Deep Learning-Based Channel Estimation

Mehran Soltani, Ali Mirzaei, Vahid Pourahmadi, Hamid Sheikhzadeh

***Abstract*—In this paper, we present a deep learning (DL) algorithm for channel estimation in communication systems. We consider the time-frequency response of a fast fading communication channel as a two-dimensional image. The aim is to find the unknown values of the channel response using some known values at the pilot locations. To this end, a general pipeline using deep image processing techniques, image super-resolution (SR) and image restoration (IR) is proposed. This scheme considers the pilot values, altogether, as a low-resolution image and uses an SR network cascaded with a denoising autoencoder as an IR network to estimate the channel. Moreover, a simple implementation of the proposed pipeline is presented. The estimation error shows that the presented algorithm is comparable to the minimum mean square error (MMSE) with full knowledge of the channel statistics and it is better than ALMMSE (an approximation to linear MMSE). The results confirm that this pipeline can be used efficiently in channel estimation.**

***Index Terms*—Channel estimation, Deep Learning, Super-resolution, Image restoration, Denoising autoencoder**

I. Introduction

Orthogonal frequency-division multiplexing (OFDM) is a modulation method that has been widely used in communication systems to address frequency-selective fading in wireless channels. In a communication channel, the received signal is usually distorted by channel characteristics. In order to recover the transmitted symbols, the channel effect must be estimated and compensated at the receiver. Generally, the receiver estimates the channel using some symbols named pilots which their positions and values in time-frequency are known to both transmitter and receiver. Depending on these pilot arrangements, three different structures can be considered: block-type, comb-type and lattice-type [1], [2]. In the block-type arrangement, pilots are transmitted periodically at the beginning of an OFDM block at all subcarriers while in comb-type, the pilots are present in few subcarriers of few OFDM symbols. In the lattice-type arrangement, pilots are inserted along both time and frequency axes with given periods in a diamond-shaped constellation.

The conventional pilot-based estimation methods, i.e., Least Square (LS) and Minimum Mean Square Error (MMSE) utilize the pilot values in time-frequency grids to find the unknown values of the channel response. These algorithms have been optimized in various conditions [3]. In contrast to the LS estimation which requires no information about the statistics of the channel, the MMSE estimation results in a better performance by utilizing the statistics of the channel and noise variance. To use the MMSE in practical scenarios, some approaches are presented which reduce the complexity of this scheme and use an estimation of the channel statistics instead of the exact information. In [4], an approximated linear version of the MMSE is proposed (ALMMSE) which its complexity is much less than the original MMSE, and it can be employed efficiently in fast fading channels. Moreover, in [5] a new approach is proposed to estimate the channel correlation matrix using a spectral smoothing filter.

Recently, Deep Learning (DL) has gained much attention in various applications like computer vision (CV), natural language processing (NLP) and automatic speech recognition (ASR) [6]. Researchers are trying to extend these algorithms to other applications including communication systems. In DL-based communication systems, some approaches have been proposed to enhance the performance of different conventional algorithms including modulation recognition [7], channel encoding and decoding [8], signal detection [9], channel equalization [10], channel state information (CSI) feedback [11] and channel estimation [12]. In [12], the communication system is considered as a black-box, and an end-to-end DL architecture is proposed for signal transmission/reception. Encoding, decoding, channel estimation and all other functionalities of a communication link are embedded in the DL-block implicitly. The black-box nature of this result limits its application in cases that we only aim to substitute some of the communication blocks and use the conventional methods for the rest. More specifically, this method is not able to explicitly find the channel time-frequency response and so not effective for applications which need to have the complete channel response.

Motivated by this, in this paper, we present a DL-based framework for channel estimation in OFDM systems. In this method, the time-frequency grid of the channel response is modeled as a 2D-image which is known only at the pilot positions. This channel grid with several pilots is considered as a low-resolution (LR) image and the estimated channel as a high-resolution (HR) one. A two-phase approach is presented to estimate the channel grid. First, an SR algorithm is proposed to enhance the resolution of the LR input. Afterwards, an IR method is utilized to remove the noise effects. For these networks, we have used deeply and fully connected architectures and evaluated the performance of the proposed scheme. The contributions of this paper are summarized as follows:

1) Model the channel time-frequency response as an image.
2) Consider the channel response in the pilot positions as a LR image and the estimated channel response as the proposed HR image.
3) Use DL-based image super-resolution and image denoising techniques to estimate the channel.

The remainder of the paper is organized as follows. Section II provides a brief survey of the channel estimation with conventional methods. Section III presents the structure of the proposed DL-base channel estimator. In section IV, simulation results are presented and finally section V concludes the paper.

II. Background

A. Channel Estimation

In an OFDM system, Considering the kth subframe and the ith subcarrier, the input-output relationship is represented as:

$$Y_{i,k} = H_{i,k} X_{i,k} + Z_{i,k}. \quad (1)$$

Considering an OFDM block of size $N_D \times N_S$, subframe index k is between $[0, N_D - 1]$ and the range of the subcarrier index i is $[0, N_S - 1]$. In (1), $Y_{i,k}$, $X_{i,k}$, and $Z_{i,k}$ are the received signal, transmitted OFDM symbol, and white Gaussian noise, respectively. $H_{i,k}$ is the (i,k) element of $\mathbf{H} \in \mathbb{C}^{N_S \times N_D}$. $\mathbf{H}$ represents time-frequency response of channel for all suibcarriers and subframes.

To estimate the channel, specifically in the channels with fading, the time domain response is divided into subframes $\mathbf{H} = \{\mathbf{h}[1], \mathbf{h}[2], ..., \mathbf{h}[N_D]\}$, where each $\mathbf{h}[N_k]$ is the channel frequency response at the kth subframe.

The LS method estimates the channel at the pilot positions. If we consider the LS estimated channel as a diagonal matrix $\mathbf{H}_p^{\mathrm{LS}} \in \mathbb{C}^{N_P \times N_P}$, $\mathbf{H}_p^{\mathrm{LS}}$ can be estimated by solving:

$$\hat{\mathbf{H}}_p^{\mathrm{LS}} = \underset{\mathbf{H}_p}{\arg\min} \|\mathbf{y}_p - \mathbf{H}_p \mathbf{x}_p\|_2^2, \quad (2)$$

where $||.||_2$ is the $\ell 2$ distance and $\hat{\mathbf{H}}_p^{\mathrm{LS}} \in \mathbb{C}^{N_P \times N_P}$ is the estimated diagonal matrix. $\mathbf{x}_p$ contains the known pilot values and $\mathbf{y}_p$ is the corresponding observations. The optimization of (2) results in $\hat{\mathbf{h}}_p^{\mathrm{LS}} = Diag(\hat{\mathbf{H}}_p^{\mathrm{LS}}) = \mathbf{y}_p / \mathbf{x}_p$. To find the channel value at an points other that pilot positions, we have to apply a two dimensional interpolation method.

A better choice than LS, is MMSE estimator which is obtained by multiplying the LS estimate at the pilot-symbol positions with a filtering matrix $\mathbf{A}_{\mathrm{MMSE}} \in \mathbb{C}^{N_D \times N_P}$ [13]:

$$\hat{\mathbf{h}}_d^{\mathrm{MMSE}} = \mathbf{A}_{\mathrm{MMSE}} \hat{\mathbf{h}}_p^{\mathrm{LS}}, \quad (3)$$

where $\hat{\mathbf{h}}_d^{\mathrm{MMSE}}$ is the MMSE estimation of channel at subframe d and $\hat{\mathbf{h}}_p^{\mathrm{LS}} = Diag(\hat{\mathbf{H}}_p^{\mathrm{LS}})$. In order to find the MMSE filtering matrix, the MSE,

$$\epsilon = \mathbb{E}\{\|\mathbf{h}_d - \mathbf{A}_{\mathrm{MMSE}} \hat{\mathbf{h}}_p^{\mathrm{LS}}\|_2^2\}, \quad (4)$$

has to be minimized. Minimizing (4) leads to

$$\mathbf{A}_{\mathrm{MMSE}} = \mathbf{R}_{\mathbf{h}_d \mathbf{h}_p} (\mathbf{R}_{\mathbf{h}_p \mathbf{h}_p} + \sigma_n^2 (\mathbf{x}\mathbf{x}^{\mathrm{H}})^{-1})^{-1}, \quad (5)$$

where the matrix $\mathbf{R}_{\mathbf{h}_d \mathbf{h}_p} = \mathbb{E}\{\mathbf{h}_d \mathbf{h}_p^{\mathrm{H}}\} \in \mathbb{C}^{N_D \times N_P}$ denotes the channel correlation matrix between desired subframe and pilot-symbols and the matrix $\mathbf{R}_{\mathbf{h}_p \mathbf{h}_p} = \mathbb{E}\{\mathbf{h}_p \mathbf{h}_p^{\mathrm{H}}\} \in \mathbb{C}^{N_P \times N_P}$ is the channel correlation matrix at the pilot-symbols. It is obvious that the MMSE will be useful only if the correlation matrix of the channel, denoted as $\mathbf{R}$, is completely known.

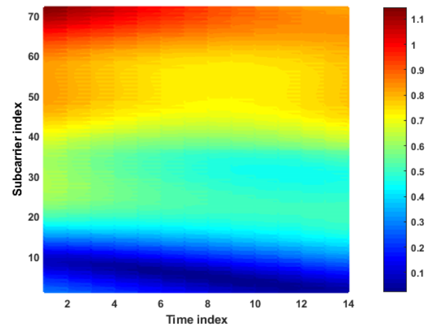

Fig. 1. channel matrix between a transmitter/receiver as a 2D image

B. Super-resolution and Image restoration

Considering a low resolution and noisy image, several techniques have been proposed to reproduce the higher resolution and less noisy image. Image super-resolution (SR) is a class of techniques used for resolution enhancement in images. DL-based algorithms, especially with deeply and fully convolutional networks, have achieved high performance in the problem of recovering the HR images from the LR image inputs. Recently, Super-resolution convolutional neural network (SRCNN) [14] is proposed to map between LR/HR images in an end-to-end manner. In [15], a Deep CNN with residual learning (DCSCN) is proposed which significantly improves the calculation cost compared to the previous methods.

Other than SR techniques, image restoration (IR) algorithms can be applied to remove/reduce the noise effect on an image. Various models have been presented for IR in the literature. For instance, in [16], a feed-forward denoising convolutional neural network (DnCNN) scheme is presented which has utilized the residual learning and batch normalization to speed up the training process.

III. ChannelNet

A. Channel Image

In this work, we consider a simple single-cell downlink Single-input, Single-output (SISO) communication system. Channel time-frequency response matrix $\mathbf{H}$ (of size $N_s \times N_d$) between the transmitter and the receiver antenna can be represented as a 2D-image with complex values. A heat map plot of the absolute values of a sample channel time-frequency grid with $N_D = 14$ time slots and $N_S = 72$ subcarriers (based on Long-Term Evolution (LTE) standard) is shown in Fig.1. This complex image can be divided into two single images which the first one contains the real values and the second one contains the imaginary values of the original image.

B. Network Structure

The overview of the proposed pipeline for DL-based channel estimation, named ChannelNet, is illustrated in Fig.2. The goal is to estimate the whole time-frequency of the channel using the transmitted pilots. Similar to LTE standard, Lattice-type pilot arrangement has been used for pilot transmission.

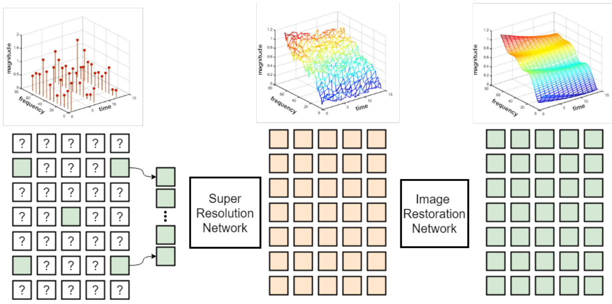

Fig. 2. The proposed pipeline for DL-based channel estimation

The estimated value of the channel at the pilot locations $\hat{\mathbf{h}}_p^{\mathrm{LS}}$ (which might be noisy) is considered as the LR and noisy version of the channel image. To obtain the complete channel image a two stage training approach is presented:

- In the first stage, an SR network is implemented which takes $\hat{\mathbf{h}}_p^{\mathrm{LS}}$ as the vectorized low resolution input image and estimates the unknown values of the time-frequency channel response $\mathbf{H}$.
- In the second stage to remove the noise effects, a denoising autoencoder is cascaded with the SR network.

For the SR architecture, we have used a simple fully connected network with three layers. The size of the input layer is equal to the size of the vector $\hat{\mathbf{h}}_p^{\mathrm{LS}}$ and the size of the output layer is equal to the number of elements in the channel response, i.e., $N_s \times N_d$. Afterwards, the output is reshaped to obtain the whole channel response. Due to the environment noise, this output is the noisy version of the actual channel response.

For the denoising network, we have used an autoencoder with three hidden layers which is cascaded with the SR network to restore the channel response.

C. Training

Lets denote the set of all network parameters by $\Theta = \{\Theta_S, \Theta_R\}$, where the Θ_S and Θ_R denote the set of parameter values for SR and IR networks, respectively. The input to the ChannelNet is the pilot values vector $\hat{\mathbf{h}}_p^{\mathrm{LS}}$ and the output is the estimated channel matrix is denoted as $\hat{\mathbf{H}}$:

$$\hat{\mathbf{H}} = f(\Theta; \hat{\mathbf{h}}_p^{\mathrm{LS}}) = f_R(f_S(\Theta_S; \hat{\mathbf{h}}_p^{\mathrm{LS}}); \Theta_R),$$

where f_S and f_R are the SR and IR functions, respectively.

The total loss function of the network is the Mean square error (MSE) between the estimated and the actual channel responses calculated as follows:

$$C = \frac{1}{\|\mathcal{T}\|} \sum_{\mathbf{h}_p \in \mathcal{T}} \|f(\Theta; \hat{\mathbf{h}}_p^{\mathrm{LS}}) - \mathbf{H}\|_2^2, \tag{6}$$

where $\mathcal{T}$ is the set of all training data and H is the perfect channel. In (6), $\|\mathcal{T}\|$ is the size of the training set.

To simplify the training process, we use a two stage training algorithm. Where in the first stage we minimize the the loss of the SR network, C_1 :

$$C_1 = \frac{1}{\|\mathcal{T}\|} \sum_{\mathbf{h}_p \in \mathcal{T}} \|\mathbf{Z} - \mathbf{H}\|_2^2, \tag{7}$$

where $\mathbf{Z} = f_S(\Theta_S; \hat{\mathbf{h}}_p^{\mathrm{LS}})$ is the output of the SR network.

In the second stage, we freeze the weights of the first network and find the parameters of the autoencoder (the denoising network) by minimizing the loss function C_2 :

$$C_2 = \frac{1}{\|\mathcal{T}\|} \sum_{\mathbf{h}_p \in \mathcal{T}} \|\hat{\mathbf{H}} - \mathbf{H}\|_2^2, \tag{8}$$

where $\hat{\mathbf{H}} = f_R(\mathbf{Z}; \Theta_D)$.

The optimal weights of the denoising network is different for different values of SNR; thus, to have a complete solution we have to re-train the denoising network for each SNR value. This approach is practically impossible to implement because the SNR value is continuous. Fortunately, however, as the results in section IV demonstrates, training networks for a few SNR values (in our case only two values) can still leads to a good performance.

IV. Simulation Results

In this section we train the network and evaluate the mean square error (MSE) over a range of SNRs and compare the results with the widely used baseline algorithms.

We consider a single antenna at the transmitter and a single antenna at the receiver. For channel modeling and pilot transmission, we have used widely used LTE simulator developed by university of Vienna, Vienna LTE-A simulator [17]. Keras and Tensorflow are used for DL architecture implementation using a GPU backend. As in LTE, in our simulations, each frame consists of 14 time slots with 72 subcarriers. For the wireless channel model, Vehicular-A (VehA) fading model with carrier frequency of 2.1 GHz, bandwidth of 1.6 MHz and UE (user equipment) speed of 50 km/h is considered.

The proposed DL network used for simulations with 48 pilots consists of a fully connected network for SR with 48, 1500, 1008 neurons and an autoencoder for IR network with 1008, 600, 400, 600, 1008 neurons.

To see the performance, we have compared the accuracy of channel estimation for the proposed method with that of three state-of-the-art algorithms i.e. ideal MMSE, estimated MMSE and ideal ALMMSE [4] when 48 pilots are used in each frame. The MSE between the estimated and the actual channel realization is considered as the performance metric. The results have been presented in Fig.3. Note that the ideal MMSE has the best performance and gives a lower bound of the achievable MSE as the correlation matrix of the channel should be known fully (without any error) which is not a valid assumption in practical applications. Estimated MMSE tries to estimate the correlation matrix based on received signals and ideal ALMMSE is a linear counterpart of ideal MMSE. Also in Fig.3, it is demonstrated that for low SNR values, the proposed ChannelNet trained by low SNR samples (denoted by deep low-SNR) has comparable performance with ideal MMSE and

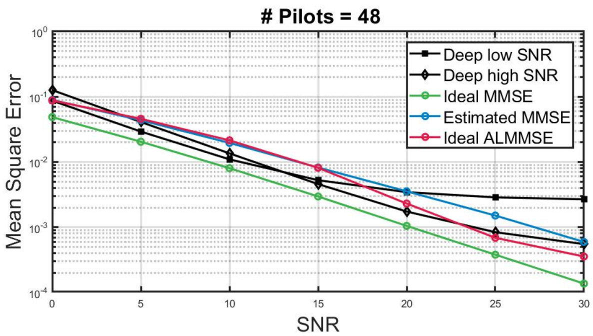

Fig. 3. Mean square error for channel estimation in terms of SNR. The number of pilots is 48

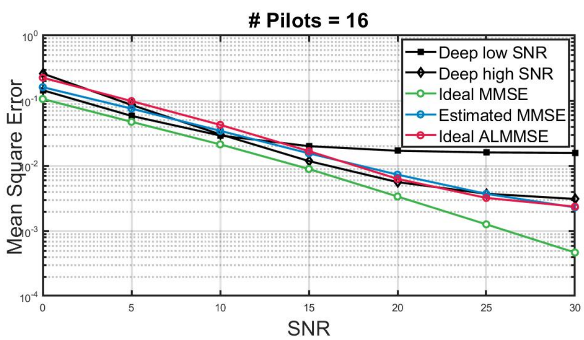

Fig. 4. Mean square error for channel estimation in terms of SNR. The number of pilots is 16

has a better performance than ideal ALMMSE and estimated MMSE. Additionally, it can be observed that after around a mid SNR value (in this example 13dB), the performance of the network trained for high SNR samples (denoted by deep high-SNR) is going to be better than the deep low-SNR. So, we divide the SNR range into two regions. When the SNR value is lower than 13 dB, we estimate the channel by the deep low-SNR network, and beyond this SNR, the deep high-SNR network will be used. It can be seen also that for the SNR values higher than 23 dB, the performance of the deep high-SNR is going to fail again and another network has to be trained; though as long as the SNR is below 20 dB, the two generated networks are sufficient.

To show the robustness of the proposed algorithm for lower number of pilots, results of simulations for 16 number of pilots have been shown in Fig.4. It is shown that in this case, the SNR value which discriminates the deep low-SNR and deep high-SNR regions is about 10 dB. To improve the performance of the proposed pipeline at the high SNRs, for example higher than 25 dB, we can train another network with 25 dB to improve the MSE value.

V. Conclusion

In this paper, we presented ChannelNet, our initial DL-based algorithm for channel estimation in communication systems. In this method, we have considered the time-frequency response of a fading channel as a 2D-image and applied SR and IR algorithms to find the whole channel state based on the pilot values. The results show that the performance of ChannelNet is highly competitive with the MMSE algorithm. The two-step network training procedure has been presented and we also discussed how multiple ChannelNets should be used to best estimate the channel. To expand these results, we aim to improve the estimation quality by applying more advance DL networks for the SR and IR algorithms.

References

[1] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in ofdm systems," *IEEE Transactions on Broadcasting*, vol. 48, pp. 223–229, Sep 2002.

[2] T.-A. Lin and C.-Y. Lee, "Predictive equalizer design for dvb-t system," in *2005 IEEE International Symposium on Circuits and Systems*, pp. 940–943 Vol. 2, May 2005.

[3] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for ofdm systems with rapid dispersive fading channels," *IEEE Transactions on Communications*, vol. 46, pp. 902–915, Jul 1998.

[4] M. Q. Abdulhasan, M. I. Salman, C. K. Ng, N. K. Noordin, S. J. Hashim, and F. B. Hashim, "Approximate linear minimum mean square error estimation based on channel quality indicator feedback in lte systems," in *2013 IEEE 11th Malaysia International Conference on Communications (MICC)*, pp. 446–451, Nov 2013.

[5] C. Mehlfuhrer, S. Caban, and M. Rupp, "An accurate and low complex channel estimator for ofdm wimax," in *2008 3rd International Symposium on Communications, Control and Signal Processing*, pp. 922–926, March 2008.

[6] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[7] T. OShea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, pp. 563–575, Dec 2017.

[8] T. Gruber, S. Cammerer, J. Hoydis, and S. t. Brink, "On deep learning-based channel decoding," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, March 2017.

[9] N. Samuel, T. Diskin, and A. Wiesel, "Deep mimo detection," in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, July 2017.

[10] D. Erdogmus, D. Rende, J. C. Principe, and T. F. Wong, "Nonlinear channel equalization using multilayer perceptrons with information-theoretic criterion," in *Neural Networks for Signal Processing XI: Proceedings of the 2001 IEEE Signal Processing Society Workshop (IEEE Cat. No.01TH8584)*, pp. 443–451, 2001.

[11] C. Wen, W. Shih, and S. Jin, "Deep learning for massive mimo csi feedback," *IEEE Wireless Communications Letters*, pp. 1–1, 2018.

[12] 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, 2018.

[13] S. Omar, A. Ancora, and D. T. M. Slock, "Performance analysis of general pilot-aided linear channel estimation in lte ofdma systems with application to simplified mmse schemes," in *2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–6, Sept 2008.

[14] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, pp. 295–307, Feb 2016.

[15] J. Yamanaka, S. Kuwashima, and T. Kurita, "Fast and Accurate Image Super Resolution by Deep CNN with Skip Connection and Network in Network," *ArXiv e-prints*, July 2017.

[16] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," *IEEE Transactions on Image Processing*, vol. 26, pp. 3142–3155, July 2017.

[17] C. Mehlführer, J. Colom Ikuno, M. Šimko, S. Schwarz, M. Wrulich, and M. Rupp, "The vienna lte simulators - enabling reproducibility in wireless communications research," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, p. 29, Jul 2011.