Parallel Network Assisted Calibrated Beam Training for mmWave Communication Systems

Jihyung Kim
Spatial Wireless Transmission Research Section
Electronics and Telecommunications Research Institute
Daejeon, South Korea
savant21@etri.re.kr

Soyoung Yoo, Junghyun Kim
Department of Artificial Intelligence
Sejong University
Seoul, South Korea
yooso0731@gmail.com, j.kim@sejong.ac.kr

***Abstract*—** This paper proposes a new structure called CNN-PN-LSTM to improve the beam prediction performance of existing deep learning-based calibrated mmWave beam training. Unlike previous works, we utilized a parallel network to effectively extract features from high-dimensional signals for model training. Simulation results show the effectiveness of the parallel network and the superior prediction performance of our model.

Keywords—beam prediction; mmWave communications; deep learning; parallel network

I. Introduction

Millimeter-wave (mmWave) communication systems have become increasingly popular for high-speed wireless data transfer due to their large bandwidths. However, the signals suffer from high attenuation and sensitivity to blockages, making beamforming critical for reliable communication links. Traditional beamforming methods rely on narrow beams, which require accurate channel information and are prone to experiencing beam misalignment due to mobility or environmental changes. Moreover, creating and measuring a large number of narrow candidate beams is necessary to cover the entire angular space, leading to a significant training overhead. One straightforward approach to alleviate this overhead is to limit the search to beams with a fixed angular spacing and employ the received signals to forecast the best beam. Nonetheless, this method's effectiveness might deteriorate under low signal-to-noise ratio (SNR) conditions, where the true angle-of-arrival (AoA) and angle-of-departure (AoD) of the main path do not coincide precisely with the middle of any candidate beam's main lobe.

Deep learning (DL) has been attracting attention in wireless communications based on its remarkable achievements in natural language processing and computer vision [1]. At the physical layer, DL has made significant strides in terms of performance and efficiency in several areas such as channel state information (CSI) feedback [2], channel denoiser [3], channel decoders [4, 5], end-to-end transceivers [6, 7], and beam prediction [8-12]. In this paper, we focus on DL-based beam prediction. A wide beam based best narrow beam prediction scheme was proposed in [8] to cover the overall angular space. The authors made a comparison between super-resolution beam prediction and super-resolution image restoration, and chose to use a convolutional neural network (CNN) for beam prediction. Ma et al. [12] also presented a CNN-based model for narrow beam prediction using received signals obtained from wide beam training. However, since the model lacks the robustness to noise, a different approach that utilizes long-short term memory (LSTM) was introduced to track user movements and calibrate beam direction using the received signals from previous beam training sessions. The combined CNN and LSTM model proposed in [12] improves robustness against noise, but still does not provide sufficient beamforming gain. Specifically, in order to extract complex features of continuous measurements, a two-layer CNN block having 64 and 256 filters is used to extend the real and imaginary two-dimensional signals to a 256-dimensional signal. However, through experiments, we found that a deeper or more efficient neural network structure is needed to effectively extract the features of the corresponding signals. In addition, there was a limit to overcome insufficient feature values with the LSTM structure.

Since the emergence of AlexNet [13], in the area of deep learning, neural networks have generally been designed to be deep in order to extract complex features with nonlinearity. However, with the proposals of backbone structures such as DenseNet [15], ResNet [14], and EfficientNet [16], research on increasing the efficiency of models using methods other than depth has become active. Recently, ParNet [17] demonstrated remarkable performance despite its shallow structure. This paper aims to improve the learning efficiency of CBT (Calibrated Beam Training) tasks by utilizing the parallel network structure of the ParNet. Through experiments conducted in the same environment as in [12], we confirmed that inserting the modified ParNet block into the combined CNN and LSTM model in [12] significantly increases the beamforming gain.

II. System Model

A. Channel Model

Our study focuses on MIMO communication in mmWave, where base station (BS) and user equipment (UE) are fitted with M_{Tx} and M_{Rx} antennas. We consider a two-dimensional channel configuration where the antennas are arranged in uniform linear arrays (ULAs). The narrowband frequency-flat channel model with a line-of-sight (LOS) path and C clusters is employed. We use $\boldsymbol{H}_{LOS}$ and $\boldsymbol{H}_{NLOS}$ to denote the LOS and non-LOS components of the channel matrix $\boldsymbol{H} \in \mathbb{C}^{M_{Rx} \times M_{Tx}}$, respectively. Each term can be expressed as

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. 2022-0-00436, Development of Standard Technologies for Reconfigurable Intelligent Surface Repeaters)

$$\boldsymbol{H}_{LOS} = \sqrt{\frac{M_{Tx}M_{Rx}}{\rho_{LOS}}}\alpha_{LOS}\boldsymbol{a}_{Rx}(\theta_{LOS})\boldsymbol{a}_{Tx}^{H}(\phi_{LOS}), \tag{1}$$

$$\boldsymbol{H}_{NLOS} = \sum_{c=1}^{C}\sqrt{\frac{M_{Tx}M_{Rx}}{\rho_c}}\sum_{l=1}^{L_c}\frac{\alpha_{c,l}}{\sqrt{L_c}}\boldsymbol{a}_{Rx}(\theta_c+\theta_{c,l})\boldsymbol{a}_{Tx}^{H}(\phi_c+\phi_{c,l}). \tag{2}$$

In this model, the LOS path has pathloss ρ_{LOS}, AoA θ_{LOS}, and AoD ϕ_{LOS}. Similarly, the NLOS path has the c-th cluster containing L_c paths has pathloss ρ_c, AoA θ_c and AoD ϕ_c, while $\alpha_{c,l}$, $\theta_{c,l}$, and $\phi_{c,l}$ are the complex gain, AoA offset, and AoD offset, respectively, corresponding to the l-th path in the c-th cluster. The antenna response vectors are represented as

$$\boldsymbol{a}_{Tx}(\phi) = \sqrt{\frac{1}{M_{Tx}}}\left[1\ e^{j2\pi d_{Tx}\sin\phi/\lambda} \dots e^{j2\pi(M_{Tx}-1)d_{Tx}\sin\phi/\lambda}\right]^T, \tag{3}$$

$$\boldsymbol{a}_{Rx}(\theta) = \sqrt{\frac{1}{M_{Rx}}}\left[1\ e^{j2\pi d_{Rx}\sin\theta/\lambda} \dots e^{j2\pi(M_{Rx}-1)d_{Rx}\sin\theta/\lambda}\right]^T, \tag{4}$$

where d_{Tx} and d_{Rx} are the antenna spacings, λ denotes the wavelength. For simplicity, we set $d_{Tx} = d_{Rx} = \lambda/2$.

B. Beam Training Model

We consider the utilization of phase shifter-based analog beamforming, where the transmitting beam of the BS is denoted by $\boldsymbol{f} \in \mathbb{C}^{M_{Tx}\times 1}$, and the receiving beam of the UE is denoted by $\boldsymbol{w} \in \mathbb{C}^{M_{Rx}\times 1}$. We employ a discrete Fourier transform (DFT) codebook, and the candidate transmitting beam $\boldsymbol{f}_m$, $m \in \{1, 2, \dots, N_{Tx}\}$, and receiving beam $\boldsymbol{w}_n$, $n \in \{1, 2, \dots, N_{Rx}\}$, can be expressed as

$$\boldsymbol{f}_m = \sqrt{\frac{1}{M_{Tx}}}\left[1\ e^{j\pi\sin\gamma_{Tx,m}} \dots e^{j\pi(M_{Tx}-1)d_{Tx}\sin\gamma_{Tx,m}}\right]^T, \tag{5}$$

$$\boldsymbol{w}_n = \sqrt{\frac{1}{M_{Rx}}}\left[1\ e^{j\pi\sin\gamma_{Rx,n}} \dots e^{j\pi(M_{Rx}-1)\sin\gamma_{Rx,n}}\right]^T, \tag{6}$$

where $\gamma_{Tx,m}$ and $\gamma_{Rx,n}$ represent the beam directions of the m-th candidate beam on the BS side and the n-th candidate beam on the UE side, respectively.

Assuming the channel matrix $\boldsymbol{H}$ and beam pair $\{\boldsymbol{f}, \boldsymbol{w}\}$ are given, the received signal y can be expressed as

$$y = \sqrt{P}\boldsymbol{w}^H\boldsymbol{H}\boldsymbol{f}x + \boldsymbol{w}^H\boldsymbol{n}, \tag{7}$$

where x is the transmitted signal with $|x| = 1$, P is the transmit power, and $\boldsymbol{n} \in \mathbb{C}^{M_{Rx}\times 1}$ denotes AWGN vector with zero mean and variance of σ^2, i.e., $\boldsymbol{n}\sim\mathcal{CN}(\boldsymbol{0}_{M_{Rx}}, \sigma^2\boldsymbol{I}_{M_{Rx}})$.

The objective of beam training is to identify the beam pair $\{\boldsymbol{f}_{m^*}, \boldsymbol{w}_{n^*}\}$ that offers the maximum gain. This can be expressed as an optimization problem shown below:

$$\{m^*, n^*\} = \arg\max_{\substack{m\in\{1,2,\dots,N_{Tx}\},\\ n\in\{1,2,\dots,N_{Rx}\}}} |\boldsymbol{w}_n^H\boldsymbol{H}\boldsymbol{f}_m|^2. \tag{8}$$

We consider a system that predicts the optimal narrow beam among N_{Tx} candidate narrow beams from measurements of N_{Tx}/s_{Tx} wide beams to reduce the overhead. Additionally, we train the prediction model to be robust to noise using continuous measurements of wide beams.

III. Parallel Network Assisted Calibrated Beam Training

A. Model Design

In [12], the authors used a combined architecture of CNN and LSTM for the CBT task. The CNN block expands the two-dimensional received signal into a 256-dimensional signal to facilitate feature extraction, while the LSTM structure utilizes the continuous input signal to be robust to noise and enables the learning of the signal changes over time. However, through experiments, we have confirmed that the two-layer CNN block is inefficient in extracting the features, and relying only on the LSTM structure has limitations in solving this issue. Therefore, we propose a new structure that can efficiently extract features while minimizing the complexity of the overall structure.

The proposed beam prediction model for the CBT task is shown in Fig. 1. After performing the t-th training, the received wide beam signals $\{\boldsymbol{y}_{w,1}, \boldsymbol{y}_{w,2}, \dots, \boldsymbol{y}_{w,t}\}$ are initially processed through the preprocessing, CNN, and ParNet modules to retrieve characteristics related to the received signals. Then, the LSTM module utilizes the received signals from current and past beam training sessions to further refine the narrow beam direction. Subsequently, the output module produces the probabilities $\{\hat{p}_{1,t}, \hat{p}_{2,t}, \dots, \hat{p}_{N_{Tx},t}\}$, and the narrow beam with the highest probability is chosen as the best beam.

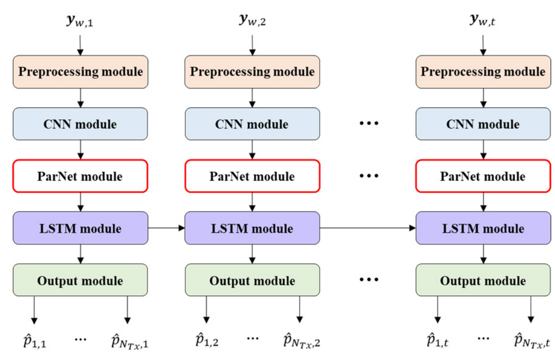

Fig. 1. Proposed CNN-PN-LSTM model for the CBT task.

B. Parallel Network

The original ParNet block includes three parallel branches: a 1×1 convolution, a 3×3 convolution, and a Skip-Squeeze-and-Excitation (SSE) component. We modified the original ParNet block using 1×3 convolution and inserted the block between CNN and LSTM modules to efficiently extract characteristics from the received signals. The proposed ParNet block is illustrated in Fig. 2.

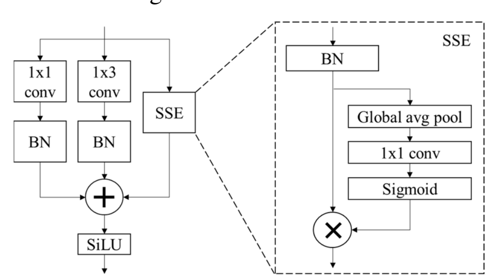

Fig. 2. Proposed ParNet block.

IV. Experimental Results

We performed simulations using the COST2100 channel model [18] as in [12]. As candidates for comparison, we considered the CNN model and the combined CNN and LSTM model in [12] as well as the following three baseline models: 1) A noise-free approach to beam prediction presented in [9], which uses N_{Tx}/s_{Tx} measurements of wide beams; 2) A DL-based beam prediction method proposed in [10], which employs N_{Tx}/s_{Tx} measurements of narrow beams uniformly sampled; 3) An adaptive and sequential beam alignment scheme introduced in [11], which leverages N_{Tx}/s_{Tx} measurements of hierarchical beams and targets a resolution of N_{Tx}.

To begin with, we investigated how the number of training sessions t for wide beams affects the attainable normalized beamforming gain G_N, as depicted in Fig. 3. It is worth noting that the CNN-assisted CBT scheme and all three baseline approaches do not depend on any prior information, so the beamforming gain does not change with t. The CNN-LSTM assisted CBT scheme demonstrates improved performance as t increases. However, when $t = 1$, it shows inferior performance compare to the CNN assisted CBT, and it can be observed that There exists a constraint on enhancing the performance beyond a certain level with only the LSTM structure. On the other hand, the proposed scheme shows the best performance for all t due to the effect of the ParNet block. In addition, the superiority of the proposed scheme is further evidenced by the cumulative distribution function (CDF) of the predicted narrow beam gain shown in Fig. 4. Both of Fig. 3 and Fig. 4 confirm that our scheme shows excellent prediction performance compared to the five previous schemes.

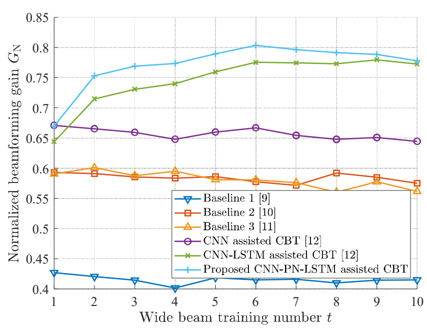

Fig. 3 Normalized beamforming gain over wide beam training number.

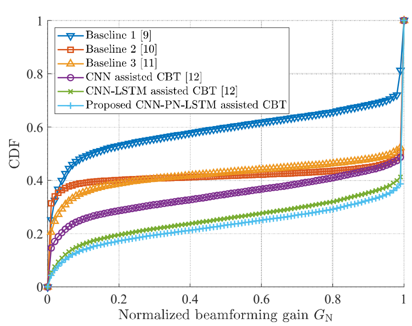

Fig. 4. CDF comparison of the predicted narrow beam gains.

V. Concluding Remarks

We addressed DL-based calibrated mmWave beam training for wireless communication systems. The use of the parallel network has significantly improved the narrow beam prediction performance by effectively extracting the features of received wide beam signals. According to the experimental results, the proposed scheme showed the best performance compared to existing schemes for all of the number of wide beam trainings. Future directions of this work include reducing the overall complexity and improving the performance of LSTM block.

References

[1] H. Lee, B. Lee, H. Yang, J. Kim, S. Kim, W. Shin, B. Shim, and H. V. Poor, "Towards 6G Hyper-Connectivity: Vision, Challenges, and Key Enabling Technologies," Journal of Communications and Networks, in press.

[2] J. Guo, C.-K. Wen, S. Jin, and X. Li, "AI for CSI Feedback Enhancement in 5G-Advanced," IEEE Wireless Communications, early access, Dec. 5, 2022, doi: 10.1109/MWC.010.2200304.

[3] S. Han, J. Kim, and H.-Y. Song, "A New Design of Channel Denoiser using Residual Autoencoder," IET Electronics Letters, vol. 59, no. 2, pp. 1-3, Jan. 2023.

[4] H. Kim, S. Oh, and P. Viswanath, "Physical Layer Communication via Deep Learning," IEEE Journal on Selected Areas in Information Theory, vol. 1, no. 1, pp. 5-18, May 2020.

[5] Y. Choukroun and L. Wolf, "Denoising Diffusion Error Correction Codes," 2022. [Online]. Available: arXiv:2209.13533.

[6] T. O'Shea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," IEEE Transactions on Cognitive Communications and Networking, vol. 3, no. 4, pp. 563-575, Oct. 2017.

[7] J. Kim, B. Lee, H. Lee, Y. Kim, and J. Lee, "Deep Learning-assisted Multi-dimensional Modulation and Resource Mapping for Advanced OFDM Systems," in Proc. IEEE Globecom Workshops, Abu Dhabi, UAE, Dec. 2018, pp. 1-6.

[8] H. Echigo, Y. Cao, M. Bouazizi, and T. Ohtsuki, "A Deep Learning-based Low Overhead Beam Selection in mmWave Communications," IEEE Trans. Veh. Tech., vol. 70, no. 1, pp. 682-691, Jan. 2021.

[9] X. Luo, W. Liu, and Z. Wang, "Calibrated Beam Training for Millimeter-wave Massive MIMO Systems," in Proc. IEEE Vehicular Technology Conference (VTC)-Fall, Honolulu, HI, USA, Sep. 2019, pp. 1-5.

[10] C. Qi, Y. Wang, and G. Y. Li, "Deep Learning for Beam Training in Millimeter Wave Massive MIMO Systems," IEEE Trans. Wireless Commun., early access, Sep. 22, 2020, doi: 10.1109/TWC.2020.3024279.

[11] S.-E. Chiu, N. Ronquillo, and T. Javidi, "Active Learning and CSI Acquisition for mmWave Initial Alignment," IEEE J. Sel. Areas Commun., vol. 37, no. 11, pp. 2474-2489, Nov. 2019.

[12] K. Ma, D. He, H. Sun, Z. Wang, and S. Chen, "Deep Learning Assisted Calibrated Beam Training for Millimeter-wave Communication Systems," IEEE Trans. Commun., vol. 69, no. 10, pp. 6706-6721, Oct. 2021.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Proc. Int. Conf. Neural Inf. Process. Syst. (NIPS), Dec. 2012, pp. 1097-1105.

[14] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 4700-4708.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770-778.

[16] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proc. Int. Conf. Mach. Learn. (ICML), Jun. 2019, pp. 6105-6114.

[17] A. Goyal, A. Bochkovskiy, J. Deng, and V. Koltun, "Non-deep Networks," 2021. [Online]. Available: arXiv:2110.07641.

[18] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. D. Doncker, "The COST 2100 MIMO Channel Model," IEEE Wireless Commun., vol. 19, no. 6, pp. 92-99, Dec. 2012.