LETTER

Deep Learning-Based CSI Feedback for Terahertz Ultra-Massive MIMO Systems

Yuling LI[†], *Student Member and* **Aihuang GUO**[†a)], *Nonmember*

SUMMARY Terahertz (THz) ultra-massive multiple-input multiple-output (UM-MIMO) is envisioned as a key enabling technology of 6G wireless communication. In UM-MIMO systems, downlink channel state information (CSI) has to be fed to the base station for beamforming. However, the feedback overhead becomes unacceptable because of the large antenna array. In this letter, the characteristic of CSI is explored from the perspective of data distribution. Based on this characteristic, a novel network named Attention-GRU Net (AGNet) is proposed for CSI feedback. Simulation results show that the proposed AGNet outperforms other advanced methods in the quality of CSI feedback in UM-MIMO systems.

key words: *CSI feedback, UM-MIMO system, terahertz communication, deep learning, channel characteristic*

1. Introduction

Terahertz (THz) band communication is envisioned as a key enabling technology, promising to provide Terabits-per-second data rates in 6G wireless communications [1]. Nevertheless, the high propagation loss at THz band limits the coverage range [2]. To combat the shortcoming, ultra-massive multiple-input multiple-output (UM-MIMO) systems with an array-of-subarray (AoSA) structure have been proposed [3]. In the AoSA structure, a large antenna array is divided into multiple subarrays (SAs), and each SA is powered by a single radio frequency (RF) chain. Based on this structure, the highly directional hybrid beamforming can be performed to enhance the capacity of UM-MIMO systems [4].

Retrieval of accurate downlink channel state information (CSI) in the base station (BS) is critical for beamforming design [5]. In frequency division duplex (FDD) systems, the downlink CSI has to be fed back from user equipment (UE). However, a dramatically increasing feedback overhead is foreseen due to the unprecedented massive number of antennas, which makes the feedback extremely challenging [6].

Recently, various techniques have been proposed for CSI feedback in massive MIMO systems. In [7], a state-of-the-art compressed sensing (CS)-based method named TVAL3 has been proposed. In [8]–[10], deep learning (DL)-based methods that borrow the idea of autoencoder architecture have been proposed. CsiNet proposed in [8] first demonstrates the advantages of DL in CSI feedback. CRNet proposed in [9] introduces a multi-resolution architecture

and an advanced training scheme. TransNet proposed in [10] adopts a two-layer Transformer architecture and greatly improves the accuracy. For CSI feedback in THz UM-MIMO systems, the CS-based methods rely on channel sparsity that is difficult to meet practically, while the DL-based methods also encounter some challenges. First, the commonly used convolutional neural network (CNN) is difficult to extract the multi-domain correlation of CSI since the convolution operations are performed locally [6]. Second, common methods of concatenating the real and imaginary parts of CSI as input will result in enormous network parameters due to the huge channel dimensions, which impede the practical deployment. Third, existing works typically exploit dedicated sparsity in the far-field to design networks. In fact, the far- and near-field paths co-exist and together form the hybrid-field channel in UM-MIMO systems [12]. The changeable channel conditions have to be considered in CSI feedback.

In this letter, a novel NN named Attention-GRU Net (AGNet) is proposed to tackle the CSI feedback problem in THz UM-MIMO systems. The hybrid-field THz UM-MIMO channel model is developed to generate CSI samples. The distribution characteristic of CSI is measured by the distance-based separability index (DSI) [11]. Based on this characteristic, a one-part training scheme is proposed to reduce the network size. The fusion of Gate Recurrent Unit (GRU) and attention mechanism is designed to flexibly extract the correlation and the changeability features of UM-MIMO CSI. Simulation results demonstrate that AGNet achieves better feedback accuracy compared to representative methods based on CS, CNN, and Transformer.

The remainder of this letter is organized as follows. The THz UM-MIMO system model is established in Sect. 2. The distribution characteristic of CSI and the design of AGNet are explained in Sect. 3. The simulation results are presented in Sect. 4, and the conclusion is given in Sect. 5.

2. System Model

We consider a THz UM-MIMO FDD system as shown in Fig. 1. The BS adopts a planar AoSA with $\sqrt{Q} \times \sqrt{Q}$ SAs for hybrid beamforming, while each SA is fed by one RF chain and contains $\sqrt{\bar{Q}} \times \sqrt{\bar{Q}}$ uniformly arranged antenna elements (AEs). The total number of antennas is $N_t = Q \times \bar{Q}$. We construct a three-dimensional Cartesian coordinate system with the origin being the first AE in the first SA and deploy the AoSA on the x-y plane. The UE is equipped with a single antenna for simplicity. Orthogonal frequency division
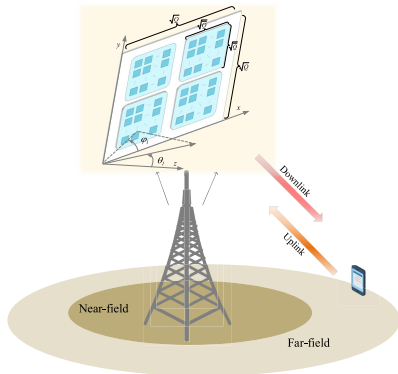
Fig. 1 A typical THz UM-MIMO FDD system.

multiplexing with N_c subcarriers is adopted, the downlink CSI matrix can be defined as

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 \cdots \mathbf{h}_{N_c} \end{bmatrix}^T \in \mathbb{C}^{N_c \times N_t} \tag{1}$$

where $\mathbf{h}_k \in \mathbb{C}^{N_t \times 1}$ denotes the channel response of the k-th subcarrier, $(\cdot)^T$ denotes the transpose operation.

In THz UM-MIMO systems, the main components of the channel are line-of-sight (LoS) and reflected paths, while the other multi-path effects such as scattering and diffraction can be ignored due to the high propagation loss [4]. Thus $\mathbf{h}_k$ can be denoted as

$$\mathbf{h}_k = \sum_{l=1}^{L} \alpha_l (f_k, r_l) \mathbf{a} (\phi_l, \theta_l, r_l) e^{-j2\pi f_k \tau_l} \tag{2}$$

where f_k is the frequency of the k-th subcarrier, $\mathbf{a}(\cdot)$ is the antenna array response, α_l is the path loss of the l-th path. ϕ_l, θ_l, r_l, and τ_l are azimuth angle of departure (AoD), elevation AoD, communication distance, and time delay of the l-th path, respectively. Assuming that $l = 1$ denotes the LoS path, $l > 1$ denotes reflected paths.

The path loss is composed of the spread loss and the molecular absorption loss, i.e.

$$\alpha_l (f_k, r_l) = |\Gamma_l| \left(\frac{c}{4\pi f_k r_l} \right)^{\frac{\gamma}{2}} e^{-\frac{1}{2} K_{abs}(f_k) r_l} \tag{3}$$

where c is the speed of light, Γ_l is the reflection factor, γ is the path loss exponent and $K_{abs}(f_k)$ is the frequency-selective molecular absorption loss [12].

Considering the notable near-field spherical propagation characteristic of the UM-MIMO channel [12], the array response adopts the hybrid-field assumption. The far-field is defined as the region where r_l is greater than or equal to the Rayleigh distance D; otherwise, it is the near-field. In the far-filed, the array response is

$$\mathbf{a}^{\text{far}} (\phi_l, \theta_l) = \begin{bmatrix} a_{1,1}^{\text{far}}, \cdots, a_{q,\bar{q}}^{\text{far}}, \cdots, a_{Q,\bar{Q}}^{\text{far}} \end{bmatrix}^T \tag{4}$$

$$a_{q,\bar{q}}^{\text{far}} = e^{-j2\pi \frac{f_k}{c} \mathbf{p}_{q,\bar{q}}^T \mathbf{t}_l} \tag{5}$$

where $\mathbf{t}_l = (sin\theta_l cos\varphi_l,\ sin\theta_l sin\varphi_l,\ cos\theta_l)^T$ is the unit vector in the AoD direction of the l-th path, $\mathbf{p}_{q,\bar{q}}$ is the three-dimensional coordinate of the $\bar{q}$-th AE in the q-th SA. In the
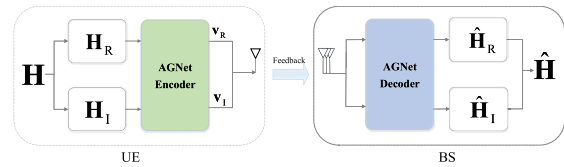

Fig. 2 Schematic diagram of AGNet aided CSI feedback workflow.

Table 1 DSI of the real and imaginary sets.

s_x	s_y	$DSI(\{R, I\})$
0.0203	0.0175	0.0189

near-filed, the array response depends on the exact distance between the AE and the UE and can be denoted as

$$\mathbf{a}^{\text{near}} (\phi_l, \theta_l, r_l) = \begin{bmatrix} a_{1,1}^{\text{near}}, \cdots, a_{q,\bar{q}}^{\text{near}}, \cdots, a_{Q,\bar{Q}}^{\text{near}} \end{bmatrix}^T \tag{6}$$

$$a_{q,\bar{q}}^{\text{near}} = e^{-j2\pi \frac{f_k}{c} \|\mathbf{p}_{q,\bar{q}} - r_l \mathbf{t}_l\|_2^2} \tag{7}$$

where $\|\cdot\|_2$ is the Euclidean norm.

3. Channel Characteristic and Design of AGNet

In this section, we explain the design of the proposed AGNet, which exploits the inherent nature of the CSI and introduces the GRU network and attention mechanism.

3.1 UM-MIMO CSI Distribution Characteristic

For complex-valued CSI, it is natural to speculate that there exists certain internal similarity between the real and imaginary parts. From the perspective of data distribution, we adopt DSI to measure this similarity. DSI is a robust separability measure that can indicate whether data belonging to different classes have the same distribution. DSI value ranges from 0 to 1, and the lower value means the greater similarity between two datasets in terms of their distributions.

In order to compute DSI, we generate 12000 CSI samples. The details of simulation parameters are provided in Sect. 4. We separate the real and imaginary parts of CSI as sets R and I, respectively. The intra-class distance (ICD) set of R is computed as $\{d_x\} = \{\|x_i - x_j\|_2 | x_i, x_j \in R; x_i \neq x_j\}$. The ICD set of I, $\{d_y\}$, is similarly computed. The between-class distance (BCD) set between R and I is computed as $\{d_{x,y}\} = \{\|x_i - y_j\|_2 | x_i \in R; y_j \in I\}$. The similarities between the ICD and BCD sets are then computed using the Kolmogorov–Smirnov (KS) test: $s_x = KS(\{d_x\}, \{d_{x,y}\})$ and $s_y = KS(\{d_y\}, \{d_{x,y}\})$. Details of the KS calculation are in [11]. DSI is the average of the two KS statistics: $DSI(\{R, I\}) = (s_x + s_y)/2$.

Table 1 presents the measure results. $DSI(\{R, I\})$ is close to 0, indicating that R and I have extremely similar distributions.

3.2 The Design of AGNet
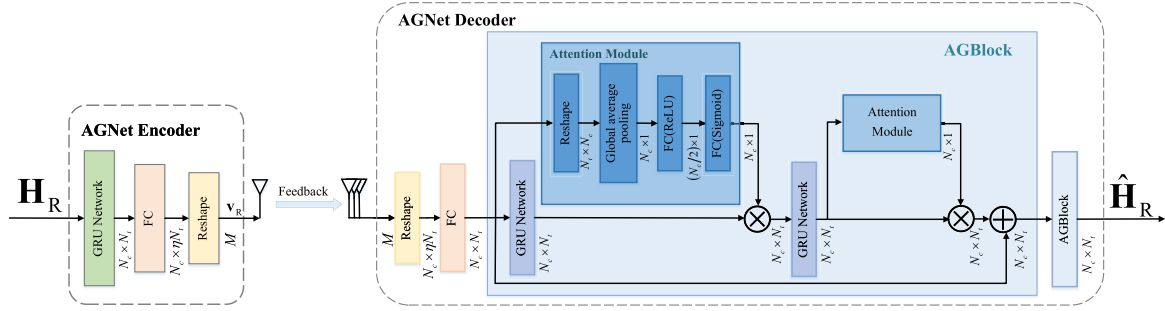
The distribution similarity between the real and imaginary

Fig. 3 Encoder and decoder design of the proposed AGNet.

parts of CSI mentioned above makes it feasible to utilize only the real part for training and apply the trained network to the imaginary part directly. Based on this finding, we design AGNet for CSI feedback.

As shown in Fig. 2, AGNet consists of an Encoder and a Decoder. The former is deployed at the UE and the latter at the BS. At the UE, both the real part $\mathbf{H}_R$ and imaginary part $\mathbf{H}_I$ of $\mathbf{H}$ share the same Encoder to generate the compressed codewords $\mathbf{v}_R$ and $\mathbf{v}_I$. At the BS, both the two codewords share the same Decoder for reconstruction. The whole feedback process can be concluded as

$$\hat{\mathbf{H}}_R = f_{de}(f_{en}(\mathbf{H}_R)) \text{ and } \hat{\mathbf{H}}_I = f_{de}(f_{en}(\mathbf{H}_I)) \qquad (8)$$

where $f_{en}(\cdot)$ and $f_{de}(\cdot)$ denote the Encoder and the Decoder of AGNet, respectively. The reconstructed complex-valued CSI matrix $\hat{\mathbf{H}}$ can be obtained by combining $\hat{\mathbf{H}}_R$ and $\hat{\mathbf{H}}_I$.

The detailed design of AGNet is shown in Fig. 3. In the Encoder, we employ GRU for feature extraction and further employ a fully connected (FC) layer for dimension compression. The real/imaginary part of the CSI is fed into GRU to generate a $N_c \times N_t$ feature map, then compressed by the FC layer and further reshaped into an M-dimensional codeword. Defining $N = N_c \times N_t$, the compression ratio is $\eta = M/N$. GRU has excellent memory function that can sensitively extract the inherent spatial correlation between adjacent antennas.

In the Decoder, the received codeword is reshaped into a $N_c \times \eta N_t$ matrix and then fed into the FC layer to recover the dimension. The initial recovered matrix is further refined by two AGBlocks for deep feature reconstruction. AGBlock is the key design of the Decoder, which contains two layers of GRU and embeds an attention module in both layers. The CSI feature maps of far- and near-field paths are significantly different. The goal of the attention module is to generate a vector that describes the importance of different feature maps. A global average pooling is first used to generate an $N_c \times 1$ vector, and then two FC layers are used to reconstruct the importance vector. The generated vector is scaled to the range $(0, 1)$ using the sigmoid activation function and then multiplied by the input feature maps. The information of far- and near-field paths is given corresponding weights by fusing the attention module, which enhances the network capacity of adapting the variable channel conditions. Additionally, an identity path is added to each AGBlock based on the idea

of residual learning, which can avoid the vanishing gradient problem and remain more effective information.

If stacking the real and imaginary parts of CSI as a real-valued input, as in the common scheme, the sizes of GRU and the FC layer are $24N^2 + 12N$ and $4NM + 2M$, respectively. By utilizing the distribution characteristic, the sizes of GRU and the FC layer are reduced to $6N^2 + 6N$ and $NM + M$, respectively. Moreover, compared with LSTM, GRU reduces one gate mechanism, namely, $2N^2 + 2N$ parameters, under the condition of achieving comparable performance.

4. Simulation Results

In our simulation, we consider a typical AoSA setting where the BS is equipped with $Q = 4$ SAs while each SA contains $\bar{Q} = 256$ AEs, and the total number of AEs at the BS is $N_t = 1024$ [3]. Considering to avoid the peak regions of molecular absorption loss, the center frequency is set as $f_c = 0.325$ THz [12]. The number of subcarriers is $N_c = 128$ and the bandwidth of the sub-band is $B = 1$ GHz. Since the multi-path effects such as scattering and diffraction can be ignored in the THz band, the number of paths is set as $L = 5$, of which 1 is the LoS path and 4 is the reflection path. The propagation distance of the LoS path is set as $r_1 = 30$ m. The geometric channel parameters including the Azimuth AoD φ_l, Elevation AoD θ_l, and NLoS paths propagation distance $r_l (l > 1)$ are generated in a distribution of $\mathcal{U}(-\pi, \pi)$, $\mathcal{U}(-\pi/2, \pi/2)$, and $\mathcal{U}(10\,\text{m}, 25\,\text{m})$. The time delay τ_l can be determined based on the distance. Specifically, the Rayleigh distance is set as $D = 20$ m, and r_l is set randomly spanning both the far-field and near-field regions to simulate the hybrid-filed propagation.

A total of 12,000 CSI samples are generated, of which 60%, 20%, and 20% are selected as training, validation, and testing datasets, respectively. The AGNet is trained for 100 epochs using Adam optimizer with a constant learning rate of 1×10^{-3}. We use the normalized mean square error (NMSE) defined in Eq. (9) to evaluate the network accuracy.

$$\text{NMSE} = \text{E}\left\{ \left\| \hat{\mathbf{H}}_{R(I)} - \mathbf{H}_{R(I)} \right\|_2^2 / \left\| \mathbf{H}_{R(I)} \right\|_2^2 \right\} \qquad (9)$$

Figure 4 presents the comparison of the feedback accuracy by TVAL3 [7], CRNet-cosine [9], TransNet [10], and AGNet, where the NMSE of three other methods is tested
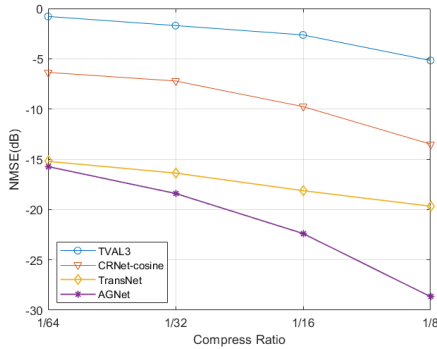
Fig. 4 NMSE comparison between AGNet and other methods at different compression ratios. We reproduce TVAL3, CRNet-cosine, and TransNet following the open source codes given in [7], [9], and [10], respectively.

Table 2 NMSE comparison under different test sets.

η	1/8	1/16	1/32	1/64
H_R	-28.6542	-22.3059	-18.3625	-15.7275
H_I	-28.6678	-22.4541	-18.4308	-15.7276

Table 3 Total number of trainable parameters.

η	1/8	1/16	1/32	1/64
AGNet-routine	76.62M	76.09M	75.83M	75.70M
ALNet	25.49M	25.36M	25.29M	25.26M
AGNet	19.19M	19.06M	18.99M	18.96M

under H, and the NMSE of AGNet is the average NMSE tested for both H_R and H_I. AGNet considerably outperforms CS-based TVAL3. Compared with CRNet-cosine, AGNet improves the feedback accuracy by over 58.15% under all compression ratios. Compared with TransNet, AGNet exhibits significant superiority at low compression ratios and improves accuracy by 87.99% and 62.51% at $\eta = 1/8$ and $\eta = 1/16$, respectively. This result is mainly attributed to the design of integrating GRU with the attention mechanism. The channel spatial correlation can be inherently retained by GRU. The distinct CSI features of the far- and near-field paths are assigned different attention weights through the attention mechanism, thereby the relevant information from different paths can be highlighted and extracted. As the compression ratio increases, the NMSE gap between AGNet and TransNet gradually narrows. In practical deployment, a simple AGNet is sufficient to meet the accuracy requirements for low compression ratios, while a more sophisticated network such as TransNet can be considered for high compression ratios.

Furthermore, the NMSE tested under H_R and H_I are summarized in Table 2. AGNet achieves similar reconstruction performance for both H_R and H_I, independent of the compression ratio. It justifies the validity of our finding of the CSI distribution characteristic.

Finally, the number of parameters of AGNet-routine, ALNet, and AGNet at different compression ratios are provided in Table 3. AGNet-routine refers to adopting the common method of concatenating the real and imaginary parts as input, and ALNet refers to replacing GRU in AGNet

with LSTM. Compared with AGNet-routine and ALNet, AGNet reduces the number of parameters by over 74.95% and 24.72% at all compression ratios, respectively. It validates that our one-part training strategy and the selection of GRU contribute to reducing memory consumption.

5. Conclusion

In this letter, a novel NN named AGNet is proposed for CSI feedback in THz UM-MIMO systems. The distribution similarity between the real and imaginary parts of CSI is verified using DSI, which is then utilized in AGNet and proven to be effective. The fusion of GRU and attention mechanism is designed to reconstruct the features of CSI. Simulation results show that the proposed AGNet significantly improves the reconstruction accuracy compared with other advanced DL approaches, especially at low compression ratios. Moreover, the distribution characteristic of CSI can be useful for other wireless communication designs in THz UM-MIMO systems, such as channel estimation and hybrid precoding.

References

[1] H. Sarieddeen, M.S. Alouini, and T.Y. Al-Naffouri, "An overview of signal processing techniques for terahertz communications," Proc. IEEE, vol.109, no.10, pp.1628–1665, 2021.

[2] M. Fujishima and S. Amakawa, "Integrated-circuit approaches to THz communications: Challenges, advances, and future prospects," IEICE Trans. Fundamentals, vol.E100-A, no.2, pp.516–523, Feb. 2017.

[3] I.F. Akyildiz and J.M. Jornet, "Realizing ultra-massive mimo (1024×1024) communication in the (0.06–10) terahertz band," Nano Communication Networks, vol.8, pp.46–54, 2016.

[4] C. Han, Y. Wang, and Y. Li, "Terahertz wireless channels: A holistic survey on measurement, modeling, and analysis," IEEE Commun. Surveys Tuts., vol.24, no.3, pp.1670–1707, 2022.

[5] X. Wang, X.-L. Hou, L. Chen, Y. Kishiyama, and T. Asai, "Deep learning-based massive MIMO CSI acquisition for 5G evolution and 6G," IEICE Trans, Commun., vol.E105-B, no.12, pp.1559–1568, Dec. 2022.

[6] J. Guo, C.K. Wen, S. Jin, and G.Y. Li, "Overview of deep learning-based CSI feedback in massive MIMO systems," IEEE Trans. Commun., vol.70, no.12, pp.8017–8045, 2022.

[7] C. Li, W. Yin, and H. Jiang, "An efficient augmented Lagrangian method with applications to total variation minimization," Comput. Optim. Appl., vol.56, no.3, pp.507–530, 2013.

[8] 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, 2018.

[9] Z. Lu, J. Wang, and J. Song, "Multi-resolution CSI feedback with deep learning in massive MIMO system," ICC 2020 - 2020 IEEE International Conference on Communications (ICC), pp.1–6, 2020.

[10] Y. Cui, A. Guo, and C. Song, "TransNet: Full attention network for CSI feedback in FDD massive MIMO system," IEEE Wireless Commun. Lett., vol.11, no.5, pp.903–907, 2022.

[11] S. Guan, M. Loew, and H. Ko, "Data separability for neural network classifiers and the development of a separability index," 2020, https://arxiv.org/abs/2005.13120

[12] S. Tarboush, H. Sarieddeen, H. Chen, M.H. Loukil, H. Jemaa, and M.-S. Alouini, "TeraMIMO: A channel simulator for wideband ultra-massive MIMO terahertz communications," IEEE Trans. Veh. Technol., vol.70, no.12, pp.12325–12341, 2021.