



专题：智能通信技术

基于深度学习的信源信道联合编码方法综述

穆天杰, 陈晓辉, 汪逸云, 马陆鹏, 刘东, 周晶, 张文逸
(中国科学技术大学, 安徽 合肥 230026)

摘要: 信息论的经典结果表明, 信源信道分离编码是渐进最优的。但现代通信系统对时延、带宽等愈发敏感, 分离设计对解码具有无限计算能力这一假设难以成立。带宽有限时, 相对于信源信道联合编码, 分离编码已被证明是次优的。传统的联合信源信道编码需要复杂的编码方案, 相较之下, 数据驱动深度学习技术则带来了新的设计思路。适时地对相关研究成果进行总结, 有助于进一步明确深度学习方法解决信源信道联合编码问题的方式, 为研究新的研究方向提供依据。首先介绍了基于深度学习的信源压缩方案和端到端收发信机模型, 随后分析不同信源类型下的两种联合编码设计思路, 最后探讨了基于深度学习的信源信道联合编码的潜在问题和未来的工作方向。

关键词: 深度学习; 图像/视频压缩; 端到端收发信机; 信源信道联合编码

中图分类号: TN929.5

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2020290

A survey on deep learning based joint source-channel coding

MU Tianjie, CHEN Xiaohui, WANG Yiyun, MA Lupeng, LIU Dong, ZHOU Jing, ZHANG Wenyi
University of Science and Technology of China, Hefei 230026, China

Abstract: Classical information theory shows that separate source-channel coding is asymptotically optimal over a point-to-point channel. As modern communication systems are becoming more sensitive to delays and bandwidth, it becomes difficult to adopt the assumption that such separate designs have unlimited computing power for encoding and decoding. Compared to joint source-channel coding, separate coding has proven to be sub-optimal when the bandwidth is limited. However, conventional joint source-channel coding schemes require complicated design. In contrast, data-driven deep learning brings new designing ideas into the paradigm. A summary of relevant research results was provided, which will help to clarify the way in which deep learning methods solve the joint source-channel coding problem and to provide an overview of new research directions. Source compression schemes and end-to-end communication system models were firstly introduced, both based on deep learning, then two kinds of joint coding designs under different types of source, and potential problems of joint source-channel coding based on deep learning and possible future research directions were introduced.

Key words: deep learning, image/video compression, end-to-end transceiver, joint source channel coding

收稿日期: 2020-09-20; 修回日期: 2020-10-10

通信作者: 张文逸, wenyizha@ustc.edu.cn

基金项目: 国家重点研发计划项目 (No.2018YFA0701603)

Foundation Item: The National Key Research and Development Program of China (No.2018YFA0701603)

1 引言

传统的通信系统通过信源编码、信道编码实现图像/视频的传输,并将信源和信道编码划分为两个模块独立设计。信源信道独立编码(separate source-channel coding, SSCC)具有设计简单、模块化的特点,但在实际通信系统中,系统可容忍的时延、复杂度和码长有限,难以满足香农分离定理中码长不受限的假设。此外,分离定理适用于点对点通信系统以及信源、信道分布已知的情况,而多用户通信和具有时变信道的移动通信等场景中分离定理并不适用。在这些场景下,采用分离设计无法保证系统的最佳性能^[1]。Gamal-Kim 从理论上阐述了多种信源信道联合编码(joint source-channel coding, JSCC)模式,可以利用多用户信源之间的相关性取得比 SSCC 更好的性能^[1]。

基于 JSCC 理论的实践方案有多种,其中一种是由 Wu 团队和 Katabi 团队在 2010 年提出的 Softcast 模式^[2-3],通过变换对输入图像或视频进行处理,并将系数直接调制为密集的星座图进行传输,无须常规的量化、熵编码和信道编码。2019 年, Wu 团队还提出了一种在瑞利衰落信道下进行无编码视频传输的 Softcast 方案^[2],针对信源数据的特点提出了一种信道分配和功率分配算法,在提升高质量数据的分集增益和减少低质量数据的损失之间进行了折中处理。

上述 Softcast 方案的局限性在于需要提前获得确定的信道模型,而真实环境下的信道特性往往是不可预测的。近年来,深度学习(deep learning)技术在图像、语音压缩等领域的发展为 JSCC 的设计提供了新的思路。自编码器(autoencoder)是一种数据降维和重构的神经网络结构,可以通过无监督学习得到原始数据的低维表示。在这种结构的启发下,近期有研究提出用深度神经网络代替传统编解码,设计端到端的信源信道联合编码结构^[4-6]。这种无监督的训练方式

不仅可以有效地避免未知信道对数据重构的影响,而且可以挖掘多用户信源之间的相关性,在特定场景下能获得很好的效果,这是 JSCC 理论的有效实践。

信源压缩的目的是在通过编码去除信源内部冗余信息的同时,在解码端尽可能准确地恢复出原始信源数据。与之不同的是,信道编码需要对原始比特序列增加校验位来实现检错、纠错的功能,以此增加比特序列传输对于噪声信道的稳健性。从设计思路上来说,信源压缩和信道编码是对立的,若将它们分离设计,则难以取得最优的折中方案。鉴于自编码器编解码器无监督联合优化的特点,可以考虑使用人工神经网络对信源压缩和信道编码进行联合设计,从而使通信系统达到端到端的最优性能。那么,如何才能合理设计自编码器的结构,在最大限度压缩信源信息的同时,提高对信道噪声的稳健性?这需从以下两个点去考虑。

其一,基于深度学习的信源编码结构可以为设计 JSCC 网络提供指导。设计 JSCC 时,需要结合不同信源的特点选择不同的网络结构,从而在网络隐层尽可能多地提取信源的特征图以提高重构质量。目前,相当一部分深度学习图像压缩算法在性能上已经能够超越 JPEG、JPEG2000、BPG 等传统方法^[7-8],而深度学习视频编码的性能在某些指标上能够达到或超越目前的一些主流标准(H.264、H.265 等)^[9-10]。

其二,与信源压缩的单模块设计不同, JSCC 是通信系统的端到端抽象,因此 JSCC 设计需要考虑多径衰落、频率选择性衰落等信道以及对应的 MIMO、OFDM 等通信系统模型。同时,如何进行信道条件未知或不可微时的联合优化也是值得考虑的问题。近些年提出的基于自编码器的端到端收发信机模型^[11-13]同样在讨论上述问题,可以为设计复杂信道环境下的信源信道联合编码提供参考。



因此，基于深度学习的信源编码和端到端收发信机的设计，能给实现信源信道联合编码带来启发与指导。本文首先分别介绍了基于深度学习的信源压缩编码方法以及端到端的收发信机模型，阐述获取信源有效低维表示与构建不同仿真环境下端到端收发信机的思路。接着介绍了基于深度学习技术的信源信道联合编码研究现状。最后，将讨论这一研究面临的问题和挑战。

2 基于深度学习的信源压缩编码

得益于深度神经网络提取信源特征的能力，深度学习技术在信源压缩编码领域取得了比传统方法更优异的效果。现有信源压缩工作研究的主要信源类型为图像和视频，但是也有一部分工作是针对语音^[14]和文本类型^[15]的压缩。因此，本节将介绍基于深度学习的图像和视频压缩编码方法，阐述对应的设计思路。

2.1 基于深度学习的图像压缩编码

在现有的基于深度学习的图像压缩编码方案中，既有沿用传统的基于卷积神经网络（convolutional neural network, CNN）、循环神经网络（recurrent neural network, RNN）的自编码器结构，也有利用生成对抗网络来进行图像重建的架构。

Ballé研究了用于图像压缩的神经网络中的激活函数，指出广义除性正规化（generalized divisive normalization, GDN）具有较好的去相关压缩能力^[16]。并通过端到端训练优化编码的率失真函数^[17]。为进一步优化，研究人员提出了基于超先验的方案^[16]和超先验结合自回归模型方案^[18]，实现了高效熵编码，在压缩效率方面超过了传统图像压缩BPG方法。循环神经网络由于其引入数据间长期或短期记忆的有效性，在图像压缩领域中取得了较为突出的效果。Google的Toderici团队在2015年提出一种基于长短期记忆（long short-term memory, LSTM）的神经网络框架来实现可变速率

的图像压缩^[19]。次年，在参考文献[19]的基础上，Toderici团队又提出一种全分辨率有损图像压缩方法^[20]，在原先LSTM网络中又加入GRU结构和残差模块，并采用熵编码进一步提高图像压缩比，在相同压缩比下取得了比JPEG更优的性能。

生成对抗网络（generative adversarial network, GAN）的生成特性可以进一步提高压缩效率。2019年，Agustsson等^[21]提出了一种基于GAN的图像压缩算法，该网络根据语义标签映射对部分图像编码；在恢复图像时，编码部分通过解码进行恢复，没有编码的部分则通过生成器自动生成，极大地提高了压缩效率，在低码率的情况下优于BPG算法。

2.2 基于深度学习的视频压缩编码

基于深度学习的视频编码可以分为两类：一类是研究基于深度神经网络的端到端编码方法，另一类是设计深度神经网络来替换传统编码方案中的某一模块^[22]。本节分别介绍该两类方法，并以DLVC视频压缩方案为例进行具体分析。

一方面，端到端的视频编码方案多以自编码器结构为基础。2019年，Lu等^[9]提出了一种端到端的深度视频编码方案，对输入视频的每一帧进行压缩，并使用光流估计模块来获取当前帧与前一帧之间的运动信息，并使用两个自编码器压缩预测残差和运动信息取得了比H.264标准更低的压缩率。Rippel等^[10]在2019年提出了一种深度视频编码方案，该方案仅采用一个自编码器同时压缩运动信息和预测残差并且加入了多帧多光流运动补偿和速率控制算法性能优于H.265参考软件。另一方面，使用神经网络替代传统视频编码中部分模块也被广泛探讨，被替代的模块包括帧内/帧间预测、变换、上下采样和编码优化等^[23]。2018年，Felix等^[12]提出了一种全连接网络用于帧内预测，使用多模型的训练策略以提高压缩率。同年，Lin等^[24]采用一组GAN

用于帧间预测, BD-rate (bjøntegaard-delta rate) 约比 HM 降低了 2%。

DLVC (deep learning video coding) 是为 JCFP (joint call for proposal) 开发的一个优于 H.265 的视频压缩方案, 属于上述第二类方法。该方案采用了两个深度编码模块: 基于 CNN 的内环路滤波器 (CNN-based in-loop filter, CNN-ILF) 和基于 CNN 的块自适应分辨率编码 (CNN-based block adaptive resolution coding, CNN-BARC), 整体结构如图 1 所示。对于 CNN-ILF 模块, DLVC 用一组自然图像训练网络, 对每个图像在不同量化参数下进行帧内编码。CNN-BARC 是一个基于下采样和上采样的编码模块, 它使用训练好的 CNN 模型作为下采样和上采样过滤器。DLVC 中每个编码树单元 (coding tree unit, CTU) 可选择直接编码或下采样编码, 而下采样编码又分为两个子模式: 基于 CNN 的上下采样或插值法。在多组视频集上的实验结果表明, 与 JEM 在随机访问和低位延配置下相比, DLVC 在 Y 通道下的平均 BD-rate 分别低 10.1% 和 11.8%。从 BD-rate 的角度来看, DLVC 是最佳方案之一。

3 端到端的通信系统模型

在计算机视觉领域, 一般难以找到特定的数学模型去表示信源特征。与之不同的是, 由于信息论的发展以及各种信道模型的建立, 现有的通信系统可以被特定的数学模型较好地描述^[25]。尽管现代通信系统的发展已经日趋成熟, 但是还有如下的几个问题需要解决^[26]。

首先, 多数的通信领域的信号处理算法拥有坚实的统计学和信息论基础, 并且可以被证明是最优的。这些算法通常是线性的、稳定的并且拥有高统计特性。但是, 在一个实际的通信系统中, 有很多的缺陷和非线性模块 (例如非线性功率放大器、量化) 只能被这些算法近似地描述。

其次, 通信系统设计的一项指导原则就是将信道处理的过程分解为一系列独立的子模块, 每个模块会执行一个独立的函数 (如信源编码、信道编码、调制、信道估计、信道均衡)。尽管这种设计的模式已经应用在几乎所有的现代通信系统中, 但是, 这种分模块独立设计的模式仍然不能保证端到端的最优性能。例如, 在实际信道或者

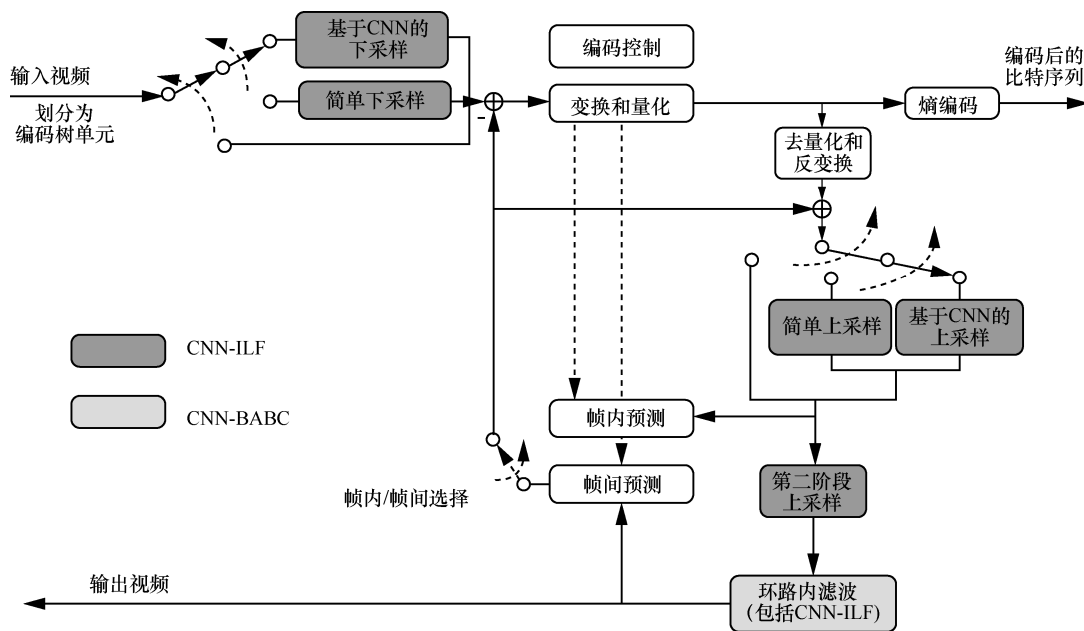


图1 DLVC 方案结构



块长度较短的情况下，信源编码和信道编码^[27]、信道编码和调制^[28]的分离设计均为次优的。若对这些子模块做联合优化，如使用因子图^[29]的编码调制一体化，会得到性能上的提升，但是也会相应地增加系统的计算复杂度。

为了解决以上传统收发信机存在的问题，2017年，O'Shea等^[30]率先使用深度学习中的自编码器来设计端到端优化的收发信机。这种设计的出发点是，通信系统物理层收发信机与自编码器在功能和结构上是相似的。自编码器的主要功能是实现数据重构，而通信系统的主要功能是在接收端恢复出发射端的信号。若把收发信机看成一种自编码器结构，则发射机与接收机分别对应于自编码器的编码器与译码器。因此，通信系统收发信机的最优化设计就转变为自编码器端到端的优化任务。参考文献[30]提出的基本框架如图2所示。

不同于传统设计方式，基于自编码器的通信系统设计是一种全新的思路^[31]。神经网络通过大量训练样本学习数据的分布，然后预测结果。在对端到端系统做联合优化时，信道需要作为一个模块融入至整个网络中。本节从信道模型出发，讨论确定信道模型与未知信道模型的端到端系统。

3.1 确定信道模型的端到端系统

确定信道模型的端到端系统大多选择高斯白噪声信道或瑞利衰落信道作为仿真信道。在参考文

献[30]中，基于自编码器的通信系统在高斯白噪声信道下自动学习从比特块对应的独热（one-hot）向量到星座点符号的映射关系，并且在接收端对接收符号进行解码，恢复出原始比特。仿真结果表明在相同的编码速率下基于全连接神经网络方案的块误码率（block error rate, BLER）低于传统的调制编码方案，如基于汉明码的BPSK。之后，O'Shea等^[32]将自编码器模型推广到了瑞利衰落信道下的MIMO系统，讨论了不同量化精度的信道状态信息对传输质量的影响，并实验证明了该方案的性能接近甚至在某些情况下优于现有的分集与复用方式。参考文献[33]又将自编码器模型扩展到有干扰的MIMO系统，通过将MIMO系统建模为两个共用同一信道的自编码器并联合训练该模型以达到消除干扰的目的。

基于自编码器的通信系统设计方法同样适用于OFDM系统。2019年，Felix等^[12]设计的系统对每个子载波都使用了自编码器进行调制编码。实验结果表明该系统的块误码率比传统的OFDM系统低2 dB（QPSK调制），并且与传统OFDM系统具有相同的优势，即单抽头均衡和针对采样同步错误的稳健性，从而能够在多径信道下进行可靠通信。

3.2 未知信道模型的端到端系统

第3.1节介绍的自编码器端到端学习过程需要信道的先验知识进行梯度下降优化。实际场景中往往难以获得精确的信道传输函数，需要考虑

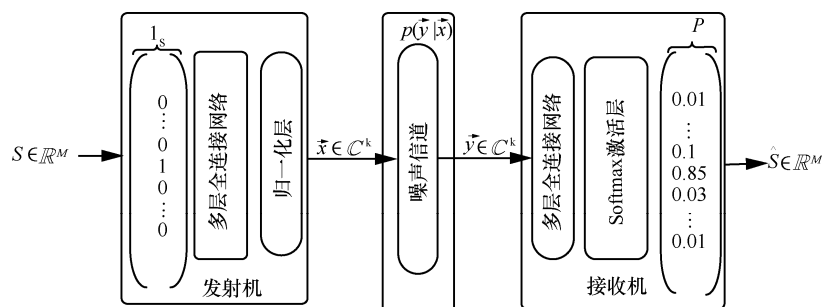


图2 基于自编码器通信系统的基本框架

如何在未知信道信息的情况下进行训练。现有的解决方法可以分为 3 类：强化学习、GAN 与元学习。

2018 年, Aoudia 等^[11]提出了一种基于强化学习的方法解决信道梯度消失的问题。它将发射机看作智能体 (agent), 将信道和接收机看作环境 (environment)。智能体不断获得环境中的反馈 (reward) 信息, 并且优化自身使反馈最大化。接收机计算端到端的损失并反馈给发射机。参考文献[11]证明这种迭代算法能在没有先验知识的情况下适应任何信道。

2019 年, Ye^[34]提出了一种基于条件 GAN 的生成网络降低数据集大小并解决信道梯度消失的问题。该方案使用条件 GAN 来生成信道传输函数, 并且将发射机的调制符号用作条件 GAN 的条件输入。参考文献[34]的仿真结果证明了该网络结构在加性高斯白噪声、瑞利衰落信道和频率选择性衰落信道下的有效性。

此外, 元学习作为一种新的研究思路, 也能在一定程度上实现信道模型未知条件下的端到端训练^[35]。在信道模型未知的情况下, 假定有一个包含一组预设信道模型的集合, 网络在该集合上执行元学习, 训练后得到的模型可以在很小的样本数或迭代次数下收敛并适应新的信道。因此, 在这种训练策略下, 即使信道参数未知, 也能以极小的代价得到一组合适的收发信机参数, 从而满足通信实时性需求。

4 深度信源信道联合编码

经典端对端无线通信系统如图 3 所示。首先, 使用信源编码 (JPEG、WebP 等) 对信源 x 去除冗余得到比特流 s ; 然后, 对 s 进行信道编码 (Turbo、LDPC 等) 得到 y , 以增加相应的校验位来抵抗信道噪声; 最后, 对比特流 y 进行调制 (BPSK、16QAM 等), 将比特映射为星座点集合中的符号, 符号的同相分量和正交分量分别由通

信链路的 I 和 Q 两路信号承载并发送^[36]。对应地, 接收端对经过信道后的符号 z 进行解调、解码操作得到 $\bar{x}$ 。

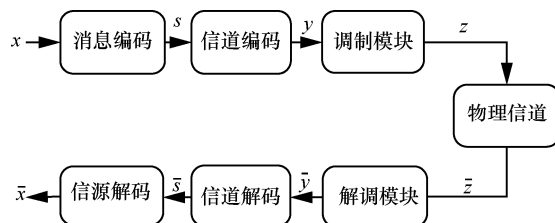


图3 经典端对端无线通信系统框架

依据定义信道的方式不同, 基于深度学习的 JSCC 可归结为两类设计思路。第一类思路受无编码传输^[3, 37-38]的启发, 将信源编码、信道编码和调制联合设计为编码器, 其系统模型如图 4 所示。也就是说, 信源符号序列会被编码器直接映射到信道符号。信道符号经过 AWGN 等连续物理信道之后, 被解码模块在信宿重建。另一类思路将通信系统中的调制、噪声信道、解调模块抽象为离散的二进制信道^[39-41], 其系统模型如图 5 所示。该模型中, 信源经过编码器联合编码为比特流, 比特流经过离散的二进制信道后, 被解码模块重建。为方便描述, 本文将前者称为基于物理信道的符号编码, 后者称为基于抽象信道的比特编码。

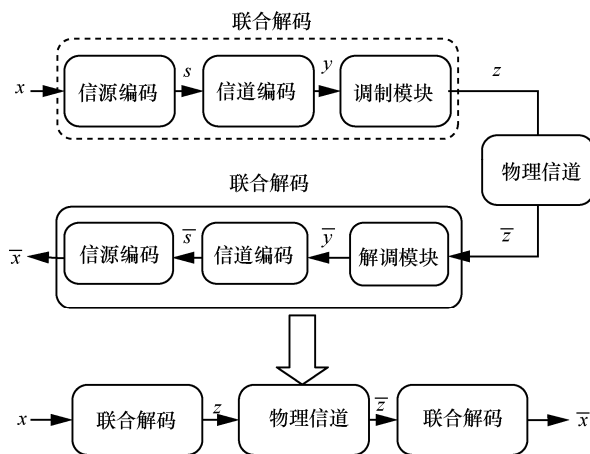


图4 基于物理信道的 Deep JSCC 模型

上述两种设计思路建立的框架中, 还需考虑编解码器的网络实现方式。深度学习领域, 针对

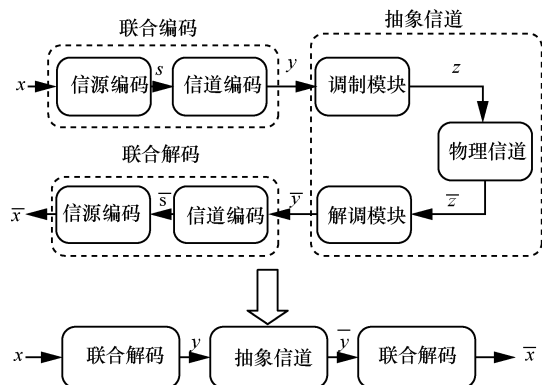


图5 基于抽象信道的 Deep JSCC 模型

不同类型的数据设计出了不同种类的结构，因此具体网络结构的实现与信源的特点密切相关。信源可以根据其是否具有结构化特征划分为两类，结构化信源（典型如图像、视频）和非结构化信源（如高斯信源）。下面分析这两种信源所适合的网络结构。

结构化信源是 Deep JSCC 的主要研究场景：由于神经网络对结构化数据强大的特征获取能力以及针对各种结构化数据设计的网络结构的出现，Deep JSCC 相较于传统设计更具有优势。图像/视频等具有空间拓扑结构信源适合 CNN 等网络结构，文本/语音等具有时间序列化结构信源适合 RNN 等网络结构。

非结构化信源情形下，Deep JSCC 则稍显羸弱。这是由于非结构化信源内部相关性弱，难以去除冗余。目前已经有部分工作对 AWGN 信道联合编解码设计，另一部分工作则专注于解码器的设计^[42-43]。

以上研究思路可以总结见表 1。本节围绕两类信道下，图像、文本及高斯信源情形，阐述已有研究工作进展。

4.1 基于物理信道的符号编码

2019 年，Gunduz 团队^[5]提出了以图像为信源

的基于 CNN 的 Deep JSCC 方案，分别研究了在给定信噪比下的 AWGN 和瑞利衰落信道下的效果。实验表明，Deep JSCC 相对于信源信道编码分离设计的方案有更优的效果，在低信噪比的信道环境下，Deep JSCC 优势尤其明显。在具有 GPU 的计算环境下，Deep JSCC 对高分辨率图片的编解码时延相对传统编解码器低了一个数量级。针对该方案信噪比适应性弱的缺陷，Gunduz 团队^[53]进一步提出将噪声反馈模块融入传输系统，以增强编解码器对变化信噪比的稳健性。解码器将一部分经过噪声信道的接收符号 $\bar{z}$ 反馈给编码器，编码器根据 $\bar{z}$ 重新计算信噪比，并对编解码网络进行调整，以适应变化的信噪比环境。

在上述基本模型的基础上，一些研究拓展了应用场景。如 Gunduz 团队^[54]设计了一种自适应带宽的联合编码方案，在不同带宽，不同信道模型情形均具有较强的自适应性。再如，Jankowki 等^[45]设计了基于 Deep JSCC 的无线图像检索方案，将信源数据直接映射到信道输入符号，从而在提高端到端图像检索精度的同时，简化、加快了编解码操作。Zarcone 等^[46]提出了在 PCM (phase-change memory) 设备上存储 Deep JSCC 压缩数据的方案，是 Deep JSCC 的硬件实践。

与上述基于结构化信源的 Deep JSCC 不同，还有一些研究者探讨了非结构化信源条件下的 Deep JSCC 方案。2019 年，Yashas^[55]提出了对 AWGN 信道下独立同分布高斯信源的联合编码的 Deep JSCC 方案。同年，Yashas 通过假设接收信号和重构信号的高斯统计特性，给出了正则化 MSE 损失的可变上限证明^[56]。进一步地，他们基于变分自编码器和拉格朗日正则器进一步改善了

表 1 JSCC 研究分类示意图

Deep JSCC	物理信道 (AWGN, 瑞利信道)	抽象信道 (BSC, BEC)
结构化信源 (图像, 文本)	Gunduz 团队 ^[5, 6, 38, 44] , Mikolaj ^[45] , Ryan ^[46]	Kristy ^[47] , Song ^[48] , Shao ^[48-50] , Nariman ^[51]
非结构化信源 (高斯信源)	Yashas ^[51] , Krishna ^[52]	Carpi ^[42] , Jiang ^[43]

Deep JSCC 的性能^[57]。

2020 年, Krishna 团队^[52]提出了基于 RNN 结构的联合设计方案, 用于高斯信源的联合编解码。相比于 Yashas 等^[55]的工作, Krishna 等的工作不需要获取信源的先验信息。除此之外, Krishna 在理论上证明了 Deep JSCC 的有效性, 并证明了基于深度学习的编码器与基于混沌动态系统 (chaotic dynamical system) 的编码函数之间的相似性, 通过实验证明该方案对信噪比不匹配的稳健性。

4.2 基于抽象信道的比特编码

与传输符号流的 Deep JSCC 方案不同, 二进制信道下传输离散比特流无法计算反向传播梯度。因此, 离散信道的嵌入也比物理信道的嵌入更为复杂。近年来神经网络离散化^[58-60]和离散自编码器^[61]的发展, 为上述难点提供了解决思路。针对离散化神经网络的问题有以下研究工作, 一个简单的方法是使用得分函数估计器代替梯度^[62]。由于该估计方差较高, 一部分工作提出了不同的公式和控制变量来解决该问题^[63-67]。另外, 为了达到使离散随机变量连续化的目的, Jang 和 Maddison 分别提出了 Gumbel-Softmax 分布^[68]和 Concrete 方案^[69]。

在基于比特流的 Deep JSCC 工作中, 针对结构化信源的研究有较大进展。对于图像信源, Kristy 等基于离散自编码器提出了 NECST (neural joint source-channel coding)^[47], 为了保留编码的硬离散性, 他们使用了 VIMCO^[69]多样本变分下界目标, 用于获得低变差梯度。实现了 BSC 信道下的联合编码。实验表明, 固定失真下, 该方案在所有测试信噪比下均比 JPEG 与理想信道编码方案性能更佳; 固定速率下, 该方案达到了与离散变分自编码器 (discrete variational autoencoder, DVAE) 和 LDPC 编码相比拟的性能。为了增强 NECST 的压缩和纠错能力, Song 等^[48]提出了新的正则化方法 IABF, 改善了 NECST 的稳健性。Song 提出了新的损失函数, 实现了网络对高维数

据更有效的优化。而 Shao 等^[50]研究了可部署到计算能力有限的移动设备中的低功耗 Deep JSCC, 提出了移动设备下的联合编码模型 Bottlenet++。BEC 信道下, Bottlenet++以精度下降 2% 的代价实现了 256 倍的位压缩率; AWGN 信道下, Bottlenet++实现了 64 倍的带宽缩减。

针对文本信源, Farsad^[51]提出了基于 RNN 结构的 Deep JSCC 方案, 实现了 BEC 信道下的联合编解码。结果表明, 编码比特块较短时, 该方案比传统方法具有更低的单词错误率 (WER)。与传统联合编码不同, 基于深度学习的编解码方案重建的文本, 能更好地保留语句的语义信息。

值得注意的是, 基于比特流的 Deep JSCC 的研究中, 针对信源为非结构化信源的工作较少。Carpi 提出了一种基于强化学习的比特位反转解码方式^[42], 让解码器由数据驱动去学习最佳的解码策略, 在某些情况下得到最佳的性能。Jiang^[43]提出了 DeepTurbo 解码器, 通过端到端的训练使 DeepTurbo 在 AWGN 信道下取得突出的性能

5 深度信源信道联合编码面临的问题和挑战

目前, 基于深度学习的信源信道联合编码处于初步发展阶段, 考虑通信系统的实时性与有效性, 往往设计系统需要考虑算法性能和复杂度之间的平衡。同时, 信源信道联合编码还面临着适应复杂且时变的实际信道的挑战。因此, 基于深度学习的信源信道联合编码走向成熟仍需解决诸多问题。

首先, Deep JSCC 方案的普适性有待提升。从信道适应性看, 数据驱动的模式导致 Deep JSCC 需要针对不同的无线信道环境提出对应的解决方案, 仍需探索如何增强 Deep JSCC 在复杂无线信道环境下的泛化性。从信源适应性角度看, 已有研究集中在定长编码, 适应视频等数据传输的变长编码、自适应编码罕见探讨, 如何利用视频特征也缺乏深入的研究。

其次, Deep JSCC 网络的设计有待细化。如



针对图像的 Deep JSCC 依赖上采样和下采样操作已避免维度灾难, 难以避免压缩带来的损伤; 广泛使用的 MSE 损失函数也往往导致重建数据较大的高频失真。寻找更好的网络结构、设计更合适的目标优化函数是已有 Deep JSCC 方案改进的自然选择。

另外, Deep JSCC 的拓展应用场景仍讨论不足。少量研究开始关注算法实现的复杂度和功耗, 进一步能结合应用场景特征做针对性设计讨论。如在物联网等计算资源不对称的情形, 数据的编码可由中心节点完成, 解码则由边缘设备完成。这意味着编码器可使用复杂的网络结构, 但解码器的网络需要尽可能简单, 此类 Deep JSCC 仍鲜见研究。

最后, 基于抽象信道的比特编码仍局限于二进制信道。为了能在实际通信系统中应用, 势必要将抽象信道模块展开为调制、实际信道、解调等模块, 以研究基于深度学习联合比特编码在实际无线信道下的效果。能否结合基于物理信道的部分已有研究成果得到了新的 JSCC 系统结构, 也有待进一步探索。

6 结束语

现代通信系统对传输时延和带宽的要求日益提升, 传统分离设计的信源信道编码方案难以保证端到端最优性能。本文从基于深度学习的信源压缩、端到端收发信机、信源信道联合编码模型几方面进行了综述。从本文分析可以看出, 深度学习在数据压缩和新型收发信机等相关研究中已有较为显著的成果, 这促使研究者们开始探索基于深度学习的信源信道联合编码方案。结合不同信道定义, 针对不同信源特征的 Deep JSCC 已取得了一些成效, 但在方案设计、应用方式及应用场景拓展等方面仍面临诸多挑战。

参考文献:

- [1] GAMAL A E, KIM Y H. Network information theory[M]. Cambridge: Cambridge University Press, 2011.
- [2] CUI H, LUO C, CHEN C W, et al. Robust uncoded video transmission over wireless fast fading channel[C]//Proceedings of IEEE INFOCOM 2014-IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2014: 73-81.
- [3] JAKUBCZAK S, KATABI D. SoftCast: clean-slate scalable wireless video[C]//Proceedings of the 2010 ACM Workshop on Wireless of the Students, by the Students, for the Students. New York: ACM Press, 2010: 9-12.
- [4] KATSAGGELOS A K. Joint source-channel coding for video communications[M]//Handbook of Image and Video Processing. London: Elsevier, Academic Press, 2005: 1065-1082.
- [5] BOURTSOULATZE E, KURKA D B, GUNDUZ D. Deep joint source-channel coding for wireless image transmission[C]//ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2019: 4774-4778.
- [6] KURKA D B, GÜNDÜZ D. DeepJSCC-f: deep joint-source channel coding of images with feedback[J]. arXiv preprint arXiv: 11174, 2019.
- [7] BALLÉ J, LAPARRA V, SIMONCELLI E P. End-to-end optimized image compression[J]. arXiv preprint arXiv:1611.01704, 2016.
- [8] JOHNSTON N, VINCENT D, MINNEN D, et al. Improved lossy image compression with priming and spatially adaptive bit rates for recurrent networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018.
- [9] LU G, OUYANG W, XU D, et al. Dvc: an end-to-end deep video compression framework[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 11006-11015.
- [10] RIPPEL O, NAIR S, LEW C, et al. Learned video compression[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 3454-3463.
- [11] AOUDIA F A, HOYDIS J. End-to-end learning of communications systems without a channel model[C]//Proceedings of 2018 52nd Asilomar Conference on Signals, Systems, and Computers. Piscataway: IEEE Press, 2018: 298-303.
- [12] FELIX A, CAMMERER S, DÖRNER S, et al. OFDM-autoencoder for end-to-end learning of communications systems[C]//Proceedings of 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). Piscataway: IEEE Press, 2018: 1-5.
- [13] O'SHEA T J, ROY T, WEST N, et al. Physical layer communications system design over-the-air using adversarial networks[C]//Proceedings of 2018 26th European Signal Processing Conference (EUSIPCO). [S.l.:s.n.], 2018: 529-532.
- [14] KANKANAHALLI S. End-to-end optimized speech coding with deep neural networks[C]//Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2018: 2521-2525.
- [15] AVINESH P, MEYER C M. Data-efficient neural text compression with interactive learning[C]//Proceedings of the 2019 Con-

- ference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). [S.l.:s.n.], 2019: 2543-2554.
- [16] BALLÉ J, MINNEN D, SINGH S, et al. Variational image compression with a scale hyperprior[J]. arXiv: 1802.01436, 2018.
- [17] BALLÉ J, LAPARRA V, SIMONCELLI E P. End-to-end optimization of nonlinear transform codes for perceptual quality[C]//Proceedings of 2016 Picture Coding Symposium (PCS). [S.l.:s.n.], 2016: 1-5.
- [18] MINNEN D, BALLÉ J, TODERICI G D. Joint autoregressive and hierarchical priors for learned image compression[C]// Proceedings of Advances in Neural Information Processing Systems. [S.l.:s.n.], 2018: 10771-10780.
- [19] TODERICI G, O'MALLEY S M, HWANG S J, et al. Variable rate image compression with recurrent neural networks[J]. arXiv: 1511.06085, 2015.
- [20] TODERICI G, VINCENT D, JOHNSTON N, et al. Full resolution image compression with recurrent neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 5306-5314.
- [21] AGUSTSSON E, TSCHANEN M, MENTZERF, et al. Generative adversarial networks for extreme learned image compression[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2019.
- [22] LIU D, LI Y, LIN J, et al. Deep learning-based video coding: a review and a case study[J]. ACM Computing Surveys (CSUR), 2020(53): 1-35.
- [23] LIU D, LI Y, LIN J, et al. Deep learning-based video coding: a review and a case study[J]. arXiv: Multimedia, 2019.
- [24] LIN J, LIU D, LI H, et al. Generative adversarial network-based frame extrapolation for video coding[C]//Proceedings of 2018 IEEE Visual Communications and Image Processing (VCIP). Piscataway: IEEE Press, 2018: 1-4.
- [25] O'SHEA T, HOYDIS J. An introduction to deep learning for the physical layer[J]. IEEE Transactions on Cognitive Communications and Networking, 2017(3): 563-575.
- [26] GÜNDÜZ D, KERRET P D, SIDIROPOULOS N D, et al. Machine learning in the air[J]. IEEE Journal on Selected Areas in Communications, 2019(37): 2184-2199.
- [27] GOLDSMITH A. Joint source/channel coding for wireless channels[C]//Proceedings of 1995 IEEE 45th Vehicular Technology Conference. Countdown to the Wireless Twenty-First Century. Piscataway: IEEE Press, 1995: 614-618.
- [28] ZEHAVERI E. 8-PSK trellis codes for a Rayleigh channel[J]. IEEE Transactions on Communications, 1992(40): 873-884.
- [29] WYMEERSCH H. Iterative receiver Design[M]. Cambridge: Cambridge University Press, 2007.
- [30] O'SHEA T, HOYDIS J. An introduction to deep learning for the physical layer[J]. IEEE Transactions on Cognitive Communications and Networking, 2017(3): 563-575.
- [31] 张静, 金石, 温朝凯, 等. 基于人工智能的无线传输技术最新研究进展[J]. 电信科学, 2018, 34(8): 46-55.
- ZHANG J, JIN S, WEN C K, et al. An overview of wireless transmission technology utilizing artificial intelligence [J]. Telecommunications Science, 2018, 34(8): 46-55.
- [32] O'SHEA T J, ERPEK T, CLANCY T C. Deep learning based MIMO communications[J]. arXiv preprint arXiv: 07980, 2017.
- [33] ERPEK T, O'SHEA T J, CLANCY T C. Learning a physical layer scheme for the MIMO interference channel[C]//Proceedings of 2018 IEEE International Conference on Communications (ICC). Piscataway: IEEE Press, 2018: 1-5.
- [34] YE H, LIANG L, LI G Y, et al. Deep learning based end-to-end wireless communication systems with conditional GAN as unknown channel[J]. arXiv: 02551, 2019.
- [35] FINN C, ABBEEL P, LEVINE S, et al. Model-agnostic meta-learning for fast adaptation of deep networks[Z]. 2017.
- [36] CAIRE G, TARICCO G, BIGLIERI E J I. Bit-interleaved coded modulation[J]. IEEE Transactions on Information Theory, 1998(44): 927-946.
- [37] GOBLICK T J. Theoretical limitations on the transmission of data from analog sources[M]. New York: ACM Press, 1965: 558-567.
- [38] TUNG T Y, GÜNDÜZ D. SparseCast: hybrid digital-analog wireless image transmission exploiting frequency-domain sparsity[J]. IEEE Communications Letters, 2018(22): 2451-2454.
- [39] NEAL R M J A I. Connectionist learning of belief networks[J]. Artificial Intelligence, 1992, 56(1): 71-113.
- [40] ROLFE J T J A P A. Discrete variational autoencoders[J]. arXiv: 1609.02200, 2016.
- [41] OORD A V D, VINYALS O. Neural discrete representation learning[C]//Proceedings of Advances in Neural Information Processing Systems. [S.l.: s.n.], 2017: 6306-6315.
- [42] CARPI F, HÄGER F C, MARTALÒ M, et al. Reinforcement learning for channel coding: learned bit-flipping decoding[C]//Proceedings of 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton). [S.l.: s.n.], 2019: 922-929.
- [43] JIANG Y, KANNAN S, KIM H, et al. Deepturbo: deep turbo decoder[C]//Proceedings of 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). Piscataway: IEEE Press, 2019: 1-5.
- [44] KURKA D B, GÜNDÜZ D. Successive refinement of images with deep joint source-channel coding[C]//Proceedings of 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). Piscataway: IEEE Press, 2019: 1-5.
- [45] JANKOWSKI M, GÜNDÜZ D, MIKOLAJCZYK K. Deep joint source-channel coding for wireless image retrieval[C]//Proceedings of ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2020: 5070-5074.
- [46] ZARCONI R, PAITON D, ANDERSON A, et al. Joint source-channel coding with neural networks for analog data compression and storage[C]//Proceedings of 2018 Data Compression Conference. [S.l.:s.n.], 2018: 147-156.
- [47] CHOI K, TATWAWADI K, WEISSMAN T, et al. NECST:



- neural joint source-channel coding[Z]. 2018.
- [48] SONG Y, XU M, YU L, et al. Infomax neural joint source-channel coding via adversarial bit flip[Z]. 2020.
- [49] JIANG F, TAO W, LIU S, et al. An end-to-end compression framework based on convolutional neural networks[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017(28): 3007-3018.
- [50] SHAO J, ZHANG J. Bottlenet++: an end-to-end approach for feature compression in device-edge co-inference systems[C]//Proceedings of 2020 IEEE International Conference on Communications Workshops (ICC Workshops). Piscataway: IEEE Press, 2020: 1-6.
- [51] FARSAD N, RAO M, GOLDSMITH A. Deep learning for joint source-channel coding of text[C]//Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2018: 2326-2330.
- [52] XUAN Z, NARAYANAN K. Analog joint source-channel coding for gaussian sources over AWGN channels with deep learning[C]//Proceedings of 2020 International Conference on Signal Processing and Communications (SPCOM). Piscataway: IEEE Press, 2020: 1-5.
- [53] KURKA D B, GÜNDÜZ D. DeepJSCC-f: deep joint source-channel coding of images with feedback[Z]. 2020.
- [54] KURKA D B, GÜNDÜZ D. Bandwidth-agile image transmission with deep joint source-channel coding[Z]. 2020.
- [55] SAIDUTTA Y M, ABDI A, FEKRI F. M to 1 joint source-channel coding of gaussian sources via dichotomy of the input space based on deep learning[C]//Proceedings of 2019 Data Compression Conference (DCC). Piscataway: IEEE Press, 2019: 488-497.
- [56] SAIDUTTA Y M, ABDI S, FEKRI F. Joint source-channel coding for gaussian sources over AWGN channels using variational autoencoders[C]//Proceedings of 2019 IEEE International Symposium on Information Theory (ISIT). Piscataway: IEEE Press, 2019: 1327-1331.
- [57] SAIDUTTA T M, ABDI A, FEKRI F. Joint source-channel coding of gaussian sources over AWGN channels via manifold variational autoencoders[C]//Proceedings of 2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton). [S.l.:s.n.], 2019: 514-520.
- [58] DUAN Y, LU J, WANG Z, et al. Learning deep binary descriptor with multi-quantization[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019(41): 1924-1938.
- [59] ZHANG Z, ZHAO Y, LIU J, et al. A general deep learning framework for network reconstruction and dynamics learning[J]. Applied Network Science, 2019(4): 1-17.
- [60] MENA G, BELANGER D, LINDERMAN S, et al. Learning latent permutations with gumbel-sinkhorn networks[J]. arXiv: 1802.08665, 2018.
- [61] VAHDAT A, ANDRIYASH E, MACREADY W. Dvae#: discrete variational autoencoders with relaxed boltzmann priors[C]//Advances in Neural Information Processing Systems. [S.l.:s.n.], 2018: 1864-1874.
- [62] WILLIAMS R J J M L. Simple statistical gradient-following algorithms for connectionist reinforcement learning[J]. Machine Learning, 1992, 8(3-4): 229-256..
- [63] WANG C, CHEN X, SMOLA A J, et al. Variance reduction for stochastic gradient optimization[C]//Proceedings of Advances in Neural Information Processing Systems. [S.l.:s.n.], 2013: 181-189.
- [64] GU S, LEVINE S, SUTSKEVER I, et al. Muprop: unbiased backpropagation for stochastic neural networks[Z]. 2015.
- [65] RUIZ F R, AUEB M T R, BLEI D. The generalized reparameterization gradient[C]//Proceedings of Advances in Neural Information Processing systems. [S.l.:s.n.], 2016: 460-468.
- [66] TUCKER G, MNIH A, MADDISON C J, et al. Rebar: low-variance, unbiased gradient estimates for discrete latent variable models[C]//Proceedings of Advances in Neural Information Processing Systems. [S.l.:s.n.], 2017: 2627-2636.
- [67] GRATHWOHL W, CHOI D, WU Y, et al. Backpropagation through the void: optimizing control variates for black-box gradient estimation[Z]. 2017.
- [68] JANG E, GU S, POOLE B J. Categorical reparameterization with gumbel-softmax[Z]. 2016.
- [69] MADDISON C J, MNIH A, THE Y W J. The concrete distribution: a continuous relaxation of discrete random variables[Z]. 2016.

[作者简介]



穆天杰（1998—），男，中国科学技术大学硕士生，主要研究方向为新型无线收发信机设计。

陈晓辉（1976—），男，中国科学技术大学副教授，主要研究方向为无线空口协议、多媒体通信等。

汪逸云（1997—），女，中国科学技术大学硕士生，主要研究方向为压缩感知、信道估计等。

马陆鹏（1994—），男，中国科学技术大学硕士生，主要研究方向为新型信源编码方法。

刘东（1983—），男，中国科学技术大学副教授，主要研究方向为图像、视频处理等。

周晶（1983—），男，博士，中国科学技术大学在读博士后，主要研究方向为信息论、无线通信等。

张文逸（1979—），男，中国科学技术大学教授，主要研究方向为信息论、无线通信等。