M.S. THESIS

Transformer-based Channel Parameter Acquisition for Terahertz Ultra-Massive MIMO Systems

테라헤르츠 초대규모 다중입출력 시스템을 위한 트랜스포머 기반 채널 파라미터 획득 기법

BY

ANHO LEE

FEBRUARY 2023

DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

M.S. THESIS

Transformer-based Channel Parameter Acquisition for Terahertz Ultra-Massive MIMO Systems

테라헤르츠 초대규모 다중입출력 시스템을 위한 트랜스포머 기반 채널 파라미터 획득 기법

BY

ANHO LEE

FEBRUARY 2023

DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Transformer-based Channel Parameter Acquisition for Terahertz Ultra-Massive MIMO Systems

테라헤르츠 초대규모 다중입출력 시스템을 위한 트랜스포머 기반 채널 파라미터 획득 기법

지도교수 심 병 효

이 논문을 공학석사 학위논문으로 제출함

2022 년 11 월

서울대학교 대학원

전기정보공학부

이 안 호

이안호의 공학석사 학위논문을 인준함

2022 년 12 월

위 원 장	최 완
부위원장	심 병 효
위 원	이 경 한

Abstract

Terahertz (THz) ultra-massive multiple-input multiple-output (UM-MIMO) is envisioned as a key technology to support ever-increasing data rates in 6G communication systems. To make the most of THz UM-MIMO systems, acquisition of accurate channel information is crucial. However, the THz channel acquisition is not easy due to the humongous pilot overhead that scales linearly with the number of antennas. In this paper, we propose a novel deep learning (DL)-based channel acquisition technique called *Transformer-based parametric THz channel acquisition* (T-PCA) for the THz UM-MIMO systems. By learning the complicated mapping function between the received pilot signal and the sparse channel parameters (e.g., angles, distances, path gains) using Transformer, a DL architecture that differently weights each input data based on the correlations between the input data, T-PCA can make a fast yet accurate channel estimation with a relatively small amount of pilot resources. Moreover, using the attention mechanism of Transformer, we can promote the correlation structure of the received pilot signals in the feature extraction, thereby improving the channel parameter estimation quality significantly. From the simulation results, we demonstrate that T-PCA is very effective in acquiring the THz channel information and reducing the pilot overhead.

Keywords: Wireless communication, Terahertz communication systems, Channel estimation, Deep neural network, Transformer

Student Number: 2021-25516

Contents

List of Figures

Chapter 1

Introduction

As a key technology to meet the demand for ever-increasing data rate in 6G, terahertz (THz) ultra-massive multiple-input multiple-output (UM-MIMO) communication has received a great deal of attention recently [1]. By exploiting the plentiful spectrum resources in the THz frequency band ($0.1 \sim 10\,\mathrm{THz}$) along with a large number of antennas, THz UM-MIMO communications can support way higher data rates than the conventional sub-6GHz and millimeter-wave wireless communication systems can offer. To maximize the potential gain of THz UM-MIMO systems, the base station (BS) needs to acquire accurate downlink THz channel information. Main challenge of the THz UM-MIMO systems is that the channel exhibits the near-field effect characteristics since the array aperture of the massive number of antenna elements is comparable to the communication distance [2]. Since the wavefront of the near-field THz signal is spherical, the THz channel can be expressed as a function of a few parameters in the spherical domain including angle of departures (AoDs), angle of arrivals (AoAs), distances, and path gains.

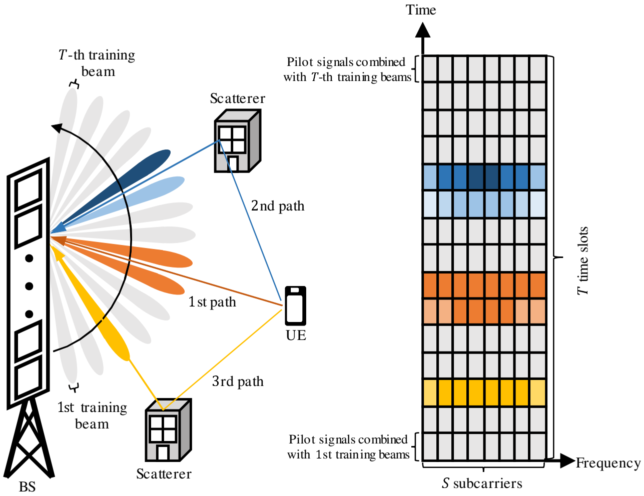


Figure 1.1: Correlation structure of the received pilot signal in THz UM-MIMO systems.

Recently, various techniques have been proposed for the acquisition of the THz channel parameters [3, 4, 5, 6, 7]. In [3, 4], compressed sensing (CS)-based channel acquisition approaches have been proposed. In [5, 6, 7], deep learning (DL)-based approaches that learn the mapping function between the received pilot signals and the channel parameters using deep neural network (DNN) have been proposed. Among various DNN architectures, a convolutional neural network (CNN) is popular due to its simplicity and ability to extract spatial features from the received pilot signals [7]. A major drawback of CNN, in the perspective of the THz channel parameter acquisition, is that it might not be effective in extracting the correlation between the spaced-apart pilot signals since the filter kernel and convolution operations are performed locally.

In the DL-based channel parameter estimator, a feature map is extracted from the DNN using the received pilot signals. By the feature map, we mean the low-dimensional vector containing core information (e.g., MIMO antenna array structure, locations of scatterers, and mobility of user equipment (UE)) of the large-dimensional input. To facilitate the feature extraction, one should deliberately handle the correlation structure of the received pilot signals. Notable characteristics of the received pilot signal of THz UM-MIMO systems are twofold; First, the received pilot signals will have meaningful power only for a few time slots. During the channel acquisition process, the BS employs the multiple sharp training beams, each of which is directed toward distinct directions [8]. Thus, the received pilot signal will have a high power only when the training beams are aligned with the direction of UE (see Fig 1.1). Second, the THz channel is determined primarily by the scattering geometry around the BS so the received pilot signals for each and every subcarrier can be expressed as functions of the same geometric parameters (e.g. angles, distances), which means that the received pilot signals, irrespective of their subcarrier positions, are highly correlated.

An aim of this paper is to propose a DL-based channel acquisition technique for the THz UM-MIMO systems. The proposed technique, dubbed as *Transformer-based parametric THz channel acquisition* (T-PCA), estimates the channel parameters (angles, distances, path gains) using *Transformer*, a DL architecture that differentially weights the significance of each input data (in our case, the received pilot signals) using the attention mechanism [9]. To make the most of the correlation structures of the received pilot signal, we employ two distinct Transformer networks, i.e., *temporal Transformer network* (TTN) and *spatial Transformer network* (STN). In TTN, using the received pilot signals as inputs, the temporally-correlated features are extracted from the product of the attention weight and the received pilot signal. Since only a small portion of received pilot signals have a meaningful power, Transformer in TTN is trained such that these dominant received pilot signals will have relatively high attention weights. Clearly, this process will facilitate the extraction of the temporally-correlated features. After that, using the low-dimensional features generated from TTN as inputs, the spatio-temporally correlated features are extracted in STN. As mentioned, the received pilot signals for all subcarriers are expressed as functions of the same channel parameters so that all received pilot signals, regardless of their positions, are correlated to each other. Main purpose of Transformer in STN is to capture the correlated features of both the adjacent and spaced-apart received pilot signals. Finally, the extracted features are converted to the channel parameters via the fully-connected network.

From the simulation results, we demonstrate that T-PCA outperforms the conventional channel acquisition schemes in terms of the normalized mean square error (NMSE). For example, T-PCA achieves more than $5\,dB$ NMSE gain over the CS-based scheme. Even when compared with the CNN-based scheme, T-PCA achieves around $2.5\,dB$ NMSE gain.

Chapter 2

Terahertz UM-MIMO System Model

We consider the THz UM-MIMO OFDM systems where a single-antenna UE transmits an uplink pilot signal to a BS equipped with a uniform linear array (ULA) of M antennas. Specifically, T time slots and S subcarriers are used for the uplink pilot transmission (see Fig 1.1). By exploiting the channel reciprocity of time-division duplexing (TDD) systems, the BS can recycle the acquired uplink channel information for the downlink data transmission.

In this setup, the received pilot signal vector $\mathbf{y}_{t,s} \in C^{N_r \times 1}$ of the s-th pilot subcarrier at t-th time slot is given by

$$\mathbf{y}_{t,s} = \mathbf{W}_t^{\mathrm{H}} \mathbf{h}_s x_{t,s} + \mathbf{W}_t^{\mathrm{H}} \mathbf{n}_{t,s} \tag{2.1}$$

$$= \sqrt{P_{tx}} \mathbf{W}_t^{\mathrm{H}} \mathbf{h}_s + \tilde{\mathbf{n}}_{t,s}, \tag{2.2}$$

where N_r is the number of RF chains in BS, $x_{t,s} = \sqrt{P_{tx}}$ is the uplink pilot, P_{tx} is the transmit power of UE, $\mathbf{h}_s \in C^{M \times 1}$ is the THz channel vector at s-th subcarrier, $\mathbf{W}_t \in C^{M \times N_r}$ is the receive beamforming matrix at the t-th time slot, and $\mathbf{n}_{t,s} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_M)$ is the additive Gaussian noise. By concatenating the received pilot signals over T time slots, we obtain the received pilot signal matrix at the s-th subcarrier $\mathbf{Y}_s = [\mathbf{y}_{1,s} \cdots \mathbf{y}_{T,s}]^{\mathrm{T}} \in C^{T \times N_r}$.

One notable characteristic of THz UM-MIMO systems is that the channel exhibits the near-field characteristics [2]. Since the signal wavefronts are spherical in the near-field channel, the phase difference between two antenna elements is affected by the distance r as well as the angle θ. Therefore, the near-field array steering vector is expressed as a function of spherical coordinates (θ, r). In fact, the near-field array steering vector $\mathbf{b}_s(\theta, r)$ at s-th subcarrier is given by [4]

$$\mathbf{b}_s(\theta, r) = [e^{-j\frac{2\pi}{\lambda}(1+\frac{f_s}{f_c})(r_1 - r)} \cdots e^{-j\frac{2\pi}{\lambda}(1+\frac{f_s}{f_c})(r_M - r)}]^{\mathrm{T}}, \tag{2.3}$$

where f_c is the carrier frequency, f_s is the baseband frequency of the s-th subcarrier, and r_m is the distance between the UE and the m-th BS antenna, given by

$$r_m = r - (m-1)d \sin\theta + (m-1)^2 \frac{d^2 \cos^2\theta}{2r}. \tag{2.4}$$

In this work, we use the near-field geometric THz channel model where the uplink channel vector $\mathbf{h}_s$ from the UE to the BS at the s-th subcarrier is expressed as

$$\mathbf{h}_s = \sum_{p=1}^{P} \alpha_p e^{-j2\pi f_s \tau_p} \mathbf{b}_s(\theta_p, r_p), \tag{2.5}$$

where P is the number of propagation paths, θ_p is the AoA, r_p is the distance, τ_p is the time delay, and α_p is the path gain of the p-th path. Let $\theta = [\theta_1 \cdots \theta_P]^{\mathrm{T}}$ and $\mathbf{r} = [r_1 \cdots r_P]^{\mathrm{T}}$ be the angle and distance vectors, respectively, and $\alpha_s = [\alpha_1 e^{-j2\pi f_s \tau_1} \ldots \alpha_P e^{-j2\pi f_s \tau_P}]^{\mathrm{T}}$ be the path gain vector for the s-th subcarrier, then $\mathbf{h}_s$ can be succinctly expressed as a function of channel parameters:

$$\mathbf{h}_s = \mathbf{B}_s(\theta, \mathbf{r}) \alpha_s, \tag{2.6}$$

where $\mathbf{B}_s(\theta, \mathbf{r}) = [\mathbf{b}_s(\theta_1, r_1) \cdots \mathbf{b}_s(\theta_P, r_P)] \in C^{M \times P}$ is the near-field array steering matrix. Note that $\mathbf{h}_s$ is parameterized by a few THz channel parameters, i.e., angles θ, distances $\mathbf{r}$, and path gains α_s, whose numbers are the same as the number of paths. Since the number of paths P (e.g., $P = 1 \sim 3$) is much smaller than the number of antennas M (e.g., $M = 256 \sim 1024$) in the THz UM-MIMO systems, one can significantly reduce the required number of measurements by estimating the sparse channel parameters instead of the full-dimensional channel vector $\mathbf{h}_s$.

Chapter 3

Transformer-based parametric Terahertz Channel Acquisition

Main goal of the proposed T-PCA is to estimate the sparse THz channel parameters (i.e., angles, distances, and path gains) using Transformer. Intriguing characteristic of T-PCA is that we extract the features of the THz UM-MIMO received pilot signals using the attention mechanism of Transformer. In essence, the attention mechanism facilitates the generation of the attention weights representing the correlations between input data. Using the product of the attention weights and the received pilot signals as input, one can extract the spatially and temporally-correlated features inherent in the THz UM-MIMO systems. Key ingredient of T-PCA is the combination of Transformer and fully-connected network to learn a complicated nonlinear mapping between the received pilot signals $\{\mathbf{Y}_s\}_{s=1}^{S}$ and the THz geometric channel parameters (θ, $\mathbf{r}$):

$$\{\hat{\theta}, \hat{\mathbf{r}}\} = g(\{\mathbf{Y}_s\}_{s=1}^{S}; \mathbf{\Gamma}), \tag{3.1}$$

where g is the mapping function and $\mathbf{\Gamma}$ are the network parameters. Once $\hat{\theta}$ and $\hat{\mathbf{r}}$ are acquired, the path gains $\{\hat{\alpha}_s\}_{s=1}^{S}$ can be easily estimated using the conventional approaches such as the least squares estimator [10]:

$$\hat{\alpha}_s = (\sqrt{P}\mathbf{W}^{\mathrm{H}}\mathbf{B}_s(\hat{\theta}, \hat{\mathbf{r}}))^{\dagger} vec(\mathbf{Y}_s^{\mathrm{T}}), \quad s = 1, \cdots, S, \tag{3.2}$$

where $\mathbf{W} = [\mathbf{W}_1 \cdots \mathbf{W}_T] \in C^{M \times TN_r}$. Using the obtained the channel parameters ($\hat{\theta}$, $\hat{\mathbf{r}}$, $\{\hat{\alpha}_s\}_{s=1}^{S}$), we can reconstruct the THz channels $\{\hat{\mathbf{h}}_s\}_{s=1}^{S}$:

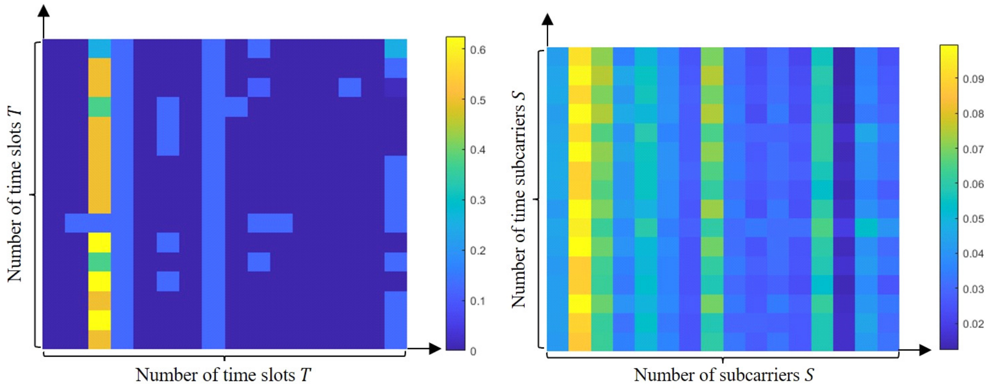
$$\hat{\mathbf{h}}_s = \mathbf{B}_s(\hat{\theta}, \hat{\mathbf{r}})\hat{\alpha}_s, \quad s = 1, \cdots, S. \tag{3.3}$$

3.1 Basic of Transformer

In the conventional CNN-based acquisition technique, the features are extracted by performing the convolution operation of a 2D/3D-shaped weight matrix (called kernel) and a part of the received pilot signal [6]. While CNN is effective in extracting the locally correlated features (e.g. correlation among antennas), it might not be efficient in extracting the globally correlated feature due to the locality of the filter kernel. Also, since the same kernel is multiplied to all input signals, the nonuniform and irregular correlation structures of the received pilot signals cannot be captured properly.

In a nutshell, Transformer extracts the features using the attention mechanism. In the attention layer of Transformer, the correlations between the input data (a.k.a., *attention weight* or *attention map*) are calculated and then multiplied to the input to generate the weighted input matrix [9]. Since the correlations between each and every elements in the input sequences (a.k.a., token) are used for the attention weight generation, Transformer can extract both the locally and globally correlated features effectively.

To be specific, using the sequence of $D \times 1$ input vectors $\mathbf{Y} = [\mathbf{y}_1 \cdots \mathbf{y}_L]^{\mathrm{T}} \in C^{L \times D}$, the attention layer constructs three different embedding matrices, i.e.,



(a) Attention map of TTN

(b) Attention map of STN

Figure 3.1: Attention maps of TTN and STN.

the query $\mathbf{Q} = \mathbf{Y}\mathbf{W}_Q$, the key $\mathbf{K} = \mathbf{Y}\mathbf{W}_K$, and the value $\mathbf{V} = \mathbf{Y}\mathbf{W}_V$ where $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V \in C^{D \times D}$ are the weight matrices and L is the number of input sequences. Since the query $\mathbf{Q}$ and the key $\mathbf{K}$ contains the features of the input data, by performing the inner product of $\mathbf{Q}$ and $\mathbf{K}$, we obtain the attention map $\mathbf{M} \in C^{L \times L}$:

$$\mathbf{M} = f_{softmax}(\mathbf{Q}\mathbf{K}^{\mathrm{T}}/\sqrt{D}), \tag{3.4}$$

where $f_{softmax}(\mathbf{Y})$ is a row-wise softmax function defined as $[f_{softmax}(\mathbf{Y})]_{i,j} = e^{\mathbf{Y}_{i,j}} / \sum_j e^{\mathbf{Y}_{i,j}}$. Finally, by multiplying the attention map $\mathbf{M}$ with the value $\mathbf{V}$, we obtain the weighted input matrix (a.k.a., *attention score*) $\mathbf{Z} \in C^{L \times D}$:

$$\mathbf{Z} = \mathbf{M}\mathbf{V}, \tag{3.5}$$

Once the attention score is obtained, the output vector passes through the fully-connected network, generating the output feature of Transformer.

To demonstrate the effect of the attention map in capturing the correlation structure of the received pilot signals, we plot the attention maps of TTN and STN in Fig 3.1. From Fig 3.1a, one can observe that the attention weights of TTN are concentrated on a small number of column vectors. Due to the extremely narrow beamwidth of THz UM-MIMO systems (i.e., pencil beam), the received pilot signals will contain the noise only when the training beams are not aligned with the signal propagation paths. This means that only a few row vectors of the received pilot signal matrix $\mathbf{Y}_s = [\mathbf{y}_{1,s} \cdots \mathbf{y}_{T,s}]^{\mathrm{T}}$ have relatively high values (same for the key $\mathbf{K}$ generated from $\mathbf{Y}_s$). Since the attention map is constructed from the inner product of $\mathbf{Q}$ and $\mathbf{K}$, the attention weights are concentrated on a few column vectors corresponding to the dominant received pilot signals.

3.2 Network Architecture of T-PCA

In T-PCA, the received pilot signal $\mathbf{y}_{t,s}$ is first separated into the real and imaginary parts $\bar{\mathbf{y}}_{t,s} = [Re(\mathbf{y}_{t,s})^{\mathrm{T}}\ Im(\mathbf{y}_{t,s})^{\mathrm{T}}]^{\mathrm{T}} \in R^{2N_r \times 1}$ and then concatenated matrices $\mathbf{Y}_s = [\bar{\mathbf{y}}_{1,s} \cdots \bar{\mathbf{y}}_{T,s}]^{\mathrm{T}} \in R^{T \times 2N_r}$ passes through the fully-connected network to generate $\mathbf{X}_s = \mathbf{Y}_s \mathbf{W}_e \in R^{T \times D}$ ($\mathbf{W}_e \in R^{2N_r \times D}$ is the weight matrix). Then a representative vector $\mathbf{x}_{0,s} \in R^{D \times 1}$, a trainable vector containing the correlated feature of the input data, is appended to the input matrices as $\bar{\mathbf{X}}_s = [\mathbf{x}_{0,s}\ \mathbf{X}_s^{\mathrm{T}}]^{\mathrm{T}} \in R^{(T+1) \times D}$ [11]. Also, to indicate the position of each element in the input data sequence, a trainable matrix called positional embedding matrix $\mathbf{W}_{pos} \in R^{(T+1) \times D}$ is added as $\tilde{\mathbf{X}}_s = \bar{\mathbf{X}}_s + \mathbf{W}_{pos}$. After that, the encoded input sequences $\{\tilde{\mathbf{X}}_s\}_{s=1}^{S}$ sequentially pass through the multiple Transformer blocks. In the last Transformer block, the temporal feature vectors $\{\mathbf{f}_s^{ttn}\}_{s=1}^{S}$ are obtained from the first row vector of the output matrix.

Once the temporal feature matrix $\mathbf{F}^{ttn} = [\mathbf{f}_1^{ttn} \cdots \mathbf{f}_S^{ttn}]^{\mathrm{T}} \in R^{S \times D}$ is obtained, $\mathbf{F}^{ttn}$ is used as an input matrix of STN. Similar to TTN, the representative vector and the positional embedding matrix are added to $\mathbf{F}^{ttn}$ and then the

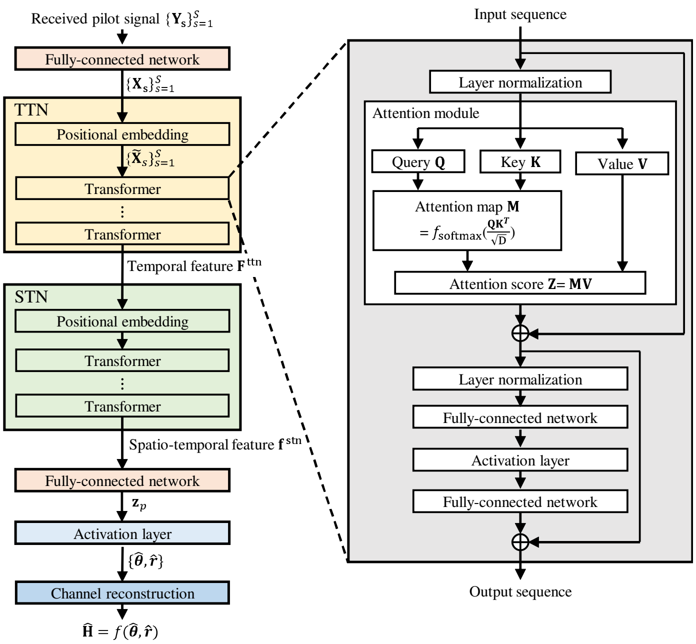


Figure 3.2: Overall structure of T-PCA

output matrix passes through multiple Transformer blocks. Then the spatio-temporal feature vector $\mathbf{f}^{stn} \in R^{D\times 1}$ is obtained from the first row vector of the output matrix of the last Transformer block.

The extracted spatio-temporal feature vector $\mathbf{f}^{stn}$ passes through the fully-connected network to generate the output vector $\mathbf{z}_p = \mathbf{W}_p\mathbf{f}^{stn} + \mathbf{b}_p \in R^{2P\times 1}$ ($\mathbf{W}_p \in R^{2P\times D}$ is the weight matrix and $\mathbf{b}_p \in R^{2P\times 1}$ is the bias vector). After that, $\mathbf{z}_p$ passes through the hyperbolic tangent layer $f_{tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ to generate the desired THz channel parameters $\{\hat{\theta}, \hat{\mathbf{r}}\}$:

$$\{\hat{\theta}, \hat{\mathbf{r}}\} = f_{tanh}(\mathbf{z}_p). \tag{3.6}$$

The overall structure of T-PCA is depicted in Fig 3.2.

Chapter 4

Simulation Result

4.1 Simulation Setup

In our simulation, we consider the THz UM-MISO OFDM systems where a BS equipped with $M = 256$ antennas and $N_r = 4$ RF chains serves a single-antenna UE. The UE is located randomly around the BS within the cell radius of $R = 50\,m$. We use the wideband THz multi-path channel model where the number of paths is $P = 1$, the carrier frequency is $f_c = 0.1\,THz$, and the channel bandwidth is $B = 1\,GHz$[1]. We set the numbers of subcarriers and time slots for pilot transmission to $S = 16$ and $T = 32$, respectively. The angles are generated randomly from $[-\pi, \pi)$ and the distances are generated based on the relative positions of BS and UE. We assume that the path gain is a complex Gaussian random variable $\alpha \sim \mathcal{CN}(0, \rho)$ where ρ is the large-scale fading coefficient accounting for the path loss and the shadow fading. Also, we

[1] In the THz systems, due to the severe path loss and directivity of THz band, the power of line-of-sight (LoS) component is almost 100 times stronger than that of the non-line-of-sight (NLoS) component [12].

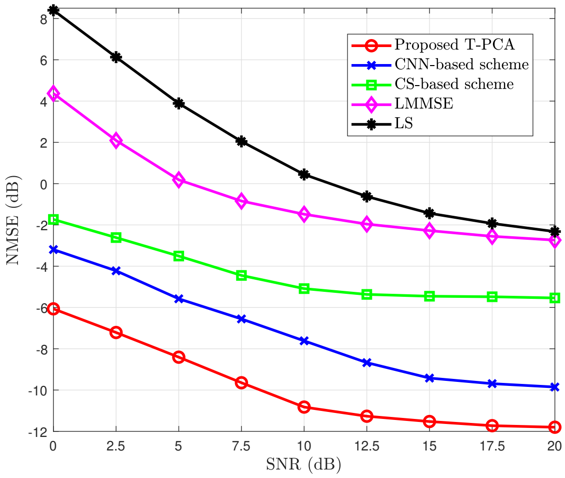


Figure 4.1: NMSE vs. SNR ($M = 256$, $N_r = 4$, $T = 32$, $S = 16$)

use the path loss model in 3GPP Rel. 16 [13].

In the proposed T-PCA, each Transformer network consists of two Transformer blocks with the embedding dimension $D = 128$. For the network parameter training, we use the unsupervised learning strategy where the network parameters Γ are updated iteratively in a way to minimize the NMSE-based loss function $J(\Gamma)$ [14]:

$$J(\Gamma) = \frac{1}{S}\sum_{s=1}^{S}\frac{||\mathbf{h}_s - \hat{\mathbf{h}}_s||^2}{||\mathbf{h}_s||^2}. \tag{4.1}$$

As a performance metric, we use the normalized mean square error (NMSE) defined as $NMSE = \frac{1}{S}\sum_{s=1}^{S}\frac{||\hat{\mathbf{h}}_s - \mathbf{h}_s||^2}{||\mathbf{h}_s||^2}$. The number of training epochs and the learning rate are set to $N_{training} = 1000$ and $\eta = 10^{-3}$, respectively. Also, for comparison, we use four benchmark channel acquisition schemes: 1) CNN-based scheme [6], 2) compressed sensing (CS)-based scheme [3], 3) linear minimum mean square error (LMMSE) estimator, and 4) least squares (LS) estimator.

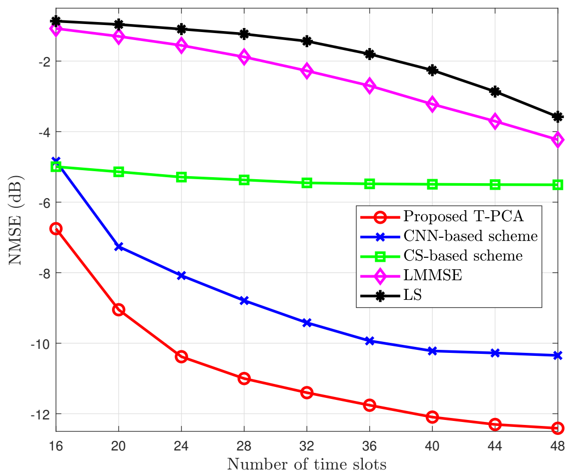


Figure 4.2: NMSE versus number of time slots ($M = 256$, $N_r = 4$, $S = 16$, $SNR = 15\,dB$)

4.2 Simulation Result

In Fig 4.1, we plot the NMSE as a function of transmit SNR. We observe that T-PCA outperforms the conventional channel estimation techniques by a large margin. For example, when $SNR = 10\,dB$, T-PCA achieves significant (more than $9\,dB$ and $11\,dB$) NMSE gains over the LMMSE and LS schemes, respectively. Even when compared with the CS-based scheme, T-PCA achieves around $6\,dB$ NMSE gain. This is because the mismatch between the true channel parameters and the quantized channel parameters is considerable in the CS-based scheme while such is not the case for T-PCA since T-PCA estimates the channel parameters in the continuous domain.

In Fig 4.2, we plot the NMSE as a function of the number of time slots. We observe that T-PCA achieves more than 33% pilot overhead reduction over the conventional schemes. For instance, to achieve the NMSE of $-10\,dB$, T-PCA requires 24 time slots while the conventional schemes require more than 36 time

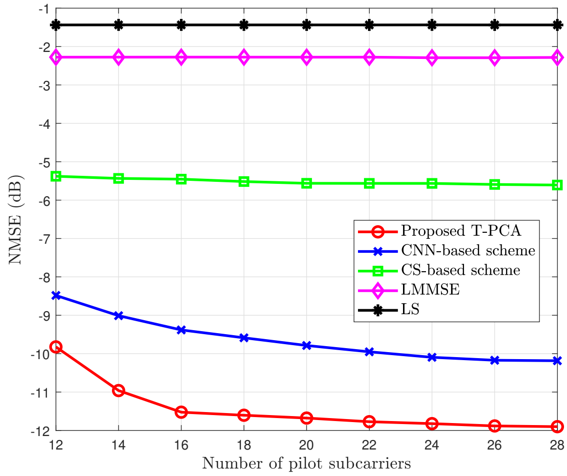


Figure 4.3: NMSE versus number of pilot subcarriers ($M = 256$, $N_r = 4$, $T = 32$, $SNR = 15\,dB$)

slots. This is not a surprise since the LMMSE and LS schemes estimate the full-dimensional THz channel vector $\mathbf{h}_s$ directly so that the required number of time slots is very large[2]. Whereas, by learning the complicated mapping between the received pilot signals and the THz channel parameters using Transformer, T-PCA can efficiently acquire the sparse THz channel parameters with a small amount of pilot resources.

In Fig 4.3, we plot the NMSE as a function of the number of pilot subcarriers. Since the proposed T-PCA promotes the correlation structure of received pilot signals using the attention mechanism of Transformer, T-PCA achieves a significant NMSE gain over the conventional schemes. For instance, when $S = 20$, T-PCA achieves more than $1.8\,dB$ and $6\,dB$ NMSE gains over the CNN and CS-based schemes, respectively. Interestingly, the NMSE gain of T-PCA over

[2]In fact, to guarantee the accurate estimation of $\mathbf{h}_s$, the number of measurements TN_r should be larger than the number of antenna elements M. For example, when $M = 256$ and $N_r = 4$, we need to allocate more than 5 subframe (more than 50% of a frame in 5G NR) just for the pilot transmission ($14\,slots/subframe \times 5\,subframe = 70 > M/N_r = 64$).

the conventional schemes increases with the number of pilot subcarriers. For example, when the number of pilot subcarriers increases from $S = 12$ to $S = 28$, the NMSE gain of T-PCA over the CNN-based scheme increases from $1.3\,dB$ to $2\,dB$.

Chapter 5

Conclusion

In recent years, a remarkable success of DL in various disciplines (e.g., image classification, speech recognition, and language translation) has stimulated increasing interest in applying this paradigm to wireless communication systems. In this paper, we proposed a DL-based channel acquisition technique for the THz UM-MIMO systems. Intriguing feature of the proposed T-PCA is that to promote the nonuniform and irregular correlation structures of the received pilot signals, we exploit Transformer, a DL architecture that differently weights each input data based on the correlations between the input data. Using the attention mechanism of Transformer, T-PCA can facilitate the extraction of spatially and temporally-correlated features inherent in the THz UM-MIMO systems. In doing so, fast yet accurate channel parameter estimation can be made with small pilot overhead. From the simulation results, we demonstrated that T-PCA achieves more than $2.5\,dB$ NMSE gain and 33% pilot overhead reduction over the conventional channel acquisition techniques. In our work, we restricted our attention to channel estimation, but there are many interesting applications of T-PCA such as channel feedback, beam tracking, and resource allocation.

Bibliography

[1] C. Han, Y. Wang, Y. Li, Y. Chen, N. A. Abbasi, T. Kürner, and A. F. Molisch, "Terahertz wireless channels: A holistic survey on measurement, modeling, and analysis," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1670–1707, 2022.

[2] X. Wei and L. Dai, "Channel estimation for extremely large-scale massive MIMO: far-field, near-field, or hybrid-field?," *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 177–181, 2022.

[3] K. Dovelos, M. Matthaiou, H. Q. Ngo, and B. Bellalta, "Channel estimation and hybrid combining for wideband terahertz massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1604–1620, 2021.

[4] J. Wu, S. Kim, and B. Shim, "Near-Field Channel Estimation for RIS-Assisted Wideband Terahertz Systems," *to appear in IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2022.

[5] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DoA estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, 2018.

[6] J. Kim, Y. Ahn, S. Kim, and B. Shim, "Parametric sparse channel estimation using long short-term memory for mmwave massive mimo systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 1397–1402, 2022.

[7] S. Bhattacharya and A. K. Gupta, "Deep learning for THz channel estimation and beamforming prediction via sub-6GHz channel," in *Proc. IEEE Int. Conf. Sig. Process Commun. (SPCOM)*, pp. 1–5, 2022.

[8] G. C. Alexandropoulos, I. Vinieratou, M. Rebato, L. Rose, and M. Zorzi, "Uplink beam management for millimeter wave cellular MIMO systems with hybrid beamforming," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, pp. 1–7, 2021.

[9] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Adv Neural Inf Process Syst*, vol. 30, 2017.

[10] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications: Useful tips and tricks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1527–1550, 2017.

[11] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[12] H. Do, S. Cho, J. Park, H.-J. Song, N. Lee, and A. Lozano, "Terahertz line-of-sight MIMO communication: Theory and practical challenges," *IEEE Commun. Mag.*, vol. 59, no. 3, pp. 104–109, 2021.

[13] "Study on channel model for frequencies from 0.5 to 100 GHz," *3GPP TR, 38.901, V16.1.0*, 2020.

[14] S. Kim, J. Son, and B. Shim, "Energy-efficient ultra-dense network using LSTM-based deep neural networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4702–4715, 2021.

Abstract (In Korean)

테라헤르츠 (terahertz; THz) 초대규모 다중 입력 다중 출력 (ultra massive-multiple input multiple output; UM-MIMO)은 6G 통신 시스템에서 증가하는 데이터 전송 속도를 지원할 수 있는 핵심 기술로 각광받고 있다. THz UM-MIMO 시스템을 최대한 활용하려면 정확한 채널 정보 획득이 중요하다. 그러나 안테나 수에 따라 선형으로 증가하는 파일럿 오버헤드로 인해 정확한 THz 채널 획득하는 것에 어려움이 있다. 본 논문에서는 THz UM-MIMO 시스템을 위한 트랜스포머 기반 THz 채널 파라미터 획득 기법 (Transformer-based parametric THz channel acquisition; T-PCA)이라는 새로운 딥 러닝 (deep learning; DL) 기반 채널 획득 기술을 제안한다. T-PCA는 입력 데이터 간의 상관 관계를 기반으로 각 입력 데이터의 가중치를 다르게 부여하는 DL 아키텍처인 트랜스포머를 사용하여 수신된 파일럿 신호와 채널 파라미터 (예: 각도, 거리, 경로 이득) 간의 복잡한 매핑 함수를 학습함으로써 상대적으로 적은 파일럿 자원으로도 빠르면서 정확한 채널 추정을 할 수 있다. 또한 트랜스포머의 주의 메커니즘 (attention mechanism)을 활용함으로써 특징 추출 (feature extraction)에 있어 수신된 파일럿 신호의 상관 구조를 충분히 반영할 수 있다. 실험을 통하여 우리는 제안하는 T-PCA가 THz 채널 정보를 획득하고 파일럿 오버헤드를 줄이는 데 매우 효과적임을 보인다.

주요어: 무선통신, 테라헤르츠 통신 시스템, 채널 추정, 심층 신경망, 트랜스포머
학번: 2021-25516