WILEY | Hindawi

Research Article

Millimeter Wave Wireless Channel Knowledge Map Construction Based on Path Matching and Environment Partitioning

Zeyang Li [iD],[1,2] **Qidong Gao** [iD],[1,2] **Wence Zhang** [iD],[1,2] **Xu Bao** [iD],[1,2] **Jing Xia** [iD],[1,2] **and Zhaowen Zheng** [iD][1,2]

[1]*School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, China*
[2]*Jiangsu Key Laboratory of Security Technology for Industrial Cyberspace, Zhenjiang, China*

Correspondence should be addressed to Wence Zhang; wencezhang@ujs.edu.cn

Key technologies in 5G and future 6G, such as millimeter wave massive multiple-input multiple-output (MIMO), relies accurate channel state information (CSI). However, when the number of base station (BS) antenna increases or the number of users is large, it is rather resource-consuming to obtain the CSI. Channel knowledge map (CKM) is an emerging environment-aware wireless communication technology, which stores the physical coordinates of BS and reference locations together with the corresponding channel path information. This makes it possible to obtain CSI with light or even without pilots, which can significantly reduce the overhead of channel estimation and improve system performance, especially suitable for quasi-static wireless environments with relatively stable channels and communication systems using millimeter waves, terahertz waves, visible light, and so on. The main challenge for CKM is how to construct an accurate CKM based on finite measurement data points at limited reference locations. In this work, we proposed a novel CKM construction method based on path matching and environmental partitioning (PMEP-CC) to address the above issues. Specifically, we first sort the propagation paths between reference locations, map them to a high-dimensional space to establish the path correlation coefficient between two reference locations. Then, the communication region are divided into different subregions based on its spatial correlation. Finally, the path information at locations where no measurements are available are estimated based on the known path information within the subregion to construct CKM. Numerical results are provided to show the performance of the proposed method over related studies.

1. Introduction

With the rapid development of wireless communication technology, new technologies such as massive multiple-input multiple-output (MIMO), millimeter wave and terahertz communication are emerging, which greatly improve the system spectrum efficiency and energy efficiency, and become potential solutions for the future 6G systems [1]. The excellent performance of the technologies such as millimeter wave massive MIMO relies on an accurate channel state information (CSI). Therefore, to obtain an accurate CSI is extremely important, which is mainly acquired through channel estimation. Massive MIMO systems which employs a large number of antennas and serves many users simultaneously results in high-dimensional channel matrix and thus the training overhead increases significantly. Therefore, efficient and accurate acquisition of CSI in the future wireless communication systems is very challenging. Channel knowledge map (CKM) is an emerging technology based on the environmental awareness [2], which retrieves accurate CSI (e.g., AoA/AOD, delay, gain, and number of paths) between transceivers based on their precise locations stored in a database. CKM-enabled environment-aware communication techniques can help to further improve the system performance by using schemes with few or even no pilots [2].

As evidenced by many applications in the recent years, mobile communication systems estimate the channel CSI mainly by transmitting pilots [3]. Both time division duplexing (TDD) and frequency division duplexing (FDD)

communication systems use the pilot to estimate the channel. At the expense of spectrum resources, the channel estimation is aided by the insertion of known symbols (pilots) in a prescribed arrangement in the transmitter's signal, and the received data symbols are recovered and corrected by the change in state of the frequency response in the channel at the receiver's end. In TDD systems, the uplink pilot is usually sent by the user equipment and the channel is estimated at the base station (BS) side, using the reciprocal characteristics of the uplink and downlink to obtain the CSI of the downlink. The length of the orthogonal pilot sequence is proportional to the number of users, and the pilot overhead becomes significantly larger with the number of users [4, 5]. Existing channel estimation schemes include time-domain MMSE methods based on subspace tracking [6], block-type pilot LS and MMSE methods [7], and compressive sensing methods [8, 9]. In the downlink of an FDD system, BS sends pilot and user equipment estimates the CSI of the downlink and feeds it back to the BS. The length of the orthogonal pilot sequence is usually proportional to the number of antennas at the BS, and the increase in the number of antennas generates a huge system overhead [10–12]. Existing channel estimation schemes include FDD-based compressive sensing methods [13] and robust super-resolution methods [14]. Although the above methods can accurately estimate the channel, as the antenna array size or the number of users becomes large, the channel dimension increases, the computational complexity grows geometrically, and the pilot overhead increases significantly, which becomes a bottleneck limiting the performance of millimeter wave massive MIMO [15] systems. The emerging CKM technology is expected to significantly reduce the channel estimation overhead and further improve the performance of technologies such as millimeter wave massive MIMO by leveraging information in the wireless environment that has been overlooked in the traditional communication system designs.

Some specific instances of CKM include channel gain map (CGM) [16], also known as radio environment map (REM), and channel path Map (CPM) [17], among others. CKM-enabled communications have recently been studied in the prediction of receive signal strength (RSS), the construction of fine radio maps, and thus improving the performance of communication systems [18–20].

Existing work mainly foucse on predicting RSS. In [21], the access point (AP) was used to divide the sensors into appropriately sized estimation groups based on the semivariogram of the measurements, and Ordinary Kriging interpolation then performed within the groups to estimate the signal strength. In the study of Sato et al. [22, 23], regression analysis was used over space and frequency domains to establish the path loss model at different frequencies. After achieving shadowing value interpolation at the target frequency, shadowing factors in the target band are spatially interpolated via Ordinary Kriging. These research methods in the above literature all rely on a path loss model, and models are built based on the probability distributions of channel parameters (e.g., small-scale fading, shadow fading, and AoA/AoD) and the scattering environment, which only consider the relative positions of transmitters and receivers, without considering current physical knowledge such as the material of obstacles between nodes and the distribution in the actual communication scenario. In addition, when fewer parameters are used to represent the channel, models are difficult to cover all features of the channel, and the estimation accuracy cannot be guaranteed. However, when there are more parameters, the system overhead is larger [24]. Therefore, the model-based channel is suitable for simple propagation environments, and its performance decreases sharply when the propagation environment is complex.

To overcome the above drawbacks, some scholars propose machine learning-based approaches to explore the information hidden between massive data using its powerful data mining capability [25] to complete the prediction of the channels. Li et al. [26] use a spatial–temporal reconstruction network to extract spatial and temporal low-resolution features, and then extracted high-level features were input to the network to complete high-precision map reconstruction. In [27], the power spectral density map estimation task is converted into an image reconstruction task and designs a map reconstruction framework based on a self-encoder and a full-convolutional network, which learns radio propagation characteristics from training data to reconstruct power spectral density map. In [28], the authors proposed a hybrid model and data-driven spectrum cartography (SC) framework that combines a radio map disaggregation model with DNN-based spatial power propagation, using DNN to represent the most complex part in the radio map model to alleviate the training and generalization difficulties. These methods do not use specifics to estimate RSS and have better prediction results in complex scattering environments. In [29], the authors proposed a radio-map learning and reconstruction approach that both revealed topology-induced structure to establish joint clustering and regression problems using maximum likelihood. The space is divided into several regions according to the distribution of obstacles, and signal strength models with the different values of parameters are built in the different regions. The above work only studies the prediction of RSS, and cannot reconstruct the channel matrix.

In CPM research, researchers aim to predict the channel path information for all remaining locations in the coverage area, such as the number of significant paths and their phases, AoA/AoD, and so on [17]. Combining the UE location and the essential parameters for the MIMO channel information stored by CPM, the MIMO channel matrix can be reconstructed for beam alignment without relying on conventional channel training. Similarly, beam index map (BIM) aims to learn the best beam pairs for all remaining locations within the transmitter coverage [30]. Compared to the CPM, although BIM does not take into account the reconstructed channel and therefore the computational overhead is light, the application scenario is more limited. Research work in this area is still in its infancy.

Due to the fact that the above-mentioned related work all use path loss models, current works do not consider the prior knowledge of the environment, which leads to mismatch problem between the model and the environment. Therefore,

the RSS estimation results are not always reliable. Methods based on the machine learning show better estimation results without specific pathloss models. However, only the channel gain is estimated in these schemes, which is sufficient for MIMO precoding or beamforming, etc. CKM contains environmentally relevant path information and is expected to enable highly accurate channel reconstruction to solve the above problems. Most existing research shows that channels exhibit sparse characteristics and often consist of a limited number of multipaths in high-frequency communication scenarios (e.g., millimeter wave, teraheriz). In order to reconstruct CKM, key path parameters such as the number of significant paths and their phases, need to be preserved. There are two major challenges in the construction of CKM. On the one hand, there is no specific model for CKM, so they cannot be constructed directly. On the other hand, the radio propagation is affected by environment, resulting in the different degrees of variation between multiple path parameters such as their phase, time delay, and gain at different locations, so it is impossible to arbitrarily select path to construct a CKM.

In order to solve the above issues, we proposed a novel CKM construction method based on path matching and environmental partitioning that does not rely on a specific pathloss model. The location of the BS is assumed to be fixed, and the majority of obstacles within its service area also has a fixed distribution. On the one hand, for a given reference location, although the variation in channel characteristics is perturbed by small-scale fading, the channel can be considered quasi-static (QS) because the obstacle distribution usually does not obviously change within a short period of time and the path information changes slowly [31]. On the other hand, the QS-CSI varies greatly at the different locations. The communication area is divided into subregions based on spatial correlation of the wireless channel. Unlike existing works which only focus on prediction of a channel single parameter, the proposed scheme takes into account multiple parameters such as AoA/AoD and delay, allowing the reconstruction of the channel matrix. The main contributions to this work are summarized as follows:

(i) The reference positions where CSI is assumed to be known are acquired through simulation or online measurements, etc., on the basis of which a model for CKM construction was developed.

(ii) A path-matching method based on high-dimensional mapping is proposed for the variability of parameters such as the number of paths, the angle, and path gain at different locations. After mapping QS-CSI to high-dimensional space with a kernel function, it then matches strongly correlated paths with the differences in path information, and calculates the spatial correlation of channels at different locations.

(iii) The coverage area is divided into subregions according to the channel correlation of different locations, and the channels within each subregion have strong correlation characteristics, while the channels between different subregions differ significantly. The QS-CSI of
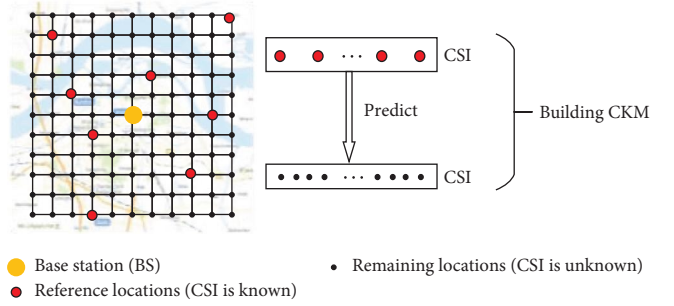

FIGURE 1: An illustration of the CKM construction process.

the location of interest is predicted using the inverse distance weighted (IDW) interpolation method in each subregion.

The remainder of this paper is organized as follows. The model of the CKM and the problem challenges are introduced in Section 2. In Section 3, we proposed a CKM construction method based on path matching and environment partitioning. In Section 4, the numerical evaluation results and analysis are reported. The conclusions are drawn in Section 5.

2. System Model and Problem Formulation

The CSI from the BS to the reference location often relies on sending pilot signals performing the channel estimation. This inevitably introduces signaling overhead, which can have a significant impact on system performance, especially when the number of antennas or users is large [32]. The CSI acquired in traditional communication is discarded after it is used and thus it is not fully utilized. And in the actual communication scenarios, the distribution of major obstacles does not change for a comparatively long time (e.g., the positions of obstacles like buildings and trees are fixed). The radio propagation environment is stable, and the CSI is almost unchanged in a short time. We proposes to mine historical CSI to estimate slow-varying CSI and construct a mapping relationship between receiver location, i.e., CKM, which can reduce or avoid signaling overhead. As shown in Figure 1, we use the CSI at the locations where is known to estimate the CSI at the remaining locations.

Denote the interested area as $\mathcal{S}$. Let $\mathbf{CI}\left(\mathbf{l}_{\mathrm{TX}}, \mathbf{l}_{\mathrm{RX}}\right)$ denote the CSI between the BS at $\mathbf{l}_{\mathrm{TX}} = [x_{\mathrm{TX}}, y_{\mathrm{TX}}, z_{\mathrm{TX}}]^{\mathrm{T}}$ and the reference location at $\mathbf{l}_{\mathrm{RX}} = [x_{\mathrm{RX}}, y_{\mathrm{RX}}, z_{\mathrm{RX}}]^{\mathrm{T}}$ where $\mathbf{l}_{\mathrm{TX}}, \mathbf{l}_{\mathrm{RX}} \in \mathcal{S}$. $\mathbf{CI}$ contains the number of valid paths and the slowly varying channel information. Existing studies have shown that when the wireless environment changes slowly, the statistical properties of small-scale fading of wireless channels in high-frequency bands do not change significantly [31], so the QS-CSI does not change significantly in a short period of time and shows strong spatial correlation. When the user's position changes, the high-speed movement causes the wireless channel highly time varying but it can still be considered a QS channel if the obstacles remains unchanged. For example, the position of a high-speed train changes rapidly, but its traveling route is fixed, and the position of the obstacles around the

route remains unchanged in general. So, it can be considered that the user is moving at high speed in a large QS region. CKM can be directly used for channel reconstruction for services that has low-CSI requirement, or be used with the assistance of a few pilots to provide high-quality CSI for high-data rate.

The expression of $\mathbf{CI} \in \mathbb{R}^{N_P \times 6}$ is given by

$$\mathbf{CI}(\mathbf{l}_{\mathrm{TX}}, \mathbf{l}_{\mathrm{RX}}) = \begin{bmatrix} \phi_1^{\mathrm{D}} & \theta_1^{\mathrm{D}} & \phi_1^{\mathrm{A}} & \theta_1^{\mathrm{A}} & g_1 & \tau_1 \\ \phi_2^{\mathrm{D}} & \theta_2^{\mathrm{D}} & \phi_2^{\mathrm{A}} & \theta_2^{\mathrm{A}} & g_2 & \tau_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_{N_P}^{\mathrm{D}} & \theta_{N_P}^{\mathrm{D}} & \phi_{N_P}^{\mathrm{A}} & \theta_{N_P}^{\mathrm{A}} & \theta_{N_P}^{\mathrm{A}} & \tau_{N_P} \end{bmatrix},$$

(1)

where each row corresponds to a path, and N_{p} denotes the number of valid paths, each containing six parameters as explained below. $\phi_i^{\mathrm{D}}, \theta_i^{\mathrm{D}} \in [0, 2\pi)$ represent the elevation and horizontal AoD of the i ($i = 1, \cdots, N_{\mathrm{p}}$)-th path, respectively. $\phi_i^{\mathrm{A}}, \theta_i^{\mathrm{A}} \in [0, 2\pi)$ represent the elevation and horizontal AoA of the path, respectively. g_i and τ_i are the path gain and time delay of the i-th path, respectively.

To construct the CKM, the reference location CSI can be acquired through simulation and/or measurements [33]. The CSI data acquisition is divided into two main parts. One is using automatic ground vehicles to measure the CSI of some locations around the building. The other one is to obtain the CSI of some locations that cannot be arranged due to the geographic environment by simulation. The specific operation of simulation method is to obtain all geographic information of the communication scenarios (e.g., the location and the height of buildings and trees, etc.), and to construct a 3D map. We use 3D map and Ray Trace simulation to obtain the CSI at the locations that cannot be measured in the actual area.

Consider the set of data $D = \{\xi^{(i)} | i = 1, ..., M\}$, where $\xi^{(i)} \triangleq \{\mathbf{l}_{\mathrm{TX}}^{(i)}, \mathbf{l}_{\mathrm{RX}}^{(i)}, \mathbf{CI}^{(i)}\}, i = 1, ..., M$. The CKM can be seen as a mapping that associates the BS and location coordinates with the corresponding $\mathbf{CI}$, which can be expressed as follows:

$$\mathcal{M}\left(\mathbf{l}_{\mathrm{TX}}^{(i)}, \mathbf{l}_{\mathrm{RX}}^{(i)}\right) \rightarrow \begin{bmatrix} \phi_1^{i\mathrm{D}} & \theta_1^{i\mathrm{D}} & \phi_1^{i\mathrm{A}} & \theta_1^{i\mathrm{A}} & g_1^i & \tau_1^i \\ \phi_2^{i\mathrm{D}} & \theta_2^{i\mathrm{D}} & \phi_2^{i\mathrm{A}} & \theta_2^{i\mathrm{A}} & g_2^i & \tau_2^i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_{N_P}^{i\mathrm{D}} & \theta_{N_P}^{i\mathrm{D}} & \phi_{N_P}^{i\mathrm{A}} & \theta_{N_P}^{i\mathrm{A}} & g_{N_P}^i & \tau_{N_P}^i \end{bmatrix},$$

(2)

the mapping is in fact $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^{N_P \times 6}$. Assuming the CSI at $\mathcal{M}^{\mathrm{RF}}$ reference locations (denoted as D_{RF}) are known and are used to predict the CSI at the $\mathcal{M}_{\mathrm{RE}}$ remaining locations (denoted as D_{RE}) in $\mathcal{S}$. Assess the precision of QS-CSI estimation at $\mathcal{M}_{\mathrm{RE}}$ through the normalized mean squared error (NMSE), which measures the difference between estimated and actual values and is given by

$$\begin{aligned} \mathrm{NMSE} &\triangleq \frac{1}{M} \sum_{j=1}^{M} \frac{\left\| \widehat{\mathbf{CI}}^{(j)} - \mathbf{CI}^{(j)} \right\|_{\mathrm{F}}^2}{\left\| \mathbf{CI}^{(j)} \right\|_{\mathrm{F}}^2} \\ &= \frac{1}{M} \sum_{j=1}^{M} \frac{\left\| \mathcal{M}\left(\mathbf{l}_{\mathrm{TX}}^{(j)}, \mathbf{l}_{\mathrm{RX}}^{(j)}\right) - \mathbf{CI}^{(j)} \right\|_{\mathrm{F}}^2}{\left\| \mathbf{CI}^{(j)} \right\|_{\mathrm{F}}^2}, \left\{ \mathbf{l}_{\mathrm{TX}}^{(j)}, \mathbf{l}_{\mathrm{RX}}^{(j)}, \mathbf{CI}^{(j)} \right\} \in D_{\mathrm{RE}}, \end{aligned}$$

(3)

where $\widehat{\mathbf{CI}}^{(j)}$ represents the estimated CSI matrix at location j. $\|\cdot\|_{\mathrm{F}}^2$ denotes the number of Frobenius norm. M represents the number of elements in D_{RE}. The CKM construction problem can be expressed as an optimization problem. The estimated CKM is estimated most accurately when the NMSE takes the smallest value, translating the above problem into

$$\begin{aligned} \mathcal{M} &= \arg \min_{\mathcal{M}} \mathrm{NMSE} \\ &= \arg \min_{\mathcal{M}} \frac{1}{M} \sum_{j=1}^{M} \frac{\left\| \mathcal{M}\left(\mathbf{l}_{\mathrm{TX}}^{(j)}, \mathbf{l}_{\mathrm{RX}}^{(j)}\right) - \mathbf{CI}^{(j)} \right\|_{\mathrm{F}}^2}{\left\| \mathbf{CI}^{(j)} \right\|_{\mathrm{F}}^2}. \end{aligned}$$

(4)

The radio propagation is impacted by obstacle distribution and propagation distance, so parameters such as path loss and shadow fading may vary significantly between different subregions, making it difficult to model and analyze in a uniform manner. In addition, the $\mathcal{M}\left(\mathbf{l}_{\mathrm{TX}}^{(j)}, \mathbf{l}_{\mathrm{RX}}^{(j)}\right)$ has no analytical model, and conventional methods cannot provide the best solution to the optimization problem (4).

To address these issues, we proposes to leverage spatial correlations among CSI data at nearby locations, and then partition the physical space and complete the construction of CKM within each subregion. The difficulties lie in the following aspects:

(i) The presence of multiple paths in $\mathbf{CI}$ implies that the order of these paths impact the correlation of CSI at different locations. Take $\mathbf{a} = [-1\ 1]$, $\mathbf{b} = [1\ -1]$, and $\mathbf{c} = [-1\ 1]$ for example. $\mathbf{b}$ and $\mathbf{c}$ have the same elements, but in the different order, which results in a different correlation coefficients with $\rho_{ab} = 0, \rho_{ac} = \sqrt{2}/2$.

(ii) Since the receivers are situated in the different physical locations, the radio propagation environment changes significantly due to different propagation distance or obstacle distribution. Therefore, the QS-CSI at different region may be significantly different. Using all data in D_{RF} to predict the CSI at a specific remaining location may lead to large prediction error and the computational complexity is very high. This can be alleviated by selecting effective data at locations in D_{RF} that experience similar propagation environment. However, dividing $\mathcal{S}$ into subregions that shows similar propagation properties is no easy task.
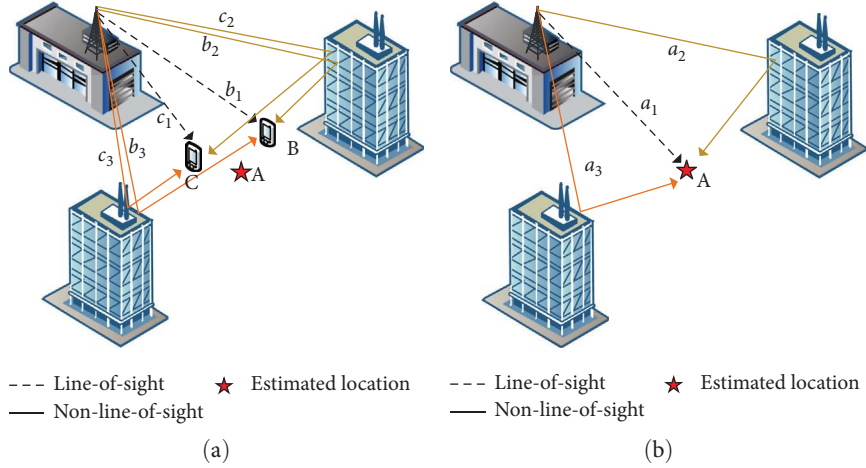
FIGURE 2: Path distribution matching map. (a) Paths information at known locations B and C. (b) Path matching map for location A.

To solve the above difficulties, we proposed a path matching and environment partitioning-based CKM construction (PMEP-CC) scheme, which will be detailed in the next section.

3. Path Matching and Environment Partitioning-Based CKM Construction Scheme

In this section, we introduce the proposed PMEP-CC scheme. First, a kernel function that maps path parameters into higher dimension is introduced to increase the differentiation between path parameters. Then an affinity propagation- (AP-) based partitioning algorithm is proposed to divide the coverage area into different subregions by utilizing spatial correlation of CSI at different locations. Finally, the CSI of the remaining locations in $\mathcal{S}$ is predicted using data from D_{RF} within the same subregion.

3.1. Path Matching Based on High-Dimensional Mapping. Unlike the previous studies on radio map (RAM) that solely consider channel gain, the proposed CKM involves multiple parameters for each path such as AoA/AoD, channel gain, and so on, i.e., **CI** for different locations. The construction of a complete CKM requires paths with similar parameters around the reference location to predict the multiple paths of the remaining location. This section proposes a path-matching solution by mapping CSI to a high-dimensional space. The matrix **CI** includes multidimensional data, making it challenging to determine the correlation between matrices accurately utilizing conventional methods. Consequently, the proposed approach measures the gap between the corresponding path parameters in the matrix **CI** at various locations when the path order changes in the multiple dimensions to decide the correlation and path matching between the matrices **CI** when the changes in gap measures are minimal. Given that the approximation of AoA/AoD in high-frequency band millimeter wave follows the normal distribution, we introduces the Gaussian Kernel (GK) as a high-dimensional mapping tool.

Consider the radio propagation path environment between transmitter A, B, and C in Figure 2, which includes both line-of-sight (LOS) and non-line-of-sight (NLOS) paths. Assuming path information is known at both locations B and C, which are used to estimate the propagation path information at location A. As depicted in Figure 2, each of the three points A, B, and C possess three propagation paths, labeled as (a_1, a_2, a_3), (b_1, b_2, b_3), and (c_1, c_2, c_3), respectively. The path information of LOS path a_1 and b_1, c_1 are similar and show high correlation, which is expected to estimate a_1 information accurately. However, if b_2 and c_2 are utilized, the estimated a_1 obtained is bound to have a large error. In order to construct the CKM, it is imperative to match the paths that have a strong correlation and estimate the path information of the remaining locations.

Due to **CI** contains the path parameters that have different orders of magnitude and units, such as a path gain of -100 dB and a delay in 10^{-7} s, which cannot be compared directly. Therefore, the unit needs to be removed and the data is normalized as follows:

$$\mathbf{y} = \frac{(\Omega_{max} - \Omega_{min}) \times [\mathbf{x} - \min(\mathbf{x})]}{\max(\mathbf{x}) - \min(\mathbf{x})} + \Omega_{min}, \quad (5)$$

where $\mathbf{y}$ and $\mathbf{x}$ represents the normalized value and the original value, respectively. $\Omega_{min}, \Omega_{max}$ represent the minimum and maximum values of the normalized range, respectively. In Equation (5), the normalization are AoA/AoD, time delay and path gain, and so on.

After normalization, the correlation coefficients for different positions **CI** are calculated in the subsequent manner. First, the order of the multiple paths at each location is determined to accomplish path matching. Fix the path order in the data samples at location B and vectorize $\mathbf{CI}_B$ as follows:

$$\widetilde{\mathbf{CI}}^B = \text{vec}\left(\mathbf{CI}_B^T\right) = \begin{bmatrix} \mathbf{w}_1^T & \mathbf{w}_2^T & \cdots & \mathbf{w}_{N_p}^T \end{bmatrix}^T, \quad (6)$$

where $\mathbf{w}_i = \begin{bmatrix} \phi_i^D & \theta_i^D & \phi_i^A & \theta_i^A & g_i & \tau_i \end{bmatrix}^T$ $(i = 1 \dots N_p)$ represents the CSI of the i-th path, $\widetilde{\mathbf{CI}}^B$ represents the order of

--- Line-of-sight   ★ Estimated locations
—— Non-line-of-sight
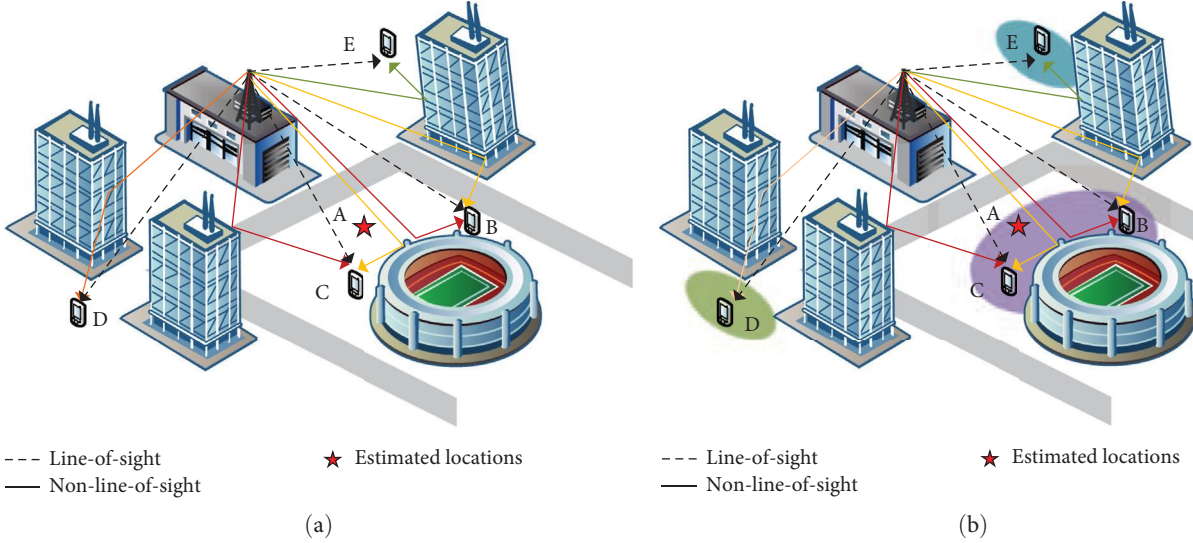
—— Non-line-of-sight

(a)

(b)

FIGURE 3: An illustration of environment partitioning based on propagation path correlations. (a) Reference locations distribution. (b) Spatial division.

paths at fixed position B. Then, the path order of the **CI** at location C can be modified and vectorized as follows:

$$\widetilde{\mathbf{CI}}_j^{\mathrm{C}} = \mathrm{vec}\left(\Pi_j \mathbf{CI}_{\mathrm{C}}^{\mathrm{T}}\right), \qquad (7)$$

where $\Pi_j \in \mathbb{R}^{N_{\mathrm{P}} \times N_{\mathrm{P}}}$ $(j = 1, \cdots, [N_{\mathrm{P}} \times N_{\mathrm{P}} - 1]/2)$ represents the permutation matrix. $\widetilde{\mathbf{CI}}_j^{\mathrm{C}}$ represents the order of different paths at position C.

$$\rho(\mathbf{CI}_{\mathrm{B}}, \mathbf{CI}_{\mathrm{C}}) \triangleq e^{-\frac{\left\|\widetilde{\mathbf{CI}}^{\mathrm{B}} - \widetilde{\mathbf{CI}}_j^{\mathrm{C}}\right\|^2}{2\sigma^2}}$$
$$\propto \sum_{n=0}^{+\infty} \left[\frac{1}{\sigma^2}\left(\widetilde{\mathbf{CI}}^{\mathrm{B}}\right)^{\mathrm{T}} \widetilde{\mathbf{CI}}^{\mathrm{C}}\right]^n, \ j = 1, \cdots, [N_{\mathrm{P}} \times N_{\mathrm{P}} - 1]/2 \qquad (8)$$

where ρ $(0 < \rho < 1)$ takes the maximum value. when $\rho(\mathbf{CI}_{\mathrm{B}}, \mathbf{CI}_{\mathrm{C}})$ is smallest, it matches the path with the strongest correlation at B, C. The larger the value of σ indicates the influence of the interaction between samples, the smoother the change of ρ and the larger the range of data sample effects, affecting more data samples. When σ is smaller, the range of effects is smaller, ρ the change is steeper and the phenomenon of overfitting is easy to occur, and the size of its value seriously affects the correlation between data samples. The optimal value of σ can be found by either robust search or solving an optimization problem which is formulated based on the data distribution [35].

3.2. Environmental Partitioning Based on AP. Due to the difference between the propagation environments, the channel parameters differ significantly between the two distant locations. Specifically, NLOS paths are affected by reflection, refraction, and diffraction, leading to variations in the path gain, delay, and AoA in different propagation environments. Similarly, LOS paths are mainly affected by the propagation

Next, the correlation coefficient between the two locations **CI** was determined when calculating the change in path order at location C. To accurately utilize the change in information in each dimension in **CI**, a Gaussian kernel function [34] was used to map the path information to a higher dimensional space to determine the matrix correlation at locations B and C, which is expressed as follows:

distance, leading to differences in path gain and delay at different locations. Therefore, if the **CI** of the remaining locations is inferred directly using CSI data without considering environment, it will lead to large errors. To address this problem, we proposes an environmental partitioning method based on AP, using the correlation coefficients of **CI** between different locations to divide the whole area into multiple subregions. Then, using the sample data to estimate the **CI** of locations that are interested within the same subregion, which can reduce the computational complexity and improve the accuracy.

After reviewing the distribution of obstacles and receivers shown in Figure 3(a), we observe that the propagation environment at locations B and C is more similar than that of location A compared to the environments at locations D and E. Additionally, there are also notable differences between the propagation environments at locations D and E. In Figure 3(b), we divide the locations A, B, C, D, E into distinct subregions based on the propagation environment. The path parameters such as AoA/AoD, time delay and path gain

of the signals within the same subregion are similar and highly correlated, and within different subregions, there are large variations in one or more path parameters, so the correlation is weak.

The correlation of the CSI matrix between receivers is obtained through high-dimensional mapping, with $\rho(\mathbf{CI}_\mathrm{B}, \mathbf{CI}_\mathrm{C})$ rewritten as $\rho(\mathbf{CI}_i, \mathbf{CI}_j)$ for universality. The environment partitioning should consider both the distance and the obstacle distribution. Therefore, we propose the following modified correlation criterion $\rho'(\mathbf{CI}_i, \mathbf{CI}_j)$, for locations i, j in $\mathcal{S}$, $\rho'(i, j) < 0$ as follows:

$$\rho'(i, j) \triangleq -\alpha_1 \rho(\mathbf{CI}_i, \mathbf{CI}_j) - \alpha_2 \frac{1}{d(i, j)}, i \neq j, \quad (9)$$

where

$$d(i, j) = \left\| \mathbf{l}_\mathrm{RX}^{(i)} - \mathbf{l}_\mathrm{RX}^{(j)} \right\|_2, \quad (10)$$

denotes the Euclidean distance between receivers after normalization. In Equation (9), $\alpha_1 + \alpha_2 = 1$ represents the weighs for channel correlation and Euclidean distance.

In this work, we use AP-based iteration algorithm for environment division to divide $\mathcal{S}$ into different subregions. First, $r(i, j)$ is initialized as follows:

$$r(i, j) = \rho'(i, j) - \max_{k \neq j} \{a(i, k) + r(i, k)\}. \quad (11)$$

Then, $a(i, j)$ is updated iteratively as follows:

$$a(i, j) = \begin{cases} \min \left\{ 0, r(j, j) + \sum_{k \neq i, j} \max(0, r(k, j)) \right\}, i \neq j \\ \sum_{k \neq j} \max(0, r(k, j)), i = j \end{cases}, \quad (12)$$

where $a(i, j)$ and $r(i, j)$ iterate over each other until the appropriate subregions are selected. $r(i, j)$ indicates the suitability of dividing the j-th reference location and the i-th reference location into the same subregion, $a(i, j)$ represents the degree to which the j-th reference location and i-th reference location are suitable for different subregions. Initialize $r(i, j) = 0$ and $a(i, j) = 0$.

3.3. CKM Construction Based on Inverse Distance Interpolation.
Within the same subregion, the radio propagation environment is similar and there is strong correlation between the **CI** at different locations. Meanwhile, the degree of correlation heavily depends on the distance between these locations. Consequently, we proposes a $\mathbf{CI} \in D_\mathrm{RF}$ to estimate $\mathbf{CI} \in D_\mathrm{RE}$ method based on inverse distance interpolation to facilitate the construction of the CKM.

Observing the specified area in Figure 4, the obstacle distribution is fixed and does not change over a long period of time, and the radio propagation environment changes
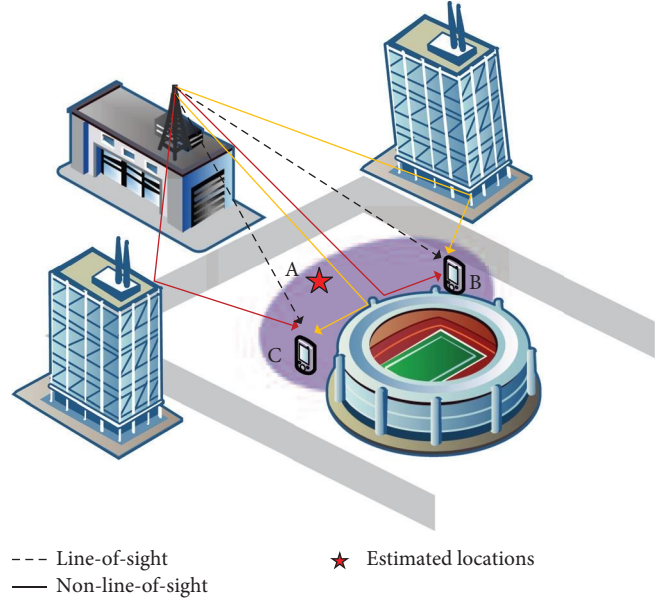

--- Line-of-sight   ★ Estimated locations
— Non-line-of-sight

FIGURE 4: Schematic diagram of QS-CSI estimation.

gradually. The QS-CSI matrices $\mathbf{CI}^\mathrm{B}, \mathbf{CI}^\mathrm{C}$ are known at B and C, and A is an arbitrary point within the subregion, using sample points B and C to estimate CSI at A. In addition, a positive correlation between path information similarity and position was observed when there were no obstacles between receivers. Therefore, using the distance from the point to be estimated to that of the known point as a weight to estimate the CSI at the point to be estimated, the **CI** at A is estimated as follows:

$$\mathbf{CI}^\mathrm{A} = \lambda_1 \mathbf{CI}^\mathrm{B} + \lambda_2 \mathbf{CI}^\mathrm{C}, \quad (13)$$

where $\lambda_1 = \frac{d_\mathrm{AC}}{d_\mathrm{AC} + d_\mathrm{AB}}$, $\lambda_2 = \frac{d_\mathrm{AB}}{d_\mathrm{AC} + d_\mathrm{AB}}$, and $d_\mathrm{AB} = \|\mathbf{l}_\mathrm{RX}^\mathrm{A} - \mathbf{l}_\mathrm{RX}^\mathrm{B}\|$, $d_\mathrm{AC} = \|\mathbf{l}_\mathrm{RX}^\mathrm{A} - \mathbf{l}_\mathrm{RX}^\mathrm{C}\|$ represents the Euclidean distances from point A to B and C, respectively.

4. Numerical Results and Analysis

In this section, the performance of the proposed PMEP-CC is evaluated through simulation. NMSE is chosen as the metric to measure the accuracy of CKM. As shown in Figure 5, an area range of 500×500 m is considered. We focus on the construction of the downlink CKM for the region shown in the red box of size 240×240 m, where the BS is located at the origin with a height of 20 m. The dense distribution of obstacles in the area leads to a large number of NLOS, which poses a strong challenge for building a full domain CKM and helps to demonstrate the performance of the proposed solution. In the simulation, the area is divided into equally spaced grids with a spacing of 6 m. Then, the Ray Tracing method and Matlab simulation are used to generate the downlink CSI, as in Equation (1), when the receiver is located at the grid point.

Due to no existing related novel work, to compare the proposed PMEP-CC performance, we use K-Means and geometric methods as benchmarks. The specific simulation

★ BS

Figure 5: Illustration of 3D simulation scene and meshing.

Table 1: Simulation parameters for communication scenarios.

Simulation parameter	Values
Building height	$10 - 30$ m
Number of grid points	41×41
High-dimensional mapping coefficient (σ)	30
Preset number of partitions for K-Means	12
Number of partitions in geometric partitioning algorithm	9
Transmitter power	10 dBm
Transmit signal frequency	2.8×10^{10} Hz
Number of transmitter	1
The accuracy of CKM	6 m

parameters for the and AP environment partitioning algorithms. The detailed simulation parameters of high-dimensional mapping, K-Means and geometric partitioning are shown in Table 1.

Figure 6 presents the results of the partitioning using geometric partitioning, K-Means, and the proposed AP algorithm, the blank space indicates that the grid point is a blind coverage area where the signal cannot be effectively received due to severe attenuation. Figures 6(a) and 6(b) show the partitioning effect of the K-Means algorithm using the original data and preprocessed (e.g., mapping) data, respectively. It can be seen from the figures, although the majority of adjacent grid points is separated into the same subregion, the partitioning results do not accurately reflect the actual geographical features and present confounding colors in certain locations. This is primarily due to limitations of the K-Means algorithm, which can only perform simple linear partitioning and struggles to divide partitions into arbitrary shapes. Meanwhile, the **CI** contains multidimensional information that cannot be directly compared without reflecting changes to the data. Also, increasing the number of partitions does not improve the K-Means algorithm's performance, as seen in the simulation results.

Figures 6(c) and 6(d) show the partitioning results obtained by the proposed AP algorithm using original and preprocessed data, respectively. In Figure 6(c), spatial grid points are divided into three regions that roughly correspond to the geographical features, but some interlocking regions with wide ranges remain. In contrast, by using preprocessed data, geographical features are better represented, clear boundaries between subregions are established, and subregions' ranges are smaller—thereby enhancing channel estimation and reducing computational complexity.

Figure 6(e) illustrates the division of the communication scene into geometric regions of similar size, without regard to the surrounding environment, as a baseline for comparison in subsequent channel reconstruction. Within the coverage area, the transmitter location is fixed and the distribution of obstacles such as buildings and trees remains static for a period of time, so the radio propagation process is largely unaffected by the obstacles and distance, and therefore the shape and extent of the subareas divided according to channel correlation also remain stable. Therefore, the communication scene can be divided into uniformly sized grids, and by simulating the channel information at all grid points, the division of the communication scene area can be completed, and the CKM can be updated after more significant changes in the environment.

In order to test the performance of the interpolation method for constructing the CKM within the partitioned subregions of the AP environment, after matching the path order between grid points proposed, the data at all grid points obtained from the simulation were divided into two parts: the known sample dataset and the estimated prediction dataset. Using the algorithm proposed, the **CI** of the remaining grid point locations is estimated from the known sample data and the NMSE performance of the algorithm is analyzed.

Figure 7 presents the CKM construction results for each grid point in the different partitioning methods. The color temperature in Figure 7 indicates the magnitude of the error,
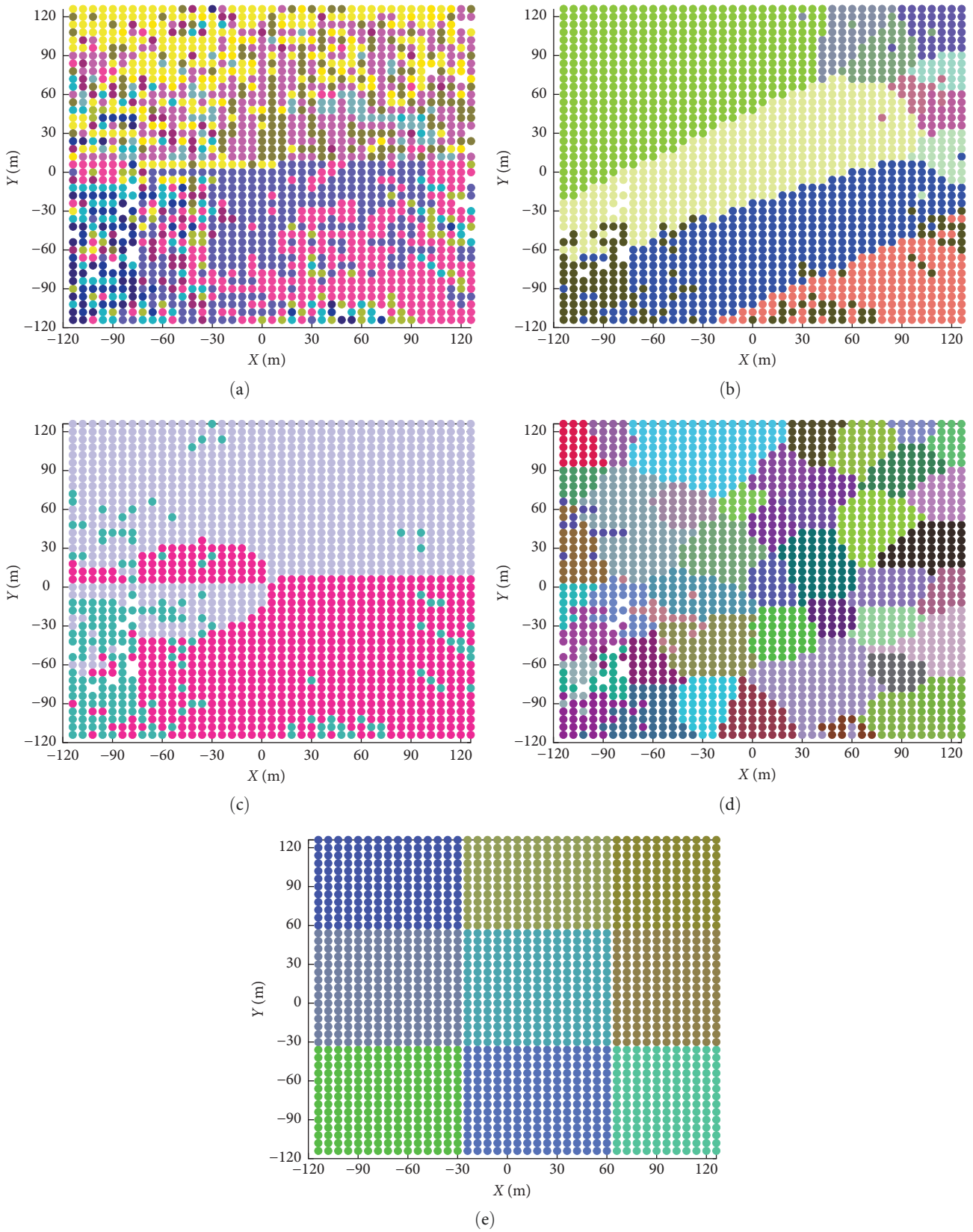
FIGURE 6: Schematic diagram of the partitioning effect and different colors indicate different partitions. (a) K-Means algorithm for partitioning the original data; (b) K-Means algorithm for partitioning the data after path matching; (c) AP environment partitioning algorithm for partitioning the original data; (d) AP environment partitioning algorithm for partitioning the data after path matching; (e) Geometric partitioning algorithm for partitioning the original data.
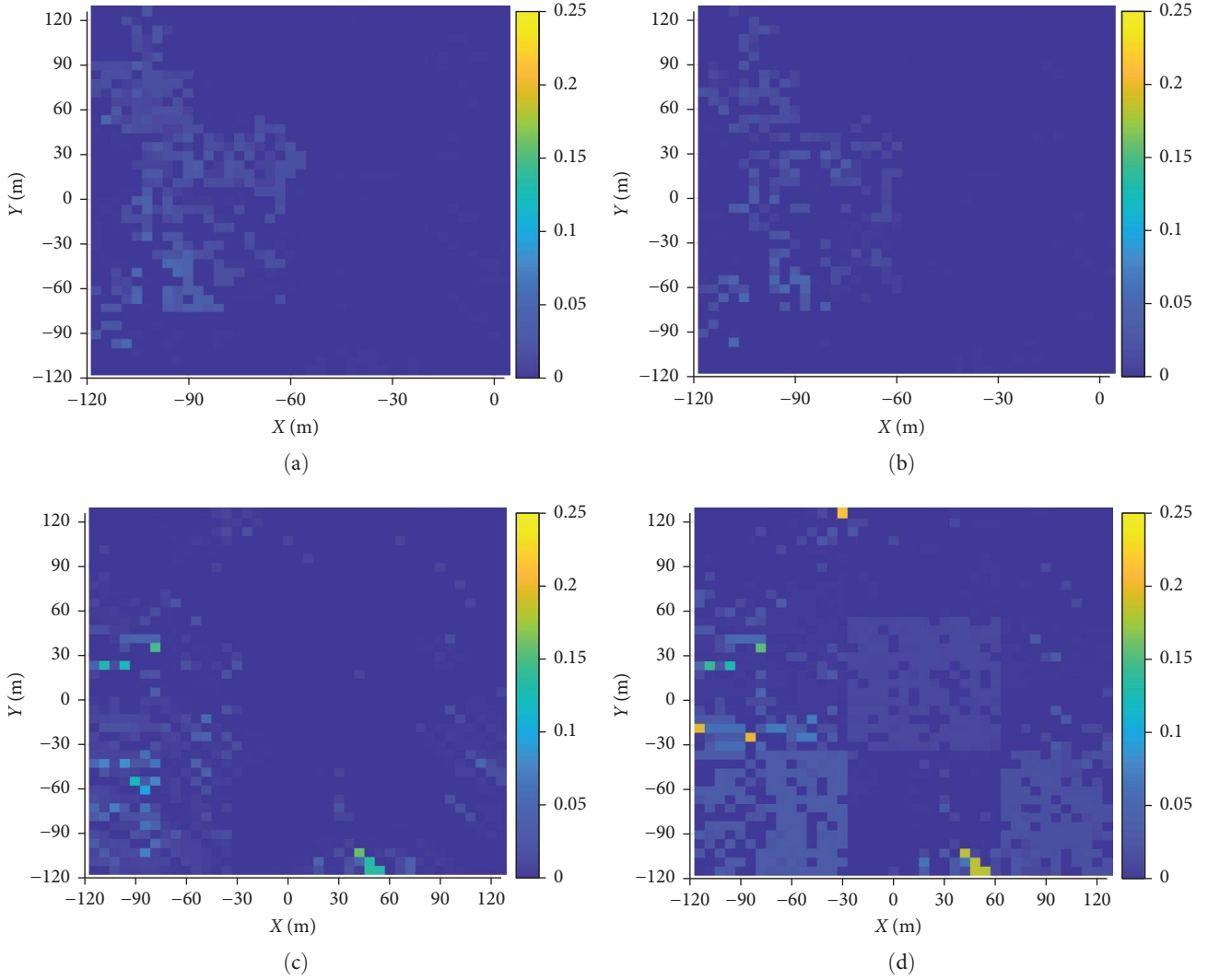
FIGURE 7: CKM construction error of different partitioning methods using data after path matching. (a) 20% D_{RF} under AP environment partitioning algorithm; (b) 50% D_{RF} under AP environment partitioning algorithm; (c) 20% D_{RF} under K-Means partitioning algorithm; (d) 20% sampling under geometric partitioning algorithm.

and the range of values on the color scale represents the range of estimation NMSE. And the location where the color obviously change indicates that the CSI estimation error is large, and the color not obvious indicates that the error change is small. Figures 7(a) and 7(b) show the reconstruction errors when 20% and 50% of the grid points in $\mathcal{S}$, respectively. It was observed that the majority of the locations with small errors, and only a small percentage of errors being large, mainly in the left part of the diagram. This is due to the dense distribution of obstacles in this region, multiple reflections of the signal and large channel variations that are difficult to estimate. Furthermore, as the number of sampled grid points increases, the overall estimation performance improves. Figures 7(c) and 7(d) show the channel reconstruction results under the K-Means partitioning algorithm and the geometric partitioning algorithm, respectively. Comparing the simulation results of interpolation errors in the area divided by different partitioning methods in Figure 7(d),

we can see that CSI estimation results under the AP environmental partitioning algorithm are more accurate.

Figure 8 shows the cumulative error distribution for estimating the CSI at the remaining 80% of the grid points under the different partitioning algorithms, where the CSI at 20% of the grid points is randomly selected as reference CSI. The results in Figure 8 show that the errors under the geometric partitioning algorithm are significantly higher than those under the AP environmental partitioning algorithm and the K-Means partitioning algorithm. This indicates the need for targeted partitioning based on environmental characteristics. Meanwhile, as can be seen from the figure, the error performance under the AP environmental partitioning algorithm proposed is excellent, with the error less than −40 dB at about 20% of the locations and below −25 dB at about 80% of the locations, outperforming the K-Means algorithm and the geometric partitioning method. This is because the proposed AP environment partitioning method compared to the methods
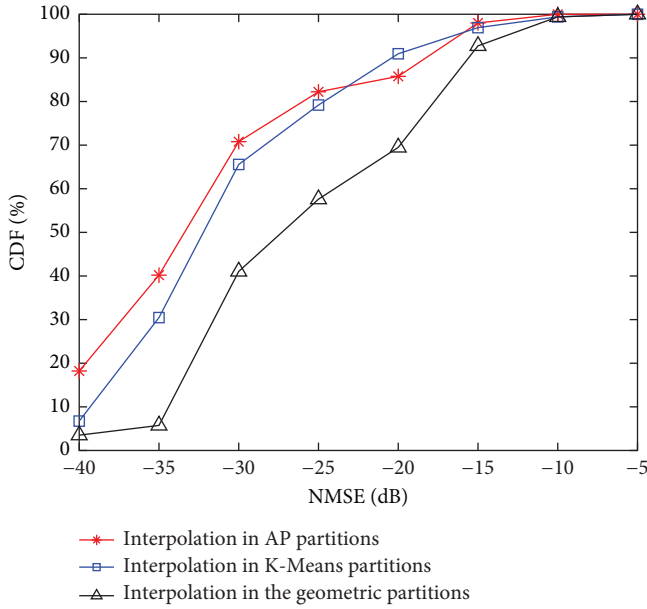
FIGURE 8: Distribution of interpolation errors within different partitioning algorithms.

such as K-Means, it can be more effective to improve estimation accuracy based on the perceived environmental changes. It should be noted that for obstacle-dense areas, the signals are reflected several times and all three partitioning methods show large errors. This is reason that the three curves get closer between −15 and −10 dB.

Figures 7 and 8 show that the CKM construction method proposed is able to accurately estimate the CSI of the remaining locations and meet the communication requirements. Therefore, we can retrieve and obtain the CSI between the transceivers from the CKM constructed in this paper combined with the precise locations of the transceivers, which can avoid or significantly reduce the pilot overhead and greatly reduce the system computational complexity. The complexity of the algorithm is divided into two main aspects, on the one hand the partitioning process is done offline with a complexity of $\mathcal{O}(6N_\mathrm{P}^2 \times M^2 + M^2 \times \log M)$, where N_P denotes the number of valid paths, M represents the number of grid points. The radio propagation environment is stable, and the partition is almost unchanged in a long period of time, without secondary zoning in a short time. On the other hand, the interpolation process is done online and the complexity of the algorithm is $\mathcal{O}((2E \times R + 6N_\mathrm{P}) \times S_\mathrm{sub})$, where R and E denote the number of reference location and points to be estimated, respectively, within the subarea, and S_sub is the number of partitions.

We have tried our best to find any relevant literature, but there is a little work on CKM construction. To the best of our knowledge, we are the first to propose using CSI to construct CKM. Among recently published related papers, most works use empirical signal models or machine learning to estimate RSS, which are significantly different from our work.

5. Conclusion

In this paper, we have proposed a CKM construction method based on PMEP-CC to estimate the CSI of the whole domain using the CSI of reference locations. Without relying on a specific path loss model, the method matches similar paths from the CSI and determines the correlation of CSI at different locations. Then, the spatial correlation is used to partition the radio propagation area into subregions and the CSI at remaining locations are estimated within the each subregion. The CSI of 20% grid points is randomly selected as the reference information to construct the CKM with an accuracy of 6 m. In the simulation results, 82% of the grid points to be estimated have NMSE under −25 dB. Based on the transceiver locations and the CKM constructed in this paper, high-accuracy CSI can be obtained without relying on pilot training, which in turn supports transmission techniques such as beamforming and precoding. The proposed method is able to construct CKM accurately, and the simulation results have verified the effectiveness of the proposed scheme.

Data Availability

Data supporting this research article are available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] W. Saad, M. Bennis, and M. Chen, "A vision of 6g wireless systems: applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2019.

[2] Y. Zeng and X. Xu, "Toward environment-aware 6G communications via channel knowledge map," *IEEE Wireless Communications*, vol. 28, no. 3, pp. 84–91, 2021.

[3] 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, no. 3, pp. 223–229, 2002.

[4] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, 2011.

[5] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 192–201, 2013.

[6] X. Hou, Z. Zhang, and H. Kayama, "Time-domain MMSE channel estimation based on subspace tracking for OFDM systems," in *2006 IEEE 63rd Vehicular Technology Conference*, pp. 1565–1569, IEEE, Melbourne, Vic., 2006.

[7] M. B. Sutar and V. S. Patil, "LS and MMSE estimation with different fading channels for OFDM system," in *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)*, pp. 740–745, IEEE, Coimbatore, India, 2017.

[8] Z. Wan, Z. Gao, and M.-S. Alouini, "Broadband channel estimation for intelligent reflecting surface aided mmWave massive MIMO systems," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, Dublin, Ireland, 2020.

[9] J. L. Paredes, G. R. Arce, and Z. Wang, "Ultra-wideband compressed sensing: channel estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 3, pp. 383–395, 2007.

[10] H. Xie, F. Gao, S. Jin, J. Fang, and Y.-C. Liang, "Channel estimation for TDD/FDD massive MIMO systems with channel covariance computing," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 4206–4218, 2018.

[11] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6169–6183, 2015.

[12] Q. K. Nguyen, V. D. Nguyen, Q. T. Nguyen, and D. T. Ha, "A channel allocation method for multi-cell OFDMA-FDD networks," in *International Conference on Communications and Electronics 2010*, pp. 232–236, IEEE, Nha Trang, Vietnam, 2010.

[13] V. Kafedziski, "Compressed sensing based channel estimation in FDD multi-user massive MIMO using angle domain sparsity and transmit antenna correlation," in *2018 European Conference on Networks and Communications (EuCNC)*, pp. 1–9, IEEE, Ljubljana, Slovenia, 2018.

[14] Y. Liu and X. Jiang, "Robust super-resolution frequency division duplex (FDD) channel estimation," in *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 1–5, IEEE, 2020.

[15] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2017.

[16] E. Dall'Anese, S.-J. Kim, and G. B. Giannakis, "Channel gain map tracking via distributed kriging," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 1205–1211, 2011.

[17] D. Wu, Y. Zeng, S. Jin, and R. Zhang, "Environment-aware and training-free beam alignment for mmWave massive mimo via channel knowledge map," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–7, IEEE, Montreal, QC, Canada, 2021.

[18] Y. Tian, S. Yuan, W. Chen, and N. Liu, "Transformer based radio map prediction model for dense urban environments," in *2021 13th International Symposium on Antennas, Propagation and EM Theory (ISAPE)*, pp. 1–3, IEEE, Zhuhai, China, 2021.

[19] H. B. Yilmaz and T. Tugcu, "Location estimation-based radio environment map construction in fading channels," *Wireless Communications and Mobile Computing*, vol. 15, no. 3, pp. 561–570, 2015.

[20] X. Yuan, Y. Hu, J. Gross, and A. Schmeink, "Radio-map-based UAV placement design for uav-assisted relaying networks," in *2021 IEEE Statistical Signal Processing Workshop (SSP)*, pp. 286–290, IEEE, Rio de Janeiro, Brazil, 2021.

[21] H. Xia, S. Zha, J. Huang, and J. Liu, "Radio environment map construction by adaptive ordinary Kriging algorithm based on affinity propagation clustering," *International Journal of Distributed Sensor Networks*, vol. 16, no. 5, pp. 1–10, 2020.

[22] K. Sato, K. Suto, K. Inage, K. Adachi, and T. Fujii, "Space-frequency-interpolated radio map," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 714–725, 2021.

[23] K. Sato, K. Inage, and T. Fujii, "Radio environment map construction with joint space-frequency interpolation," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, pp. 051–054, IEEE, Fukuoka, Japan, 2020.

[24] C. Phillips, D. Sicker, and D. Grunwald, "Bounding the error of path loss models," in *2011 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, pp. 71–82, IEEE, Aachen, Germany, 2011.

[25] M. Srinivas and L. M. Patnaik, "Genetic algorithms: a survey," *Computer*, vol. 27, no. 6, pp. 17–26, 1994.

[26] K. Li, J. Chen, B. Yu, Z. Shen, C. Li, and S. He, "Supreme: fine-grained radio map reconstruction via spatial–temporal fusion network," in *2020 19th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pp. 1–12, IEEE, Sydney, NSW, Australia, 2020.

[27] X. Han, L. Xue, and Y. Xu, "A PSD maps estimation algorithm for underlay cognitive radio networks based on the convolutional neural network," *Mathematical Problems in Engineering*, vol. 2020, Article ID 4024631, 10 pages, 2020.

[28] S. Shrestha, X. Fu, and M. Hong, "Deep spectrum cartography: completing radio map tensors using learned neural models," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1170–1184, 2022.

[29] J. Chen, U. Yatnalli, and D. Gesbert, "Learning radio maps for UAV-aided wireless networks: a segmented regression approach," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, Paris, France, 2017.

[30] V. Va, J. Choi, T. Shimizu, G. Bansal, and R. W. Heath, "Inverse multipath fingerprinting for millimeter wave V2I beam alignment," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4042–4058, 2018.

[31] Z. Xiao, Z. Zhang, C. Huang, X. Chen, C. Zhong, and M. Debbah, "C-GRBFnet: a physics-inspired generative deep neural network for channel representation and prediction," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 8, pp. 2282–2299, 2022.

[32] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: how many antennas do we need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, 2013.

[33] D. Ding, D. Wu, Y. Zeng, S. Jin, and R. Zhang, "Environment-aware beam selection for IRS-aided communication with channel knowledge map," in *2021 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, IEEE, Madrid, Spain, 2021.

[34] B. Scholkopf, K.-K. Sung, C. J. C. Burges et al., "Comparing support vector machines with Gaussian kernels to radial basis function classifiers," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2758–2765, 1997.

[35] A. Anaissi, A. Braytee, and M. Naji, "Gaussian kernel parameter optimization in one-class support vector machines," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, Rio de Janeiro, Brazil, 2018.