

高速移动场景下基于自适应Transformer网络的MIMO-OFDM信道估计

廖希^① 何香妮^① 张哲^{*①②} 王洋^{①②}

^①(重庆邮电大学通信与信息工程学院 重庆 400065)

^②(中国人民解放军31401部队 长春 130012)

摘要: 高速移动场景下显著多普勒频移引发信道快速时变特性及子载波间干扰加剧, 导致MIMO-OFDM系统信道估计精度下降及鲁棒性不足, 本文提出一种基于特征线性调制(Feature-wise Linear Modulation, FiLM)的自适应Transformer信道估计模型—AdaFiT。首先, 采用可分离二维线性上采样获取全时频域信道估计, 并利用点卷积在通道维度对多天线复数信道的实部与虚部进行联合建模, 结合卷积特征增强模块提取多尺度时频特征; 然后, 设计信道自适应特征调制模块, 将信道参数作为先验条件信息, 基于FiLM机制生成特征层面的缩放系数与偏移系数, 仿射调制Transformer网络的块嵌入序列, 实现对信道统计变化的条件自适应; 最后, 通过二维可学习位置编码的Transformer编码器全局建模块级特征, 并结合残差重建结构融合局部与全局特征, 完成高精度信道重建。结果表明, 所提AdaFiT模型在CDL-C和CDL-A两种信道模型下均优于LS-Bilinear、LMMSE及现有自适应方法, 对不同信道模型具有较好的适应性。其中, 在CDL-C信道模型下的动态信噪比、多普勒频移及不同时延扩展条件下, 均方根误差最大可提升7 dB。

关键词: MIMO-OFDM系统; 信道估计; Transformer网络; 自适应特征调制; 特征线性调制

中图分类号: TN911.23

文献标识码: A

文章编号: 1009-5896(2026)00-0001-12

DOI: 10.11999/JEIT260075

CSTR: 32379.14.JEIT260075

1 引言

多输入多输出正交频分复用(Multiple-Input Multiple-Output Orthogonal Frequency Division Multiplexing, MIMO-OFDM)技术凭借其高频谱效率和抗多径衰落能力, 已成为第五代及未来无线通信系统的关键物理层技术^[1,2]。在MIMO-OFDM系统中, 精确的信道状态信息(Channel State Information, CSI)是实现可靠信号检测、高效波束赋形以及自适应资源分配的基础^[3,4], 其估计精度直接制约系统的整体性能。

传统信道估计算法通常基于信道在相干时间内保持准静态的假设, 在低速移动场景下是有效的。然而, 在高铁、无人机、车联网等高速移动场景下, 由收发端高速相对运动产生的显著多普勒频移, 使得信道不仅具有多径传播导致的频率选择性, 更呈现出强烈的时变特性, 形成复杂的时频双选择性衰落^[5,6]。这种快时变特性破坏了OFDM子载波间的正交性, 引致严重的载波间干扰^[7], 极大地增加了信道估计的复杂度与难度。面对这一挑战, 以最小二乘(Least Squares, LS)和线性最小均

方误差(Linear Minimum Mean Square Error, LMMSE)为代表的传统线性估计算法, 局限性日益凸显。LS算法通过最小化接收信号与估计信号之间的均方误差, 计算复杂度低, 但未利用信道的统计先验^[8], 在低信噪比环境下估计精度严重恶化。LMMSE算法虽可利用信道二阶统计特性实现优异的噪声抑制, 但在高动态环境下, 信道统计特征的快速时变且难以获取, 导致其性能严重恶化^[9,10]。为此, 设计一种能够适应高速移动环境下信道快时变特性, 兼具高精度与强鲁棒性的信道估计方法至关重要。

近年来, 深度学习凭借其强大的数据驱动非线性建模能力及对显式信道统计先验的弱依赖性, 为突破传统方法瓶颈提供了新范式。现有研究主要聚焦于信道时频特征的有效提取、关键特征的增强辨识以及模型对动态环境的自适应能力。早期工作主要利用卷积神经网络(Convolutional Neural Network, CNN)强大的局部特征提取能力来恢复信道矩阵^[11]。例如, 文献^[12]提出一种Dual CNN架构, 分别在空-频域和角-时域进行卷积操作, 利用不同域的互补稀疏性提升估计精度。文献^[13]进一步拓展至三维Dual-CNN, 以同时捕获空间-频率-时间三维相关性, 并引入元学习框架以提升模型对新环境的快速适应能力。然而, 标准CNN结构平等对待所有特征通道, 缺乏对关键信息的辨别能力, 在强噪声环境下易受干扰。为增强特征筛选能

收稿日期: 2026-01-21; 改回日期: 2026-07-13; 网络出版: 2026-07-24

*通信作者: 张哲 19628069@qq.com

基金项目: 重庆市自然科学基金(No.CSTB2025YITP-QCRC0045)

Foundation Item: Chongqing Natural Science Foundation (No.CSTB2025YITP-QCRC0045)

力, 研究者们将注意力机制引入CNN。文献[14]提出的AttRNet在残差块中嵌入注意力模块, 自适应地调整特征通道的权重, 有效抑制LS初始估计中的噪声成分。文献[15]引入压缩激励(Squeeze-and-Excitation, SE)通道注意力机制, 增强了残差网络的特征表达能力。但其依赖静态特征相关性分配注意力权重, 对动态信道变化的自适应能力仍有限。此外, CNN受局部感受野限制, 难以建模长距离依赖, 因此具备全局建模能力的Transformer架构被引入信道估计领域。文献[16]串行融合CNN与Transformer模型, 以兼顾局部细节与全局上下文建模, 但其性能提升较依赖于高信噪比数据的加权训练, 在低信噪比条件下抗噪鲁棒性有限。为提升模型对动态复杂信道的适应能力, 近期研究开始将信道统计特征(如信噪比、多普勒频移、时延扩展等)作为先验信息注入网络, 实现信道自适应估计。文献[17]提出的CE-ViT结合Vision Transformer处理时频图块, 并引入包含信道参数的特定标记Token, 在固定500 Hz多普勒频移的动态信噪比测试下获得约5 dB性能增益, 展现良好的鲁棒性。文献[18]提出的AdaFortiTran通过信道自适应模块将信道统计量编码为嵌入向量, 用以指导Transformer进行条件特征提取, 低信噪比和高移动速度下获得6 dB性能增益, 并在高多普勒频移、时延扩展变化等动态场景下保持稳定估计性能。然而, 上述研究大多局限于单天线OFDM场景, 难以直接应对高速移动MIMO-OFDM系统中因天线间耦合、空间维度扩展以及更复杂的时频干扰所带来的挑战。此外, 现有方法多采用标记拼接或嵌入相加等方式融合先验信息, 依赖网络自主学习复杂映射, 易产生计算冗余, 且难以精准调控中间特征分布。

针对上述问题, 本文提出一种面向高速移动MIMO-OFDM系统的自适应特征调制Transformer信道估计网络AdaFiT。该模型采用CNN与Transformer协同的混合架构, 充分利用CNN提取局部时频相关性与Transformer建立全局及天线间依赖关系的优势。不同于现有方法将信道参数编码后与网络特征进行融合, 本文引入基于特征线性调制(Feature-wise Linear Modulation, FiLM)的显式自适应模块[19]。该模块能够根据输入的信道物理参数(如信噪比、时延扩展和最大多普勒频移), 动态生成仿射变换参数, 对深层特征进行逐通道条件调制, 这种机制使模型能够敏锐感知并适应信道环境的变化。最后, 通过残差连接融合多尺度特征, 实现鲁棒的高精度信道重建。

2 系统模型

2.1 信号传输模型

本文考虑一个具有 N_{tx} 根发射天线和 N_{rx} 根接收天线的MIMO-OFDM系统。假设一个子帧中有 N_t 个OFDM符号和 N_f 个子载波, 由第 i 根天线发射的第 m 个OFDM符号向量表示为 $\mathbf{x}_{i,m} = [x_{i,m}(0), x_{i,m}(1), \dots, x_{i,m}(N_f - 1)]^T$, 其中 $x_{i,m}(n), n \in \{0, 1, \dots, N_f - 1\}$ 表示第 m 个OFDM符号在第 n 个子载波上的发射信号。接收天线 $j \in \{1, 2, \dots, N_{\text{rx}}\}$ 接收的OFDM向量表示为

$$\mathbf{y}_{j,m} = \sum_{i=1}^{N_{\text{tx}}} \mathbf{G}_{i,j,m} \cdot \mathcal{F}^{-1}\{\mathbf{x}_{i,m}\} + \mathbf{z}_{j,m}, 1 \leq j \leq N_{\text{rx}},$$

$$m = 0, 1, \dots, N_t - 1 \quad (1)$$

其中, $\mathbf{y}_{j,m} = [y_{j,m}(0), y_{j,m}(1), \dots, y_{j,m}(N_f - 1)]^T$ 表示时域接收序列, $\mathbf{G}_{i,j,m} \in \mathbb{C}^{N_f \times N_f}$ 表示第 i 根发射天线与第 j 根接收天线之间的时域信道矩阵, $\mathbf{z}_{j,m} \in \mathbb{C}^{N_f \times 1}$ 表示加性白噪声, $\mathcal{F}^{-1}\{\cdot\}$ 表示逆傅里叶变换运算。频域接收信号表示为

$$Y_{j,m}(n) = \sum_{i=1}^{N_{\text{tx}}} \left(H_{i,j,m}(n, n) x_{i,m}(n) + \sum_{k=0, k \neq n}^{N_f - 1} H_{i,j,m}(n, k) x_{i,m}(k) \right) + Z_{j,m}(n) \quad (2)$$

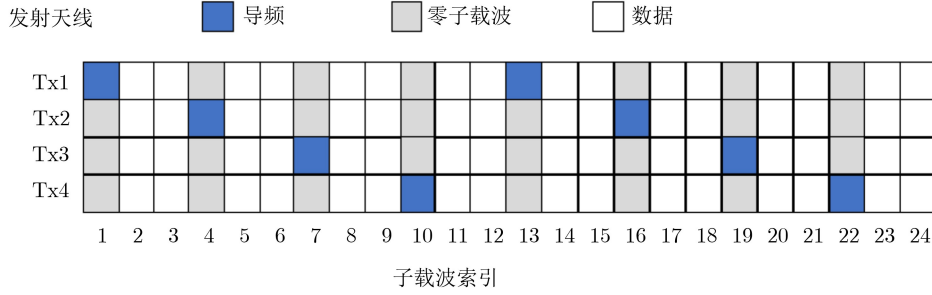
其中, $Y_{j,m}(n)$ 、 $Z_{j,m}(n)$ 分别表示 $\mathbf{Y}_{j,m} \in \mathbb{C}^{N_f \times 1}$ 、 $\mathbf{Z}_{j,m} \in \mathbb{C}^{N_f \times 1}$ 在第 n 个子载波上的频域分量, $\mathbf{Y}_{j,m}$ 和 $\mathbf{Z}_{j,m}$ 分别表示 $\mathbf{y}_{j,m}$ 和 $\mathbf{z}_{j,m}$ 的离散傅里叶变换, $\mathbf{H}_{i,j,m} \in \mathbb{C}^{N_f \times N_f}$ 表示频域信道响应矩阵, $H_{i,j,m}(n, n)$ 表示第 n 个子载波的频域信道增益, $H_{i,j,m}(n, k)$ 表示第 k 个子载波对第 n 个子载波的频域耦合系数。当 $k \neq n$ 时, 该项表征子载波间干扰。

2.2 导频模式与帧结构

本文基于5G NR标准设计帧结构与导频图案, 如图1所示。频域上, 单个OFDM符号内每根发射天线包含 N_p 个均匀分布的导频, 相邻导频间隔为 D , 导频索引记为 $k, 0 \leq k \leq N_p - 1$ 。为避免多天线传输时的导频干扰, 采用基于频分复用的正交导频设计。相对于基准天线($i=1$), 第 i 根天线的导频图案偏移 $\Delta_i = (i - 1) \times 4$ 个子载波。因此, 第 i 根天线的第 k 个导频所对应的子载波索引 $n_{i,k}$ 表示为

$$n_{i,k} = kD + \Delta_i, 1 \leq i \leq N_{\text{tx}} \quad (3)$$

时域上, 一个子帧配置2个OFDM符号用于传输导频, 因此每根发射天线的导频数量为 $2N_p$, 整个系统总导频开销为 $N_{p_total} = 2N_{\text{tx}}N_p$ 。

图 1 导频模式($N_t = 4, D = 12, N_p = 2$)

3 基于AdaFiT的MIMO-OFDM信道估计方法

3.1 网络架构

本文设计一种面向高速移动MIMO-OFDM系统的自适应特征调制Transformer信道估计网络,称为AdaFiT,如图2所示。该网络旨在从稀疏的导频观测中高精度地恢复出完整的信道状态信息。具体而言,网络输入主要包含两个部分:一是在导频位置处获取LS估计复数矩阵 $\hat{\mathbf{H}}_p^{LS}$;二是作为先验知识的信道统计物理参数集,包含信噪比 s 、时延扩展 τ_{ds} 以及最大多普勒频移 f_d 。网络输出为重建后的全维度信道矩阵 $\hat{\mathbf{H}}$ 。AdaFiT网络主要由六个核心子模块组成,依次为上采样子模块、特征增强子模块、块嵌入子模块、信道自适应特征调制子模块、Transformer编码器子模块以及信道重建子模块。

(1)上采样子模块:不同于文献[18]中将导频位置处的LS估计 $\hat{\mathbf{H}}_p^{LS}$ 展平成向量 $\hat{\mathbf{h}}_p^{LS}$ 的线性上采样方法,本文采用一种可分离的二维线性上采样方法。首先,将输入的 $\hat{\mathbf{H}}_p^{LS} \in \mathbb{C}^{N_{tx} \times N_{rx} \times N_{f,p} \times N_{t,p}}$ 分解为实部 $\text{Re}\{\hat{\mathbf{H}}_p^{LS}\} \in \mathbb{R}^{N_{tx} \times N_{rx} \times N_{f,p} \times N_{t,p}}$ 与虚部 $\text{Im}\{\hat{\mathbf{H}}_p^{LS}\} \in \mathbb{R}^{N_{tx} \times N_{rx} \times N_{f,p} \times N_{t,p}}$,并作为两个独立通道。其中 $N_{f,p}$ 和 $N_{t,p}$ 分别表示导频在频域和时域的数量, N_{tx} 和 N_{rx} 分别为发射和接收天线数。随后,利用频域映射矩阵 $\mathbf{W}_f \in \mathbb{R}^{N_f \times N_{f,p}}$ 与时域映射矩阵 $\mathbf{W}_t \in \mathbb{R}^{N_t \times N_{t,p}}$,分别沿频域和时域进行线性上采样,最终得到完整OFDM网格的信道估计矩阵 $\hat{\mathbf{H}}_{up} \in \mathbb{R}^{2 \times N_{tx} \times N_{rx} \times N_t}$ 。

(2)特征增强子模块:为利用多天线间的空间相关性并融合复数域信息,首先引入一个复数混合层作为前置处理。将上采样模块的输出 $\hat{\mathbf{H}}_{up}$ 通过一个轻量级 1×1 卷积网络,实现所有天线的实部与虚部特征全通道加权组合,学习天线间的耦合关系,得到混合特征 $\mathbf{H}_{mix} \in \mathbb{R}^{C_{mix} \times N_f \times N_t}$,其中 C_{mix} 为 1×1 卷积后的通道数, $C_{mix} = 2N_{tx}N_{rx}$ 。随后,将 $\mathbf{H}_{mix}$ 经输入投影层映射至高维特征空间,得到 $\mathbf{H}_{stem} \in \mathbb{R}^{C_{base} \times N_f \times N_t}$,其中 C_{base} 为中间特征通道数;进一

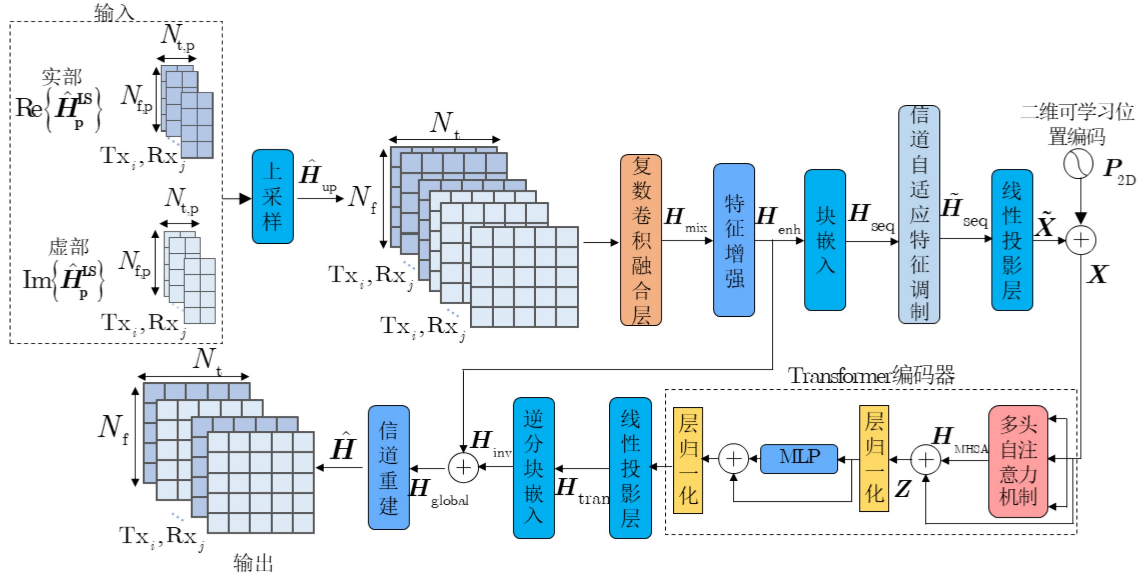
步将 $\mathbf{H}_{stem}$ 输入堆叠的多尺度卷积块以构建核心表征,每个卷积块内并行包括三个不同感受野的分支,标准 3×3 卷积,膨胀 3×3 卷积和 1×1 卷积,其特征融合表示为:

$$\mathbf{H}_{ms} = \psi(\text{BN}(\mathbf{H}_{3 \times 3} + \mathbf{H}_{dilated} + \mathbf{H}_{1 \times 1})) \in \mathbb{R}^{C_{base} \times N_f \times N_t} \quad (4)$$

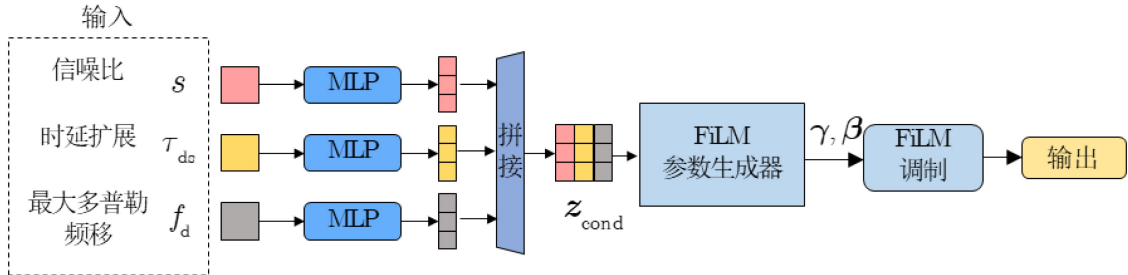
其中, $\text{BN}(\cdot)$ 表示批归一化(Batch Normalization, BN)操作, $\psi(\cdot)$ 为高斯误差线性单元(Gaussian Error Linear Unit, GELU)激活函数,堆叠两层多尺度卷积结构得到深层特征 $\mathbf{H}_{deep}$ 。随后,引入压缩-激励(Squeeze-and-Excitation, SE)机制,自适应生成通道权重向量 $\mathbf{w} \in \mathbb{R}^{C_{base}}$,并对 $\mathbf{H}_{deep}$ 进行通道重标定。加权后的特征经投影层恢复原始通道维度后,与原始混合特征 $\mathbf{H}_{mix}$ 通过全局残差连接融合,最终得到 $\mathbf{H}_{enh} \in \mathbb{R}^{C_{mix} \times N_f \times N_t}$,为后续Transformer提供局部特征嵌入。

(3)块嵌入子模块:该模块用于将特征增强子模块输出的二维特征图转换为适配Transformer架构的序列化表示,同时保留多天线信道在频域与时域上的局部相关性。与视觉Transformer中沿空间维度进行块划分的方式不同,该模块在时频平面上进行分块,并在通道维度上对各局部块进行拼接,以形成结构化的序列表示。具体而言,采用尺寸为 3×2 的滑动窗口对 $\mathbf{H}_{enh}$ 进行非重叠划分,将特征图在空间上离散化为 N 个局部块,其中序列长度 $N = N_f N_t / 6$ 。针对每一个局部块,将各通道内的 3×2 局部矩阵展平为长度为6的向量,并在通道维度上进行串行拼接,使每一个局部块对应一个维度为 $6C_{mix}$ 高维特征向量。最终,整个OFDM时频网格被重塑成特征序列 $\mathbf{H}_{seq} \in \mathbb{R}^{(N_f N_t / 6) \times 6C_{mix}}$ 。

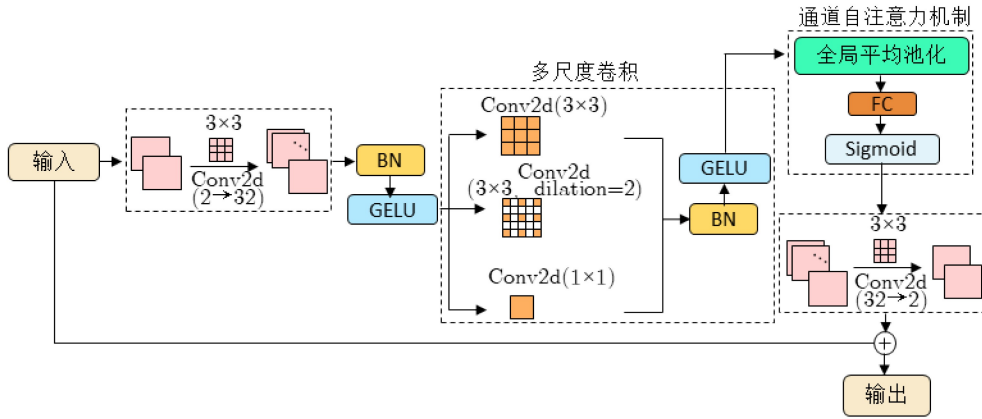
(4)信道自适应特征调制子模块:受FiLM在多任务学习及概率深度学习领域的应用启发,本文引入基于FiLM的自适应机制。该模块首先利用三个独立的多层感知机(Multi-Layer Perceptron, MLP)分别将信噪比 s 、时延扩展 τ_{ds} 、最大多普勒频移 f_d 进行嵌入编码,得到三个高维的条件向量,再将



(a) AdaFiT模型总架构



(b) 信道自适应特征调制子模块



(c) 特征增强及信道重建子模块

图 2 AdaFiT模型架构及子模块

三个条件向量进行拼接，形成融合多维信道环境信息的条件向量 z_{cond} 。随后，利用FiLM参数生成器将 z_{cond} 映射为缩放调制向量 $\gamma \in \mathbb{R}^{6C_{mix}}$ 和偏移向量 $\beta \in \mathbb{R}^{6C_{mix}}$ ，并沿序列维广播，对特征序列 H_{seq} 进行逐通道线性调制：

$$\tilde{H}_{seq} = (1 + \gamma) \odot H_{seq} + \beta \quad (5)$$

其中 $\tilde{H}_{seq} \in \mathbb{R}^{(N_f N_t / 6) \times 6C_{mix}}$ ， $\odot$ 表示哈达码积。借助这一机制，网络可依据当前的信道环境，自适应地

重标定不同特征通道响应，从而提升特征表示对不同信道条件的适应性。

(5) Transformer编码器子模块：首先将FiLM调制的特征序列 $\tilde{H}_{seq}$ 应用线性投影 $W_p \in \mathbb{R}^{(6C_{mix}) \times D_{model}}$ 及偏置向量 $b_p \in \mathbb{R}^{(N_f N_t / 6) \times D_{model}}$ ，得到高维特征向量 $\tilde{X} \in \mathbb{R}^{(N_f N_t / 6) \times D_{model}}$ 作为初始输入特征。由于标准自注意力机制本身不具备显式的位置感知能力，直接输入序列化特征会导致时频网格的空间结构信

息丢失, 本文引入二维可学习位置编码, 定义行嵌入参数矩阵 $\mathbf{P}_{\text{row}} \in \mathbb{R}^{1 \times N_{\text{row}} \times 1 \times D_{\text{model}}}$ 与列嵌入参数矩阵 $\mathbf{P}_{\text{col}} \in \mathbb{R}^{1 \times 1 \times N_{\text{col}} \times D_{\text{model}}}$, 其中 N_{row} 和 N_{col} 表示频域和时域网格块数量, D_{model} 表示模型维度, 通过广播相加得到二维位置编码 $\mathbf{P}_{2D}$ 。随后, 将 $\hat{\mathbf{X}}$ 重塑为 $N_{\text{row}} \times N_{\text{col}}$ 二维特征网格, 与 $\mathbf{P}_{2D}$ 逐元素相加后重新展平, 得到融合时频位置信息特征序列 $\mathbf{X}$ 。随后, $\mathbf{X}$ 被输入由 L 层堆叠而成的Transformer编码器。每层包含多头自注意力(Multi-Head Self-Attention, MHSA)机制与前馈神经网络(Feed-Forward Network, FFN), 并辅以残差连接与层归一化操作。MHSA机制通过并行计算 M 个注意力头来捕获不同子空间的信息。对于第 l 层的第 h 个注意力头, 其输出通过缩放点积注意力计算:

$$\begin{aligned} & \mathbf{A}_h \\ &= \text{softmax} \left(\frac{(\mathbf{X}^{l-1} \mathbf{W}_Q^h + \mathbf{B}_Q^h)(\mathbf{X}^{l-1} \mathbf{W}_K^h + \mathbf{B}_K^h)^T}{\sqrt{d_k}} \right) \\ & \quad \cdot (\mathbf{X}^{l-1} \mathbf{W}_V^h + \mathbf{B}_V^h) \end{aligned} \quad (6)$$

其中, $\mathbf{X}^{l-1}$ 表示第 l 个编码层的输入特征, $\mathbf{W}_Q^h, \mathbf{W}_K^h, \mathbf{W}_V^h \in \mathbb{R}^{D_{\text{model}} \times d_k}$ 和 $\mathbf{B}_Q^h, \mathbf{B}_K^h, \mathbf{B}_V^h \in \mathbb{R}^{(N_f N_t / 6) \times d_k}$ 分别为第 h 个头的投影矩阵和偏置矩阵, $d_k = D_{\text{model}} / M$ 为缩放因子。所有头的输出被拼接成形状 $\mathbb{R}^{(N_f N_t / 6) \times D_{\text{model}}}$, 然后通过线性变换 $\mathbf{W}_O \in \mathbb{R}^{D_{\text{model}} \times D_{\text{model}}}$ 和偏置矩阵 $\mathbf{B}_O \in \mathbb{R}^{(N_f N_t / 6) \times D_{\text{model}}}$ 进行线性投影得到多头注意力输出 $\mathbf{H}_{\text{MHSA}} \in \mathbb{R}^{(N_f N_t / 6) \times D_{\text{model}}}$ 。输出经过残差连接与归一化:

$$\mathbf{Z} = \text{LN}(\mathbf{H}_{\text{MHSA}} + \mathbf{X}) \quad (7)$$

$$\mathbf{H}_{\text{tran}} = \text{LN}(\mathbf{Z} + \text{MLP}(\mathbf{Z})) \quad (8)$$

其中, $\text{LN}(\cdot)$ 代表层归一化(Layer Normalization, LN), $\text{MLP}(\mathbf{Z}) = \psi(\mathbf{Z} \mathbf{W}_{t,1} + \mathbf{B}_{t,1}) \mathbf{W}_{t,2} + \mathbf{B}_{t,2}$, $\mathbf{W}_{t,1} \in \mathbb{R}^{D_{\text{model}} \times 2D_{\text{model}}}$ 和 $\mathbf{W}_{t,2} \in \mathbb{R}^{2D_{\text{model}} \times D_{\text{model}}}$ 作为线性投影, $\mathbf{B}_{t,1} \in \mathbb{R}^{(N_f N_t / 6) \times 2D_{\text{model}}}$ 和 $\mathbf{B}_{t,2} \in \mathbb{R}^{(N_f N_t / 6) \times D_{\text{model}}}$ 作为偏置矩阵, $\psi(\cdot)$ 表示 GELU 函数, 本文设置 $L = 6$, $M = 4$ 以及 $D_{\text{model}} = 128$ 以平衡性能与复杂度的权衡。

(6) 信道重建子模块: 编码器输出 $\mathbf{H}_{\text{tran}} \in \mathbb{R}^{(N_f N_t / 6) \times D_{\text{model}}}$ 经过 $\mathbf{W}_2 \in \mathbb{R}^{D_{\text{model}} \times 6C_{\text{mix}}}$ 线性变换和偏置矩阵 $\mathbf{B}_2 \in \mathbb{R}^{(N_f N_t / 6) \times 6C_{\text{mix}}}$ 还原到形状 $\mathbb{R}^{(N_f N_t / 6) \times 6C_{\text{mix}}}$, 随后进行块划分逆操作生成 $\mathbf{H}_{\text{inv}} \in \mathbb{R}^{C_{\text{mix}} \times N_f \times N_t}$, 将其与特征增强模块的输出 $\mathbf{H}_{\text{enh}}$ 逐元素相加得到全局信道估计 $\mathbf{H}_{\text{global}}$, 最后通过多层卷积细化 $\mathbf{H}_{\text{global}}$ 得到最终的信道估计矩阵 $\hat{\mathbf{H}} \in \mathbb{R}^{2 \times N_{\text{tx}} \times N_{\text{rx}} \times N_f \times N_t}$ 。

总体而言, AdaFiT 以导频位置处的 LS 信道估

计结果为输入, 依次通过稀疏观测上采样、局部特征增强、块级序列建模、信道条件调制、全局依赖建模以及残差信道重建等过程, 端到端地学习从稀疏导频观测到完整信道响应之间的映射关系。模型首先将导频位置处的稀疏 LS 估计上采样至完整 OFDM 时频网格, 获得初始信道表示; 随后利用卷积特征增强模块提取局部时频相关性与多天线耦合特征, 并通过块嵌入模块将二维时频特征转换为适配 Transformer 处理的序列化表示; 在此基础上, 信道自适应特征调制模块根据信噪比、时延扩展和最大多普勒频移等信道物理参数对块级特征进行条件调制, 使网络能够根据输入信道先验动态调整特征表示; 最后, Transformer 编码器进一步建模全局时频依赖关系, 并由信道重建模块融合局部与全局特征, 输出完整信道估计结果。

3.2 模型训练

AdaFiT 采用端到端联合训练方式。训练过程中, 上述各子模块均在同一损失函数约束下协同优化, 所有可训练参数通过反向传播统一更新, 不进行模块预训练、分阶段训练或级联式独立训练。网络参数集用 θ 表示。AdaFiT 的输入为导频位置处的 LS 估计 $\hat{\mathbf{H}}_{\text{p}}^{\text{LS}}$, 信噪比 s 、时延扩展 τ_{ds} 和最大多普勒频移 f_{d} 。最终, 估计的信道响应定义为

$$\hat{\mathbf{H}} = f(\hat{\mathbf{H}}_{\text{p}}^{\text{LS}}, s, \tau_{\text{ds}}, f_{\text{d}}; \theta) \quad (9)$$

其中, $f(\cdot; \theta)$ 代表由 $\theta \in \mathbb{R}$ 参数化的 AdaFiT 网络对应的映射函数, 通过最小化数据集上的均方误差 (Mean Squared Error, MSE) 进行优化。训练目标定义为

$$\min_{\theta} \left\{ \frac{1}{M} \sum_{i=1}^M \left\| \mathbf{H}_i - \hat{\mathbf{H}}_i \right\|_2^2 \right\} \quad (10)$$

其中, $\mathbf{H}_i$ 表示训练数据集中第 i 个样本的真实信道响应, M 代表训练样本总数。

AdaFiT 采用 Adam 优化器训练 500 轮, 配合早停机制以防止模型过拟合。批次大小为 128, 初始学习率为 0.001, 本文采用指数学习率衰减策略, 衰减因子为 0.995。

3.3 复杂度分析

为系统评估不同方法的复杂度, 本文从参数规模与推理计算开销两方面进行度量, 并采用 $O(\cdot)$ 描述渐近时间复杂度。结果如表 1 所示, 其中, N_{tx} 和 N_{rx} 分别代表发射和接收天线的数量, N_f 代表总子载波数量, N_t 代表总符号数。

对于传统方法, LS-Bilinear 不含可训练参数, 其计算复杂度主要来源于双线性插值, 复杂度随天

线规模和时频资源线性增长, 为 $O(N_{\text{tx}}N_{\text{rx}}N_{\text{f}}N_{\text{t}})$ 。LMMSE方法需要进行与子载波维度相关的协方差矩阵运算, 复杂度为 $O(N_{\text{tx}}N_{\text{rx}}N_{\text{f}}^3N_{\text{t}})$, 因此在大规模子载波场景下计算开销较高。在深度学习方法中, AdaFiT、AdaFiT去除自适应模块后的对比模型与AdaFortiTran均包含天线空间特征构造过程, 其共同开销为 $O((N_{\text{tx}}N_{\text{rx}})^2N_{\text{f}}N_{\text{t}})$ 。无自适应模块的AdaFiT模型利用标准的Transformer结构对序列化特征进行建模, 其自注意力和前馈网络的复杂度分别为 $O(S^2D)$ 和 $O(SD^2)$, 其中 S 表示输入序列的长度, D 表示每个序列元素的特征维度。AdaFiT模型在此基础上引入基于FiLM的自适应调制机制, 其额外复杂度为 $O((N_{\text{tx}}N_{\text{rx}})N_{\text{f}}N_{\text{t}})$, 该项开销相较于自注意力模块可忽略不计。相比之下, AdaFortiTran将 k 维高阶特征向量与原始特征拼接, 输入维度扩展为 $D+k$, 其自注意力机制与前馈网络的计算复杂度分别增大为 $O(S^2(D+k))$ 和 $O(S(D+k)^2)$, 实际计算开销相应增加。

综合表1可见, 三类深度学习模型在计算复杂度上均低于子载波维度呈立方增长的LMMSE方法。其中, 无自适应模块的AdaFiT模型的参数规模与计算开销最低, AdaFortiTran因特征维度扩展具有更高的计算与存储开销, 而AdaFiT仅引入有限的附加复杂度, 即可获得更优的估计性能与鲁棒性, 在性能与复杂度之间实现较好的权衡。

4 仿真结果

4.1 数据集构建

数据集是模型训练与性能评估的基础。本文基于第二节所建立的MIMO-OFDM系统模型构建仿真数据集, 用于模型的训练、验证与测试。系统参数参照3GPP NR标准的典型配置进行设置。其中, 信道模型采用CDL-C模型, 系统采用 4×4 MIMO结构, 收发端均配置为4阵元均匀线性阵列(Uniform Linear Array, ULA), 阵元间距为 0.5λ , 天线单元极化角为 -45° ; CDL信道模型结合簇级角度参数与收发端阵列响应生成具有空间相关性的MIMO信道矩阵。载波频率设置为3.5 GHz, 子载波间隔为15 kHz。数据生成过程中对信噪比、时延扩展以及最大多普勒频移参数进行随机配置, 构建具有多样统计特性的信道样本。系统详细参数如表2所示。

训练数据集共包含30 000个独立信道样本, 其中, 80%样本用于模型训练, 20%样本用于模型验证。测试数据集在相同信道模型与系统配置下构建, 并针对关键参数变化设计了多种测试场景。动态信噪比测试集包括信噪比取值为 $\{0, 5, 10, 15, 20, 25, 30\}$ dB的500个测试通道, 时延扩展设为300 ns, 最大多普勒频移设为972 Hz。动态时延扩展测试集包括时延扩展取值为 $\{100, 200, \dots, 700\}$ ns的500个测试通道, 信噪比固定为20 dB, 最大多普勒频移为500 Hz。动态最大多普勒频移测试集包含最大多普

表 1 模型复杂度对比

模型	参数量	复杂度
LS-Bilinear	-	$O(N_{\text{tx}}N_{\text{rx}}N_{\text{f}}N_{\text{t}})$
LMMSE	-	$O(N_{\text{tx}}N_{\text{rx}}N_{\text{f}}^3N_{\text{t}})$
AdaFortiTran	2.738M	$O((N_{\text{tx}}N_{\text{rx}})^2N_{\text{f}}N_{\text{t}} + S^2(D+k) + S(D+k)^2)$
AdaFiT(无自适应模块)	0.268M	$O((N_{\text{tx}}N_{\text{rx}})^2N_{\text{f}}N_{\text{t}} + S^2D + SD^2)$
AdaFiT(本文模型)	0.958M	$O((N_{\text{tx}}N_{\text{rx}})^2N_{\text{f}}N_{\text{t}} + S^2D + SD^2 + (N_{\text{tx}}N_{\text{rx}})N_{\text{f}}N_{\text{t}})$

表 2 MIMO-OFDM系统参数

参数	符号	数值	参数	符号	数值
发射天线数	N_{tx}	4	接收天线数	N_{rx}	4
载波频率	f_{c}	3.5 GHz	信道模型	-	CDL-C; CDL-A
子载波间隔	Δf	15 kHz	系统带宽	B	5 MHz
OFDM符号数	N_{t}	14	总子载波数	N_{f}	300
单天线导频数	N_{p}	50	导频间隔	D	12
数据调制	-	QPSK	信噪比	s	$[0, 5, 10, 15, 20, 25, 30]$ dB
时延扩展	τ_{ds}	$[50, 150, \dots, 650]$ ns	最大多普勒频移	f_{d}	$[100, 200, \dots, 1\,000]$ Hz

勒频移取值为 $\{200, 400, \dots, 1400\}$ Hz, 信噪比固定为20 dB, 时延扩展为300 ns。

4.2 性能评估与鲁棒性分析

为系统评估所提AdaFiT模型的有效性与鲁棒性, 本文选取传统信道估计算法LS-Bilinear和LMMSE作为基线方法, 并引入文献[18]中的AdaFortiTran模型作为已有自适应信道估计模型进行对比, AdaFortiTran将信道参数编码为高维向量与Transformer特征进行拼接或嵌入融合, 而本文所提AdaFiT基于FiLM机制, 将信道物理先验显式映射为特征缩放与偏移参数, 实现对中间特征的自适应调制。同时, 为评估自适应模块对模型性能的影响, 本文通过消融实验分别构建含自适应模块与不含自适应模块的模型配置, 对比分析各模型在不同信噪比(Signal-to-Noise Ratio, SNR)、时延扩展以及最大多普勒频移条件下MSE。

图3展示了CDL-C信道下各模型在动态信噪比测试集下的MSE随信噪比变化的对比曲线。结果表明, 基于深度学习的估计算法在性能上显著优于传统的LS-Bilinear和LMMSE算法。其中, AdaFiT在整个信噪比变化范围内均保持最低的MSE, 展现出最优的整体估计性能。在消融实验中, AdaFiT相比无自适应模块的基线, 最高获得约2.1 dB的MSE下降, 这验证了所设计的信道自适应特征调制模块能够有效提升模型对动态环境的鲁棒性。与同类自适应模型AdaFortiTran对比, 在低信噪比(SNR=0 dB)条件下, AdaFiT的MSE降低约2.5 dB。随着信噪比升高, 性能优势进一步扩大, 在高信噪比(SNR=30 dB)条件下, AdaFiT实现约7 dB的MSE性能提升。这一对比结果表明, 相较于将信道参数简单映射为高维向量后直接拼接的融合方式, 本文基于FiLM机制的物理先验显式地转化为

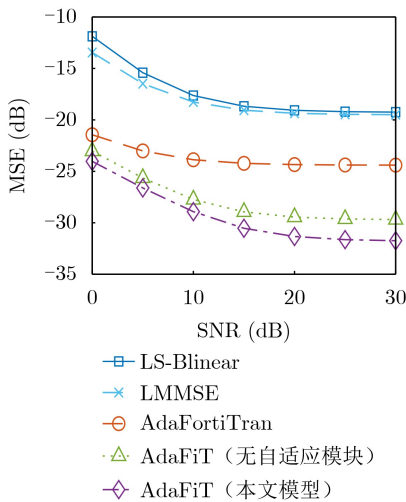


图3 不同信噪比下各模型的MSE

对特征分布的自适应缩放与偏移, 实现特征层动态重标定。这种设计不仅强化了先验信息与深层特征的交互效率, 也显著提升了模型在宽信噪比范围内的估计精度与适应性。

图4展示了CDL-C信道下动态最大多普勒频移测试集下, 各模型的MSE随多普勒频移变化的对比情况。可以看出, 当多普勒频移从200 Hz增大至1400 Hz时, AdaFiT在整个变化区间内始终保持最低MSE, 体现了其对快时变信道条件的优越适应能力与稳定性。特别是在高多普勒区间(1000–1400 Hz), AdaFiT相较于无自适应模块的模型能稳定取得约2 dB的MSE增益; 在同一区间内, 较AdaFortiTran也实现约3.5 dB的性能提升。这些结果证实, AdaFiT在高多普勒频移主导的高速移动场景下, 仍能维持较高估计精度与鲁棒性。

图5展示了CDL-C信道下动态时延扩展测试集

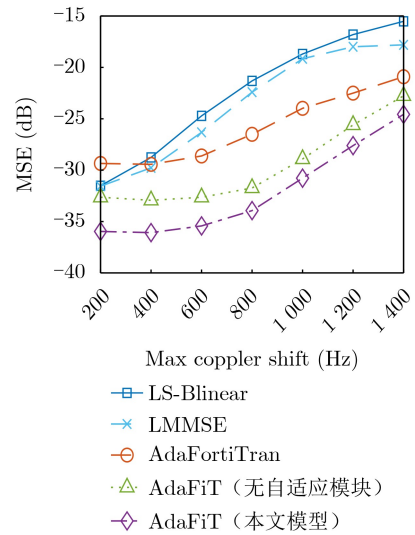


图4 不同最大多普勒频移下各模型的MSE

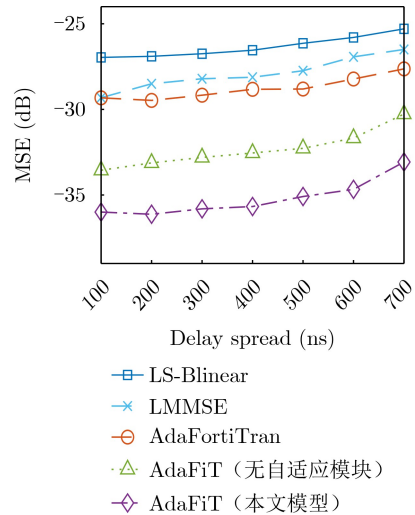


图5 不同时延扩展下各模型的MSE

下各模型的MSE随时延扩展变化的对比情况。观察可知,在100–700 ns的时延扩展范围内,各模型的MSE整体变化较为平稳,而AdaFiT始终保持最低MSE,显示出对多径条件变化的良好稳定性。在全区间内,AdaFiT相比无自适应模块的模型可稳定获得约3 dB的MSE增益,这说明所引入的自适应机制能有效抑制频率选择性衰落带来的不利影响。同时,AdaFiT相较AdaFortiTran始终保持约5.5 dB的性能优势,进一步表明所提模型在时延扩展动态变化的场景下,不仅具备应对多径变化的鲁棒性,也拥有对强频率选择性衰落的高精度估计能力。

4.3 不同信道模型下的性能分析

为进一步评估AdaFiT模型在不同信道模型下的性能表现,本文基于3GPP CDL-A信道开展仿真分析。CDL-A信道代表具有严重多径衰落的非视距(NLOS)场景,其功率延迟谱与空间角度分布与CDL-C信道存在差异。图6展示了在CDL-A信道下,各算法均方误差(MSE)随不同信噪比、最大多普勒频移以及时延扩展变化的对比结果。整体而言,在不同信噪比、最大多普勒频移以及时延扩展条件下,AdaFiT均取得较低MSE,并优于各对比算法,表明所提模型在CDL-A信道条件下仍具有较好的估计性能。

如图6(a)所示,在低信噪比(0–5 dB)条件下,AdaFiT较LMMSE获得约4–5 dB的性能增益;在高信噪比条件(20–30 dB)下,AdaFiT的MSE较LMMSE降低7 dB。与同样引入先验信息的AdaFortiTran模型相比,AdaFiT在低信噪比条件下约获得2 dB的MSE增益,并在较高信噪比条件下优势扩大至约5.5 dB。尽管CDL-A与CDL-C信道在空间散射特征和多径结构方面存在差异,AdaFiT仍能够在不同信噪比条件下保持较为稳定的估计性

能,表明其在不同标准化CDL信道配置下具有一定的适应能力。

如图6(b)所示,随着最大多普勒频移增大,LS和LMMSE的MSE整体呈上升趋势,表明传统方法在快时变信道条件下估计性能的退化。相比之下,AdaFiT在200 Hz至800 Hz的中低多普勒频移范围内保持较平稳的性能,其MSE维持在–38 dB至–37 dB之间,较AdaFortiTran保持约5 dB的优势。当最大多普勒频移超过1000 Hz后,受信道快速时变及子载波间干扰加剧的影响,AdaFiT的MSE有所上升;但在1400 Hz条件下,其估计性能较LMMSE仍保持约6 dB的增益,并优于AdaFortiTran。上述结果表明,在具有不同多普勒扩展特性的CDL-A信道条件下,AdaFiT能够利用最大多普勒频移先验对中间特征进行条件调制,并保持相对稳定的信道估计性能,体现出对快时变信道环境的一定适应能力。

由图6(c)可知,面对CDL-A复杂的频率选择性衰落,AdaFiT在整个时延扩展区间内始终保持最优估计性能,相较传统LMMSE算法可获得约4至6 dB的性能增益。同时,相较于AdaFortiTran模型,AdaFiT在全范围的时延扩展下也稳定保持约4至5.5 dB的性能领先。这表明,本文所提的AdaFiT网络并不依赖于单一特定信道模型的统计经验。在面对具有不同功率延迟谱与多径结构的CDL-A信道时,该架构依然能够有效利用时延扩展等先验信息实现高精度的信道估计,体现出较好的适用性与动态环境适应能力。

综合图6的仿真结果可知,在CDL-A信道下,AdaFiT在不同信噪比、最大多普勒频移以及时延扩展条件下均保持最优的信道估计性能。结合前述CDL-C信道下的仿真结果可知,所提方法的性能

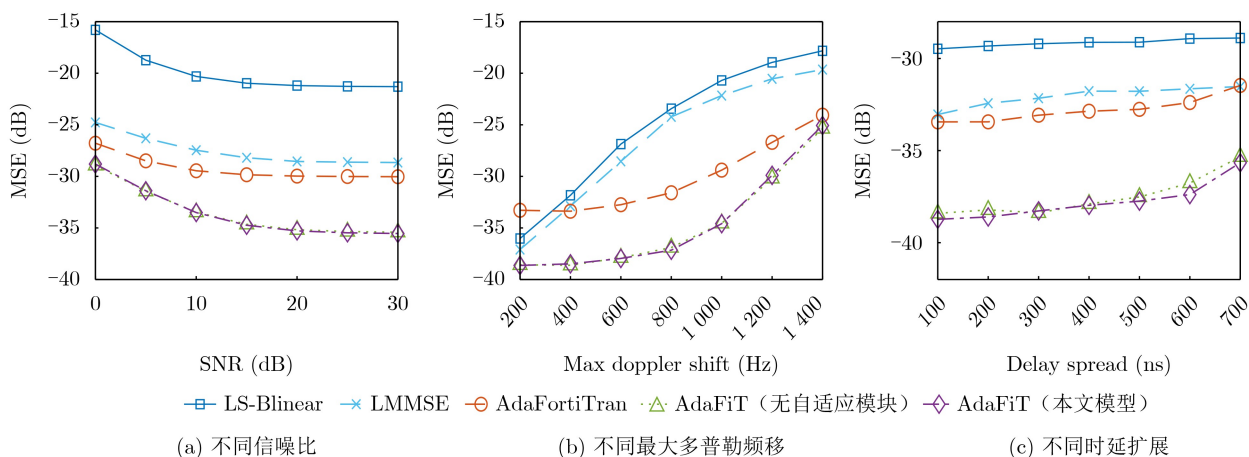


图6 CDL-A信道下各模型在不同信道条件下的MSE性能对比

优势并非仅限于单一CDL-C信道模型。由于CDL-A与CDL-C在功率延迟谱、簇级角度参数及多径结构方面存在差异，二者对应的MIMO信道空间统计特性也有所不同。因此，上述结果表明，AdaFiT在本文所考虑的不同标准化CDL信道配置下具有一定的泛化能力和性能稳定性。

4.4 先验失配鲁棒性分析

在实际通信系统中，信噪比、最大多普勒频移及时延扩展等信道物理参数可基于导频符号、参考信号或已知训练序列进行估计。通常，信噪比可由导频接收功率与噪声方差估计得到；最大多普勒频移可由相邻OFDM符号或相邻帧之间导频信道估计的时间相关性获得，也可结合终端移动速度和载波频率进行近似计算；时延扩展则可由导频或训练序列恢复得到的信道冲激响应及其功率时延谱进行估计。本文研究中假设上述信道参数可通过系统测量或估计准确获得，重点评估其与AdaFiT信道估计网络融合后对估计性能的影响。

由于实际系统中参数估计不可避免地存在误差，输入网络的先验信息可能与真实信道参数不完全一致。为评估算法在先验信息失配条件下的鲁棒性，本文开展了先验失配仿真分析。该仿真仅在模型评

估阶段对输入信道自适应模块的先验参数进行扰动，而保持导频观测输入与真实信道标签不变，以模拟先验信息存在估计偏差的情况。具体而言，对评估数据集的每个样本，在其SNR上加入范围为 $[-1,1]$ dB的随机加性误差；对时延扩展和最大多普勒频移分别乘以 $[0,1]$ 范围内的随机缩放因子，以模拟相关先验参数与真实信道参数不完全匹配的情况，并对比准确先验和错误先验条件下的均方误差(MSE)性能。

如图7所示，当先验信息存在失配时，所提算法在不同信噪比、最大多普勒频移和时延扩展条件下的MSE性能均有所下降，但整体性能损失较为有限，且未随信道条件变化呈现明显发散趋势。具体而言，相较于准确先验条件，失配先验条件下所提算法在不同信噪比条件下的MSE性能下降约1–3 dB，在不同最大多普勒频移条件下的MSE性能下降约1–2 dB，在不同时延扩展条件下的MSE性能下降约1.5–2 dB。该结果表明，先验信息误差会在一定程度上影响模型的信道估计精度，但所提AdaFiT模型在先验失配条件下仍能保持较为稳定的估计性能，未出现明显性能退化或估计失稳现象，体现出其对先验信息不完全匹配的容错能力以及在快时变信道环境下的鲁棒性。

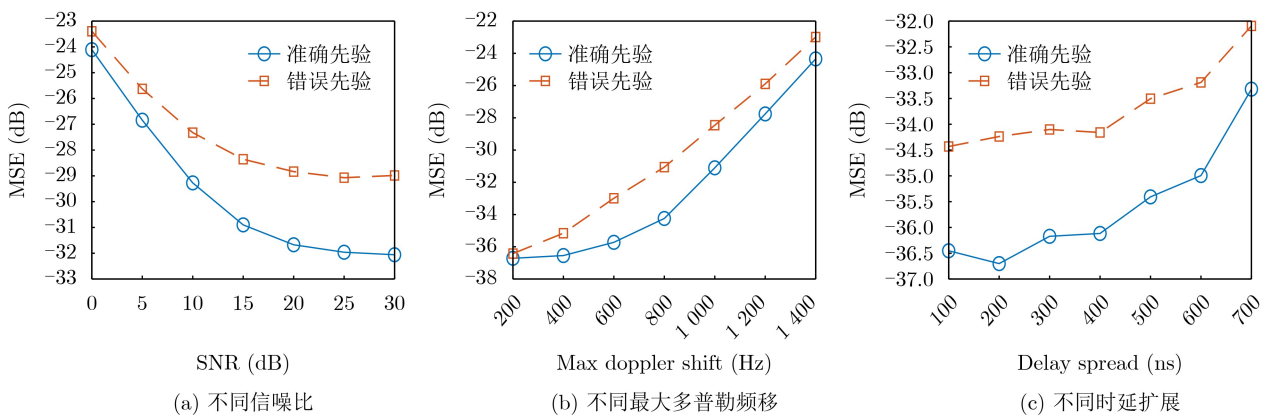


图7 先验信息失配对AdaFiT信道估计性能的影响

5 结束语

本文针对高速移动场景下MIMO-OFDM信道快速时变与子载波间干扰增强导致的估计性能受限问题，提出一种融合物理先验的自适应信道估计模型AdaFiT。该方法通过FiLM机制实现基于物理参数的特征级动态调制，并结合多尺度卷积与Transformer结构协同建模局部与全局信息，从而提升信道重建精度与鲁棒性。仿真结果表明，在CDL等典型信道模型及不同信道条件下，该方法均表现出优异的估计性能与稳定性。未来将进一步探索其在大规模MIMO稀疏信道估计中的扩展应用。

参考文献

- [1] SINGH S, KUMAR S, MAJHI S, *et al.* Blind carrier frequency offset estimation techniques for next-generation multicarrier communication systems: Challenges, comparative analysis, and future prospects[J]. *IEEE Communications Surveys & Tutorials*, 2025, 27(1): 1–36. doi: [10.1109/COMST.2024.3472109](https://doi.org/10.1109/COMST.2024.3472109).
- [2] GUO Jiajia, CHEN Tong, JIN Shi, *et al.* Deep learning for joint channel estimation and feedback in massive MIMO systems[J]. *Digital Communications and Networks*, 2024, 10(1): 83–93. doi: [10.1016/j.dcan.2023.01.011](https://doi.org/10.1016/j.dcan.2023.01.011).
- [3] XU Jiarui, LI Lianjun, ZHENG Lizhong, *et al.* Learning to

- Estimate: A real-time online learning framework for MIMO-OFDM channel estimation[J]. *IEEE Transactions on Wireless Communications*, 2025, 24(4): 2634–2646. doi: [10.1109/TWC.2024.3487111](#).
- [4] CAI Jun, YIN Chuan, and DING Youwei. Optimization of resource allocation in FDD massive MIMO systems[J]. *Digital Communications and Networks*, 2024, 10(1): 117–125. doi: [10.1016/j.dcan.2022.11.017](#).
- [5] 李一兵, 汤云鹤, 简鑫, 等. 面对高速移动场景的OTFS系统导频设计方法[J]. 电子与信息学报, 2025, 47(2): 490–497. doi: [10.11999/JEIT240349](#).
- LI Yibing, TANG Yunhe, JIAN Xin, *et al.* Pilot design method for OTFS system in high-speed mobile scenarios[J]. *Journal of Electronics & Information Technology*, 2025, 47(2): 490–497. doi: [10.11999/JEIT240349](#).
- [6] 蒲旭敏, 刘雁翔, 宋米雪, 等. 基于模型驱动深度学习的OTFS信道估计[J]. 电子与信息学报, 2024, 46(2): 680–687. doi: [10.11999/JEIT230072](#).
- PU Xumin, LIU Yanxiang, SONG Mixue, *et al.* Orthogonal time frequency space channel estimation based on model-driven deep learning[J]. *Journal of Electronics & Information Technology*, 2024, 46(2): 680–687. doi: [10.11999/JEIT230072](#).
- [7] 廖勇, 罗渝, 荆亚昊. 6G新型时延多普勒通信范式: OTFS的技术优势、设计挑战、应用与前景[J]. 电子与信息学报, 2024, 46(5): 1827–1842. doi: [10.11999/JEIT231133](#).
- LIAO Yong, LUO Yu, and JING Yahao. 6G new time-delay Doppler communication paradigm: Technical advantages, Design challenges, applications and prospects of OTFS[J]. *Journal of Electronics & Information Technology*, 2024, 46(5): 1827–1842. doi: [10.11999/JEIT231133](#).
- [8] KIM D, PARK S, KANG J, *et al.* Block-fading non-stationary channel estimation for MIMO-OFDM systems via meta-learning[J]. *IEEE Communications Letters*, 2022, 26(12): 2924–2928. doi: [10.1109/LCOMM.2022.3204763](#).
- [9] GHEREKHLOO S, ARDAH K, and HAARDT M. SALSA: A sequential alternating least squares approximation method for MIMO channel estimation[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(5): 7430–7435. doi: [10.1109/TVT.2023.3347290](#).
- [10] GIZZINI A K and CHAFII M. A survey on deep learning based channel estimation in doubly dispersive environments[J]. *IEEE Access*, 2022, 10: 70595–70619. doi: [10.1109/ACCESS.2022.3188111](#).
- [11] DOHA S R and ABDELHADI A. Deep learning in wireless communication receivers: A survey[J]. *IEEE Access*, 2025, 13: 113586–113605. doi: [10.1109/ACCESS.2025.3584000](#).
- [12] JIANG Peiwen, WEN Chaokai, JIN Shi, *et al.* Dual CNN-based channel estimation for MIMO-OFDM systems[J]. *IEEE Transactions on Communications*, 2021, 69(9): 5859–5872. doi: [10.1109/TCOMM.2021.3085895](#).
- [13] LEE J, AHN S, PARK S I, *et al.* Alternative meta-learning with 3D dual-CNN for MIMO channel estimation[J]. *IEEE Wireless Communications Letters*, 2025, 14(11): 3650–3654. doi: [10.1109/LWC.2025.3600061](#).
- [14] GAO Wei, ZHANG Wei, LIU Libin, *et al.* Deep residual learning with attention mechanism for OFDM channel estimation[J]. *IEEE Wireless Communications Letters*, 2025, 14(2): 250–254. doi: [10.1109/LWC.2022.3232378](#).
- [15] FOLA E, LUO Yang, and LUO Chunbo. AttenReEsNet: Attention-aided residual learning for effective model-driven channel estimation[J]. *IEEE Communications Letters*, 2024, 28(8): 1855–1859. doi: [10.1109/LCOMM.2024.3412802](#).
- [16] LI Jinbao and PENG Qi. Lightweight channel estimation networks for OFDM systems[J]. *IEEE Wireless Communications Letters*, 2022, 11(10): 2066–2070. doi: [10.1109/LWC.2022.3193199](#).
- [17] LIU Fangyu, ZHANG Jing, JIANG Peiwen, *et al.* CE-ViT: A robust channel estimator based on vision transformer for OFDM systems[C]. GLOBECOM 2023-2023 IEEE Global Communications Conference, Kuala Lumpur, Malaysia, 2023: 4798–4803. doi: [10.1109/GLOBECOM54140.2023.10436847](#).
- [18] GULER B and JAFARKHANI H. AdaFortiTran: An adaptive transformer model for robust OFDM channel estimation[C]. ICC 2025-IEEE International Conference on Communications, Montreal, Canada, 2025: 3797–3802. doi: [10.1109/ICC52391.2025.11160810](#).
- [19] TURKOGLU M O, BECKER A, GÜNDÜZ H A, *et al.* FiLM-Ensemble: Probabilistic deep learning via feature-wise linear modulation[C]. The 36th Annual Conference on Neural Information Processing Systems (NeurIPS 2022), New Orleans, USA, 2022: 22229–22242.
- 廖 希: 女, 博士, 教授, 研究方向为面向6G的信道测量与建模、广域无线通信等, 邮箱为 liaoqi@cqupt.edu.cn.
- 何香妮: 女, 硕士生, 研究方向为OFDM系统信道估计.
- 张 哲: 男, 硕士, 工程师, 研究方向为无线通信, 邮箱为 19628069@qq.com.
- 王 洋: 男, 博士, 教授, 研究方向为第六代移动通信技术、毫米波太赫兹测量与建模、涡旋电磁波和智能反射面等.

责任编辑: 余 蓉

Channel Estimation for MIMO-OFDM Based on Adaptive Transformer Network in High-Speed Mobile Scenarios

LIAO Xi^① HE Xiangni^① ZHANG Zhe^{①②} WANG Yang^{①②}

^①(College of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

^②(PLA Unit 31401, Changchun 130012, China)

Abstract:

Objective Accurate channel state information (CSI) is essential for coherent detection, beamforming, and adaptive resource allocation in multiple-input multiple-output orthogonal frequency division multiplexing (MIMO-OFDM) systems. In high-mobility environments, large Doppler shifts and multipath propagation jointly cause doubly selective fading, destroy subcarrier orthogonality, and aggravate inter-carrier interference. Consequently, conventional least squares (LS) and linear minimum mean square error (LMMSE) estimators suffer substantial performance degradation. Existing deep learning estimators provide strong nonlinear modeling capability, but often show insufficient adaptability to variations in signal-to-noise ratio (SNR), delay spread, and Doppler shift, or inefficiently incorporate physical channel priors. To address these problems, a Transformer network with adaptive feature modulation, termed AdaFiT, is proposed for channel estimation in high-mobility MIMO-OFDM systems.

Methods The proposed AdaFiT framework performs channel estimation by jointly exploiting local time-frequency correlations, global dependencies, and explicit channel-aware adaptation. It takes least squares (LS) estimates at pilot positions, together with signal-to-noise ratio (SNR), delay spread, and maximum Doppler shift, as inputs. A separable two-dimensional linear upsampling module first interpolates the sparse pilot estimates to the full OFDM time-frequency grid by independently processing the real and imaginary components along the frequency and time dimensions. A convolutional feature enhancement module then extracts robust local representations: a complex feature mixing layer fuses multi-antenna real and imaginary components, while multi-scale convolutional blocks with channel attention capture short-range time-frequency correlations and suppress noise. Next, a feature-wise linear modulation (FiLM)-based channel-adaptive module embeds the three channel parameters through independent multilayer perceptrons and combines them into a channel-state representation. This representation generates scaling and shifting coefficients to dynamically recalibrate the block-embedded sequences according to varying channel statistics. Finally, the modulated sequences are processed by a Transformer encoder with learnable two-dimensional positional encoding to model long-range dependencies across the time-frequency grid and antenna dimensions. A residual reconstruction structure fuses the global features with locally enhanced representations, yielding accurate channel estimates with global consistency and preserved local details.

Results and Discussions Simulation analyses are conducted based on the CDL-C and CDL-A channel models. The LS-based bilinear interpolation method, the LMMSE method, the AdaFortiTran model, and the AdaFiT model without the adaptive module are selected as comparison schemes. The mean squared error (MSE) performance of these models is compared under different signal-to-noise ratio (SNR), maximum Doppler shift, and delay spread conditions. Under the CDL-C channel model, AdaFiT achieves the lowest MSE over the entire SNR range (Fig. 3). At an SNR of 0 dB, its MSE is approximately 2.5 dB lower than that of AdaFortiTran, and the performance gain increases to approximately 7 dB at an SNR of 30 dB. Compared with the AdaFiT model without the adaptive module, AdaFiT achieves a maximum MSE gain of approximately 2.1 dB, which verifies the effectiveness of the proposed channel-adaptive feature modulation module. When the maximum Doppler shift increases from 200 Hz to 1400 Hz, AdaFiT consistently maintains the lowest MSE (Fig. 4). In the high-Doppler range of 1000–1400 Hz, AdaFiT outperforms the AdaFiT model without the adaptive module and AdaFortiTran by approximately 2 dB and 3.5 dB, respectively, demonstrating improved robustness against rapid channel variations. For delay spreads ranging from 100 ns to 700 ns, AdaFiT also

maintains the lowest MSE over the entire range (Fig. 5), achieving approximately 3 dB and 5.5 dB gains over the AdaFiT model without the adaptive module and AdaFortiTran, respectively. These results demonstrate that the proposed adaptive feature modulation mechanism effectively improves the adaptability of the model to time-selective and frequency-selective fading. To further evaluate the performance of the AdaFiT model under different channel models, simulation analysis is conducted based on the 3GPP CDL-A channel model. At SNRs of 0–5 dB, AdaFiT outperforms LMMSE by approximately 4–5 dB, while the performance gain reaches approximately 7 dB at SNRs of 20–30 dB (Fig. 6(a)). Compared with AdaFortiTran, AdaFiT achieves an MSE gain of approximately 2 dB under low-SNR conditions, and the gain increases to approximately 5.5 dB under high-SNR conditions. In the maximum Doppler shift experiment, the MSE of AdaFiT remains between approximately –38 dB and –37 dB over the range of 200–800 Hz and is approximately 5 dB lower than that of AdaFortiTran (Fig. 6(b)). Although the MSE of AdaFiT increases when the maximum Doppler shift exceeds 1 000 Hz, it still achieves an approximately 6 dB gain over LMMSE at 1 400 Hz and remains superior to AdaFortiTran. Over the entire delay spread range, AdaFiT achieves approximately 4–6 dB gain over LMMSE and approximately 4–5.5 dB gain over AdaFortiTran (Fig. 6(c)).

Conclusions The proposed AdaFiT framework performs channel estimation by jointly exploiting local time-frequency correlations, global dependencies, and an explicit channel-aware adaptation mechanism. Simulation results under the CDL-C and CDL-A channel models show that AdaFiT consistently achieves lower MSE than the LS-based bilinear interpolation method, LMMSE, AdaFortiTran, and the AdaFiT model without the adaptive module under different SNR, maximum Doppler shift, and delay spread conditions. These results verify the effectiveness of the proposed adaptive feature modulation mechanism and demonstrate that AdaFiT maintains high estimation accuracy and stable performance under different channel models, indicating good adaptability to varying channel environments.

Key words: MIMO-OFDM Systems; Channel estimation; Transformer network; Adaptive feature modulation; Feature-wise Linear Modulation(FiLM)