

基于隐式神经表示与物理约束的室内声场重建方法

胡泽宁¹, 田益民^{2*}, 王金戈¹

¹北京印刷学院信息工程学院, 北京

²北京印刷学院基础部, 北京

收稿日期: 2025年12月26日; 录用日期: 2026年1月23日; 发布日期: 2026年2月2日

摘要

在实际室内声学应用中, 受测点数量、布置条件及测量成本等因素限制, 往往只能获得稀疏采样条件下的声场观测数据, 给声场空间分布的准确重建带来挑战。针对这一问题, 本文将室内稳态声场重建视为由空间坐标到声压幅值的函数拟合问题, 提出一种融合隐式神经表示与物理约束的室内声场重建方法。该方法以多层感知机为基础, 引入位置编码与正弦激活函数以增强网络对复杂空间振荡特征的表达能力, 并进一步结合Helmholtz方程构建物理一致性约束, 引导网络在稀疏采样条件下学习符合声波传播规律的声场空间分布。在不同采样率与频率条件下对所提方法进行了系统验证。实验结果表明, 与传统MLP及仅引入位置编码的基线模型相比, 该方法在稀疏采样条件下能够显著提升声场重建精度与稳定性, 有效抑制未观测区域中的非物理振荡现象。相关研究为有限观测条件下的室内声场重建提供了一种可行的建模思路。

关键词

室内声场, 稀疏采样, 物理约束, 隐式神经表示, 声场重建

A Method for Indoor Sound Field Reconstruction Based on Implicit Neural Representations and Physical Constraints

Zening Hu¹, Yimin Tian^{2*}, Jing Wang¹

¹School of Information Engineering, Beijing Institute of Graphic Communication, Beijing

²Foundation Department, Beijing Institute of Graphic Communication, Beijing

Received: December 26, 2025; accepted: January 23, 2026; published: February 2, 2026

*通讯作者。

文章引用: 胡泽宁, 田益民, 王金戈. 基于隐式神经表示与物理约束的室内声场重建方法[J]. 计算机科学与应用, 2026, 16(2): 90-101. DOI: [10.12677/csa.2026.162041](https://doi.org/10.12677/csa.2026.162041)

Abstract

In practical indoor acoustics applications, the number of measurement points is often limited by measurement conditions, sensor deployment, and acquisition costs, resulting in sparsely sampled sound field observations and posing challenges to accurate reconstruction of spatial sound field distributions. To address this problem, indoor steady-state sound field reconstruction is formulated as a function approximation task that maps spatial coordinates to sound pressure amplitudes. Based on this formulation, a sound field reconstruction method integrating implicit neural representations and physical constraints is proposed. The proposed method is built upon a multilayer perceptron architecture and incorporates positional encoding and sinusoidal activation functions to enhance the network's capability in representing complex spatial oscillatory patterns. Furthermore, a physics-consistent constraint derived from the Helmholtz equation is introduced to guide the network toward learning sound field distributions that conform to acoustic wave propagation characteristics under sparse sampling conditions. The proposed method was systematically validated under different sampling rates and frequency conditions. The results demonstrate that, compared with conventional multilayer perceptron models and baseline models employing only positional encoding, the proposed method achieves higher reconstruction accuracy and improved stability under sparse sampling, while effectively suppressing non-physical oscillations in unobserved regions. The proposed approach provides a feasible modeling framework for indoor sound field reconstruction under limited observation conditions.

Keywords

Indoor Sound Field, Sparse Sampling, Physical Constraints, Implicit Neural Representation, Sound Field Reconstruction

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在室内声学分析与工程应用中,准确获取声场在空间中的分布特征对于建筑声学设计、声环境评估及声场控制具有重要意义。然而,在实际测量过程中,受测点数量、布置条件及测量成本等因素限制,往往只能获得稀疏采样条件下的声场观测数据,使得声场连续空间分布的准确重建面临较大挑战。因此,在有限观测条件下实现稳定、可靠的室内声场重建,已成为当前声场建模与分析研究中的重要问题。

早期声场重建研究主要基于声学理论模型与解析展开方法展开,其典型代表包括平面波展开法、等效源法、球谐函数展开法、模态展开法以及各类空间插值方法等[1]-[3]。该类方法通常假设声场可由有限数量的基函数线性叠加表示,通过估计展开系数实现声场重建。在规则空间或采样较为密集的条件下,上述方法能够较好地恢复声场整体分布特征,且模型结构清晰、物理意义明确。然而,在复杂边界条件、非规则空间结构或稀疏采样条件下,其计算复杂度与建模难度显著增加,重建精度和稳定性也受到一定限制。

为降低对高采样密度与传感器数量的依赖,部分研究引入统计建模与机器学习思想对声场重建问题进行拓展。例如, Hu 等[4]在压缩感知理论框架下对声场进行稀疏表示估计,实现了欠采样条件下的声场重建; Caviedes-Nozal 等[5]将高斯过程回归方法引入声场建模,通过协方差函数刻画声场在空间中的相关性结构; Horiuchi 等[6]和 Ribeiro 等[7]进一步引入正则化与自适应核学习策略,以提升模型在复杂声场

条件下的稳定性与泛化能力。尽管上述方法在一定程度上改善了稀疏采样条件下的重建性能, 但其建模效果仍依赖于核函数选取及模型规模设置, 在高维空间或复杂声场结构中仍存在一定局限。

近年来, 随着深度学习技术的发展, 基于神经网络的声场重建方法逐渐受到关注。Lluis 等[8]基于 U-Net 结构对封闭空间内的声场分布进行了重建探索; Kristoffersen 等[9]在此基础上实现了声场幅值与相位的联合建模; 随后, Pezzoli 等[10]、刁威宇等[11]和 Fernandez-Grande 等[12]分别从卷积神经网络、生成对抗网络及多模型融合等角度对声场重建方法进行了拓展。Karakonstantis 和 Fernandez-Grande [13]进一步结合物理建模思想, 将平面波模型引入 GAN 训练过程, 以增强模型对声学先验的利用能力。上述研究表明, 神经网络在声场重建任务中具有较强的表示能力和建模灵活性, 但纯数据驱动方法在稀疏采样条件下仍可能出现预测不稳定及物理一致性不足的问题。

为提升模型的物理合理性与泛化性能, 研究者开始在神经网络中引入物理约束思想。Karakonstantis 等[14]基于物理信息引导的深度学习框架, 对有限测量条件下的房间脉冲响应重建问题进行了系统研究; Li 等[15]进一步将该类方法拓展至三维双耳房间脉冲响应建模任务。除基于波动方程的约束方式外, Damiano 等[16]提出了引入边界积分算子的物理信息神经网络方法; Chen 等[17]针对频域声场重建问题, 提出了融合多源物理信息的神经网络模型, 在低频及多频组合场景下表现出较好的泛化能力。然而, 现有研究多集中于时域脉冲响应或特定频段声场建模, 对于稳态室内声场在稀疏采样条件下的系统研究仍相对有限。

基于上述研究背景, 本文将室内稳态声场重建问题建模为由空间坐标到声压幅值的函数拟合问题, 提出一种融合隐式神经表示与物理约束的室内声场重建方法。该方法以多层感知机为基础, 引入位置编码与正弦激活函数以增强网络对复杂空间振荡结构的表达能力, 并结合 Helmholtz 方程构建物理一致性约束, 引导模型在稀疏采样条件下学习符合声波传播规律的声场空间分布特征。通过在不同采样率与频率条件下的数值实验, 对所提方法的重建性能与稳定性进行了系统验证。

2. 原理

2.1. 室内声场重建问题建模

在稳态条件下, 室内声场在空间中的分布可视为由空间坐标决定的连续函数。针对有限测量条件下的声场重建问题, 本文将室内稳态声场重建建模为一个函数拟合问题, 即通过有限采样点处的声压观测值, 学习空间坐标到声压幅值之间的非线性映射关系。具体而言, 设二维空间中任意位置的坐标为 $X = (x, y)$ 。对应的稳态声压幅值为 $p(x)$, 声场重建的目标是在有限采样点条件下, 构建近似函数 $\hat{p}(x)$, 以恢复整个空间内声压幅值的连续分布。在稀疏采样条件下, 传统插值或低阶函数拟合方法难以有效刻画声场中普遍存在的多尺度振荡结构, 容易在未观测区域产生过度平滑或非物理振荡现象。因此, 需要引入具有更强表达能力的建模方法, 以提升声场空间结构的重建精度与稳定性。

2.2. 基于隐式神经表示的声场建模

为增强模型对复杂空间振荡结构的表达能力, 本文采用隐式神经表示(Implicit Neural Representation, INR)对室内声场进行建模。隐式神经表示通过神经网络直接学习空间坐标与物理量之间的映射关系, 能够以连续形式表达高维空间函数, 适用于声场等连续物理场的建模任务。

本文以多层感知机(MLP)作为基础网络结构, 将空间坐标作为网络输入, 声压幅值作为网络输出。为克服传统 MLP 在表达高频空间变化特征方面的不足, 在网络输入端引入位置编码机制, 对原始空间坐标进行高维特征映射。位置编码通过引入多尺度正弦与余弦函数, 将低维坐标映射至包含丰富频率信息的高维特征空间, 从而增强网络对声场中驻波结构及局部细节变化的建模能力, 其基本形式可表示为:

$$\gamma(x) = [\sin(2^0 \pi x), \cos(2^0 \pi x), \dots, \sin(2^{L-1} \pi x), \cos(2^{L-1} \pi x)] \quad (1)$$

在此基础上, 网络隐含层采用正弦激活函数, 以进一步提升模型对周期性与振荡性空间结构的表达能力。与常规非线性激活函数相比, 正弦激活函数在表示高频变化特征时具有更好的连续性与稳定性, 有助于隐式神经表示对复杂声场结构的拟合。网络整体映射关系和第 i 层的非线性映射形式可表示为:

$$\begin{aligned}\Phi(x) &= W_n(\phi_{n-1} \circ \phi_{n-2} \circ \cdots \circ \phi_0)(x) + b_n, \\ \phi_i(x_i) &= \sin(W_i x_i + b_i)\end{aligned}\quad (2)$$

2.3. 物理约束的引入

尽管基于隐式神经表示的模型具有较强的函数逼近能力, 但在稀疏采样条件下, 仅依赖数据拟合仍可能导致模型在未观测区域出现不稳定预测或非物理振荡。为提升模型的物理一致性与泛化能力, 本文进一步在训练过程中引入物理约束, 对网络输出的空间变化特性进行引导。

在理想均匀介质条件下, 室内稳态声场的空间分布满足 Helmholtz 方程:

$$\nabla^2 p(x) + k^2 p(x) = 0 \quad (3)$$

其中 k 为波数, 与声源频率及介质参数相关。基于上述物理背景, 本文在不显式求解声学控制方程的前提下, 将 Helmholtz 方程作为物理一致性约束, 引入网络训练过程。利用自动微分技术对网络输出进行空间二阶导数计算, 构建物理约束损失项, 用以刻画网络预测声场与理想声波传播规律之间的偏离程度。该约束以“弱约束”的形式参与模型训练, 与基于观测数据的拟合损失共同构成优化目标, 从而在保证数据拟合精度的同时, 引导模型学习符合声学传播特性的空间分布结构。

综合数据拟合约束与物理一致性约束, 本文构建的总体损失函数可表示为:

$$L = L_{\text{data}} + \lambda L_{\text{phys}} \quad (4)$$

其中 L_{data} 表示基于观测点声压幅值的均方误差损失, L_{phys} 表示基于 Helmholtz 方程残差构建的物理约束损失项, λ 为权重系数, 用于平衡数据拟合精度与物理一致性之间的关系。通过该联合优化策略, 模型能够在稀疏采样条件下实现对室内声场空间分布的稳定重建。

3. 算法设计

3.1. 总体框架

针对稀疏采样条件下室内声场重建中存在的空间振荡结构难以刻画及未观测区域预测不稳定等问题, 本文构建了一种融合隐式神经表示与物理一致性约束的声场重建算法, 其总体技术路线如图 1 所示。该方法以连续函数建模为核心, 通过数据驱动建模与物理先验约束的协同作用, 实现对室内稳态声场空间分布的稳定重建。

在算法框架中, 首先将声场重建问题表述为由空间坐标到声压幅值的函数映射关系。网络输入为二维空间坐标 (x, y) (及对应频率参数), 输出为相应位置处的稳态声压幅值。为增强模型对复杂空间结构的表征能力, 在输入端引入多尺度位置编码机制, 将低维空间坐标映射至包含丰富频率成分的高维特征空间, 从而为后续网络学习提供更充分的空间频率信息。

在特征表示层, 采用以多层感知机为主体的隐式神经表示模型对编码后的输入特征进行建模。网络隐含层引入正弦激活函数, 使模型在表达周期性与振荡性空间结构方面具备更好的适应性, 有利于刻画室内声场中普遍存在的驻波特征与局部高频变化。通过该隐式表示方式, 声场空间分布可被视为由网络参数隐式定义的连续函数, 从而避免传统网格化建模带来的离散误差。在模型训练阶段, 除基于观测点声压数据的拟合约束外, 算法进一步引入物理一致性约束, 以增强模型预测结果的物理合理性。具体而言, 通过对网络输出进行自动微分计算, 引入与 Helmholtz 方程相关的残差约束, 对声压场的空间变化特

性进行引导。该物理约束以正则化项的形式参与网络优化，在不显式求解控制方程的前提下，引导模型在稀疏采样条件下学习符合声波传播规律的声场空间分布。

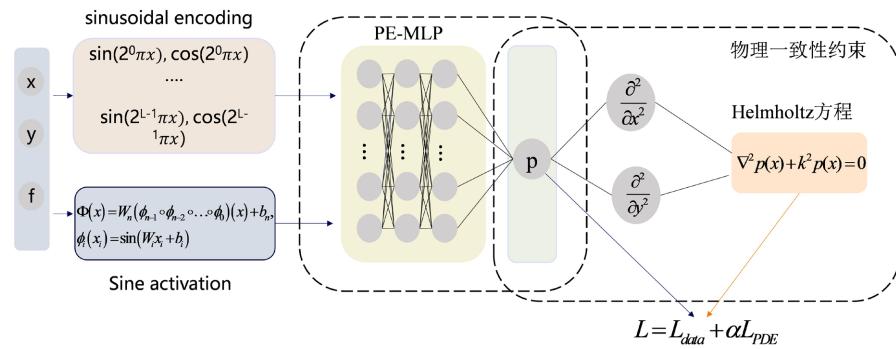


Figure 1. Overall technology roadmap
图 1. 总体技术路线图

综上，所提出的算法通过“位置编码 - 隐式神经表示 - 物理一致性约束”三者的协同设计，在保持模型结构简洁性的同时兼顾了数据驱动方法的灵活性与物理模型的约束能力，为稀疏采样条件下的室内声场重建提供了一种稳定有效的实现路径。

3.2. 数据集构建与采样率设置

本文所用室内声场数据集基于 k-Wave 数值仿真工具生成，用于为神经网络模型提供稳态声场的监督信号。通过对典型室内空间进行声学数值模拟，房间尺寸及参数，如表 1 所示，获得目标频率下声压幅值在空间中的分布结果，并作为声场重建任务的参考解。

Table 1. Classroom size and reverberation time

表 1. 教室尺寸与混响时间

房间尺寸(m ³)	混响时间——T ₆₀ (s)					
	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
9.6 × 7.7 × 3.1 m ³	1.70	1.79	2.13	1.79	1.80	1.57

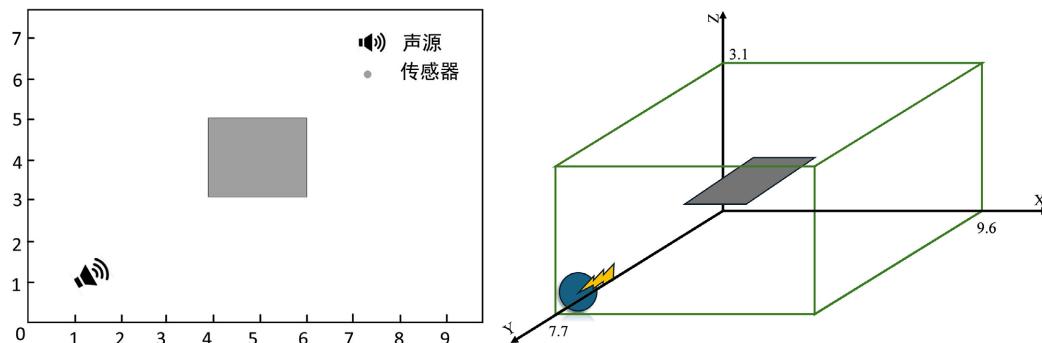


Figure 2. Sound source and sensor layout scheme
图 2. 声源及传感器布置方案

在数据生成过程中，首先构建室内空间的几何模型，并设置声源与接收平面位置，对目标频率下的稳态声场进行数值仿真。图 2 给出了声源、接收区域及观测平面的空间布置示意，其中声场重建任务限

定在选定的二维接收平面内进行。该设置用于明确声场数据的空间来源及采样区域范围，为后续重建任务提供统一的空间参考。

在获得完整声场分布后，对接收平面内的声压幅值进行规则网格化采样，形成连续声场参考结果。随后，从完整网格点集中随机选取部分空间位置作为观测点，构建不同采样密度下的稀疏声场观测数据，用以模拟实际测量条件下有限传感器布置所获得的声场信息。未被选取为观测点的空间位置仅用于模型测试阶段的整体重建效果评估。在具体实验中，选取 125 Hz、250 Hz 和 500 Hz 三个典型频率作为研究对象。针对每个频率分别生成对应的室内稳态声场数据，并在相同采样策略下构建训练与测试样本，用以分析所提方法在不同频率条件下的重建性能。

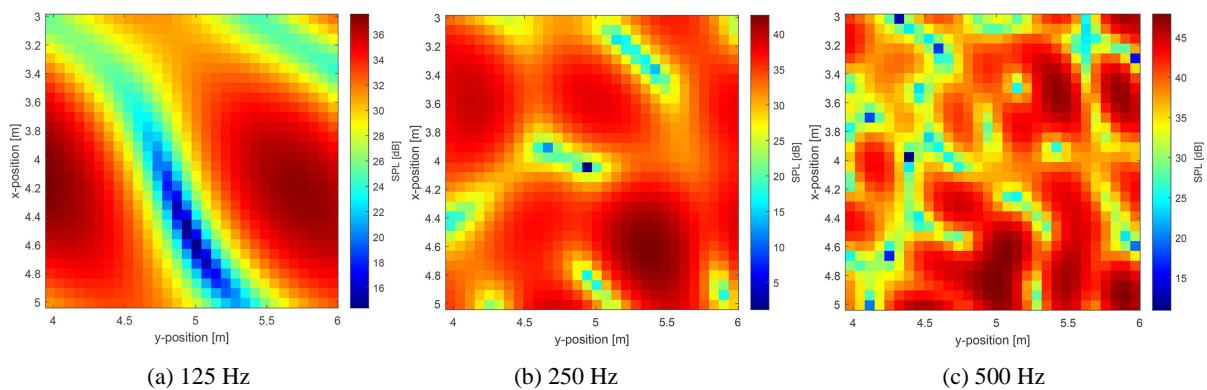


Figure 3. Steady-state sound pressure level distribution diagram in the mid-to-low frequency band

图 3. 中低频段稳态声压级分布图

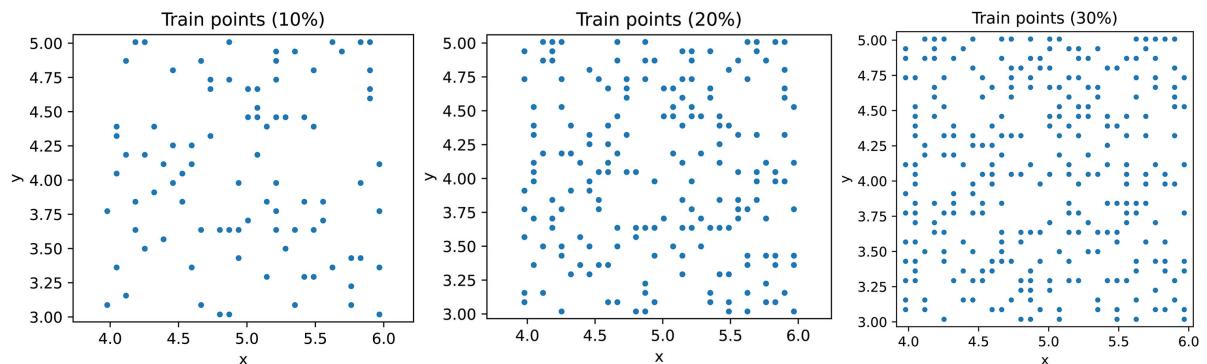


Figure 4. Training data sampling point diagram

图 4. 训练数据采样点示意图

图 3 给出了不同频率条件下的室内稳态声场幅值分布示例。本章中声场结果主要以声压级形式进行展示与分析，以便直观反映不同频率条件下室内声场的空间分布特征及其物理意义。可以看出，随着频率的升高，声场空间分布由相对平滑的低频模态结构逐渐演变为更为复杂的干涉与振荡形态，空间变化特征明显增强。上述声场数据为后续基于隐式神经表示的建模与稀疏采样条件下的重建性能评估提供了统一的数据基础。在后续实验中，所有重建模型均在相同的声场数据与采样策略下进行训练与测试，以保证实验结果的可比性。

在室内声场重建任务中，观测点数量及其空间分布对模型重建性能具有直接影响。为模拟实际测量条件下有限传感器布置所获得的声场观测数据，并系统评估所提方法在不同稀疏采样条件下的重建能力，本文在完整声场数据的基础上构建不同采样率的稀疏观测数据集。

设接收平面内完整声场网格点总数为 N , 本文在接收平面上采用 30×30 的规则网格进行采样, 因此 $N = 900$ 。为模拟实际工程中“有限麦克风数量”约束, 并提升稀疏度定义的可解释性, 本文以传感器绝对数量 M 来表征采样稀疏度, 并在括号中给出其相对于全域网格点的比例 $\rho = M/N$ 。具体而言, 本文设置 $M \in \{90, 180, 270\}$, 分别对应约 10%、20% 和 30% 采样, 如图 4 所示。对应的训练样本数量为 $N_{\text{train}} = M$ 。在数据划分过程中, 所有模型在相同采样率条件下采用一致的训练与测试划分方式, 以保证不同方法之间对比的公平性。上述设置用于分析模型在不同稀疏采样条件下的重建性能变化, 并验证所提方法在有限观测条件下的稳定性与适应性。

3.3. 网络训练与实现过程

在实验实现中, 本文采用基于隐式神经表示的多层感知机结构对室内声场进行建模。网络输入为空间坐标 (x, y) 及对应频率参数, 输出为预测的声压幅值。网络由 3 个隐含层构成, 每层包含 64 个神经元; 同时对更大网络规模(如每层 128 个神经元)进行了对比实验, 结果表明在当前稀疏采样与物理约束设置下, 继续增大网络规模未能带来稳定的性能提升, 反而在部分频率条件下出现轻微退化。因此, 在综合重建精度与模型复杂度后, 本文选取 64 作为主要网络配置。

在网络训练阶段, 本文采用基于梯度下降的优化策略对模型参数进行更新。训练过程中同时考虑基于观测数据的拟合误差与物理一致性约束项, 以引导网络在拟合有限观测数据的同时, 保持合理的空间结构特征。网络采用多层感知机结构, 输入为空间坐标及对应频率参数, 输出为预测的声压幅值。模型参数通过 Adam 优化算法进行更新, 学习率设为 1×10^{-4} , 最大训练轮数为 4000。数据拟合项采用均方误差作为损失函数, 同时引入基于 Helmholtz 方程残差的物理一致性约束作为正则化项。物理约束损失项的权重系数设为 $\lambda = 1 \times 10^{-4}$, 用于平衡数据拟合精度与物理一致性之间的关系。声波数由 $k = 2\pi f/c_0$ 给出, 其中声速取 343 m/s。在训练过程中, 为避免梯度爆炸对模型收敛造成不利影响, 对梯度范数进行裁剪, 裁剪阈值设为 1.0。此外, 训练过程中以验证集上的数据拟合误差作为模型选择依据, 保存验证误差最小的网络参数, 并在训练结束后加载该最优模型用于测试集评估及全场声场重建。

通过上述训练策略设计, 模型能够在有限采样条件下实现稳定收敛, 并逐步学习室内声场的空间分布特征, 为后续重建性能评估提供可靠基础。

4. 实验结果与分析

4.1. 实验设置

本文所有实验均在统一的计算环境下完成, 实验环境配置如表 2 所示, 后续实验结果均在该环境下获得。

Table 2. Experimental environment configuration
表 2. 实验环境配置

配置	名称	具体信息
硬件环境	CPU	Intel (R) Core (TM) i9-14900HX
	GPU	NVIDIA RTX 4060
软件环境	OS	Windows 11
	语言	python 3.12
	框架	PyTorch 2.7
	环境配置	CUDA 11.8

为定量评估不同模型在室内声场重建任务中的性能, 本文从预测精度与整体重建能力两个方面对模型进行评价。考虑到声场重建问题本质上属于连续空间函数逼近问题, 本文选取决定系数 R^2 以及基于 L2 范数定义的相对误差作为主要评价指标。

决定系数 R^2 用于衡量模型预测结果与参考声场之间的整体拟合程度, 其定义为:

$$R^2 = 1 - \frac{\sum_{i=1}^N (p_i - \hat{p}_i)^2}{\sum_{i=1}^N (p - \bar{p})^2} \quad (5)$$

其中, p_i 表示参考声场在第 i 个采样点处的声压幅值, \hat{p}_i 表示模型预测结果, \bar{p} 为参考声场的均值, N 为样本点数量。 R^2 越接近 1, 表明模型对声场空间分布的拟合效果越好。为进一步反映模型预测误差相对于真实声场幅值整体水平的大小, 本文引入基于 L2 范数定义的相对误差指标, 其表达式为:

$$\text{RelErr}_{L2} = \frac{\|\hat{\mathbf{P}} - \mathbf{P}\|_2}{\|\mathbf{P}\|_2} \quad (6)$$

其中, \mathbf{P} 与 $\hat{\mathbf{P}}$ 分别表示参考声场与模型预测声场在所有评价点处的声压幅值向量。该指标数值越小, 表示声场重建误差越低。在具体实验中, 上述评价指标均在测试集及完整声场重建结果上计算, 用以评估模型在未观测区域的泛化能力以及整体声场分布的重建效果。后续实验结果与对比分析均基于上述评价指标展开。

为在保证分析可比性的同时控制实验规模, 后续消融实验与频率对比实验统一选取 20% 采样率作为代表性实验设置; 同时, 通过在 10%、20% 和 30% 三种采样率条件下进行敏感性分析, 系统评估不同方法在稀疏采样条件下的性能变化趋势。

4.2. 消融实验

为分析所提方法中物理一致性约束对声场重建性能的影响, 本文在相同实验设置下设计消融实验, 以引入位置编码的 PE-MLP 模型作为数据驱动基线, 对比分析在是否引入物理约束条件下的重建结果差异。

在消融实验中, 本文选取 125 Hz、采样率为 20% 的实验结果进行可视化对比分析。首先, 图 5 给出了基础多层次感知机(MLP)在该条件下的重建结果。可以观察到, MLP 在一定程度上能够学习声压幅值的整体分布趋势, 但在空间振荡结构与局部细节刻画方面存在明显不足, 重建结果呈现出较强的平滑特征。误差分布结果显示, MLP 在声场中部及振荡较为剧烈的区域出现系统性偏差, 说明仅依赖二维空间坐标输入的基础 MLP 难以有效表达室内声场中的高频空间变化特征。该结果主要用于说明声场重建任务本身的难度, 而非作为后续消融分析的基线模型。

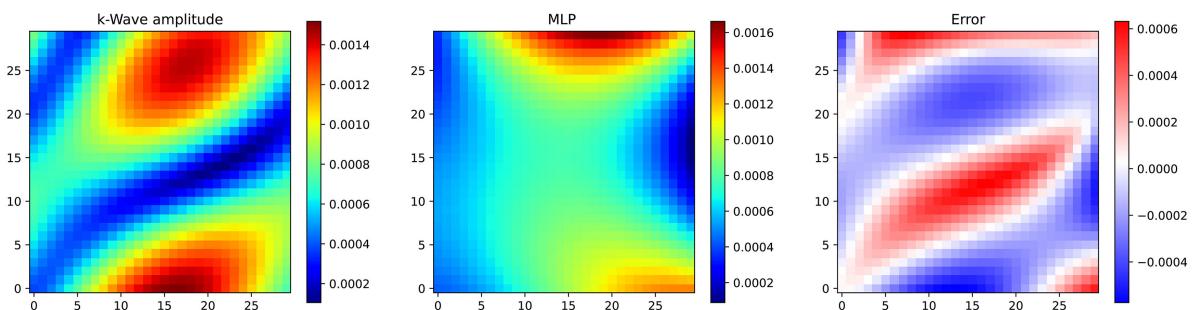


Figure 5. Comparison of MLP reconstruction results at 125 Hz and 20% sampling rate

图 5. 125 Hz 20% 采样率下的 MLP 重建效果对比

在此基础上, 本文引入位置编码机制构建 PE-MLP 基线模型, 其重建结果如图 6 所示。相比基础 MLP, PE-MLP 能够更清晰地恢复声场中的主要振荡模式, 整体结构与参考声场更加一致, 表明位置编码机制有效增强了模型对高频空间特征的表达能力。然而, 从误差分布仍可观察到, 在未观测区域内, PE-MLP 的预测结果存在一定程度的局部不稳定现象, 表现为非物理的局部波动。

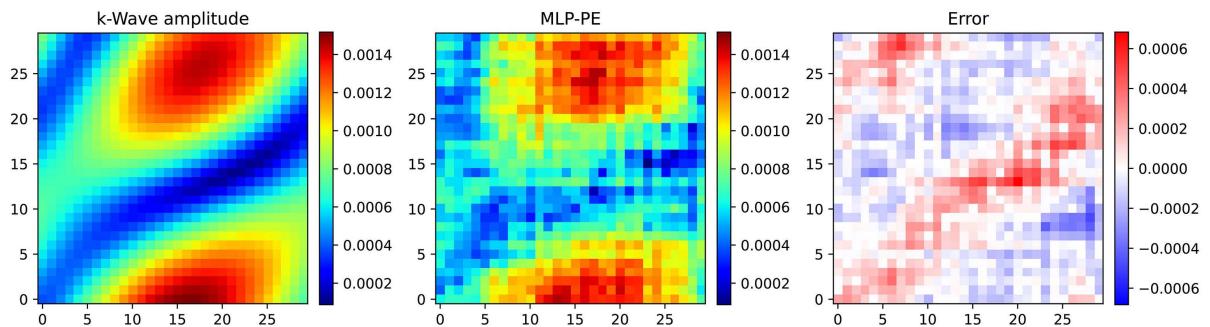


Figure 6. Comparison of PE-MLP reconstruction results at 125 Hz and 20% sampling rate

图 6. 125 Hz 20% 采样率下的 PE-MLP 重建效果对比

为进一步定量分析各关键模块对模型性能的影响, 本文在 PE-MLP 基线模型的基础上, 引入周期激活函数与物理一致性约束, 并通过消融实验对其作用进行评估。消融实验结果汇总于表 3, 其中 “test” 表示测试集指标, “full” 表示全场重建指标。

Table 3. Ablation experiment

表 3. 消融实验

PE-MLP	周期激活函数	物理启发约束	test RelErr	full relerr	test R^2	full R^2
√	✗	✗	0.741	0.793	21.24%	18.71%
√	√	✗	0.941	0.951	10.08%	9.08%
√	✗	√	0.937	0.949	10.43%	9.20%
√	√	√	0.978	0.982	6.10%	5.48%

从消融结果可以看出, 仅采用 PE-MLP 作为基线模型时, 其在测试集与全场重建中的相对误差较高, 且决定系数 R^2 明显偏低, 表明在稀疏采样条件下, 单纯依赖位置编码的模型仍难以稳定刻画声场的复杂空间振荡结构。

在 PE-MLP 基础上引入周期激活函数后, 模型在各项指标上均获得显著提升。测试集与全场重建的相对误差明显降低, R^2 值同步提高, 说明周期激活函数有助于增强网络对声场周期性与高频空间特征的表达能力, 从而改善整体拟合效果。

当仅在 PE-MLP 基线模型上引入物理一致性约束时, 模型同样表现出明显的性能提升。相比仅引入周期激活函数的情形, 物理约束在抑制预测结果中的非物理振荡方面表现出更强的稳定作用, 使得测试集与全场重建结果在误差与拟合度上均保持较好一致性。

当同时引入周期激活函数与物理一致性约束时, 模型在所有评价指标上均取得最优结果。测试集与全场重建的相对误差进一步降低, R^2 值接近 1, 表明模型不仅能够准确拟合观测数据, 还能在未观测区域中稳定恢复声场的整体空间分布。

综合消融实验结果可知, 周期激活函数主要提升模型对声场高频振荡结构的表达能力, 而物理一致性约束则有效增强模型在稀疏采样条件下的稳定性与物理合理性。二者协同作用, 共同构成了所提方法

性能提升的关键因素。

4.3. 与现有方法对比

为进一步验证所提方法在室内声场重建任务中的性能优势, 本文在相同数据集、采样策略与评价指标条件下, 将所提方法与多种典型声场重建方法进行对比分析。对比方法包括双线性插值(Bilinear)、平面波分解方法(PWD)、高斯过程回归(GPR)以及径向基函数神经网络(RBFNN)。对比结果汇总于表4, 涵盖不同频率(125 Hz、250 Hz、500 Hz)及不同采样率条件下的决定系数 R^2 与相对误差 RelErr_{L2}。其中, 双线性插值作为一种经典的空间插值方法, 不涉及模型训练过程, 其重建结果完全由已知观测点的空间插值关系确定。平面波分解方法采用基于 Tikhonov 正则化的最小二乘求解方式, 并针对不同采样稀疏度条件, 通过验证集对正则化参数及平面波基函数数量进行搜索与优化, 以确保其在稀疏观测条件下达到最优重建性能。高斯过程回归与径向基函数神经网络均基于监督学习框架实现。其中, GPR 采用常见的径向基核函数, 并通过最大化边缘似然的方式确定超参数; RBFNN 的隐含节点数量及相关参数通过验证集进行选取。上述方法的参数设置均遵循原方法的通用实现策略, 避免引入针对特定测试场景的额外调优, 从而保证对比结果的客观性。

Table 4. Comparison of reconstruction performance of different network models
表 4. 不同网络模型的重建性能对比

频率/Hz	采样率	Bilinear		PWD		GPR		RBFNN		本文方法	
		R^2	RelErr	R^2	RelErr	R^2	RelErr	R^2	RelErr	R^2	RelErr
125	10%	0.3351	33.51%	0.6601	23.96%	0.9544	8.78%	0.9215	11.52%	0.9518	9.02%
	20%	0.6000	25.86%	0.6755	23.41%	0.9726	6.80%	0.9424	9.87%	0.9822	5.48%
	30%	0.7490	20.59%	0.6761	23.39%	0.976	6.36%	0.9438	9.74%	0.9891	4.29%
250	10%	0.2728	35.47%	0.307	34.62%	0.5844	26.81%	-	-	0.8896	13.82%
	20%	0.5332	28.42%	0.4159	31.79%	0.6476	24.69%	0.629	25.33%	0.9451	9.75%
	30%	0.7196	22.02%	0.4373	31.20%	0.6778	23.61%	0.7191	22.04%	0.9932	3.43%
500	10%	0.2001	43.26%	0.0289	47.66%	0.2064	43.09%	-	-	0.2882	40.81%
	20%	0.4931	34.43%	0.151	44.57%	0.4383	36.25%	0.2095	43.00%	0.8150	20.80%
	30%	0.6319	29.34%	0.2145	42.86%	0.4904	34.53%	0.3844	37.95%	0.9170	13.93%

在 125 Hz 条件下, 各方法整体重建难度相对较低, 但在稀疏采样场景中仍存在明显性能差异。当采样率为 10% 时, Bilinear 与 PWD 方法的 R^2 分别仅为 0.3351 和 0.6601, 对应相对误差超过 23%, 表明传统插值方法与基于平面波假设的方法在稀疏采样条件下难以准确恢复声场结构。其中, PWD 方法基于有限数量平面波的线性叠加模型, 其重建效果依赖于声场中平面波成分的主导性, 在仅利用幅值信息且采样点较少的条件下, 模型表达能力受到一定限制。GPR 与 RBFNN 在该条件下表现较好, R^2 分别达到 0.9544 和 0.9215。所提方法在 10% 采样率下取得 $R^2 = 0.952$, 相对误差为 9.02%, 整体性能与 GPR 接近, 略优于 RBFNN。随着采样率提高至 20% 与 30%, 所提方法的 R^2 分别提升至 0.982 与 0.989, 相对误差降至 5.48% 与 4.29%, 在各对比方法中保持最优或接近最优水平。需要指出的是, 在低频且采样率较高条件下, 不同先进方法之间的性能差距逐渐缩小, 所提方法的优势主要体现在低采样率场景。

在 250 Hz 条件下, 声场空间结构复杂度明显提高, 各方法性能差异更加显著。Bilinear 与 PWD 方法

在不同采样率条件下均表现出较大的重建误差, 10%采样率下 R^2 分别仅为 0.2728 与 0.307。随着频率升高, 声场空间结构更加复杂, 局部干涉现象增强, 使得基于简化声学假设的平面波展开方法难以通过有限基函数有效表征声场细节, 尤其在稀疏观测条件下该问题更为突出。GPR 与 RBFNN 在中等采样率条件下具备一定建模能力, 但在低采样率条件下性能下降明显。相比之下, 所提方法在 10%、20% 和 30% 采样率条件下分别取得 $R^2=0.8896$ 、0.9451 和 0.9932, 相对误差对应为 13.82%、9.75% 和 3.43%。可以看出, 在中频条件下, 所提方法在不同采样率范围内均保持较稳定的重建性能, 尤其在 20% 及以上采样率条件下优势更为明显。但同时也应注意, 在 10% 采样率下, 其相对误差仍保持在 10% 以上, 说明在中频、极稀疏采样条件下, 模型对局部高频振荡结构的恢复仍存在一定困难。

在 500 Hz 高频条件下, 声场空间分布呈现出更为复杂的干涉与起伏结构, 对重建方法的表达能力与稳定性提出了更高要求。整体来看, 各对比方法在该频段下的重建性能均出现不同程度的下降, 且在稀疏采样条件下这一趋势尤为明显。传统插值方法与基于简化声学假设的模型在高频条件下面临明显困难, 其重建结果难以有效刻画声场中的高频空间振荡特征, 表现为整体平滑化严重、局部结构缺失。基于统计建模的 GPR 方法在高频条件下同样受到样本数量与空间分布限制, 模型稳定性随采样率降低而明显下降。需要指出的是, 500 Hz 已体现出高频条件下声场空间振荡增强与干涉结构复杂化等典型特征, 对重建方法的表达能力与稳定性提出更高要求。若进一步提升至 1000 Hz 及以上, 声场空间特征尺度将继续减小, 对采样密度与函数表示能力的需求显著提高; 同时, 在基于数值仿真数据的设定下, 更高频率通常需要更细的网格分辨率以避免数值色散影响, 从而带来更高的数据生成与训练成本。因此, 本文选取 500 Hz 作为高频代表工况, 其结果能够反映方法在高频条件下面临的主要挑战与趋势。

对于 RBFNN 方法, 在 250 Hz 和 500 Hz 的部分采样率条件下未给出有效结果。这主要是由于径向基函数网络在中高频声场建模中对中心点分布与核函数尺度参数高度敏感。在稀疏采样与高频叠加的情况下, RBFNN 容易出现数值不稳定或训练难以收敛的问题, 导致模型无法获得可靠的全场重建结果。该现象也反映出基于固定核函数展开的方法在复杂高频声场条件下的适用性存在一定局限。相比之下, 本文方法在高频条件下仍能够保持相对稳定的重建趋势。尽管在极低采样率下, 其对高频细节结构的恢复能力仍受到限制, 但随着采样率的提高, 模型对声场整体空间结构的刻画能力明显增强, 重建结果在连续性与一致性方面优于其他对比方法。这表明引入隐式神经表示与物理一致性约束, 有助于缓解高频声场中由稀疏采样带来的不稳定问题, 但并不能完全消除高频重建对观测密度的依赖。

综合不同频率与采样率条件下的对比结果可以看出, 传统插值方法与基于简化声学假设的模型在稀疏采样和高频条件下性能受限; GPR 与 RBFNN 在中低频、较高采样率条件下具有一定优势, 但稳定性随频率升高明显下降。然而, 表 4 中 500 Hz 且采样率为 10% 的结果也表明, 在高频且极稀疏观测条件下, 所提出方法的重建误差仍然较大, 说明仅依赖有限观测数据与物理一致性约束, 仍难以完全恢复复杂声场中的高频细节结构。这一现象从侧面反映出, 高频声场重建对传感器数量与空间分布具有更高要求, 若进一步提升至更高频率(如 1000 Hz 及以上), 该挑战将更加突出。这一局限性也反映在在实际测量条件下, 高频声场重建对传感器密度与空间分布仍具有较高要求。需要指出的是, 本文方法主要面向特定声学场景下的声场重建任务, 网络参数针对给定房间几何、边界条件及声源位置进行训练, 其目标是在有限观测点条件下高精度逼近特定声场函数, 而非学习可直接泛化至不同声源位置或不同场景的通用预测模型。从这一角度看, 该方法更侧重于场景内的函数重建能力, 而非跨场景泛化能力。在实际应用中, 这种建模方式意味着训练时间在声源位置或场景频繁变化时可能构成一定制约; 然而, 在场景参数相对固定、允许离线建模的应用中, 模型仅需训练一次即可完成声场函数的学习, 训练完成后声场重建仅涉及前向推理计算, 具有较高效率, 适用于对同一场景进行重复分析与评估的应用需求。

基金项目

本研究得到以下两个项目支持：杜云飞，北京印刷学院基础教育学院，北京 102600，项目：北京市教育委员会科技一般项目(KM202110015001)；北京印刷学院重点教学改革项目——工程认证背景下的工科数学教学改革对大学生创新思维与创业能力培养的研究与实践。

参考文献

- [1] Hu, D., Li, H., Hu, Y. and Fang, Y. (2018) Sound Field Reconstruction with Sparse Sampling and the Equivalent Source Method. *Mechanical Systems and Signal Processing*, **108**, 317-325. <https://doi.org/10.1016/j.ymssp.2018.02.031>
- [2] Talagala, D.S., Zhang, W. and Abhayapala, T.D. (2014) Efficient Multi-Channel Adaptive Room Compensation for Spatial Soundfield Reproduction Using a Modal Decomposition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, **22**, 1522-1532. <https://doi.org/10.1109/taslp.2014.2339195>
- [3] Wang, Y. and Chen, K. (2017) Sound Field Reconstruction within an Entire Cavity by Plane Wave Expansions Using a Spherical Microphone Array. *The Journal of the Acoustical Society of America*, **142**, 1858-1870. <https://doi.org/10.1121/1.5006057>
- [4] Hu, D., Liu, X., Xiao, Y. and Fang, Y. (2019) Fast Sparse Reconstruction of Sound Field via Bayesian Compressive Sensing. *Journal of Vibration and Acoustics*, **141**, Article ID: 041017. <https://doi.org/10.1115/1.4043239>
- [5] Caviedes-Nozal, D., Riis, N.A.B., Heuchel, F.M., Brunskog, J., Gerstoft, P. and Fernandez-Grande, E. (2021) Gaussian Processes for Sound Field Reconstruction. *The Journal of the Acoustical Society of America*, **149**, 1107-1119. <https://doi.org/10.1121/10.0003497>
- [6] Horiuchi, R., Koyama, S., Ribeiro, J.G.C., Ueno, N. and Saruwatari, H. (2021) Kernel Learning for Sound Field Estimation with L1 and L2 Regularizations. 2021 *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, 17-20 October 2021, 261-265. <https://doi.org/10.1109/waspaa52581.2021.9632731>
- [7] Ribeiro, J.G.C., Koyama, S. and Saruwatari, H. (2023) Kernel Interpolation of Acoustic Transfer Functions with Adaptive Kernel for Directed and Residual Reverberations. *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, 4-10 June 2023, 1-5. <https://doi.org/10.1109/icassp49357.2023.10095429>
- [8] Lluís, F., Martínez-Nuevo, P., Bo Møller, M. and Ewan Shepstone, S. (2020) Sound Field Reconstruction in Rooms: Inpainting Meets Super-resolution. *The Journal of the Acoustical Society of America*, **148**, 649-659. <https://doi.org/10.1121/10.0001687>
- [9] Kristoffersen, M.S., Møller, M.B., Martínez-Nuevo, P., et al. (2021) Deep Sound Field Reconstruction in Real Rooms: Introducing the Isobel Sound Field Dataset.
- [10] Pezzoli, M., Perini, D., Bernardini, A., Borra, F., Antonacci, F. and Sarti, A. (2022) Deep Prior Approach for Room Impulse Response Reconstruction. *Sensors*, **22**, Article No. 2710. <https://doi.org/10.3390/s22072710>
- [11] 刁威宇. 基于组合神经网络的混响水池声场重建技术研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工程大学, 2023.
- [12] Fernandez-Grande, E., Karakontantis, X., Caviedes-Nozal, D. and Gerstoft, P. (2023) Generative Models for Sound Field Reconstruction. *The Journal of the Acoustical Society of America*, **153**, 1179-1190. <https://doi.org/10.1121/10.0016896>
- [13] Karakontantis, X. and Fernandez-Grande, E. (2023) Generative Adversarial Networks with Physical Sound Field Priors. *The Journal of the Acoustical Society of America*, **154**, 1226-1238. <https://doi.org/10.1121/10.0020665>
- [14] Karakontantis, X., Caviedes-Nozal, D., Richard, A. and Fernandez-Grande, E. (2024) Room Impulse Response Reconstruction with Physics-Informed Deep Learning. *The Journal of the Acoustical Society of America*, **155**, 1048-1059. <https://doi.org/10.1121/10.0024750>
- [15] Li, Y., Wang, L. and Reiss, J. (2025) Binaural Room Impulse Responses Interpolation Using Physics-Informed Neural Networks in Three Dimensions. *51st German Annual Conference on Acoustics (DAGA)*, Copenhagen, 17-20 March 2025. <https://doi.org/10.71568/dasdaga2025.112>
- [16] Damiano, S. and Van Waterschoot, T. (2025) Sound Field Reconstruction Using Physics-Informed Boundary Integral Networks. *2025 33rd European Signal Processing Conference (EUSIPCO)*, Palermo, 8-12 September 2025, 76-80. <https://doi.org/10.23919/eusipco63237.2025.11226620>
- [17] Chen, F., Xiao, Y., Yu, L. and Jiang, L. (2026) Frequency-Domain Physics-Informed Neural Network for Accurate Reconstruction of 3D Acoustic Fields under Sparse and Multi-Frequency Measurements. *Neural Networks*, **197**, Article ID: 108476. <https://doi.org/10.1016/j.neunet.2025.108476>