



研究与开发

基于MIMO-CSI预测的轻量化信源信道联合编码方法

于创宇, 徐彦彦, 潘少明

(武汉大学测绘遥感信息工程全国重点实验室, 湖北 武汉 430079)

摘要: 高效的深度信源信道联合编码 (deep joint source-channel coding, DeepJSCC) 是实现带宽受限场景下语义通信的关键技术, 然而在车联网或无人机等终端资源受限的场景中, 现有方法难以适应多输入多输出 (multiple-input multiple-output, MIMO) 信道的动态变化, 且模型庞大难以部署。为此, 提出一种轻量化 DeepJSCC 框架 (VxLJSCC)。首先, 基于扩展长短期记忆网络的语义提取网络实现轻量化的高质量语义特征提取; 然后, 采用 MIMO-信道状态信息 (channel state information, CSI) 预测来解决语义通信系统使用过时 CSI 而导致系统性能下降的问题; 最后, 为使语义信息充分适应时变 MIMO 信道质量, 设计了基于信道预测的特征分配与自适应模块, 结合语义特征的重要性, 为不同特征分配合适的传输信道和时隙, 并对特征进行调整, 从而提升图像重建的语义精度。实验表明, 相较于先进的 DeepJSCC-MIMO 方法, VxLJSCC 在节省最多 61.67% 模型存储和 77.86% 计算量的情况下, 仍能提供高达 2.972 dB 的信道增益。

关键词: 信源信道联合编码; 语义通信; 信道状态信息; 多输入多输出; 图像传输; 特征分配

中图分类号: TN929.5

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2025126

Lightweight joint source-channel coding method based on MIMO-CSI prediction

YU Chuangyu, XU Yanyan, PAN Shaoming

State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University,
Wuhan 430079, China

Abstract: Efficient deep joint source-channel coding (DeepJSCC) is a key technology for enabling semantic communication in band-width-constrained scenarios. However, in resource-limited environments such as vehicular networks or unmanned aerial vehicles, existing methods struggle to adapt to the dynamic characteristics of multiple-input multiple-output (MIMO) channels and are challenging to deploy due to their large model size. To address these issues, a lightweight DeepJSCC framework (VxLJSCC) was proposed. A semantic extraction network based on extended long short-term memory was proposed to achieve a lightweight, high quality semantic feature extraction.

收稿日期: 2024-12-12; 修回日期: 2025-04-24

基金项目: 国家重点研发计划项目 (No.2022YFB3903404); 湖北省技术创新计划科技重大项目 (No.2024BAA011)

Foundation Items: The National Key Research & Development Program of China (No.2022YFB3903404), Hubei Province Major Science and Technology Innovation Program (No.2024BAA011)



Then, MIMO-channel state information (CSI) prediction was employed to address the performance degradation caused by CSI aging in semantic communication systems. Finally, to adapt the semantic information to the time-varying quality of MIMO channels, a channel adaptive module was designed. This module assigns appropriate transmission subchannels and time slots to different features based on their importance, thereby enhancing the semantic accuracy of image reconstruction. Simulation results show that, compared to the best method DeepJSCC-MIMO, Vx-LJSCC saves up to 61.67% in model storage and 77.86% in computational cost, while still providing up to 2.972 dB channel gain.

Key words: joint source-channel coding, semantic communication, channel state information, MIMO, image transmission, feature allocation

0 引言

随着多媒体和网络技术的发展, 车载终端或无人机终端等设备能够采集到海量观测数据。然而, 在车联网和无人机网络等通信资源受限场景中, 难以在有限的频谱资源上以较低时延传输海量图像或视频数据^[1]。语义通信利用近年来迅速发展的深度学习技术从原始数据中提取语义信息, 在保留主要含义的同时可大幅减少传输数据量, 降低频谱带宽占用^[2-3]。其中, 深度信源信道联合编码 (deep joint source-channel coding, DeepJSCC) 结合语义通信与传统信源信道联合编码技术, 将图像的语义信息直接映射到通道符号中, 将传输单位从语法比特转化为语义比特, 并考虑信道质量对语义信息的影响, 使得在低信噪比 (signal-to-noise ratio, SNR) 条件下, 系统也能保持很好的鲁棒性, 因此该技术近年来取得了突破性进展^[4-10]。

基于深度学习模型提取语义信息, 是语义通信的基础。文献[4]首次提出了用于图像传输的 DeepJSCC, 利用卷积神经网络作为语义提取的骨干网络, 其性能优于分离编码方案。但该网络难以捕捉图像深层次特征和细节。文献[5]提出基于 Swin Transformer 的 DeepJSCC (SwinJSCC), 利用 Swin Transformer 取代卷积神经网络, 提升了系统的语义表达能力, 但注意力机制的引入导致计算复杂度更高。文献[6]提出基于广义状态间

空模型的自适应性 DeepJSCC (MambaJSCC), 降低了模型复杂度, 但仍然难以适用于边缘设备等资源受限的应用场景中^[11]。

现有的 DeepJSCC 方案大多针对单传输通道开展研究, 聚焦于单输入单输出 (single-input single-output, SISO) 的应用场景。文献[7]基于正交频分复用 (orthogonal frequency division multiplexing, OFDM) 技术设计与显式 OFDM 块联合训练的 DeepJSCC 方案, 以解决 SISO 的多径衰落问题。文献[8-9]通过将语义重要性引入传输任务, 以最小化端到端语义通信系统损失, 从而提升整体性能。然而, 上述方法未考虑多输入多输出 (multiple-input multiple-output, MIMO) 系统的空间分集特性。文献[10]首次将 DeepJSCC 应用到 MIMO 系统中, 提出了 DeepJSCC-MIMO。该模型通过信道状态信息 (channel state information, CSI) 评估信道状态, 将输入信号映射到多个天线, 并通过多个信道恢复信息, 充分利用了 MIMO 系统的空间分集优势。但该方法依赖估计的 CSI, 有可能由于信道反馈信息延迟而错误估计 CSI, 导致性能下降^[12]。

针对上述问题, 本文提出一种轻量化且高效的 DeepJSCC 方案 (VxLJSCC), 利用改进的视觉长短期记忆 (vision long short-term memory, ViL) 网络^[13]提取语义信息, 在提升语义信息表征能力的同时降低模型的计算复杂度; 使用扩展长短期记忆 (extended long short-term memory, xLSTM)

网络^[14]模型作为 MIMO-CSI 预测骨干网络, 利用循环结构和分块机制实现轻量化的信道预测器, 可有效捕捉信道的时域和频域变化特性, 并预测信道变化, 避免错误估计 CSI 对系统造成影响; 提出基于预测的多通道自适应特征分配机制, 通过梯度下降在线评估特征的重要性, 将关键语义特征优先分配至高质量子信道和合适时隙, 减轻信道波动对特征传输的影响。此外, 根据预测的信道信息, 自适应调整特征和传输码率, 在动态信道下实现稳健的特征传输。

1 基于 MIMO-CSI 预测的 DeepJSCC 框架

1.1 数学模型

给定源图像 $\mathbf{x} \in \mathbf{R}^{H \times W \times 3}$, H 、 W 和 3 分别表示图像的高度、宽度和 RGB 通道数, 定义 $n = H \times W \times 3$ 为图像传输所需要的源带宽^[4]。本文提出的基于 MIMO-CSI 预测的 DeepJSCC 方法, 结合有损 MIMO 信道, 设计端到端的 DeepJSCC 模型 $f(\cdot)$, 使得通过 $f(\cdot)$ 重建后的图像 $\hat{\mathbf{x}} \in \mathbf{R}^{H \times W \times 3}$ 能够得到更好的重建效果, 如下。

$$\hat{\mathbf{x}} = f(\mathbf{x}) = f_d(\mathbf{H}f_e(\mathbf{x}, \bar{\mathbf{H}}) + \mathbf{N}) \quad (1)$$

其中, $f_e(\cdot)$ 、 $f_d(\cdot)$ 分别是 DeepJSCC 的编解码器函数, 可实现图像语义信息与语法信息的相互转换; $\mathbf{H} \in \mathbf{C}^{N_r \times N_t \times N_f}$ 表示有损 MIMO 信道的三维增益矩阵; $\bar{\mathbf{H}} \in \mathbf{C}^{N_r \times N_t \times N_f}$ 表示预测的信道增益矩阵; $\mathbf{N} \in \mathbf{C}^{N_r \times N_t \times N_f}$ 为三维信道噪声矩阵, 表示实际传输对图像重建产生的干扰; N_r 、 N_t 和 N_f 分别代表接收端天线、发送端天线和子载波的个数。

1.2 VxLJSCC 整体框架

图 1 展示了 VxLJSCC 整体框架。VxLJSCC 包括位于发送端的编码器、接收端的解码器和位于接收端的 MIMO-CSI 预测器 (MIMO-CSI predictor, MCP)。本文提出的 MCP 由位于接收端的多阶段 xLSTM 构成, 通过接收历史 CSI 数据来预测未来时刻的 CSI, 并将预测结果反馈至发送端,

为编解码器的特征分配和自适应编码提供 CSI 支持。本文提出的编码器 $f_e(\cdot)$ 由 VxLJSCC 编码器、信道特征分配模块和信道自适应编码模块组成。在 VxLJSCC 编码器中, 通过分阶段的 xLSTM-V 语义提取模块, 捕获不同尺度的图像语义信息, 并通过挖掘不同输入顺序的图像子块的相互关系, 高效融合多样化特征, 提取更丰富的图像语义信息; 在信道特征分配模块中, 首先根据图像损失梯度评估语义特征的重要性, 基于 MCP 提供的 CSI 计算 MIMO-信干噪比 (signal-to-interference-plus-noise ratio, SINR) 以对 MIMO 信道质量进行评估, 通过重要性和信道质量进行匹配, 将更重要的语义特征分配至高信道质量的通道中, 同时生成二进制向量掩码 $\mathbf{M}$, 用于表征语义特征与信道之间的分配关系, 辅助解码器重建语义特征; 在信道自适应编码模块中, 根据 MIMO-SINR 对语义特征进行编码调整, 生成与信道质量相关的语义表示 $\hat{\mathbf{y}}$ 。该语义表示, 能够表征复杂图像语义并有效对抗信道噪声、衰落等传输干扰。本文提出的解码器 $f_d(\cdot)$ 与编码器呈对称结构, 分别由 VxLJSCC 解码器、信道特征重建模块和信道自适应解码模块组成。信道特征重建模块通过语义二进制向量掩码 $\mathbf{M}$ 重建不同信道时隙得到的语义特征, 继而从语义特征中恢复原始图像信息。

1.3 xLSTM-V 块结构

基于注意力机制的语义提取模型能够挖掘深层次的语义信息, 但由于其计算复杂度较高难以在实际应用中部署。为解决这一问题, 本文基于 xLSTM 中可并行化的矩阵长短期记忆 (matrix long short-term memory, mLSTM) 网络块^[14]构建 xLSTM-V 块结构。xLSTM 本身具有自回归性质, 依赖于顺序图像块的输入方式, 缺少对整个输入图像的全局感知能力。为此该模块以序列 VxL (sequence VxL, S-VxL) 为特征提取网络,



图1 VxLJSCC整体框架

通过扫描拓展方法建立多方向的图像块遍历顺序,使 xLSTM 能够在不同输入组合中实现全局图像块的相互感知,达到语义提取能力的提升。同时,可并行化的 mLSTM 结构有效提高了处理不同顺序的速度。对于每个 xLSTM-V 块,首先通过层归一化对输入 token 进行标准化处理,然后通过全连接 (full-connection, FC) 层将 token 映射到更高维度的特征空间。映射后的高维 token 被输入 S-VxL 中,用于进行高维特征提取。

S-VxL 模块的计算流程如图 2 所示。假设第 i 个 token 对应的矩阵为 $z_i \in \mathbf{R}^{h_i \times w_i}, i=1, 2, \dots, c$, 其中 h_i, w_i 和 c 分别表示第 i 个 token 的高度、宽度和特征维度。为兼顾性能和效率,本文通过两种方式对每个矩阵 z_i 进行从二维至一维的平展,生成相应的两个输入向量 $s_i^j \in \mathbf{R}^{S \times 1}, j=1, 2, S=h_i \times w_i$, 以模拟非重叠块之间不同的输入顺序关系,捕获序列层面的特征信息,从而实现最佳的语义特征提取效果:

$$\begin{cases} s_i^1 = \text{vec}(z_i) \\ s_i^2 = R(\text{vec}(z_i)) \end{cases} \quad (2)$$

其中, $\text{vec}(\cdot)$ 和 $R(\cdot)$ 分别表示二维矩阵平展为一维向量和向量元素顺序的反转。当上述两个向量分别传输至两个独立的 S-VxL 模块后,使用 mLSTM 进一步进行处理,以提取全局的图像信息。对于每一个向量 s_i^j , 按照 mLSTM 结构进行构建和处理。S-VxL 通过捕获当前向量的图像块之间

的空间关系,以实现空间特征的提取,先使用交替块进行并行处理,利用 mLSTM 的 6 个可学习矩阵对特征学习进行调整^[14], 并生成最终的特征状态 $h_i^j \in \mathbf{R}^{S \times 1}$, 然后将两个输出进行融合,生成综合特征向量:

$$h_i = h_i^1 + R\left(\left(h_i^2\right)^T\right) \quad (3)$$

同时,通过 FC 层提取的特征经过 Swish 激活函数进行非线性化。该输出与 S-VxL 的输出逐元素相乘,以融合两路信息。接着,另一个 FC 层将相乘结果投影回原始维度,以确保与输入维度一致。为增强信息流动性,本文通过残差连接将结果与初始输入相加,生成最终的特征表示。

1.4 基于时变信道的 MIMO 信道预测

本节讨论基于 xLSTM 的轻量化 MIMO-CSI 的预测模型 MCP。在该 MIMO 系统中,单个基站配备了 N_t 个天线的均匀线性阵列,用户端配备 N_r 个全向天线。系统采用 OFDM,由 N_f 个子载波组成。在 t 时刻,信号通过 M 条路径传播,每条路径具有特定的传播延迟 τ_m^t 、基站天线到达角 θ_m^t 和复数振幅增益 α_m^t 。该系统中,第 j 个发送天线与第 i 个接收天线在 t 时刻的第 l 个子载波 f_l 上的无线信道可以描述为^[15]:

$$h_{i,j}^t[l] = \sum_{m=1}^M \alpha_m^t a(\theta_m^t) e^{-j2\pi f_l \tau_m^t + j\phi^t(m,l)} \quad (4)$$

其中, $a(\theta_m) = \left[1, e^{-\frac{j2\pi d \cos(\theta_m)}{\lambda}}, e^{-\frac{j2\pi 2d \cos(\theta_m)}{\lambda}}, \dots\right]$,

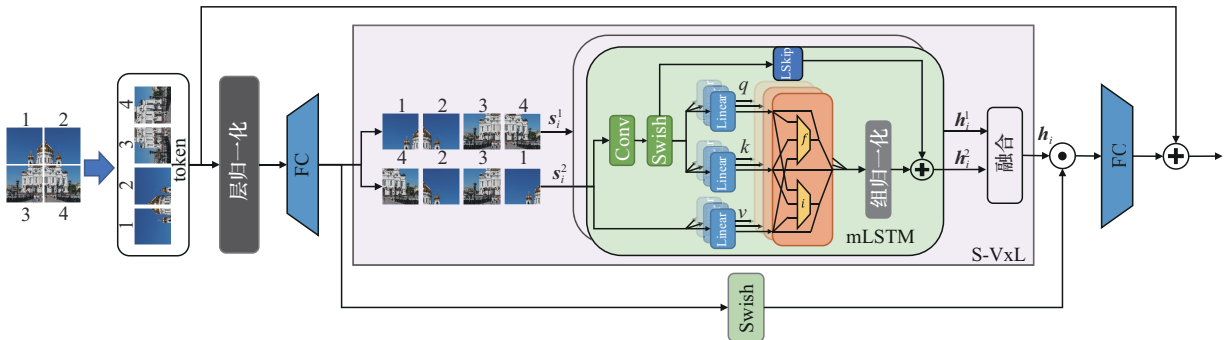


图2 S-VxL 模块的计算流程



$e^{-\frac{j2\pi(N_t-1)d\cos(\theta_m^t)}{\lambda}}$ 表示均匀线性阵列的阵列响应向量 (d 为天线间隔距离), $\phi^l(m, l)$ 为信号传播

过程中产生的附加相位。因此 t 时刻基站与用户之间的三维 CSI 矩阵 $\mathbf{H}(t) \in \mathbb{C}^{N_t \times N_t \times N_f}$ 可以表示为:

$$\mathbf{H}(t) = \begin{bmatrix} h_{1,1}^t[1, 2, \dots, N_f] & h_{1,2}^t[1, 2, \dots, N_f] & \dots & h_{1,N_t}^t[1, 2, \dots, N_f] \\ h_{2,1}^t[1, 2, \dots, N_f] & h_{2,2}^t[1, 2, \dots, N_f] & \dots & h_{2,N_t}^t[1, 2, \dots, N_f] \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_t,1}^t[1, 2, \dots, N_f] & h_{N_t,2}^t[1, 2, \dots, N_f] & \dots & h_{N_t,N_t}^t[1, 2, \dots, N_f] \end{bmatrix} \quad (5)$$

为实现一种更高效的CSI预测,本文使用基于标量长短期记忆 (scalar long short-term memory, sLSTM) 网络结构^[14]的 xLSTM-C 块作为 MIMO-CSI 预测模块的骨干,通过引入改进的门控更新规则,实现更灵活的存储决策。图3

展示了 MCP 工作流程,通过 N 个 xLSTM-C 块串行递进,实现对过去时隙规律的挖掘。假设使用前 T 个时隙的 CSI 来预测未来 L 个时隙的 CSI,则模型的输入、输出可以分别表示为:

$$\begin{cases} \mathbf{H}_N^{(T)}(t) = [\mathbf{H}(t-T+1), \mathbf{H}(t-T), \dots, \mathbf{H}(t)]_{\mathbb{C}^{N_t \times N_t \times N_f \times T}} \\ \bar{\mathbf{H}}_N^{(L)}(t) = [\mathbf{H}(t+1), \mathbf{H}(t+2), \dots, \mathbf{H}(t+L)]_{\mathbb{C}^{N_t \times N_t \times N_f \times L}} \end{cases} \quad (6)$$

MCP 按照串行顺序从左到右依次处理输入的 CSI。为了实现从短期到长期的多尺度特征提取, MCP 采用渐进式分块策略,对前 T 个时隙的 CSI

矩阵 $\mathbf{H}_N^{(T)}(t)$ 进行时间窗口划分。对于第 k 个块 ($k \in [1, N]$), 其处理的时间窗口 CSI 矩阵 $\mathbf{H}_k^{(T)}(t)$ 可表示如下。

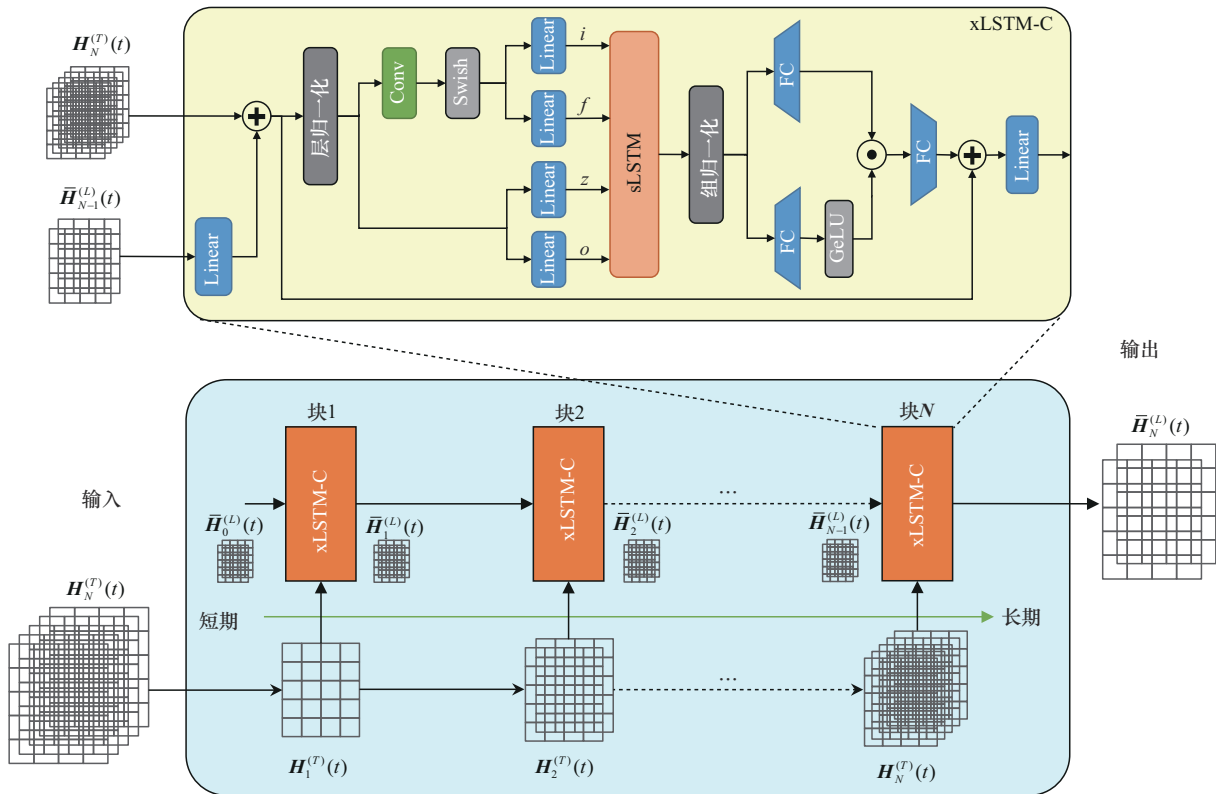


图3 MCP工作流程

$$\mathbf{H}_k^{(T)}(t) = \left[\mathbf{H}\left(t - k \cdot \left\lceil \frac{T}{N} \right\rceil + 1\right), \mathbf{H}\left(t - k \cdot \left\lceil \frac{T}{N} \right\rceil\right), \dots, \mathbf{H}(t) \right]_{\mathbf{C}^{N_r \times N_t \times N_f \times k \cdot \left\lceil \frac{T}{N} \right\rceil}} \quad (7)$$

这种设计使得每个后续块的时间窗口长度都比前一个块增加 $\lceil T/N \rceil$ 个时隙，从而使得更靠后的xLSTM-C块能够观察和学习到更长时间跨度的CSI变化规律。本文通过使用上取整 T/N 来确定基础窗口长度，保持了时间特征提取的连续性。

在特征传递方面，每个xLSTM-C块接收两个输入：一是来自前一个块的输出 $\bar{\mathbf{H}}_{k-1}^{(L)}(t)$ ，用于继承已提取的CSI特征；二是对应时间窗口的CSI矩阵 $\mathbf{H}_k^{(T)}(t)$ ，用于学习特定时间尺度上的特征。对于第一个块，由于没有前序输出，使用零矩阵 $\bar{\mathbf{H}}_0^{(L)}(t)$ 和其相应的CSI矩阵 $\mathbf{H}_1^{(T)}(t)$ 作为输入。最后一个块的输出 $\bar{\mathbf{H}}_N^{(L)}(t)$ 即模型对未来 L 个时隙CSI的预测结果。每个xLSTM-C块通过两个线性层保持输出与最终预测结果相同的维度，确保特征传递的一致性。

这种分块递进的设计使MCP预测结果能够融合不同时间尺度下的CSI特征：前面的块主要学习短期快速变化的特征；而后面的块通过累积前序特征并结合更长时间窗口的数据，提取长期稳定的CSI特征。多尺度特征学习可显著提升模型对CSI变化的理解和预测能力。

当得到未来 L 个时隙的CSI预测矩阵后，为减小MIMO系统信道之间的相互干扰，本文考虑使用基于接收端信道估计的MIMO模型。最小化均值误差均衡是通过接收器的后处理减轻信道干扰的一种方法^[16]。当信道存在噪声时，均衡器可以找到一个三维矩阵 $\mathbf{G}$ 使接收信号 $\hat{\mathbf{y}}$ 和源图像与重建图像的误差 e 相关性为0：

$$\mathbf{G} = \mathbf{H}^H \left(\mathbf{H} \mathbf{H}^H + \frac{\sigma_n^2}{P} \mathbf{I} \right)^{-1} \quad (8)$$

其中， P 为发送功率， σ_n^2 为MIMO信道的噪声方差， $\mathbf{I}$ 为单位矩阵， $\mathbf{G}$ 为计算得到的最佳匹配滤

波器。由于多个天线之间存在干扰，直接使用CSI来计算SNR往往不够准确，因此本文使用SINR来估计信道质量。第 j 个发射天线表示为 TX_j ，其接收到的SINR可以计算为：

$$\text{SINR}(\text{TX}_j) = \frac{\|\mathbf{g}_j \mathbf{h}_j \mathbf{x}_j\|^2}{\sum_{k \neq j} \|\mathbf{g}_j \mathbf{h}_k \mathbf{x}_j\|^2 + \|\mathbf{g}_j \mathbf{N}_j\|^2} = \frac{P \|\mathbf{g}_j \mathbf{h}_j\|^2}{P \sum_{k \neq j} \|\mathbf{g}_j \mathbf{h}_k\|^2 + \sigma^2 \|\mathbf{g}_j\|^2} \quad (9)$$

其中， $\mathbf{g}_j \in \mathbf{C}^{1 \times N_r \times N_f}$ 表示矩阵 $\mathbf{G}$ 沿第一维的第 j 个二维切片， $\mathbf{h}_j \in \mathbf{C}^{N_r \times 1 \times N_f}$ 表示矩阵 $\mathbf{H}$ 沿第二维的第 j 个二维切片， $\mathbf{h}_k \in \mathbf{C}^{N_r \times 1 \times N_f}$ 表示矩阵 $\mathbf{H}$ 沿第二维与 j 不相同的第 k 个二维切片， $\mathbf{x}_j$ 表示从天线 TX_j 发送的信号信息， $\mathbf{N}_j \in \mathbf{C}^{N_r \times 1 \times N_f}$ 表示噪声矩阵 $\mathbf{N}$ 沿第二维的第 j 个二维切片， P 和 σ^2 分别代表天线 TX_j 的功率和噪声的方差。由于 P 和 σ^2 是常数，SINR受发射天线 TX_j 和信道矩阵 $\mathbf{H}$ 的影响，因此，进行特征分配的信道选择时要优先考虑信道质量更好的信道，避免重要语义信息受到信道的影响。

为确保特征分配的鲁棒性，本文考虑MIMO系统的时域与频域变化特性，引入MIMO信道分配块来表征一段时间MIMO信道的时间与空间上的质量关系。首先，根据 $\bar{\mathbf{H}}^{(L)}(t)$ 计算对应 L 个时隙的最佳匹配滤波器 $\bar{\mathbf{G}}^{(L)}(t) \in \mathbf{C}^{N_t \times N_r \times N_f \times L}$ ，计算对应时隙的信道分配块矩阵 $\bar{\mathbf{Q}}^{(L)}(t) \in \mathbf{C}^{N_t \times L \times N_f}$ ，如下。

$$\bar{\mathbf{Q}}^{(L)}(t) = \begin{bmatrix} \bar{\mathbf{q}}_{1,t+1} & \gamma \bar{\mathbf{q}}_{1,t+2} & \cdots & \gamma^{L-1} \bar{\mathbf{q}}_{1,t+L} \\ \bar{\mathbf{q}}_{2,t+1} & \gamma \bar{\mathbf{q}}_{2,t+2} & \cdots & \gamma^{L-1} \bar{\mathbf{q}}_{2,t+L} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\mathbf{q}}_{N_t,t+1} & \gamma \bar{\mathbf{q}}_{N_t,t+2} & \cdots & \gamma^{L-1} \bar{\mathbf{q}}_{N_t,t+L} \end{bmatrix} \quad (10)$$



其中, 一维向量 $\bar{q}_{j,i}$ ($j=1,2,\dots,N_t, i=1,2,\dots,L$) 表示第 i 个时隙发送天线 TX_j 对应所有子载波上的 SINR。由于预测精度会随预测步长增加而下降, 因此引入衰减因子 γ 以减小多步预测的影响。信道分配块能够捕捉信道在频域的多天线分集特性和时间维度上的时变特性, 从而关联不同时间点和天线的信道质量。通过这种方式, 系统能够更好地应对信道条件的波动, 确保特征分配的准确性。最后, 矩阵 $\bar{Q}^{(L)}(t)$ 被展平为 SINR 向量 $\mathbf{u} \in \mathbf{R}^D$, 其中 $D=N_t \times L \times N_f$ 。展平操作采用时间维度优先顺序, 即首先沿矩阵 $\bar{Q}^{(L)}(t)$ 的第二维度逐行展平, 然后依次展平剩余维度。此操作旨在将三维数据线性化, 以便后续处理。由于向量 $\mathbf{u}$ 将传递至发送端用于辅助调整对语义特征, 发送端需要按 SINR 降序排列向量 $\mathbf{u}$, 并获取对应的索引。考虑 MIMO 信道中天线数与载波数较多导致 D 值较大, 且引入衰减因子, 多个时间步后的预测 SINR 可能较低, 因此采用时间维度优先顺序展平可以减少对向量 $\mathbf{u}$ 的排序时间, 从而避免了排序对后续优化带来的干扰。

1.5 基于信道与语义重要性的特征分配模块

SISO 场景通常使用一条信道进行传输。然

而, 在 MIMO 场景中, 需要考虑多天线信道质量对特征传输的影响。因此, 在分析 MIMO 信道时变特性的基础上将探讨如何利用这些预测信息优化信道传输效率, 以实现时变 MIMO 信道下的特征分配与信道适应。

1.5.1 特征语义重要性

在语义通信系统中, 不同语义特征对系统性能的影响程度各不相同。为量化这些特征的重要性, 本文采用梯度加权类激活映射 (gradient-weighted class activation mapping, Grad-CAM) 算法, 通过分析系统损失相对于每个特征的梯度来计算其重要性^[17]。损失函数反映了系统的实际性能, 而梯度则衡量了特征对损失的影响。因此, 梯度大小直接表明特征在整个系统中的重要性。基于梯度的语义重要性计算如图4所示, 首先, 计算使用语义提取模型得到的特征张量 $\mathbf{y}'$ (将维度简化为 $h \times w \times c$), 对于其第 k 个特征 $\mathbf{Y}_k$ 计算其关于系统损失 $\mathcal{L}(x, \hat{x})$ 的梯度矩阵:

$$\mathbf{T}_k = \nabla_{\mathbf{Y}_k} \mathcal{L}(x, \hat{x}) = \left[\frac{\partial \mathcal{L}(x, \hat{x})}{\partial y_{k,ij}} \right] \quad (11)$$

其中, $\mathbf{T}_k \in \mathbf{R}^{h \times w}$, $y_{k,ij}$ 为 $\mathbf{Y}_k$ 的第 i 行第 j 列元素。通过对得到的 $\mathbf{T}_k$ 进行全局平均池化 AP($\cdot$) 可以将其变为一维向量, 得到 $\mathbf{Y}_k$ 的语义重要性:

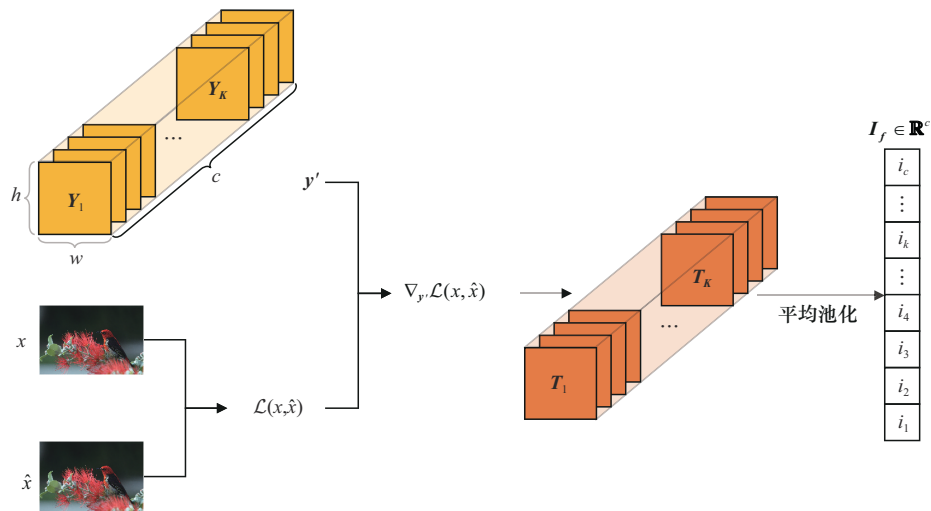


图4 基于梯度的语义重要性计算

$$i_k = \text{AP}(\mathbf{T}_k) = \frac{1}{hw} \sum_{i=0}^{h-1} \sum_{j=0}^{w-1} \frac{\partial \mathcal{L}(x, \hat{x})}{\partial y_{k,ij}} \quad (12)$$

最后关于特征张量 $\mathbf{y}'$ 的重要性分布向量 $\mathbf{I}_f$ 可以通过排列表示为:

$$\mathbf{I}_f = [i_1, i_2, \dots, i_c]^T \quad (13)$$

1.5.2 基于时变信道的特征分配

复杂多变的 MIMO 信道会在时域和频域上波动。采用简单随机特征分配方法可能会导致关键特征被分配到信道质量较差的子通道,从而遭遇信道衰落,导致重要特征损坏,进而引发整体系统性能显著下降。

为解决 MIMO 系统中时域频域信道波动所带来的影响,本文设计了一个时变信道下的特征分配模块。分配器结合 $\mathbf{u}$ 和进行特征的时域频域分配。信道特征分配模块如图 5 所示。特征分配模块遵循特定的操作顺序,首先,接收来自接收端的信道状态向量 $\mathbf{u}$,得到当前 MIMO 系统中未来 L 个时隙的信道状态 $c_i (i=1, 2, \dots, D)$;然后将得到的 c_i 和重要性分布向量 $\mathbf{I}_f$ 进行降序排列,得到两个索引排序 I_c 和 $\hat{I}_f$ 。其中,元素值越大在索引向量中的排名越高;随后,分配器通过匹配索引向量中的元素,将相等级级的信道与特征对应,此过程将每个特征分配到合适的传输时隙,重新组织特征顺序;最终,应用新的顺序至原始特征张量,生成维度为 $\mathbf{R}^{N_t \times N_f \times L}$ 的二进制向量掩码 M ,用于表征特征的传输情况,没有传输特征的位置

以 0 填充,如下:

$$M = F(\mathbf{u}, \mathbf{I}_f, \mathbf{y}') = \text{Match}(\text{Sort}(\mathbf{u}, \mathbf{I}_f), \mathbf{y}') \quad (14)$$

其中, $\text{Match}(\cdot)$ 和 $\text{Sort}(\cdot)$ 分别表示匹配重建函数和排序函数。

信道带宽比 (channel bandwidth ratio, CBR) [18]

用于表示传输数据与信道带宽之间的关系,可确定每个源符号有多少信道符号可用,通常被定义为: $\text{CBR} = K/n$, 其中 K 为信道带宽, n 为数据传输量。通常情况下 $\text{CBR} < 1$ 。为适应动态变化的信道质量,有效地调整编码器输出的数据带宽大小,可根据链路质量动态调节 CBR: 通过裁剪不重要的特征,以实现 CBR 与信道质量的匹配。在接收端,信道特征重建器利用该向量掩码,从对应时隙的信息中解析对应的特征向量,并将其按原有顺序重排,重建出原始特征张量。

1.6 基于信道预测的时变信道自适应编码

受文献[5]的启发,本框架中集成了无须依赖梯度下降的自适应信道编码模块。模块在编解码器中实现,通过分通道方式调制信道特征分配模块输出,并根据预测的 MIMO-SINR 调制相应的特征向量。

信道自适应编码模块结构如图 6 所示。在发送端,首先将经过信道特征分配后的语义特征 $\mathbf{y}'$ 输入该模块,通过多个串行的特征调制层结合 MCP 模块的预测信息 $\mathbf{u}$, 实现与信道条件的自适应。在接收端,接收到的信道符号 $\hat{\mathbf{y}}$ 同样经过该模

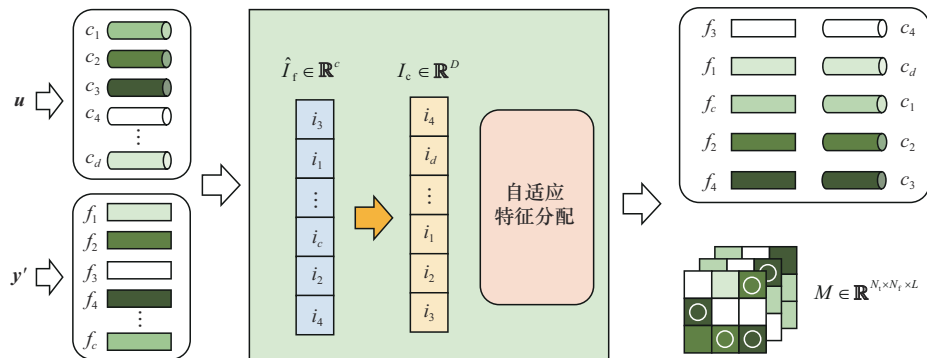


图5 信道特征分配模块

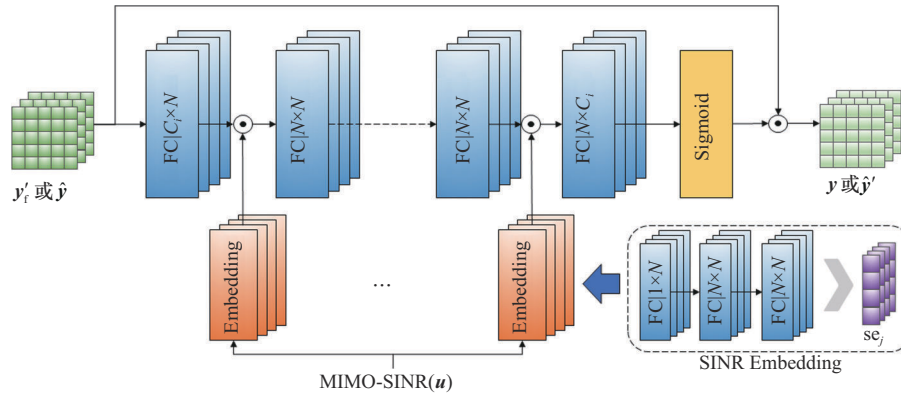


图6 信道自适应编码模块结构

块的逆向调制过程，逐步解耦并恢复出对应子信道的语义特征 $\hat{y}'$ ，从而完成端到端的语义特征重建。

信道自适应编码模块结构如图6所示，包括信道状态嵌入调制和特征调制两部分。模块由多个并行FC层和嵌入层交替组成。每个信道对应一个FC层和嵌入层。嵌入层通过三层全连接结构将信道的SINR转换为N维向量，并按顺序对子信道进行调制。在第*i*个信道上从 c_i 到 $se_{i,j}$ 的映射过程如下。

$$se_{i,j}^{(1)} = \text{ReLU}\left(W^{(1)} \cdot c_i + b^{(1)}\right) \quad (15)$$

$$se_{i,j}^{(2)} = \text{ReLU}\left(W^{(2)} \cdot se_{i,j}^{(1)} + b^{(2)}\right) \quad (16)$$

$$se_{i,j} = \text{Sigmoid}\left(W^{(3)} \cdot se_{i,j}^{(2)} + b^{(3)}\right) \quad (17)$$

其中，Sigmoid为激活函数，ReLU为修正线性单元激活函数， $W^{(k)}$ 和 $b^{(k)}$ 分别为映射函数的可学习矩阵及其偏置。因此，信道状态信息与 $se_{i,j}$ 相关联。

2 仿真实验与分析

2.1 实验设置

(1) 数据集。VxLJSCC模型使用DIV2K数据集^[19]进行训练。图像被随机裁剪为256×256大小的块。使用Kodak数据集(521×768)和CLIC2021测试集(近似2K分辨率)来评估性能。

(2) 比较方案。将提出的VxLJSCC与SwinJSCC^[5]、MambaJSCC(w/o CA)^[6]、DeepJSCC-

MIMO^[10]以及基于信源信道的编码方案进行比较。对于传统分离编码方案，即基于更好的可移植图形(better portable graphics, BPG)图像压缩和低密度奇偶校验(low density parity check, LDPC)编码方案(BPG+LDPC)，使用BPG作为信源编码器，使用LDPC^[20]作为信道编码器。对于传统联合编码方案，使用双LDPC(double LDPC, D-LDPC)^[21]作为信源信道联合编码器。由于SwinJSCC和MambaJSCC基于SISO，而DeepJSCC-MIMO基于MIMO，因此实验分别在SISO和MIMO场景下进行。对于SwinJSCC，使用无SNR调整模块和速率调整模块的版本SwinJSCC w/o SA&RA作为对比，对于MambaJSCC，使用无信道调整模块的版本MambaJSCC w/o CA作为对比。本文将包含时变特征分配和信道自适应编码的方法称为“VxLJSCC w/ FA&CA”，仅具有时变特征分配的称为“VxLJSCC w/ FA”，仅有信道自适应的称为“VxLJSCC w/ CA”，基线版本称为“VxLJSCC w/o FA&CA”，其中“FA”和“CA”分别表示时变特征分配和信道自适应。实验过程中，由于其余方法并没有设计对应的信道感知模块，因此测试过程中所有方法均使用基于MCP预测的CSI来指引方法的调整和分配。

(3) 评估指标。本文使用像素指标峰值信噪比(peak signal-to-noise ratio, PSNR)^[22]和感知指标多尺度结构相似性(multi-scale structural simi-

larity, MS-SSIM) 进行评估^[23]。当使用 PSNR 时, 本文通过计算原始图像和重建图像的均方误差来优化模型损失; 当使用 MS-SSIM 时, 通过损失函数 $d=1-\text{MS-SSIM}$ 来优化模型。由于 MS-SSIM 的值通常高于 0.9, 因此本文使用 $\text{MS-SSIM}(\text{dB})=-10\lg(1-\text{MS-SSIM})$ 来表示。同时使用乘法累加运算 (multiply-accumulate operation, MAC)、参数量和推理时延 (inference delay, ID) 来评估模型的复杂度、大小与推理时间。

(4) 模型训练细节。VxLJSCC 采用两阶段的结构 $[N_1, N_2]=[4, 8/12]$, $[C_1, C_2]=[128, 256]$ 。模型在 Pytorch 2.3.1 中使用两个 RTX 4090 GPU 实现, 利用 Adam 优化器, 学习率为 5×10^{-5} 。MIMO 信道使用 8 个信道来进行训练, 平均 SNR 从 1 dB 到 18 dB 均匀分布。

(5) 信道环境。本文主要针对 3 种信道场景进行仿真模拟——加性高斯白噪声 (additive white Gaussian noise, AWGN) 信道、Rayleigh 信道和 5G NR 标准中的 CDL-C 信道。式 (5) 定义了时刻 t 的 CSI 矩阵 $\mathbf{H}(t)$, 其中每个元素 $h'_{i,j}[l]$ 表示第 i 个接收天线与第 j 个发送天线之间的信道增益, 且在子载波 l 处有所不同。对于 AWGN 信道模型, 信道增益 $h'_{i,j}[l]$ 可以被视为常数 1 dB, 信道的变化主要受到如式 (1) 定义的噪声 N 的影

响, 在 AWGN 信道中 N 通常表现为高斯随机分布; 而在 Rayleigh 衰落信道下, 信道增益 $h'_{i,j}[l]$ 是由多径传播引起的, 通常呈现为 Rayleigh 分布的随机值, 主要影响体现在式 (4); 当使用 CDL-C 信道时, 为方便生成对应的信道增益矩阵 $\mathbf{h}'_{i,j}[l]$, 本文使用 MATLAB 自带的 5G NR 套件输出对应的 $\mathbf{h}'_{i,j}[l]$, 信道噪声依旧定义为高斯噪声。因此, 式 (5) 中的 CSI 矩阵 $\mathbf{H}(t)$ 会依据信道模型的不同而有所不同, 从而影响信道增益的表现和信号的传输质量。

2.2 MIMO-CSI 预测分析

为了验证 MCP 的有效性, 使用公开数据集 DeepMIMO^[24] 进行训练, 该数据集基于室外场景。图 7 展示了 DeepMIMO 信道预测场景, 参数设置见表 1。

其中, 信道 CSI 矩阵 $\mathbf{H} \in \mathbf{C}^{N_r \times N_t \times N_f}$, 由于信道为复数信道, 在预测时将实数与虚数部分拆分开进行预测, 实际 $\tilde{\mathbf{H}} \in \mathbf{R}^{N_r \times N_t \times 2N_f}$ 。按照表 1 进行场景设置, 该数据集总共生成 50 000 个样本, 其中 35 000 个用于训练, 15 000 个用于测试。由于一般情况下 $N_t \times N_r \ll K$, 其中 K 为信道带宽, 因此对未来的预测时隙 L 取值为 1 和 2, 而每个样本使用过去 50 步进行预测, 预测衰减因子取 $\gamma=0.05$, MCP 中 xLSTM-C 块数 $N=5$ 。

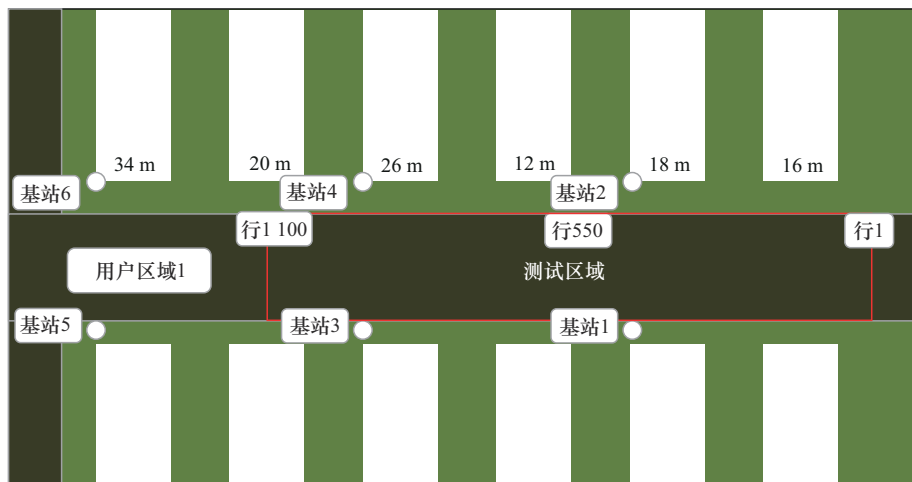


图7 DeepMIMO 信道预测场景



表1 参数设置

参数	取值
基站编号	1
天线间距 (波长)	0.5
信道带宽/MHz	10
基站天线数量 N_t	64
用户天线数量 N_r	1
多径数量	5
用户位置行数	1~1 100
子载波数 N_F	128

为验证 MCP 在准确性和模型规模上的表现, 本文基于 iTransformer^[25]实现了一个与文献[26]类似的 MIMO 信道预测模型, 以比较本模型的实际效果。比较方式参考文献[26]采用归一化均方误差 (normalized mean square error, NMSE) 和余弦相似度 ρ 来验证预测结果的准确性。实验进行多次并取平均值, 预测模型实验结果见表2。

实验结果表明, 基于 xLSTM 的 MCP 在更轻量化的模型规模下实现了更优的信道预测效果, 能够在节省大约 71.48% 参数量的同时达到最高 2.13 dB 的 NMSE 提升, 实现了一种轻量化的高效预测模型。特征分配与调节均在 MCP 的预测基础上实现。

2.3 SISO 系统结果分析

在 SISO 系统中, 由于仅存在一条信道, 本文采用 VxLJSCC w/ CA 和 VxLJSCC w/o FA&CA 作为验证方法, 对比各方法在不同测试条件下的表现。

2.3.1 PSNR 性能对比

不同信道条件下 PSNR 性能对比如图 8 所示,

图 8 展示了本方法在 AWGN 信道和 Rayleigh 信道下与其他两种 DeepJSCC 方法以及传统信源信道编码方式的性能对比。图 8 中的模型均在同一 SNR 下进行训练, 可以看出本方法在 CBR=1/16 时, 无论是在 Kodak 数据集还是 CLIC2021 数据集上的性能表现都要优于作为基线的 4 种方法, 能够提供 1 dB 左右的信道增益。尤其是在低 SNR 环境下, 本方法能够展示出更为优异的重建能力。相较于两种传统信源信道编码方案, 由于深度学习强大的语义信息提取能力, 在不同的测试环境下 DeepJSCC 都取得了更好的表现。值得注意的是, 当测试时的 SNR 低于模型训练时使用的 SNR, VxLJSCC 虽然也像其他 DeepJSCC 模型一样出现了性能退化, 但仍然保持着更好的传输性能。

2.3.2 MS-SSIM 性能对比

为了对本文提出的模型进行更全面的评估, 本文进行了进一步的实验, 实验使用 MS-SSIM 指标来评估其性能, 该指标能获取更好的人类视觉感知。Kodak 数据集在不同信道下 MS-SSIM 性能对比如图 9 所示。

本节对比了两种基于 DeepJSCC 的方法和传统信源信道编码方法在 Kodak 数据集上使用固定 SNR 训练后模型的 MS-SSIM 表现, 可以看到各个方法在 AWGN 信道上的表现都优于 Rayleigh 信道上的, 同时所有基于 DeepJSCC 的方法都有着远超基于传统信源信道编码的性能, 这体现出了 DeepJSCC 方法的优越性。同时在低 SNR 环境下, VxLJSCC w/CA 提供的信道增益要高于 VxLJSCC w/o FA&CA。

表2 预测模型实验结果

模型	MAC/GB	参数量/MB	NMSE /dB		ρ	
			$L=1$	$L=2$	$L=1$	$L=2$
iTransformer	1.72	10.31	22.19	21.24	0.996 7	0.995 2
MCP	0.59	2.94	24.05	23.37	0.998 1	0.997 0

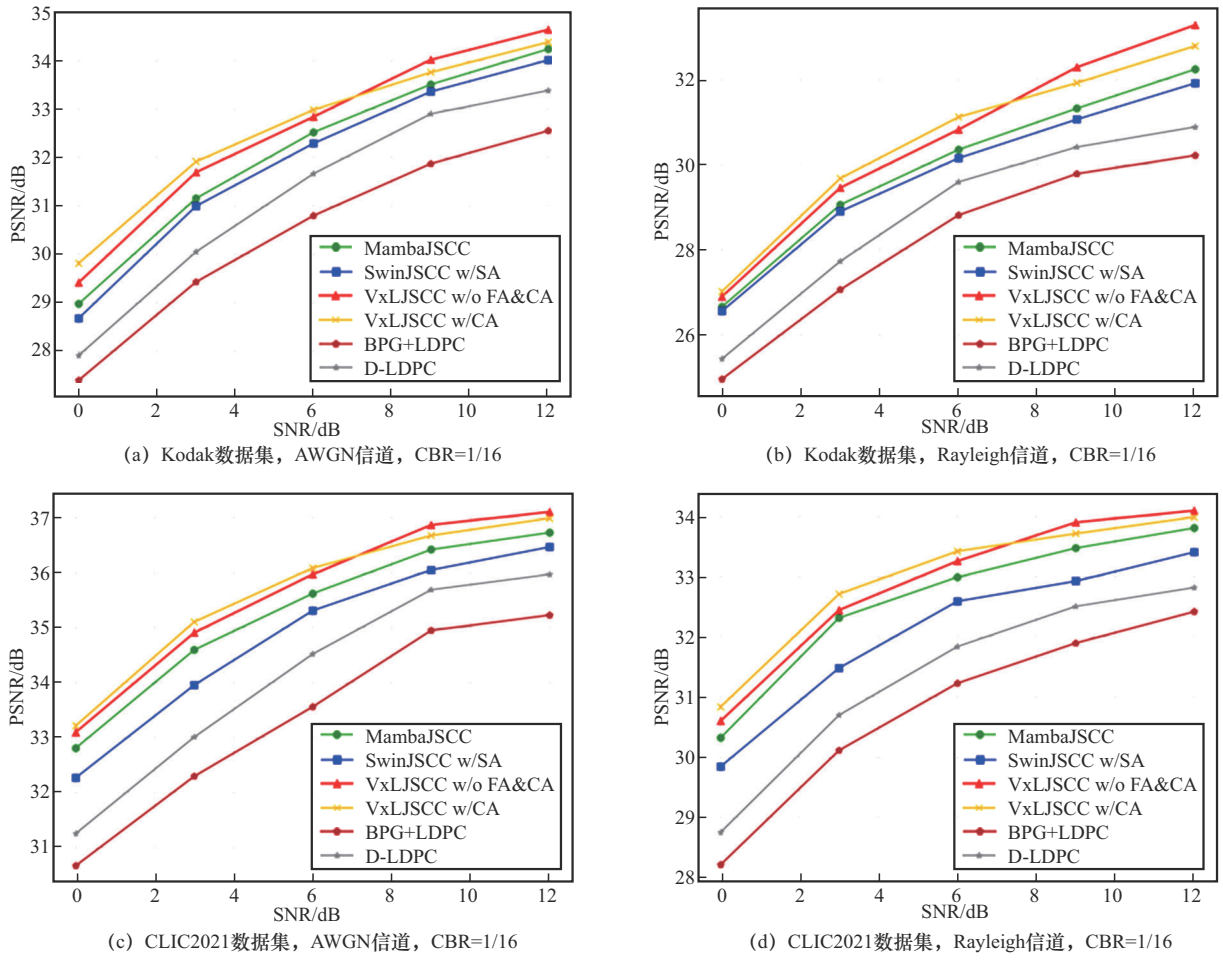


图8 不同信道条件下PSNR性能对比

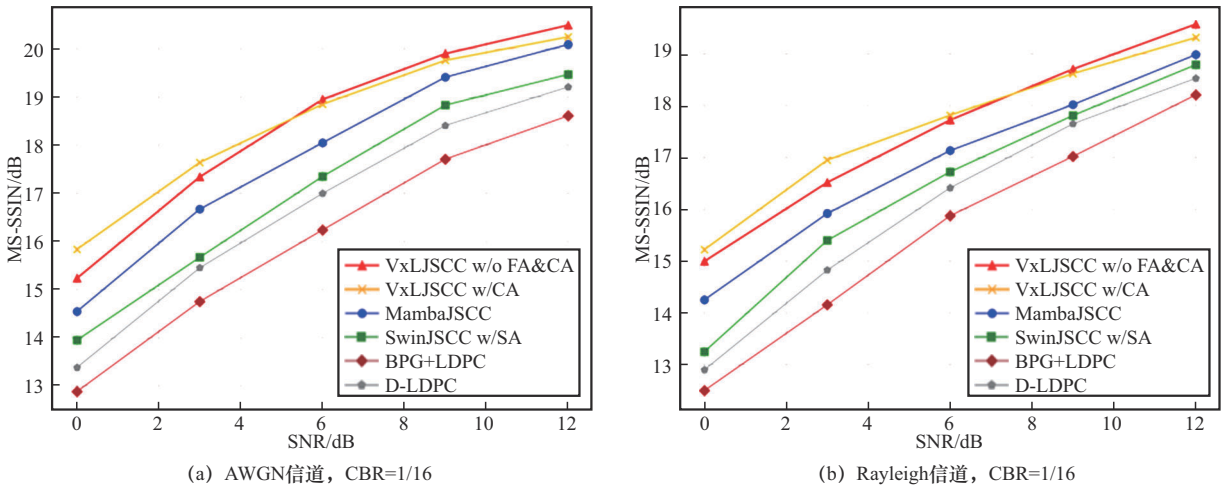


图9 Kodak数据集在不同信道下MS-SSIM性能对比

2.3.3 模型参数与计算量分析

本节主要分析本文所提模型在相同条件下与其他模型在参数量等方面的对比,以研究本方法

在性能和轻量化上的优势。

本节评估了这些方案在不同信道下的传输性能。在Kodak数据集上进行10次实验,批量大小



为1, 以计算每张图像的平均编解码时间。实验验证 VxLJSCC 通过两阶段特征提取即可达到效果, 对比方法则按原有设计使用四阶段。针对 SwinJSCC 和 MambaJSCC, 按如下设置进行复现: $[N_1, N_2, N_3, N_4] = [2, 2, N_s / N_m, 2]$, $[C_1, C_2, C_3, C_4] = [128, 192, 256, 320]$, 其中 N_s 和 N_m 分别表示两种方法的块大小。SISO 下模型参数与性能对比见表3。训练和测试环境的 SNR 均为 6 dB, 使用 1/16 的 CBR 进行压缩。如表3所示, 本节对比了不同块结构下各方法在相同信道下的模型规模与性能, 由于 VxLJSCC 基于线性运算, 其参数量和计算开销相比基于 Transformer 的 SwinJSCC 有明显减少。同时, 由于 VxLJSCC 仅需两阶段的模型训练, 即使是同为线性计算空间的 MambaJSCC, 在轻量化方面也表现出更好的效果。此外, 尽管引入 CA 模块增加了额外的参数量和计算开销, 但带来了性能提升, 尤其在 SNR 低于训练 SNR 的情况下更加明显。不论是否使用 CA 模块, VxLJSCC 都能获得较大的性能增益。如在 AWGN 信道下, 最优的 VxLJSCC 可以实现 0.53 dB 的平均信道增益, 因此本方法在显著降低计算复杂度和开销的同时, 能够有效提升性能。

2.4 MIMO 系统结果分析

本文在基于 MIMO 的模型中使用 VILJSCC w/o FA、VILJSCC w/ CA、VILJSCC w/o FA&CA 和

VILJSCC w/o FA&CA 作为对比方法, 主要对比在不同的 CBR 下的 PSNR 表现情况, 由于存在多条子信道, 取多条子信道的平均 SNR 作为实验结果, 训练和测试 SNR 均为 13 dB。

2.4.1 PSNR 性能对比

Kodak 数据集在不同 CBR 下的 PSNR 性能如图10所示。本文在 AWGN、Rayleigh 和 CDL-C3 种信道下对 Kodak 数据集进行了训练和测试, 主要对比本文所提模型与 DeepJSCC-MIMO、BPG+LDPC 和 D-LDPC 的实际性能。可以看出, 在相同的信道环境和 CBR 约束下, 本文提出的4种方法均表现出更优的性能, 均超过了 DeepJSCC-MIMO 方法。具体而言, 如图10(a)、图10(d)和图10(g)所示, 在 CBR 为 1/8, 信道平均 SNR=1 dB 时, 本文方法相比 DeepJSCC-MIMO 实现了 1.28 dB、1.575 dB 和 1.583 dB 的性能增益, 而在如图10(c)、图10(f)和图10(i)所示的 CBR 为 1/48 条件下, 本文方法获得了 2.394 dB、2.14 dB 和 2.972 dB 的性能提升。

如图10所示, 当 SNR 和 CBR 均处于较低水平时, 本文基于 FA 的两种方法表现出显著优于现有 JSCC 和基于分离编码的最优方法的性能。这是因为在 CBR 较低的情况下, 保留的特征对于图像重建尤为重要, 通过将这些关键特征映射到更好的信道中, 进一步降低了信道条件对图像重建产生的负面影响。

表3 SISO 下模型参数与性能对比

模型	块大小	参数量/MB	MAC /GB	ID /ms	PSNR (AWGN) /dB	PSNR (Rayleigh) /dB
SwinJSCC	$N_s=2$	12.27	26.34	56.72	26.89	26.01
	$N_s=6$	19.01	33.49	73.56	27.16	26.04
MambaJSCC	$N_m=6$	9.72	15.75	30.25	27.31	26.17
	$N_m=12$	16.59	24.67	46.38	27.46	26.20
VxLJSCC w/o FA&CA	$N_v^1, N_v^2=4, 8$	9.49	7.15	35.56	27.67	26.33
	$N_v^1, N_v^2=4, 12$	12.66	9.41	42.02	27.85	26.39
VxLJSCC w/ CA	$N_v^1, N_v^2=4, 8$	9.49+3.15	7.15+3.97	77.43	27.83	26.35
	$N_v^1, N_v^2=4, 12$	12.66+3.15	9.41+3.97	81.69	27.99	26.43

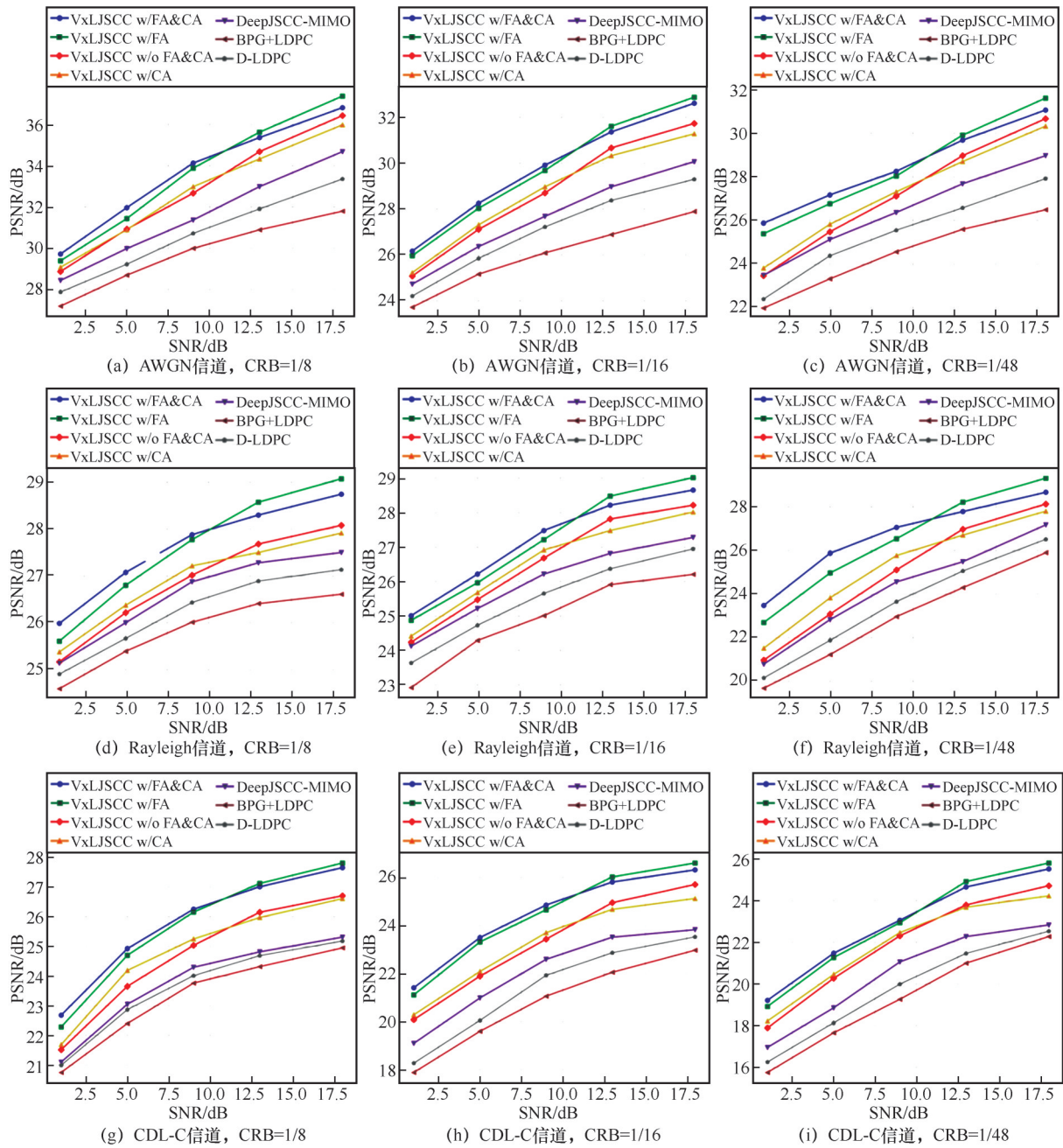


图10 Kodak数据集在不同CBR下的PSNR性能

2.4.2 MS-SSIM性能对比

Kodak数据集在不同CBR下的MS-SSIM性能如图11所示。由图11可知，所提出的模型在CDL-C信道下的性能优于其他方法，尤其在低CBR区域表现出更多性能提升。基于学习的方案在性能上优于传统信源信道编码方法，因为BPG压缩是通过手工设计的约束来优化平方误差，而非直接优化感

知质量。此外，使用FA的VxLJSCC模型的性能优于其他模型，展现出更优异的图像重建效果。

由图11可以看出，当SNR和CBR均处于较低水平时，本文基于时变特征分配的两种方法显著优于现有DeepJSCC方法和传统信源信道分离和联合编码方法的性能。这是因为在CBR较低的情况下，保留的特征对于图像重建尤为重要，这

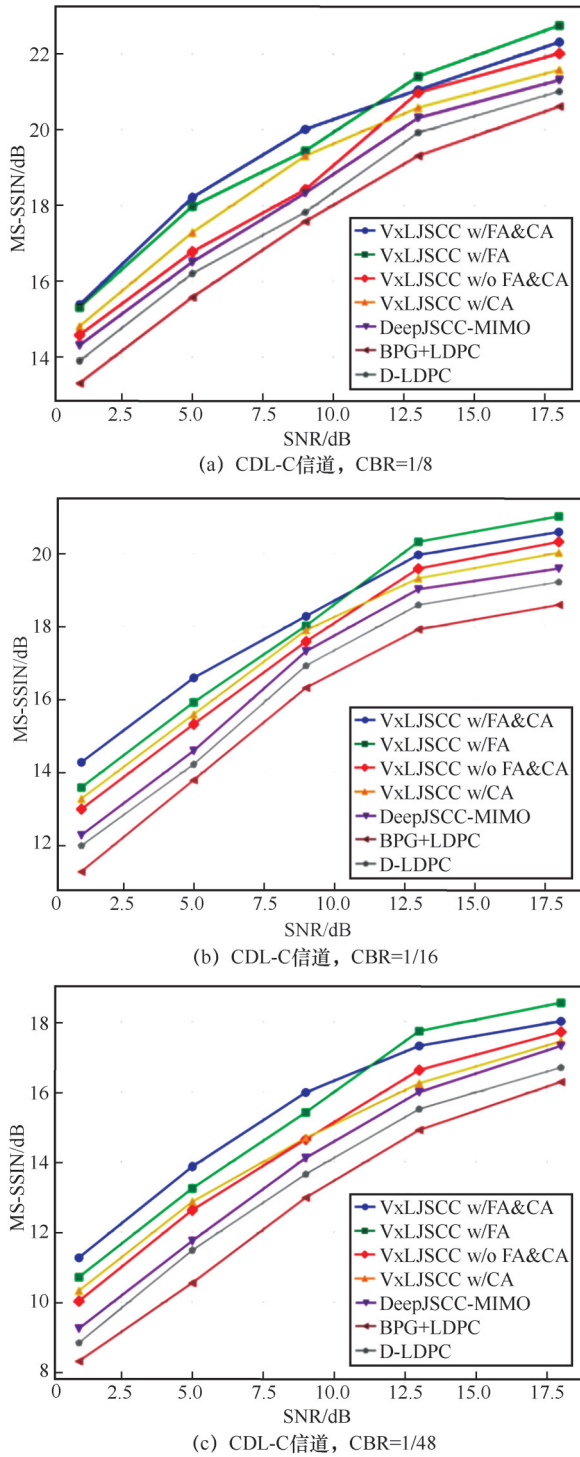


图11 Kodak数据集在不同CBR下的MS-SSIM性能

些关键特征映射到更好的信道中,进一步降低了信道条件对图像重建产生的负面影响。

2.4.3 模型参数与计算量分析

MIMO场景下本方法与DeepJSCC-MIMO模

型的性能对比见表4。针对DeepJSCC-MIMO,本节按照如下设计进行复现: $[N_1, N_2, N_3, N_4] = [2, 2, N_d, 2]$, $[C_1, C_2, C_3, C_4] = [128, 192, 256, 320]$, 其中 N_d 表示可变块大小,以区分不同的模型结构。使用1/16的CBR进行压缩。由表4可知,本文提出的模型在相同测试条件下最多节省61.67%的模型存储开销和77.86%的计算开销,并且实现了1.41 dB的信道增益。随着块数的增加,本文提出的方法在模型规模和计算量的节省方面更加明显。这主要是因为基于Transformer的方法在处理更长的序列或更大的输入尺寸时,表现出二次计算复杂性,限制了此类模型在实际应用中的部署可能性。

本文提出的4种模式中, VxLJSCC w/FA&CA实现了最高的信道增益,相较于对比方法,在AWGN信道下可获得1.99 dB的增益,同时模型存储和计算开销也显著降低,这种性能提升主要源于两个方面:一方面是MCP提供的MIMO信道CSI预测,可用于指导VxLJSCC动态调整信道;另一方面,FA和CA模块在MCP的引导下,将重要特征分配至质量较好的子信道和时隙,并根据信道条件对特征进行调整。这两方面的优化共同提升了MIMO场景下的整体性能。

2.5 MCP对模型效果的性能分析

本节主要讨论MCP对VxLJSCC模型的性能影响,比较方案如下。

(1) VxLJSCC w/o FA&CA (NC): 不使用任何基于信道感知的调整策略。

(2) VxLJSCC (PC): 使用MCP指导VxLJSCC工作。

(3) VxLJSCC (AC): 假定已知未来MIMO-CSI的准确值,将其作为理想的基线方法。

CLIC2021数据集上MCP对模型效果的性能影响如图12所示。所有方法都取得了较高PSNR,展示了出色的信道适应性。在MIMO系统中,FA

表4 MIMO场景下本方法与DeepJSCC-MIMO模型的性能对比

模型	块大小	参数量/MB	MAC /GB	ID /ms	PSNR (AWGN) /dB	PSNR (Rayleigh) /dB
DeepJSCC-MIMO	$N_d=6$	18.85	28.57	84.37	26.87	25.21
	$N_d=12$	24.76	32.30	92.20	27.05	25.39
VxLJSCC w/ FA	$N_v^1, N_v^2=4, 8$	9.49	7.15	35.58	28.80	26.75
	$N_v^1, N_v^2=4, 12$	12.66	9.41	42.04	28.87	26.93
VxLJSCC w/ FA&CA	$N_v^1, N_v^2=4, 8$	9.49+3.15	7.15+3.97	95.25	28.99	26.96
	$N_v^1, N_v^2=4, 12$	12.66+3.15	9.41+3.97	101.80	29.04	27.09

注：“VxLJSCC w/o FA&CA”和“VxLJSCC w/ CA”结果与表3类似，不在此展示

模块通过将重要特征分配到质量较好的子信道，显著提升了VxLJSCC的整体性能。与不使用信道预测的VxLJSCC w/o FA&CA (NC) 方法相比，所有基于信道预测的方法均取得了显著的性能提升。可以看出，使用MCP的VxLJSCC(PC)方法在性能上略低于基于准确信道质量的VxLJSCC(AC)方法。这一差异源于MCP无法避免预测误差，导致一定的性能损失。然而，考虑实际中完美CSI难以获取^[12]，本方法通过牺牲部分准确性换取了更快的响应时间，具有重要的实际意义。

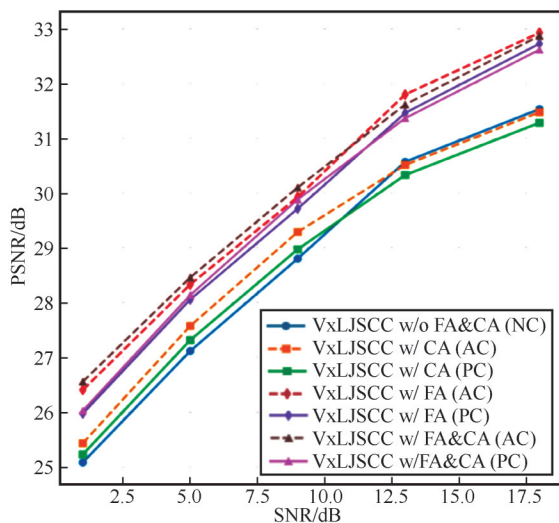


图12 CLIC2021数据集上MCP对模型效果的性能影响
(Rayleigh信道, CBR=1/16, SNR=13 dB)

此外，当实际信道条件低于训练时的信道条件时，CA模块能够提供更优的重建效果，因为它能够根据CSI调整特征以适应当前环境。然而，当信道质量与训练条件相似或更好时，CA

模块可能引入特征损失，导致性能下降。这种变化反映了模型在不同信道条件下的适应性与调整机制之间的权衡。

3 结束语

本文提出了一种用于MIMO系统无线图像传输的轻量化的DeepJSCC框架VxLJSCC。VxLJSCC使用xLSTM来学习图像的语义表示，同时进行更为精准且轻量化的MIMO-CSI时空预测，将预测得到的信道信息集成到框架中，以适应动态变化的MIMO信道场景。此外，本文将重要语义特征分配至质量较好的子信道，充分利用MIMO系统的时域和频域关联性，以识别特征的重要性并合理分配至合适的时空信道中，避免信道质量对语义传输的负面影响。实验结果表明，由于VxLJSCC具有线性计算空间的优势，在保证性能相似甚至更优的情况下，可减少模型大小和计算量。相较于其他DeepJSCC方法，VxLJSCC提高了模型在实际部署中的可行性。同时，基于MIMO预测的特征分配方案，避免了使用过时CSI或错误信道分配方式造成的性能下降问题。因此，VxLJSCC为MIMO系统中构建轻量化且高性能的语义通信提供了前景广阔的解决方案。未来的研究将进一步探讨预测对推理时延的影响和信道自适应编码模块在不同信道环境下的泛化能力。



参考文献:

- [1] 江沸波, 彭于波, 董莉. 面向6G的深度图像语义通信模型[J]. 通信学报, 2023, 44(3): 198-208.
JIANG F B, PENG Y B, DONG L. Deep image semantic communication model for 6G[J]. Journal on Communications, 2023, 44(3): 198-208.
- [2] CALVANESE STRINATI E, BARBAROSSA S. 6G networks: Beyond Shannon towards semantic and goal-oriented communications[J]. Computer Networks, 2021, 190: 107930.
- [3] SHI G M, XIAO Y, LI Y Y, et al. From semantic communication to semantic-aware networking: model, architecture, and open problems[J]. IEEE Communications Magazine, 2021, 59(8): 44-50.
- [4] BOURTSOULATZE E, BURTH KURKA D, GÜNDÜZ D. Deep joint source-channel coding for wireless image transmission[J]. IEEE Transactions on Cognitive Communications and Networking, 2019, 5(3): 567-579.
- [5] YANG K, WANG S X, DAI J C, et al. SwinJSCC: taming swin transformer for deep joint source-channel coding[J]. IEEE Transactions on Cognitive Communications and Networking, 2025, 11(1): 90-104.
- [6] WU T, CHEN Z Y, TAO M X, et al. MambaJSCC: deep joint source-channel coding with visual state space model[J]. arXiv preprint, 2024, arXiv:2405.03125.
- [7] YANG M Y, BIAN C H, KIM H S. OFDM-guided deep joint source channel coding for wireless multipath fading channels[J]. IEEE Transactions on Cognitive Communications and Networking, 2022, 8(2): 584-599.
- [8] SUN Q Z, GUO C L, YANG Y, et al. Deep joint source-channel coding for wireless image transmission with semantic importance[C]//Proceedings of the 2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall). Piscataway: IEEE Press, 2022: 1-7.
- [9] XIAO Z X, YAO S S, DAI J C, et al. Wireless deep speech semantic transmission[C]//Proceedings of the ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2023: 1-5.
- [10] WU H T, SHAO Y L, BIAN C H, et al. Deep joint source-channel coding for adaptive image transmission over MIMO channels[J]. IEEE Transactions on Wireless Communications, 2024, 23(10): 15002-15017.
- [11] PENG J C, XING H L, LI Y, et al. Task-oriented multi-user semantic communication with lightweight semantic encoder and fast training for resource-constrained terminal devices[J]. IEEE Wireless Communications Letters, 2024, 13(9): 2427-2431.
- [12] BURGHAL D, LI Y, MADADI P, et al. Enhanced AI-based CSI prediction solutions for massive MIMO in 5G and 6G systems[J]. IEEE Access, 2023, 11: 117810-117825.
- [13] ALKIN B, BECK M, PÖPPEL K, et al. Vision-LSTM: xLSTM as generic vision backbone[J]. arXiv preprint, 2024, arXiv: 2406.04303.
- [14] BECK M, PÖPPEL K, SPANRING M, et al. xLSTM: extended long short-term memory[J]. arXiv preprint, 2024, arXiv: 2405.04517.
- [15] KANG H, HU Q Y, CHEN H X, et al. Cross-shaped separated spatial-temporal UNet transformer for accurate channel prediction[C]//Proceedings of the IEEE INFOCOM 2024 - IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2024: 2079-2088.
- [16] YANG S S, HANZO L. Fifty years of MIMO detection: the road to large-scale MIMOs[J]. IEEE Communications Surveys & Tutorials, 2015, 17(4): 1941-1988.
- [17] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2017: 618-626.
- [18] KURKA D B, GÜNDÜZ D. DeepJSCC-f: deep joint source-channel coding of images with feedback[J]. IEEE Journal on Selected Areas in Information Theory, 2020, 1(1): 178-193.
- [19] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: dataset and study[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE Press, 2017: 1122-1131.
- [20] RICHARDSON T, KUDEKAR S. Design of low-density parity check codes for 5G new radio[J]. IEEE Communications Magazine, 2018, 56(3): 28-34.
- [21] FRESIA M, PERÉZ-CRUZ F, POOR H V, et al. Joint source

- and channel coding[J]. IEEE Signal Processing Magazine, 2010, 27(6): 104-113.
- [22] HORÉ A, ZIOU D. Image quality metrics: PSNR vs. SSIM[C]//Proceedings of the 2010 20th International Conference on Pattern Recognition. Piscataway: IEEE Press, 2010: 2366-2369.
- [23] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment[C]//Proceedings of the Thirtieth Asilomar Conference on Signals, Systems & Computers, 2003. Piscataway: IEEE Press, 2003: 1398-1402.
- [24] ALKHATEEB A. DeepMIMO: a generic deep learning dataset for millimeter wave and massive MIMO applications[J]. arXiv preprint, 2019, arXiv:1902.06435.
- [25] LIU Y, HU T G, ZHANG H R, et al. iTransformer: inverted transformers are effective for time series forecasting[J]. arXiv preprint, 2023, arXiv: 2310.06625.
- [26] ZHANG Y B, WANG J, SUN J L, et al. CV-3DCNN: complex-valued deep learning for CSI prediction in FDD massive MIMO systems[J]. IEEE Wireless Communications Letters, 2021, 10(2): 266-270.

[作者简介]



于创宇 (2000-), 男, 武汉大学测绘遥感信息工程全国重点实验室硕士生, 主要研究方向为语义通信与联邦学习。



徐彦彦 (1974-), 女, 博士, 武汉大学测绘遥感信息工程全国重点实验室教授, 主要研究方向为云计算安全与智能网络通信。



潘少明 (1972-), 男, 博士, 武汉大学测绘遥感信息工程全国重点实验室副教授, 主要研究方向为多媒体网络与通信技术和数据挖掘与建模。