

面向未来的语义通信：基本原理与实现方法

张平¹, 牛凯¹, 姚圣时², 戴金晟²

(1. 北京邮电大学网络与交换技术国家重点实验室, 北京 100876;

2. 北京邮电大学泛网无线通信教育部重点实验室, 北京 100876)

摘 要: 阐述了语义通信的基本原理与实现方法。首先, 提出了语义通信的系统模型, 梳理语义通信的基本概念与术语, 对比经典通信与语义通信的主要区别, 总结了语义通信的技术优势。在此基础上, 回顾了语义信息论的发展, 构建了语义信息的度量指标体系, 提出了基于归一化条件复杂度的新型度量, 用于评估语义压缩的极限, 探讨了语义典型序列编码的性质, 揭示语义编码的渐近性能。进一步地, 归纳了语义编码传输的 2 种典型方法——直接编码与变换编码, 分析了它们的基本原理, 并针对文本、语音与图像信源展示了语义编码传输的优越性能。最后, 总结了语义通信的研究难点与问题, 指出未来的研究方向。

关键词: 语义通信; 语义信息论; 归一化条件复杂度; 语义编码传输; 非线性变换编码

中图分类号: TN911.22

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2023079

Semantic communications for future: basic principle and implementation methodology

ZHANG Ping¹, NIU Kai¹, YAO Shengshi², DAI Jincheng²

1. The State Key Laboratory of Networking and Switching Technology,

Beijing University of Posts and Telecommunications, Beijing 100876, China

2. The Key Laboratory of Universal Wireless Communications, Ministry of Education,

Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract: The basic principle of semantic communications and the associated implementation methodology were introduced. First, the systematic model of semantic communications was proposed and basic concepts and terminology were deduced by comparing the fundamental differences between classic communications and semantic communications. Thus, the technological advantages of semantic communications were concluded. Then, the development of semantic information theory was retrospected and the measurement system of semantic information was established. Furthermore, the normalized conditional complexity (NCC) was proposed to evaluate the limit of semantic compression and the property of semantic typical sequence coding was discussed so as to reveal the corresponding asymptotic performance. In addition, the semantic coding transmission methods were classified into two typical schemes, that is, direct coding and transform coding, and the basic principles of these schemes were presented. For the text, speech, and image sources, the superior performance of the semantic coding techniques were demonstrated. Finally, the difficulty and open issues of the semantic communications were further concluded and further research directions were pointed out.

Keywords: semantic communications, semantic information theory, normalized conditional complexity, semantic coded transmission, nonlinear transform coding

收稿日期: 2023-01-09; 修回日期: 2023-03-10

通信作者: 牛凯, niukai@bupt.edu.cn

基金项目: 国家自然科学基金资助项目 (No.62293481, No.92067202, No.62071058)

Foundation Item: The National Natural Science Foundation of China (No.62293481, No.92067202, No.62071058)

0 引言

信息与通信技术是现代社会的科技基础，信息产业深刻地改变了人类社会的形态。1948 年，Shannon^[1]发表了经典文献，为信息与通信奠定了理论基础。经典信息论指导下的通信技术已经日臻完善，例如，以 Huffman 编码、算法编码、矢量量化、变换编码为代表的信源编码技术已经逼近信源熵/率失真函数，以低密度奇偶校验 (LDPC) 码、极化码为代表的先进信道编码技术已经逼近信道容量。这些先进技术推动以互联网、5G 为代表的现代通信科技取得了突飞猛进的发展。但是，经典信息论在研究范畴、研究层次与研究维度方面仍然存在局限。

从认识论观点看，信息分为 3 个层次：语法、语义和语用。语法信息是最简单、最基本的层次。语义学的概念最初由 Morris^[2]提出，他在符号论中提出了符号的语法-语义-语用的三元概念。Shannon 和 Weaver^[3]意识到了语义的重要性，指出通信的语义问题以及有效性问题，引入了一般通信的 3 个级别，即 Level A~Level C。其中，Level A 级通信解决技术问题，即通信符号如何保证正确传输；Level B 级通信解决语义问题，即发送的符号如何传递确切的含义；Level C 级通信解决有效性问题，即接收的含义如何以期望的方式影响系统行为。长期以来，经典信息论局限在语法信息传输层次，即 Level A 级通信。但早在 1956 年 Brillouin^[4]就指出，经典信息论忽略信息的含义是因为通信工程的特殊需要，并不意味着人们要永远忽视语义信息。

多年来，人们一直在推进语义信息论的研究。Carnap 等^[5]和 Bar-Hillel^[6]提出用逻辑概率而非经典信息论中的统计概率度量语义信息量，得到语义信息熵的表达式。Barwise 与 Perry^[7]将语义信息论扩展到情景逻辑领域。Floridi^[8]提出了强语义信息理论，修正了语义矛盾的语句具有无穷大信息量的问题。语义信息的研究已超越了 Carnap 提出的框架。Kolchinsky 和 Wolpert^[9]将语义信息定义为描述系统与其所处环境之间关系的因果作用信息。Kountouris 与 Pappas^[10]对通信系统不同层次的语义信息给出了多粒度定义，并使用 Rényi 熵^[11]对语义信息进行度量。但这些工作仍然停留在语义信息的初步探讨，语义信息的定义与度量尚未形成统一观

点，语义通信的理论框架与实现方法还有待深入。

近年来，人工智能与算力技术的兴起为通信系统处理语义信息提供了技术底座，语义通信的理论与方法研究取得了进一步发展。本文作者张平院士深入分析语义信息特征，提出语义基 (Seb, semantic base) 模型^[12]，指出语义信息可以用高维空间的特征参量 Seb 进行表征。张平院士提出“智简 (Intellicise)”理念^[13]，进一步提出模型驱动的语义通信框架，实现通信系统由传统传输比特演进为传输“模型”，该模型即信源信道联合语义处理得到的新特征，如语义基等。牛凯等^[14-15]深入研究了从经典通信到语义通信的范式转变，提出了语义通信系统的基本框架。石光明等^[16]提出了语义通信的基本框架，探讨了语义信息表征与压缩编码问题^[17-18]。秦志金等^[19]探讨了深度学习赋能的语义通信理论、框架和系统模型，提出了任务驱动的多用户语义通信框架^[20]。

当前，信息与通信科学面临又一个重要关口，经典信息论指导下的通信技术已经制约了未来通信的发展。构建与完善语义信息论、建立语义通信方法体系，将极大地扩展现代通信的研究范畴。语义通信引入的新概念与新思想、提出的新理论与方法为通信系统的进一步优化提供了广阔空间，具有重要的科学意义与应用价值。本文从信息基础理论和原理入手，总结语义通信理论与方法的研究进展。

1 语义通信模型与理论

多年来，学术界一直在探索语义通信的基础理论。但与经典信息论相比，语义信息的定义与度量尚未达成一致，语义信息的理论框架也远未成熟。同时，在语义通信系统的建模问题上，研究者提出了语义通信的基础框架^[10,19-24]。这些框架各有特点，但目前学术界还未取得共识。

本节在现有工作基础上，提出语义通信的系统模型、概念与术语，构建语义信息的度量指标体系，对于语义通信的性能极限进行探讨分析。

1.1 系统模型

本文提出的语义通信系统模型如图 1 所示，参照 Shannon 与 Weaver^[3]的思想，分为 Level A 技术级通信与 Level B 语义级通信 2 个层级，由信源、语义知识库、语义编码器、语法编码器、信道、语义译码器、语法译码器、信宿等部分组成。

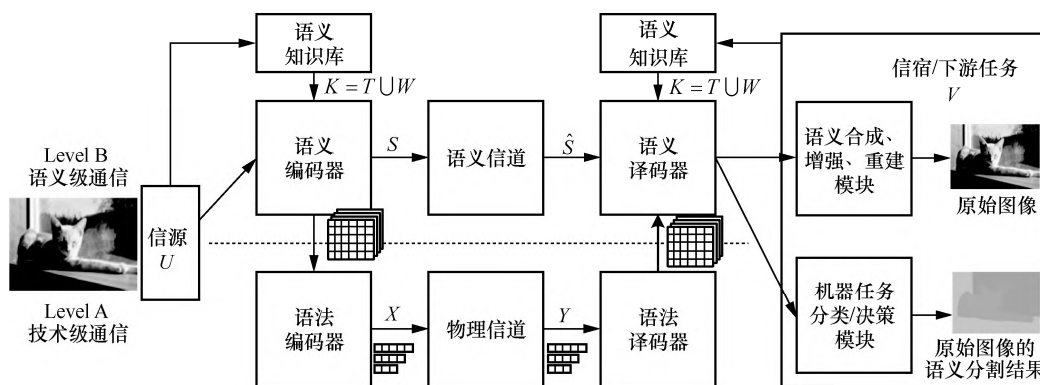


图1 语义通信系统模型

图1所示的系统模型中,技术级通信包括信源、语法编码器、物理信道、语法译码器、信宿5个模块,这是Shannon^[1]在经典文献中提出的点到点通信系统模型。语法编/译码器实际上就是经典通信模型中的编/译码器,主要处理语法信息。语义级通信是叠加在经典通信模型上的系统,引入了语义知识库、语义编码器以及语义译码器,并扩展了信道与信宿,各模块的功能简述如下。

1) 语义知识库。语义知识库是语义通信区别于经典通信而引入的重要模块。它从信源或信宿中提取语义背景知识,从信道传播环境中提取语义特征知识,背景/特征知识作为先验信息为语义编/译码提供辅助指导。背景/特征知识具有多种表示形式,如知识图谱^[25]、语义标签、下游任务相关知识、经过训练/优化的参数模型或非参数模型、信道模型及传播环境特征等。

2) 语义编码器。语义编码器在语义知识库的辅助下,提取信源消息的语义相关特征以及与传输任务有关的特征,而非概率信息。此外,语义编码器根据信源语义特性和信道特性,指导语法编码器对语义特征进行适当的编码来对抗传输中的干扰和噪声。因此,语义编码器既关注信源的语义特征,也关注信道的语义特征。也有学者把提取与任务有关特征定义为语义过滤或语义后处理^[10,19],指在全局特征中选择适配下游任务的语义特征,但语义特征提取和语义过滤的逻辑关系还有待进一步研究。

3) 语义译码器。根据信宿的传输需求,语义译码器选择重建信源消息,即面向人类感知,或者执行下游智能任务,即面向机器或其他意识体。在执行下游任务时,通常需要综合多种不同环境信息做出决策,如自动驾驶环境感知算法,因此需要在语义知识库的辅助下,根据不同模态信源的语义特征

进行多模态语义信息融合^[25-26],再通过决策模型做出决策。

4) 信道与信宿。信道与信宿包括物理信道与语义信道2个单元。语义信道是在物理信道上的抽象,对语义信息传输中的差错与损失进行建模。信宿中的下游任务对通信对象与通信目标进行了大幅度扩展。通信对象不限于人类用户,还可以是机器或物品,通信目标也不限于信号检测与估计,还可以是感知^[27]、基于知识图谱的推理^[28-29]、分类等各种任务。

在语义知识库辅助下,语义编码器首先采用深度神经网络等非线性处理模块提取图像的语义特征,获取图像的空间拓扑结构,表示高维语义特征空间的嵌入。然后,语法编码器对语义特征编码,产生物理信道中的传输符号。当然,也可以将语义编码器和语法编码器合并为一个联合处理模块。在接收侧,语法译码器恢复语义特征空间中的语义特征,进而在面向人的通信场景中,语义译码器利用重建的语义特征进行合成、增强,最终生成原始图像,以提高人的感知质量,而在面向下游任务场景中,信宿直接利用重建语义特征执行下游任务,图1右侧展示了原始图像的语义分割结果。

与经典通信系统相比,语义通信系统含有语义知识库、语义编码器以及语义译码器3个重要模块,这是两者的关键区别。语义通信可以理解为附着在经典通信之上的高层系统,既依赖但又高于经典通信系统。语义知识库为收发两端提供语义信息处理的指导,发射机和接收机共享知识库使语义通信成为可能。语义知识库不仅“感知”到信源的语义特征,还与具体的传输任务和传输条件有关,这使在“感知”信道状态信息的情况下能够实现非平衡传输,将更多资源分配到更重要的语义特征传输上,

同时,一些语义层级的轻微错误很容易被已有知识识别和纠正^[30-31],从而便于人或机器在语义理解上屏蔽这些错误。

1.2 语义通信系统概念与术语

经典通信系统和语义通信系统的概念与术语对比如表 1 所示。

下面简要分析经典通信系统与语义通信系统的区别。

1) 作用域不同

经典通信系统的作用域是信号空间,而语义通信系统的作用域为语义空间。语义空间中的语义特征描述了信号的语义信息,即消息的含义。信源消息的语义信息和信源的数据组织方式存在一对多的映射关系。例如,世界各地的语言各不相同,但都遵循规则结构(音韵、形态、句法、语义和语用)来保持内在秩序,因此语义指的是交流过程中的单词和句子的实际含义。

2) 处理对象和先验知识不同

经典通信系统的处理对象是比特流,比特本身并不具有含义;而语义通信系统的处理对象是语义对象,可以用语义基等语义特征表征,其本身具有含义。语义通信系统需要以知识库作为先验知识,通过收发两端共享的先验语义知识库,辅助实现语义编码和解码。

3) 处理方法不同

经典通信系统将信源压缩和信号传输分离,而语义通信系统面向下游任务及信道环境,采用语义特征提取、语义内容识别、多模态语义融合等技术提取并编码语义对象,从而提升端到端的通信质量。

如图 1 所示,语义通信的层级(Level-B)高于经典通信的层级(Level-A),但又附着于经典通信。语义信息的提取、语义特征的编码及传输存在协同作用,而经典通信系统中的多载波传输、多址接入与多天线等技术,可有机融合到语义通信系统中,提高系统整体性能。

4) 性能准则和优化目标不同

经典通信系统关注的是比特信息传输的可靠性,具体性能准则有最小均方误差等,优化目标有误比特率、误块率等,而信源压缩问题是另一独立问题,经典通信系统这一分离式设计和优化使系统评价方式和优化方式单一。而语义通信系统的准则和评价方式与信源的模态、传输的目的或任务有关,特点是关注端到端语义误差,在面向人的感知的语义通信场景中,性能准则和评价方式为优化人的感知体验,包括视觉感知质量指标 LPIPS (learned perceptual image patch similarity)^[32]、听觉感知质量指标 PESQ (perceptual evaluation of speech quality)^[33]等;在面向下游机器任务的语义通信场景中,性能准则转变为给定传输带宽下语义传输所达到的效用分数,优化目标为速率-效用函数。

综合以上分析,相比经典通信系统,语义通信系统具有三方面性能优势。

1) 深度挖掘——信息处理层次

语义通信突破了传统通信对于信息处理的层次限制,借助神经网络的非线性信息处理能力,语义信息的提取和处理成为可能。挖掘更深层次的信息含义,甚至信息的用途,将进一步提升未来通信系统的信息传输效率。

2) 广度扩展——通信应用范畴

语义通信极大地扩展了通信应用的范畴。经典通信系统,特别是编码传输系统,主要追求比特级信息的准确传输,研究范畴相对单一。而语义通信更关注接收端获取发送端信息含义的准确性,即满足“达意通信”的目的。由此,语义传输的可靠性取决于人的主观感受或者下游任务的各种性能指标,能灵活适应不同应用场景及传输条件。

3) 方法创新——系统端到端优化

语义通信系统的设计采用性能指标的端到端

表 1

经典通信系统和语义通信系统概念与术语对比

通信系统	作用域	处理对象	先验知识	处理方法	性能准则	优化目标
语义通信系统	语义空间	语义基等语义特征	知识库	语义特征提取、语义内容识别、多模态语义融合、语义纠错、语义非平衡传输	端到端语义误差,包括感知质量指标、任务执行度(效用分数)等	速率-效用函数
经典通信系统	信号空间	比特流	无	信源编码、信道编码、链路自适应技术、多载波传输技术、多址接入、多天线技术	最小均方误差(MMSE)等	错误率 (误比特率、误块率)

优化方法, 是系统整体优化观点的体现。经典通信系统采用分离式设计和优化, 已逐渐出现性能瓶颈, 通信系统的性能难以进一步提升。而语义通信系统采用信源语义信息提取以及语义信息传输的联合优化设计, 为通信系统的整体最优设计提供了新的解决思路与实现方法。

1.3 语义信息论简述

多年来, 语义信息度量与语义通信极限的研究一直持续进行。Bao 等^[34]考虑了收发两端存在本地知识库, 根据逻辑蕴含关系得到给定背景知识下的语义信息熵表达式, 并得到语义失真理论度量及语义信道容量。Luca 与 Termini^[35-36]基于隶属度函数得到了模糊集合信息熵的表达式以度量模糊事件的信息量。Körner^[37]将随机变量表征为概率图, 定义了图熵, 用以度量信源符号序列的最优编码^[38]。基于图熵, Alon 和 Orlitsky^[39]探索了接收端存在边信息时的信源编码极限。

吴伟陵教授^[40]提出了广义信息熵理论, 将香农信息熵的概念移植到模糊集合中, 证明了广义熵同样具有香农信息熵的 4 种重要性质, 为语义信息表征和语义通信奠定了早期理论基础。钟义信教授^[41-43]在语义信息研究方面的代表性著作中阐明了信息科学的概念、原理和方法。鲁晨光^[44]讨论了广义熵与广义互信息的编码意义。Fan 等^[45]和 She 等^[46]定义了信息重要性测度表征随机过程的不确定性。这些工作从多个角度、多个层次探讨了语义信息的内涵, 富有启发意义, 但现有语义通信的理论框架还不够完善与统一, 还需要进一步深入研究。

本节总结了语义信息的度量指标体系, 如表 2 所示, 包括语义熵、语义互信息、平均语义失真、语义信道容量、语义率失真函数。

表 2 语义信息的度量指标体系

指标	含义	表达式
语义熵	平均语义信息量	$\tilde{H}(S T)$
语义互信息	一个对象包含关于另一个对象的语义信息量	$\tilde{I}(S; \hat{S} T, W)$
平均语义失真	语义通信导致的语义信息损失	$\mathbb{E}[d_s(s, \hat{s})]$
语义信道容量	特定语义失真下的最大传输速率	$C_s(\mathcal{E}_s)$
语义率失真函数	特定语义失真下的最小编码速率函数	$R_s(D_s)$

参考图 1 的语义通信系统模型, 信源集合为 U , 语义信息集合为 S , 语义知识库为 $K = T \cup W$, 其中, T 是信源语义集合, W 是信道语义集合。语法信息集合为 X , 经过物理信道接收到的语法信息序列集合为 Y , 等价的接收端语义信息集合为 \hat{S} , 最终重建的信宿消息集合为 V 。

给定信源语义集合 $T \subset U$, 对于任意符号 $t \in T$, 其概率分布为 $p(t)$, 由该符号提取生成的语义信息为 $s \in S$, 映射函数为 $\varphi(s|t)$, 则集合 S 上的语义熵表示为

$$\tilde{H}(S|T) = -\sum_{t \in T} \sum_{s \in S} \varphi(s|t) p(t) \log \varphi(s|t) \quad (1)$$

一般地, 映射函数包含了信源符号到语义空间的各种复杂映射, 可能是一对一映射, 也可能是多对多的模糊/主观映射, 难以用客观的概率映射描述, 往往需要考虑主观模糊性, 可以用隶属度函数描述。如果 $|S| \leq |U|$, 映射函数能够客观反映信源语法信息到语义信息的映射, 则语义熵小于信息熵, 即满足

$$\tilde{H}(S|T) \leq H(U) \quad (2)$$

当然, 如果 $|S| > |U|$, 且映射函数存在很大偏差, 则语义熵也可能大于信息熵。

类似地, 给定信源语义集合 T 与信道语义集合 W , 定义语义互信息 $\tilde{I}(S; \hat{S}|T, W)$ 为

$$\begin{aligned} \tilde{I}(S; \hat{S}|T, W) = & -\sum_{w \in W} \sum_{s \in S} \sum_{t \in T} \sum_{s \in S} p(t) p(w) \cdot \\ & \varphi(s|t) p(\hat{s}|s, t, w) \log \frac{p(\hat{s}|t, w)}{p(\hat{s}|s, t, w)} \end{aligned} \quad (3)$$

语义互信息表征了给定知识库条件下, 接收方获得的语义信息量。

借助语义相似度或语义距离度量, 定义平均语义失真为

$$\begin{aligned} \mathbb{E}[d_s(s, \hat{s})] = & \sum_{t \in T} \sum_{w \in W} \sum_{s \in S} \sum_{s \in S} p(t) p(w) \varphi(s|t) p(\hat{s}|s, t, w) d_s(s, \hat{s}) \end{aligned} \quad (4)$$

语义信道容量可以定义为实现语义失真不超过 \mathcal{E}_s 的最大传输速率, 即

$$C_s(\mathcal{E}_s) = \sup_{\varphi(s|t) \in P_{\mathcal{E}_s}} \tilde{I}(S; \hat{S}|T, W) \quad (5)$$

其中, $P_{\varepsilon_s} = \{\varphi(s|t) | \mathbb{E}[d_s(s, \hat{s})] \leq \varepsilon_s\}$ 为平均语义失真不超过 ε_s 的语义信源集合。

而语义通信模型中的语义率失真函数定义为平均语义失真不超过 D_s 条件下的最小编码速率, 即

$$R_s(D_s) = \inf_{p(\hat{s}|s, t, w) \in P_{D_s}} \tilde{I}(S; \hat{S} | T, W) \quad (6)$$

其中, $P_{D_s} = \{p(\hat{s}|s, t, w) | \mathbb{E}[d_s(s, \hat{s})] \leq D_s\}$ 为平均语义失真不超过 D_s 的语义信道集合。

借鉴算法信息论 (AIT, algorithm information theory) [47], 本文对语义编码的压缩极限进行分析, 进一步借鉴经典信息论的渐近分析方法, 对语义编码的典型序列进行讨论。

1.3.1 语义压缩极限分析

如前文所述, 基于概率模型的香农信息熵并不适用于量化语义信息量, 而适用于衡量消息概率集合中的不确定性, 但语义信息不能仅用概率行为来刻画, 它还存在其他不确定度, 如模糊性。

复杂的结构性数据 (如音频、图像、视频) 不能假设为伯努利信源或者稳态遍历信源。语义传输的极限应定义为某个特殊的信源消息集合, 而不是反映总体消息集合的平均性能指标。信息计算、操作和传输的范式越来越多地从面向随机变量转变为面向个体对象。信源消息的语义信息量, 即语义编码传输需要的最少资源, 可以由该消息个体的语义信息量来度量。

所谓语义, 从自然语言处理的角度解释, 即一段文本中所包含的信息, 它与文本字符出现的概率无关, 而与这个字符串表达的含义有关。一般意义上, 语义信息可以用通用图灵机建模, 而最简表述可以衡量该字符串的语义信息量下限。

柯尔莫哥洛夫复杂性^[47]是算法信息论中的一个关键指标, 它建立在通用图灵机的理念上, 用于描述计算机程序或算法序列的最短长度, 可以刻画单个有限长序列的描述复杂度。图灵证明了一切消息或算法程序都可以在通用图灵机上进行合理计算。一个字符串为 s 的柯尔莫哥洛夫复杂性 $C(s)$ 是这个字符串的最短描述长度。

如图 2 所示, 二进制序列 11.001001000011... 看起来是一个随机且无意义的序列, 但它代表数学常数 π 的二进制表示。基于圆周率的先验知识, 可以采用简单且长度有限的高斯-勒让德算法表示这

个序列。因此, 表示 π 的二进制序列的描述复杂度就由高斯-勒让德算法的复杂度来量化。而对于无限长随机序列, 柯尔莫哥洛夫复杂性的期望趋近于香农熵。但遗憾的是, 因为通用图灵机的理想性, 找到一个算法的最短描述长度在理论上不可实现。因此, 柯尔莫哥洛夫复杂性具有不可计算性, 一般只能通过近似计算进行逼近。

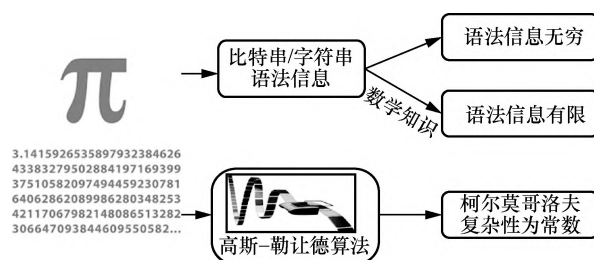


图 2 柯尔莫哥洛夫复杂性示意

牛凯等^[14]探索了有限知识库条件下的语义压缩极限, 提出了归一化条件复杂度 (NCC, normalized conditional complexity) 作为柯尔莫哥洛夫复杂性的近似, 用于对语义压缩极限的估计。NCC 定义为在有/无知识库的情况下特定数据集各个消息编码的额外所需的复杂度, 由联合复杂度和知识库本身的复杂度给出, 并用信源消息序列长度归一化, 其表示形式为

$$NCC(S|T) = \mathbb{E}_{s \in S} \left[\frac{C(s, T) - C(T)}{l(s)} \right] \quad (7)$$

其中, T 为知识库即信源语义集合, S 为语义信息集合, C 为柯尔莫哥洛夫复杂性, $l(s)$ 为信源消息序列的长度。如需计算单个数据集的归一化复杂度 (即知识库由自身建立), 则需对不同的数据集划分方式进行枚举并求期望。

由此, NCC 可以衡量在特定先验语义知识的情况下对某个信源消息集合进行语义编码/压缩的极限, 是语义压缩编码下界的估计。

不同文本信源上语义压缩率的评估结果如图 3 所示, 数据集分别取自小说集、法律文献等英文文本数据集, 结果均为平均每单词编码比特数。由图 3 可见, 语义编码的压缩性能介于香农信息熵与 NCC 界之间。NCC 界显示了文本信息有进一步压缩的潜力, 但它只是对语义压缩极限的初步探索, 未来还需要从多个角度继续深入研究语义编码极限。

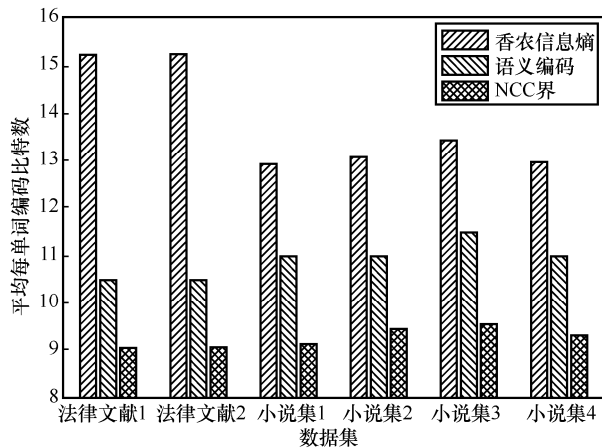


图3 不同文本信源的语义压缩率的评估结果

1.3.2 典型序列渐近分析

类比经典信息论的典型序列分析方法^[1], 语义信息论也可以采用类似方法进行渐近分析。依据渐近等分割 (AEP) 特性, 当序列长度 n 足够大时, 一个独立同分布观察序列的概率近似等于 2^{-nH} , H 为单符号熵。由此将所有编码序列构成的集合划分成 2 个子集, 其中一个子集为典型集合, 其样本为典型序列, 样本熵接近于香农信息熵; 另一个子集为非典型集合, 包含几乎不可能出现的序列。对于任意 $\delta > 0$, 只要序列长度 n 足够大, 则典型序列出现的概率满足

$$\left| \frac{1}{n} \log p(X^n) - H(X) \right| < \delta \quad (8)$$

经典编码传输和语义编码传输的典型序列对比如图 4 所示。对于经典编码传输, 每个编码长度为 n 的码字 X^n 都对应于信源信号空间长度为 m 的典型序列 U^m 。经过信道传输后, 总的接收序列 Y^n 包含约 $2^{nH(Y)}$ 个序列。每个典型序列 X^n 对应接收序列 Y^n 构成的一个子集, 这个子集中的序列和 X^n 构成了联合典型序列。子集中包含约 $2^{nH(Y|X)}$ 个等概率序列, 其中 $H(Y|X)$ 为条件熵。为了区分不同的典型序列 X^n , 需要将它们映射到互不相交的子集中, 因此子集数量不超过 $2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}$ 个, 这意味着最多能发送 $2^{nI(X;Y)}$ 个长度为 n 的序列。

而对于语义编码传输, 每个压缩后的语义编码序列 S^n 和信源序列 U^m 的子集对应, 它们构成了联合典型序列。这个子集中的信源序列具有相同的语义含义, 经过压缩编码的码字是相同的。从经典信源编码的角度出发, 这种合并的过程显

然会导致有损的数据压缩。然而, 因为语义通信系统中背景知识的存在, 这个合并过程在语义层次上是“无损”的, 语义信息能够无差错传输, 或者完美执行下游任务。基于语义合并操作, 语义空间中的典型序列数量将显著少于信源信号空间中的序列数量。

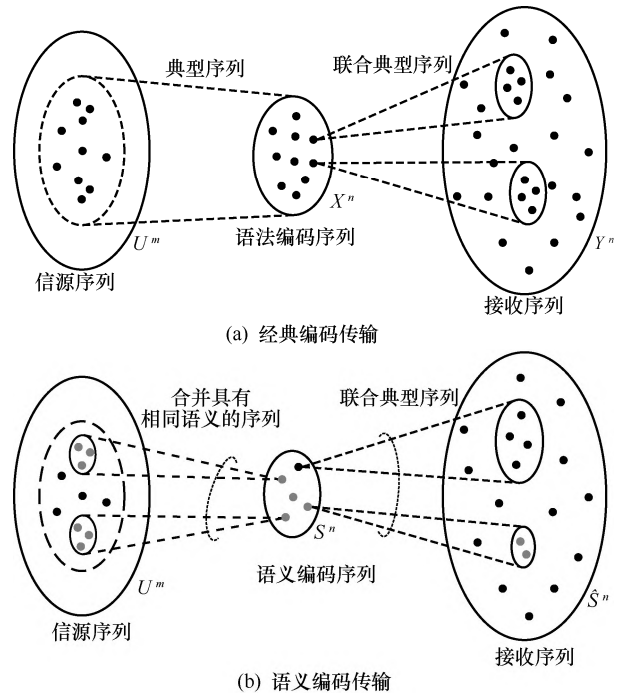


图4 经典编码传输和语义编码传输的典型序列对比

在语义通信中, 语义空间中表征语义特征的码字具有不等重要性, 这个重要性不仅取决于信源本身, 也取决于语义传输的目的与任务, 对应不同的语义知识库。因为这种不等重要性, 它们对应的接收序列子集大小也是不等的。对于一个更重要的语义特征 S^n , 它对应的接收序列 \hat{S}^n 数量相对较少, 也就是说不确定性更小。由此可见, 引入语义知识库和采用语义编码, 在渐近意义上能够提高信息传输的效率。

2 语义编码传输方法

与经典通信相比, 语义通信更关注语义信息的传输误差, 而不再强调比特传输误差。总结现有的语义编码方法, 大致可以归纳为两类, 第一类是直接编码方法, 第二类是变换编码方法。

直接编码方法的基本思路是采用神经网络, 直接从信源序列中提取语义特征并映射为信道中的传输符号, 属于信源信道联合编码。Bourtsoulatz

等^[48]针对图像和视频信源,利用深度神经网络实现了端到端信源信道联合编码方案 DeepJSCC,端到端图像质量随信噪比的降低而平缓降低,不存在悬崖效应,传输性能表现优越。Tung 等^[49]加入调制模块,提出了基于数字通信的模型 DeepJSCC-Q,利用固定的调制星座实现与 DeepJSCC 类似的性能。Xie 等^[50]提出了一种语义通信方案 DeepSC,针对文本信源使用神经网络提取语义特征,采用信源信道联合编/解码技术在接收端恢复文本语义信息。类似地,Weng 和 Qin^[51]针对语音信源构建语义通信方案 DeepSC-S,提高了语音信号的重建质量。Xie 和 Qin^[52]还针对物联网设计了精简分布式语义通信方案 L-DeepSC,对物联网云端设备的文本数据进行低复杂度语义压缩与传输。

变换编码方法的基本思路是通过非线性变换将信源数据映射到隐空间,反映信源数据的语义信息,在隐空间中更容易实现高效的语义表征,然后进行信源信道联合编码。Ballé 等^[53]提出了非线性变换编码方案,并应用于端到端图像压缩^[54]、语音波形编码^[55],具有显著的性能优势。在此基础上,Dai 等^[56-57]提出非线性变换信源信道联合编码(NTSCC)框架,建立了变换语义编码传输的基本优化准则,并应用于无线图像的语义传输^[58-59]、语音语义编码和传输^[60]、工业互联网场景数据处理^[61],与传统编码方案相比,NTSCC 能够获得显著的性能增益。

2.1 直接编码方法

直接编码方法的代表是 DeepJSCC 方案^[48]以及 DeepSC 方案^[50],它们的基本思想是采用神经网络,实现从信源数据到信道中传输符号的映射,且该编码过程与信道条件相关,下面简要介绍其基本原理。

2.1.1 编码方案设计与分析

直接编码语义通信模型如图 5 所示,信源向量 $\mathbf{u} \in \mathbb{R}^m$ 经过编码神经网络 f_e 映射,得到语义编码向量 $\mathbf{s} \in \mathbb{R}^k$,即满足编码约束 $\mathbf{s} = f_e(\mathbf{u}; \phi_f)$ 。一般地,编码器采用卷积神经网络(CNN), ϕ_f 是模型参数。信道的转移概率为 $P(\mathbf{y}|\mathbf{s})$,对于加性白高斯噪声(AWGN)信道,接收信号模型为

$$\mathbf{y} = \mathbf{s} + \mathbf{n} \quad (9)$$

其中, $\mathbf{n} \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I}_k)$ 是独立同分布的高斯噪声向量, σ_n^2 是噪声方差, \mathbf{I}_k 是单位矩阵。

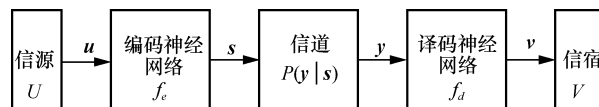


图 5 直接编码语义通信模型

接收端经过联合译码网络处理,得到输出向量 $\mathbf{v} = f_d(\mathbf{y}; \theta_f)$ 。为了实现端到端联合优化,编码与译码神经网络模型需要联合学习,即满足式(10)所示优化模型。

$$(\phi_f^*, \theta_f^*) = \arg \min_{\phi_f, \theta_f} \mathbb{E}_{\mathbf{u} \sim p(\mathbf{u})} \mathbb{E}_{\mathbf{v} \sim p(\mathbf{v}|\mathbf{u})} [d_s(\mathbf{u}, \mathbf{v})] \quad (10)$$

其中, $d_s(\mathbf{u}, \mathbf{v})$ 是信源与信宿符号之间的语义失真。

DeepJSCC 直接优化端到端的语义失真,但由于语义特征提取不明确,语义信息难以度量,因此只能通过提高模型复杂度来提升性能,难以获得编码增益,无法实现高效语义通信。

2.1.2 性能评估

对于图像信源,采用 DeepJSCC 方案^[48]在 AWGN 信道 SNR=10 dB 条件下进行语义编码传输,峰值信噪比(PSNR)性能如图 6 所示。由图 6 可知,在相同信噪比条件下,DeepJSCC 方案在低带宽比下的 PSNR 性能优于基于 JPEG、JPEG2000 编码的传统方案。在相同信噪比与 PSNR 条件下,DeepJSCC 方案的频谱利用率显著优于传统方案,并且没有悬崖效应。

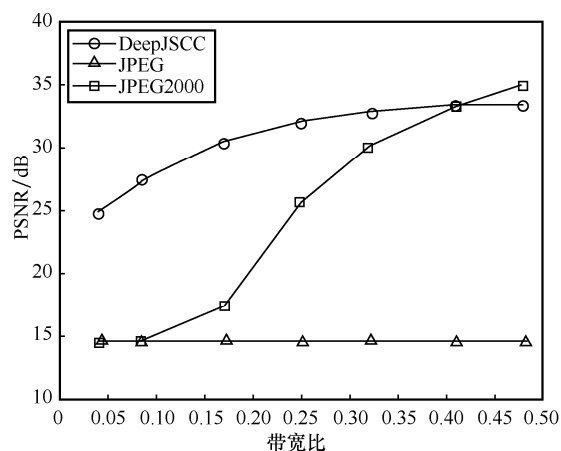


图 6 图像信源 DeepJSCC 方案的 PSNR 性能

对于文本信源,DeepSC^[50]使用 Transformer 神经网络结构^[62]提取文本的语义特征,输出与上下文有关的文本表征向量。AWGN 信道 SNR=12 dB 条件下,不同编码方案语句相似度性能对比如表 3 所示。由表 3 所知,在低信噪比条件下,DeepSC 方案的语句相似度优于 Huffman 编码级联 RS (Reed-Solomon) 码、Turbo 码等信道编码的传统编码方案。

表3 文本信源不同编码方案语句相似度性能对比

编码方案	文本语句相似度
Huffman 编码级联 RS 码	0.03
Huffman 编码级联 Turbo 码	0.82
DeepSC 方案	0.96

这一方案可以扩展到语音信源^[51]，在卷积神经网络中引入“伸展-挤压”模块，用于语音信源的语义传输。在 AWGN 信道 SNR=9 dB 条件下，不同方案的 PESQ 性能对比如表 4 所示。从表 4 可知，直接语义编码传输方案的 PESQ 性能优于 PCM 编码级联 Turbo 码的传统编码方案。但直接语义编码传输方案传输带宽较高。

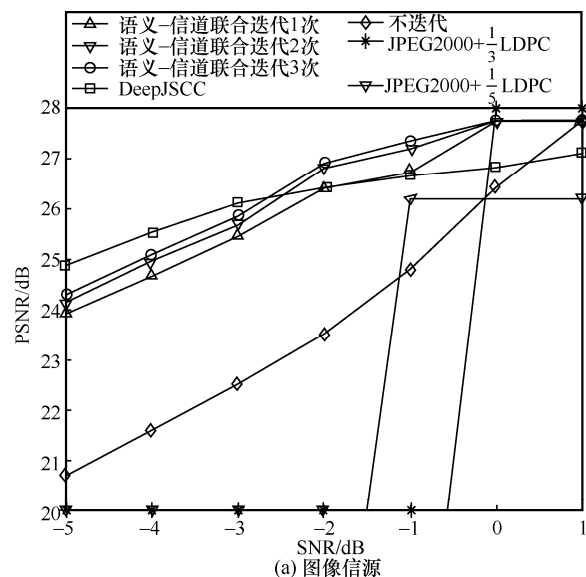
表4 语音信源不同方案 PESQ 性能对比

编码方案	PESQ
PCM 编码级联 Turbo 码	2.04
直接语义编码传输方案	4.3

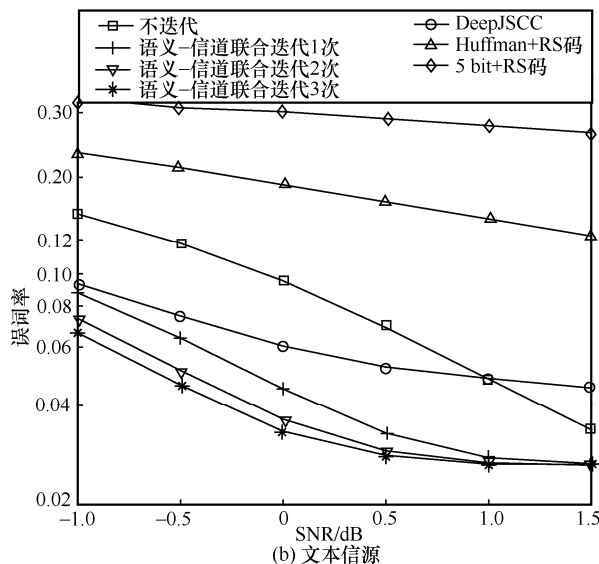
在此基础上，本文将语义编码与经典通信系统中的信道编码与调制级联，并设计了语义-信道联合迭代算法^[63-64]，每次迭代语义译码器为信道译码器提供额外的先验信息，提升整体性能。图 7 给出了 AWGN 信道下图像和文本信源采用语义-信道联合迭代接收方案的 PSNR-信噪比以及误词率-信噪比性能曲线（误词率采用对数坐标），并与传统方案进行对比。其中，JPEG2000+ $\frac{1}{3}$ LDPC 和 JPEG2000+ $\frac{1}{5}$ LDPC 分别

表示 JPEG2000 级联码率分别为 $\frac{1}{3}$ 和 $\frac{1}{5}$ 的 LDPC 码，Huffman+RS 码表示 Huffman 编码级联 RS 码，5 bit+RS 码表示 5 bit 定长编码级联 RS 码。

由图 7 可知，通过 1~2 次语义-信道联合迭代译码，对于图像信源，在低信噪比的条件下，PSNR 指标显著优于传统的 JPEG2000 级联 LDPC 码方案与深度学习 DeepJSCC 方案；对于文本信源，误词率 WER 远低于传统的编码方案（Huffman 编码级联 RS 码、5 bit 定长编码级联 RS 码）与 DeepJSCC 方案语义。辅助的迭代接收方案实现了语义编码和经典通信技术的松耦合，能够更好地与数字通信系统兼容，并且具有显著的编码性能增益。



(a) 图像信源



(b) 文本信源

图7 语义-信道联合迭代接收方案性能

2.2 变换编码方法

直接编码方法将信源序列作为整体进行压缩编码，难以区分信源数据序列上不同部分的语义差异，也即无法从细粒度上刻画语义信息的多少，无法实现高效语义编码。为了进一步提高传输效率，本文应用了非线性变换编码的思想^[53-54]，设计了基于变换域的语义编码方案 NTSCC^[57]，大幅度提高了语义传输的效率与可靠性，下面简要介绍 NTSCC 的基本原理。

2.2.1 编码方案设计与分析

变换编码语义通信模型如图 8 所示，其由非线性变换与信源信道联合编译码组成，其基本思想是信源向量不直接映射到语义编码序列，而是首先进行非线性解析变换，提取隐向量表示，然后在隐空间中进行信源信道联合编码。

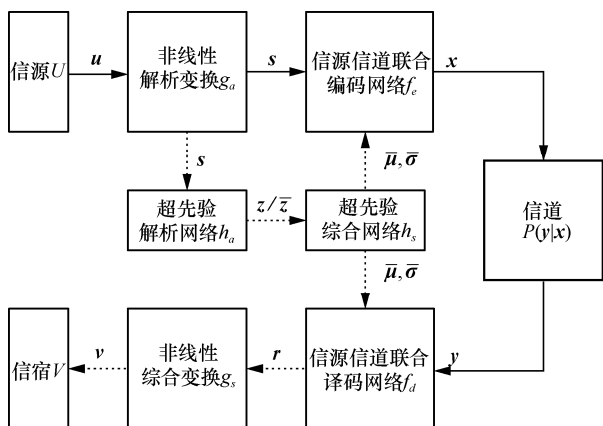


图8 变换编码语义通信模型

如图8所示,信源向量 \mathbf{u} 经过非线性解析变换,得到隐空间语义表征向量 \mathbf{s} ,即 $\mathbf{s} = \mathbf{g}_u(\mathbf{u}; \phi_g)$,其中 ϕ_g 是模型参数。在隐空间中,可以更好地对语义信息进行表征和编码。具体而言,引入超先验解析/综合网络,刻画语义表征向量在时间/空间维度的相关性,通过变分建模用简单的高斯分布均值向量 μ 与标准差向量 σ ,建模编码和译码概率分布所需参数。携带语义信息的隐向量 \mathbf{s} 送入信源信道联合编码网络 f_e ,得到编码向量 \mathbf{x} ,即 $\mathbf{x} = f_e(\mathbf{s}; \phi_f)$ 。经过量化与压缩送入超先验综合网络 h_s ,产生量化版本的均值与标准差向量 $(\bar{\mu}, \bar{\sigma}) = h_s(\bar{\mathbf{z}}; \theta_h)$ 可作为额外的边信息对恢复的隐空间向量进行优化,提高重建可靠性,得到 $\mathbf{r} = f_d(\mathbf{y}; \theta_d)$ 。最后,送入非线性综合变换,得到重建信源向量 $\mathbf{v} = \mathbf{g}_s(\mathbf{r}; \theta_g)$ 。

与直接编码不同,变换编码方法引入端到端的语义率失真优化模型,优化损失函数为

$$L = \mathbb{E}_{\mathbf{u} \sim p(\mathbf{u})} \left\{ \lambda \left[\underbrace{-\eta \log P(\bar{\mathbf{s}}|\bar{\mathbf{z}})}_{k_s} - \underbrace{\frac{\log P(\bar{\mathbf{z}})}{C_z}}_{k_z} \right] + d_s(\mathbf{u}, \mathbf{v}) \right\} \quad (11)$$

其中, $\mathbf{s} = \mathbf{g}_u(\mathbf{u}; \phi_g)$, $\mathbf{z} = h_u(\mathbf{s}; \phi_h)$, $\bar{\mathbf{s}} = Q(\mathbf{s})$ 与 $\bar{\mathbf{z}} = Q(\mathbf{z})$ 表示均匀量化, $\mathbf{v} = \mathbf{g}_s(\mathbf{r}; \theta_g)$, k_s 是语义隐向量 \mathbf{s} 的维度, k_z 是边信息 \mathbf{z} 的维度, C_z 是相应的信道容量,参数 η 控制语义信息编码速率,拉格朗日因子 λ 控制 $k_s + k_z$ 与端到端的语义失真 d_s 之间的折中关系。

变换编码方案根据语义信息条件熵 $\log P(\bar{\mathbf{s}}|\bar{\mathbf{z}})$ 精细控制语义编码的速率,能够更准确地调节信源信道联合编码速率,高效压缩信源冗余,适应信道

变化。因此,相比直接编码,变换编码在语义传输性能与压缩比之间能够达到更好的折中。

2.2.2 性能评估

对于图像信源,使用Swin transformer网络^[65]作为非线性解析和综合变换网络的基础框架,图9展示了基于变换域的语义编码方案在不同图像数据集的端到端率失真性能,图9(a)展示了PSNR-带宽比性能,图9(b)采用了主观感知质量评价指标LPIPS^[32](LPIPS值越低,图像质量越高)。与传统方案BPG信源编码+LDPC码以及直接编码DeepJSCC方案对比,NTSCC方案在相同传输带宽比条件下,主观感知质量更好,接收端恢复图像的质量越高,且在主观感知质量评价指标下使用感知质量优化方法相比均方误差损失函数能获得更好的主观感受。

对于语音信源,首先使用卷积神经网络将分帧后的语音波形信号从时域映射到隐空间;然后,利用卷积神经网络提取隐向量之间的时间相关性;最后,对帧之间的相关性通过Transformer网络建模,从而对不同帧根据内容复杂度差异实现不等长度编码。

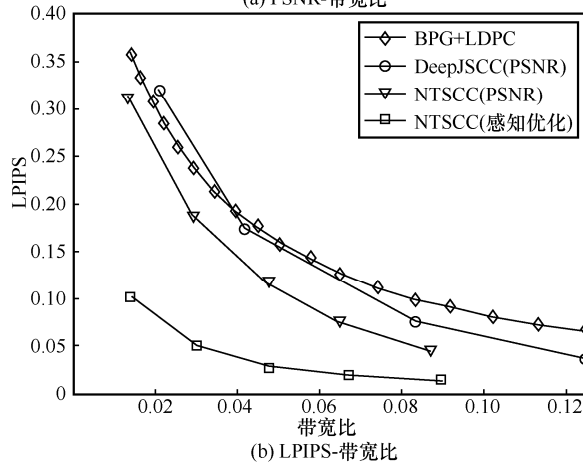
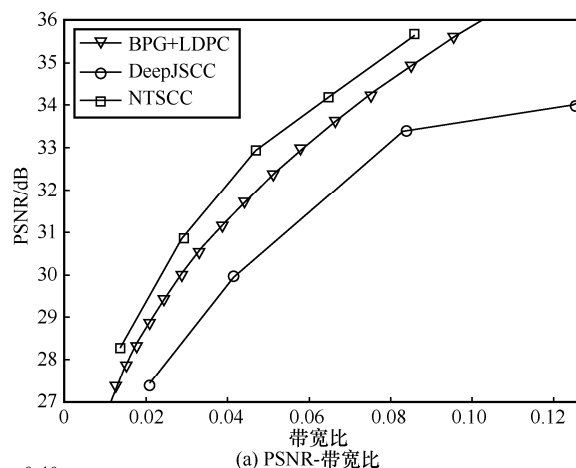


图9 基于变换编码域的语义编码方案端到端率失真性能

宽带语音信源在 AWGN 信道和 COST2100 衰落信道^[66]下的客观感知质量如图 10 所示,采用 PESQ^[33]作为客观感知质量评价指标。图 10(a)中 SNR=6 dB,图 10(b)中传输带宽 $K=10$ kHz。在相同感知质量指标下,NTSCC 相较于传统语音编码自适应多速率宽带 (AMR-WB, adaptive multi-rate wideband) 语音编码器^[67]、Opus 编码器^[68]级联 5G 信道编码,以及直接语义编码方案 DeepSC-S,所需带宽更少,且在衰落信道中也具备出色的稳健性,优于对比方案。

宽带语音信源在 AWGN 信道和 COST2100 衰落信道下的主观感知质量如图 11 所示,本文采用评估有损音频的主观评价方法 MUSHRA 测试^[69],MUSHRA 得分表示参与对比的若干编码方案的相对得分,得分越高,生成的音频质量更高。图 11(a)中 SNR=6 dB;图 11(b)中 DeepSC-S 方案传输带宽

为 12 kHz,其他方案传输带宽为 10 kHz,SNR=2 dB。与传统方案和直接编码方案相比,NTSCC 方案得益于语音帧内语义信息量和帧间相关性建模,既能够保持较高的主观感知质量,又显著节省了信号传输带宽,体现出双重性能优势。

3 未来研究方向

作为一个前沿研究方向,语义通信的理论与方法还远未成熟,需要进一步探索。本文列出了未来语义通信的一些研究难点与问题,供研究者参考。

3.1 语义信息的度量分析

经典信息论对于语法信息的度量建立在概率论基础上,具有客观定量的特点。但语义信息天然具有主观性与模糊性,无法用概率进行准确度量,需要引入新的数学工具。并且,语义信息具有复杂

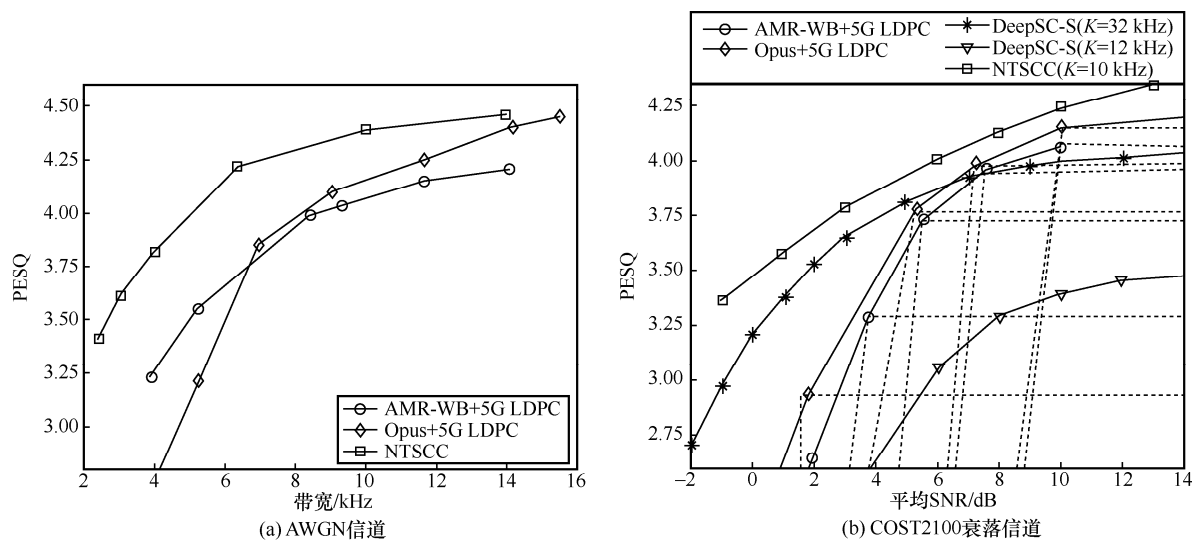


图 10 宽带语音信源在不同信道下的客观感知质量

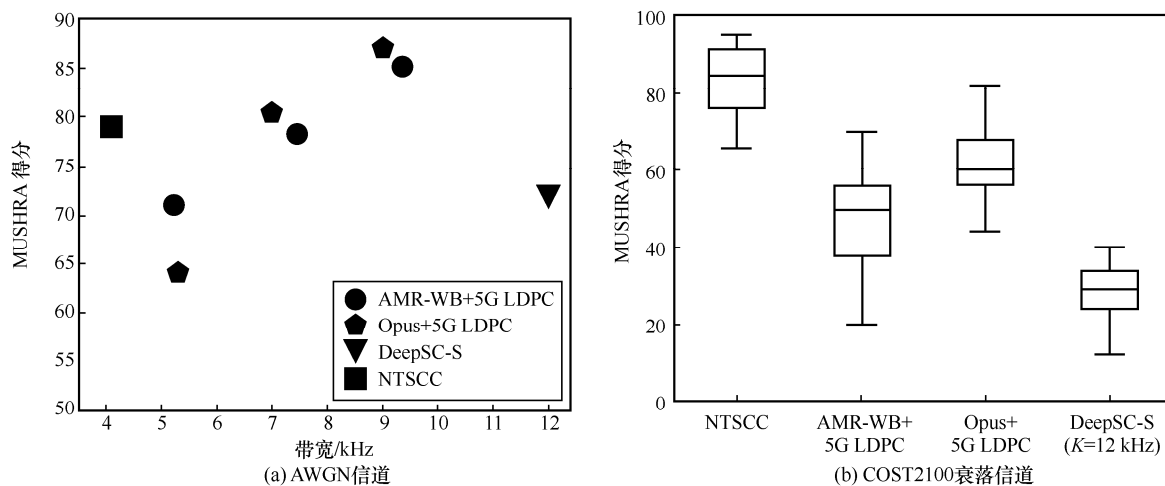


图 11 宽带语音信源在不同信道下的主观感知质量

的层次关系,无法用简单的统计模型表征,需要引入全新的分析方法。尽管神经网络模型采用数据驱动的方法,可以拟合一部分语义信息的复杂特征,但基于贝叶斯概率推断的框架并不能完全表征语义信息的本质特征。特别是在多模态语义处理、语法与语义信息混合处理等复杂场景下如何定量分析语义信息,还需要进一步完善理论框架。

3.2 语义信息论极限分析

虽然语义信息论很早就被提出,但研究工作比较零散。多年来,经典信息论的光芒完全掩盖了语义信息论,由于理论研究远滞后于实际需求,语义信息论未能形成完整的理论体系。特别是在语义压缩与传输极限方面,虽然有初步的探索,但还没有形成业界公认的理论成果。与经典信息论类似,语义信息论也应当在语义无失真压缩、语义信道容量、语义限失真编码、多用户语义通信等方面建立牢固的理论基础,从而指导通信系统的优化设计,推动语义通信技术的真正落地。

3.3 语义编码方法的通用性

现有语义编码多数采用神经网络模型与深度学习方法,需要使用数据集进行大规模训练,其泛化能力与通用性受到一定限制。未来,需要深入研究基于迁移学习的语义编码,增强语义编码方法的通用性,适应不同信道特性与不同模态的信源特征。另一方面,也可以考虑设计语义编码的预训练大模型,利用大模型的通用能力,适配不同场景与信源的通信需求。

3.4 语义编码传输的实用化

目前的语义通信系统都偏向于理想传输场景,很少考虑量化损失、功放非线性、器件非理想、同步误差等实际通信非理想因素的影响,并且语义编码依赖神经网络模型,需要消耗大量计算资源。考虑这些问题,语义通信的实用化方案初期在网络侧/基站端部署的可能性较大,随着 IC 芯片性能的进一步提升,成熟期的语义通信技术才会逐步在终端侧应用。因此,语义编码传输的实用化还有众多问题有待深入探讨。

4 结束语

本文对语义通信的理论与方法进行了系统探讨与总结。首先提出了语义通信的系统模型,总结了语义通信概念与术语,强调了与经典通信的区别。然后,类比经典信息论,归纳了语义信息的度

量体系,基于算法信息论提出归一化柯尔莫哥洛夫复杂性,对于语义压缩极限进行了深入探讨,并进一步讨论了语义编码的典型序列特征。进一步地,介绍了 2 种语义编码传输方法——直接编码与变换编码的基本原理,并进行了文本、语音与图像传输等不同模态下语义编码传输的性能对比分析。最后,探讨了未来语义通信的研究难点与问题。

综上所述,语义通信打开了通信领域一扇新的大门,相信在学术界与工业界研究人员的共同努力下,现代通信将会揭开新的一页。

参考文献:

- [1] SHANNON C E. A mathematical theory of communication[J]. Bell System Technical Journal, 1948, 27(4): 623-656.
- [2] MORRIS C W. Foundations of the theory of signs[M]. Chicago: The University of Chicago Press, 1938.
- [3] SHANNON C E, WEAVER W. The mathematical theory of communication[M]. Urbana: The University of Illinois Press, 1971.
- [4] BRILLOUIN L. Science and information theory[M]. New York: Academic Press, 1956.
- [5] CARNAP R, BAR-HILLEL Y. An outline of a theory of semantic information[R]. 1952.
- [6] BAR-HILLEL Y, CARNAP R. Semantic information[J]. The British Journal for the Philosophy of Science, 1953, 4(14): 147-157.
- [7] BARWISE J, PERRY J. Situations and attitudes[J]. The Journal of Philosophy, 1981, 78(11): 668-691.
- [8] FLORIDI L. Outline of a theory of strongly semantic information[J]. Minds and Machines, 2004, 14(2): 197-221.
- [9] KOLCHINSKY A, WOLPERT D H. Semantic information, autonomous agency and non-equilibrium statistical physics[J]. Interface Focus, 2018, 8(6): 20180041.
- [10] KOUNTOURIS M, PAPPAS N. Semantics-empowered communication for networked intelligent systems[J]. IEEE Communications Magazine, 2021, 59(6): 96-102.
- [11] RENYI A. On measures of entropy and information[C]//Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. [S.l.:s.n.], 1960: 547-561.
- [12] ZHANG P, XU W, GAO H, et al. Toward wisdom-evolutionary and primitive-concise 6G: a new paradigm of semantic communication networks[J]. Engineering, 2022, 8(1): 60-73.
- [13] ZHANG P, XU X, DONG C, et al. Intellicise communication system: model-driven semantic communications[J]. The Journal of China Universities of Posts and Telecommunications, 2022, 29(1): 2-12.
- [14] NIU K, DAI J C, YAO S S, et al. A paradigm shift toward semantic communications[J]. IEEE Communications Magazine, 2022, 60(11): 113-119.
- [15] 牛凯, 戴金晟, 张平, 等. 面向 6G 的语义通信[J]. 移动通信, 2021, 45(4): 85-90.
NIU K, DAI J C, ZHANG P, et al. 6G-oriented semantic communications[J]. Mobile Communications, 2021, 45(4): 85-90.
- [16] SHI G M, XIAO Y, LI Y Y, et al. From semantic communication to

- semantic-aware networking: model, architecture, and open problems[J]. IEEE Communications Magazine, 2021, 59(8): 44-50.
- [17] 石光明, 李莹玉, 谢雪梅. 语义通讯: 智能时代的产物[J]. 模式识别与人工智能, 2018, 31(1): 91-99.
- SHI G M, LI Y Y, XIE X M. Semantic communications: outcome of the intelligence era[J]. Pattern Recognition and Artificial Intelligence, 2018, 31(1): 91-99.
- [18] 石光明, 肖泳, 李莹玉, 等. 面向万物智联的语义通信网络[J]. 物联网学报, 2021, 5(2): 26-36.
- SHI G M, XIAO Y, LI Y Y, et al. Semantic communication networking for the intelligence of everything[J]. Chinese Journal on Internet of Things, 2021, 5(2): 26-36.
- [19] QIN Z J, TAO X M, LU J H, et al. Semantic communications: principles and challenges[J]. arXiv Preprint, arXiv: 2201.01389, 2022.
- [20] XIE H Q, QIN Z J, TAO X M, et al. Task-oriented multi-user semantic communications[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(9): 2584-2597.
- [21] SHI G M, GAO D H, SONG X D, et al. A new communication paradigm: from bit accuracy to semantic fidelity[J]. arXiv Preprint, arXiv: 2101.12649, 2021.
- [22] GÜNDÜZ D, QIN Z J, ESTELLA A I, et al. Guest editorial special issue on beyond transmitting bits: context, semantics, and task-oriented communications[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(1): 1-4.
- [23] CHACCOUR C, SAAD W, DEBBAH M, et al. Less data, more knowledge: building next generation semantic communication networks[J]. arXiv Preprint, arXiv: 2211.14343, 2022.
- [24] LUO X W, CHEN H H, GUO Q. Semantic communications: overview, open issues, and future research directions[J]. IEEE Wireless Communications, 2022, 29(1): 210-219.
- [25] ZHUANG Z W, LI R, JIA K, et al. Perception-aware multi-sensor fusion for 3D LiDAR semantic segmentation[C]//Proceedings of IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2022: 16260-16270.
- [26] MOHD T K, NGUYEN N, JAVAID A Y. Multi-modal data fusion in enhancing human-machine interaction for robotic applications: a survey[J]. arXiv Preprint, arXiv: 2202.07732, 2022.
- [27] LIU F, CUI Y H, MASOUIROS C, et al. Integrated sensing and communications: toward dual-functional wireless networks for 6G and beyond[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(6): 1728-1767.
- [28] ZHOU F H, LI Y H, ZHANG X Y, et al. Cognitive semantic communication systems driven by knowledge graph[C]//Proceedings of IEEE International Conference on Communications. Piscataway: IEEE Press, 2022: 4860-4865.
- [29] ZHOU K, ZHAO W X, BIAN S Q, et al. Improving conversational recommender systems via knowledge graph based semantic fusion[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2020: 1006-1014.
- [30] LIU P F, YUAN W Z, FU J L, et al. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing[J]. ACM Computing Surveys, 2023, 55(9): 1-35.
- [31] LI W B, LIN Z, ZHOU K, et al. MAT: mask-aware transformer for large hole image inpainting[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 10748-10758.
- [32] ZHANG R, ISOLA P, EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 586-595.
- [33] RIX A W, BEERENDS J G, HOLLIER M P, et al. Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs[C]//Proceedings of 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Piscataway: IEEE Press, 2002: 749-752.
- [34] BAO J, BASU P, DEAN M K, et al. Towards a theory of semantic communication[C]//Proceedings of 2011 IEEE Network Science Workshop. Piscataway: IEEE Press, 2011: 110-117.
- [35] LUCA A D, TERMINI S. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory[J]. Information and Control, 1972, 20(4): 301-312.
- [36] LUCA A D, TERMINI S. Entropy of L-fuzzy sets[J]. Information and Control, 1974, 24(1): 55-73.
- [37] KÖRNER J. Coding of an information source having ambiguous alphabet and the entropy of graphs[C]//Proceeding of the 6th Prague Conference on Information Theory. [S.l.:s.n.], 1973: 411-425.
- [38] KORNER J, MARTON K. Random access communication and graph entropy[J]. IEEE Transactions on Information Theory, 1988, 34(2): 312-314.
- [39] ALON N, ORLITSKY A. Source coding and graph entropies[J]. IEEE Transactions on Information Theory, 1996, 42(5): 1329-1339.
- [40] 吴伟陵. 广义信源与广义熵[J]. 北京邮电大学学报, 1982, 5(1): 29-41.
- WU W L. Generalized information source and generalized entropy[J]. Journal of Beijing University of Posts and Telecommunications, 1982, 5(1): 29-41.
- [41] 钟义信. 信息科学原理[M]. 北京: 北京邮电大学出版社, 1996.
- ZHONG Y X. Theory of information science[M]. Beijing: Beijing University of Posts and Telecommunications Press, 1996.
- [42] 钟义信. 自然语言理解的全信息方法[J]. 北京邮电大学学报, 2004, 27(4): 1-12.
- ZHONG Y X. Comprehensive information based methodology for natural language understanding[J]. Journal of Beijing University of Posts and Telecommunications, 2004, 27(4): 1-12.
- [43] 钟义信. 信息生态学与语义信息论[J]. 图书情报知识, 2017(6): 4-11.
- ZHONG Y X. Information ecology and semantic information theory[J]. Document, Information & Knowledge, 2017(6): 4-11.
- [44] 鲁晨光. 广义熵和广义互信息的编码意义[J]. 通信学报, 1994, 15(6): 37-44.
- LU C G. Meanings of generalized entropy and generalized mutual information for coding[J]. Journal of China Institute of Communications, 1994, 15(6): 37-44.
- [45] FAN P Y, DONG Y Q, LU J X, et al. Message importance measure and its application to minority subset detection in big data[C]//Proceedings of 2016 IEEE GLOBECOM Workshops. Piscataway: IEEE Press, 2017: 1-5.
- [46] SHE R, LIU S Y, FAN P Y. Recognizing information feature variation: message importance transfer measure and its applications in big data[J]. Entropy, 2018, 20(6): 401.
- [47] LI M, VITÁNYI P. An introduction to Kolmogorov complexity and its applications[M]. New York: Springer, 2008.

- [48] BOURTSOULATZE E, BURTH K 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.
- [49] TUNG T Y, KURKA D B, JANKOWSKI M, et al. DeepJSCC-Q: channel input constrained deep joint source-channel coding[C]//Proceedings of IEEE International Conference on Communications. Piscataway: IEEE Press, 2022: 3880-3885.
- [50] 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.
- [51] WENG Z Z, QIN Z J. Semantic communication systems for speech transmission[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(8): 2434-2444.
- [52] XIE H Q, QIN Z J. A lite distributed semantic communication system for Internet of things[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(1): 142-153.
- [53] BALLÉ J, CHOU P A, MINNEN D, et al. Nonlinear transform coding[J]. IEEE Journal of Selected Topics in Signal Processing, 2020, 15(2): 339-353.
- [54] BALLÉ J, LAPARRA V, SIMONCELLI E P. End-to-end optimized image compression[J]. arXiv Preprint, arXiv: 1611.01704v3, 2016.
- [55] YAO S S, XIAO Z X, DAI J C, et al. Variational speech waveform compression to catalyze semantic communications[C]//Proceedings of IEEE Wireless Communications and Networking Conference. Piscataway: IEEE Press, 2023: 1-6.
- [56] DAI J C, ZHANG P, NIU K, et al. Communication beyond transmitting bits: semantics-guided source and channel coding[J]. IEEE Wireless Communications, 2022: doi.org/10.1109/mwc.017.2100705.
- [57] 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.
- [58] WANG S X, DAI J C, LIANG Z J, et al. Wireless deep video semantic transmission[J]. IEEE Journal on Selected Areas in Communications, 2022, 41(1): 214-229.
- [59] WANG J, WANG S X, DAI J C, et al. Perceptual learned source-channel coding for high-fidelity image semantic transmission[C]//Proceedings of 2022 Global Communications Conference. Piscataway: IEEE Press, 2022: 3959-3964.
- [60] XIAO Z X, YAO S S, DAI J C, et al. Wireless deep speech semantic transmission[C]//Proceedings of 2023 International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2023: 1-5.
- [61] YAO S X, WANG S X, DAI J C, et al. Semantic information processing in industrial networks[J]. The Journal of China Universities of Posts and Telecommunications, 2022, 29(1): 41-49.
- [62] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM Press, 2017: 6000-6010.
- [63] WANG S X, DAI J C, YAO S S, et al. A novel deep learning architecture for wireless image transmission[C]//Proceedings of 2021 IEEE Global Communications Conference. Piscataway: IEEE Press, 2022: 1-6.
- [64] YAO S S, NIU K, WANG S X, et al. Semantic coding for text transmission: an iterative design[J]. IEEE Transactions on Cognitive Communications and Networking, 2022, 8(4): 1594-1603.
- [65] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2022: 9992-10002.
- [66] LIU L F, OESTGES C, POUTANEN J, et al. The COST2100 MIMO channel model[J]. IEEE Wireless Communications, 2012, 19(6): 92-99.
- [67] BESSETTE B, SALAMI R, LEFEBVRE R, et al. The adaptive multi-rate wideband speech codec (AMR-WB)[J]. IEEE Transactions on Speech and Audio Processing, 2002, 10(8): 620-636.
- [68] VALIN J, VOS K, TERRIBERRY T B. Definition of the opus audio codec[J]. RFC, 2012, 6716: 1-326.
- [69] SCHOEFFLER M, STÖTER F R, EDLER B, et al. Towards the next generation of web-based experiments: a case study assessing basic audio quality following the ITU-R recommendation BS. 1534 (MUSHRA)[C]//Proceedings of the 1st Web Audio Conference. [S.l.:s.n.], 2015: 1-6.

[作者简介]



张平 (1959—)，男，陕西汉中，人，博士，中国工程院院士，北京邮电大学教授、博士生导师，主要研究方向为先进移动通信系统等。



牛凯 (1976—)，男，山西文水，人，博士，北京邮电大学教授、博士生导师，主要研究方向为信息论与信道编码。



姚圣时 (1997—)，男，江苏常熟，人，北京邮电大学博士生，主要研究方向为信息论、信源信道编码和语义通信。



戴金晟 (1992—)，男，湖南常德，人，博士，北京邮电大学副教授、博士生导师，主要研究方向为无线通信信号处理、信息论与信道编码。