

基于深度压缩感知的联合信源信道编码方法研究

董浩¹, 李劭辉², 阚诺文¹, 郑紫阳¹, 戴文睿^{1*}, 熊红凯¹

(1. 上海交通大学电子信息与电气工程学院, 上海 200240; 2. 清华大学深圳国际研究生院, 广东深圳 518055)

摘要: 无线图像传输面临着带宽和计算资源的双重挑战, 在节点计算能力有限的物联网等应用场景中尤为突出. 联合信源信道编码(Joint Source-Channel Coding, JSCC)能够同时优化信源和信道编码, 逐渐成为无线图像传输中一个重要研究方向. 近年来, 基于深度学习的JSCC方法受到广泛关注, 其通过端到端训练实现编码器与解码器的联合优化. 然而, 大多数基于深度学习的JSCC方法的编码器涉及大量的线性与非线性运算, 导致计算复杂度较高, 难以应用于物联网边缘计算节点等计算资源受限的设备. 为实现轻量化的编码过程, 本文提出了一种基于深度压缩感知的联合信源信道编码方法BCS-JSCC(Block Compressive Sensing-Joint Source Channel Coding), 实现对于编解码器的端到端优化. 该方法在编码端设计可学习尺度二值化测量的压缩感知采样, 实现噪声环境下匹配解码器的轻量化编码方法; 在解码端, 基于MMSE(Minimum Mean Squared Error)准则求解测量值传输的线性逆问题, 获得信道噪声敏感的初始重建, 抑制噪声对参数复用重建网络的影响. 与现有的基于深度学习的JSCC方法相比, 在保持编码端每像素浮点计算次数(Floating Point operations per pixel, FLOPs per pixel)相同的条件下, 本文所提出的BCS-JSCC方法在高信噪比条件下可以取得更好的传输性能. 在低算力(0.10 K FLOPs/pixel)情况下, 优势更为明显. 本文提出的BCS-JSCC方法编码器构造简单、计算量低, 适用于物联网边缘计算节点等低算力设备部署.

关键词: 无线图像传输; 联合信源信道编码; 深度学习; 编码器; 压缩感知

基金项目: 国家自然科学基金(No.62320106003)

中图分类号: TN919.8

文献标识码: A

文章编号: 0372-2112(2025)07-2178-15

电子学报URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250002

Research on Joint Source-Channel Coding Method Based on Deep Compressive Sensing

DONG Hao¹, LI Shao-hui², KAN Nuo-wen¹, ZHENG Zi-yang¹, DAI Wen-rui^{1*}, XIONG Hong-kai¹

(1. School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;

2. Tsinghua Shenzhen International Graduate School, Shenzhen, Guangdong 518055, China)

Abstract: Wireless image transmission faces the dual challenges of bandwidth and computing resources, which is particularly prominent in application scenarios such as the Internet of Things where node computing power is limited. Joint source-channel coding (JSCC) can optimize both source and channel coding, and has gradually become an important research direction in wireless image transmission. In recent years, deep learning-based JSCC methods have received widespread attention, which achieve joint optimization of encoders and decoders through end-to-end training. However, most encoders of deep learning-based JSCC methods involve a large number of linear and nonlinear operations, resulting in high computational complexity and difficulty in application to devices with limited computing resources such as edge computing nodes in the Internet of Things. In order to achieve a lightweight coding process, this paper proposes a joint source channel coding method BCS-JSCC (Block Compressive Sensing-Joint Source Channel Coding) based on deep compressed sensing to achieve end-to-end optimization of the codec. This method designs a compressed sensing sampling of learnable scale binary measurement at the encoding end to realize a lightweight encoding method that matches the decoder in a noisy environment; at the decoding end, the linear inverse problem of measurement value transmission is solved based on the MMSE (Minimum Mean Squared Error) criterion to obtain the initial reconstruction sensitive to channel noise and suppress the influence of noise on the parameter reuse reconstruction network. Compared with the existing JSCC method based on deep

learning, the BCS-JSCC method proposed in this paper can achieve better transmission performance under high signal-to-noise ratio conditions while keeping the number of floating point operations per pixel (FLOPs per pixel) at the encoding end the same. The advantage is more obvious under low computing power (0.10 K FLOPs/pixel). The encoder of the BCS-JSCC method proposed in this paper has a simple structure and low computational complexity, and is suitable for deployment on low computing power devices such as edge computing nodes of the Internet of Things.

Key words: wireless image transmission; joint source-channel coding; deep learning; encoder; compressive sensing
Foundation Item(s): National Natural Science Foundation of China (No.62320106003)

1 引言

传统的通信系统通过信源编码、信道编码两步编码过程实现对图像的编码:(1)信源编码如 JPEG (Joint Photographic Experts Group)、BPG (Better Portable Graphics) 等对图像进行压缩以去除冗余信息,减少传输的信息量;(2)信道编码如 LDPC (Low-Density Parity-Check)、RS (Reed-Solomon) 等对压缩后的比特流增加冗余,以避免信道噪声的干扰. 这一编码方法称为分离式信源信道编码 (Separate Source-Channel Coding, SSCC). 基于香农分离定理^[1],码字长度趋于无穷长时,SSCC 在理论上可以达到最优性能. 但是,实际通信系统中的编码长度有限,SSCC 无法达到无限码长下的理论最优性能. 此外,从设计思路上看,信源编码减少冗余而信道编码增加冗余,如果分别设计,难免产生额外的计算开销. 因此,可以将信源编码、信道编码二者联合进行设计,以实现传输性能及计算量的优化,此种编码方法称为联合信源信道编码 (Joint Source-Channel Coding, JSCC).

联合信源信道编码涵盖多种设计思路,包括级联式信源信道编码器^[2](信源编码与信道编码之间存在信息交互以实现协同优化)、非均匀差错保护策略^[3,4](根据信源符号的重要性施加不同等级的冗余保护),以及一体式架构^[5-7](将信源压缩与信道抗噪融合为单一编码模块). 近年来,基于深度学习技术设计 JSCC 系统引起了广泛的关注^[8],其从设计思路上来说属于一体式 JSCC 架构. 基于深度学习的 JSCC 方法利用神经网络统一替换信源编码和信道编码模块,其不再需要对信源编码、信道编码进行显式的模块划分,通过数据驱动的方式学习信源到信道符号的最优紧耦合特征映射. 在这一过程中,神经网络不仅学习如何去除信源中的冗余信息实现有效压缩,还在同一映射过程中融入对信道特性的鲁棒性处理,从而隐式实现信源压缩与信道抗噪的联合优化. 编码器所输出的图像语义特征,本质上是信源统计特性与信道传输约束下共同作用的联合最优表征.

传统的 JSCC 方法通常仅在编解码器一方进行单端优化,而未考虑编码器和解码器之间的协同优化,这种单端优化策略往往无法达到整体的最优性能. 基于深度学习的 JSCC 方法通过端到端训练对整个系统进行联

合优化,相比于传统的 JSCC 方法独立优化编码器与解码器,能够达到更优的性能. 目前,已有大量的工作对基于深度学习的 JSCC 方法展开了研究^[9-19]. 然而,现有的研究工作大多侧重于提高图像信号的重构质量,而忽视了编码器的计算复杂度. 编码器通常包含大量的卷积、注意力机制等线性与非线性操作,导致计算复杂度较高. 对于计算资源受限的场景(如无线传感器网络等),信号采集设备只能执行低复杂度的编码操作,上述计算复杂的 JSCC 系统在此类应用场景下并不适用.

压缩感知^[20] (Compressive Sensing, CS) 的非对称计算架构则为此类场景下的 JSCC 系统提供了一种新的设计思路. 压缩感知的信号采样过程非常简单,仅通过一次矩阵乘法就可以获取高维信号的简洁表示,适用于无线传感网络等计算资源受限的场景^[21,22]. 目前已有一些研究工作在 JSCC 系统的设计中引入压缩感知,Chen 等人^[23]在编码端依次对信号进行压缩感知采样与 LDPC 编码,在解码端通过合并压缩感知与 LDPC 的因子图进行迭代联合解码. Zheng 等人^[24]将图像信号经块压缩感知采样后,将测量值量化为三元符号,通过三元 SC-LDGM (Systematic Convolutional Low Density Generator Matrix) 编码进行信道编码. 上述研究工作虽然考虑了压缩感知与传统编码技术的结合,但仍局限于编码器或解码器的单端优化,无法充分发挥系统的整体性能. 受现有的基于深度学习的压缩感知方法启发,本文提出基于块压缩感知的联合信源信道编码方法 BCS-JSCC (Block Compressive Sensing-Joint Source Channel Coding),联合优化编解码器. 在编码端,BCS-JSCC 仅通过压缩感知的采样过程实现对信号的联合编码,将采样所得的测量值作为 JSCC 传输的信道符号;在解码端,基于压缩感知的重建方法实现对信号的联合解码. 实验表明,本文所提出的方法在编码端计算资源受限的场景下,可以实现更优的图像传输性能.

本文的主要贡献如下:

(1) 提出了基于深度压缩感知模型的联合信源信道编码系统,通过端到端训练实现编解码器的联合优化;

(2) 在编码端,设计了压缩感知采样的轻量化编码方法,在噪声环境下通过可学习尺度的二值化测量匹

解编码器,维持传输性能的同时降低计算复杂度;

(3)在解码端,根据信号与噪声的统计特性,基于最小均方误差(Minimum Mean Squared Error, MMSE)准则求解测量值传输的线性逆问题,设计信道噪声敏感初始重建的参数复用重建网络,有效抑制噪声影响.

在保持编码器每像素浮点计算次数(Floating Point Operations per pixel, FLOPs per pixel)相同的条件下,与现有的基于深度学习的JSCC方法进行对比,本文所提出的BCS-JSCC方法在高信噪比条件下具有显著优势.在编码器低算力(0.10 K FLOPs/pixel)条件下,BCS-JSCC在各类信噪比环境下均展现出优于对比模型的整体性能.

2 相关工作

过去几年,深度学习技术被广泛应用于联合信源信道编码方法设计中. Bourtsoulatze 等人^[9]最先将深度学习技术应用于图像的联合信源信道编码的设计中,其提出的DeepJSCC模型的图像传输性能要优于传统的分离式信源信道编码方法. Kurka 等人^[10]在DeepJSCC模型的基础上,通过分层处理与反馈调整,实现了在变化的信噪比环境中鲁棒的图像传输. Kurka 等人^[11]考虑了图像随时间或频率逐步分层传输的场景,实现图像的自适应带宽传输. Xu 等人^[12]提出了一种基于通道注意力机制的ADJSCC(Attention DL based JSCC)模型,通过引入注意力模块,可以自适应各种信噪比条件. Yang 等人^[13]通过引入策略网络动态生成二进制掩码,基于图像内容与信道条件自适应控制速率. Wang 等人^[14]研究了分布式场景下图像源的联合信源信道编码方法,在解码器端采用交叉注意力模块来计算不同源之间的相关性. Dai 等人^[15]采用非线性变换方法学习图像的隐表示,并构建嫡模型作为该隐表示的先验,从而隐式地逼近真实源的概率分布. Wang 等人^[16]在文献[15]的基础上进一步改进,通过引入上下文嫡模型来更好地捕捉隐表示的空间相关性. 此外,他们将带宽比和信道状态信息嵌入到编解码器参数中,从而支持不同的带宽比和信道状态. Erdemir 等人^[17]在解码器引入生成对抗网络(Generative Adversarial Nets, GAN),获得了相比于传统的DeepJSCC方案更好的感知质量指标. Wu 等人^[18]基于Vision Transformer(ViT)设计了面向MIMO(Multiple-Input Multiple-Output)系统的JSCC模型ViT-MIMO,能够自适应地基于原图像与信道条件学习特征映射与功率分配. Yang 等人^[19]基于Swin Transformer设计联合编码器与解码器,通过引入通道适应模块与码率适应模块,基于信道状态信息和目标传输速率缩放隐表示,从而在单个模型中灵活适应不同通道状态和传输速率.

尽管上述基于深度学习的JSCC系统在不同的信道条件下可以提供良好的图像传输效果,但其编码器通常包含大量卷积、注意力机制等线性与非线性操作,导致计算复杂度较高,这对于计算资源有限的编码设备而言并不适用. 在此背景下,如何降低编码器的计算复杂度成为关键问题. 压缩感知技术在编码过程仅需执行矩阵乘法操作,其在此类场景下提供了一种具有较高可行性的信号编码解决方案.

近年来,深度学习技术与压缩感知的结合引起了学术界的广泛关注^[25]. 基于深度学习的压缩感知技术通过端到端的数据驱动学习方式,自适应提取信号的结构特征并实现高效的信号重建. Mousavi 等人^[26]提出了基于全连接架构的SDA(Stacked Denoising Autoencode)网络模型,首次将深度学习方法应用于压缩感知重建当中. Kulkarni 等人^[27]设计了基于卷积神经网络的ReconNet模型,实现了压缩感知的高效重建. Shi 等人^[28]将测量矩阵与压缩感知重建层联合优化,通过端到端训练实现自适应采样与重建.

上述基于深度学习的压缩感知模型通常将神经网络视为“黑盒”进行训练,这种方法缺乏对网络内部机制的解释,导致模型的可解释性较差. 相比之下,优化思想启发的深度压缩感知模型通过神经网络对传统优化算法进行展开,这样的设计不仅使网络具备了可解释性,同时也进一步提高了图像的重建质量. Zhang 等人^[29]提出了一种结合近似消息传递算法(Approximate Message Passing, AMP)的压缩感知重建模型AMP-Net,通过神经网络结构模拟AMP算法的迭代过程. 裴翰奇等人^[30]使用可学习卷积层替换传统平滑投影Landweber算法(Smooth Projected Landweber, SPL)中的人工设计参数. Zhang 等人在文献[31]中将迭代收缩阈值算法(Iterative Shrinkage-Thresholding Algorithm, ISTA)的迭代过程使用深度神经网络展开学习,提出了ISTA-Net网络模型. 基于这一优化思想,Zhang 等人^[32]进一步改进了ISTA-Net,提出了OPINE-Net(OPtimization-INSpired Explicable deep Network)网络模型,该模型结合了优化算法与深度学习的优点,进一步提升了图像的重建质量.

3 系统模型

考虑图1所示的点对点图像无线传输系统. 联合编码器将输入图像信号 $\mathbf{x} \in \mathbb{R}^n$ 通过联合编码器 $f_\theta: \mathbb{R}^n \rightarrow \mathbb{C}^k$ 映射到复值向量 $\mathbf{s} \in \mathbb{C}^k$,编码过程可以表示为

$$\mathbf{s} = f_\theta(\mathbf{x}) \quad (1)$$

其中, $k(k < n)$ 表示信道符号的长度, θ 表示联合编码器的参数,比率 $R = k/n$ 称之为信道带宽比.

编码出的信道符号 \mathbf{s} 通过无线信道 $\eta: \mathbb{C}^k \rightarrow \mathbb{C}^k$ 进行

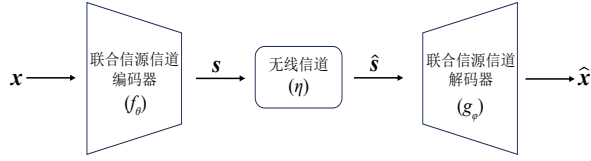


图1 点对点无线传输系统

传输,本文考虑加性高斯白噪声(Additive White Gaussian Noise, AWGN)信道中信号的传输. 此时联合解码器接收到的通道输出符号 $\hat{s} \in \mathbb{C}^k$ 可以表示为

$$\hat{s} = \eta(s; \nu) = s + \mathbf{n} \quad (2)$$

其中, ν 表示信道参数,在 AWGN 信道中, ν 表示信道信噪比,可以在解码端估计得到; $\mathbf{n} \in \mathbb{C}^k$ 为高斯白噪声,其由圆对称复高斯分布(circularly symmetric complex Gaussian distribution)的独立同分布样本组成,即 $\mathbf{n} \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I}_k)$,其中, σ_n^2 表示噪声的平均功率, \mathbf{I}_k 为 k -维单位矩阵.

在解码端,联合解码器通过映射 $g_\phi: \mathbb{C}^k \rightarrow \mathbb{R}^n$ 从含有噪声的信道符号 \hat{s} 重构出原始图像的估计 \hat{x} ,解码过程可以表示为式(3),其中 ϕ 表示联合解码器的参数:

$$\hat{x} = g_\phi(\hat{s}, \nu) = g_\phi(\eta(f_\theta(x); \nu), \nu) \quad (3)$$

在信道带宽比 R 一定条件下,编解码器联合设计以最小化原始图像 x 与重建图像 \hat{x} 之间的平均失真:

$$(\theta^*, \phi^*) = \underset{\theta, \phi}{\operatorname{argmin}} \mathbb{E}_{\nu \sim p_\nu} \mathbb{E}_{x \sim p_x} [d(x, \hat{x})] \quad (4)$$

其中, $d(x, \hat{x})$ 为 x 与 \hat{x} 的失真测量,本文采用均方误差(Mean Squared Error, MSE)作为失真度量. θ^*, ϕ^* 分别为最优编码器参数与最优解码器参数. 接下来将介绍如何使用深度压缩感知网络对联合编码器与联合解码器进行建模.

4 基于深度压缩感知的联合信源信道编码框架

4.1 BCS-JSCC 基本架构

BCS-JSCC 模型架构如图2所示,输入的图像信号 x 首先在编码器进行压缩感知采样得到压缩感知测量值 y ,将测量值 y 的前半部分与后半部分分别作为 I, Q 分量形成复值信道符号 s ,并通过 AWGN 信道传输至解码端. 解码端估计出信道的信噪比参数(Signal-to-Noise Ratio, SNR)对接收到的信号 \hat{s} 进行重塑,得到含有噪声的测量值信号 \hat{y} . 初始重建网络基于 SNR 与 \hat{y} 进行初始重建,得到初步重建的图像,并将其输入到后续的深度重建网络,经多层次子网络处理后得到重建图像 \hat{x} . 接下来将按照模型的结构,分别讨论压缩感知采样、初始重建网络以及深度重建网络三部分网络结构的设计.

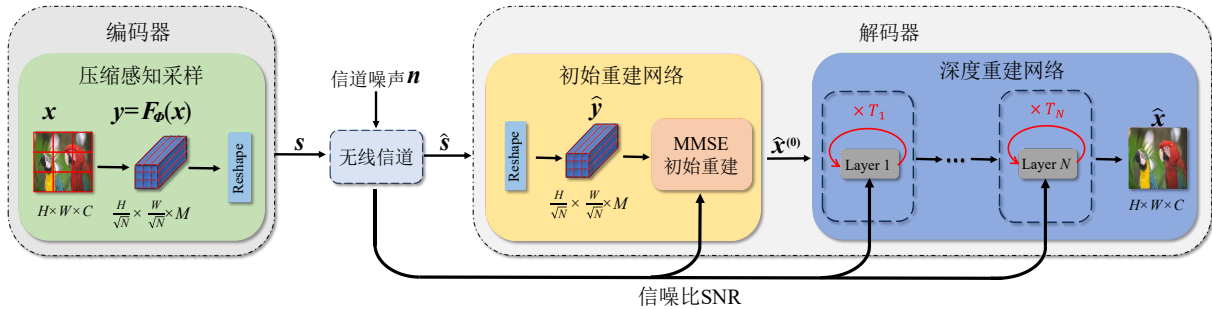


图2 BCS-JSCC 模型基本架构

4.2 压缩感知采样

BCS-JSCC 基于分块压缩感知(Block Compressive Sensing, BCS)对图像信号进行编码. 对于输入大小为 $H \times W \times C$ 的图像,BCS-JSCC 首先将图像等分为 N 个 $B \times B$ 大小的图像块,并对每个图像块使用相同的测量矩阵进行采样. 对于单个图像块,BCS 压缩感知采样过程可以表示为

$$\mathbf{y}_B = \Phi_B \mathbf{x}_B \quad (5)$$

其中, $\mathbf{x}_B \in \mathbb{R}^{B^2 C}$ 表示子图像块; $\Phi_B \in \mathbb{R}^{M \times B^2 C}$ 表示测量矩阵 ($M < B^2 C$); $\mathbf{y}_B \in \mathbb{R}^M$ 表示该图像块对应的压缩感知测量值. 令 Φ 表示整幅图像的测量矩阵,则采样过程可以表示为

$$\mathbf{y} = \Phi \mathbf{x} = \begin{bmatrix} \Phi_B & 0 & 0 & 0 \\ 0 & \Phi_B & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Phi_B \end{bmatrix} \mathbf{x} \quad (6)$$

上述操作可以使用卷积进行等效的表示,具体而言,可以将 Φ_B 的每一行重塑为 $C \times B \times B$ 的卷积核 \mathbf{W}_ϕ ,此时采样过程可以由步长为 B 的无偏置卷积进行描述. 本文使用 $F_\phi(\cdot)$ 表示使用无偏置卷积对图像进行分块压缩感知采样的操作,则采样过程可表示为

$$\mathbf{y} = F_\phi(\mathbf{x}) = \mathbf{W}_\phi * \mathbf{x} \quad (7)$$

为了进一步降低编码器的存储需求,BCS-JSCC 对测量矩阵 Φ_B 的元素进行了离散化处理. 具体而言,首

先构建各个元素均为连续值的测量矩阵 Φ_B^* , 并对 Φ_B^* 进行如下的二值化操作:

$$\Phi_B = \alpha \cdot \text{sign}(\Phi_B^*) \quad (8)$$

其中, α 是一个可学习的比例因子参数; $\text{sign}(\cdot)$ 为符号函数, 其将大于等于 0 的元素映射为 1, 将小于 0 的元素映射为 -1. 由于 $\text{sign}(\cdot)$ 函数不可微, 为了使模型可训练, 在训练过程令其梯度等于 1. 在模型训练过程中, 实际优化的参数为 Φ_B^* 与 α , 实际对图像信号进行压缩感知采样的为二值化矩阵 Φ_B . 编码器只需存储浮点数 α 与二值化矩阵 Φ_B 的符号信息, 可以很大程度上减少编码器所需的存储资源, 从而优化其硬件部署.

4.3 初始重建网络

在解码端, 信道的信噪比 SNR 可以通过导频符号、谱分析等方法进行估计, 并将估计出的 SNR 作为信道状态信息输入到后续的初始重建网络和深度重建网络中. 初始重建网络首先对接收到的信道符号 \hat{s} 分别提取实部与虚部, 并以此将 \hat{s} 重塑为测量值 \mathbf{y} 的形状, 得到含有噪声的测量值 $\hat{\mathbf{y}}$. $\hat{\mathbf{y}}$ 包含有每个图像块 \mathbf{x}_B 的测量值 $\hat{\mathbf{y}}_B$, 其满足

$$\hat{\mathbf{y}}_B = \mathbf{y}_B + \mathbf{n}_B = \Phi_B \mathbf{x}_B + \mathbf{n}_B \quad (9)$$

其中, \mathbf{n}_B 为加入到测量值 \mathbf{y}_B 中的高斯白噪声. BCS-JSCC 使用最小均方误差估计方法, 根据 $\hat{\mathbf{y}}_B$ 对各个子图像块进行初始重建. 具体而言, 我们希望找到一个初始

估计矩阵 \mathbf{W}_B , 以最小化原始信号 \mathbf{x}_B 与估计信号 $\hat{\mathbf{x}}_B = \mathbf{W}_B \hat{\mathbf{y}}_B$ 的均方误差

$$\mathbf{W}_B = \underset{\mathbf{W}}{\text{argmin}} \mathbb{E}_{(\mathbf{x}_B, \hat{\mathbf{y}}_B) \sim p(\mathbf{x}_B, \hat{\mathbf{y}}_B)} \left[\left\| \mathbf{x}_B - \mathbf{W} \hat{\mathbf{y}}_B \right\|_2^2 \right] \quad (10)$$

其中, $p(\mathbf{x}_B, \hat{\mathbf{y}}_B)$ 为原始图像块 \mathbf{x}_B 与含有噪声的测量值 $\hat{\mathbf{y}}_B$ 的联合分布. 通过求解式 (10), 可以得到 \mathbf{W}_B 的解析表达式

$$\mathbf{W}_B = \mathbf{R}_{\mathbf{x}_B} \Phi_B^T (\Phi_B \mathbf{R}_{\mathbf{x}_B} \Phi_B^T + \sigma_{\mathbf{n}_B}^2 \mathbf{I}_M)^{-1} \quad (11)$$

其中, \mathbf{I}_M 为单位矩阵, $\sigma_{\mathbf{n}_B}^2$ 为 \mathbf{n}_B 的功率, 其可以通过信噪比 SNR 以及接收测量值 $\hat{\mathbf{y}}_B$ 的功率得到; $\mathbf{R}_{\mathbf{x}_B} = \mathbb{E}_{\mathbf{x}_B \sim p_{\mathbf{x}_B}} [\mathbf{x}_B \mathbf{x}_B^T]$ 为图像块 \mathbf{x}_B 的协方差矩阵. 接收端无法获知 \mathbf{x}_B 的统计信息, BCS-JSCC 通过将 $\mathbf{R}_{\mathbf{x}_B}$ 设置为可学习参数, 自适应地学习协方差矩阵 $\mathbf{R}_{\mathbf{x}_B}$. 初始重建网络对每个图像块使用 \mathbf{W}_B 进行初始化, 得到初始重建图像 $\hat{\mathbf{x}}^{(0)}$.

4.4 深度重建网络

BCS-JSCC 的深度重建网络基于 OPINE-Net^[32] 网络模型的基本架构, 并在其基础上进行了优化与改进, 其网络架构如图 3 所示. 深度重建网络基于 ISTA 算法的思想对图像进行压缩感知重建, 其优化函数为

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\text{argmin}} \frac{1}{2} \left\| \Phi \mathbf{x} - \mathbf{y} \right\|_2^2 + \lambda \left\| \Psi \mathbf{x} \right\|_1 \quad (12)$$

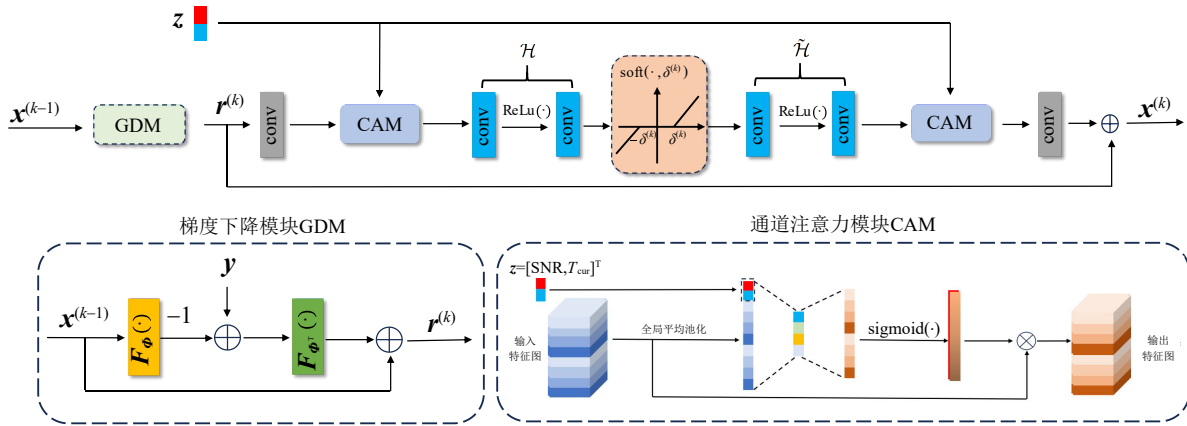


图3 深度重建网络第k层网络架构

其中, $\Psi \mathbf{x}$ 表示 \mathbf{x} 相对于变换 Ψ 下的变换系数; λ 为正优化参数. Zhang 等人在文献[32]中将式 (12) 展开为如下的迭代过程:

$$\mathbf{r}^{(k)} = \mathbf{x}^{(k-1)} - \rho \Phi^T (\Phi \mathbf{x}^{(k-1)} - \mathbf{y}) \quad (13)$$

$$\mathbf{x}^{(k)} = \mathbf{r}^{(k)} + \mathcal{G} \left(\tilde{\mathcal{H}} \left(\text{soft} \left(\mathcal{H} \left(\mathcal{D} \left(\mathbf{r}^{(k)} \right) \right), \theta \right) \right) \right) \quad (14)$$

式中, $\mathbf{x}^{(k)}$ 表示第 k 次迭代重建的图像; $\Phi^T (\Phi \mathbf{x}^{(k-1)} - \mathbf{y})$

是信号保真项 $\frac{1}{2} \left\| \Phi \mathbf{x} - \mathbf{y} \right\|_2^2$ 在 $\mathbf{x}^{(k-1)}$ 处的梯度; ρ 表示步长; $\mathcal{D}(\cdot)$ 、 $\mathcal{H}(\cdot)$ 、 $\mathcal{G}(\cdot)$ 分别表示非线性变换, 均使用卷积与激活函数进行实现, $\tilde{\mathcal{H}}$ 为 \mathcal{H} 的逆变换, 满足 $\tilde{\mathcal{H}} \circ \mathcal{H} = \mathbf{I}$, \circ 表示复合运算, \mathbf{I} 表示单位算子; $\text{soft}(\cdot, \theta)$ 表示软阈值函数, 其解析形式为

$$\text{soft}(\mathbf{x}, \theta) = \text{sign}(\mathbf{x}) (|\mathbf{x}| - \theta) \quad (15)$$

深度重建网络包含为多个子网络, 每一层子网络

对应于式(13)和式(14)表示的迭代过程. 随着子网络数量的增加,深度重建网络会表现出更好的性能. 然而,这种性能的提高是归因于网络迭代次数的增加,而不是归因于重构子网络数量的增加. 与 OPINE-Net 通过增加重建子网络数量以提高迭代次数不同,BCS-JSCC 为每一层子网络设计了循环结构^[33](如图 2 子网络块中红色标注所示). 对于第 k 层子网络,其输出将被反馈回其输入端,如此的循环过程将重复 T_k 次. 此类循环架构能够充分利用各层重构子网络的容量,在保持图像重建质量的同时,显著减少重构子网络的总数量,从而有效降低解码器的存储需求.

为了使模型可以适应变化的信噪比环境,BCS-JSCC 为每一层子网络设计了通道注意力模块(Channel Attention Module, CAM),其结构可见图 3 中的 CAM 块. CAM 接受特征图作为输入,并通过全局平均池化进行处理,池化后得到的向量与输入向量 \mathbf{z} 进行拼接,从而得到上下文信息. 其中 \mathbf{z} 是一个包含 2 个元素的向量,表示为 $\mathbf{z} = [\text{SNR}, T_{\text{cur}}]^T$, SNR 即为解码器估计得到的通道信噪比, T_{cur} 为到目前为止已经完成的子网络迭代次数. 上下文信息经过一个多层感知机(MultiLayer Perceptron, MLP)后,利用 sigmoid 函数对各元素进行归一化,并与输入的特征图进行逐通道相乘,从而得到校准后的特征图. 通过将信噪比 SNR 信息嵌入到 CAM,模型能够适应变化的信噪比条件,实现对多种信噪比环境的统一建模,无需为每种特定信噪比单独训练模型.

4.5 损失函数

BCS-JSCC 基于文献[32]中的损失函数对模型进行优化,其表达式如下:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{discrepancy}} + \gamma \mathcal{L}_{\text{symmetry}} + \mu \mathcal{L}_{\text{orth}} \quad (16)$$

其中:

$$\mathcal{L}_{\text{discrepancy}} = \frac{1}{N_x L} \sum_{i=1}^L \|\hat{\mathbf{x}}_i - \mathbf{x}_i\|_2^2 \quad (17)$$

$$\mathcal{L}_{\text{symmetry}} = \frac{1}{N_h} \sum_{i=1}^L \sum_{k=1}^{T_{\text{tot}}} \|\tilde{\mathcal{H}}(\mathcal{H}(\mathbf{h}_i^{(k)})) - \mathbf{h}_i^{(k)}\|_2^2 \quad (18)$$

$$\mathcal{L}_{\text{orth}} = \frac{1}{M^2} \|\Phi_B \Phi_B^T - \mathbf{I}_M\|_2^2 \quad (19)$$

其中, $\mathcal{L}_{\text{discrepancy}}$ 为重建损失函数, N_x 为 \mathbf{x} 中的元素个数, L 表示训练集中的图像数目; $\mathcal{L}_{\text{symmetry}}$ 为对称损失函数, $\mathbf{h}_i^{(k)}$ 表示子网络在经历 k 次迭代后,输入到 $\mathcal{H}(\cdot)$ 变换中的特征图, N_h 为其元素总个数, T_{tot} 表示所有子网络迭代的总次数; $\mathcal{L}_{\text{orth}}$ 为正交损失函数; γ, μ 为网络正则化参数,在训练过程中均设置为 0.01. BCS-JSCC 网络通过端到端训练,以最小化损失函数 $\mathcal{L}_{\text{total}}$.

5 实验与分析

5.1 训练与实现细节

本文依次在低分辨率数据集 CIFAR-10^[34](Canadian Institutes For Advanced Research)和高分辨率数据集 Kodak^[35]上评估所提出的方法. CIFAR-10 数据集由 50 000 个训练图像和 10 000 个测试图像组成,每幅图像的分辨率大小是 32×32 像素,且具有 3 个通道(RGB 格式). 实验采用 Adam 优化器^[36]以及 64 的批大小训练模型,整个训练经历 150 个 epochs:前 100 个 epochs 学习率为 1×10^{-3} ,后 50 个 epochs 学习率下降到 1×10^{-4} . 整个模型在 AWGN 信道环境下训练,在 0 dB 至 20 dB 之间均匀采样信噪比 SNR. 为进一步验证模型对于高分辨率图像的传输性能,本文在高分辨率数据集 DIV2K^[37](Diverse 2K resolution high quality images)上对模型进行训练,并在 Kodak 数据集上进行了测试,其包含了 24 张分辨率大小为 512×768 的 RGB 图像. DIV2K 数据集每幅图像被随机裁剪成 100 个 128×128 大小的补丁,得到 90 000 张大小相同的图像补丁. 批量大小设置为 32,共计训练 150 个 epochs,前 100 个 epochs 学习率为 1×10^{-4} ,后 50 个 epochs 学习率下降到 1×10^{-5} . BCS-JSCC 的性能根据重建图像的峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)进行量化评估. 其定义如下:

$$\text{PSNR} = 10 \lg \frac{\text{MAX}^2}{d(\mathbf{x}, \hat{\mathbf{x}})} \quad (20)$$

其中, MAX 为图像像素的最大可能值,对于 8 位 RGB 彩色图像, MAX 等于 255; $d(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n} \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2$ 为原始图像与 \mathbf{x} 重建图像 $\hat{\mathbf{x}}$ 之间的均方差.

实验保持深度重建网络中每个子网络的循环次数一致,并通过调整网络层数与每层子网络的循环次数,训练得到了三个模型:BCS-JSCC^S、BCS-JSCC^M、BCS-JSCC^L. 这三个模型在深度重建的子网络层数与循环次数如表 1 所示. 需要注意的是,三者仅在深度重建网络的结构上存在差异,其余设置完全一致.

表 1 模型的网络层数与循环次数

模型名称	子网络层数 Layer	每层子网络的循环次数 T
BCS-JSCC ^S	1	9
BCS-JSCC ^M	3	3
BCS-JSCC ^L	9	1

整个系统使用 Pytorch 框架^[38]进行构建,实验环境为:Windows11 64 位操作系统, CPU (Central Processing Unit) 为 Intel(R) Core(TM) i7-14700HX 2.10 GHz, GPU (Graphics Processing Unit) 为 Nvidia GeForce RTX 4070.

5.2 对比模型设置

实验选取 DeepJSCC^[9] (DJSCC)、ADJSCC^[12] 两种基于深度学习的 JSCC 模型进行对比, 使用每像素浮点运算次数对编码器需要的计算资源进行评估. 对于输入分辨率大小为 $H \times W$ 的 RGB 格式图像, 其每像素浮点运算次数定义为

$$\text{FLOPs/pixel} = \frac{\text{FLOPs}}{H \times W} \quad (21)$$

其中, FLOPs 为完成一次编码所需要的总浮点计算次数. DJSCC 和 ADJSCC 均是基于 CNN 的 JSCC 模型, 其编解码器包含多层的卷积操作与激活函数. 在对比三种模型 (包括 BCS-JSCC) 在编码器相同 FLOPs/pixel 条件下的性能时, 通过调整卷积层的通道数、网络层数以及卷积核大小这些超参数, 对 DJSCC 和 ADJSCC 的编码器进行剪枝, 而解码器结构保持不变. DJSCC、ADJSCC 的

模型容量 (model capacity), 即模型的表达能力, 与输入图像的分辨率、通道数量、模型深度密切相关^[39]. 在上述因素中, 模型的通道数量是影响其模型容量的最重要参数^[19]. 更多的通道可以使模型学习更丰富的特征表示, 从而提高复杂任务的性能. 因此, 在对上述模型进行剪枝时, 优先保证每层的通道数量, 其次再调整模型的层数和卷积核大小等其他参数.

在实验中, 所有模型的带宽压缩比 R 均设置为 1/6. 针对 BCS-JSCC, 将压缩感知的块大小分别设置为 4、8、16, 设计了三组实验. 在每种块大小设置下, 对 DJSCC 和 ADJSCC 的编码器进行剪枝, 以确保三个模型的编码器的每像素浮点运算次数近似一致. 具体的编码器网络结构如表 2 所示. 在每组实验中, DJSCC 和 ADJSCC 均重新进行训练, 直至各自损失函数达到收敛, 以适配其调整后的编码器结构.

表 2 各模型的编码器结构

组别	模型	编码器结构	编码器 FLOPs/pixel/K
1	BCS-JSCC	Conv(4 × 4 × 16 4)	0.10
	DJSCC	Conv(3 × 3 × 3 2)+PReLU; Conv(3 × 3 × 16 2)+PReLU	0.10
	ADJSCC	Conv(3 × 3 × 3 2)+PReLU+AF; Conv(3 × 3 × 16 2)+PReLU	0.10
2	BCS-JSCC	Conv(8 × 8 × 64 8)	0.38
	DJSCC	Conv(3 × 3 × 12 2)+PReLU; Conv(3 × 3 × 16 2)+PReLU	0.39
	ADJSCC	Conv(3 × 3 × 12 2)+PReLU+AF; Conv(3 × 3 × 16 2)+PReLU	0.39
3	BCS-JSCC	Conv(16 × 16 × 256 16)	1.54
	DJSCC	Conv(3 × 3 × 17 2)+PReLU; Conv(3 × 3 × 35 2)+PReLU; Conv(3 × 3 × 16 1)+PReLU	1.54
	ADJSCC	Conv(3 × 3 × 25 2)+PReLU+AF; Conv(3 × 3 × 25 2)+PReLU+AF; Conv(3 × 3 × 16 1)+PReLU	1.52

注: Conv($F \times F \times C|S$) 表示卷积核大小为 F , 步长为 S , 输出通道数为 C 的卷积层, 不同层之间以分号隔开; PReLU 为非线性激活函数, AF 为 ADJSCC 中的注意力组件.

5.3 CIFAR-10 数据集实验结果

5.3.1 AWGN 信道实验结果

图 4 展示了在 AWGN 信道, 带宽压缩比为 1/6 条件下, BCS-JSCC、ADJSCC、DJSCC 三种模型在 CIFAR-10 测试集下的 PSNR-SNR 性能曲线, 其中 DJSCC 模型分别在信噪比为 4 dB、10 dB、16 dB 条件下分别进行训练与评估. 三种模型均在相同的训练集 CIFAR-10 上进行训练, 直到各自的损失函数收敛.

在三组实验条件下, BCS-JSCC 的三种变体 (BCS-JSCC^S、BCS-JSCC^M、BCS-JSCC^L) 在不同信噪比条件下表

现出相近的 PSNR 性能, 且 PSNR 性能依次提升, 但最大提升不超过 1 dB. 尽管这三种变体的性能相近, 其解码器结构却存在差异: 解码器中深度重建网络的层数依次增加, 导致模型大小逐渐升高. 第 5.5 节对三种变体在解码端所需要的存储空间大小进行了详细比较, 并进行分析.

在第一组实验 (编码器 FLOPs/pixel 等于 0.10 K) 条件下, 所提出的方法 BCS-JSCC 在不同信噪比条件下表现均优于 DJSCC、ADJSCC, 并且随着信噪比的升高, 性能差异逐渐扩大, 在 SNR=20 dB 处 PSNR 性能超出

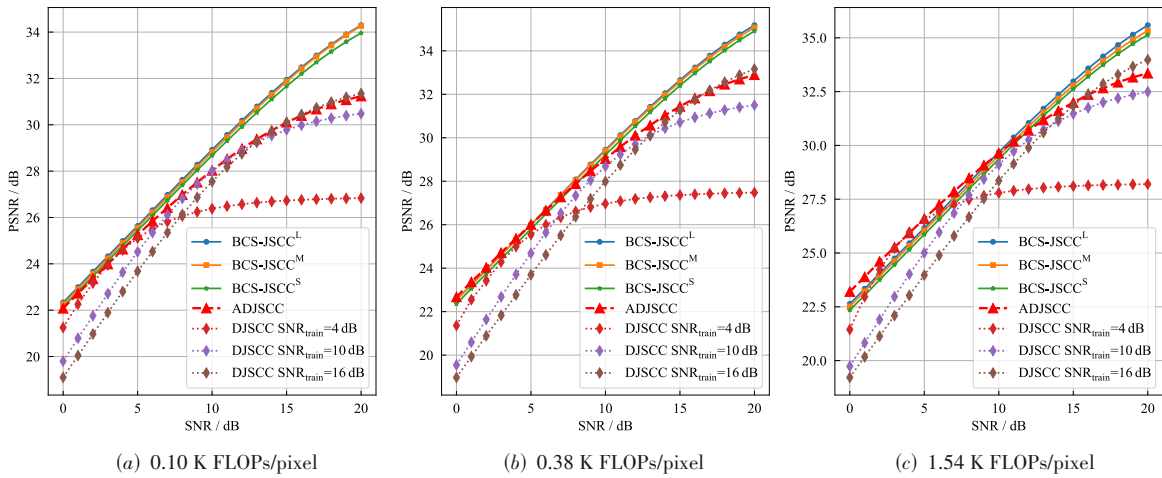


图4 CIFAR-10数据集:不同方法在AWGN信道,带宽压缩比1/6条件下PSNR-SNR性能曲线

DJSCC和ADJSCC约3 dB. 由于通道注意力模块CAM的引入,BCS-JSCC对不同的信噪比条件具有适应性,其PSNR性能随信噪比SNR增加而平稳升高,而这一点与DJSCC在一定信噪比后性能趋于饱和的现象形成了鲜明对比.

在第二和第三组实验条件下(编码器FLOPs/pixel分别为0.38 K和1.54 K),BCS-JSCC在低信噪比(0~5 dB)下的性能接近或低于ADJSCC和DJSCC,但性能差距最大不超过1 dB. 随着信噪比的增加,BCS-JSCC的性能呈现明显的提升,并超过ADJSCC和DJSCC. 在信噪比为20 dB时,BCS-JSCC的性能超出ADJSCC和DJSCC约2 dB.

图5展示了在信道信噪比分别为4 dB、10 dB和16 dB时,BCS-JSCC、ADJSCC和DJSCC三种模型的PSNR性能随编码器FLOPs/pixel变化的曲线,其中主要关注PSNR随FLOPs/pixel降低而下降的速率. 由图5可以看出,随着编码器的计算量逐渐降低,三种模型的PSNR性能均有所下降,但下降的速率存在差异. 在所有信噪比条件下,BCS-JSCC的PSNR性能下降速率较ADJSCC与DJSCC更为缓慢,特别是在FLOPs/pixel从0.38 K下降到0.10 K的过程中,BCS-JSCC的下降速率明显慢于两种对比模型,表明BCS-JSCC在编码器计算量降低时能够更好地维持较高的图像传输性能.

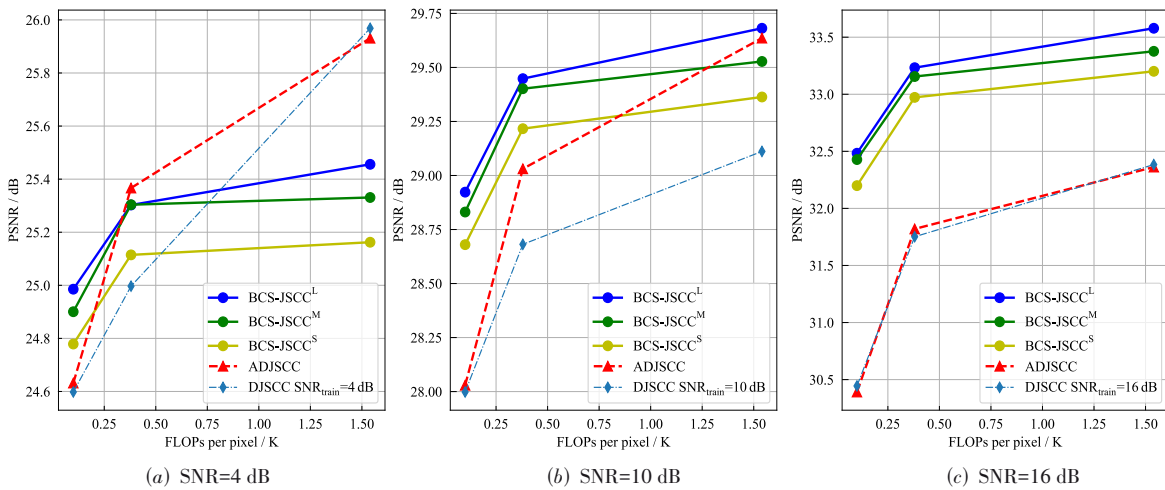


图5 CIFAR-10数据集:不同方法在AWGN信道,带宽压缩比1/6条件下的PSNR-FLOPs/pixel性能曲线

由上述实验结果可见,当编码器计算量一致时,BCS-JSCC在高信噪比条件下具有更为显著的性能优势. 此外,低算力(0.10 K FLOPs/pixel)条件下,BCS-JSCC在各信噪比环境下均具有更优的性能表现. 随着编码器计算复杂度的降低,BCS-JSCC的性能下降趋势

较弱,体现出BCS-JSCC模型在编码器低算力条件下具有更好的性能维持能力.

我们同样尝试与传统的分离式编码方案进行对比. 具体而言,信源编码采用BPG编码^[40]算法. 针对信道编码,构建了基于低密度奇偶校验码的编码方案,编

码率涵盖 1/3、1/2、3/5、2/3、3/4 及 5/6, 同时为码字配置了多阶正交幅度调制 (Quadrature Amplitude Modulation, QAM), 包含 4-QAM、16-QAM、64-QAM. 通过将特定的 LDPC 编码码率与 QAM 调制阶数相结合, 形成了多组预定义的调制编码方案. 图 6 展示了 AWGN 信道, 带宽压缩比为 1/6 的条件下, 本文所提出的 BCS-JSCC 方法与传统的分离式编码方法 BPG+LDPC 的性能差异曲线图. 实验数据表明, BPG+LDPC 在其信噪比阈值处的性能与 BCS-JSCC 接近, 但其表现出典型的“悬崖效应”(Cliff Effect) 特征: 当 SNR 降低至特定阈值以下时, 重建图像质量急剧恶化. 这种现象的成因可归结于, 当实际 SNR 低于预设目标信噪比时, BPG 编码出的码字出现比特错误, 导致整幅图像无法正确解码, 而该目标信噪比的取值则与 LDPC 编码的码率选择及调制阶数密切相关. 与之形成鲜明对比的是, 本文提出的 BCS-JSCC 方案展现出渐进式的性能衰减特性: 在低信噪比区域呈现平缓的质量劣化曲线, 同时在高信噪比区间避免了性能饱和现象的出现. 这种特性表明联合编码方案相比传统分离式编码架构能更有效地适应信道条件的动态变化.

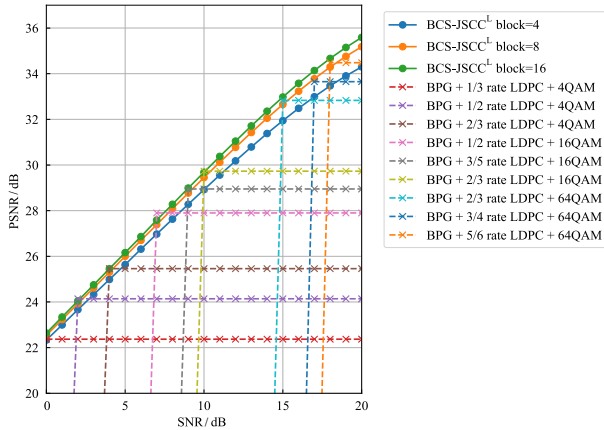


图 6 CIFAR-10 数据集: AWGN 信道, 带宽压缩比为 1/6 条件下, BCS-JSCC 与数字传输方案 BPG+LDPC 的 PSNR 性能对比曲线

5.3.2 TDL 信道实验结果

为了进一步验证 BCS-JSCC 在复杂信道环境的传输性能, 本文进一步在抽头延迟线 (Tapped Delay Line, TDL) 多径信道模型下进行实验验证. TDL 信道中, 发射信号通过不同的反射路径到达接收端, 各条路径具有不同的衰减与延迟. 需特别指出, 实验设定中假设接收端已知 TDL 信道的冲激响应, 这为模型在后续的深度重建中提供了关键的先验信息. 在 BCS-JSCC 模型中, 图像块 \mathbf{x}_B 经压缩感知采样、调制为复数信道符号, TDL 信道传输过程中, 接收端所接收信号 $\hat{\mathbf{s}}_B$ 满足:

$$\hat{\mathbf{s}}_B = \mathbf{HT}\Phi_B \mathbf{x}_B + \mathbf{n}_B \quad (22)$$

其中, \mathbf{H} 为 TDL 信道的等效信道矩阵, 其构建基于接收

端已知的信道冲激响应参数; $\mathbf{T} = [\mathbf{I}_{M/2} \quad \mathbf{jI}_{M/2}]$ 为调制矩阵, 其作用是将图像块 \mathbf{x}_B 对应的压缩感知测量值 $\Phi_B \mathbf{x}_B$ 的前后两半分别映射为复信号的 I, Q 分量. 通过将接收信号 $\hat{\mathbf{s}}_B$ 分解为实部与虚部, 可推导出等效的实数域传输模型:

$$\begin{aligned} \hat{\mathbf{y}}_B &= \begin{bmatrix} \text{Re}\{\hat{\mathbf{s}}_B\} \\ \text{Im}\{\hat{\mathbf{s}}_B\} \end{bmatrix} \\ &= \begin{bmatrix} \text{Re}\{\mathbf{HT}\} \\ \text{Im}\{\mathbf{HT}\} \end{bmatrix} \Phi_B \mathbf{x}_B + \begin{bmatrix} \text{Re}\{\mathbf{n}_B\} \\ \text{Im}\{\mathbf{n}_B\} \end{bmatrix} \\ &= \tilde{\Phi}_B \mathbf{x}_B + \tilde{\mathbf{n}}_B \end{aligned} \quad (23)$$

其中, $\tilde{\Phi}_B = [\text{Re}\{\mathbf{HT}\} \quad \text{Im}\{\mathbf{HT}\}]^T \Phi_B$ 为等效的压缩感知测量矩阵. 基于上述推导, TDL 信道下的传输模型在数学形式上与式 (9) 所描述的 AWGN 信道模型具有等价性, 因此可以借助 AWGN 信道的处理方法对 TDL 信道进行有效分析与处理.

实验采用 TDL-A 信道模型, 载波频率设置为 3.5 GHz, 采样率设置为 10 MHz, 时延扩展选取为 30 ns. 为保障对比公平性, 对比 DJSCC、ADJSCC 同样假设接收端已知 TDL 信道冲激响应, 并采用 MMSE 均衡算法进行信道补偿. 图 7、图 8 分别给出了 TDL 信道, 带宽压缩比为 1/6 的条件下, 不同方法的 PSNR-SNR 以及 PSNR-FLOPs/pixel 曲线. 实验结果表明, 基于等效的信道矩阵变换, BCS-JSCC 在 TDL 信道环境下仍保持着良好的传输性能, 特别是在编码器低计算复杂度条件下 (对应于第一组实验, FLOPs/pixel 等于 0.10 K), 在各信噪比环境下取得了优于对比模型的普遍性能. 此外, 随着编码器计算复杂度的降低, BCS-JSCC 在 TDL 信道中依然表现出较强的性能维持能力, 验证了其在计算资源受限场景下的适用性.

本节接下来在 TDL 信道环境下对 BCS-JSCC 与传统分离式编码 BPG+LDPC 方法进行了性能对比. 实验设定中, BPG+LDPC 方案同样假设接收端已知 TDL 信道冲激响应, 并采用 MMSE 均衡算法以抑制多径干扰. 图 9 展示了在带宽压缩比为 1/6 的条件下, BCS-JSCC 与 BPG+LDPC 数字传输方案的 PSNR 性能对比曲线. 为提升结果的可读性, 实验从 LDPC 编码码率与调制方案的所有可能组合中选取了典型配置进行对比分析. 实验结果表明, 在 TDL 信道环境下, BPG+LDPC 数字传输方案的性能显著低于所提出的 BCS-JSCC 方法. 虽然传统的 BPG+LDPC 编码方法在 AWGN 信道下具有良好的传输性能, 然而在 TDL 信道环境下, 信道不稳定性以及缺乏针对衰落特性的联合优化机制, 导致其性能显著退化. 相比之下, BCS-JSCC 通过联合优化信源与信道编码, 展现出更强的信道适应性, 能够有效应对 TDL 信道中的复杂传播条件.

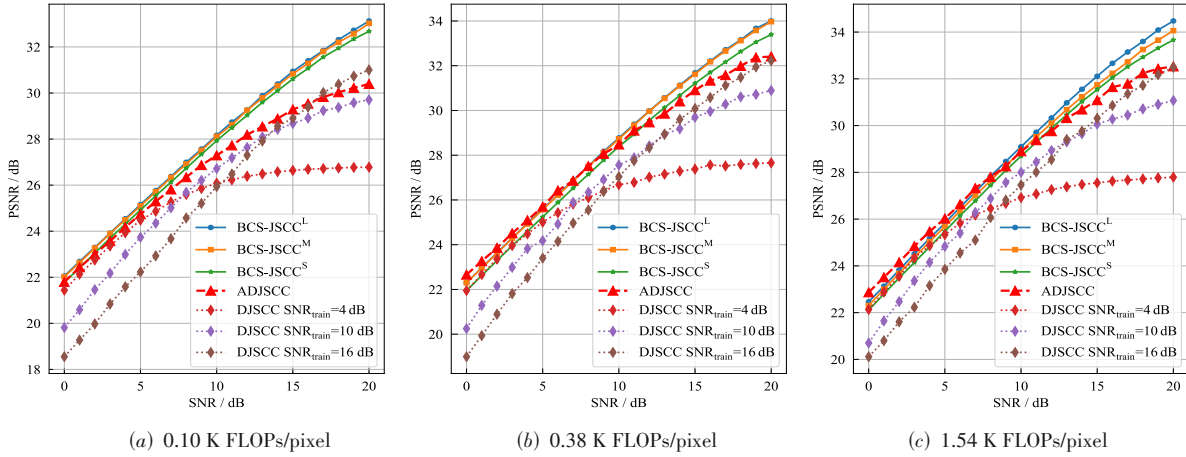


图7 CIFAR-10数据集:不同方法在TDL信道,带宽压缩比1/6条件下PSNR-SNR性能曲线

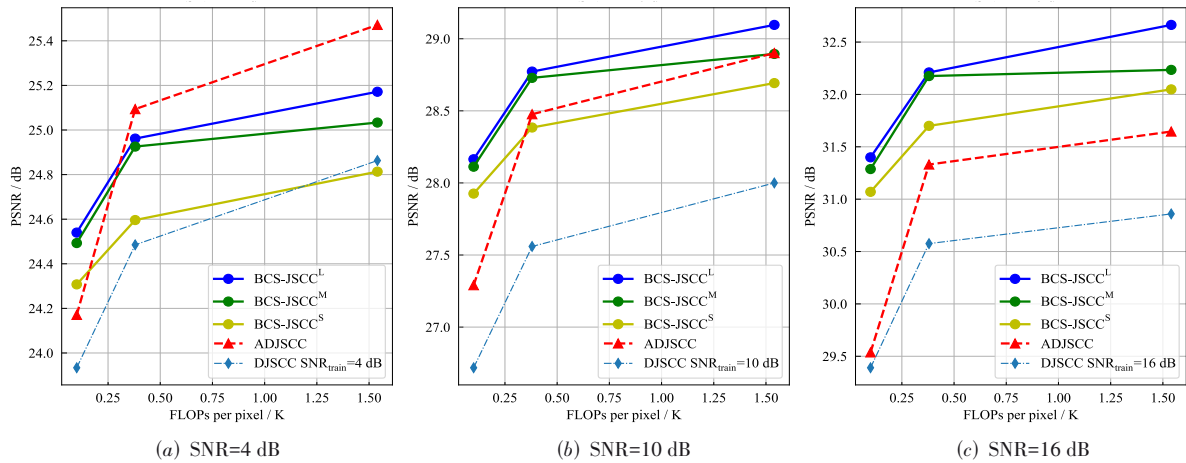


图8 CIFAR-10数据集:不同方法在TDL信道,带宽压缩比1/6条件下的PSNR-FLOPs/pixel性能曲线

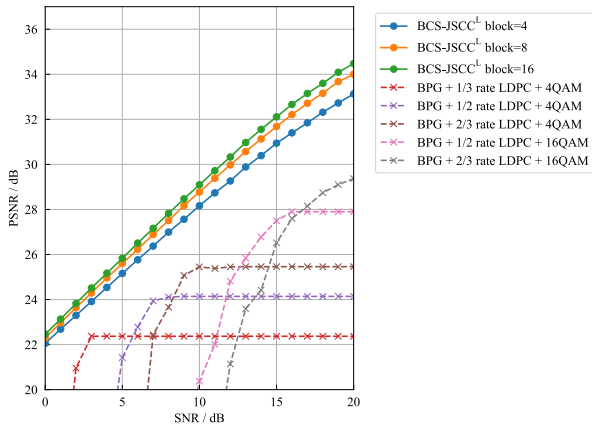


图9 CIFAR-10数据集:TDL信道,带宽压缩比为1/6条件下,BCS-JSCC与数字传输方案BPG+LDPC的PSNR性能对比曲线

5.4 Kodak数据集实验结果

为验证所提出的BCS-JSCC在高分辨率图像无线传输场景下的性能表现,本文接下来在Kodak数据集上进行实验评估.图10给出了BCS-JSCC、DJSCC以及ADJSCC三种模型在Kodak数据集上的PSNR-SNR性能

曲线.三种模型均在相同的高分辨率数据集DIV2K数据集上进行训练,直到各自的损失函数收敛.由图10可见,在三组实验条件下,BCS-JSCC在高信噪比条件下相比于对比模型均具有更为明显的优势.与图4中CIFAR-10数据集的结果不同,BCS-JSCC模型相比于其他两种模型在高信噪比条件下的差距更为显著.尤其是在第一组实验条件下(编码器FLOPs/pixel等于0.10 K),BCS-JSCC在SNR=20 dB处的PSNR性能可超ADJSCC约5 dB.

图11展示了在信道信噪比分别为4 dB、10 dB和16 dB时,BCS-JSCC、ADJSCC和DJSCC的PSNR性能随编码器FLOPs/pixel变化的曲线.与图5给出的结果相比,BCS-JSCC在编码器计算量下降时,PSNR性能的降低趋势明显较其他两种模型更为缓慢,特别是在FLOPs/pixel从0.38 K下降至0.10 K的过程中.

5.5 模型存储空间占用

表3列出了BCS-JSCC的三种变体BCS-JSCC^L、BCS-JSCC^M、BCS-JSCC^S以及DJSCC、ADJSCC的各自的编解码器存储空间占用情况,表中使用粗体突出显示存储

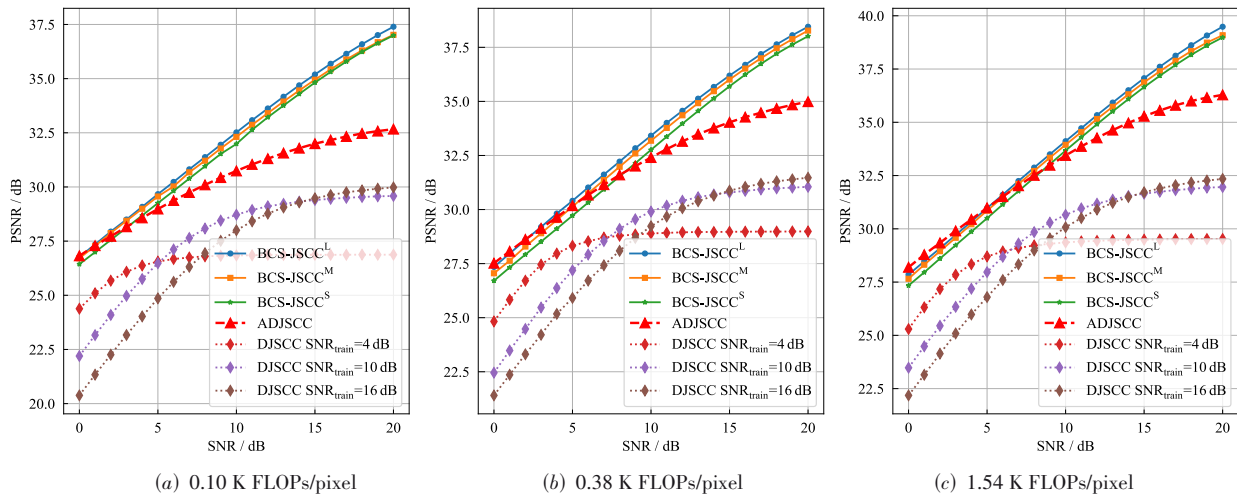


图 10 Kodak 数据集:不同方法在 AWGN 信道,带宽压缩比 1/6 条件下 PSNR-SNR 性能曲线

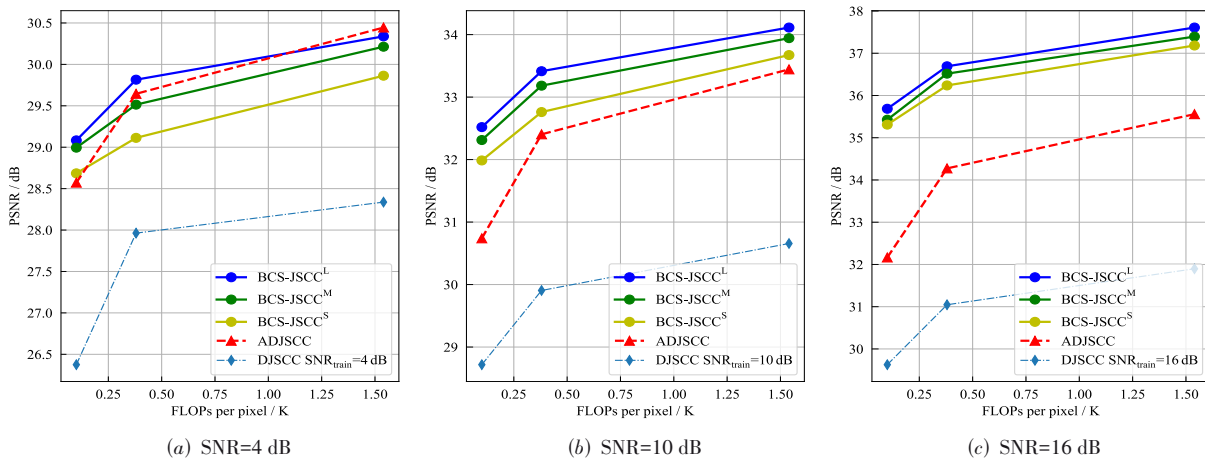


图 11 Kodak 数据集:不同方法在 AWGN 信道,带宽压缩比 1/6 条件下的 PSNR-FLOPs/pixel 性能曲线

空间占用最小的模型对应的数据. 共统计了三组实验, 每一组实验的模型结构均与表 2 所示内容相对应. 如表 3 所示, 所提出的模型 BCS-JSCC 在编码器方面的存储需求显著低于 DJSCC 和 ADJSCC. 这是因为在 BCS-JSCC 中, 编码器的测量矩阵基于式 (8) 进行了二值化处理, 并将二值化后的测量矩阵以比特序列的形式保存. 因此, BCS-JSCC 的编码器不需要存储大量的浮点数, 很大程度上减少编码器所需的存储资源, 从而降低了模型的部署难度.

在解码器存储需求层面, 所提出模型 BCS-JSCC 的三种变体 BCS-JSCC^L、BCS-JSCC^M、BCS-JSCC^S 随着深度重建子网络层数的减少, 存储需求也逐渐减小, 但三者总子网络迭代次数保持相同. 由表 3 与图 4 可见, BCS-JSCC^S 模型在存储空间占用方面优于其他两种变体, 同时在 PSNR 性能方面相比于其他两种变体未见明显下降.

在第一组实验条件下 (编码器 FLOPs/pixel 等于

表 3 各模型的编解码器的存储空间占用

组别	模型	编码器大小/KB	解码器大小/KB
1	BCS-JSCC ^L	0.1	2 144.8
	BCS-JSCC ^M		721.3
	BCS-JSCC ^S		247.1
	ADJSCC	5.8	19 989.7
	DJSCC	4.5	309.7
2	BCS-JSCC ^L	1.5	2 279.8
	BCS-JSCC ^M		856.3
	BCS-JSCC ^S		381.7
	ADJSCC	12.2	19 989.7
	DJSCC	10.5	309.7
3	BCS-JSCC ^L	24.0	4 440.0
	BCS-JSCC ^M		3 016.4
	BCS-JSCC ^S		2 542.1
	ADJSCC	49.4	19 989.7
	DJSCC	45.8	309.7

0.10 K), BCS-JSCC^s 解码器大小低于 ADJSCC 和 DJSCC. 而随着编码器计算复杂度的升高, BCS-JSCC 的存储需求逐渐增大, 在第二、三组实验中均高于解码器结构更为简单的 DJSCC 模型. 这一现象可归因于 BCS-JSCC 模型压缩感知块的增大, 导致初始重建网络中协方差矩阵 \mathbf{R}_{x_b} 的增大, 从而提高了解码器存储需求.

5.6 消融实验

5.6.1 初始重建子网络的影响

为验证 BCS-JSCC 模型中初始重建子网络方案的有效性, 本实验将初始重建子网络中由 MMSE 方法推导出来的重建矩阵替换为可学习的初始重建矩阵 $\mathbf{W}_{\text{learned}}$, 其余设置与原模型一致. 实验中, 压缩感知块大小设置为 4, 深度重建子网络的层数为 3, 每层子网络的循环次数 T 等于 3. 使用 CIFAR-10 数据集作为训练集和测试集, 在 AWGN 信道条件下对两种方法的性能进行了比较, 两种方法的 PSNR-SNR 性能曲线如图 12 所示. 实验结果表明, 基于 MMSE 的初始重建方法在不同信噪比条件下的 PSNR 性能普遍优于可学习初始重建方法, 且在 SNR 为 20 dB 时, PSNR 值超出可学习方法约 0.6 dB. 分析表明, 虽然可学习初始重建矩阵能够自适应调整, 但未能考虑信道噪声的影响; 而基于 MMSE 的初始重建方法通过考虑噪声因素, 在不同信噪比条件下可以最小化初始重建图像与原图像之间的均方误差, 从而实现了更优的重建性能.

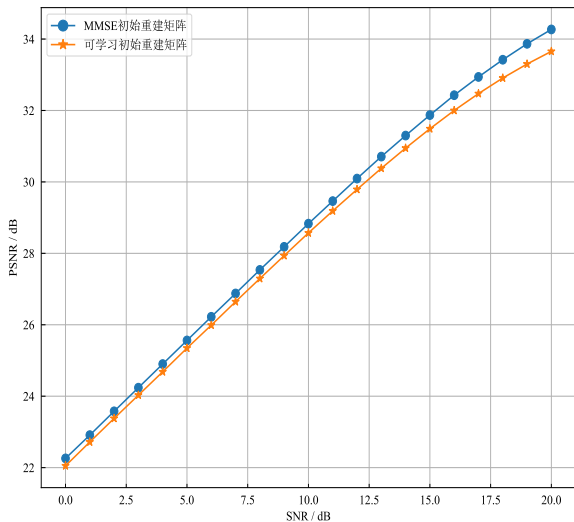


图 12 MMSE 初始重建方法与可学习初始重建方法的性能比较

5.6.2 深度重建网络迭代次数的影响

在主要结果部分, BCS-JSCC 的总子网络迭代次数被固定为 9 次. 为了进一步研究子网络迭代次数对 BCS-JSCC 的影响, 本实验训练了五种具有不同迭代次数的模型. 在这些模型中, 深度重建子网络层数均设置为 3, 每层子网络的循环次数分别设置为 1、2、3、4、

5, 总的迭代次数则从 3 开始, 以步长 3 增加, 直到 15. 实验使用 CIFAR-10 数据集作为训练集和测试集, 在 AWGN 信道条件下评估上述五种模型在不同信噪比条件下的 PSNR 性能, 实验结果如图 13 所示. 由图 13 的结果可以看出, 随着子网络迭代次数的增加, BCS-JSCC 的性能有所提高, 但这种提升逐渐减小. 当迭代次数超过 9 时, 每增加 3 次迭代, 所得到的 PSNR 增益小于 0.1 dB, BCS-JSCC 的性能趋于收敛. 此外, 随着迭代次数的增加, 模型的解码延迟也显著上升. 当迭代次数过高时, 虽然性能有所增加, 但解码延迟的增幅过大, 导致性能提升的收益难以抵消解码延迟的增加. 因此, 过多的迭代次数在实际应用中可能得不偿失.

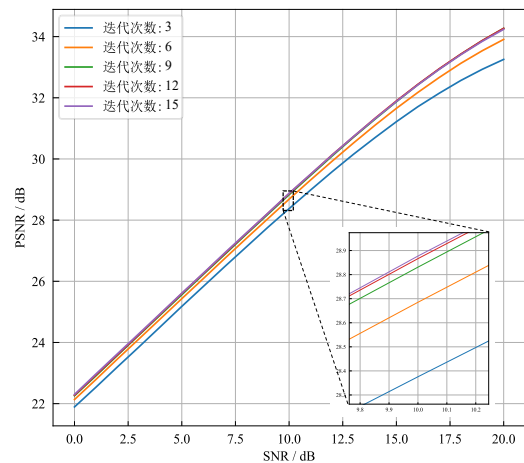


图 13 BCS-JSCC 模型不同迭代次数对应的 PSNR 性能曲线

6 结论

针对现有基于深度学习的 JSCC 方法在编码端存在较高计算复杂度, 难以在物联网边缘计算节点等计算资源受限场景中应用的问题, 本文提出了一种基于深度压缩感知的联合信源信道编码方法 BCS-JSCC. 在编码端, 通过压缩感知采样实现信号的联合编码, 并通过测量矩阵二值化减少编码器的存储需求; 在解码端, 通过基于 ISTA 算法的展开网络实现信号的联合解码, 并引入循环架构以充分利用重构子网络的网络容量, 降低解码器的存储需求. 实验结果表明, 在编码端计算量相同的条件下, BCS-JSCC 在高信噪比条件下相比于现有的基于深度学习的 JSCC 方法具有显著优势. 在低算力 (0.10 K FLOPs/pixel) 情况下, BCS-JSCC 在所有信噪比环境下均具有更优的性能表现. 此外, 由于测量矩阵的二值化操作, BCS-JSCC 的编码器大小显著低于对比模型. 解码器的循环架构在保持性能无明显下降的同时, 有效降低了模型的存储空间占用. 本文所提出的方法具有编码器简单而解码器复杂的特点, 编码器可

部署于物联网节点,而解码器则需在复杂终端上运行,适用于物联网节点采集图像并将测量值发送至终端处理的场景.

参考文献

- [1] SHANNON C E. A mathematical theory of communication[J]. ACM SIGMOBILE Mobile Computing and Communications Review, 2001, 5(1): 3-55.
- [2] CAMICIOTTI L, LAMY C, MEILHAC L, et al. Joint-source-channel coding for 4G multimedia streaming[C]// 2nd WWRP Meeting. Helsinki, Finland: WWRP, 2001.
- [3] KIM J, MERSEREAU R M, ALTUNBASAK Y. Error-resilient image and video transmission over the Internet using unequal error protection[J]. IEEE Transactions on Image Processing, 2003, 12(2): 121-131.
- [4] CAO L. On the unequal error protection for progressive image transmission[J]. IEEE Transactions on Image Processing, 2007, 16(9): 2384-2388.
- [5] PETTIJOHN B D, HOFFMAN M W, SAYOOD K. Joint source/channel coding using arithmetic codes[J]. IEEE Transactions on Communications, 2001, 49(5): 826-836.
- [6] KATABI D, RAHUL H, JAKUBCZAK S. Softcast: One video to serve all wireless receivers[R/OL]. (2009-02-07) [2024-12-03]. <https://dspace.mit.edu/handle/1721.1/44585>.
- [7] 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.
- [8] 穆天杰, 陈晓辉, 汪逸云, 等. 基于深度学习的信源信道联合编码方法综述[J]. 电信科学, 2020, 36(10): 56-66.
MU T J, CHEN X H, WANG Y Y, et al. A survey on deep learning based joint source-channel coding[J]. Telecommunications Science, 2020, 36(10): 56-66. (in Chinese)
- [9] 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.
- [10] 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.
- [11] 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.
- [12] XU J L, AI B, CHEN W, et al. Wireless image transmission using deep source channel coding with attention modules[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(4): 2315-2328.
- [13] YANG M Y, KIM H S. Deep joint source-channel coding for wireless image transmission with adaptive rate control[C]//ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE, 2022: 5193-5197.
- [14] WANG S X, YANG K, DAI J C, et al. Distributed image transmission using deep joint source-channel coding[C]// ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE, 2022: 5208-5212.
- [15] 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.
- [16] WANG S X, DAI J C, QIN X Q, et al. Improved nonlinear transform source-channel coding to catalyze semantic communications[J]. IEEE Journal of Selected Topics in Signal Processing, 2023, 17(5): 1022-1037.
- [17] ERDEMIR E, TUNG T Y, DRAGOTTI P L, et al. Generative joint source-channel coding for semantic image transmission[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(8): 2645-2657.
- [18] WU H T, SHAO Y L, BIAN C H, et al. Vision transformer for adaptive image transmission over MIMO channels[C]// ICC 2023 - IEEE International Conference on Communications. Piscataway: IEEE, 2023: 3702-3707.
- [19] 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.
- [20] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [21] 孙泽宇, 李传锋, 阎奔. 传感网中基于压缩感知的丢包匹配数据收集算法[J]. 电子学报, 2020, 48(4): 723-733.
SUN Z Y, LI C F, YAN B. CS-MDGA: A packet loss matching data gathering algorithm in sensor networks based on compressive sensing[J]. Acta Electronica Sinica, 2020, 48(4): 723-733. (in Chinese)
- [22] 杨浩, 王喜玮. 基于区域化压缩感知的无线传感器网络数据收集方法[J]. 计算机学报, 2017, 40(8): 1933-1945.
YANG H, WANG X W. Data gathering based on regionalized compressive sensing in WSN[J]. Chinese Journal

- of Computers, 2017, 40(8): 1933-1945. (in Chinese)
- [23] CHEN J, SHAO S, WANG T Y, et al. LDPC coded compressive sensing for joint source-channel coding in wireless sensor networks[J]. IEEE Transactions on Vehicular Technology, 2023, 72(2): 2145-2160.
- [24] ZHENG L Q, ZHU T T, MA X. Block compressed sensing-based joint source-channel coding for wireless image transmission[C]//2020 International Conference on Wireless Communications and Signal Processing. Piscataway: IEEE, 2020: 13-18.
- [25] LI L X, FANG Y, LIU L W, et al. Overview of compressed sensing: Sensing model, reconstruction algorithm, and its applications[J]. Applied Sciences, 2020, 10(17): 5909.
- [26] MOUSAVI A, PATEL A B, BARANIUK R G. A deep learning approach to structured signal recovery[C]//2015 53rd Annual Allerton Conference on Communication, Control, and Computing. Piscataway: IEEE, 2016: 1336-1343.
- [27] KULKARNI K, LOHIT S, TURAGA P, et al. ReconNet: Non-iterative reconstruction of images from compressively sensed measurements[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 449-458.
- [28] SHI W Z, JIANG F, LIU S H, et al. Image compressed sensing using convolutional neural network[J]. IEEE Transactions on Image Processing, 2020, 29: 375-388.
- [29] ZHANG Z H, LIU Y P, LIU J N, et al. AMP-net: Denoising-based deep unfolding for compressive image sensing[J]. IEEE Transactions on Image Processing, 2021, 30: 1487-1500.
- [30] 裴翰奇, 杨春玲, 魏志超, 等. 基于 SPL 迭代思想的图像压缩感知重构神经网络[J]. 电子学报, 2021, 49(6): 1195-1203.
- PEI H Q, YANG C L, WEI Z C, et al. Image compressive sensing reconstruction network based on iterative SPL theory[J]. Acta Electronica Sinica, 2021, 49(6): 1195-1203. (in Chinese)
- [31] ZHANG J, GHANEM B. ISTA-net: Interpretable optimization-inspired deep network for image compressive sensing[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 1828-1837.
- [32] ZHANG J, ZHAO C, GAO W. Optimization-inspired compact deep compressive sensing[J]. IEEE Journal of Selected Topics in Signal Processing, 2020, 14(4): 765-774.
- [33] ALHEJAILI R, ALFARRAJ M, LUQMAN H, et al. Recursions are all you need: Towards efficient deep unfolding networks[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2023: 4705-4714.
- [34] KRIZHEVSKY A. Learning multiple layers of features from tiny images[R/OL]. (2009-08-08)[2024-12-03]. <https://www.cs.utoronto.ca/~kriz/learning-features-2009-TR.pdf>.
- [35] Eastman Kodak Company. Kodak lossless true color image suite[EB/OL]. (1999-01-01)[2024-12-03]. <http://r0k.us/graphics/kodak/>.
- [36] KINGMA D P, BA J. Adam: A method for stochastic optimization[EB/OL]. (2017-01-30)[2024-12-03]. <https://arXiv.org/abs/1412.6980>.
- [37] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: Dataset and study[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2017: 1122-1131.
- [38] PASZKE A, GROSS S, MASSA F, et al. PyTorch: An imperative style, high-performance deep learning library[EB/OL]. (2019-12-03)[2024-12-03]. <https://arXiv.org/abs/1912.01703>.
- [39] TAN M X, LE Q V. EfficientNet: Rethinking model scaling for convolutional neural networks[EB/OL]. (2020-09-11)[2024-12-03]. <https://arXiv.org/abs/1905.11946>.
- [40] BELLARD F. The BPG image format[EB/OL]. (2015-09-20)[2024-12-03]. <http://bellard.org/bpg/>.

作者简介



董浩 男, 2002年2月出生于辽宁省葫芦岛市. 2024年获得北京航空航天大学电子信息工程专业学士学位. 现为上海交通大学集成电路学院硕士研究生. 主要研究方向为图像与视频编码.

E-mail: 124034910027@sjtu.edu.cn



李劭辉 男, 1994年5月出生于山东省兰陵县. 2016年获得大连理工大学电子信息工程专业学士学位, 并于2019年和2022年分别获得上海交通大学信息与通信工程专业硕士和博士学位. 现为清华大学深圳国际研究院助理研究员. 主要研究方向为图像视频编码、信号处理.

E-mail: lishaohui@sz.tsinghua.edu.cn



阚诺文 男, 1996年12月出生于湖南省邵阳市. 2024年获得上海交通大学计算机科学与技术博士学位. 现为上海交通大学信息与通信工程流动站博士后. 主要研究方向为多媒体通信、深度强化学习和基于深度学习的网络优化. 中国电子学会会员编号: E190187389M.
E-mail: kannw_1230@sjtu.edu.cn



郑紫阳 男, 1996年1月出生于安徽省蒙城县. 2016年获得西安交通大学应用物理学学士学位, 2024年获得上海交通大学信息与通信工程专业博士学位. 现为上海交通大学集成电路学院助理教授. 主要研究方向为多媒体信号处理、物理神经网络等. 中国电子学会会员编号: E190186001M.
E-mail: zhengziyang@sjtu.edu.cn



戴文睿 男, 1983年5月出生于上海市. 上海交通大学教授、博士生导师. 主要研究方向为多媒体信号处理与编码、机器学习. 中国电子学会会员编号: E190014727M.
E-mail: daiwenrui@sjtu.edu.cn



熊红凯 男, 1973年7月出生于江西省南昌市. 上海交通大学特聘教授、博士生导师. 主要研究方向为信号表示与小波分析、图像视频编码、多媒体网络通信、计算机视觉、机器学习. 中国电子学会会员编号: E190013028F.
E-mail: xionghongkai@sjtu.edu.cn