宽SNR多信道环境下稳定潜在表示学习的语义通信方法

杨颜博[1]，高晓婷[1]，张嘉伟[2]，马建峰[2]，高丽丽[1]
（1.内蒙古科技大学数智产业学院，内蒙古 包头 014010；2.西安电子科技大学网络与信息安全学院，陕西 西安 710071）

摘要：针对宽信噪比（SNR）及多信道环境下潜在语义表示的分布漂移问题，提出一种信道感知稳定语义通信框架，通过信道状态辅助调制与跨信道表示约束增强潜在表示一致性，并结合亮度–色度（YUV）颜色空间解耦重建机制和动态特征自适应优化策略，提升复杂信道条件下的语义重建能力与训练稳定性。为评估潜在表示的跨信道稳定特性，从分布一致性、结构一致性、信道相关性及潜空间聚类特征等多个维度开展实验分析，并采用切片Wasserstein距离、Fréchet距离、中心化核对齐、互信息估计及t分布随机邻域嵌入等指标进行验证。实验结果表明，在加性白高斯噪声、瑞利衰落和莱斯衰落等典型信道环境下，所提方法在保持较好图像重建性能的同时，能够有效降低潜在表示分布偏移，增强跨信道结构一致性与潜空间聚类紧致性，表现出较好的跨信道稳定性与泛化能力。

关键词：语义通信；信道感知；宽信噪比；YUV分离重建
中图分类号：TN929.5
文献标志码：A
doi：10.11959/j.issn.1000–436x.TXXB260356

Semantic communication with stable latent representation learning in wide-SNR multi-channel environments

Yang Yanbo[1], Gao Xiaoting[1], Zhang Jiawei[2], Ma Jianfeng[2], Gao Lili[1]
1. School of Digital and Intelligent Industry, Inner Mongolia University of Science & Technology, Baotou 014010, China
2. School of Cyber Engineering, Xidian University, Xi'an 710071, China

Abstract: To address latent semantic distribution shifts across wide signal-to-noise ratio (SNR) ranges and diverse channels, a channel-aware stable semantic communication framework was proposed. It integrated channel-state-conditioned modulation, cross-channel consistency constraints, YUV-based decoupled reconstruction, and dynamic feature adaptation to enhance representation stability and reconstruction robustness. Stability was evaluated using sliced Wasserstein distance (SWD), Fréchet distance (FD), centered kernel alignment (CKA), mutual information (MI), and t-SNE. Experiments over AWGN, Rayleigh, and Rician channels demonstrate reduced distribution shifts, improved structural consistency, and enhanced cross-channel generalization while maintaining high reconstruction quality.

Key words: semantic communication, channel awareness, wide SNR, YUV decoupled reconstruction

收稿日期：2026–06–17；**修回日期：**2026–08–05
通信作者：高晓婷，gaoxiaoting@stu.imust.edu.cn
基金项目：内蒙古自治区自然科学基金委联合项目（No.2024LHS06005）；国家自然科学基金资助项目（No.U2541201，No.U24A20243）；内蒙古自治区自然科学基金（No.2026QC0413）
Foundation Items: Inner Mongolia Autonomous Region Natural Science Foundation Joint Project (No.2024LHS06005), The National Natural Science Foundation of China (No.U2541201, No.U24A20243), Natural Science Foundation of Inner Mongolia Autonomous Region (No.2026QC0413)

0　引言

随着无线通信向智能化、自适应化方向发展，传统以比特级可靠传输为核心的香农通信范式已难以满足未来智能业务对语义理解、任务协同和低时延传输的需求。语义通信通过提取并传输与接收任务相关的语义信息，在有限通信资源和复杂信道条件下提高信息传输效率与任务鲁棒性，被认为是面向6G的重要发展方向[1]。张平等[2]从语义通信系统模型、基本概念、语义信息度量以及直接编码与变换编码等方面系统阐述了语义通信的基本原理和实现方法。刘传宏等[3]进一步面向智能任务研究语义信息度量、语义编码及联合信源信道编码，并从信息瓶颈角度分析了语义保持与信息压缩之间的关系。上述研究为语义通信系统的理论建模和技术实现提供了重要基础。牛凯等[4]从同义映射出发构建了语义信息度量体系，讨论了语义熵、语义互信息、语义信道容量及语义率失真函数，并给出了语义通信编码理论的数学描述，为语义通信系统的理论分析提供了进一步支撑。

在语义通信体系架构和应用研究方面，石光明等[5]提出面向语义信息直传的通信架构，强调从完整信息中筛选对信宿具有实际信息价值的语义内容进行传输，以降低通信资源和信息处理开销。秦志金等[6]从文本、图像、视频及多模态信息等角度系统总结了语义通信的研究进展，指出语义表征、模态融合、信道鲁棒传输及语义安全是语义通信进一步发展的重要方向。

在图像语义通信领域，基于深度学习的联合信源信道编码（joint source-channel coding, JSCC）已成为重要技术路线。Bourtsoulatze等[7]提出深度联合信源信道编码（deep joint source-channel coding, DeepJSCC），利用端到端神经网络直接将图像映射为信道符号，缓解了传统数字通信系统在低信噪比（signal-to-noise ratio, SNR）条件下的悬崖效应；Kurka等[8]通过引入信道反馈增强模型对动态信道的适应能力；Xie等[9]提出深度语义通信（deep semantic communication, DeepSC）框架，实现语义特征提取与信道编码的联合优化；Yang等[10]无线图像传输Transformer（wireless image transmission transformer, WITT）模型，利用移位窗口Transformer（shifted window transformer, swin transformer）增强图像全局语义和局部结构建模能力。国内学者也开展了相关研究，江沸菠等[11]提出深度图像语义通信模型DeepISC（deep image semantic communication），利用基于视觉Transformer（vision transformer, ViT）的自编码器提取图像语义特征，并结合信道自编码器和判别器辅助训练，提高了不同信道下的图像重建与语义恢复能力；张平等[12]从非线性变换编码角度研究面向语义通信的特征压缩与编码传输，为可学习语义编码提供了统一的变换编码思路；罗鹏等[13]通过融合全局与局部上下文特征增强语义特征表达，提高了端到端语义通信系统的传输可靠性。上述研究验证了深度JSCC及语义特征建模在图像语义传输中的有效性。

针对实际无线环境中信噪比和通信资源动态变化的问题，研究者进一步开展了信道自适应与灵活编码研究。Kurka等[14]提出Bandwidth-agile Deep-JSCC，通过可变传输维度实现带宽自适应；Ding等[15]利用信道状态信息调节网络特征，提高单一模型在不同SNR条件下的适应能力；Xie等[16]利用超网络根据信道状态动态生成模型参数；Peng等[17]通过多尺度视觉Transformer和交叉注意力增强复杂信道条件下的语义恢复能力；杨舒涵等[18]提出基于深度灵活编码与性能预测的联合信源信道编码（deeply flexible joint source-channel coding，DF-JSCC）系统，通过灵活特征编码和传输性能预测提高动态通信环境下的图像重建性能和通信资源利用效率。上述方法表明，引入信道状态感知、特征增强和灵活资源分配能够有效提高图像语义通信系统的信道适应能力。

然而，现有方法主要以重建质量、任务准确率、带宽消耗或信道适应性能为优化目标。信道状态辅助调制虽然能够使模型针对当前信道条件调整编码策略，但并不能直接保证同一语义内容在不同信噪比（SNR）和不同信道类型下获得的潜在表示具有一致的统计结构。在宽SNR及多信道环境中，噪声强度、衰落增益和信道统计参数的变化可能使潜在语义特征包含不同程度的信道相关信息，导致同一图像对应的潜在分布随信道状态发生偏移。当训练信道与测试信道不一致时，这种跨信道分布漂移可能进一步引起语义结构退化和泛化性能下降。现有研究对该问题的显式约束以及从分布一致性、结构一致性、信道依赖性

和潜空间聚类特性等多个维度开展的统一评价仍相对有限。

此外，在低 SNR 及衰落信道条件下，信道噪声容易造成图像结构模糊、纹理缺失和颜色漂移。传统像素级重建损失虽然能够提高峰值信噪比，但可能产生过度平滑现象，难以同时兼顾结构保持和视觉感知质量[19]。由于人类视觉系统对亮度结构变化较为敏感，而色度分量对信道噪声和重建误差具有不同响应特性[20]，在解码阶段对亮度结构和色度信息进行差异化恢复，有助于改善低 SNR 条件下的结构保持和颜色稳定性。

基于上述分析，本文提出一种面向宽 SNR 多信道环境的信道感知稳定语义通信系统（channel-aware stable semantic communication system, CASCS）。该系统以稳定潜在表示学习为核心，通过信道状态条件建模和跨信道一致性约束协调信道适应性与表示稳定性，并结合亮度–色度解耦重建和动态优化策略，提高复杂信道条件下的图像语义恢复能力。本文的主要贡献如下。

（1）针对宽 SNR 和多信道环境下潜在语义表示容易发生分布漂移的问题，提出一种稳定语义表示学习机制。通过构建信道状态嵌入与跨信道特征一致性约束，降低潜在表示对非必要信道状态变化及统计扰动的敏感性，从而提高不同信道和 SNR 条件下潜在表示的一致性。

（2）针对低 SNR 条件下结构退化和颜色噪声耦合问题，提出信道感知结构解码机制。该机制在亮度–色度颜色空间（luminance-chrominance color space, YUV），并结合重建可靠性估计与动态自适应滤波决策，根据当前信道质量、图像局部结构及恢复可信度动态调节滤波强度，在抑制色度噪声的同时减少边缘和纹理信息损失。

现有研究多关注重建性能提升或局部信道适配，而对宽 SNR 多信道环境下潜在表示分布漂移缺乏系统性分析与显式约束。本文在加性白高斯噪声（additive white Gaussian noise, AWGN）[21]、瑞利衰落（Rayleigh fading）[22]与莱斯衰落（Rician fading）[23]信道环境下开展系统实验，采用峰值信噪比（peak signal-to-noise ratio, PSNR）[24]、结构相似性（structural similarity, SSIM）[25]与学习感知图像块相似性（learned perceptual image patch similarity, LPIPS）[26]评价重建质量，并结合切片 Wasserstein 距离（sliced Wasserstein distance, SWD）[27]、Fréchet 统计距离（Fréchet distance, FD）[28]分析、中心化核对齐（centered kernel alignment, CKA）[29]结构一致性分析、高斯近似（Gaussian approximation）[30]与神经互信息估计（mutual information neural estimation, MINE）[31]互信息估计以及采用 t 分布随机邻域嵌入（t-distributed stochastic neighbor embedding, t-SNE）[32]等稳定性分析指标上表现出更好的跨信道一致性。

1 跨信道潜在表示稳定性问题分析

现有图像语义通信研究主要关注端到端重建性能，而对宽 SNR、多信道环境下潜在语义表示的分布漂移关注不足。复杂无线环境中，噪声、衰落及信道统计特性的变化会导致潜在特征空间偏移，降低语义表示一致性和系统鲁棒性。本文基于潜在表示分析，研究宽 SNR 条件下潜在特征的分布变化规律，并从表示稳定性角度分析复杂信道环境下的语义表征行为。

1.1 跨信道潜在表示问题描述

图像语义通信利用深度神经网络将输入图像映射为潜在语义表示，并通过无线信道传输实现端到端语义恢复。然而，现有方法主要通过改进网络结构或增强重建能力提升传输性能，较少关注信道动态变化对潜在表示的影响。当模型仅在单一信道条件下训练时，潜在表示容易包含信道相关信息，导致同一图像在不同信道条件下产生不同的特征分布。

从表示学习角度看，理想的语义通信系统应学习与信道状态解耦的潜在语义表示，使同一输入样本的特征主要反映图像语义内容，而不随信道条件变化发生明显漂移。若潜在表示过度依赖信道状态，则不同信道环境下的特征分布将发生偏移，导致解码器难以保持稳定的语义恢复能力，并进一步削弱系统的跨信道泛化性能。因此，降低潜在表示对信道状态的依赖、学习具有跨信道一致性的稳定潜在表示，是提升宽 SNR 多信道语义通信鲁棒性的关键。

1.2 跨信道分布漂移分析

设输入图像表示为x，不同信道状态表示为c_i，编码器输出潜在语义表示为z_i。对于同一输入图像，当信道状态发生变化时，由于噪声统计特性、

衰落系数以及信道增益的差异，潜在表示可能产生不同程度的偏移，即：

$$z_i = f(x, c_i) \tag{1}$$

其中，f表示编码映射函数。理想情况下，潜在表示应主要表征图像语义内容，因此在不同信道状态下应保持较高的一致性。然而，由于网络在训练过程中需要同时适应图像特征与信道扰动，潜在表示往往会包含一定的信道相关信息。当信道条件发生变化时，这部分信息将导致潜在特征分布产生偏移，即跨信道分布漂移现象。

从特征空间角度来看，不同SNR及信道条件对应不同的潜在表示分布。低SNR时噪声主导特征表达，导致特征离散度增大和语义边界模糊；随着SNR提高，潜在表示逐渐向真实语义分布收敛。因此，不同信道统计特性容易造成潜在空间分布不连续，削弱模型的跨信道泛化能力。

因此，跨信道分布漂移本质上反映了潜在表示对信道状态的敏感性。构建具有一致性、连续性和鲁棒性的潜在表示空间，是提升宽SNR多信道环境下语义通信性能的重要基础。

1.3　跨信道潜在表示变化分析

为表征语义通信系统在复杂无线环境下的跨信道鲁棒性，本文引入稳定潜在表示的概念。对于同一输入图像，不同信道条件下获得的潜在表示虽然允许存在局部扰动，但其整体统计结构应保持相对一致，并能够稳定表征图像语义信息。设潜在表示为z，信道状态集合为C，则理想情况下应满足：

$$D(P(z|c_i), P(z|c_j)) \approx 0, \forall c_i, c_j \in C \tag{2}$$

其中，$D(\cdot)$表示分布差异度量。该条件表明，同一图像在不同信道状态下对应的潜在表示应具有相近的统计特征。

稳定潜在表示应同时满足语义保持性和信道不变性，即在保留图像语义信息的同时降低对信道状态变化的敏感性，从而使解码器在不同信道条件下接收到更加一致的潜在特征，提高跨信道泛化能力。

当潜在表示同时具备上述特性时，解码器能够在不同信道条件下接收到更加一致的特征分布，从而提升系统的跨信道泛化能力与语义恢复稳定性。因此，构建稳定潜在表示是缓解跨信道分布漂移、增强宽SNR多信道环境下语义通信鲁棒性的关键基础。

1.3.1　潜在表示分布变化描述

在宽SNR及多信道环境下，同一图像样本在不同信道条件下可能对应不同的潜在表示分布。为刻画潜在表示随信道变化的统计特性，本文采用分布距离作为潜在特征一致性的度量指标。

设不同信道条件下的潜在表示分别为：

$$z_i \sim P_i(z), z_j \sim P_j(z) \tag{3}$$

其中，P_i与P_j分别表示不同信道条件下潜在表示对应的统计分布。为保证潜在空间具有良好的语义表达能力，除满足跨信道一致性外，还应保持不同图像样本之间的特征可分性，即：

$$z_i \neq z_j, i \neq j \tag{4}$$

因此，稳定语义表示学习可视为类内稳定、类间可分的联合约束问题。其中，类内稳定性用于减弱信道变化带来的分布漂移，类间可分性用于维持潜在空间的语义判别能力。为定量评估跨信道分布偏移程度，定义跨信道分布距离为：

$$D_{\text{shift}} = D(P(z|c_i), P(z|c_j)) \tag{5}$$

其中，$D(\cdot)$表示分布距离度量函数，下标shift表示跨信道分布漂移。

考虑到潜在表示通常位于高维连续空间，本文采用SWD衡量不同信道条件下的分布差异：

$$D_{\text{SWD}}(P_i P_j) = \frac{1}{L}\sum_{l=1}^{L} W_1(\Pi_l P_i, \Pi_l P_j) \tag{6}$$

其中，$P_i = P(z|c_i)$，$P_j = P(z|c_j)$，$W_1(\cdot)$表示一阶Wasserstein距离，Π_l表示第l个随机投影方向，L表示随机投影次数。SWD值越小，表明不同信道条件下的潜在表示分布越接近，即具有更强的跨信道一致性。

1.3.2　潜在表示与信道状态关系分析

受信息瓶颈（information bottleneck, IB）[33]理论及其深度变分实现（deep variational information bottleneck, DVIB）[34]思想启发，本文从信息相关性角度分析潜在表示与信道状态之间的关系。在复杂无线环境中，潜在表示不仅需要保留图像语义信息，还可能包含部分信道相关特征。设潜在表示为z，信道状态为c，P表示相应概率密度函数，则二者之间的互信息可表示为：

$$I(z; c) = \int p(z; c) \log \frac{p(z, c)}{p(z)\, p(c)} \mathrm{d}z\mathrm{d}c \tag{7}$$

互信息用于衡量潜在表示对信道状态的依赖程

度。较大的$I(z;c)$表明潜在表示包含更多信道相关信息，而较小的$I(z;c)$则意味着潜在特征能够有效抑制信道扰动，更集中地保留与语义任务相关的信息。因此，降低潜在表示与信道状态之间的互信息，有助于提高跨信道稳定性和泛化能力。

1.4 潜在表示稳定性分析方法

为评估不同模型在宽SNR及多信道环境下的潜在表示稳定性，本文从分布统计特性、信息相关性、潜空间结构及聚类特征等方面开展实验分析。

（1）分布统计变化分析

为刻画不同信道条件下潜在表示的分布偏移程度，本文采用Fréchet距离衡量不同信道条件下潜在表示整体统计分布差异，其定义为：

$$\mathrm{FD}(P_i,P_j)=\left\|\mu_i-\mu_j\right\|_2^2+\mathrm{Tr}\left(\sum i+\sum j-2\left(\sum i\sum j\right)^{\frac{1}{2}}\right) \tag{8}$$

其中，μ_i与μ_j分别表示潜在特征均值，$\sum i$与$\sum j$表示对应协方差矩阵。FD越小，表示分布一致性越高。

（2）潜在表示互信息分析

为分析潜在表示与信道状态之间的相关性，本文采用Gaussian Approximation与MINE两种方法估计潜在表示与信道状态之间的互信息。语义保持能力基本相当的前提下，较低的非必要信道相关互信息通常表示潜在特征对信道扰动的敏感性较低，但互信息不能单独作为模型优劣的判定依据，仍需结合重建质量和结构一致性综合分析。

（3）潜在表示低维可视化分析

为直观观察跨信道条件下潜在特征的聚类行为，本文利用t-SNE将高维潜在表示映射至二维空间，用于观察不同信道条件下潜在表示的聚类结构及跨信道一致性。

（4）类内与类间距离分析

为定量评估潜在表示的稳定性与判别性，本文定义了类内距离与类间距离。

类内距离用于衡量同一图像在不同SNR条件下潜在表示的平均距离，设第i张图像在不同信道条件下对应的潜在特征表示为：$\left\{Z_i^{(1)},Z_i^{(2)},\cdots,Z_i^{(K)}\right\}$，其中$K$表示不同SNR条件数量，则类内距离定义为：

$$D_{\mathrm{intra}}=\frac{1}{N}\sum_{i=1}^{N}\frac{2}{K(K-1)}\sum_{p<q}\left\|z_i^{(p)}-z_i^{(q)}\right\|_2 \tag{9}$$

其中，N表示测试图像总数，$\|\cdot\|_2$表示欧氏距离。

设不同图像对应的潜在表示均值分别为：

$$\bar{z}_i=\frac{1}{K}\sum_{k=1}^{K}z_i^{(k)} \tag{10}$$

进一步定义类间距离为：

$$D_{\mathrm{inter}}=\frac{2}{N(N-1)}\sum_{i<j}\left\|\bar{z}_i-\bar{z}_j\right\|_2 \tag{11}$$

类内距离衡量同一图像在不同信道条件下潜在表示的一致性；类间距离衡量不同图像之间的可分性。较低的类内距离和较高的类间距离分别对应更好的稳定性和判别能力。

（5）综合稳定性分析指标

为联合衡量潜在表示的稳定性与判别能力，构建综合评价指标：

$$\mathrm{Score}=\frac{D_{\mathrm{inter}}}{D_{\mathrm{intra}}+\varepsilon} \tag{12}$$

其中，ε为极小常数。较高Score值表明模型在跨信道条件下可能同时具有较好的潜在表示稳定趋势与特征可分性。

为全面评价宽SNR多信道环境下潜在表示的稳定性，本文构建多维稳定性分析方法，从分布一致性、统计依赖性、结构一致性和聚类特性4个方面进行综合评价。其中，SWD和FD用于衡量不同信道条件下潜在表示的分布一致性；高斯近似和MINE用于分析潜在表示与信道状态之间的统计依赖关系；CKA用于评价潜在表示的结构一致性；t-SNE结合类内距离、类间距离及Score指标用于分析潜在空间的聚类紧致性和判别能力。上述指标从多个互补维度评价模型的跨信道稳定性，不存在评价冗余。

2 信道感知语义通信系统

2.1 系统总体框架

针对宽SNR及复杂多信道环境下潜在表示分布漂移和跨信道泛化能力不足的问题，本文提出一种稳定语义通信系统CASCS。该框架以稳定语义表示学习为核心，联合信道感知、结构引导解码和动态优化机制，提高复杂无线环境下的语义传输与图像重建性能。

系统整体框架如图1所示，包括发送端语义编码、信道感知稳定语义调制、无线信道传输、结构引导语义解码及动态稳定优化5个模块。发送端结合信道状态信息（channel state information, CSI）

对潜在语义特征进行条件调制，接收端利用信道感知结构解码模块（channel-aware structure decoder, CASD）完成图像重建，动态特征自适应优化（dynamic feature adaptive optimization, DFAO）负责动态协调训练过程中的多目标优化。

本文假设系统能够获得理想CSI。需要说明的是，本文所使用的CSI并非传统通信中的瞬时信道矩阵，而是由SNR、信道类型及对应统计参数共同构成的信道状态描述信息。与传统通信中主要利用CSI进行信道均衡不同，本文将其作为网络的条件信息参与端到端语义通信模型建模，引导模型学习不同信道条件下的语义特征映射关系，从而增强宽SNR、多信道环境下的信道自适应能力，具体的信道状态嵌入方式将在第2.2.1节介绍。

接收端采用CASD，在YUV空间分别恢复结构与颜色信息，并结合动态滤波缓解低SNR条件下的结构退化与颜色失真。此外，本文引入DFAO机制，通过动态调整各损失项权重，提高宽SNR下的训练稳定性。与传统方法相比，CASCS不仅关注图像重建性能，还关注潜在表示的跨信道一致性，并结合SWD、FD、互信息及t-SNE等指标进行综合评价。

2.2　稳定语义表示学习机制

为降低宽SNR及多信道环境下潜在表示对信道状态的敏感性，本文提出稳定语义表示学习机制（stable semantic representation learning, SSRL），通过信道状态感知、跨信道特征对齐及连续信道建模约束潜在语义空间。

2.2.1　信道状态感知嵌入

宽SNR、多信道环境下，不同信道类型及其统计参数会导致潜在语义表示产生不同程度的分布偏移，使传统固定编码策略难以兼顾不同信道条件下的传输性能。为了提高模型对不同信道环境的自适应能力，需要将当前信道状态作为条件信息引入编码过程，使网络能够根据不同信道条件动态调整语义特征表示。为此，本文设计了信道状态嵌入网络。具体而言，将由SNR、信道类型及对应统计参数构成的信道状态描述信息编码为统一的信道状态向量，并映射至连续特征空间，获得统一的连续信道表示。

设当前信道状态为：

$$c = \{s,t,\theta\} \tag{13}$$

其中，s表示当前SNR，t表示信道类型，θ表示对应信道的统计参数。对于AWGN信道，θ表示噪声方差；对于Rayleigh信道，θ表示衰落统计参数；对于Rician信道，θ进一步包含Rice因子等统计信息。上述信息共同构成统一的信道状态描述向量：

$$\boldsymbol{c} = [s,t,\theta] \tag{14}$$

随后，将信道状态向量输入信道状态嵌入网

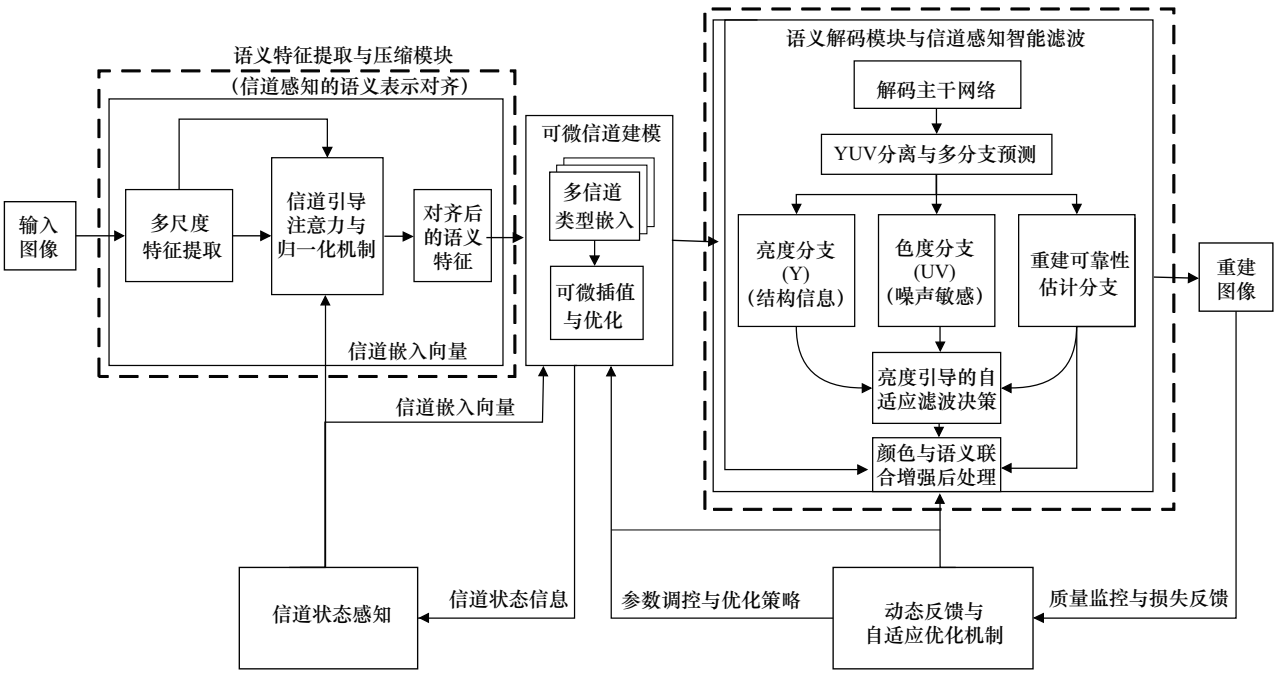


图1　系统总体框架

络，通过全连接映射获得连续信道嵌入表示：

$$e_c = f_{\text{emb}}(c) \tag{15}$$

其中，$f_{\text{emb}}(\cdot)$表示信道嵌入函数，e_c表示连续信道嵌入向量。

为了增强不同信道条件下编码器的自适应能力，本文进一步利用连续信道嵌入向量对编码器潜在特征进行条件调制，其表达式为：

$$\boldsymbol{z} = \boldsymbol{z} \odot \sigma(\boldsymbol{W}\boldsymbol{e}_c + \boldsymbol{b}) \tag{16}$$

其中，$\boldsymbol{z}$表示编码器提取的潜在语义特征，$\sigma(\cdot)$表示Sigmoid激活函数，$\boldsymbol{W}$和$\boldsymbol{b}$分别表示为可学习参数和偏置参数，$\odot$表示逐元素乘积。通过门控调制方式，不同信道状态对应不同的调制权重，从而使模型能够根据当前信道条件动态调整潜在语义特征表达。

最终，编码器输出的潜在语义表示同时包含输入图像的语义信息和对应的信道条件信息，为后续连续信道建模及稳定语义表示学习提供统一的条件表示。

2.2.2 语义特征提取与跨信道对齐

为增强复杂信道环境下的语义表达能力，本文采用卷积与Transformer相结合的层次化编码结构。首先利用残差卷积提取局部纹理特征：

$$\boldsymbol{f}_l = H(\boldsymbol{x}) + \boldsymbol{x} \tag{17}$$

其中，$H(\cdot)$表示卷积映射。

随后利用窗口自注意力机制建模全局语义关系：

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \text{Soft}\max\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d}}\right)\boldsymbol{V} \tag{18}$$

其中，$\boldsymbol{Q}$、$\boldsymbol{K}$和$\boldsymbol{V}$分别表示查询、键和值矩阵。

为抑制跨信道分布漂移，进一步引入跨信道特征对齐约束：

$$L_{\text{sem}} = \left\| z_i - z_j \right\|_2^2 \tag{19}$$

其中，z_i与z_j分别表示同一图像在不同信道条件下的潜在语义表示。通过约束跨信道特征距离，促使模型学习与信道状态弱相关的语义表示，从而增强潜在空间的一致性与稳定性。

2.2.3 可微信道插值建模

传统方法通常针对离散信道分别训练模型，难以学习不同信道之间共享的统计特征。为此，本文提出可微信道插值建模（differentiable channel interpolation layer, DCIL），对信道状态嵌入进行连续建模。

DCIL以信道状态嵌入为输入，在连续嵌入空间内完成可微插值与条件调制。DCIL连续建模的对象并非AWGN、Rayleigh、Rician等离散信道类型，而是对应的信道状态嵌入。首先，根据当前信道类型及其统计参数构建统一的信道状态向量，设不同信道的参数化映射表示为：

$$\boldsymbol{s} = [c,\rho,\theta] \tag{20}$$

其中，c表示信道类型编码，ρ表示SNR，θ表示对应信道的统计参数，如噪声方差、衰落参数以及Rice K因子等。随后，通过信道状态嵌入网络映射得到连续信道表示，对于任意两个信道状态，其对应嵌入分别表示为$\boldsymbol{e}_i$和$\boldsymbol{e}_j$。DCIL在连续嵌入空间内进行可微插值，生成中间信道表示，并利用插值后的连续信道表示对潜在语义特征进行条件调制：

$$\boldsymbol{e}_\lambda = (1-\lambda)\boldsymbol{e}_i + \lambda\boldsymbol{e}_j, \lambda \in [0,1] \tag{21}$$

$$\tilde{z} = \phi(z,\boldsymbol{e}_\lambda) \tag{22}$$

其中，$\tilde{z}$表示编码器输出的潜在语义表示，$\phi(\cdot)$表示特征调制函数。DCIL通过连续建模信道状态嵌入，实现不同信道间的平滑过渡，提高宽SNR、多信道及未知信道环境下的泛化能力。

DCIL通过连续建模信道状态嵌入，实现不同信道间的平滑过渡，因此，即使训练过程中仅覆盖有限数量的信道条件，模型仍能够学习不同信道之间共享的语义映射规律，提高宽SNR、多信道环境下的信道自适应能力和未知信道条件下的泛化能力。

2.2.4 基于信息瓶颈的稳定语义表示约束分析

为学习稳定潜在表示，本文引入IB理论，在保持语义信息的同时抑制与信道状态相关的冗余信息，其优化目标定义为：

$$\max I(\boldsymbol{z};\boldsymbol{x}) - \beta I(\boldsymbol{z};\boldsymbol{c}) \tag{23}$$

其中，$I(\boldsymbol{z};\boldsymbol{x})$表示潜在表示与输入图像之间的互信息，用于衡量语义保持能力；$I(\boldsymbol{z};\boldsymbol{c})$表示潜在表示与信道状态之间的互信息，用于描述潜在特征对信道条件的依赖程度；β为平衡系数。

由于互信息难以直接优化，其优化过程依赖互信息估计器，训练复杂度较高且容易产生估计误差。因此，本文采用跨信道一致性约束作为信息瓶

颈目标的代理优化目标。虽然DCIL能够利用信道条件学习不同信道环境下的自适应语义映射，提高模型对不同信道条件的适应能力，但其优化目标仍然是条件建模，而非潜在表示一致性。不同信道统计特性及随机噪声仍可能导致同一图像在不同信道条件下产生不同的潜在表示，因此仅依赖DCIL仍难以保证潜在表示具有跨信道一致性。为降低潜在表示中的信道相关冗余信息，本文引入基于信息瓶颈思想的稳定语义表示学习，其核心思想是在保持语义信息不变的前提下，通过跨信道一致性约束减弱潜在表示对信道变化的敏感性，使模型不仅具有信道自适应能力，而且具有跨信道稳定表示能力。对于同一输入图像x，分别在不同信道条件c_i和c_j下获得对应潜在表示：

$$L_{\text{con}} = \left\| z_i - z_j \right\|_2^2 \tag{24}$$

其中，z_i和z_j分别表示同一图像在不同信道条件下获得的潜在表示。最小化一致性损失能够减少潜在表示随信道变化产生的分布偏移，从而间接降低潜在表示与信道状态之间的互信息$I(z;c)$。同时，由于约束作用于同一图像，不会破坏其语义信息，因此能够较好地保持潜在表示与输入图像之间的互信息$I(z;x)$。

综合重建任务与稳定表示约束，最终训练目标定义为：

$$L_{\text{SSRL}} = L_{\text{task}} + \lambda L_{\text{con}} \tag{25}$$

其中，L_{task}表示图像重建损失，λ为一致性约束权重。通过上述代理优化，模型无需直接估计互信息即可近似实现信息瓶颈目标，从而提高潜在语义表示在宽SNR、多信道动态环境下的稳定性与泛化能力。

因此，DCIL与稳定语义表示学习分别作用于两个不同层面。DCIL利用信道条件建立连续语义映射，提高模型对不同信道环境的自适应能力；稳定语义表示学习进一步约束不同信道条件下潜在表示的一致性，降低潜在表示中的信道相关冗余信息，提高跨信道泛化能力。

2.3　结构引导语义重建机制

2.3.1　语义特征解码主干网络

为缓解低SNR条件下语义退化和结构失真问题，本文在接收端设计结构引导语义解码模块CASD，采用轻量化卷积解码网络对接收语义特征进行恢复，得到中间语义表示，为后续亮度、色度及重建可靠性估计提供共享特征基础。其中，$\boldsymbol{F}$表示解码主干输出的共享特征图。

2.3.2　三分支预测头

为实现结构与颜色信息的协同恢复，CASD采用亮度预测、色度预测和重建可靠性估计3分支结构。

（1）亮度预测分支

亮度分支输出单通道亮度图：

$$\boldsymbol{Y} = \sigma(\psi_Y(\boldsymbol{F})), \boldsymbol{Y} \in [0,1]^{H \times W} \tag{26}$$

其中，$\psi_Y(\cdot)$为卷积映射，$\sigma(\cdot)$表示Sigmoid函数。亮度通道承载主要结构信息，该分支用于提供稳定的结构恢复约束。

（2）色度预测分支

色度分支输出UV两通道颜色信息，利用亮度引导滤波对色度预测进行约束，以提升颜色恢复的稳定性和视觉一致性。为避免将UV理解为U与V的乘积，本文将两路色度分量拼接得到的联合色度表示记为$\boldsymbol{C}_{\text{UV}}$，即$\boldsymbol{C}_{\text{UV}} = \text{cat}(U,V)$，其中cat表示通道拼接。

（3）重建可靠性估计分支

为评估不同空间位置的重建可信程度，引入重建可靠性估计分支，其输出为一张空间可靠性图：

$$\boldsymbol{R} = \sigma(f_R(\boldsymbol{F}_d)) \tag{27}$$

其中，$f_R(\cdot)$表示轻量卷积预测头，下标d表示解码特征，R表示重建可靠性图，其数值反映对应区域恢复结果的可信程度。可靠性较高的区域说明预测结果稳定，而可靠性较低的区域通常对应信道扰动较强或恢复困难的位置。该可靠性图作为动态自适应滤波决策机制（dynamic adaptive filtering decision mechanism, DAFDM）的反馈信息，用于引导动态滤波参数调整，实现对不同区域的自适应优化，提高复杂信道环境下的整体重建质量。

2.3.3　动态自适应滤波决策机制

为适应不同信道条件下的噪声分布变化，本文提出DAFDM。首先对当前SNR进行归一化处理，构建噪声压力特征，以表征当前信道环境下噪声干扰程度；随后结合亮度分量的局部方差和梯度信息生成结构强度图：

$$\boldsymbol{S} = f(\text{Var}(\boldsymbol{Y}), \nabla \boldsymbol{Y}) \tag{28}$$

其中，$\text{Var}(\cdot)$和$\nabla(\cdot)$分别表示局部方差与梯度算子。

随后，将噪声压力特征、亮度结构强度图及重建可靠性估计结果进行融合，并输入轻量级卷积决策网络，生成空间滤波强度图：

$$\boldsymbol{M} = \sigma(\mathbb{G}(\boldsymbol{N},\boldsymbol{S},\boldsymbol{R})) \tag{29}$$

其中，$\mathbb{G}(\cdot)$表示卷积决策网络，$\boldsymbol{M}$表示空间滤波强度图，$\boldsymbol{N}$表示噪声压力特征，$\boldsymbol{S}$表示由亮度方差与梯度得到的结构引导特征，$\boldsymbol{R}$表示可靠性图。

最终，根据当前信道质量、亮度结构信息及重建可靠性，自适应调节引导滤波半径、正则化系数及去噪融合权重，使模型能够在不同信道条件下动态平衡纹理细节保持与噪声抑制，提高图像重建质量和跨信道鲁棒性。

2.3.4 颜色与语义联合增强后处理模块

为进一步提升低SNR条件下的颜色恢复质量，本文设计了颜色与语义联合增强模块，包括色度神经去噪和亮度引导滤波两个阶段。

（1）色度神经去噪

针对色度通道易出现噪声放大和颜色伪影的问题，首先利用轻量神经网络对色度预测进行净化，并采用加权融合方式得到增强后的色度表示：

$$\boldsymbol{C}'_{\mathrm{UV}} = \alpha\boldsymbol{C}_{\mathrm{UV},d} + (1-\alpha)\boldsymbol{C}_{\mathrm{UV}} \tag{30}$$

其中，$\boldsymbol{C}_{\mathrm{UV},d}$表示去噪结果，$\alpha$为动态融合权重，下标UV表示联合色度张量，下标$d$表示预净化结果，撇号表示融合后的结果。

（2）亮度引导滤波

色度预净化后使用亮度Y作为引导图像对色度通道执行引导滤波，以抑制色度噪声跨边缘扩散并增强结构一致性。色度通道$X \in \{U',V'\}$，引导滤波在局部窗口ω内满足线性模型，线性系数由最小二乘估计得到。将亮度与滤波后的色度拼接，随后将YUV转化为RGB。

2.4 动态稳定优化机制

2.4.1 动态特征自适应优化

在宽SNR及多信道环境下，为协调不同信道条件下的训练难度，本文提出DFAO机制，根据历史训练性能动态调整任务恢复和稳定表示学习之间的权重。设第c类信道与第s个SNR条件下的历史性能集合为：

$$H_{c,s} = \{P_1,P_2,\cdots,P_T\} \tag{31}$$

其中，P_T表示当前训练阶段对应的重建性能指标。本文采用滑动窗口对不同信道状态下的历史性能进行统计，其平均性能定义为：

$$\overline{P}_{c,s} = \frac{1}{T}\sum_{t=1}^{T} P_t \tag{32}$$

为刻画不同信道条件之间的训练失衡程度，定义全局训练性能均值：

$$\overline{P}_{\mathrm{global}} = \frac{1}{N}\sum_{c,s}\overline{P}_{c,s} \tag{33}$$

其中，N表示当前参与训练的信道状态数量。随后，根据当前信道状态与全局训练状态之间的性能偏差定义信道困难度：

$$\Delta_{c,s} = \overline{P}_{\mathrm{global}} - \overline{P}_{c,s} \tag{34}$$

当$\Delta_{c,s}$较大时，说明当前信道条件下仍存在较强的语义恢复困难与分布漂移问题，需要增强稳定性约束；反之则适当提高细节恢复权重。据此构建任务恢复分支与语义稳定分支分别构建动态调节因子：

$$\lambda_{\mathrm{task}} = \mathrm{Clip}(1+\alpha\Delta_{c,s},\lambda_{\min},\lambda_{\max}) \tag{35}$$

$$\lambda_{\mathrm{sem}} = \mathrm{Clip}(1+\beta\Delta_{c,s},\lambda_{\min}^{\mathrm{sem}},\lambda_{\max}^{\mathrm{sem}}) \tag{36}$$

其中，α与β分别表示任务恢复分支与语义稳定分支的调节系数。为提高训练稳定性，采用指数滑动平均（exponential moving average, EMA）对动态权重进行平滑更新：

$$\lambda_t = (1-\eta)\lambda_{t-1} + \eta\hat{\lambda}_t \tag{37}$$

其中，η表示平滑更新系数，$\hat{\lambda}_t$表示当前时刻计算得到的动态权重。

2.4.2 总体优化目标

为同时保证重建质量、语义一致性和跨信道稳定性，本文构建多约束联合优化目标：

$$L_{\mathrm{total}} = \lambda_{\mathrm{rec}}L_{\mathrm{rec}} + \lambda_{\mathrm{sem}}L_{\mathrm{sem}} + \lambda_{\mathrm{freq}}L_{\mathrm{freq}} + \lambda_{\mathrm{edge}}L_{\mathrm{edge}} + \lambda_{\mathrm{stable}}L_{\mathrm{stable}} \tag{38}$$

其中，λ_{rec}、λ_{sem}、λ_{freq}、λ_{edge}、$\lambda_{\mathrm{stable}}$分别表示对应损失项的动态权重系数。

为保证基础重建质量并降低复杂信道条件下局部失真对整体潜在表示的影响，本文首先采用像素重建损失约束恢复图像与原始图像之间的差异，其定义为：

$$L_{\mathrm{rec}} = \frac{1}{N}\sum_{i=1}^{N}\left\| x_i - \hat{x}_i \right\|_1 \tag{39}$$

其中，x_i与$\hat{x}_i$分别表示原始图像与重建图像。

为增强不同信道条件下高层语义特征的一致性，本文引入语义一致性损失，其定义为：

$$L_{\text{sem}} = 1 - \frac{f(x)\cdot f(\hat{x})}{\|f(x)\|_2 \|f(\hat{x})\|_2} \tag{40}$$

其中，$f(\cdot)$表示语义特征提取函数。

考虑到复杂无线环境容易导致高频纹理信息退化，本文引入频域一致性损失，其定义为：

$$L_{\text{freq}} = \frac{1}{N}\sum_{i=1}^{N} \| \mathcal{F}(\boldsymbol{x}_i) - \mathcal{F}(\hat{\boldsymbol{x}}_i) \|_1 \tag{41}$$

其中，$\mathcal{F}(\cdot)$表示傅里叶频域变换。

此外，为维持图像结构轮廓连续性并增强复杂信道环境下的结构鲁棒性，本文引入边缘结构约束损失：

$$L_{\text{edge}} = \frac{1}{N}\sum_{i=1}^{N} \left\| \nabla x_i - \nabla \hat{x}_i \right\|_1 \tag{42}$$

其中，$\nabla(\cdot)$表示图像梯度算子。

为缓解不同信道条件下潜在特征分布偏移问题，本文引入跨信道稳定性约束项：

$$L_{\text{stable}} = \frac{1}{M}\sum_{m=1}^{M} \left\| z_m^{(c_1)} - z_m^{(c_2)} \right\|_2^2 \tag{43}$$

其中，$z_m^{(c_1)}$与$z_m^{(c_2)}$分别表示同一图像样本在不同信道条件下对应的潜在语义表示，M表示该稳定性损失所用的图像样本数。

为适应不同信道状态下的训练需求，各损失项权重由信道困难度动态调节：

$$\lambda_i = \lambda_i^0 (1 + \alpha_i \Delta_{c,s}) \tag{44}$$

其中，λ_i^0表示第i个损失项的基础权重系数，α_i表示对应的动态调节因子。

为避免训练过程中因瞬时信道波动导致动态权重产生剧烈振荡，本文进一步采用指数滑动平均（exponential moving average, EMA）机制对动态权重进行平滑更新：

$$\lambda_t = (1 - \eta)\lambda_{t-1} + \eta \hat{\lambda}_t \tag{45}$$

其中，η表示平滑更新系数，$\hat{\lambda}_t$表示当前时刻估计得到的动态权重。

通过上述动态联合优化策略，模型能够根据不同信道条件自适应调整训练重点，在保证重建性能的同时增强潜在语义表示的跨信道一致性与稳定性，从而提升复杂无线环境下的鲁棒性和泛化能力。

3　实验设置

3.1　通信资源与潜在表示容量定义

为分别衡量通信资源消耗与潜在表示能力，本文采用信道带宽压缩率（channel bandwidth ratio, CBR）[35]和潜在表示率（latent representation ratio, LRR）作为评价指标。实验在统一CBR条件下比较通信性能，并利用LRR分析潜在表示容量对稳定性的影响。

CBR用于表征系统实际占用的通信资源，定义为：

$$\text{CBR} = \frac{N_c}{N_p} \tag{46}$$

其中，N_c表示发送端实际传输的复信道符号数，N_p表示输入图像像素总数。后续重建性能实验均在统一CBR约束下进行，以保证不同方法之间的公平比较。

为评估潜在表示的信息承载能力，定义LRR：

$$\text{LRR} = \frac{d_z}{H \times W} \tag{47}$$

其中，d_z表示编码器输出潜在表示的维度，H和W分别表示输入图像的高度和宽度。LRR用于刻画潜在空间的表示容量，其值越大，潜在特征的语义表达能力越强。

3.2　实验配置

使用的数据集为BSDS300[36]和DIV2K[37]，平台配置为Ubuntu 18.04.5、NVIDIA GeForce RTX 2080 Ti、Python 3.10及PyTorch 2.5.1。实验在AWGN、Rayleigh和Rician这3类典型无线信道下进行，Rayleigh和Rician信道参数采用连续插值策略，以模拟实际无线信道统计特性的连续变化。SNR范围设置为−20~20 dB。对比模型包括Deep-JSCC、WITT、基于移位窗口Transformer的联合信源信道编码（Swin Transformer joint source-channel coding, SwinJSCC）、基于扩散模型的联合信源信道编码（diffusion-based joint source-channel coding, Diffusion-JSCC）和DF-JSCC。训练尺寸设置为128×128，DFAO采用20步历史窗口统计训练状态，并利用EMA系数0.1平滑更新动态权重。实验实现参数如表1所示。

表1　　实验实现参数

参数	设置
训练裁剪尺寸	128×128
批量大小	8
优化器	AdamW
训练轮数	300
特征通道数	64
语义表示维度	256
DFAO历史窗口长度	20
DFAO指数滑动平均系数	0.1
瑞利信道相关系数	0.9
瑞利衰落参数	0.15~2.5
莱斯K因子	0.1~10.5
瑞利信道噪声惩罚系数	1.30
莱斯信道噪声惩罚系数	1.15

3.3　重建性能与稳定表示评价框架

实验从重建性能和稳定性分析两个方面展开。

（1）重建性能评价

重建质量采用PSNR、SSIM和LPIPS进行评价。

① 固定CBR条件下不同信道的重建性能比较。

② 不同CBR与不同数据集的鲁棒性分析。

③ 系统关键模块消融实验。

④ 重建图像、频谱误差图和边缘误差图的可视化。

（2）潜在表示稳定性评价

为全面评估潜在表示稳定性，统一设置LRR为$\frac{1}{3}$，从分布一致性、统计相关性、结构一致性及聚类特性4个维度进行分析。

① 采用SWD和FD评估分布漂移。

② 利用高斯近似和MINE分析语义—信道互信息。

③ 基于CKA评价结构一致性。

④ 结合t-SNE及类内、类间距离等指标分析潜在表示的紧致性与可分性。

4　实验结果与分析

4.1　不同模型性能测试

为验证所提方法的重建性能，在CBR=$\frac{1}{3}$条件下，对不同模型在不同信道上的性能进行测试，采用PSNR、SSIM和LPIPS作为评价指标，结果如图2~图4所示。随着SNR提高，各模型重建性能持续改善。在低SNR区域，DeepJSCC和WITT受噪声及衰落影响较为明显，重建性能相对较低；SwinJSCC和Diffusion-JSCC通过增强特征建模改善了结构与感知恢复，但在部分衰落信道及低SNR条件下仍存在性能退化。DF-JSCC依据信道状态实施灵活特征编码，在部分SNR区间取得了较好的PSNR或LPIPS结果，体现出一定的信道自适应优势。相比之下，CASCS尤其在低SNR和衰落信道条件下表现出更稳定的结构恢复与感知质量。

4.2　不同CBR、数据集下的性能测试

为评估模型在不同带宽资源条件下的适应能力，分别在BSDS300和DIV2K数据集上设置CBR

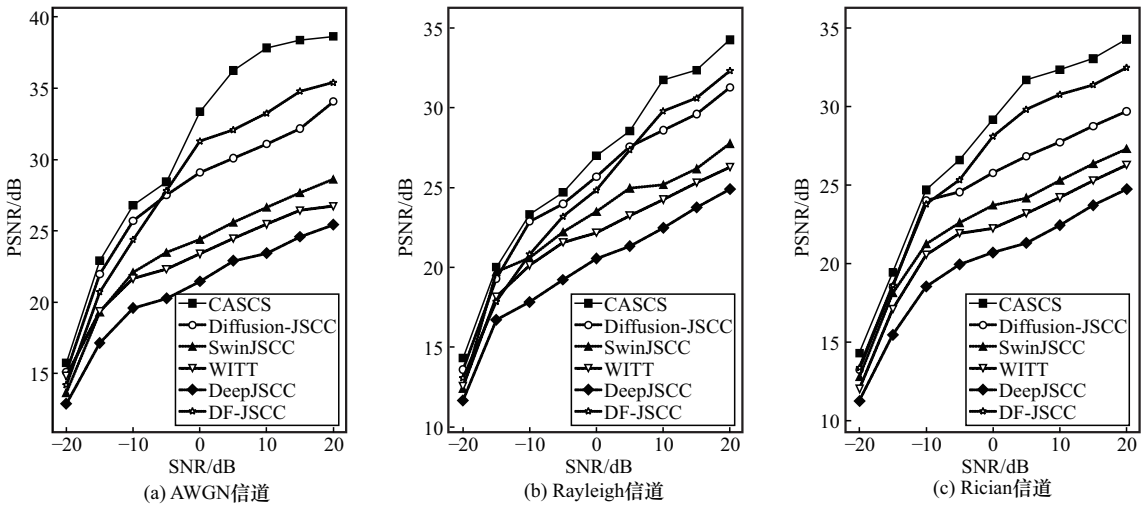


图2　不同模型的PSNR随SNR在不同信道上的变化

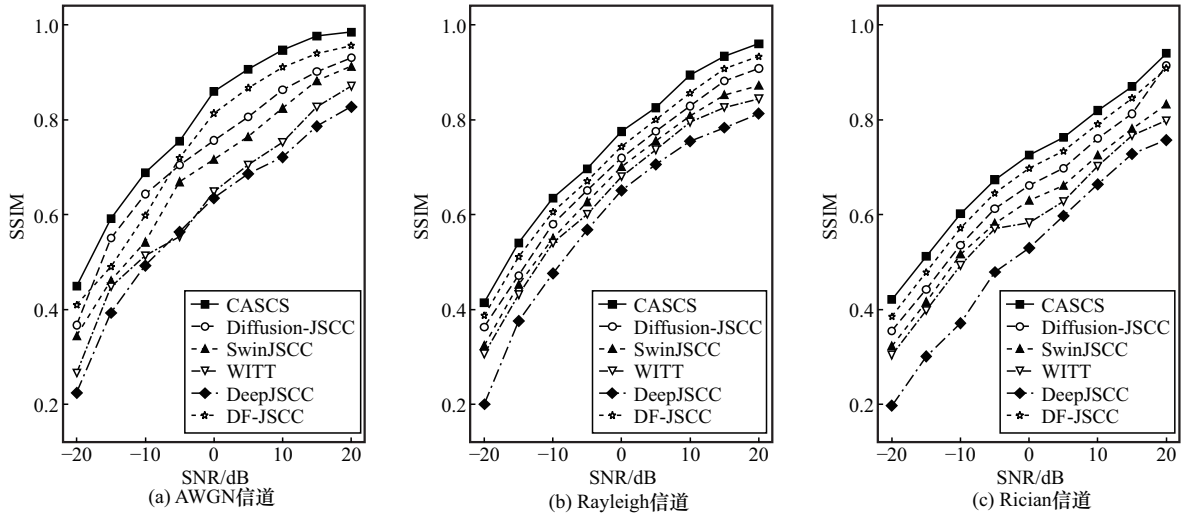


图3　不同模型的SSIM随SNR在不同信道上的变化

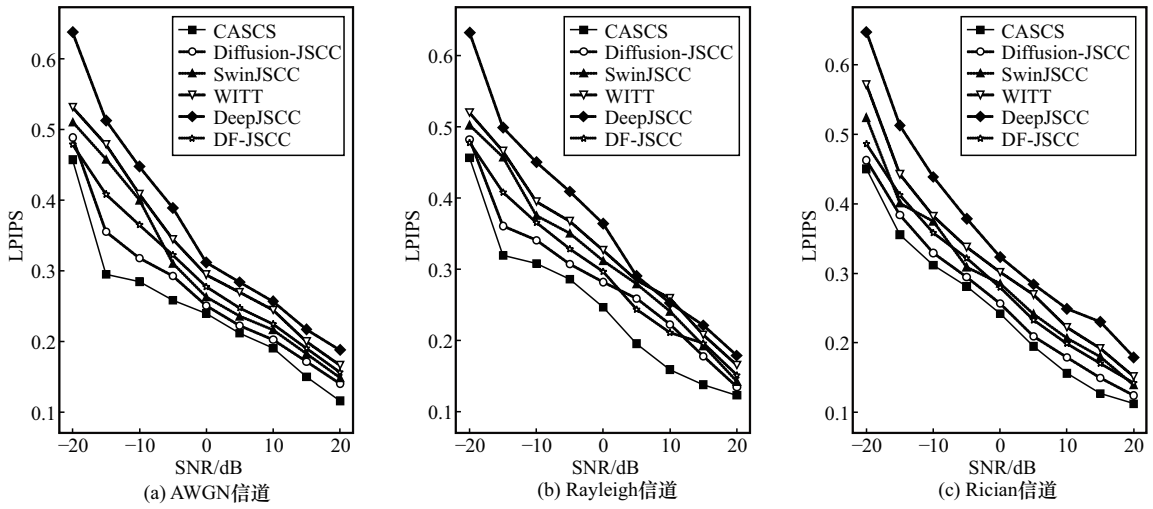


图4　不同模型的LPIPS随SNR在不同信道上下的变化

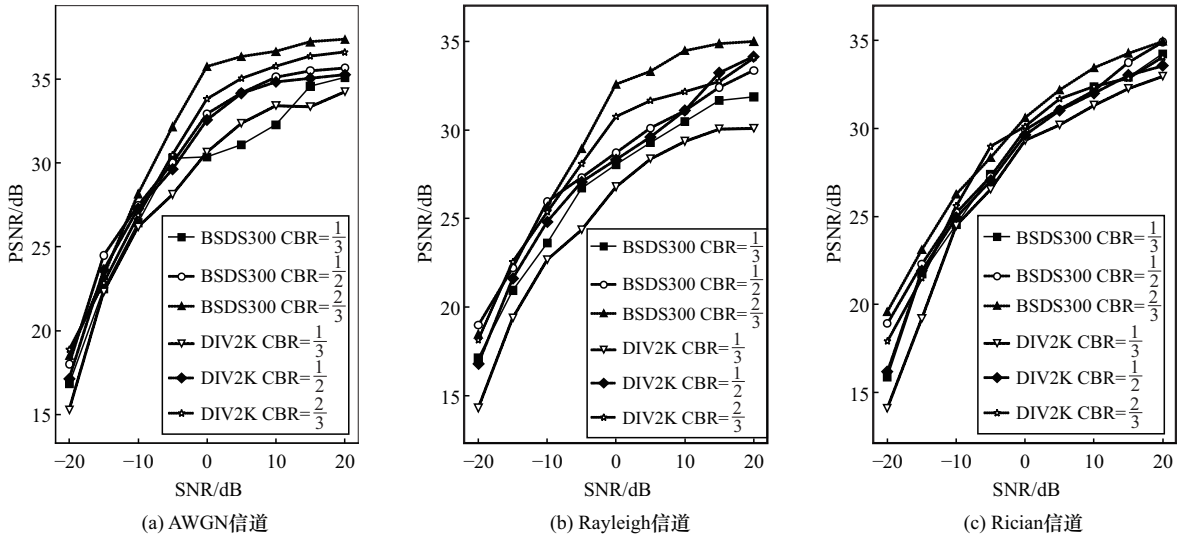
为$\frac{1}{3}$、$\frac{1}{2}$和$\frac{2}{3}$，并在3种信道及−20~20 dB的SNR范围内开展实验，结果如图5~图7所示。图5~图7表明，随着CBR增大，可用信道传输资源增加，各项重建性能持续提升。当CBR进一步增大时，性能增益逐渐趋缓，说明通信资源投入与重建性能之间存在边际收益递减现象。不同信道和数据集下的性能变化趋势基本一致，说明CASCS具有较好的带宽适应性、跨信道鲁棒性和跨数据集泛化能力。

4.3　关键模块消融与稳定潜在表示机制分析

图8给出完整系统、YUV解码、仅SSRL、仅频域约束、仅DFAO、RGB解码和基础模型这7种模型的消融结果。SSRL-only整体性能优于Freq-only和DFAO-only，说明稳定语义表示学习能够有效减弱跨信道分布漂移，提高语义一致性。Freq-only在PSNR和SSIM方面优于DFAO-only，表明频域约束能够增强纹理恢复能力。DFAO虽然性能略低，但曲线更加平滑，说明其有助于改善复杂信道环境下的训练稳定性。CASD(YUV)明显优于RGB Decode，验证了亮度–色度解耦对结构恢复和颜色保持的有效性，但单独优化解码端仍难以完全消除跨信道语义偏移。综合来看，各模块具有良好的互补性，共同提升了CASCS在复杂信道环境下的整体性能。

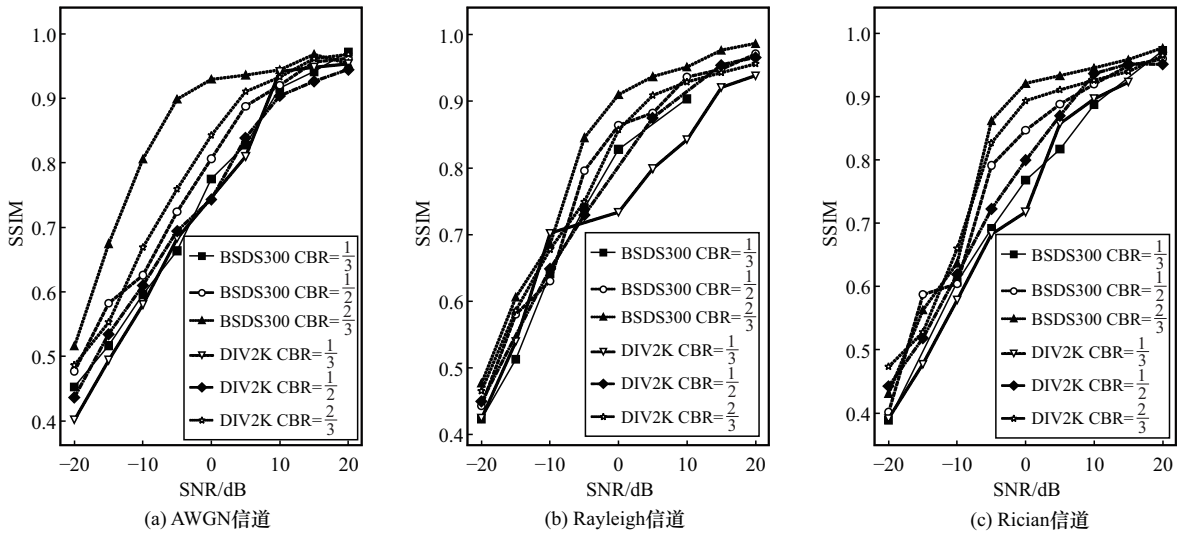
4.4　重建误差可视化

本文从空间域、亮度–色度域、边缘域和频率域开展联合误差分析，评价像素误差、Y/UV通道误差、边缘误差及频谱差异。图9(a)~图9(e)分别对应CASCS、DeepJSCC、WITT、Diffusion-JSCC和



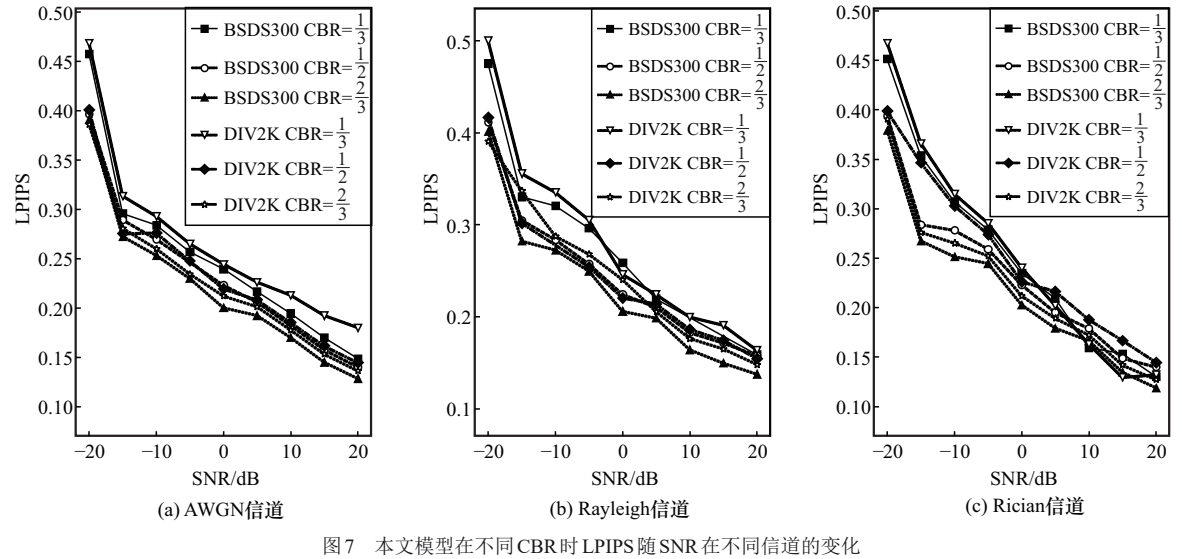
(a) AWGN信道　　(b) Rayleigh信道　　(c) Rician信道

图5　本文模型在不同CBR时PSNR随SNR在不同信道的变化



(a) AWGN信道　　(b) Rayleigh信道　　(c) Rician信道

图6　本文模型在不同CBR时SSIM随SNR在不同信道的变化



(a) AWGN信道　　(b) Rayleigh信道　　(c) Rician信道

图7　本文模型在不同CBR时LPIPS随SNR在不同信道的变化

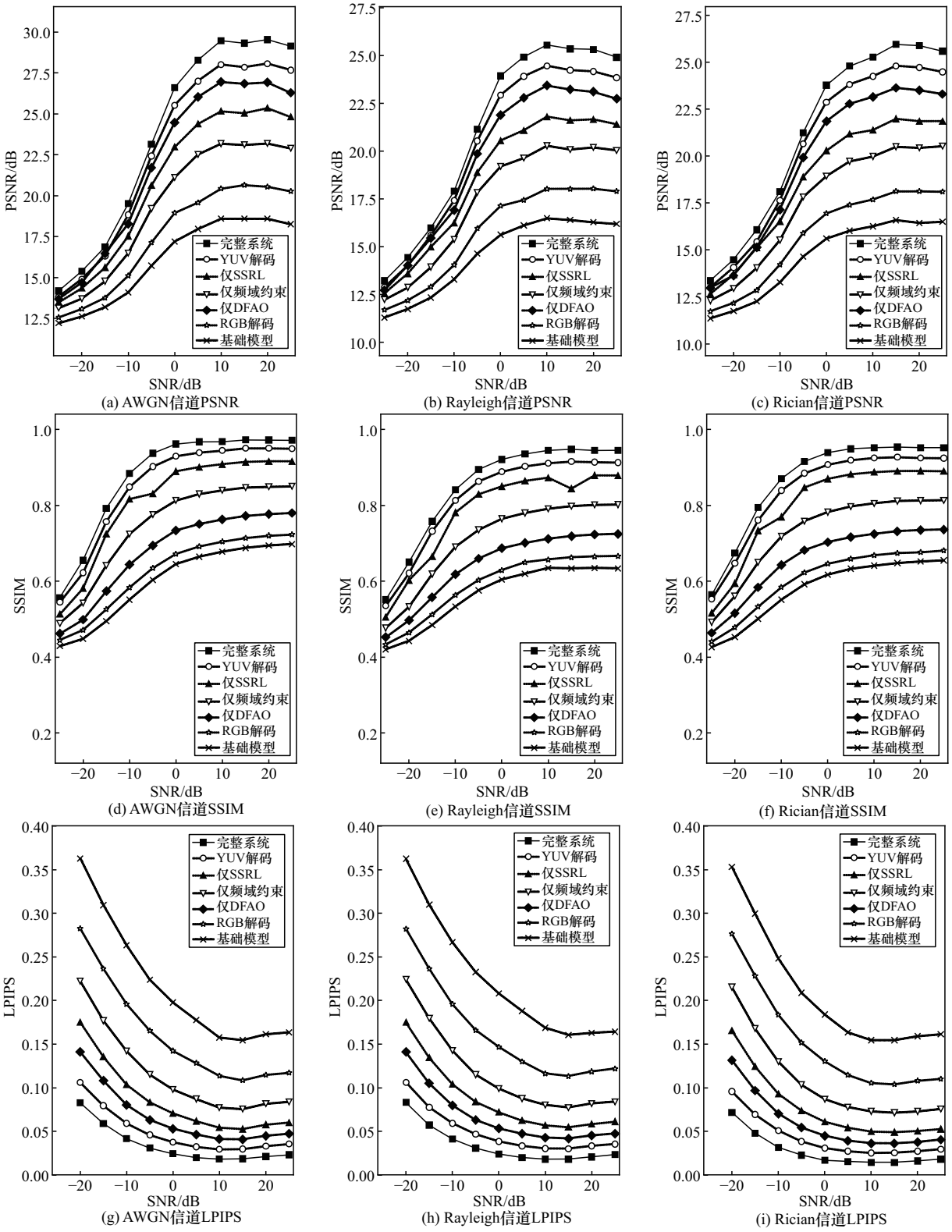


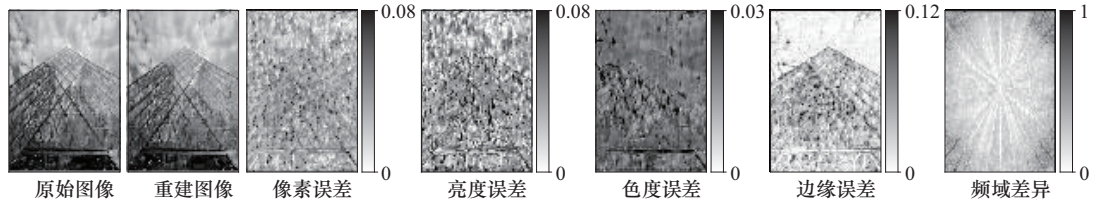
图8　不同信道条件下关键模块消融性能对比

DF-JSCC。图9(a)中，CASCS重建结果与原图在主体轮廓和颜色分布上较为一致，像素及Y/UV通道误差主要集中于玻璃框架和纹理交界区域，体现出较好的结构–颜色协同恢复能力。图9(b)中，Deep-JSCC虽能保持基本轮廓，但存在较明显的颜色偏移和纹理模糊，像素、边缘及频域误差相对较大。

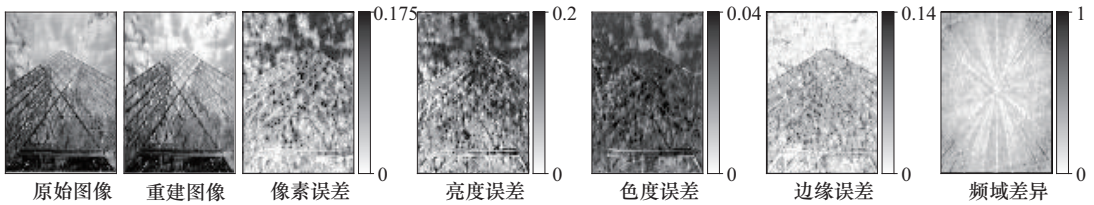
图9(c)中，WITT的频域差异相对较低，但Y/UV通道及边缘区域仍存在较明显误差，说明其频率恢复优势未完全转化为结构感知质量。图9(d)中，Diffusion-JSCC具有较低的频域差异，但局部纹理存在一定平滑现象。图9(e)中，DF-JSCC在感知质量和部分边缘区域恢复方面表现较好，其LPIPS略低于CASCS，但像素误差和亮度通道误差仍较为集中。

4.5　潜在表示稳定性分析

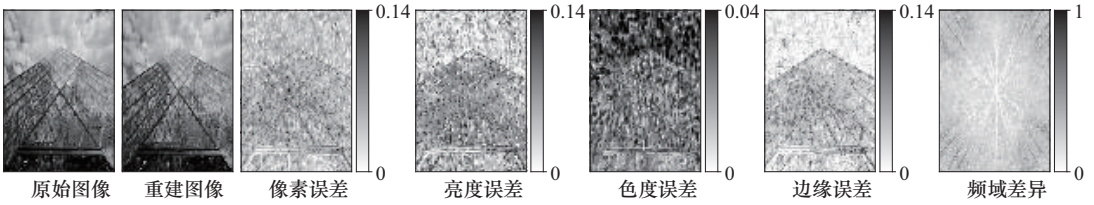
为验证所提方法在宽SNR、多信道环境下学习稳定潜在表示的有效性，本节采用分层评价框架开展训练后离线分析。对于表征跨信道分布漂移与内部结构保持能力的SWD、FD和CKA指标，将DF-JSCC纳入统一比较，以考察信道自适应编码机制对潜在表示稳定性的影响。对于Gaussian Approximation、MINE、t-SNE、类内−类间距离及相关性分析，鉴于相关指标对潜在特征的信道条件含义、维度尺度和表征形式较为敏感，选取潜在表示定义及特征提取层级具有较高可比性的模型进行机制分析。DF-JSCC以信道状态驱动的灵活编码与传输性能预测为主要特征，其潜在表示中的信道相关信息



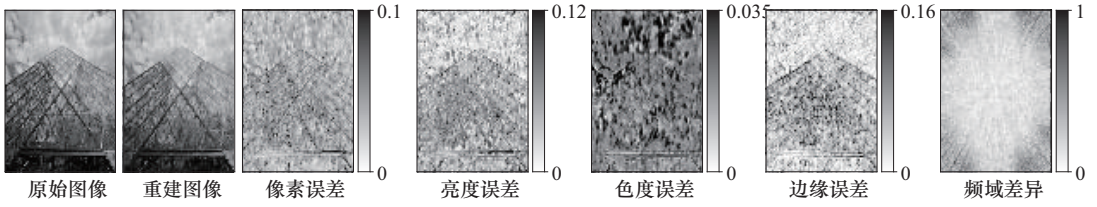
(a) CASCS (PSNR=27.38 dB, SSIM=0.918, LPIPS=0.086 HFER=0.884)



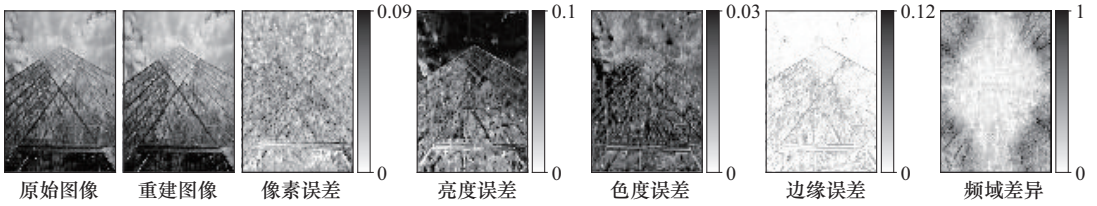
(b) DeepJSCC (PSNR=19.85 dB, SSIM=0.826, LPIPS=0.203, HFER=1.555)



(c) WITT (PSNR=23.78 dB, SSIM=0.747, LPIPS=0.129, HFER=0.667)



(d) Diffusion-JSCC (PSNR=25.39 dB, SSIM=0.842, LPIPS=0.107, HFER=0.523)



(e) DF-JSCC (PSNR=22.47 dB, SSIM=0.910, LPIPS=0.083, HFER=0.623)

图9　不同模型在CBR=$\frac{1}{3}$、SNR=−10 dB、AWGN信道下的多域重建误差对比

同时承担自适应调节作用，因而不宜将其与以降低信道依赖为目标的方法进行同尺度对比。上述指标从分布一致性、结构一致性、统计依赖性及潜空间组织特性等互补维度评价稳定表示机制，不作为模型综合性能的唯一判定依据，且均不参与梯度计算及网络参数更新。

4.5.1　潜在对齐分析

采用SWD和FD评价不同信道状态下潜在表示的分布一致性，两项指标越小，表明模型相对于参考信道状态的潜在分布漂移越弱。由图10和图11可知，DeepJSCC、Diffusion-JSCC和SwinJSCC在跨越较大SNR区间时表现出较高的SWD与FD，说明其潜在特征对信道状态变化较为敏感；WITT在部分区间取得较低的分布距离，但不同区间的表现存在一定波动。DF-JSCC依据信道状态进行灵活特征编码，在若干中等跨度SNR区间获得了低于CASCS的SWD，体现出较好的局部分布适应能力，但其FD在多数区间仍高于CASCS，表明潜在分布在全局均值与协方差结构上的一致性仍有提升空间。相比之下，CASCS在两项指标上整体保持较低水平，且未随信道状态跨度增大出现显著异常波动，表现出更均衡的跨信道分布保持能力。

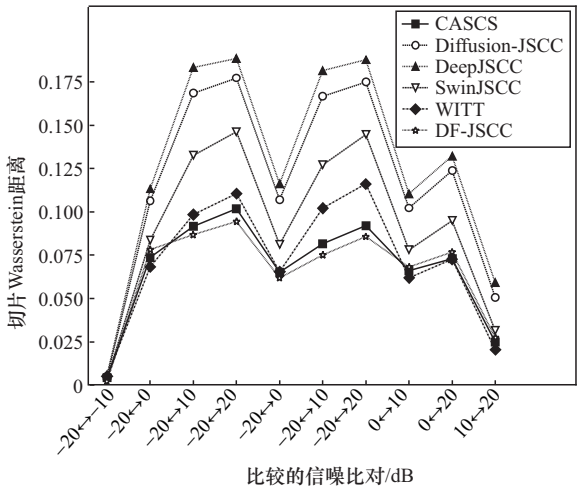


图10　不同模型SWD的变化趋势

综合来看，DF-JSCC的信道自适应机制能够在部分条件下减弱局部潜在分布差异，而CASCS通过信道状态建模与跨信道一致性约束，在局部几何分布和整体统计结构之间取得了更稳定的平衡，从而为宽SNR、多信道环境下的语义重建提供较为一致的潜在特征。

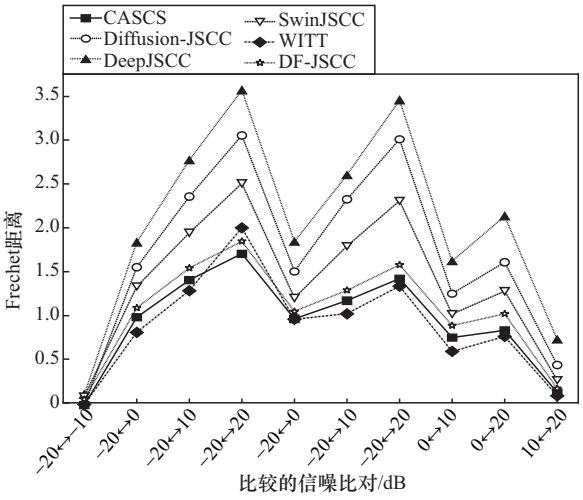


图11　不同模型FD值的变化趋势

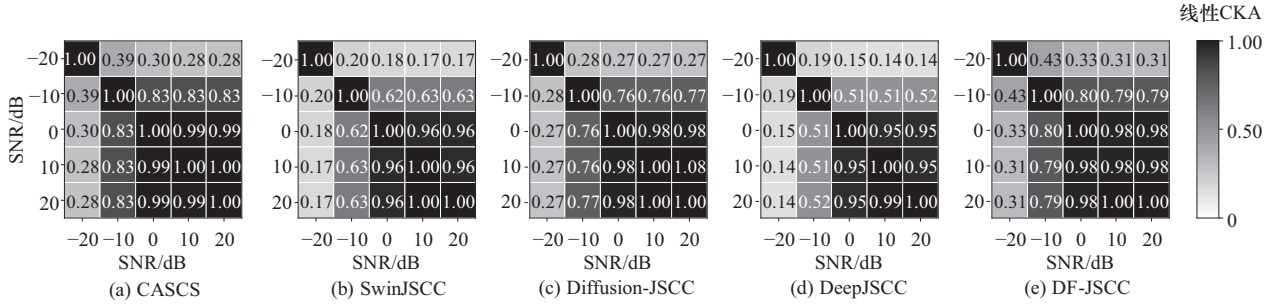
4.5.2　潜在表示结构一致性分析

本文采用线性CKA评价不同SNR条件下潜在表示内部结构关系的保持程度，CKA值越高，表示潜在特征的结构一致性越强。图12~图14表明，各模型在相邻中高SNR区间均保持较高的CKA，而低SNR与中高SNR之间的一致性明显下降，说明强噪声是引起潜在结构变化的主要因素。

具体而言，DeepJSCC在3类信道下的整体CKA较低，表明其潜在结构对信道变化较为敏感；SwinJSCC有所改善，但在低SNR区间仍存在明显结构偏移；Diffusion-JSCC在中高SNR下保持了较好的结构连续性，整体优于前两种方法。DF-JSCC凭借灵活编码与信道适应机制，在AWGN和Rician信道的部分区域取得略高的CKA，但在Rayleigh衰落下的一致性有所下降。相比之下，CASCS在3类信道下均保持较高且相对均衡的结构一致性，尤其在Rayleigh信道下表现更为稳定，其跨信道平均CKA略高于DF-JSCC。结合SWD和FD结果可知，CASCS不仅能够减小潜在表示的分布漂移，还能较好地保持不同信道条件下的内部结构关系，从分布一致性与结构一致性两个层面验证了稳定语义表示学习机制的有效性。

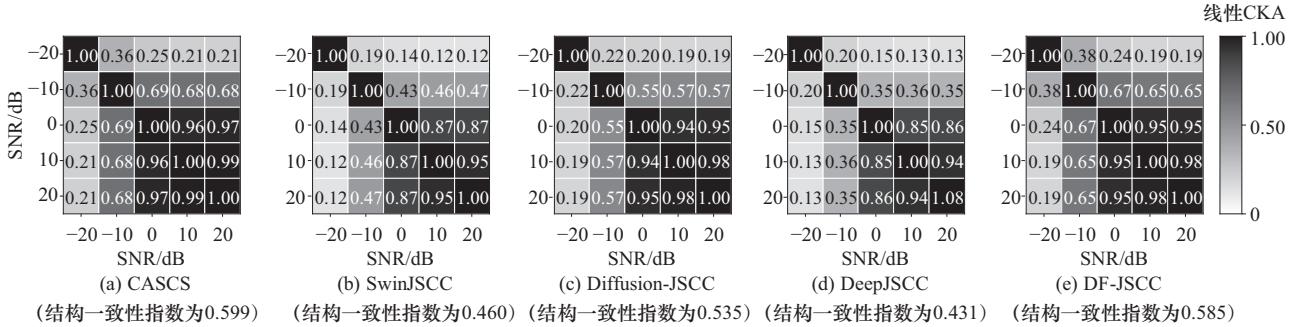
4.5.3　基于互信息的信道不变表示分析

采用高斯近似和MINE估计潜在表示与信道状态之间的互信息，用于衡量潜在表示对信道状态的依赖程度。图15表明，DeepJSCC具有最高互信息，说明其潜在表示仍保留较多信道相关信息；WITT和SwinJSCC能够降低这种相关性，而CASCS在两种估计方法下均进一步降低互信息，其中MINE较



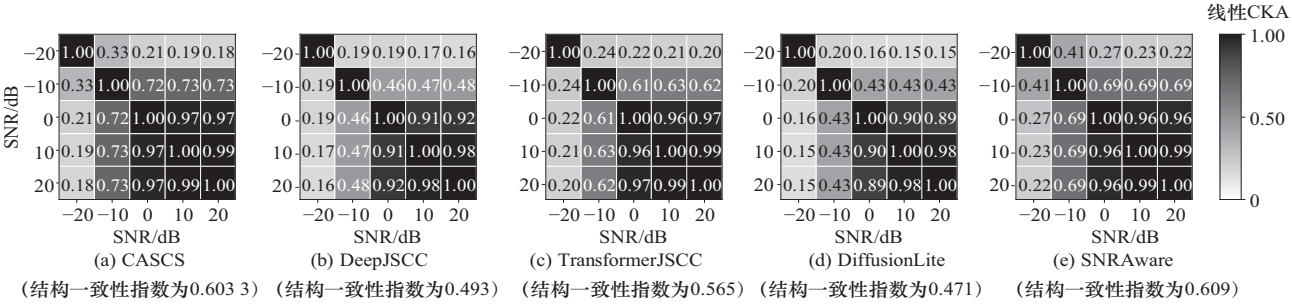
（结构一致性指数为0.671）（结构一致性指数为0.550）（结构一致性指数0.634） （结构一致性指数为0.505） （结构一致性指数为0.673）

图12 不同模型在AWGN信道类型下CKA矩阵



（结构一致性指数为0.599） （结构一致性指数为0.460）（结构一致性指数为0.535）（结构一致性指数为0.431） （结构一致性指数为0.585）

图13 不同模型在Rayleigh信道类型下CKA矩阵



（结构一致性指数为0.603 3） （结构一致性指数为0.493） （结构一致性指数为0.565）（结构一致性指数为0.471） （结构一致性指数为0.609）

图14 不同模型在Rician信道类型下CKA矩阵

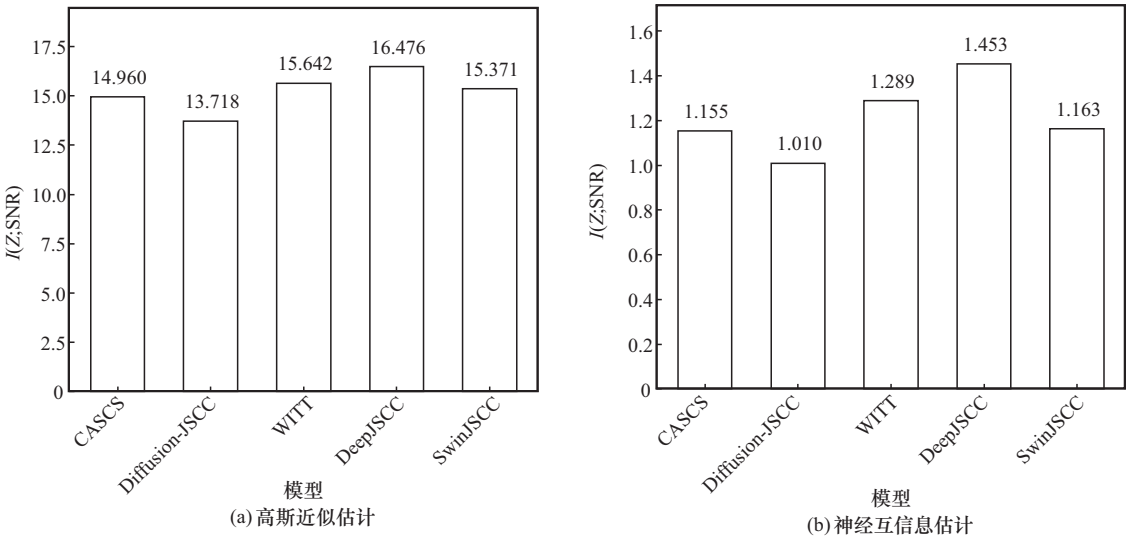


图15 不同的模型高斯近似估计与神经互信息估计

WITT和SwinJSCC分别降低10.40%和0.69%，说明SSRL能够有效减弱潜在表示对信道状态的依赖。

需要指出的是，Diffusion-JSCC虽然获得最低互信息，但重建性能并非最优，说明互信息并非越低越好。CASCS在降低信道相关性的同时仍保持较优的PSNR、SSIM和LPIPS，表明其在语义保持与信道解耦之间取得了更好的平衡。

4.5.4　t-SNE语义特征分布形态

图16给出了不同模型潜在表示的t-SNE可视化，其中颜色表示图像类别，标记表示不同SNR。DeepJSCC中同一样本随SNR变化形成明显轨迹，说明潜在表示易受信道扰动；WITT存在类别重叠；Diffusion-JSCC聚类更加紧凑，但类别边界仍不清晰。相比之下，CASCS能够保持类内紧凑、类间分离的分布结构，体现出更好的跨信道一致性和判别能力。

为进一步定量比较不同模型潜在表示的聚类特性，表2统计了各模型对应的Intra、Inter及Score。DeepJSCC由于Intra最大，说明跨信道漂移最为明显；WITT虽然降低了Intra，但Inter较小，导致Score最低；Diffusion-JSCC模型进一步减小了Intra，但Inter同步下降，因此综合稳定性提升有限。相比之下，CASCS在保持较低Intra为5.164 1的同时获得最大的Inter为21.760 1，最终取得最高Score为4.214，说明所提出的稳定语义表示学习机制能够同时兼顾跨信道稳定性和潜在空间判别能力。

表2　不同模型潜空间聚类评价结果

Model	Intra ↓	Inter ↑	Score ↑
DeepJSCC	33.957 2	8.763 9	0.258
WITT	22.387 6	2.652 0	0.118
Diffusion-JSCC	4.057 7	1.081 8	0.267
CASCS	5.164 1	21.760 1	4.214

综合SWD、FD、CKA、互信息及t-SNE结果，CASCS能够有效降低跨信道分布偏移和信道相关性，同时保持良好的结构一致性和类别可分性，进一步验证了稳定语义表示学习机制的有效性。

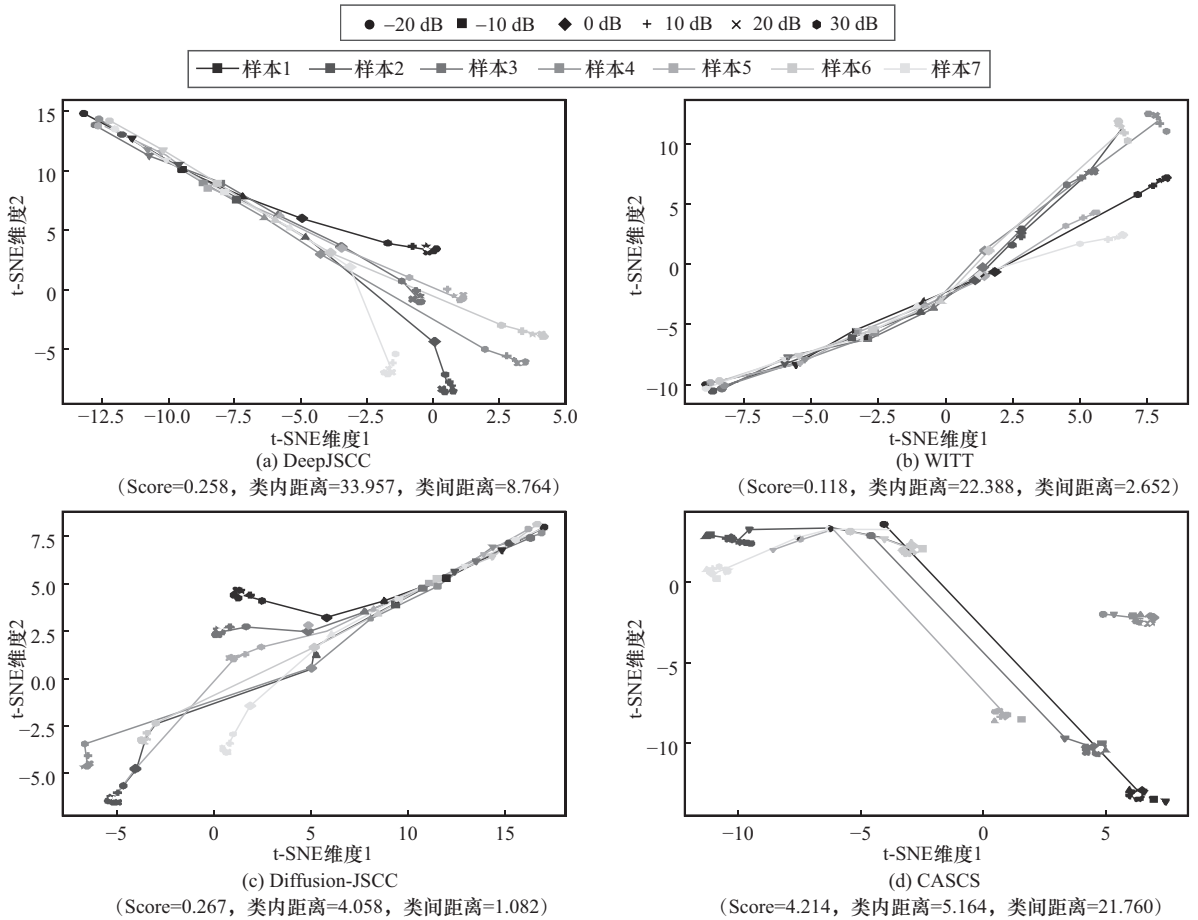


图16　不同模型的t-SNE潜空间分布

表3 不同模型稳定性指标与重建性能指标Pearson相关系数

模型	SWD-PSNR	FD-PSNR	SWD-SSIM	FD-SSIM	MINE-LPIPS	MINE-SSIM	CKA-PSNR
CASCS	−0.61	−0.59	−0.89	−0.88	−0.46	0.45	0.78
DeepJSCC	−0.89	−0.79	−0.91	−0.81	−0.48	0.48	0.98
SwinJSCC	−0.84	−0.75	−0.97	−0.90	−0.54	0.54	0.91
WITT	−0.96	−0.88	−0.96	−0.89	−0.44	0.46	0.99
Diffusion-JSCC	−0.58	−0.54	−0.64	−0.60	0.31	−0.28	0.56
DF-JSCC	−0.74	−0.69	−0.78	−0.73	—	—	0.71

4.5.5 稳定性指标与重建性能相关性分析

为分析潜在表示稳定性与重建性能之间的关系，本文采用Pearson相关系数分析不同稳定性指标与重建指标之间的线性相关关系，结果如表3所示，其中，—表示本文未对DF-JSCC开展MINE互信息估计，因而未计算MINE与LPIPS、SSIM之间的Pearson相关系数。Pearson相关系数仅反映变量之间的线性关联程度，不表示模型性能排序，也不能用于推断因果关系。6种模型的SWD和FD与PSNR、SSIM均呈负相关，表明较小的跨信道潜在分布漂移通常对应较好的图像重建质量；CKA与PSNR均呈正相关，说明较高的潜在结构一致性与较好的重建性能具有一致的变化趋势。

不同模型的MINE与重建性能呈现不同的相关方向。CASCS、DeepJSCC、WITT和SwinJSCC的MINE与LPIPS呈负相关、与SSIM呈正相关，说明一定程度的信道相关信息可能参与信道自适应过程；Diffusion-JSCC的MINE与LPIPS呈弱正相关、与SSIM呈弱负相关，表明其信道统计依赖与重建性能之间的线性关系相对较弱。上述结果表明，信道统计依赖与重建性能并不存在统一的单调关系，模型需要在降低非必要信道依赖、保留有效信道适应信息和保持语义内容之间取得平衡。综合表3及前述实验，潜在分布一致性、结构一致性及信道统计依赖程度与重建性能呈现不同程度的关联趋势，为稳定潜在表示与宽SNR、多信道环境下图像重建性能之间的关系提供支持。

4.6 模型复杂度分析

为评估模型部署能力，本文从参数量、计算复杂度（用GFLOPs衡量）、显存占用和推理时延4个方面，对CASCS与DeepJSCC、SwinJSCC和Diffusion-JSCC进行比较，测试条件为输入256×256、批量大小为1。表4结果表明，CASCS的参数量和计算复杂度高于DeepJSCC，但明显低于Diffusion-JSCC，整体上取得了重建性能、稳定性与计算开销之间的折中。

表4 不同模型复杂度对比分析

模型	参数量	计算复杂度	显存占用/GB	推理时延/ms
DeepJSCC	6×10^6	25	1.2	5
SwinJSCC	20×10^6	120	4.5	18
Diffusion-JSCC	80×10^6	480	11	120
CASCS	23×10^6	82	4.7	38

5 结束语

本文围绕宽SNR与复杂信道环境下语义通信图像重建的稳定性问题，提出基于信道感知调控的语义通信图像重建框架。本文从潜在语义表示稳定性的角度出发，结合信道状态辅助调制、YUV结构解耦恢复以及动态稳定优化机制，对复杂无线环境下的跨信道语义一致性问题进行了初步探索。实验结果表明，所提出方法在AWGN、Rayleigh及Rician等多种信道条件下能够保持较稳定的重建性能，并在潜在特征分布一致性、互信息耦合程度和t-SNE潜空间形态等方面表现出一定的跨信道稳定性趋势。

本文采用完美CSI假设，主要用于分析不同信道统计特性对潜在表示稳定性的影响，而非用于传统物理层均衡。未来将进一步研究非理想CSI条件下的语义通信系统，将信道估计、信道补偿等物理层技术与稳定语义表示学习相结合，实现物理层信道补偿与语义层表示优化的协同设计，并进一步扩展至时变信道、多天线及未知信道环境，以提升复杂无线环境下语义通信系统的鲁棒性与泛化能力。

参考文献：

[1] Gündüz D, Qin Z J, Aguerri I E, et al. Beyond transmitting bits: context, semantics, and task-oriented communications[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(1): 5-41.

[2] 张平, 牛凯, 姚圣时, 等. 面向未来的语义通信: 基本原理与实现方法[J]. 通信学报, 2023, 44(5): 1-14.
Zhang P, Niu K, Yao S S, et al. Semantic communications for future: basic principle and implementation methodology[J]. Journal on Communications, 2023, 44(5): 1-14.

[3] 刘传宏, 郭彩丽, 杨洋, 等. 面向智能任务的语义通信: 理论、技术和挑战[J]. 通信学报, 2022, 43(6): 41-57.
Liu C H, Guo C L, Yang Y, et al. Intelligent task-oriented semantic communications: theory, technology and challenges[J]. Journal on Communications, 2022, 43(6): 41-57.

[4] 牛凯, 张平. 语义通信的数学理论[J]. 通信学报, 2024, 45(6):7-59.
Niu K, Zhang P. A mathematical theory of semantic communication[J]. Journal on Communications, 2024, 45(6): 7-59.

[5] 石光明, 杨旻曦, 高大化, 等. 面向语义信息直传的通信架构[J]. 通信学报, 2023, 44(5): 15-27.
Shi G M, Yang M X, Gao D H, et al. Communication framework for directed transmission of informative semantic[J]. Journal on Communications, 2023, 44(5): 15-27.

[6] 秦志金, 赵茭茭, 李凡, 等. 多模态语义通信研究综述[J]. 通信学报, 2023, 44(5): 28-41.
Qin Z J, Zhao T T, Li F, et al. Survey of research on multimodal semantic communication[J]. Journal on Communications, 2023, 44(5): 28-41.

[7] 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.

[8] 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.

[9] Xie H Q, Qin Z J, Li G Y, et al. Deep learning enabled semantic communication systems[J]. IEEE Transactions on Signal Processing, 2021, 69: 2663-2675.

[10] Yang K, Wang S X, Dai J C, et al. WITT: a wireless image transmission transformer for semantic communications[C]//Proceedings of the ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2023: 1-5.

[11] 江沸菠, 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.

[12] 张平, 戴金晟, 张育铭, 等. 面向语义通信的非线性变换编码[J]. 通信学报, 2023, 44(4): 1-14.
Zhang P, Dai J S, Zhang Y M, et al. Nonlinear transform coding for semantic communications[J]. Journal on Communications, 2023, 44(4): 1-14.

[13] 罗鹏, 刘月玲, 张聿远, 等. 高效融合全局和局部上下文特征的语义通信系统[J]. 通信学报, 2023, 44(7): 14-25.
Luo P, Liu Y L, Zhang Y Y, et al. Semantic communication system with efficient integration of global and local context features[J]. Journal on Communications, 2023, 44(7): 14-25.

[14] Kurka D B, Gündüz D. Bandwidth-agile image transmission with deep joint source-channel coding[J]. IEEE Transactions on Wireless Communications, 2021, 20(12): 8081-8095.

[15] Ding M Z, Li J H, Ma M Y, et al. SNR-adaptive deep joint source-channel coding for wireless image transmission[C]//Proceedings of the ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2021: 1555-1559.

[16] Xie S J, He H T, Li H R, et al. Deep learning-based adaptive joint source-channel coding using hypernetworks[C]//Proceedings of the 2024 IEEE International Mediterranean Conference on Communications and Networking (MeditCom). Piscataway: IEEE Press, 2024: 191-196.

[17] Peng X, Qin Z J, Tao X M, et al. A robust image semantic communication system with multi-scale vision transformer[J]. IEEE Journal on Selected Areas in Communications, 2025, 43(4): 1278-1291.

[18] 杨舒涵, 申滨, 黄晓舸. 基于深度灵活编码策略与性能预测的无线语义图像传输系统[J]. 通信学报, 2025, 46(5): 29-46.
Yang S H, Shen B, Huang X G. Wireless semantic image transmission system based on deep flexible coding strategy and performance prediction[J]. Journal on Communications, 2025, 46(5): 29-46.

[19] Blau Y, Michaeli T. The perception-distortion tradeoff[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 6228-6237.

[20] Poynton C. Digital video and HD[M]. Waltham: Morgan Kaufmann, 2012.

[21] Shannon C E. A mathematical theory of communication[J]. Bell System Technical Journal, 1948, 27(3): 379-423.

[22] Goldsmith A. Wireless communications[M]. New York: Cambridge University Press, 2005.

[23] Rappaport T S. Wireless communications: principles and practice[M]. 2nd ed. Upper Saddle River: Prentice Hall PTR, 2000.

[24] Huynh-Thu Q, Ghanbari M. Scope of validity of PSNR in image/video quality assessment[J]. Electronics Letters, 2008, 44(13): 800-801.

[25] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.

[26] Zhang R, Isola P, Efros A A, et al. The unreasonable effectiveness of deep features as a perceptual metric[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 586-595.

[27] Kolouri S, Rohde G K, Hoffmann H. Sliced Wasserstein distance for learning Gaussian mixture models[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 3427-3436.

[28] Heusel M, Ramsauer H, Unterthiner T, et al. GANs trained by a two time-scale update rule converge to a local Nash equilibrium[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM Press, 2017: 6629-6640.

[29] Kornblith S, Norouzi M, Lee H, et al. Similarity of neural network representations revisited[C]//Proceedings of the 36th International Conference on Machine Learning. New York: ACM Press, 2019: 3519-3529.

[30] Cover T M, Thomas J A. Elements of information theory[M]. 2nd ed. Hoboken: John Wiley & Sons Inc, 2006.

[31] Belghazi M I, Baratin A, Rajeshwar S, et al. Mutual information neural estimation[C]//Proceedings of the 35th International Conference on Machine Learning. New York: ACM Press, 2018: 531-540.

[32] Laurens V D M, Hinton G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9: 2579-2605.

[33] Tishby N, Pereira F C, Bialek W. The information bottleneck method [PP]. V1. (2000-04-24) [2026-06-17]. arXiv: arXiv.physics/0004057.

[34] Alemi A A, Fischer I, Dillon J V, et al. Deep variational information bottleneck[PP]. V7. (2019-10-23) [2026-06-17]. arXiv: arXiv. 1612.00410.

[35] Dai J C, Wang S X, Tan K L, et al. Nonlinear transform source-channel coding for semantic communications[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(8): 2300-2316.

[36] Martin D, Fowlkes C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]//Proceedings of the Eighth IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2001: 416-423.

[37] 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.

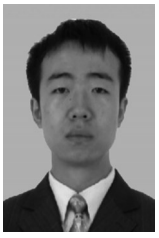
[作者简介]



杨颜博（1983– ），男，博士，内蒙古科技大学副教授、硕士生导师，主要研究方向为大数据、人工智能算法的理论与应用、网络编码理论与应用等。



高晓婷（2000– ），女，内蒙古科技大学硕士生，主要研究方向为语义通信。



张嘉伟（1985– ），男，博士，西安电子科技大学副教授，主要研究方向为数据安全、网络安全、云计算安全和区块链等。



马建峰（1964– ），男，博士，西安电子科技大学教授、博士生导师，主要研究方向为应用密码学、无线网络安全、数据安全、移动智能系统安全。



高丽丽（1976– ），女，内蒙古科技大学副教授，主要研究方向为无线通信、室内定位技术。