

Diffraction and Scattering Aware Radio Map and Environment Reconstruction using Geometry Model-Assisted Deep Learning

Wangqian Chen and Junting Chen

Abstract—Machine learning (ML) facilitates rapid channel modeling for 5G and beyond wireless communication systems. Many existing ML techniques utilize a city map to construct the radio map; however, an updated city map may not always be available. This paper proposes to employ the received signal strength (RSS) data to jointly construct the radio map and the virtual environment by exploiting the geometry structure of the environment. In contrast to many existing ML approaches that lack of an environment model, we develop a virtual obstacle model and characterize the geometry relation between the propagation paths and the virtual obstacles. A multi-screen knife-edge model is adopted to extract the key diffraction features, and these features are fed into a neural network (NN) for diffraction representation. To describe the scattering, as oppose to most existing methods that directly input an entire city map, our model focuses on the geometry structure from the local area surrounding the TX-RX pair and the spatial invariance of such local geometry structure is exploited. Numerical experiments demonstrate that, in addition to reconstructing a 3D virtual environment, the proposed model outperforms the state-of-the-art methods in radio map construction with 10%–18% accuracy improvements. It can also reduce 20% data and 50% training epochs when transferred to a new environment.

Index Terms—Radio map, Environment sensing, Diffraction features, Scattering-aware, Deep learning, Neural network

I. INTRODUCTION

RADIO maps find extensive applications in wireless communications, such as network planning, localization and cognitive radio systems [1]–[4]. For example, radio maps have been constructed and exploited to assist for low altitude unmanned aerial vehicle (UAV) communications [5]–[8], in which radio maps can help UAV seek line-of-sight (LOS) opportunities to serve for the ground users that are probably blocked by obstacles in a dense urban environment.

The work was supported in part by the National Science Foundation of China (NSFC) under Grant No. 62293482, by the Basic Research Project No. HZQB-KCXYZ-2021067 of Hetao Shenzhen-HK S&T Cooperation Zone, by NSFC Grant No. 62171398, by the Shenzhen Science and Technology Program under Grant No. JCYJ20210324134612033 and No. KQTD2020090911473003, by Guangdong Research Projects No. 2019QN01X895, No. 2017ZT07X152, and No. 2019CX01X104, by the Shenzhen Outstanding Talents Training Fund 202002, by the Guangdong Provincial Key Laboratory of Future Networks of Intelligence (Grant No. 2022B1212010001), by the National Key R&D Program of China with grant No. 2018YFB1800800, and by the Key Area R&D Program of Guangdong Province with grant No. 2018B030338001.

W. Chen and J. Chen are with the School of Science and Engineering, and the Future Network of Intelligence Institute (FNii), The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China (email: wangqianchen@link.cuhk.edu.cn; juntingc@cuhk.edu.cn).

Classical radio map construction can be roughly categorized into ray tracing (RT) approaches, interpolation approaches, and deep learning based approaches. RT approaches [9]–[12] construct the propagation channel by launching propagation rays via geometry analysis based on 3D digital environment models. Fine-grained propagation models including reflection models, diffraction models, and scattering models as well as the electromagnetic coefficients of the materials in the environment are required for evaluating the channel coefficients. However, obtaining an accurate digital environment map with sufficient details for an accurate RT is very challenging. It is also computationally intense to simulate the propagation rays due to the complexity of an actual environment. Moreover, based on the classical RT models, it is almost prohibitive to solve the inverse problem that reconstructs the geometry of the environment based on the radio measurements.

Interpolation-based radio map construction techniques exploit real measurements taken at various locations of transmitters (TXs) and receivers (RXs) without explicitly exploiting the geometry structure of the environment. Some representative interpolation methods for radio map construction include k -nearest neighbor (KNN) interpolation [13], inverse distance weighted (IDW) interpolation [14], matrix completion [15], dictionary-based compressive sensing [16], and Kriging [17], etc. These methods are based on the spatial correlation of measurements, but they cannot differentiate the corresponding propagation conditions, such as LOS and non-line-of-sight (NLOS). To address this issue, a segmented propagation model was proposed in [18], and the measurement data is clustered into different categories corresponding to the different propagation conditions using subspace clustering, where different propagation models are applied locally for constructing radio map. In general, many interpolation-type methods assume that the TX locations are fixed. When the TX location also varies, conventional interpolation-type methods may suffer from the curse of dimensionality issue, where a prohibitive amount of data is needed to learn the spatial structure of the radio map.

Deep learning approaches for radio map construction were attempted in [19]–[21], where the features that are fed into the neural network (NN) include building heights, propagation distances, and the environment maps. In [22, 23], a convolutional neural network (CNN) based structure was proposed to learn the channel path loss based on 3D city maps. Some studies also proposed to fuse the information from satellite images to learn the characteristics of the propagation environment [24]–[26].

TX变化
需要海量数据
来学习
空间
结构

In addition, with an adequate number of radio measurements, deep completion autoencoders [27] and generative adversarial network [28] were also adopted for radio map construction.

To summarize, most machine learning approaches directly interpolate the radio map from a large number of measurements or they fuse the environment information to assist for the interpolation, without explicitly exploiting the geometry structure of the environment. Although some existing CNN based approaches employ city maps as input, the maps are usually treated as images rather than 3D structure. As a result, it is challenging to build a radio map with full spatial degrees of freedom where both the TX position and the RX position can be varying over the whole 3D space; instead, existing approaches usually require at least one coordinate of the TX and RX locations to be fixed. Moreover, little is known on how to reconstruct the propagation environment from pure radio measurements when a city map is not available or inaccurate.

This paper attempts to construct a radio map with full spatial degrees of freedom for the large-scale channel quality between variable TX and RX locations based on received signal strength (RSS) measurements. In contrast to many existing machine learning approaches that lack of an environment model, we aim at constructing a propagation model that explicitly characterizes the propagation attenuation as a function of the geometry structure of the propagation environment. In addition, we propose to jointly reconstruct the geometry of the propagation environment and the radio map. Towards this end, we develop a virtual obstacle model and we characterize the propagation attenuation by analyzing the geometry relation between the propagation rays and the virtual obstacles. While our earlier attempts [29, 30] only considered signal attenuation due to blockage along the direct path, some significant factors including diffraction and reflection were ignored, resulting in an oversimplified propagation model. In a related approach, simultaneous localization and mapping (SLAM) usually assumes at most one reflection and no diffraction as a compromise with the complexity. In this work, we attempt to also model the effect of diffraction and scattering for a more accurate radio map construction as well as a virtual environment reconstruction.

Specifically, we adopt the virtual obstacle concept developed in [29, 30], but the virtual obstacle constructed in this paper not only models the signal blockage of the direct path, but also captures the signal attenuation due to diffraction and scattering. A multi-screen knife-edge model is adopted to extract the key diffraction parameters of the propagation path, and these parameters are fed into an NN for diffraction representation. To describe the scattering, as oppose to most existing methods that directly map an entire city map or satellite image to a radio map, our model only focuses on the geometry structure from the local area surrounding the TX-RX position pair, resulting in efficient data augmentation and model parameter reduction.

The novelty and contribution are summarized as follows:

- We propose a deep learning model to jointly construct a 6D radio map and a virtual geometry of the propagation environment from RSS measurements. The model consists of an LOS branch, a diffraction branch, and a

scattering branch that explicitly or implicitly characterize the geometry relation between the propagation attenuation and the virtual environment.

- In the diffraction branch, we adopt a transformer structure to approximate the Vogler's method for an efficient computation of the diffraction coefficient. In the scattering branch, we develop a rotation invariant and scale invariant mapping for data augmentation to learn the geometry structure of the local scatters, enhancing the convergence of the CNN model parameters.
- Simulations demonstrate that the proposed model outperforms the state-of-the-art methods in radio map construction with 10%–18% accuracy improvements and it can be generalized well to the whole 3D space without any additional data to fine-tune the parameters.
- Transferability of the proposed model is demonstrated with nearly 20% data reduction and 50% training epoch reduction compared with the non-transferred cases.
- We present an application of radio-map-based UAV relay placement in which the proposed model accelerates the UAV relay positioning with reducing 99% search distance with the classical radio-map-based methods.

The rest of the paper is organized as follows. The system model are reviewed in Section II. Section III discusses the proposed learning framework. Design examples are presented in Section IV and conclusions are drawn in Section V.

II. SYSTEM MODEL

Denote a wireless channel $\tilde{\mathbf{p}} = (\mathbf{p}_t, \mathbf{p}_r) \in \mathbb{R}^6$ using the positions $\mathbf{p}_t, \mathbf{p}_r \in \mathbb{R}^3$ of the TX and RX, respectively. Consider a blockage-aware channel model that describes the channel attenuation in logarithmic scale for $\tilde{\mathbf{p}}$ as

$$g_{[\text{dB}]}(\tilde{\mathbf{p}}, \mathbf{H}) = g_0(\tilde{\mathbf{p}}, \mathbf{H}) + g_d(\tilde{\mathbf{p}}, \mathbf{H}) + g_s(\tilde{\mathbf{p}}, \mathbf{H}) + \xi(\tilde{\mathbf{p}}) \quad (1)$$

where \mathbf{H} is a parameter to describe the propagation environment. The first term $g_0(\tilde{\mathbf{p}}, \mathbf{H})$ captures the path loss if the link $\tilde{\mathbf{p}}$ is in the LOS condition, the second term $g_d(\tilde{\mathbf{p}}, \mathbf{H})$ describes the additional loss due to diffraction when the link is in NLOS condition, the third term $g_s(\tilde{\mathbf{p}}, \mathbf{H})$ describes the fluctuation of the attenuation due to the other scattering effects, and finally, the term $\xi(\tilde{\mathbf{p}})$ is a random component that captures the uncertainty due to the small-scale fading.

The radio map studied in this paper captures the large-scale channel information for each link $\tilde{\mathbf{p}}$ in the area of interest, i.e.,

$$G(\tilde{\mathbf{p}}) = g_0(\tilde{\mathbf{p}}, \mathbf{H}) + g_d(\tilde{\mathbf{p}}, \mathbf{H}) + g_s(\tilde{\mathbf{p}}, \mathbf{H}). \quad (2)$$

Note that the environment parameter \mathbf{H} is a hidden variable that can be learned from the data.

We measure only the location-labeled RSS $y(\tilde{\mathbf{p}}) = G(\tilde{\mathbf{p}}) + n$ at various locations $\tilde{\mathbf{p}}$ for the construction of the radio map, where the small-scale fading $\xi(\tilde{\mathbf{p}})$ in (1) is assumed to be partially averaged and absorbed as the measurement noise n .

A. Blockage Model

We adopt a virtual obstacle model to describe the propagation environment. Partition the ground area of interest into

M grid cells. The virtual obstacle on the m th grid cell is modeled as a cube that occupies the m th grid with height h_m as shown in Fig. 1. Denote \mathbf{h} as a vector for the collection of all h_m , $m = 1, 2, \dots, M$, and denote \mathbf{H} as the matrix form of h_m such that $\mathbf{h} = \text{vec}(\mathbf{H})$, where the (i, j) th entry of \mathbf{H} represents the virtual obstacle at the (i, j) th grid cell. Hence, the variable \mathbf{H} completely characterizes the geometry of the virtual environment. Note that these virtual obstacles do not have to exactly match with the city map, but they are to be fitted such that the channel quality based on the radio map model (2) matches with the radio measurements.

For each position pair $\tilde{\mathbf{p}}$, denote $\mathcal{B}(\tilde{\mathbf{p}})$ as the set of grid cells that are covered by the line segment joining \mathbf{p}_t and \mathbf{p}_r . For each grid cell $m \in \mathcal{B}(\tilde{\mathbf{p}})$, denote $z_m(\tilde{\mathbf{p}})$ as the height of the line segment of $\tilde{\mathbf{p}}$ that passes over the m th grid cell. Denote $\tilde{\mathcal{D}}_0$ as the set of channels $\tilde{\mathbf{p}}$ which are not blocked by any virtual obstacle, and \mathcal{D}_0 is termed as the (virtual) *LOS region* in this paper. Mathematically, for an LOS case $\tilde{\mathbf{p}} \in \tilde{\mathcal{D}}_0$, we have $h_m < z_m(\tilde{\mathbf{p}})$ for all grid cells $m \in \mathcal{B}(\tilde{\mathbf{p}})$; for an NLOS case $\tilde{\mathbf{p}} \notin \tilde{\mathcal{D}}_0$, we have $h_m \geq z_m(\tilde{\mathbf{p}})$ for at least one grid cell $m \in \mathcal{B}(\tilde{\mathbf{p}})$, i.e., the direct path from \mathbf{p}_t to \mathbf{p}_r is blocked by at least one of the virtual obstacles along the path. Thus, a model for the LOS region $\tilde{\mathcal{D}}_0(\mathbf{H})$ given the parameter \mathbf{H} for the heights of the virtual obstacles can be formulated as

$$\mathbb{I}\{\tilde{\mathbf{p}} \in \tilde{\mathcal{D}}_0(\mathbf{H})\} = \prod_{m \in \mathcal{B}(\tilde{\mathbf{p}})} \mathbb{I}\{h_m < z_m(\tilde{\mathbf{p}})\} \quad (3)$$

where $\mathbb{I}\{A\}$ is an indicator function that takes value 1 if condition A is true, and 0, otherwise.

As a result, when a communication link is not blocked by any obstacle, the channel attenuation is modeled as

$$g_0(\tilde{\mathbf{p}}, \mathbf{H}) = (\beta_0 + \gamma_0 \log_{10} \|\mathbf{p}_t - \mathbf{p}_r\|_2) \mathbb{I}\{\tilde{\mathbf{p}} \in \tilde{\mathcal{D}}_0(\mathbf{H})\} \quad (4)$$

where β_0 and γ_0 are the path loss parameters to be learned for $\tilde{\mathbf{p}} \in \tilde{\mathcal{D}}_0$, and $\|\cdot\|_2$ represents the Euclidean norm.

B. Diffraction Model

In propagation theory, when the direct path is blocked, the diffraction mechanism allows the electromagnetic wave to propagate around obstacles and contribute to the RX located in the shadow region. We adopt a multi-screen knife-edge model reference, which assumes that the diffraction effect mainly occurs on the edge of the obstacles, and the obstacles are simplified as knife-edge shapes. We thus build a connection between the diffraction and the virtual environment \mathbf{H} .

Specifically, as illustrated in the example shown in Fig. 1, consider the set of grid cells $\mathcal{B}(\tilde{\mathbf{p}})$. When the line segment $\overline{A_0B}$ joining \mathbf{p}_t and \mathbf{p}_r intersects with at least one virtual obstacle, i.e., equation (3) equals to 0, then construct a concave piecewise linear curve $\overline{A_0A_1 \cdots A_N B}$ with the shortest length, such that each vertex A_1, A_2, \dots, A_N touches the top of a virtual obstacle in $\mathcal{B}(\tilde{\mathbf{p}})$ and no segment $\overline{A_i A_{i+1}}$ intersects with any virtual obstacles. The virtual obstacles where the vertexes of the curve touch on are modeled as the multiple knife edges that mainly contribute to the diffraction attenuation along the direct path, while all the other obstacles below the height of the segment $\overline{A_i A_{i+1}}$ are neglected. The curve

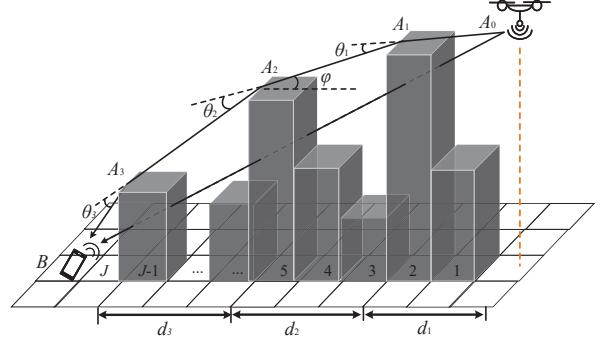


Figure 1. Modeling the diffraction between point A_0 and point B by exploiting the geometry relation of distances d_i and angles θ_i from the virtual obstacles (gray cubes).

represents the main diffraction propagation path around the obstacles from \mathbf{p}_t to \mathbf{p}_r .

Define d_i as the horizontal (ground) distance between vertices A_i and A_{i-1} , and θ_i as the diffraction angle between adjacent line segments $\overline{A_i A_{i-1}}$ and $\overline{A_{i+1} A_i}$. Vogler's method [31] suggests that the distances d_i and the angle θ_i in each segment are the key parameters that affect the diffraction attenuation. See Appendix A for a specific expression reported in the literature. The diffraction effect can be modeled as a function $f_d(d, \theta)$ that only depends on the distance variable $\mathbf{d}(\tilde{\mathbf{p}}, \mathbf{H}) = (d_1, d_2, \dots, d_N)$ and the angle variable $\boldsymbol{\theta}(\tilde{\mathbf{p}}, \mathbf{H}) = (\theta_1, \theta_2, \dots, \theta_N)$. Considering that diffraction mainly occurs under the NLOS condition, the diffraction component in the proposed radio map model (2) is expressed as

$$g_d(\tilde{\mathbf{p}}, \mathbf{H}) = f_d(\mathbf{d}, \boldsymbol{\theta})(1 - \mathbb{I}\{\tilde{\mathbf{p}} \in \tilde{\mathcal{D}}_0(\mathbf{H})\}). \quad (5)$$

C. Scattering Model

We adopt an ellipse model that only focuses on the local scatters in an elliptical area in the neighborhood of the ground positions of \mathbf{p}_t and \mathbf{p}_r . This is because remote scatters are likely blocked by other obstacles and the corresponding propagation distance is also longer, resulting in a relatively small impact on the channel quality. In addition, to compromise with the modeling complexity, we need the following properties for our scattering model:

- *Rotation invariance*: Since only the relative positions of TX and RX in the virtual environment matter with the propagation, the same attenuation can be observed if one rotates the entire environment.
- *Scale invariance*: Recall that the first two terms in the radio map model (2) depends on the propagation distance. As the scattering component $g_s(\tilde{\mathbf{p}}, \mathbf{H})$ adds to the first two terms in log-scale, it aims at capturing the relative attenuation. Thus, a distance-independent modeling may yield a good approximation.

This philosophy is demonstrated in Fig. 2, where there are three wireless channels from one TX to three RXs at different locations. As shown in Fig. 2, the three channels share a similar local propagation environment, where each channel consists of a slightly obstructed direct path with diffraction

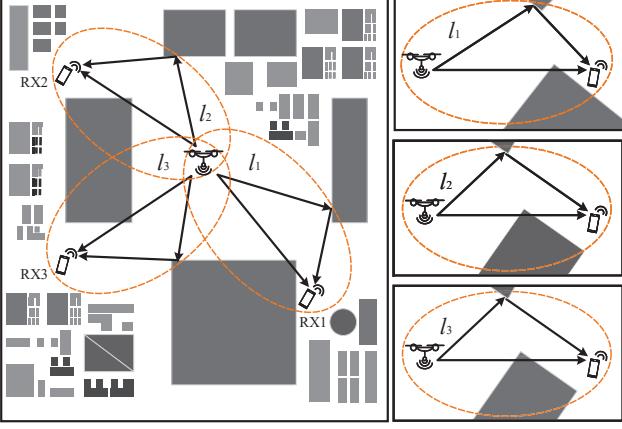


Figure 2. While the propagation environments from the TX to three RXs are different (left figure), the local geometry structures are similar after proper rotation and scaling (right figures). Thus, the same propagation mechanism should be learned.

and a strong reflected path. As such, one would expect that the three channels share the same scattering component in (2).

Specifically, the scattering component $g_s(\tilde{\mathbf{p}}, \mathbf{H})$ is constructed as follows. Recall our notation for the TX position $\mathbf{p}_t = (\bar{\mathbf{p}}_t, p_{t,3}) \in \mathbb{R}^3$, where $\bar{\mathbf{p}}_t = (p_{t,1}, p_{t,2}) \in \mathbb{R}^2$ is the ground projected location of the TX and $p_{t,3}$ is the altitude; similar notation $\mathbf{p}_r = (\bar{\mathbf{p}}_r, p_{r,3})$ is defined for the RX. Denote $d_0 = \|\bar{\mathbf{p}}_t - \bar{\mathbf{p}}_r\|_2$ as the ground distance. Define an elliptical area using $\bar{\mathbf{p}}_t$ and $\bar{\mathbf{p}}_r$ as the foci with a given eccentricity parameter e . As a result, the major and minor axes of the ellipse are given by d_0/e and $(d_0/e)\sqrt{1-e^2}$, respectively. Define $\mathcal{B}_E(\tilde{\mathbf{p}}, \mathbf{H})$ as the set of grid cells within the elliptical area. The scattering component in (2) is modeled as

$$g_s(\tilde{\mathbf{p}}, \mathbf{H}) = f_s(\mathcal{B}_E(\tilde{\mathbf{p}}, \mathbf{H})) (1 - \mathbb{I}\{\tilde{\mathbf{p}} \in \tilde{\mathcal{D}}_0(\mathbf{H})\}) \quad (6)$$

where $f_s(\mathcal{B}_E(\tilde{\mathbf{p}}, \mathbf{H}))$ is a mapping that depends on a subset of entries of \mathbf{H} based on $\mathcal{B}_E(\tilde{\mathbf{p}}, \mathbf{H})$ and such a mapping f_s satisfies the rotation invariance and scale invariance properties.

III. NETWORK ARCHITECTURE AND TRAINING

In this section, a deep NN representation is constructed to model the relationship between the radio map $G(\tilde{\mathbf{p}})$ and the virtual obstacle environment \mathbf{H} . Specifically, we need to address the following technical challenges:

- How to construct a deep learning architecture for (2) such that the virtual obstacle map \mathbf{H} can be efficiently constructed from the RSS measurements? For example, it could be challenging to find a NN representation for the blockage relation (3) such that the variable \mathbf{H} can be efficiently updated by back-propagation for training.
- How to construct a NN expression for the Vogler diffraction model to exploit the geometry structure of diffraction? As is seen in Section II-B and Appendix A, the Vogler expression has a special mathematical structure, and a conventional fully connected NN (FCN) structure may not be a good choice for balancing the representation capability and the learning efficiency.

- How to design a CNN with the rotation invariance and scale invariance properties for the scattering component? While CNN is a common structure that can exploit both local and global features, it still needs to extend its capability for the rotation and scale invariance.

In this section, we will explore specialized network structures, such as multi-head attention and spatial transformation to address the above issues. The overall architecture of the proposed network is shown in Fig. 3.

A. Overall Architecture

As illustrated in Fig. 3, the proposed network consists of three main modules, Area Filter, Area Focusing, and Path Loss Prediction. For each module, there are two to three branches focusing on the LOS component g_0 , the diffraction component g_d , and the scattering component g_s in (2). The virtual obstacle environment that has been modeled as a 2D virtual obstacle map \mathbf{H} is processed by the Area Focusing module.

1) Area Filter: Recall that the signal propagation for a wireless channel $\tilde{\mathbf{p}}$ is mostly determined by the local area surrounding the locations of the TX \mathbf{p}_t and the RX \mathbf{p}_r . The Area Filter module thus aims at determining the relevant local areas of $\tilde{\mathbf{p}}$ for the LOS component, the diffraction component, and the scattering component.

Specifically, for a given $\tilde{\mathbf{p}}$, the Area Filters are formed as $M_1 \times M_2$ matrices to describe a 3D line structure \mathbf{L} and an ellipse mask \mathbf{M} , where the 3D line structure \mathbf{L} aims at computing the level of obstruction for the wireless channel $\tilde{\mathbf{p}}$ due to the virtual obstacle map \mathbf{H} and selecting the relevant virtual obstacles h_m to form the set $\mathcal{B}(\tilde{\mathbf{p}})$ in (3); and the ellipse mask \mathbf{M} aims at selecting the virtual obstacles in the surrounding of $\tilde{\mathbf{p}}$ to construct $\mathcal{B}_E(\tilde{\mathbf{p}}, \mathbf{H})$ and learn f_s in (6).

For the line structure, we project the altitude of the direct line connecting \mathbf{p}_t and \mathbf{p}_r to the ground. Recall that $\mathcal{B}(\tilde{\mathbf{p}})$ is the set of grid cells where the line segment joining \mathbf{p}_t and \mathbf{p}_r passes through. Thus, for $m \in \mathcal{B}(\tilde{\mathbf{p}})$, let $\bar{\mathbf{c}}_m \in \mathbb{R}^2$ be the corresponding ground location of the m th grid. The m th element of the matrix \mathbf{L} is given by $L_m = (p_{t,3} - p_{r,3}) \frac{\|\bar{\mathbf{c}}_m - \bar{\mathbf{p}}_r\|_2}{\|\bar{\mathbf{p}}_t - \bar{\mathbf{p}}_r\|_2} + p_{r,3}$, if $m \in \mathcal{B}(\tilde{\mathbf{p}})$; and $L_m = 0$, for $m \notin \mathcal{B}(\tilde{\mathbf{p}})$. L: 沿直达路径高度

For the ellipse mask, the entry of the matrix \mathbf{M} takes value 1 if the entry is covered by the ellipse with an eccentricity parameter e and foci $\bar{\mathbf{p}}_t$ and $\bar{\mathbf{p}}_r$, and it takes value 0 otherwise.

2) Area Focusing: The Area Focusing module is to select a subset of virtual obstacle variables h_m and form the local environment feature for the wireless channel $\tilde{\mathbf{p}}$. As a result, only the subset of the relevant variables h_m are updated through back-propagation during the training process.

Based on the line structure \mathbf{L} , the line feature map of Area Focusing can be formulated as

$$\mathbf{O}_L = \text{ReLU}((\mathbf{H} - \mathbf{L}) \odot \text{sign}(\mathbf{L})) \quad (7)$$

where $\text{ReLU}(\cdot)$ and $\text{sign}(\cdot)$ are element-wise operators, with $\text{ReLU}(x) = \max(x, 0)$ and $\text{sign}(x) = \mathbb{I}\{x > 0\}$, and \odot is the Hadamard product. It can be verified that the non-zero elements of \mathbf{O}_L indicate the locations where the propagation is blocked along the direct path. In addition, the value of the non-zero entries of \mathbf{O}_L may represent the level of blockage.

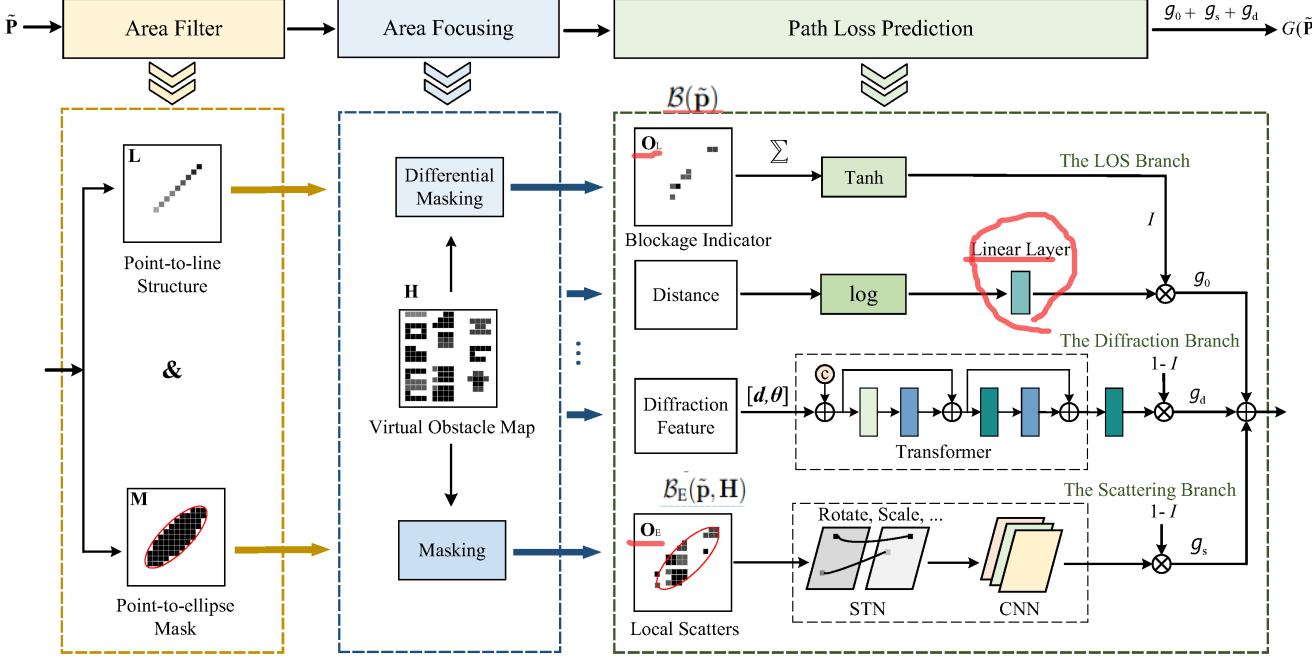


Figure 3. The overall architecture to exploit the geometry structure of the environment for joint 6D radio map and virtual environment map reconstruction.

Based on the ellipse mask, the ellipse feature map of Area Focusing is calculated as

$$\mathbf{O}_E = \mathbf{H} \odot \mathbf{M} \quad (8)$$

where only the obstacle variable h_m that locates inside the ellipse indicated by the mask \mathbf{M} are selected for \mathbf{O}_E .

3) *Path Loss Prediction*: The Path Loss Prediction module consists of three branches. As in Fig. 3, the first two branches from the top take the feature \mathbf{O}_L to predict the path loss from the LOS component and the diffraction component. The third branch takes the feature \mathbf{O}_E to predict the additional attenuation from the scattering. These are illustrated as follows.

B. The LOS Branch

The LOS branch is to implement the indicator function in (4), which is approximated by a tanh function for a non-degenerated gradient in the training phase. We thus have the following soft indicator function as

$$I = 1 - \tanh(\text{sum}(\mathbf{O}_L)) \quad (9)$$

where $\text{sum}(\mathbf{A}) = \sum_{ij} a_{ij}$ and $\tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$.

It follows that when the blockage indicator matrix \mathbf{O}_L contains all zeros, implying that the propagation is not blocked by any virtual obstacle, we have the soft indicator $I = 1$. On the other hand, when \mathbf{O}_L contains non-zeros values and $\text{sum}(\mathbf{O}_L) > 2$, we have $I \leq 0.04$, implying that the propagation is very likely blocked. Here, the 2-meter margin is to match with the intuition from real measurements that there is a soft transition from LOS to NLOS.

The path loss model in (4) is implemented by a regression layer as seen from Fig. 3.

C. The Diffraction Branch

The diffraction branch is designed to mimic the computation structure of the Vogler's method illustrated in Appendix A. The Vogler expression has a special recursive structure between the computation of the N th diffraction F_N and the $(N-1)$ th diffraction F_{N-1} . The diffraction distances d_i and angles θ_i play a key role in determining the diffraction attenuation. However, such inherently recursive and sequential nature precludes parallelization in practical implementation. The transformer network [32] processes all features using a multi-head attention mechanism which models the dependencies of recursive structure without the need of the order of time and make them highly parallelizable. Therefore, we propose to design a transformer network to learn the diffraction by exploiting the diffraction features of distances d_i and angles θ_i .

1) *Determine d_i and θ_i :* Using the example depicted in Fig. 1, the curve $A_0A_1 \dots A_3B$ characterizes the diffraction path and the distances d_i and angles θ_i are related to the locations of the vertexes. To locate the vertexes, order the grid cells in $\mathcal{B}(\tilde{\mathbf{p}})$ as $1, 2, \dots, J$ according to the distance away from \mathbf{p}_t . The coordinate and height of the virtual obstacle in j th grid cell are given as $\bar{\mathbf{r}}_j$ and h_j , respectively. To find the vertex A_1 , take A_0 as the base point and connect A_0 to the vertexes of all the virtual obstacles in $\mathcal{B}(\tilde{\mathbf{p}})$. Geometrically, the vertex A_1 is located in the j th grid cell where the segment joining the vertex of the virtual obstacle and A_0 generates the smallest elevation angle at the virtual obstacle, which is denoted as φ in Fig. 1. More generally, assume A_i is obtained in the K th grid cell with $K < j$. To search A_{i+1} , connect A_i and the vertex of the virtual obstacle in j th grid cell to generate the segments for all $j > K$, and calculate the elevation angle φ_j of each segment at the virtual obstacle. The vertex A_{i+1} is located in the l th grid cell where $l = \{j \mid j = \text{argmin}(\varphi_j), \forall j > K\}$.

or zenith at A_0

Algorithm 1 Virtual obstacle selecting algorithm for determining diffraction distances d_i and angles θ_i .

```

1: Input  $p_t$ ,  $p_r$ ,  $L$  and  $O_L$ 
2: Generate indexes for grid cells in  $\mathcal{B}(\tilde{p})$  with  $1, 2, 3, \dots, J$ 
3: Initialize  $\mathcal{I} = \{\bar{p}_t\}$ ,  $\mathcal{H} = \{p_{t,3}\}$ ,  $\mathcal{S} = \mathcal{B}(\tilde{p})$ ,  $l = 0$ ,  $h_0 = p_{t,3}$  and  $h_{J+1} = p_{t,3}$ 
4: while ( $\mathcal{S}$  is not empty) do
5:   for  $j = l+1$  to  $J+1$  do
6:      $\varphi_j = \arctan\left(\frac{h_l - h_j}{j}\right)$ 
7:   end for
8:    $l = \{j | j = \operatorname{argmin}(\varphi_j), \forall j > l\}$ 
9:   if  $v = J+1$  then
10:    Add  $\bar{r}_{J+1}$  to  $\mathcal{I}$  and add  $h_{J+1}$  to  $\mathcal{H}$ , go to step 15
11:   else
12:    Add  $\bar{r}_l$  to  $\mathcal{I}$  and add  $h_l$  to  $\mathcal{H}$ 
13:   end if
14: end while
15: for  $i = 1$  to  $\operatorname{length}(\mathcal{I}) - 1$  do
16:    $d_i = \|\mathcal{I}_i - \mathcal{I}_{i+1}\|_2$  and
       $\theta_i = \arctan\left(\frac{\mathcal{H}_i - \mathcal{H}_{i+1}}{\|\mathcal{I}_i - \mathcal{I}_{i+1}\|_2}\right) - \arctan\left(\frac{\mathcal{H}_{i-1} - \mathcal{H}_i}{\|\mathcal{I}_i - \mathcal{I}_{i-1}\|_2}\right)$ 
17: end for
18: Output  $d_i$  and  $\theta_i$ 

```

With these vertexes, the diffraction distances d_i and angles θ_i are determined. The details of the virtual obstacle selecting algorithm for determining d_i and θ_i are summarized in Algorithm 1. The extracted distances d_i and angles θ_i are input into the transformer network to learn the diffraction $f_d(d, \theta)$.

2) *Transformer representation of diffraction:* The transformer network is shown in Fig. 4. The whole network consists of a stack of three identical blocks followed with a FCN to generate the diffraction component. Each block has a multi-head attention layer, a feed forward layer and two normalization layers. Residual connection is employed after the layer normalization. To explore the sequential features of the distances d_i and angles θ_i , positional encoding is implemented as an additional input associated with the distances d_i and angles θ_i , which are all input to an attention layer to learn their dependencies between different positions to approximate the computation structure of the Vogler expression.

D. The Scattering Branch

The scattering branch consists of a spatial transformation network (STN) to capture the rotation and scaling invariance and a CNN to learn the geometry structure of the local scatters. In STN, the local scatters in the feature map of Area Focusing O_E will be translated, rotated and scaled under a definitive rule. The relative position information with geometry structure of the local scatters is extracted. Followed with such operations, a CNN takes these features of the local geometry structures to predict the scattering component.

The STN is implemented by coordinate transformation and sampling [33].

1) *Coordinate Transformation:* Define the output feature map as $\mathbf{F} \in \mathbb{R}^{M_1 \times M_2}$ that is translated, rotated and scaled from the input feature map O_E . Let $(x_s^{(i)}, y_s^{(i)})$ be the i th pixel

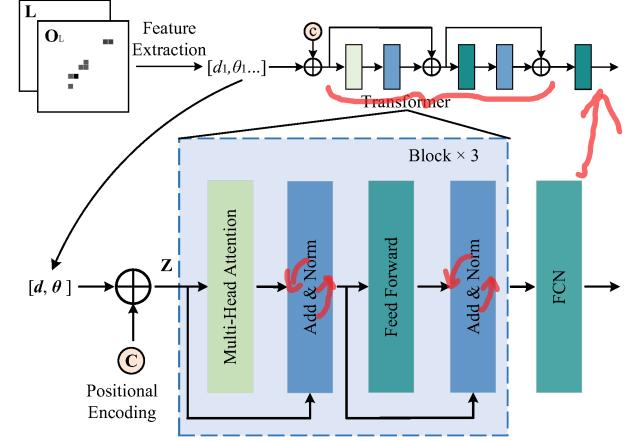


Figure 4. The transformer to learn the diffraction mechanism by exploiting the geometry structures of the diffraction distances d_i and angles θ_i .

coordinates in the output feature map \mathbf{F} for $i = 1, 2, \dots, M_1 M_2$. The corresponding source coordinates $(x_s^{(i)}, y_s^{(i)})$ of the input feature map O_E under transformation is formulated as

$$\begin{bmatrix} x_s^{(i)} \\ y_s^{(i)} \end{bmatrix} = \begin{bmatrix} c_1 \cos(\omega) & -c_2 \sin(\omega) \\ c_1 \sin(\omega) & c_2 \cos(\omega) \end{bmatrix} \begin{bmatrix} x_t^{(i)} \\ y_t^{(i)} \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (10)$$

where Δx and Δy are translation parameters of x and y coordinates, c_1 and c_2 are the scaling factors, and ω is the rotation angle. Since the mapping above is from integer value to real value, most pixels in the output feature map \mathbf{F} cannot find their exact counterparts in the input feature map O_E . Therefore, a bilinear sampling kernel is applied for calculating the values of these pixels.

2) *Bilinear Sampling:* Using bilinear sampling, the pixel value $F(x_t^{(i)}, y_t^{(i)})$ in the coordinates $(x_t^{(i)}, y_t^{(i)})$ of the output feature map \mathbf{F} will be calculated by weighting the nearest four neighbor pixels around $(x_s^{(i)}, y_s^{(i)})$ in the input feature map O_E , which is formulated as

$$F(x_t^{(i)}, y_t^{(i)}) = \sum_n^M \sum_m^M O_E(n, m) w(n, m, x_s^{(i)}, y_s^{(i)}) \quad (11)$$

where $w(n, m, x_s^{(i)}, y_s^{(i)})$ is the weight coefficient and $O_E(n, m)$ is the pixel value of the input feature map O_E in the coordinates (n, m) . The weights of the four nearest pixels calculated by the bilinear sampling are given as $w(n, m, x_s^{(i)}, y_s^{(i)}) = (1 - |x_s^{(i)} - n|)(1 - |y_s^{(i)} - m|)$. For the other pixels, the weights are zeros. Geometrically, the weights of these pixels are zeros where the coordinate distances away from $(x_s^{(i)}, y_s^{(i)})$ are larger than 1. The weight coefficient for any pixel in $(x_s^{(i)}, y_s^{(i)})$ can be consistently reformulated as

$$w(n, m, x_s^{(i)}, y_s^{(i)}) = \max(0, 1 - |x_s^{(i)} - n|) \max(0, 1 - |y_s^{(i)} - m|). \quad (12)$$

Using (12), a weight coefficient vector $\mathbf{S}_i \in \mathbb{R}^{1 \times M_1 M_2}$ for the i th pixel of the output feature map \mathbf{F} can be defined as

$$\mathbf{S}_i = [w(1, 1, x_s^{(i)}, y_s^{(i)}), w(1, 2, x_s^{(i)}, y_s^{(i)}), \dots, w(n, m, x_s^{(i)}, y_s^{(i)})]$$

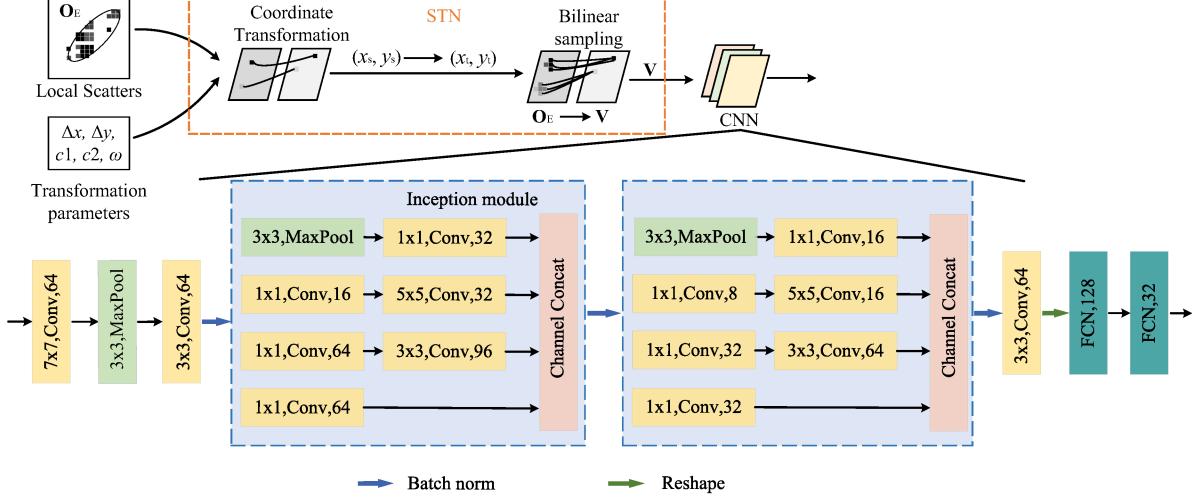


Figure 5. The CNN to learn the geometry structure of the local environment by exploiting the rotation invariance and scale invariance properties.

and therefore the output feature map \mathbf{F} can be consistently sampled as

$$\text{vec}(\mathbf{F}) = \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \\ \vdots \\ \mathbf{S}_{M_1 M_2} \end{bmatrix} \text{vec}(\mathbf{O}_E) \quad (13)$$

where $\text{vec}(\cdot)$ is the column vectorization operation. Taking \mathbf{O}_E with (8), the rotation and scaling properties can be applied to the virtual obstacle map \mathbf{H} , which helps reconstruct the geometry structure of the virtual environment. Using (13), it is also possible to downsample or oversample a feature map, as one can define the output size to be different to the input size, which can adjust the feature map in different scales flexibly for CNN layer.

3) Parameter Generation: The knowledge of how to transform each feature map is solved by the positions of the TX and RX. For position pair $\tilde{\mathbf{p}}$, consider the line segment connecting the ground position of the TX and RX for reference. The line segment is translated to the center of the feature map and rotated to the horizontal direction while keeping the relative locations of all scatters. The whole feature map will be scaled based on the length of the line segment. As a result, the parameters represent translation, scaling and rotation can be computed, which are formulated as

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} p_{t,1} + p_{r,1} - M_1 \\ M_1 \\ p_{t,2} + p_{r,2} - M_2 \\ M_2 \end{bmatrix}, \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \|\bar{\mathbf{p}}_t - \bar{\mathbf{p}}_r\|_2 \\ M_1 \\ \|\bar{\mathbf{p}}_t - \bar{\mathbf{p}}_r\|_2 \\ M_2 \end{bmatrix}$$

and $\omega = -\arctan((p_{t,2} - p_{r,2})/(p_{t,1} - p_{r,1}))$ if $p_{t,1} \geq p_{r,1}$; otherwise, $\omega = -\arctan((p_{t,2} - p_{r,2})/(p_{t,1} - p_{r,1})) - \pi$.

After the STN extracting the local scatter structures with rotation and scaling invariance, the CNN takes the transformed scatter feature map \mathbf{F} as input to learn the scattering channel attenuation, i.e., $g_s(\tilde{\mathbf{p}}, \mathbf{H}) = f_{\text{CNN}}(\mathbf{F})$. The architecture of the proposed CNN is plotted in Fig. 5. Two inception modules [34] are implemented in this work to process and aggregate

the scatter feature map at various scales, so that the next stage can abstract higher-level features from the different scales simultaneously for deeper scatter structure learning. Any other CNN architecture is available in our work. Therefore, the scattering branch combines the STN with CNN to learn the geometry structure of the local environment by exploiting the rotation and scaling invariance, which can improve the generalization ability of the proposed network.

E. Neural Network Training

Given a set of RSS measurements $\{y^{(i)}\}$ taken at position pair $\{\tilde{\mathbf{p}}^{(i)}\}$, the designed network is trained with the mean-squared error (MSE) loss function as

$$\underset{\Theta, \mathbf{H}}{\text{minimize}} \quad \frac{1}{N} \sum_{i=1}^N (y^{(i)} - g(\tilde{\mathbf{p}}^{(i)}, \Theta, \mathbf{H}))^2 \quad (14)$$

where Θ is a collection of all parameters for each component. We implement the designed network using the deep learning library Pytorch and the Adam optimizer is used.

To obtain an initial virtual obstacle map \mathbf{H} , a simplified clustering algorithm is applied to classify the training dataset into 2 categories (i.e. LOS and NLOS). A new dataset with category labels, i.e., $(\tilde{\mathbf{p}}^{(i)}, 0)$ for the LOS category and $(\tilde{\mathbf{p}}^{(i)}, 1)$ for the NLOS category, is constructed to train the virtual obstacle map \mathbf{H} by equation (9), as a binary classification task. The details of the initialization algorithm are summarized in Algorithm 2.

IV. SIMULATIONS

In this section, numerical results are presented to demonstrate the effectiveness of the proposed model for joint radio map and virtual obstacle map construction. To train and evaluate the proposed model, we collect the data of channel measurements from a reference 3D RT model.

In the simulation, we consider an air-to-ground communication scenario in an urban environment of Shanghai and Beijing, China. The region of interest is a square area of side 650 meters and the aerial TXs are randomly distributed with altitudes

Algorithm 2 Learning an initial virtual obstacle map \mathbf{H} .

```

1: Input training dataset  $(\tilde{\mathbf{p}}^{(i)}, y^{(i)})$ ,  $i = 0, 1, \dots, N$ 
2: Initialize categories  $\mathcal{C}_0$ ,  $\mathcal{C}_1$  and parameters  $\beta_1, \gamma_1, \beta_2, \gamma_2$ 
3: for  $n = 1$  to  $K$  do
4:   Empty  $\mathcal{C}_0$  and  $\mathcal{C}_1$ 
5:   for  $i = 0$  to  $N$  do
6:      $\hat{g}_1^{(i)} = \beta_1 + \gamma_1 \log_{10} \|\mathbf{p}_t^{(i)} - \mathbf{p}_r^{(i)}\|_2$ 
7:      $\hat{g}_2^{(i)} = \beta_2 + \gamma_2 \log_{10} \|\mathbf{p}_t^{(i)} - \mathbf{p}_r^{(i)}\|_2$ 
8:     if  $|y^{(i)} - \hat{g}_1^{(i)}| \leq |y^{(i)} - \hat{g}_2^{(i)}|$  then
9:        $\tilde{\mathbf{p}}^{(i)} \rightarrow \mathcal{C}_0$ 
10:    else
11:       $\tilde{\mathbf{p}}^{(i)} \rightarrow \mathcal{C}_1$ 
12:    end if
13:   end for
14:   Minimize $||y^{(i)} - \hat{g}_1^{(i)}||$ ,  $\forall \tilde{\mathbf{p}}^{(i)} \in \mathcal{C}_0$ 
15:   Minimize $||y^{(i)} - \hat{g}_2^{(i)}||$ ,  $\forall \tilde{\mathbf{p}}^{(i)} \in \mathcal{C}_1$ 
16:   Update  $\beta_1, \gamma_1, \beta_2, \gamma_2$ 
17: end for
18: if mean( $\hat{g}_1^{(i)}$ )  $\leq$  mean( $\hat{g}_2^{(i)}$ ) then
19:    $\mathcal{C}_0$  is the LOS category (label with 0)
20:    $\mathcal{C}_1$  is the NLOS category (label with 1)
21: else
22:    $\mathcal{C}_1$  is the LOS category;  $\mathcal{C}_0$  is the NLOS category
23: end if
24: Construct a new dataset  $\{\tilde{\mathbf{p}}^{(i)}, c^{(i)}\}$  where  $c^{(i)} \in \{0, 1\}$ 
25: Train  $\mathbf{H}$  by equation (9) using category labels as
   minimize $-\frac{1}{N} \sum_{i=1}^N c^{(i)} \log(I) + (1 - c^{(i)}) \log(1 - I)$ 

```

ranging from 50 to 200 meters. The heights of the ground RXs are fixed at 1.5 meters. The 3D geometry of the surroundings is obtained from the recent city map database. The channel measurements are generated using Remcom Wireless Insite, with up to 6 reflections and 1 diffraction. The waveform is chosen as sinusoidal signal at 5.9 GHz with 10MHz bandwidth and there is 3 dB noise to the RXs.

Based on the above setting, we construct three datasets from different environments. Dataset I and Dataset II are generated from the city areas in Shanghai and Beijing, China, respectively. There are 600k measurements at a TX altitude of 50 meters and nearly 190k measurements in the other altitudes for both datasets. Dataset III is referred to as the *diffraction dataset* that consists of only the NLOS propagation scenarios where the position pair $\tilde{\mathbf{p}}$ is blocked by at least one obstacle. In this dataset, the channel measurements are generated using Remcom Wireless Insite with the setting of up to 3 diffraction and no reflection which contains 300k measurements.

The proposed model is compared with the following baselines, MLP [19], RadioUNet [22], PMNet [23] and k -nearest neighbor (KNN), which are summarized as follows:

1) MLP [19]: This method is a data-fitting multilayer perceptron (MLP). The network partitions the samples into three clusters and fit the data in each cluster using NN with seven dense layers. It receives the TX-RX position pair $\tilde{\mathbf{p}}$ as input and returns the estimation of the path loss.

2) RadioUNet [22]: The network consists of two UNets.

Each consists of 9 encoder layers, and each layer consists of convolution, ReLU, and Maxpool layer. Similarly, there are 9 decoder layers, each consisting of transposed convolution followed by ReLU. The first UNet takes the city map and the TX position as inputs and outputs a coarse radio map. The second UNet takes the city map and the TX position plus the coarse radio map to generate the final radio map.

3) PMNet [23]: The network consists of 5 ResNet-based encoder layers and 7 ResNet-based decoder layers. Each layer applies several parallel atrous convolutions, Maxpooling, and ReLU. It takes the city map and the TX position map as input to generate the path loss radio map as well.

4) KNN: In KNN, the algorithm selects 6 measurement samples that are closest to $\tilde{\mathbf{p}}$ from the training set and then forms the neighbor set as $\mathcal{N}(\tilde{\mathbf{p}})$. The channel path loss at $\tilde{\mathbf{p}}$ is determined by $g(\tilde{\mathbf{p}}) = \mu^{-1} \sum_{i \in \mathcal{N}(\tilde{\mathbf{p}})} \omega(\tilde{\mathbf{p}}, \tilde{\mathbf{p}}^{(i)}) y^{(i)}$, where $\omega(\tilde{\mathbf{p}}, \tilde{\mathbf{p}}^{(i)}) = \exp[-\|\tilde{\mathbf{p}} - \tilde{\mathbf{p}}^{(i)}\|_2^2 (2s^2)]$ with a properly chosen parameter $s = 50$ meters and $\mu = \sum_{i \in \mathcal{N}(\tilde{\mathbf{p}})} \omega(\tilde{\mathbf{p}}, \tilde{\mathbf{p}}^{(i)})$ is a normalizing factor.

A. Performance of the Proposed Diffraction Component

In this experiment, we separate the diffraction component g_d from the proposed model and train it individually using the diffraction Dataset III. The diffraction dataset is split into three parts, where the first part is used for training, the second part is used for validation test and the last part is used for generalization test in the cases of single obstacle and multiple obstacles. The environment map is provided to train the proposed component g_d and a baseline CNN in this experiment. Here we consider a standard CNN architecture consisting of convolution, ReLU, and Maxpool layers with a comparable parameter quantity as the proposed model. Another classical large-scale distance fitting model is also considered as a baseline, i.e., $PL(d) = A + B \log_{10}(d)$, where A and B are the path loss exponents. The mean absolute error (MAE) between the estimated and the actual channel realization is considered as the performance metric.

As an illustrative example, a special case of diffraction propagation for the NLOS scenario with a single obstacle is presented at Fig. 6. In the first experiment, the positions of the obstacle and TX are fixed. The RX moves away from the virtual obstacle. The prediction curves of the path loss generated by the proposed model and the baselines are plotted in Fig. 6(b). The curve of the proposed model approximates the ground truth while the CNN and the distance fitting fails to capture the path loss variation in this NLOS scenario. It is surprisingly observed that the path loss is decreasing as the increase of the distance, which is contrary to the common large-scale distance model. It is because for the RX further away from the obstacle, the signal is less attenuated due to the diffraction effect.

In the second experiment, the positions of the TX and RX are fixed. The prediction curves in terms of the increasing obstacle heights are plotted in Fig. 6(c). With the increase of the obstacle height, the predicted path loss of both the proposed model and the CNN increase as the true one, but the proposed model is more accurate compared with the CNN.

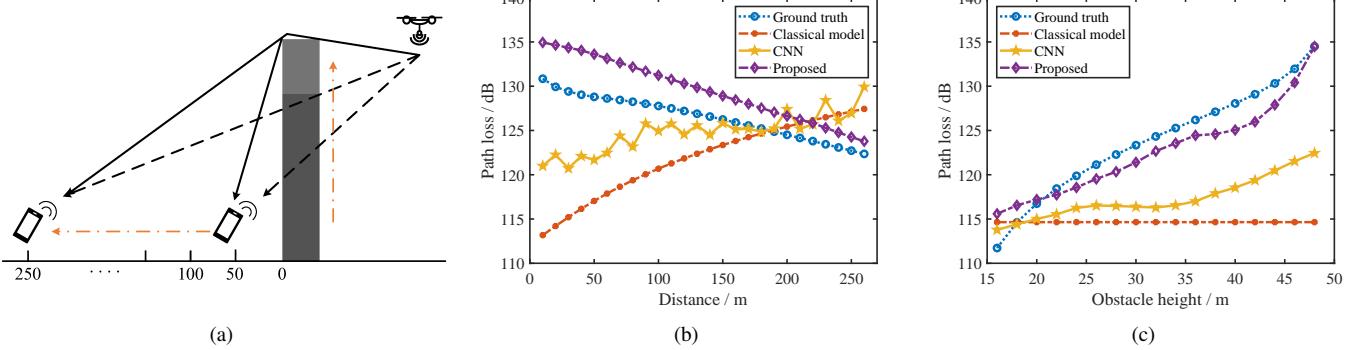


Figure 6. a) A special example of the diffraction propagation scenario with a single obstacle; b) With the RX moving away from the obstacle, the path loss is decreasing as the increase of the distance due to the diffraction effect; c) With the increase of the obstacle height, the diffraction effect diminishes resulting in larger channel path loss.

Table I
COMPARISON RESULTS IN THE DIFFRACTION SCENARIOS.

Scheme	Validation	Test	
		Single obstacle	Multiple obstacles
Classical model	8.98	9.25	9.22
CNN	7.82	5.55	8.23
Proposed	7.36	5.01	7.41

The classical model is insensitive to the change of the obstacle height and fails to capture the diffraction effect in the NLOS scenario. By contrast, the proposed model shows a great superiority in the NLOS scenario by exploring the geometry structures from the diffraction distances and angles.

In the third experiment, we evaluate the MAE of the proposed model, the CNN and the classical model for the path loss prediction in the Dataset III, which are summarized in Table I. It can be observed that the proposed model can reduce the MAE by 1.5 dB and 0.5 dB compared with the classical model and the CNN, respectively. In the test cases of the single obstacle or the multiple obstacles, the prediction performance of the classical model deteriorates. By contrast, the proposed model generates the best radio map performance, especially in single obstacle scenarios.

B. Performance Improvement by the Scattering Component

In this experiment, we evaluate the performance of the radio map construction of the proposed model under different components, i.e., the LOS component g_0 , the diffraction component g_d and the scattering component g_s , using Dataset I and Dataset II. We fix the altitude of the TX at 50 meter in this experiment and the accuracy of the radio map construction is assessed by the normalized mean absolute error (NMAE). The radio map accuracy comparison of the proposed full model ($g_0 + g_d + g_s$), the proposed partial model ($g_0 + g_d$) and the proposed partial model ($g_0 + g_s$) are presented in Table II. A performance increase is observed with the proposed full model ($g_0 + g_d + g_s$), as opposed to just utilizing the proposed partial model ($g_0 + g_d$) or ($g_0 + g_s$). The model without the scattering

Table II
COMPARISON OF RADIO MAP ACCURACY WITH DIFFERENT METHODS.

Scheme	NMAE	
	Dataset I	Dataset II
MLP [19]	0.0775	0.0765
RadioUNet [22]	0.0841	0.0831
PMNet [23]	0.0782	0.0754
KNN	0.0779	0.0772
Proposed (full model) $(g_0 + g_d + g_s)$	0.0697	0.0677
Proposed (partial model) $(g_0 + g_d)$	0.0916	0.0906
Proposed (partial model) $(g_0 + g_s)$	0.0812	0.0790

component g_s has the worst performance, where an decrease in the predictive NMAE is observed to be nearly 35%, which demonstrates the significance of the proposed scattering component g_s . As a whole, each proposed component can provide a performance improvement in the proposed model of radio map construction.

C. Reconstruction of Virtual Obstacle Environment

We further demonstrate the recovered geometry of the surroundings that represent the radio propagation environment captured by the proposed virtual obstacle map \mathbf{H} . We present the results of the radio map construction in Dataset I.

The real city map of the Dataset I and its corresponding 3D terrain model are presented in Fig. 7(a) and Fig. 7(b). Fig. 7(c) shows the virtual obstacle map \mathbf{H} generated by the proposed model without the scattering component g_s and Fig. 7(d) is the virtual obstacle map \mathbf{H} reconstructed by the proposed full components. It can be seen that the geometry and distribution of the virtual obstacles quite coincide with the true city map, which can demonstrate the ability of the proposed model to uncover the hidden environment information from the pure channel measurements. However, it should be noted

that the virtual obstacle map is not equivalent to the true city map even though there is a morphological similarity between them. Compared with Fig. 7(c) and Fig. 7(d), the virtual obstacle map ignoring scattering seems more discriminative in terms of the obstacle shapes and overall outline. With scattering, the virtual obstacle map gets vaguer and unsmooth even in some obvious LOS regions. It is because the virtual obstacle map serves as a geometry interpretation of the radio propagation environment, where the scattering effect incurs more complicated radio propagation characteristics and thus leads to the complexity of radio environment.

The proposed virtual obstacle map represents the radio propagation environment and provides blockage-aware ability of the channel condition. Such ability can be used to predict the LOS/NLOS channel conditions over the area of interest which has a great benefit in the radio map construction and applications, which will be further discussed in this paper.

D. Comparison to Other State-of-the-arts

We compare the performance of radio map construction of the proposed model and the baselines in Dataset I and Dataset II. For a fair comparison, we fix the altitude of the TX at 50 meters and NMAE is used to evaluate the performance.

A 2D slice of the path loss radio map constructed by the proposed model as well as the corresponding ground true radio map are presented in Fig. 8. It can be observed that the radio geometry characteristics in a dense urban area can be very complicated due to the distribution of a large number of obstacles. The radio geometry of the proposed radio map is very closed to the ground true one, which demonstrates that the proposed model is able to capture the radio geometry characteristics in a complex environment for radio map construction.

In Table II, we also present the NMAE of different design methods in terms of the radio map construction. It is observed that the proposed model outperforms all the baselines in both Dataset I and Dataset II, which offers a performance of 10%-18% accuracy improvements. In addition to accuracy improvements, the proposed model does not require any city map compared with the RadioUNet and PMNet, or a huge amounts of data to be stored compared with KNN, which shows a great superiority in training and implementation.

E. Transferability

1) *Generalize to Different Altitudes over 3D Space:* In this experiment, we demonstrate the proposed model is capable of learning a high-dimensional radio map with full spatial degrees of freedom where the TX can freely move over the 3D space as example. In this experiment, the altitudes of the aerial TXs are fixed in 50 meters in the training phase while the altitudes will range from 50 to 200 meters in the test phase. The proposed model trained in 50 meters is generalized to predict the radio map in different altitudes over the 3D space. We add two input channels to MLP model with TX and RX heights, making it feasible for high-dimensional radio map construction. The training data is from the Dataset I and the MAE is used as the performance metric as well.

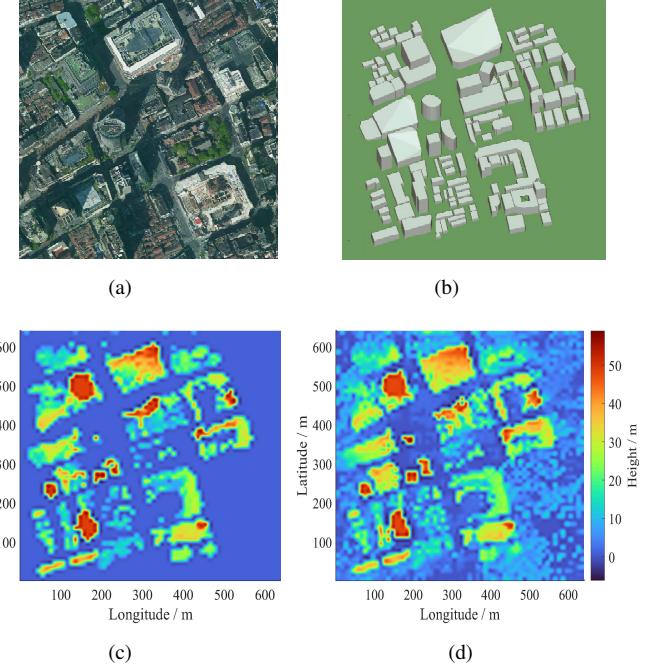


Figure 7. (a) The terrain map of an urban area of Nanjing Road in Shanghai; (b) The 3D terrain of the city map; (c) The virtual obstacle map generated by the proposed model without scattering component g_s ; (b) The virtual obstacle map generated by the proposed model with full components.

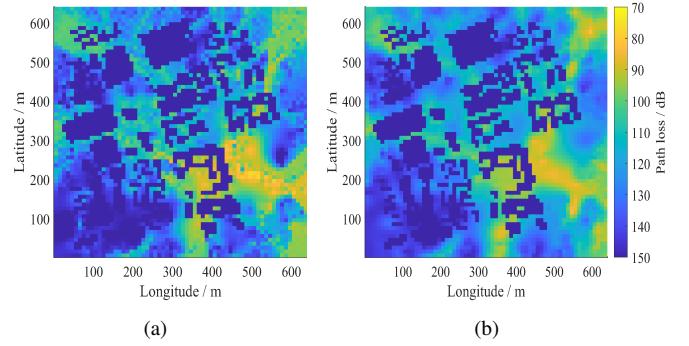


Figure 8. A radio map of the path loss constructed from Dataset I at a fixed TX position: a) The ground true radio map; b) The radio map generated by the proposed model.

The comparison results of the radio map construction in different TX altitudes in Dataset I are plotted in Fig. 9. It is obvious that the proposed model is superior to the baselines in all TX altitudes over the 3D space. With the increase of the TX altitudes, the MAE of the MLP, RadioUNet, PMNet and KNN increases greatly. They cannot be generalized to predict the radio maps beyond 50 meters. By contrast, there are merely 1 dB deviation among all altitudes in the proposed model without any additional measurement. Therefore, the proposed model can be efficiently generalized to the radio map construction with full spatial degrees of freedom without any additional data to fine-tune the parameters, which have superiority in the practical implementation.

2) *Transfer to a New Environment:* We further demonstrate the transfer ability of the proposed model to a new environment. The notion of *transfer learning* [35] is applied as

Table III
COMPARISON RESULTS OF RADIO MAP CONSTRUCTION WITH TRANSFER LEARNING.

Scheme	Transferred cases			Non-transferred cases		
	10%	30%	50%	10%	30%	50%
Epoch / NMAE						
RadioUNet [22]	90 / 0.120	90 / 0.092	65 / 0.086	100 / 0.202	90 / 0.201	85 / 0.198
PMNet [23]	100 / 0.107	97 / 0.082	92 / 0.075	100 / 0.132	100 / 0.107	91 / 0.106
Proposed	50 / 0.106	38 / 0.089	30 / 0.075	100 / 0.125	67 / 0.102	50 / 0.081

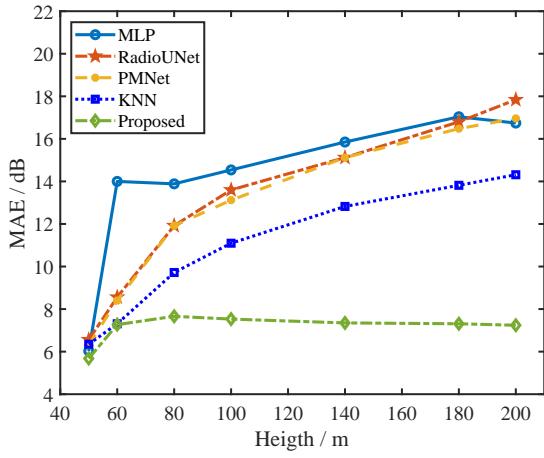


Figure 9. Comparison results of the generalization to different TX altitudes over the whole 3D space when the TX altitude in training phase is fixed.

follows: first learn an initial parameters in an environment and use the learned parameters as initialization for model learning in a new environment. The impact of such initialization is that the result of the model in the new environment will be generally closer to a better local optimum than if a random initialization were adopted.

The proposed model is trained in Dataset I and transferred to the Dataset II. In the case of transfer learning, the model is initialized with the pre-trained parameters in Dataset I and different data volumes with 10%, 30% and 50% of Dataset II are utilized to fine-tune the model, respectively. For comparison, we also set up a case of non-transfer learning, where the model is initialized randomly and trained with the same data volumes of Dataset II. The proposed model is compared with the RadioUNet and the PMNet in both cases of transfer learning and non-transfer learning.

The comparison results of the training epoch and NMAE with/without transfer learning are summarized in Table III. It is observed that the transfer learning does improve the radio map performance compared with the non-transfer learning in all cases. With 50% training data, the radio map accuracy of all schemes under transfer learning approximates the optimal radio map accuracy shown in Table. II. By contrast, the proposed method outperforms the baselines in all transferred/non-transferred cases except the case of the transfer learning with 30% training data. Compared with the non-transferred cases, the proposed transferred model converges faster with almost

half of reduction in training epoch, which shows a good transfer ability of the proposed model to a new environment.

F. Application of UAV-assisted Wireless Communication

In this example, the superiority of the proposed model with the virtual environment is demonstrated by using an application of UAV-assisted wireless communication.

Consider to deploy a low altitude UAV to establish LOS channels for two users in deep shadow in a dense urban area. Denote $\bar{g}(\mathbf{p}, \mathbf{q})$ as the large-scale channel gain that captures both the path loss and the shadowing between a terminal at position \mathbf{p} and a terminal at position \mathbf{q} . The objective of the UAV relay placement problem is formulated as

$$\begin{aligned} & \underset{\mathbf{p} \in \mathbb{R}^3}{\text{maximize}} \quad \min\{\bar{g}(\mathbf{p}, \mathbf{p}_1), \bar{g}(\mathbf{p}, \mathbf{p}_2)\} \\ & \text{subject to} \quad (\mathbf{p}, \mathbf{p}_i) \in \tilde{\mathcal{D}}_0, \quad \forall i \in \{1, 2\}. \end{aligned} \quad (15)$$

In the radio-map-based UAV relay placement problem, there is no any additional information of environment except the radio map constructed from the channel measurements. Because the classical radio-map-based methods supply the pure radio map over an area of interest, only the exhaustive 2D search or 3D search are available for UAV relay positioning. By contrast, the proposed model can utilize the virtual environment to provide blockage-aware ability of the channel condition which can be adapted to a more intelligent searching strategy. Based on the geometry property of the virtual obstacle model, if a UAV position is double-LOS for two users, then all UAV positions perpendicularly above it are double-LOS. The work [36] has demonstrated that it is empirically the best option to search on the middle perpendicular plane of the two users for a nearly-optimal UAV relay position. The UAV search strategy with the proposed model is summarized in two iterative steps:

- Search downward when UAV is at double-LOS position;
- Search along the circle with a fixed radius $r(\mathbf{p}')$ when UAV is at non-double-LOS position, where $r(\mathbf{p}') = \|\mathbf{p}' - \mathbf{o}\|_2$ is the radius from a point \mathbf{p}' on the perpendicular plane to the midpoint of two users \mathbf{o} .

The proposed UAV relay placement is compared with the exhaustive 2D search (with fixed height at 50 meters) and exhaustive 3D search using the same radio map generated by the proposed model. Another radio map is generated by MLP [19] offline. Such radio map also requires 3D search for the best UAV relay position, denoted as MLP_{3D} .

The average optimized channel gain of 100 user position pairs as well as the search distance generated by the proposed

Table IV
COMPARISON RESULTS OF THE UAV RELAY PLACEMENT.

Schemes	Channel Gain (dB)	Search Distance (km)
Proposed	-98.79	1.371
2D Search	-100.91	422.5
3D Search	-98.42	6760
MLP _{3D}	-99.83	6760

method, 2D search, 3D search and MLP_{3D} are summarized in Table IV. It is observed that the proposed method outperforms 2D search and MLP_{3D} in terms of channel performance and search complexity. The 3D search and MLP_{3D} utilize exhaustive search for the optimal UAV position with high expense, but the performance of MLP_{3D} is still worse than that of the proposed method due to the deficient accuracy of the radio map. It demonstrates that the proposed virtual environment is able to accelerate the UAV relay positioning. The performance gap in channel gain compared with 3D search is within 0.5 dB, but the proposed method shows great advantage in reducing more than 99% search distance, which further demonstrates the efficiency and superiority of the proposed method.

V. CONCLUSION

This paper developed a geometry model-assisted NN to explicitly exploit the geometry structure of the environment for joint 6D radio map and virtual obstacle map learning from the pure RSS data. In contrast to many existing machine learning approaches that lack of an environment model, a virtual obstacle model is developed to characterize the geometry structure of the virtual environment. To capture the diffraction, a transformer network is designed to mimic the computation structure of the Vogler expression and learn the diffraction mechanism from the key features of the diffraction distances and angles. In addition, instead of directly mapping an entire city map to a radio map, the proposed model only focuses on the geometry structure of the local area surrounding the TX-RX pair to describe the scattering. A CNN combined with a STN is designed to exploit the rotation invariance and scale invariance of the local geometry structures for data augmentation and scattering learning. Design examples show that the proposed model can reconstruct the 3D geometry of the virtual environment and provide 10%–18% accuracy improvement of the radio map construction. It also shows a notable generalizability to the whole 3D space and transferability to a new environment with 20% training data reduction and 50% training epoch reduction. The usefulness is further illustrated by an application of the radio-map-based UAV relay placement, where the proposed model can reduce 99% searching distance for UAV relay positioning with the guarantee of an equivalent channel performance.

APPENDIX A THE VOGLER EXPRESSION WITH RECURSIVE IMPLEMENTATION [31]

Consider the geometry of N knife-edge diffraction, where $\{\theta_i\}_{n=1}^N$ and $\{d_i\}_{n=1}^N$ are N diffraction angles and separation

distances between knife and edge, respectively. Then, the diffraction attenuation F_N for N knife-edges is determined by

$$F_N = \frac{1}{2^N} C_N e^{\sigma_N} \left(\frac{2}{\sqrt{\pi}} \right) \int_{\beta_1}^{\infty} \cdots \int_{\beta_N}^{\infty} e^{2F} \prod_{i=1}^N e^{-u_i^2} du_1 \cdots du_N \quad (16)$$

where

$$\begin{aligned} C_N &= \begin{cases} 1 & \text{for } N = 1 \\ \sqrt{\frac{(\sum_{i=1}^{N+1} d_i) \prod_{i=1}^N d_i}{\prod_{i=1}^N (d_i + d_{i+1})}} & \text{for } N \geq 2 \end{cases} \\ \sigma_N &= \sum_{i=1}^{N-1} \beta_i^2 \\ F &= \begin{cases} 0 & \text{for } N = 1 \\ \sum_{i=1}^{N-1} \psi_i & \text{for } N \geq 2 \end{cases} \\ \psi_i &= \phi_i(u_i - \beta_i)(u_{i+1} - \beta_{i+1}) \\ \beta_i &= \theta_i \left[\frac{j\pi d_i d_{i+1}}{\lambda(d_i + d_{i+1})} \right]^{1/2} \\ \phi_i &= \left[\frac{d_i d_{i+1}}{(d_i + d_{i+1})(d_{i+1} + d_{i+2})} \right]^{1/2} \end{aligned}$$

with λ denoting the wavelength and $j = \sqrt{-1}$ being an imaginary number.

Consider $N > 2$, by exploiting the fact that

$$\frac{2}{\sqrt{\pi}} \int_{\beta}^{\infty} (u - \beta)^m e^{-u^2} du = m! I(m, \beta) \quad (17)$$

where $m!$ refers to the factorial of m and $I(m, \beta)$ defines the repeated integrals of the complementary error function. A general solution for F_N is given as

$$F_N = \frac{1}{2^N} C_N e^{\sigma_N} \sum_{m=0}^{\infty} I_m \quad (18)$$

where

$$I_m = 2^m \sum_{m_1=0}^m \cdots \sum_{m_{N-2}=0}^{m_{N-3}} \prod_{i=1}^N \frac{(m_{i-1} - m_{i+1})!}{(m_i - m_{i+1})!} \phi_i^{m_{i-1} - m_i} I(n_i, \beta_i) \quad (19)$$

with, by using notion $m_0 = m$

$$n_i = \begin{cases} m_0 - m_1, & i = 1 \\ m_{i-2} - m_i, & 2 \leq i \leq N-1 \\ m_{N-2} - m_{N-1}, & i = N \end{cases} \quad (20)$$

Define

$$\begin{aligned} C(N-1, m_{N-2}, m_{N-3}) &= (m_{N-3})! \phi_{N-1}^{m_{N-2}} I(m_{N-3}, \beta_{N-1}) I(m_{N-2}, \beta_N). \end{aligned} \quad (21)$$

Using the notion

$$\begin{aligned} i &= m_{N-L}, j = m_{N-L-1}, k = m_{N-L-2} \\ 2 \leq L &\leq N-2, N \geq 4 \end{aligned} \quad (22)$$

and the recursive relationship

$$C(N-L, j, k) = \sum_{i=0}^j \frac{(k-i)!}{(j-i)!} \phi_{N-L}^{j-i} I(k-i, \beta_{N-L}) C(N-L+1, i, j) \quad (23)$$

the recursive computation of I_m is given as

$$I_m = 2^m \sum_{m_1=0}^{m_0} \phi_1^{m-m_1} I(m-m_1, \beta_1) C(2, m_1, m). \quad (24)$$

Therefore, the computation of the Vogler expression can be represented in a recursive manner, where the computation of N th diffraction F_N can depend on the computation of the $(N-1)$ th diffraction F_{N-1} . To emulate the computational structure inherent in the Vogler expression, the utilization of dependencies and sequence natures is paramount in the design of NN. The *transformer* architecture stands as an exemplar in this regard by harnessing the self-attention mechanism to discern and model intricate relationships across the sequence, which renders it particularly well-suited for tasks reliant on understanding intricate dependencies and sequences, aligning with the requisites of the computational structure of the Vogler expression.

REFERENCES

- [1] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Intelligent reflecting surface enhanced indoor robot path planning: A radio map-based approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4732–4747, 2021.
- [2] C. Wu, Z. Yang, and C. Xiao, "Automatic radio map adaptation for indoor localization using smartphones," *IEEE Trans. on Mobile Computing*, vol. 17, pp. 517–528, 2017.
- [3] Y. Zeng and X. Xu, "Toward environment-aware 6G communications via channel knowledge map," *IEEE Wireless Commun.*, vol. 28, no. 3, pp. 84–91, 2021.
- [4] X. Han, L. Xue, Y. Xu, and Z. Liu, "A two-phase transfer learning-based power spectrum maps reconstruction algorithm for underlay cognitive radio networks," *IEEE Access*, vol. 8, pp. 81 232–81 245, 2020.
- [5] S. Zhang and R. Zhang, "Radio map-based 3D path planning for cellular-connected UAV," *IEEE Trans. on Wireless Commun.*, vol. 20, no. 3, pp. 1975–1989, 2021.
- [6] X. Mo, Y. Huang, and J. Xu, "Radio-map-based robust positioning optimization for UAV-enabled wireless power transfer," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 179–183, 2020.
- [7] Y. Zeng, X. Xu, S. Jin, and R. Zhang, "Simultaneous navigation and radio mapping for cellular-connected UAV with deep reinforcement learning," *IEEE Trans. on Wireless Commun.*, vol. 20, no. 7, pp. 4205–4220, 2021.
- [8] Q. Hu, Y. Cai, A. Liu, G. Yu, and G. Y. Li, "Low-complexity joint resource allocation and trajectory design for UAV-aided relay networks with the segmented ray-tracing channel model," *IEEE Trans. on Wireless Commun.*, vol. 19, no. 9, pp. 6179–6195, 2020.
- [9] Q. Zhu, K. Mao, M. Song, X. Chen, B. Hua, W. Zhong, and X. Ye, "Map-based channel modeling and generation for U2V mmWave communication," *IEEE Trans. Veh. Tech.*, vol. 71, no. 8, pp. 8004–8015, 2022.
- [10] Y.-G. Lim, Y. J. Cho, M. S. Sim, Y. Kim, C.-B. Chae, and R. A. Valenzuela, "Map-based millimeter-wave channel models: An overview, data for B5G evaluation and machine learning," *IEEE Trans. Wireless Commun.*, vol. 27, no. 4, pp. 54–62, 2020.
- [11] N. Suga, R. Sasaki, M. Osawa, and T. Furukawa, "Ray tracing acceleration using total variation norm minimization for radio map simulation," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 522–526, 2021.
- [12] N. Suga, Y. Maeda, and K. Sato, "Indoor radio map construction via ray tracing with RGB-D sensor-based 3D reconstruction: Concept and experiments in WLAN systems," *IEEE Access*, vol. 11, pp. 24 863–24 874, 2023.
- [13] Y. Zhang and S. Wang, "K-Nearest neighbors Gaussian process regression for urban radio map reconstruction," *IEEE Commun. Lett.*, vol. 26, no. 12, pp. 3049–3053, 2022.
- [14] C. Phillips, M. Ton, D. Sicker, and D. Grunwald, "Practical radio environment mapping with geostatistics," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw.*, 2012, pp. 422–433.
- [15] H. Sun and J. Chen, "Propagation map reconstruction via interpolation assisted matrix completion," *IEEE Trans. Signal Process.*, vol. 70, pp. 6154–6169, 2022.
- [16] A. C. M. Austin and M. J. Neve, "Efficient field reconstruction using compressive sensing," *IEEE Trans. Antennas Propag.*, vol. 66, no. 3, pp. 1624–1627, 2018.
- [17] Y. Hu and R. Zhang, "A spatiotemporal approach for secure crowd-sourced radio environment map construction," *IEEE/ACM Trans. Netw.*, vol. 28, no. 4, pp. 1790–1803, 2020.
- [18] J. Chen, U. Yatnalli, and D. Gesbert, "Learning radio maps for UAV-aided wireless networks: A segmented regression approach," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, May 2017.
- [19] K. Saito, Y. Jin, C. Kang, J. ichi Takada, and J.-S. Leu, "Two-step path loss prediction by artificial neural network for wireless service area planning," *IEICE Commun. Exp.*, vol. 8, no. 12, pp. 611–616, 2019.
- [20] S. I. Popoola, A. Jefia, A. A. Atayero, O. Kingsley, N. Faruk, O. F. Oseni, and R. O. Abolade, "Determination of neural network parameters for path loss prediction in very high frequency wireless channel," *IEEE Access*, vol. 7, pp. 150 462–150 483, 2019.
- [21] L. Wu, D. He, B. Ai, J. Wang, H. Qi, K. Guan, and Z. Zhong, "Artificial neural network based path loss prediction for wireless communication network," *IEEE Access*, vol. 8, pp. 199 523–199 538, 2020.
- [22] R. Levie, Ā. Yapar, G. Kutyniok, and G. Caire, "RadioUNet: Fast radio map estimation with convolutional neural networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 4001–4015, 2021.
- [23] J.-H. Lee, O. G. Serbetci, D. P. Selvam, and A. F. Molisch, "PMNet: Robust pathloss map prediction via supervised learning," *arXiv:2211.10527*, 2022.
- [24] H. F. Ates, S. M. Hashir, T. Baykas, and B. K. Gunturk, "Path loss exponent and shadowing factor prediction from satellite images using deep learning," *IEEE Access*, vol. 7, pp. 101 366–101 375, 2019.
- [25] J. Thrane, D. Zibar, and H. L. Christiansen, "Model-aided deep learning method for path loss prediction in mobile communication systems at 2.6 GHz," *IEEE Access*, vol. 8, pp. 7925–7936, 2020.
- [26] O. Ahmadien, H. F. Ates, T. Baykas, and B. K. Gunturk, "Predicting path loss distribution of an area from satellite images using deep learning," *IEEE Access*, vol. 8, pp. 64 982–64 991, 2020.
- [27] Y. Teganya and D. Romero, "Deep completion autoencoders for radio map estimation," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1710–1724, 2022.
- [28] T. Hu, Y. Huang, J. Chen, Q. Wu, and Z. Gong, "3D radio map reconstruction based on generative adversarial networks under constrained aircraft trajectories," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8250–8255, 2023.
- [29] W. Liu and J. Chen, "UAV-aided radio map construction exploiting environment semantics," *IEEE Trans. on Wireless Commun.*, vol. 22, no. 9, pp. 6341–6355, 2023.
- [30] P. Zeng and J. Chen, "UAV-aided joint radio map and 3D environment reconstruction using deep learning approaches," in *Proc. IEEE Int. Conf. Commun.*, 2022, pp. 5341–5346.
- [31] V. D. Nguyen, H. Phan, A. Mansour, A. Coatanhay, and T. Marsault, "On the proof of recursive vogler algorithm for multiple knife-edge diffraction," *IEEE Trans. Antennas Propag.*, vol. 69, no. 6, pp. 3617–3622, 2021.
- [32] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2017, pp. 5998–6008.
- [33] M. Jaderberg, K. Simonyan, A. Zisserman, and k. kavukcuoglu, "Spatial transformer networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015.
- [34] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2015, pp. 1–9.
- [35] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [36] Y. Zheng and J. Chen, "Geography-aware optimal UAV 3D placement for LOS relaying: A geometry approach," *arXiv:2209.15161*, 2022.