO-OFDM downlinks for high-speed railway communication systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 2958–2972, Apr. 2018.

[2] M. A. Morsy and A. S. Alsayyari, "Performance analysis of coherent BPSK-OCDMA wireless communication system," *Wireless Netw.*, vol. 26, no. 6, pp. 4491–4505, Aug. 2020.

[3] J. Wu and P. Fan, "A survey on high mobility wireless communications: Challenges, opportunities and solutions," *IEEE Access*, vol. 4, pp. 450–476, 2016.

[4] Z. Sheng, H. D. Tuan, H. H. Nguyen, and Y. Fang, "Pilot optimization for estimation of high-mobility OFDM channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 8795–8806, Oct. 2017.

[5] A. Riadi, M. Boulouird, and M. M. Hassani, "Performance of massive-MIMO OFDM system with M-QAM modulation based on LS channel estimation," in *Proc. Int. Conf. Adv. Syst. Emergent Technol. (IC ASET)*, Mar. 2019, pp. 74–78.

[6] D. Neumann, T. Wiese, and W. Utschick, "Learning the MMSE channel estimator," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2905–2917, Jun. 2018.

[7] K.-C. Hung and D. W. Lin, "Pilot-based LMMSE channel estimation for OFDM systems with power–delay profile approximation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 150–159, Jan. 2010.

[8] X. Ye, G. Zheng, A. Zhang, L. You, and X. Gao, "DFT-based low-complexity channel estimation method for millimeter-wave MIMO systems," *Wireless Pers. Commun.*, vol. 107, no. 1, pp. 205–216, Jul. 2019.

[9] P. Ge and B. Hu, "An efficient channel estimation algorithm for PDM CO-OFDM WDM system," in *Proc. IEEE 3rd Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC)*, Mar. 2019, pp. 2527–2530.

[10] F. Gu, Y. Fan, L. Wang, X. Tan, and J. Wei, "A universal channel estimation algorithm based on DFT smoothing filtering," *IEEE Access*, vol. 7, pp. 129883–129891, 2019.

[11] Y. Kang, K. Kim, and H. Park, "Efficient DFT-based channel estimation for OFDM systems on multipath channels," *IET Commun.*, vol. 1, no. 2, p. 197, 2007.

[12] J. K. Tugnait, S. He, and H. Kim, "Doubly selective channel estimation using exponential basis models and subblock tracking," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1275–1289, Mar. 2010.

[13] J. T. Dias, R. C. de Lamare, and Y. V. Zakharov, "BEM-based channel estimation for 5G multicarrier systems," in *Proc. 22nd Int. ITG Workshop Smart Antennas (WSA)*, Bochum, Germany, Mar. 2018, pp. 1–5.

[14] N. M. Idrees, X. Yu, and A. Springer, "Optimal tracking of doubly-selective radio channels for OFDM based modern wireless systems," *Phys. Commun.*, vol. 35, Aug. 2019, Art. no. 100739.

[15] Y.-H. Jan, "Iterative joint channel estimation and signal detection for OFDM system in double selective channels," *Wireless Pers. Commun.*, vol. 99, no. 3, pp. 1279–1294, Apr. 2018.

[16] A. Djouama, M.-S. Lim, and F. Y. Ettoumi, "Channel estimation in long term evolution uplink using minimum mean square error-support vector regression," *Wireless Pers. Commun.*, vol. 79, no. 3, pp. 2291–2304, Dec. 2014.

[17] A. Charrada and A. Samet, "Joint interpolation for LTE downlink channel estimation in very high-mobility environments with support vector machine regression," *IET Commun.*, vol. 10, no. 17, pp. 2435–2444, Nov. 2016.

[18] M. Zhang, X. Zhou, and C. Wang, "A novel noise suppression channel estimation method based on adaptive weighted averaging for OFDM systems," *Symmetry*, vol. 11, no. 8, p. 997, Aug. 2019.

[19] B. Clerckx and J. Kim, "On the beneficial roles of fading and transmit diversity in wireless power transfer with nonlinear energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7731–7743, Nov. 2018.

[20] H. Hu, M. Senkaya, and Y. Zheng, "A novel measurement of the surface wave dispersion with high and adjustable resolution: Multi-channel nonlinear signal comparison," *J. Appl. Geophys.*, vol. 160, pp. 236–241, Jan. 2019.

[21] Z. Jiang, S. Chen, A. F. Molisch, R. Vannithamby, S. Zhou, and Z. Niu, "Exploiting wireless channel state information structures beyond linear correlations: A deep learning approach," *IEEE Commun. Mag.*, vol. 57, no. 3, pp. 28–34, Mar. 2019.

[22] Y. Liao, G. Sun, Z. Cai, X. Shen, and Z. Huang, "Nonlinear Kalman filter-based robust channel estimation for high mobility OFDM systems," *IEEE Trans. Intell. Transp. Syst.*, early access, Jul. 15, 2020, doi: 10.1109/TITS.2020.3004587.

[23] Y. Yang, F. Gao, X. Ma, and S. Zhang, "Deep learning-based channel estimation for doubly selective fading channels," *IEEE Access*, vol. 7, pp. 36579–36589, 2019.

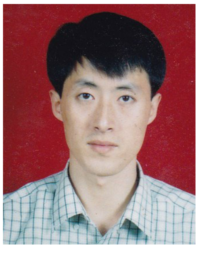
[24] D. Gupta, "Training primal K-nearest neighbor based weighted twin support vector regression via unconstrained convex minimization," *Int. J. Speech Technol.*, vol. 47, no. 3, pp. 962–991, Oct. 2017.
[25] Y. Xu, Z. Yang, Y. Zhang, X. Pan, and L. Wang, "A maximum margin and minimum volume hyper-spheres machine with pinball loss for imbalanced data classification," *Knowl.-Based Syst.*, vol. 95, pp. 75–85, Mar. 2016.
[26] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer, 1995.
[27] Y. Xu, X. Li, X. Pan, and Z. Yang, "Asymmetric ν-twin support vector regression," *Neural Comput. Appl.*, vol. 2, pp. 1–16, Dec. 2018.
[28] M. Sanchez-Fernandez, M. de-Prado-Cumplido, J. Arenas-Garcia, and F. Perez-Cruz, "SVM multiregression for nonlinear channel estimation in multiple-input multiple-output systems," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2298–2307, Aug. 2004.
[29] X. Peng, "TSVR: An efficient twin support vector machine for regression," *Neural Netw.*, vol. 23, no. 3, pp. 356–372, 2010.
[30] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, "New support vector algorithms," *Neural Comput.*, vol. 12, no. 5, pp. 1207–1245, May 2000.
[31] Y. Shao, C. Zhang, Z. Yang, L. Jing, and N. Deng, "An ε-twin support vector machine for regression," *Neural Comput. Appl.*, vol. 23, pp. 175–185, Jul. 2013.
[32] L. Wang, C. Gao, and N. Zhao, "Wavelet transform-based weighted $\nu\nu$-twin support vector regression," *Int. J. Mach. Learn. Cyber.*, vol. 11, no. 6, pp. 95–110, 2019.
[33] G. L. Stuber, *Principles of Mobile Communication*. Boston, MA, USA: Kluwer, 1996.

SHUO LI was born in Jilin, China, in 1976. She received the B.S. degree from the Jilin Institute of Technology, in 1999, and the M.S. degree from the Chongqing University of Posts and Telecommunication, China, in 2005. She is currently pursuing the Ph.D. degree with Northeastern University, China. Since 2005, she has been with the University of Science and Technology Liaoning. Her research interests include pattern recognition, machine learning, and digital signal processing.

WEIMIN LEI (Member, IEEE) received the B.E. degree in computer software from Nankai University, in 1992, the M.E. degree in computer software from the Chinese Academy of Sciences, in 1995, and the Ph.D. degree from the Dalian University of Technology, in 1999. He is currently working as a Professor with the School of Computer Science and Engineering, Institute of Communication and Information System, Northeastern University, China. He has published more than 80 articles. His recent research interests include real-time multipath transmission optimization, real-time signal processing, and future network architecture.

WEI ZHANG is currently a Lecturer with the School of Computer Science and Engineering, Institute of Communication and Information System, Northeastern University, China. Her current research interests include multimedia communication, protocols and services in next generation networks, and network architecture of cloud computing.

XIAOYU WANG was born in Liaoning, China, in 1978. She received the B.S. degree from Liaoning University, China, in 2001, and the master's degree from the University of Science and Technology Liaoning, China, in 2006. Since 2001, she has been with the University of Science and Technology Liaoning. Her research interests include signal processing and wireless communication systems.



LIDONG WANG was born in Liaoning, China, in 1971. He received the B.S. degree from the Harbin Institute of Technology, in 1993, the M.S. degree from the University of Science and Technology Liaoning, in 2004, and the Ph.D. degree from Gyeongsang National University, South Korea, in 2008. He is currently working as a Professor with the University of Science and Technology Liaoning. His research interests include pattern recognition, machine learning, and nonlinear system control.