CSI feedback compression based on deep learning using wavelet transform

Yong Jin Kwon, Anseok Lee and Heesoo Lee
Terrestrial and Non-Terrestrial Integrated Telecommunications Research Laboratory
Electronics and Telecommunications Research Institute (ETRI)
Daejeon, Republic of Korea
{yjkwon, alee, heelee}@etri.re.kr

***Abstract*—Recently, machine learning approaches have been widely applied in the mobile communication field. This paper presents a novel deep learning (DL) model for Channel State Information (CSI) feedback compression using wavelet transform. Specifically, the proposed model incorporates wavelet transform into the DL model. And, the model is trained with a combined loss function to effectively preserve high-frequency components. Based on extensive simulations using New Radio (NR) channel model, the CSI reconstruction performance of the proposed model is improved compared to an existing DL-based method for CSI feedback compression.**

***Index Terms*—CSI feedback Compression, Deep Learning, Wavelet Transform**

I. Introduction

In current communication systems, massive multiple-input multiple-output (MIMO) is essential to provide higher throughput to serving users [1]. To guarantee performance of multiple users in Downlink, it is important to obtain precise channel information. However, obtaining precise wireless channel information requires more overhead transmitted from the User Equipment (UE) to the base station (BS). Therefore, the trade-off between overhead reduction and precision of channel feedback could be an interesting research topic.

With the development of machine learning (ML) methods, various research efforts in academia and industry are actively conducted to apply ML to mobile communication systems. For example, 3rd Generation Partnership Project (3GPP), an international standardization organization, is conducting a study on applying ML technology to the air interface of 5G New Radio (NR) systems [2]. In this study item, CSI feedback compression is discussed as a representative use case to increase efficiency using ML technology.

In the direction of applying ML to CSI feedback compression, an Autoencoder, one of the deep learning (DL) structure, is commonly investigated to obtain a compressed latent space for MIMO channel. It was shown that CSI feedback can be effectively compressed by treating the wireless channel as an image and training it with Convolutional Neural Network (CNN) based autoencoder [3]. Compression using eigenvectors extracted from the wireless channel, rather than the full channel matrix itself, is superior to the existing 3GPP codebook-based method [4]. However, existing research has not considered a more diverse and complicated channel environment. In this paper, we performed diverse simulations using the 3GPP NR channel model. The neural network structure for CSI feedback compression should consider various channel environment during the training period to improve performance. Therefore, we aimed to present a novel deep learning structure that can be utilized in real-world wireless system.

Wavelet transform is a method of decomposing an arbitrary signal using functions known as wavelets. Wavelet transform is suitable for representing signals and images using only a small number of coefficients, and it is commonly used to reduce noise or compress images especially in the field of image processing. In [5], wavelet based autoencoder was applied to image compression in order to preserve various frequency domains of images. For the method of performing wavelet transformation, a lifting scheme, a second-generation wavelet generation method, is proposed in [6]. CNN using the lifting scheme shows comparable performance in image reconstruction [7].

In the general case of an autoencoder structure for CSI feedback compression, if the dimension of the latent space is not large enough, components corresponding to high frequencies may not be preserved during compression. Therefore, reconstruction performance can be degraded. Moreover, in the case of CNN-based neural networks which are widely used in the CSI feedback compression, this structure was developed in a situation where continuity with surrounding data was secured to some extent such as pictures. On the other hand, channel information such as eigenvectors has relatively low continuity with surrounding data. Therefore, there is a need to resolve this difference. To overcome this situation, we propose a deep autoencoder model for CSI feedback compression using wavelet transform in this paper. With the use of wavelet transform, an original channel image can be decomposed into multiple channel images of different frequencies.

The remainder of this paper is organized as follows, In Section II, we describe the system model in which CSI feedback is performed. Section III explains the wavelet transform and the proposed autoencoder structure using wavelet transform. In Section IV, the simulation results are presented and the excellence of the proposed method is investigated. Finally, we conclude this paper in Section V.

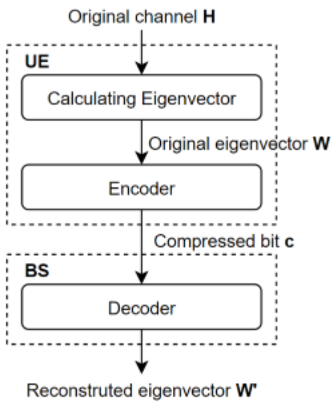

Fig. 1. Autoencoder architecture for CSI feedback compression.

II. SYSTEM MODEL

In this paper, we consider a typical massive MIMO transmission system. The BS which has N_{tx} transmission antenna ports transmits data to the UE with N_{rx} receiver antenna ports. It is assumed that the UE measures CSI feedback using the pilot signal transmitted by the BS. The frequency domain of the measured channel consists of K subbands, with each subband including 4 resource blocks (RB). The Downlink channel of can be written as

$$H = [H_1, H_2, ..., H_K], \quad (1)$$

where $H_k \in \mathbb{C}^{N_{tx} \times N_{rx}}, 1 \leq k \leq K$ indicates Downlink channel of the Kth subband.

In this study, it is assumed that the eigenvector calculated from downlink channel H is compressed and transmitted. This is because, feedback of eigenvector can reduce overhead compared to feedback of the full channel matrix. Moreover, we consider actual CSI information transmitted from the UE to the BS in 5G NR standard. The eigenvector is calculated by using Singular Vector Decomposition (SVD) which is widely used in MIMO system.

And, the eigenvector of the DL channel in K subbands can be written as

$$W = [W_1, W_2, ..., W_K], \quad (2)$$

where $W \in \mathbb{C}^{N_{tx} \times K}$ and w_k with normalization $\left\|w_k^2\right\| = 1$.

In this paper, NR channel models defined in 3GPP for system-level simulations are utilized. Specifically, Urban Micro (UMi) represents typical outdoor scenarios. More details about these NR channel models can be founded in 3GPP specification 38.901 [8].

In CSI feedback compression, the basic configuration of autoencoder is displayed in Fig. 1. The UE compresses the measured eigenvector using an encoder.

$$c = f_{enc}(W) \quad (3)$$

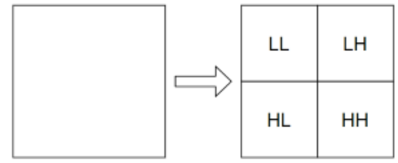

Fig. 2. Two-dimensional wavelet transform.

The compressed bits c are transmitted to the BS through the air interface. Then, the BS decompresses received bits using a decoder and reconstructs the eigenvector.

$$W = f_{dec}(c) \quad (4)$$

III. PROPOSED METHOD

A. Wavelet transform

With wavelet transform, the original image is divided into different images in multiple frequency bands. Fig. 2 shows an example of an image transformed by two-dimensional wavelet transform. The entire image on the left side is decomposed into four regions which has the same total overall size. And, each region has the half height and width of original images. On the right side, L means low-frequency band information and H means high-frequency band information. Therefore, the LL part includes low-frequency band information in both the horizontal and vertical directions. Conversely, images of different frequency bands (LL, HL, LH, HH) can be combined with the original image using inverse wavelet transform.

The lifting scheme, the second generation of making wavelet transform, consists of three steps: split, prediction and update. In the split step, the input signal is divided into two separate signals, which are generally separated into odd and even parts, x_e and x_o. In prediction process, coefficients in a high-frequency region, x_d, are calculated by taking the difference between odd part and the result obtained from the even part after filtering.

$$x_d = x_o - P(x_e), \quad (5)$$

where $P(\cdot)$ represents prediction filter. Then, in the update process, the even part and the filtering result from the odd part are added to obtain coefficients in the low-frequency region.

$$x_c = x_e + U(x_d), \quad (6)$$

where $U(\cdot)$ indicates update filter.

B. Proposed DL model

Wavelet transform described above is incorporated into the DL model for CSI feedback compression. As shown in Fig. 3, the architecture of proposed DL model is detailly explained. In this structure, wavelet transform and inverse wavelet transform are included in the first part of the encoder and the last part of the decoder.

Input joint eigenvector W is consists of K eigenvectors from each subband. And, each eigenvector composed of real

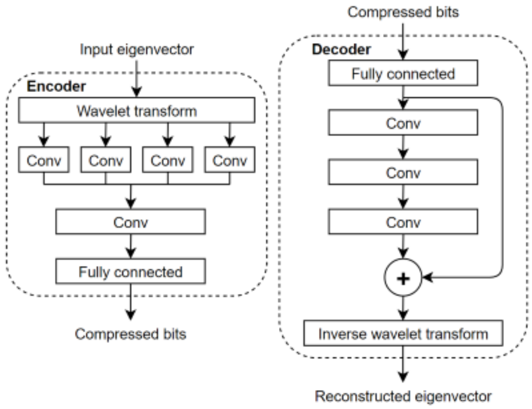

Fig. 3. Proposed DL model.

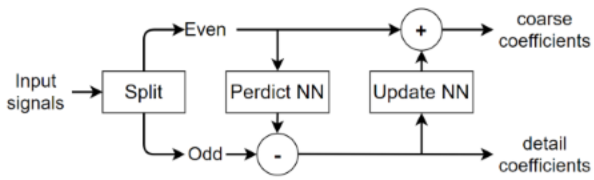

Fig. 4. Lifting scheme with neural networks.

and imaginary parts. Thus, input dimension is $2 \times K \times N_{tx}$. In the encoder, wavelet transform is performed on the input eigenvector, and the results are converted to compressed bits. More specifically, the input eigenvector is partitioned into four smaller images by wavelet transform, and the dimension of each image is $2 \times K/2 \times N_{tx}/2$. Afterward, each image passes through 2D convolutional layer with 2 filters and 3x3 kernel size, and they are merged into one image with 8 feature channels. In the above process, the leaky Rectified linear unit (ReLU) activation function is used. Next, the image is converted into compressed bits through a fully connected (FC) layer.

In the decoder, eigenvector reconstruction from compressed bits is performed using a more complicated structure compared to the encoder. First, the FC layer is used to change compressed bits back to the original dimension. The output of FC layer is then reshaped into $K/2 \times N_{tx}/2$ image with 8 feature channels. After that, 2D convolution layers are successively utilized to capture deep features. In this part, we modify RefineNet structure proposed in CsiNet [3], and the well-known shortcuts structure is also used. After repeated use of RefineNet, inverse wavelet transform is adopted to convert the image of 8 feature channels into an image of the original size with two feature channels.

Instead of using existing wavelets, wavelet transform that is trained on the currently used input can be included in the deep learning model. In the lifting scheme, parts corresponding to prediction and update are replaced to the neural network as shown in Fig. 4. The predict and update neural networks learn wavelet transform and inverse wavelet transform suitable for the current input during the training period. These networks consist of multiple 2D convolution layers with 3×3 kernel size.

C. Loss function

For the training of the proposed DL model, we utilize a combined loss function that incorporates both the MSE loss of original image and the transformed image. The first MSE loss of original eigenvector indicates the difference between the joint eigenvector and the restored eigenvector from the decoder.

$$L_{MSE_1}(W, W') = \frac{1}{\|W\|} \sum \|W - W'\|^2, \tag{7}$$

where $\|\cdot\|$ represents Euclidian norm.

Θ from the second MSE loss represent high-frequency components after wavelets transform such as LH, LH, and HH. By using the loss function of high-frequency parts, the model can preserve more high-frequency information.

$$L_{MSE_2}(\Theta, \Theta') = \frac{1}{\|\Theta\|} \sum \|\Theta - \Theta'\|^2 \tag{8}$$

The combined loss function is written as

$$L = \alpha_1 L_{MSE_1}(W, W') + \alpha_2 L_{MSE_2}(\Theta, \Theta'), \tag{9}$$

where α_1 and α_2 represents the weights of the loss function.

IV. Simulation Results

A. Simulation environment

In order to train the proposed DL model and evaluate its performance, channel dataset is created using NR channel model. In NR Urban Micro (UMi) environment, 19 BSs are located in 3 tiers, and 30 UEs are uniformly generated in each BS. In one drop, samples of channel data for 570 UEs are created. After multiple drops, these samples were used for training and testing. During training period, 200 epochs are conducted using the Adam optimizer with an adaptive learning rate, starting from $1e^{-3}$ and ending at $1e^{-4}$. Additionally, the batch size set to 128. The detailed simulation parameters are summarized in Table I.

We compared the proposed DL model with CsiNet [3], one of the existing DL methods for CSI feedback compression. For a fair comparison, we modified CsiNet to have the same

TABLE I
Simulation Parameters

Parameters	Value
Carrier frequency	4GHz
Bandwidth	10MHz
Subcarrier spacing	15KHz
Subband number	12
Tx Antenna port	32
Rx Antenna port	4
Channel model	NR UMi
Number of train samples	27,000
Number of test samples	3,000

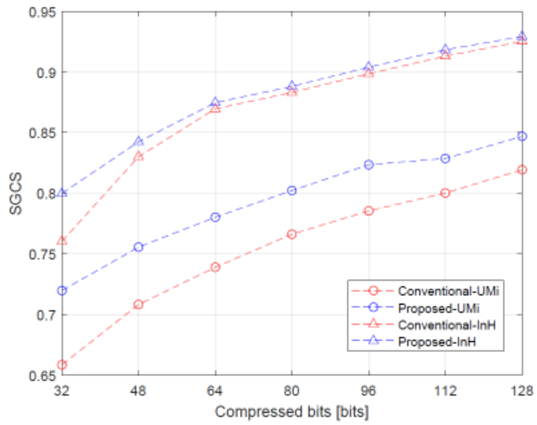

Fig. 5. Performance comparison of the proposed and conventional methods for SGCS.

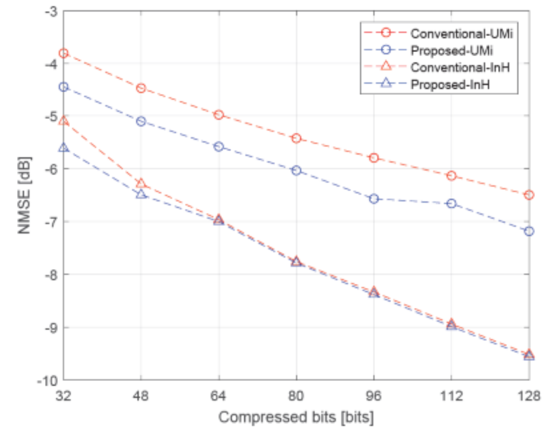

Fig. 6. Performance comparison of the proposed and conventional methods for NMSE.

structure as our proposed model except for wavelet transform. Two performance metrics are used to measure the performance of the reconstructed eigenvector. The first one is normalized MSE (NMSE), which quantifies the difference between the restored and the original eigenvector. Additionally, Squared Generalized Cosine Similarity (SGCS) is used to compare the direction of the eigenvector. And, SGCS can be written as

$$\sigma_k^2 = \left(\frac{\left\| w_k^H w_k' \right\|}{\left\| w_k^H \right\| \left\| w_k' \right\|} \right)^2 \tag{10}$$

SGCS has a value between 0 and 1, where a value of 1 indicates a perfect match with the original direction.

B. Channel reconstruction performance

As shown in Fig. 5, a performance comparison is presented between the proposed model using wavelet transform and the conventional DL method for SGCS. This comparison was conducted by varying the size of compressed bits. As the size of compressed bits decreases, the compression rate increases compared to the dimension of the original eigenvector. However, the overall performance is reduced due to the loss of more information during compression.

According to the SGCS results, the UMi scenario exhibits an approximate 3.4% performance improvement compared to the conventional method. In contrast, the indoor scenario shows a relatively modest performance improvement of 0.5%. This is because of the characteristics of channel model in the indoor scenario, which is relatively easier to train and where Non Line-of-Sight (NLOS) conditions frequently occur. On the other hand, in the UMi environment, Line-of-Sight (LOS) conditions also occurs. This setting results in eigenvectors having a sharper appearance and containing high-frequency components. By utilizing the proposed model with wavelet transform, these high-frequency components are better preserved, consequently leading to enhanced performance.

NMSE results in Fig. 6 demonstrate the same trend as the as SGCS results. It is evident that the performance gain is at least -0.52dB in the UMi scenario. Conversely, in the indoor environment, the performance gain decreases to below -0.05dB.

V. Conclusion

In this paper, we have proposed the DL model for CSI feedback compression using wavelet transform. Specifically, we utilized wavelet transform and the modified loss function to preserve high-frequency components. And, the simulation results with NR channel model demonstrate that the proposed model can improve CSI reconstruction performance compared to the existing DL-based method. Therefore, this method offers overhead reduction and performance improvement in MIMO environment.

Acknowledgment

This work was supported by Institute of Information & communications Technology Planning and Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-00972, Development of Intelligent Wireless Access Technology)

References

[1] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," IEEE J. Sel. Topics Signal Process., vol. 8, no. 5, pp. 742–758, Oct. 2014.
[2] RP-213599, "New SI: Study on Artificial Intelligence (AI)/Machine Learning (ML) for NR air interface," Qualcomm, 3GPP RAN #94-e, December 2021.
[3] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep learning for massive MIMO CSI feedback," IEEE Wireless Commun. Lett., vol. 7, no. 5, pp. 748–751, Oct. 2018.
[4] W. Liu, W. Tian, H. Xiao, S. Jin, X. Liu, and J. Shen, "EVCsiNet: Eigenvector-based CSI feedback under 3GPP link-level channels," IEEE Wireless Commun. Lett., vol. 10, no. 12, pp. 2688–2692, Dec. 2021.
[5] D. Mishra, S. K. Singh, and R. K. Singh, , "Wavelet-based deep auto encoder-decoder (WDAED)-based image compression," IEEE Trans. Circuits Syst. Video Technol., vol. 31, no. 4, pp. 1452–1462, Apr. 2021.
[6] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," Appl. Comput. Harmon. Anal., vol. 3, no. 2, pp. 186–200, 1996
[7] H. Ma, D. Liu, R. Xiong and F. Wu, "iWave: CNN-Based Wavelet-Like transform for image compression," in IEEE Transactions on Multimedia, vol. 22, no. 7, pp. 1667-1679, July 2020.
[8] "3rd generation partnership project; technical specification group radio access network; NR; physical layer procedures for data (release 16)," 3GPP, Sophia Antipolis, France, 3GPP Rep. TR 38.901.