

特色专题

语义通信网络关键技术研究

罗良杰, 刘向南, 张海君*

摘要 传统通信系统难以满足智能通信场景对传输效率的迫切需求,综述了语义通信作为一种新型通信范式,在智能通信场景下能够理解和传递信息本质,并进一步探讨了其在未来无线通信技术中的发展前景。首先,基于数据的模态将语义通信分为单模态语义通信、跨模态语义通信、多模态语义通信3类,并梳理了不同模态下的语义通信研究现状。其次,针对语义通信网络在性能评估度量、收发端设计及资源管理策略等方面分析了相关研究,并提出了各项技术面临的挑战,以及未来发展中要解决的关键问题。最后,对语义通信与未来通信网络的深度融合进行展望与总结,未来语义通信亟需构建统一的语义性能评估体系,深化对跨模态语义融合与语义通信资源管理策略进一步研究,并推动语义通信与现有网络架构的不断融合,实现“语义赋能、智慧互联”的未来通信愿景。

关键词 语义通信;多模态;语义度量;资源管理

面向未来虚拟宇宙、数字孪生、智能工厂等多样化应用场景及海量智能设备的大规模接入,未来通信网络不再仅仅局限于高容量、高可靠性、低延迟,以及高准确性的传输需求,还需要具备语义感知、识别、分析、理解和推理能力,在人与人交互、人与机器交互的基础上实现万物智联。语义通信技术作为一种全新的通信范式,将助力未来通信网络实现由数据驱动向语义驱动的范式转变^[1]。

语义通信理论首次由 Shannon^[2]提出,语义通信被分为语法层次、语义层次和语用层次3个级别,其中,语法层次的根本目的是准确无误地传输消息的符号/比特,其主要问题是解决发射机到接收机的符号传输。语法层次由香农信息理论驱动,是现代通信系统中不可或缺的基石,但是语法层次局限于符号/比特级别的传输机制,因此,不具备对消息上下文含义的解析能力。而语义层次则聚焦于传输符号如何准确地

传达预期含义,此层次上实现的通信过程称为语义通信。语义通信能够使通信双方在信息的“含义”上相互传输,其本质上是发送端提取语义,并在接收端以最小的语义差错在不同的信道条件下进行恢复,其主要问题是实现收发端语义的准确传输。而语用层次考虑的则是接收到的信息对于接收者目标的效用,更加注重传输的语义信息如何高效地实现预期目标,其主要目标是解决语义信息的效用性问题,此层次的通信一般称为目标导向的通信。

语义通信能够从源数据中提取与目标高度相关的语义特征信息,通过最小化语义上不相关信息,能够有效降低传输延迟,减少带宽与功耗等通信资源的开销。具体来说,语义编码器从源数据中提取相关的语义特征并进行语义压缩。在通过信道编码与物理信道后,这些失真的语义特征由信道解码器和语义解码器恢复。与传统通信仅聚焦于数据传输与交换不同,语义通信实现了从语法层到语义层的范式跃迁。通过对语义信息的压缩,通信效率显著提升,而基于

北京科技大学计算机与通信工程学院,北京 100083

收稿日期: 2025-05-05; 修回日期: 2025-12-04

作者简介: 罗良杰, 硕士研究生, 研究方向为无线网络资源管控, 电子信箱: 1358133781@qq.com; 张海君(通信作者), 教授, 研究方向为 6G 移动通信, 电子信箱: zhanghaijun@ustb.edu.cn

引用格式: 罗良杰, 刘向南, 张海君. 语义通信网络关键技术研究[J]. 科技导报, 2026, 44(9): 62-74; doi:10.3981/j.issn.1000-7857.2025.05.00006

语义理解的通信机制也能有效地增强信息解析能力,在6G(第6代移动通信)及未来移动通信网络中,语义通信展现出广阔的应用前景,尤其适用于海量数据处理、元宇宙、增强/虚拟现实、智能工厂等亟需高效、智能的通信新范式以应对数据规模与复杂度呈指数级增长的新兴场景。

1 语义通信分类

传统的通信系统只考虑语法层次,在Shannon^[2]确定的3个通信层次中,语义通信的目标是将语义层次和语用层次整合到现有的通信系统的设计中,因此,可以考虑将语义通信分为3个类别,即面向语义的通信、面向目标的通信、语义感知通信^[3]。这种分类方式是针对语义通信在具体应用中的作用,并没有体现出源数据的模态间的差别,如图1所示,基于数据的模态将语义通信分为单模态语义通信、跨模态语义通信及多模态融合语义通信3个类别。

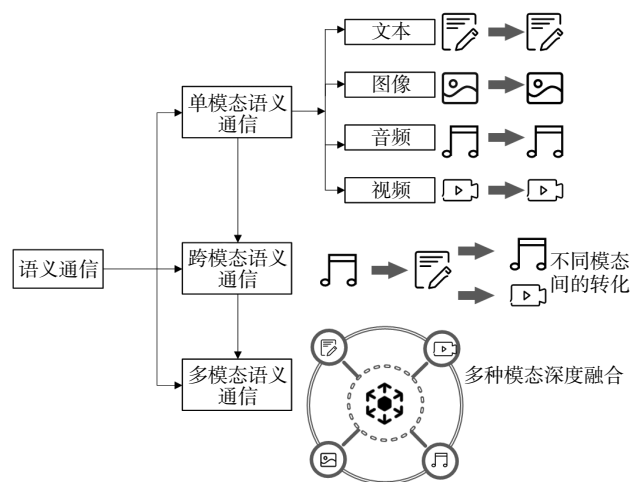


图1 语义通信分类

1.1 单模态语义通信

单模态语义通信是指基于单一类型数据的语义信息传输系统,就目前语义通信的研究而言,围绕典型的文本、图像、语音、视频等模态。

在文本通信领域,语义通信技术突破了传统字符编码传输的局限,特别是在低带宽环境下,仅传输关键语义信息,而非完整句子,能够显著提升通信效率。文本语义通信由Farsad等^[4]首次提出,使用递归神经网络作为编码器和解码器,能够提取语义信息并从中恢复文本,但是该算法使用了固定的比特长度来

编码不同长度的句子,会造成带宽浪费以及信息损失等问题。在此基础上,Rao等^[5]进一步开发出一种用于文本的可变长度联合信源和信道编码方案,能够将输入文本动态地编码为可变长度的传输符号。受到深度学习端到端通信与自然语言处理技术的启发,Xie等^[6]提出了一种基于Transformer的语义通信系统,通过恢复句子的含义而不是传统通信中的比特或符号错误来最大化系统容量并最小化语义错误。Xie等^[7]继续使用模型冗余修剪以及模型量化等技术对深度学习模型进一步压缩,以实现快速模型传输和低成本物联网设备的部署。Jiang等^[8]提出了一种混合自动重复请求方法,该方法能够进一步减少文本传输中的语义错误。Lu等^[9]设计了一种基于置信度的蒸馏机制,并提出了一种利用强化学习解决语义鸿沟的语义文本通信系统。Zhang等^[10]提出了一种基于掩码学习的多速率文本语义通信方案,通过引入目标速率作为边信息,实现具有高语义恢复性能的任意速率传输。

语音语义通信方面,Weng等^[11]设计了一种基于深度学习的语音语义通信系统,提取和传输全局语义特征,该方法将语音谱的每一帧看作图像,并利用卷积神经网络来压缩语音谱;Xiao等^[12]通过引入信道信噪比自适应机制,使单个模型也能够应用于各种信道状态。Tong等^[13]提出了一种基于联邦学习的方法,能够进一步提高语义提取的准确性。Weng等^[14]提出了一种基于注意力的语义通信系统,其中语音信号被转录为文本并在接收器处合成,即使在信噪比较低的环境中仍能保持清晰的语义传达。进一步地,Weng等^[15]还提出一种高鲁棒性的语音语义通信系统,通过传递基本的语义信息来实现语音传输。Zhou等^[16]提出一种端到端语音语义通信系统,在接收端对语义信号进行多层信息的重构和整合,在保留原始语义信息的同时有效地消除语义信号噪声。

深度学习的不断发展让图像领域也实现了智能化的语义特征提取,而基于注意力机制的深度学习技术也能够自动聚焦图像中的语义关键区域,如人脸、躯体、文字等,实现自适应的语义编解码,为远程医疗、自动驾驶等场景提供新的可能性。Bourtsoulatze等^[17]提出了一种用于支持无线图像传输的联合信源信道编码的卷积神经网络模型,该模型可以在有限带宽

和低信噪比条件下自适应地调整图像的重建质量; Kurka 等^[18]则设计了一种具有多个不同压缩比的细化层的分层无线图像传输方案; Zhang 等^[19]提出了一种用于图像传输的多级语义感知通信,其显示了诸如图像字幕信息的高级语义信息的重要性。Erdemir 等^[20]研究了一种基于生成模型的图像语义传输联合信源信道编码方案, Huang 等^[21]考虑了用于图像传输的语义通信系统。Peng 等^[22]设计了一种利用多尺度语义信息来大大减轻语义损伤并增强语义保真度的图像语义通信系统。

视频语义通信技术则突破了时空冗余处理的瓶颈,通过动作识别和事件检测算法,系统能够识别“人物跳跃、移动”等高层语义,仅传输关键语义帧而非连续画面。Jiang 等^[23]提出了一种语义视频会议系统,通过用关键点表示人脸的语义来减少传输负载。Huang 等^[24]针对点云视频数据量身定制的创新语义通信系统,采用轻量级的编解码器和关注区域选择,在资源受限的场景下实现实时解码和渲染。

1.2 跨模态语义通信

多模态服务通过整合文本、图像、音频、视频及触觉信号,为用户提供低延迟、高可靠性的沉浸式体验。尽管这些多模态信号在数据结构、传输延迟和抖动特性上存在差异,但它们之间潜在的跨模态关联性为信号协同处理提供了可能。语义通信作为一种全新的通信范式,通过提取多模态信号中的紧凑语义信息,能够实现不同模态间的语义关联,进而实现高效的跨模态语义通信。但是,语义本身具有多义性与歧义性,难以满足跨模态语义通信的可靠性要求。现有研究^[25]通过挖掘模态间潜在关联及构建语义公共信息库,已初步实现跨模态语义通信。Chen 等^[26]提出如图 2 所示的跨模态语义通信框架,该框架通过整合跨模态语义关联,可以融合不同模态的数据进行传输。

跨模态语义通信通过挖掘不同模态间的内在关联来实现语义特征的高效表征与编码,从而突破模态间的传输壁垒并显著提升通信性能^[27-28]。相较于传统单模态通信,这种技术能够通过动态语义对齐与融合机制更准确地理解和传递多模态信息^[29]。Weng 等^[30]开发了一种语音识别语义通信系统,其通过传输文本相关的语义特征来在接收器处重构语音信号的文本转录,实现语音数据与文本数据的跨模态通信。

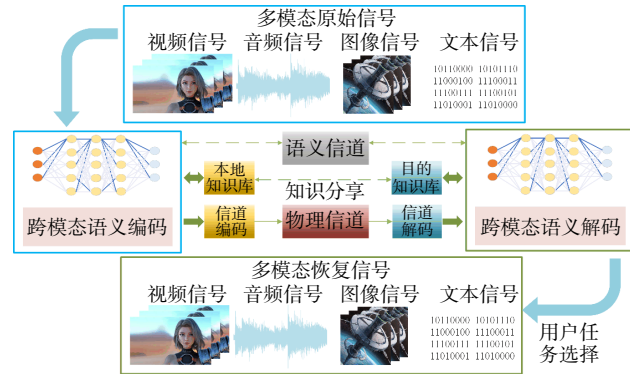


图2 跨模态语义通信框架

Weng 等^[14]还设计了一个用于实现语音识别和合成任务的语音语义通信系统,并将机器翻译模块融合进该系统中,进一步探索了语音到文本翻译和语音到语音翻译任务。Xie 等^[31]通过融合文本和视觉语义特征,为机器翻译和视觉问答任务设计了一个面向任务的语义通信系统,另外, Xie 等^[32]也设计了一种能够针对任务进行多模态数据传输的视觉问答语义通信系统。Li 等^[33]提出了一种跨模态知识图谱,用于进行语义编码和解码;针对无线信道中信号传输的特点, Luo 等^[34]提出了一种信道级信息融合的多模态信息融合方案,实现了在多种无线信道中的可靠传输。Pokhrel 等^[35]则提出了一种域自适应散列方法,能够从多模态数据中提取语义。Qin 等^[36]提出了一种能够从多模态源和无线信道中提取语义的新方法。Xie 等^[37]设计了一种用于多模态数据的记忆辅助语义通信系统,以提高传输效率。Wang 等^[29]提出了一个具有跨模态校准的深度学习语义通信系统,有效地利用了多模态信号之间的相关性来增强传输的鲁棒性; Jiang 等^[38]设计了基于视觉语言模型构建的跨模态语义通信系统,实现基于文本语义引导的图像重构以降低带宽消耗;而另一项可编辑语义通信技术则支持用户对文本-图像语义进行个性化编辑与传输^[39]。

然而现有研究在将跨模态信息深度整合到联合信源信道编码(joint source-channel coding, JSCC)框架方面仍存在明显不足,这使得开发支持深度 JSCC 的跨模态语义通信机制成为当前亟待解决的关键问题。

1.3 多模态语义通信

多模态深度融合语义通信主要以多模态语义对齐技术为核心,实现跨模态语义提取、关联建模、自

适应编码等功能。在未来 6G 场景下,多模态语义通信通过文本、语音、图像、视频等模态的深度融合、协同编码,推动人机交互从符号传输向单模态语义传输再到多模态融合语义传输的跃迁。从一个新的跨域融合角度,通过共享和传递多种媒介获得的多种模态信息和通感算存跨域资源,更好地减轻云—边—端三侧、通感算存多域,以及空天地无线传输中与任务无关的复杂负面影响^[40]。尽管基于单模态的语义通信已显著提升传输效率,但它难以完整表达多模态间语义关联特性,也难以实现面向更加多样化场景的智能交互,因此,需要深度融合多模态语义来构建更完备的通信体系。GPT-4^[41]、Gemini^[42]、Llama^[43]等多模态大语言模型(multimodal large language models, MLLM)为多模态语义通信提供了统一的语义表征平台,能够整合来自文本、图像、音频等多种模态的信息^[44],由于 MLLM 具有强大的推理能力和丰富的世界知识^[45],因此,它能够语义交流提供强大的知识库和多模态信息处理能力,实现跨模态的语义对齐与生成。Wang 等^[46]提出了一种集成 LLM 的端到端可学习语义通信模型,该模型利用 LLM 的各种结构来设计语义编码器,在语义保真度、跨场景泛化和复杂性方面表现出优异的性能。通过将 MLLM 和生成模型的最新进展融入新兴的语义通信范式,Nam 等^[47]提出了一种新的面向语言的语义通信框架,通过在上下文中学

习侦听器的语言风格来生成侦听器定制的提示。Zhang 等^[48]将预训练的基于 LLM 的编码器与定制的图形解码器相结合来生成语义。Zhao 等^[49]引入了一个由 LLMs 驱动的语义通信系统,通过使用多模态特征重建原始视觉信息来扩展单峰传输系统并增强其泛化能力,进而提高传输质量。

多模态深度融合的语义通信系统可以通过语义图谱建模不同模态间的语义依存关系,实现不同模态语义特征融合。针对现有多模态语义通信研究中忽略了对多模态数据间推理关联性的探索的问题,Xing 等^[50]提出了一种基于知识图谱的多模态语义表示与融合模型,语义信息被提取并映射成一个双层次语义架构,能够利用知识图谱的结构优势融合多模态语义信息。而 Li 等^[51]则是提出了一种基于有向图的跨模态特征互补算法,缓解多模态融合中的异质性差距问题。

2 语义通信的使能技术

根据语义通信系统整体设计流程,语义通信的关键实现技术可以分为性能评估度量、收发端设计、资源管理 3 个方面,这些技术将支撑语义通信实现通信双方在语义层次上相互沟通,推动现有通信从传统的符号/比特传输升级为语义传输,如图 3 所示。

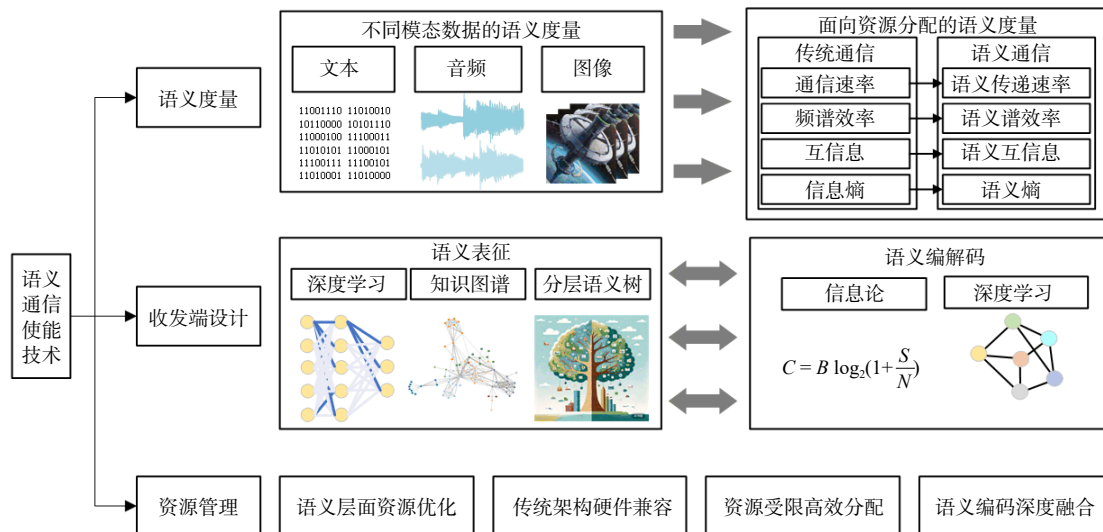


图3 语义通信关键技术

2.1 语义通信性能评估度量

通信系统性能评估是通信系统设计的关键组成

部分,依赖于通信系统的性能指标,可以定量比较不同通信优化算法之间的性能优劣,并根据特定场景的

性能需求实现差异化的算法优化。与传统通信不同,语义通信的核心在于传递语义信息,其系统架构和通信策略均发生根本性的转变,传统性能评估指标无法完全适用,需要建立新的评价体系来准确衡量语义通信系统的性能。

2.1.1 不同模态数据的语义度量指标

由于传输数据模态的不同,用于衡量语义通信系统性能的度量存在着细微差别。针对语义文本传输,现有许多文本语义度量方法对文本语义系统进行评估。如在一定程度上评估语义相似度的语义距离与词错误率(word error rate, WER)^[52],以及衡量 2 个语句之间的语义相似度的双语评估替补(bilingual evaluation understudy, BLEU)^[53]。WER 不适用于语义文本传输,因为 2 个不同词语也可能具有很高的语义相似度;BLEU 最初是为了评估机器翻译文本的质量,后来被用于评估文本语义通信系统的有效性^[54]。但是,由于 BLEU 分数仅比较 2 条文本消息之间的差异,无法区分单词中更细微的差异,如多义词和同义词之间的差别。因此,提出了用句子相似度计算原始发送句子和恢复句子之间的语义相似度^[55]。另外,从传统通信的传输比特角度,Jiang 等^[8]也引入了每句话的平均比特消耗作为文本语义通信度量。

针对语音数据而言,一般使用以下度量:WER、字符错误率(character error rate, CER)、信号失真率(signal-to-distortion ratio, SDR),以及语音质量感知评价(perceptual evaluation of speech quality, PESQ)。与 WER 在词语级别评估不同,CER 在字符层次评估语音识别的准确性^[30]。而 SDR 则是语音传输的常用度量,用于衡量源语音数据与重构语音数据之间的均方误差^[56]。对于语义通信而言,用户体验(quality of experience, QoE)也是一个十分重要的指标,而 PESQ 是国际电信联盟制定的度量标准,能够用于评估语音语义通信的 QoE^[57]。

对于图像数据而言,已经提出许多用于图像语义传输的语义度量方法。像素级别的峰值信噪比^[17]用于量化期望图像信号的最大可能功率与期望重构信号的噪声功率之间的比率,结构级别的结构相似度^[58]用于评估源图像与重构图像的整体质量,多尺度结构相似度^[59]用于在各种分辨率下捕获图像的细节,以及学习感知图像块相似度^[60]、视觉转换分数^[61]均用于在语

义级别上评估图像相似度。Peng 等^[22]重点关注图像中由对抗性扰动引起的语义损伤,提出图像语义损伤强度来量化语义损伤的程度,Fan 等^[62]提出语义相似度评分来量化图像之间语义级的差异。由于视频是由多个图像帧构成的,一些图像数据的度量指标其实也能够用于视频数据,如峰值信噪比(peak signal-to-noise ratio, PSNR)^[52]等。

2.1.2 面向资源分配的语义度量指标

依赖于通信系统的性能指标,可以比较所提出的通信优化算法相较于现有优化算法的优劣情况,能够针对特定场景的性能需求实现差异化的通信资源分配,以实现单个或多个网络的优化。单模态数据的性能指标显然不能满足未来整个语义通信系统的设计需求。基于传统通信架构的性能指标,结合语义通信的本质特点,可以考虑以下的面向资源分配的通用语义度量指标。

1) 语义相似度。在之前单模态数据中能够通过语义距离、WER、CER,以及图像相似度在一定程度上评估源数据与重构数据之间的语义相似度。语义相似度是衡量 2 段数据在含义上的相似程度的指标,不是简单地考虑词语间的匹配性,而是从语义层次评估两者是否表达相同或相近的含义。

2) 语义传递速率。在传统通信中,单位时间内传输的数据量即通信速率,用于衡量通信链路的效率。同样地,在语义通信中,使用语义传递速率来衡量语义通信系统在单位时间内传递有效语义信息的效率,Yan 等^[63]将语义传递速率定义为每秒内有效传输的语义信息量。

3) 语义谱效率。频谱效率是衡量传统通信系统在单位带宽内传输信息能力的核心指标,定义为每赫兹带宽可传输的比特数,是传统无线通信优化的重点研究方向。而与之类似地,语义谱效率则被定义为语义信息可以在单位带宽上成功传输的速率^[63]。

4) 语义互信息。语义互信息旨在量化特定任务语义压缩过程中存在的语义级失真。当处理输入图像的像素级信息时,可以获取特征级信息,进而获取语义级信息。语义级信息是最终被任务理解并包含在感知结果中的含义^[64]。

5) 语义熵。传统通信系统中,信息熵基于信源符号的统计特性来度量信息,而语义熵则是直接量化信

源的语义信息,在资源分配中对语义通信效率的量化发挥着重要作用。Yan 等^[65]提出了一种语义熵的近似度量,Xin 等^[66]则提出一种基于语义扩展和知识冲突的语义通信系统度量,能够衡量离散源的语义熵,在此基础上,Rong 等^[67]提出一种应用于图像的语义熵度量。

现有语义通信度量指标主要针对单一模态数据进行度量,另外,针对传统通信的性能指标的语义化以适应未来语义通信也是值得研究的方向,在未来语义通信网络技术中仍然需要提出更加通用的性能评估度量,以适应语义通信技术的发展。

2.2 语义通信收发端设计

2.2.1 语义表征

语义通信的本质是语义信息的传输,通过只传输语义信息,语义通信能够实现用更少的数据生成更多的知识,而其中语义表征则是语义信息传输的第 1 步。在通信系统中,原始数据往往包含大量与目标任务无关的冗余信息。为提高信息处理效率,必须对数据进行语义层面的有效过滤和特征提取。早期的语义表征研究受限于计算能力,主要采用手动特征设计方法,如尺度不变特征变换^[68]和基于组件的模板匹配技术^[69]。然而,这类方法无法建立底层特征与高层语义之间的映射关系,导致语义表征的语义泛化能力较差,无法适应动态变化的通信场景。

随着语义通信的不断发展,语义信息可以根据语义表征算法的不同,表示为 3 种形式:基于深度学习的语义特征向量、基于知识图谱的实体关系网络,以及基于分层语义树的实体间层次网络^[70]。

1) 基于深度学习的语义特征向量。早期特征设计方法的泛化能力差,无法适应动态变化的通信场景,随着设备性能的提升,基于深度学习的语义特征提取开始受到越来越多的关注。通过深度学习技术,可以将图像、文本等数据映射到语义空间,但基于深度学习的语义表征是需要针对特定的任务进行训练的,对于非目标任务的语义表征泛化能力较差,需要针对不同的任务进行相关深度学习模型的训练。

2) 基于知识图谱的实体关系网络。知识图谱是一种用于描述现实世界中的实体、属性及其相互关系的数据结构。知识图谱通过图结构组织知识,其基本单元是“实体-关系-实体”或者“实体-属性-值”的三

元组。实体对齐技术是知识融合技术的重要组成部分,旨在将不同的知识图谱中指向同一对象的不同实体进行连接^[71]。通过实体对齐,能够实现多模态语义信息的跨源一致性和交互高效率。另外,可以将多模态原始数据的语义信息建模为知识图谱,进而序列化为三元组序列,以支持高效的知识传输与推理^[70]。Hu 等^[72]提出了一种基于知识图谱的语义通信系统的示例,而 Li 等^[33]则将知识图谱用于语义编码和解码。

3) 基于分层语义树的实体间层次网络。分层语义树(hierarchical semantic tree, HST)将实体组织成树形结构,根据实体之间的关系,分层语义树可以分为 2 类:一类是基于本体的 HST,只包含严格的属性关系;另一类则是基于分类的 HST,包含上下文关系。智能体若要基于层次化语义树实现高效语义通信,其核心前提是确保 HST 叶节点所承载的元语义信息具备严格的形式化定义。Wilks^[73]给出了此类元语义信息需满足的 5 大约束条件:有限性、全面性、独立性、非循环性、连续性。Zhong^[74]提出了一种基于 HST 的语义通信系统,该系统通过由 n 维元语义信息组成的向量来表示消息的语义信息。而 Shi 等^[75]则提供了一个包括视觉、听觉、嗅觉、味觉和触觉等人类基础感官的元语义信息的示例。Zhang 等^[76]将元语义信息定义为语义通信网络的“语义基”,并解释了多模态数据的语义信息可以通过多级变换最终由语义基序列表示。

2.2.2 语义编解码

语义编码通过捕获传输数据的潜在含义实现信息的高效传输,与传统通信的编码技术不同,语义编码考虑了数据中的语义和关系,可以在保持基本内容和减少冗余的同时获得更高的压缩比。在语义表征的框架下,语义编解码需要更加有效地构建信号与语义之间的映射^[77]。

早期的语义编码研究主要集中于从信息论的角度实现语义编解码,Juba 等^[78]基于信源和信宿处符号的先验概率分布可能不同的事实,通过信息论论证了模糊压缩的必要性,Güler 等^[79]从语义相似度的角度,利用博弈论方法设计了一个有损编码框架,目标是最小化端到端平均语义错误。

随着深度学习领域的不断发展,语义编解码开始向着深度学习的方向发展。Farsad 等^[4]提出了一种基

于深度学习的文本联合信源信道编码,它实现了较低的单词错误率并保留了句子的语义信息,而 Bourtsoulaze 等^[17]则将图像像素值直接映射到复值信道输入符号,并学习噪声弹性编码表示,因此,在所有信噪比下都优于基于分离的数字通信。Xu 等^[80]提出基于注意力机制的联合信源信道编码,成功地在不同的信噪比水平下实现图像传输。Li 等^[81]则是提出了一种具有可扩展性的语义编码框架,将非结构化图像分解为层

次结构的语义特征,以增强图像中不同语义目标的处理能力,从而提高了语义表示的准确性。Huang 等^[21]提出了一种基于强化学习的自适应语义编码方法,实现了对图像进行像素级以上的编码。为了减轻突发噪声的影响,Barka 等^[82]提出了一种脉冲噪声自适应的深度语义通信架构。另外,神经网络技术也是语义编码中不可或缺的一环,如通过卷积网络提取语义特征^[83],引入注意力机制的图注意力网络^[84]等(表 1)。

表 1 语义通信编解码方案对比

参考文献	编解码技术	信源模态	技术方案
[79]	贝叶斯博弈	文本	将语义通信问题表述为贝叶斯博弈,并研究贝叶斯纳什均衡存在的条件,最小化端到端平均语义错误
[4]	递归神经网络	文本	使用递归神经网络编解码器实现更低的单词错误率
[81]	注意力机制	图像	提出了一种可伸缩的语义编码框架,以增强对图像中不同语义目标的处理
[83]	卷积神经网络	视频+文本	文本通过相对语义进行卷积神经网络编码,视频通过标准编解码器编码

总体而言,语义通信收发端设计的核心挑战在于构建兼具语义保真度与场景适应性的通用化模型。早期工作主要基于信息论框架与特征设计方法对语义冗余问题进行理论探索,而随着深度学习的突破,基于深度神经网络的端到端语义特征提取与编码已成为可能。未来需要对多模态深度融合下的语义关联以及联合信源信道编码做进一步探索。

2.3 语义通信资源管理策略

在传统通信下的资源分配的核心目标是在有限的网络资源条件(如频谱、功率等)下,优化网络性能指标(如能效、谱效、吞吐量等),主要围绕物理层参数进行优化。而语义通信下的资源分配,则是在语义层次上提高通信效率,优化语义通信系统的性能指标,对应用层的任务性能提出了新要求。现有的传统通信的服务质量定义不再适用于面向任务的语义通信,精确的语义资源分配仍然是一个棘手的挑战。

基于 Wang 等^[85]提出的 3 个问题,本文将其拓展深化为以下 4 个语义资源分配亟待解决的关键问题。

2.3.1 语义层面实现准确的资源优化

针对语义层次的资源优化问题,Yan 等^[63]提出的资源分配方案能够在语义级别上优化资源分配,但需要预先训练的映射来指导网络资源分配,其中传输度量和语义度量之间的关系是预先记录的,难以应对动态化的通信场景,而 Wang 等^[85]提出了一种采用语义

比特量化的资源分配方案,解决了由于语义度量与传输度量之间的额外映射关系而导致的环境感知不准确问题。

2.3.2 与传统无线通信架构下的硬件设备兼容

针对硬件设备兼容问题,Hu 等^[86]基于码本的语义知识库,实现了连续语义信息的离散化,使得语义信息可以在无线通信中通过比特传输。但是,基于码本的量化效率较低,需要一种更加高效通用的语义比特量化方法来适应当前无线传输硬件设备。针对高动态的车载网络,Su 等^[87]提出了一种优化每层语义信息的访问控制和功率控制的方法,这个系统能够得到满足用户低延迟和高可靠性的结果,在同信道干扰和信道不确定性的高动态环境中能够表现出良好的性能。针对语义和超可靠低延迟通信业务共存的融合系统,Ding 等^[88]提出了一种动态复用和协同调度方案,在满足业务延迟要求的同时,最大化语义服务的效用。针对多小区多任务上行蜂窝网络,Yan 等^[89]提出了一种基于体验质量的多任务网络资源感知分配方法,解决了信道分配和发射功率分配问题。

2.3.3 在资源受限条件下实现通信资源的高效分配

在资源受限情况下,如何最大化语义通信的效率也成为一个问题,Mu 等^[90]提出了一种语义驱动的上行链路非正交多址接入框架,利用语义通信在低信噪比或无线资源有限的情况下也能够提供良好

的性能。Zhang 等^[91]引入了一种基于深度强化学习的面向任务的语义通信网络动态资源分配方案,对具有基本语义信息的数据进行优先级排序,以提高资源受限的无线网络中的任务性能。为了进一步提高频谱利用率和能量可持续性,Zhang 等^[92]以用户体验质量为评价指标,提出了一种能量收集、认知无线电和非正交多址接入联合场景下的语义感知资源分配方案,实现语义信息丰富的数据优先占用资源。Chen 等^[93]提出了一种基于深度强化学习的最大化长期平均语义能量效率的新算法,该算法用于能量收集驱动的语义通信系统,能够实现能量受限场景的语义通信。Wu 等^[94]提出了正交频分复用传输中的信道自适应 JSCC 方案,引入了双重注意机制,以使用在各种信噪比环境中估计的信道状态信息来确定适当子载波的特征和功率分配。Tung 等^[95]介绍了第 1 种 E2E JSCC

视频传输方法,提出了一种用马尔可夫决策过程表示的优化方法,用于为图像组的每一帧分配可用带宽。

2.3.4 语义资源分配与现有语义编码技术深度融合

针对语义资源分配与现有的语义编码深度融合的问题,传统的数学方法难以支持实时计算并与 AI 驱动的语义编码协同工作。对此,Wang 等^[96]通过计算基于知识图谱的语义信息结构的重要性,进而优化资源块的分配与语义信息的编码,实现语义编码与语义资源分配的协同。另外,为了解决支持多用户通信场景下语义压缩与语义失真相互矛盾的问题,Liu 等^[97]提出了一种基于语义重要性自适应确定压缩比,同时优化资源分配和用户选择的算法,并对该算法进行了仿真验证,同时 Zhang 等^[91]也提出了一种语义压缩比、发射功率和带宽分配的联合优化策略(表 2)。

表 2 语义通信资源管理策略对比

参考文献	面向问题	优化目标	学习方法	技术方案
[85]	语义层面实现准确地资源优化	语义量化效率、传输延迟	深度强化学习	采用语义比特量化,解决了由于语义度量与传输度量之间的额外映射关系而导致的环境感知不准确问题
[89]	与传统无线通信架构下的硬件设备兼容	语义熵、语义保真度、语义传递速率	深度神经网络	定义了语义熵来量化不同任务用户的语义信息,分析了语义熵与香农熵之间的关系
[93]	在资源受限条件下实现通信资源的高效分配	语义保真度、语义能量效率	深度强化学习	在保证语义保真度的前提下,最大化能量收集驱动的语义感知网络的能量效率
[97]	语义资源分配与现有的语义编码技术深度融合	语义压缩率、传输延迟	深度神经网络	优化压缩比和资源分配,以最大限度地提高任务的成功概率

3 未来语义通信的研究展望

未来语义通信的研究可以考虑以下方向,以推动语义通信向更高效、更智能、更可靠、低冗余的方向发展,为现有通信范式由“比特传输”向“语义交互”转变提供支撑。同时,使能技术的突破也将拓展语义通信在自动驾驶、工业互联网、远程医疗等领域的应用深度。

1) 在语义通信数据模态方面,研究多模态语义通信的深度融合,考虑多模态语义统一表征,实现将不同模态的数据映射到统一的语义空间,消除模态间的异构性;研究多模态深度融合下动态语义关联建模新方法,开发动态自适应融合机制,捕捉多模态数据间的动态语义依赖关系。针对远程医疗场景,融合医学影像、诊断报告和语音描述,实现多模态数据融合的

语义关联辅助诊断。针对自动驾驶场景,融合激光雷达、视频图像,以及 GPS(全球定位系统)数据,实现复杂环境的多模态语义理解与高效传输。

2) 在语义通信性能评估方面,设计面向自动驾驶、远程医疗等不同应用场景的基准测试集,形成标准化的语义通信评估框架。在远程医疗场景中,重点考虑语义压缩准确性、语义传输实时性、语义理解精确性对于医疗诊断的影响,实现远程医疗领域的语义通信性能评估;在自动驾驶方面,量化评估语义提取与语义理解的实时性、精确性,实现自动驾驶领域的通信性能评估。面向多样化的应用场景,考虑构建多维度的语义通信评估体系,实现语义通信指标与语义通信性能的统一。

3) 在语义通信资源管控方面,研究智能协同的语

义通信资源优化机制,能够面向异构算力与动态网络需求,构建语义驱动的跨域资源管控统一框架,实现算力与网络深度融合下的高效语义资源分配。针对医学数据对语义通信高保真、低延迟,以及隐私保护的需求,考虑采用动态语义卸载与算力分配策略,针对医疗数据的风险特性,实现语义任务的动态卸载与资源管控;针对自动驾驶场景下多源环境语义通信的低延迟与高可靠需求,实现终端轻量化语义提取、边缘局部语义融合、云端全局语义优化的分层语义通信资源管控架构。

4 结语

围绕语义通信网络关键技术展开分析,探讨了语义通信的层次化递进类别及使能技术。语义通信作为新一代智能通信范式,在未来通信中具有巨大的应用潜力,能够重构当前传统通信网络架构。但是,要实现语义通信的广泛应用,仍须攻克语义通信的多模态深度融合问题、跨域资源管控问题,以及设计更加通用合理的性能评估度量体系。针对这些挑战,未来研究可聚焦于多模态语义对齐、语义资源感知及语义信息基础理论,通过理论创新与算法优化,构建支持6G网络的高效语义通信体系。希望能为推动通信范式变革提供新思路,助力实现“语义赋能、智慧互联”的未来通信愿景。

参考文献(References)

- [1] 张海君, 陈安琪, 李亚博, 等. 6G移动网络关键技术[J]. 通信学报, 2022, 43(7): 189–202.
- [2] Shannon C E. A mathematical theory of communication[J]. The Bell System Technical Journal, 1948, 27(3): 379–423.
- [3] Yang W T, Du H Y, Liew Z Q, et al. Semantic communications for future internet: Fundamentals, applications, and challenges[J]. IEEE Communications Surveys & Tutorials, 2023, 25(1): 213–250.
- [4] Farsad N, Rao M, Goldsmith A. Deep learning for joint source–channel coding of text[C]//Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2018: 2326–2330.
- [5] Rao M, Farsad N, Goldsmith A. Variable length joint source–channel coding of text using deep neural networks[C]//Proceedings of IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). Piscataway, NJ: IEEE, 2018: 1–5.
- [6] 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.
- [7] 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.
- [8] Jiang P W, Wen C K, Jin S, et al. Deep source–channel coding for sentence semantic transmission with HARQ[J]. IEEE Transactions on Communications, 2022, 70(8): 5225–5240.
- [9] Lu K, Zhou Q Y, Li R P, et al. Rethinking modern communication from semantic coding to semantic communication[J]. IEEE Wireless Communications, 2023, 30(1): 158–164.
- [10] Zhang M T, Zhang H X. Mask learning enabled multi–rate text semantic communications[C]//Proceedings of IEEE/CIC International Conference on Communications in China (ICCC Workshops). Piscataway, NJ: IEEE, 2024: 260–264.
- [11] 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.
- [12] Xiao Z X, Yao S S, Dai J C, et al. Wireless deep speech semantic transmission[C]//Proceedings of ICASSP 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2023: 1–5.
- [13] Tong H N, Yang Z H, Wang S H, et al. Federated learning for audio semantic communication[J]. Frontiers in Communications and Networks, 2021, 2: 734402.
- [14] Weng Z Z, Qin Z J, Tao X M, et al. Deep learning enabled semantic communications with speech recognition and synthesis[J]. IEEE Transactions on Wireless Communications, 2023, 22(9): 6227–6240.
- [15] Weng Z Z, Qin Z J, Li G Y. Robust semantic communications for speech transmission[C]//Proceedings of ICASSP 2025 – 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2025: 1–5.
- [16] Zhou Z L, Zheng S L, Chen J, et al. Speech semantic communication based on swin transformer[J]. IEEE Transactions on Cognitive Communications and Networking, 2024, 10(3): 756–768.
- [17] Boursoulatz 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.

- [18] 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.
- [19] Zhang Z G, Yang Q Q, He S B, et al. Wireless transmission of images with the assistance of multi-level semantic information[C]//Proceedings of International Symposium on Wireless Communication Systems (ISWCS). Piscataway, NJ: IEEE, 2022: 1–6.
- [20] 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.
- [21] Huang D L, Gao F F, Tao X M, et al. Toward semantic communications: Deep learning-based image semantic coding[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(1): 55–71.
- [22] Peng X, Qin Z J, Tao X M, et al. A robust image semantic communication system with multi-scale vision transformer[J]. IEEE Journal on Selected Areas in Communications, 2025, 43(4): 1278–1291.
- [23] Jiang P W, Wen C K, Jin S, et al. Wireless semantic communications for video conferencing[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(1): 230–244.
- [24] Huang Y K, Bai B Y, Zhu Y W, et al. ISCom: Interest-aware semantic communication scheme for point cloud video streaming on metaverse XR devices[J]. IEEE Journal on Selected Areas in Communications, 2024, 42(4): 1003–1021.
- [25] Zhou L, Wu D, Wei X, et al. Cross-modal stream scheduling for ehealth[J]. IEEE Journal on Selected Areas in Communications, 2020, 39(2): 426–437.
- [26] Chen M K, Liu M H, Wang W J, et al. Cross-modal semantic communications in 6G[C]//Proceedings of IEEE/CIC International Conference on Communications in China (ICCC). Piscataway, NJ: IEEE, 2023: 1–6.
- [27] Wei X, Wu D, Zhou L, et al. Cross-modal communication technology: A survey[J]. Fundamental Research, 2025, 5(5): 2256–2267.
- [28] Wu D, Zhou L. Cross-modal stream transmission: Architecture, strategy, and technology[J]. IEEE Wireless Communications, 2024, 31(1): 134–140.
- [29] Wang W J, Liu M H, Chen M K. CA_DeepSC: Cross-modal alignment for multi-modal semantic communications[C]//Proceedings of GLOBECOM 2023 IEEE Global Communications Conference. Piscataway, NJ: IEEE, 2023: 5871–5876.
- [30] Weng Z Z, Qin Z J, Li G Y. Semantic communications for speech recognition[C]//Proceedings of IEEE Global Communications Conference (GLOBECOM). Piscataway, NJ: IEEE, 2021: 1–6.
- [31] 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.
- [32] Xie H Q, Qin Z J, Li G Y. Task-oriented multi-user semantic communications for VQA[J]. IEEE Wireless Communications Letters, 2022, 11(3): 553–557.
- [33] Li A, Wei X, Wu D, et al. Cross-modal semantic communications[J]. IEEE Wireless Communications, 2022, 29(6): 144–151.
- [34] Luo X W, Gao R B, Chen H H, et al. Multimodal and multiuser semantic communications for channel-level information fusion[J]. IEEE Wireless Communications, 2024, 31(2): 117–125.
- [35] Pokhrel S R, Choi J. Understand-before-talk (UBT): A semantic communication approach to 6G networks[J]. IEEE Transactions on Vehicular Technology, 2023, 72(3): 3544–3556.
- [36] Qin Z J, Gao F F, Lin B, et al. A generalized semantic communication system: From sources to channels[J]. IEEE Wireless Communications, 2023, 30(3): 18–26.
- [37] Xie H Q, Qin Z J, Li G Y. Semantic communication with memory[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(8): 2658–2669.
- [38] Jiang F B, Tang C G, Dong L, et al. Visual language model-based cross-modal semantic communication systems [J]. IEEE Transactions on Wireless Communications, 2025, 24(5): 3937–3948.
- [39] Yu W B, Chen B, Zhang Q S, et al. Editable-DeepSC: Cross-modal editable semantic communication systems[C]//Proceedings of IEEE 99th Vehicular Technology Conference (VTC2024-Spring). Piscataway, NJ: IEEE, 2024: 1–5.
- [40] 任超, 丁思颖, 张晓奇, 等. 6G无线多模态通信技术[J]. 电子与信息学报, 2024, 46(5): 1658–1671.
- [41] Mai L, Carson-Berndsen J. Enhancing conversation smoothness in language learning chatbots: An evaluation of GPT4 for ASR error correction[C]//Proceedings of ICASSP 2024 – 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2024: 11001–11005.
- [42] Islam R, Ahmed I. Gemini—the most powerful LLM: Myth or truth[C]//Proceedings of 5th Information Communication Technologies Conference (ICTC). Piscataway, NJ: IEEE, 2024: 303–308.

- [43] Lin D F, Wen Y, Wang W S, et al. Enhanced sentiment intensity regression through LoRA fine-tuning on llama 3[J]. *IEEE Access*, 2024, 12: 108072–108087.
- [44] Liang T, Lin G S, Wan M Y, et al. Expanding large pre-trained unimodal models with multimodal information injection for image-text multimodal classification[C]// *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway, NJ: IEEE, 2022: 15471–15480.
- [45] Jiang Y L, Weng J W, Zhang X T, et al. A CNN-based born-again TSK fuzzy classifier integrating soft label information and knowledge distillation[J]. *IEEE Transactions on Fuzzy Systems*, 2023, 31(6): 1843–1854.
- [46] Wang Y, Sun Z, Fan J P, et al. On the uses of large language models to design end-to-end learning semantic communication[C]// *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*. Piscataway, NJ: IEEE, 2024: 1–6.
- [47] Nam H, Park J, Choi J, et al. Language-oriented communication with semantic coding and knowledge distillation for text-to-image generation[C]// *Proceedings of ICASSP 2024–2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Piscataway, NJ: IEEE, 2024: 13506–13510.
- [48] Zhang M, Chen W L, Duan X Y, et al. Improving graph-based dependency parsing models with dependency language models[J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2013, 21(11): 2313–2323.
- [49] Zhao Y R, Yue Y, Hou S L, et al. LaMoSC: Large language model-driven semantic communication system for visual transmission[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2024, 10(6): 2005–2018.
- [50] Xing C L, Lü J, Luo T, et al. Representation and fusion based on knowledge graph in multi-modal semantic communication[J]. *IEEE Wireless Communications Letters*, 2024, 13(5): 1344–1348.
- [51] Li J, Wang X P, Lü G Q, et al. GraphCFC: A directed graph based cross-modal feature complementation approach for multimodal conversational emotion recognition[J]. *IEEE Transactions on Multimedia*, 2023, 26: 77–89.
- [52] 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.
- [53] Papineni K, Roukos S, Ward T, et al. Bleu: A method for automatic evaluation of machine translation[C]// *Proceedings of the Proceedings of the 40th annual meeting of the Association for Computational Linguistics*. Stroudsburg, PA, USA: ACL, 2002: 311–318.
- [54] Zhou Q, Li R, Zhao Z, et al. Semantic communication with adaptive universal transformer[J]. *IEEE Wireless Communications Letters*, 2022, 11(3): 453–457.
- [55] Xie H Q, Qin Z J, Li G Y, et al. Deep learning based semantic communications: An initial investigation[C]// *Proceedings of GLOBECOM 2020–2020 IEEE Global Communications Conference*. Piscataway, NJ: IEEE, 2020: 1–6.
- [56] Vincent E, Gribonval R, Fevotte C. Performance measurement in blind audio source separation[J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2006, 14(4): 1462–1469.
- [57] 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 IEEE International Conference on Acoustics, Speech, and Signal Processing*. Proceedings. Piscataway, NJ: IEEE, 2001: 749–752.
- [58] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: From error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600–612.
- [59] Wang Z, Simoncelli E P, Bovik A C. Multiscale structural similarity for image quality assessment[C]// *Proceedings of Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003. Piscataway, NJ: IEEE, 2003: 1398–1402.
- [60] Zhang R, Isola P, Efros A A, et al. The unreasonable effectiveness of deep features as a perceptual metric[C]// *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2018: 586–595.
- [61] Zhu T T, Peng B, Liang J F, et al. How to evaluate semantic communications for images with ViTScore metric?[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2024, 10(5): 1744–1758.
- [62] Fan S R, Bao Z C, Dong C, et al. Semantic similarity score for measuring visual similarity at semantic level[J]. *IEEE Internet of Things Journal*, 2025, 12(9): 12034–12047.
- [63] Yan L, Qin Z J, Zhang R, et al. Resource allocation for text semantic communications[J]. *IEEE Wireless Communications Letters*, 2022, 11(7): 1394–1398.
- [64] Getu T M, Kaddoum G, Bennis M. Making sense of meaning: A survey on metrics for semantic and goal-oriented communication[J]. *IEEE Access*, 2023, 11: 45456–45492.
- [65] Yan L, Qin Z J, Zhang R, et al. QoE-aware resource allocation for semantic communication networks[C]// *Proceedings of GLOBECOM 2022–2022 IEEE Global Communications Conference*. Piscataway, NJ: IEEE, 2022: 3272–3277.
- [66] Xin G T, Zhu Z Q, Fan P Y. Information framework expansion

- sion meets knowledge collision for semantic communications[C]//Proceedings of ICC 2023 – IEEE International Conference on Communications. Piscataway, NJ: IEEE, 2023: 6096–6101.
- [67] Rong Y K, Nan G S, Zhang M W, et al. Semantic entropy can simultaneously benefit transmission efficiency and channel security of wireless semantic communications[J]. IEEE Transactions on Information Forensics and Security, 2025, 20: 2067–2082.
- [68] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91–110.
- [69] Crandall D, Felzenszwalb P, Huttenlocher D. Spatial priors for part-based recognition using statistical models[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Piscataway, NJ: IEEE, 2005: 10–17.
- [70] Wang Y N, Han H, Feng Y X, et al. Semantic communication empowered 6G networks: Techniques, applications, and challenges[J]. IEEE Access, 2025, 13: 28293–28314.
- [71] 马浩然, 王金华. 基于图结构与元素信息融合的实体对齐技术[J]. 科技导报, 2024, 42(18): 98–109.
- [72] Hu L S, Li Y H, Zhang H, et al. Robust semantic communication driven by knowledge graph[C]//Proceedings of 9th International Conference on Internet of Things: Systems, Management and Security (IOTSMS). Piscataway, NJ: IEEE, 2022: 1–5.
- [73] Wilks Y. An intelligent analyzer and understander of English[J]. Communications of the ACM, 1975, 18(5): 264–274.
- [74] Zhong Y X. A theory of semantic information[J]. China Communications, 2017, 14(1): 1–17.
- [75] Shi G, Li Y, Xie X. Semantic communications: Outcome of the intelligence era[J]. Pattern Recognition and Artificial Intelligence, 2018, 31(1): 91–99.
- [76] Zhang P, Xu W J, Gao H, et al. Toward wisdom-evolutionary and primitive-concise 6G: A new paradigm of semantic communication networks[J]. Engineering, 2022, 8: 60–73.
- [77] Zhang P, Xu W J, Liu Y M, et al. Intellicise wireless networks from semantic communications: A survey, research issues, and challenges[J]. IEEE Communications Surveys & Tutorials, 2025, 27(3): 2051–2084.
- [78] Juba B, Sudan M. Universal semantic communication I[C]//Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing. New York: ACM, 2008: 123–132.
- [79] Güler B, Yener A, Swami A. The semantic communication game[J]. IEEE Transactions on Cognitive Communications and Networking, 2018, 4(4): 787–802.
- [80] 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.
- [81] Li H X, Gao D H, Yang M X, et al. A scalable coding method with semantic decomposition for semantic communication[J]. IEEE Wireless Communications Letters, 2025, 14(7): 2009–2013.
- [82] Barka H, Kaddoum G, Sahabul Alam M, et al. Beyond Shannon capacity: Deep semantic communication for enhanced performance in impulsive noise environments[J]. IEEE Wireless Communications Letters, 2025, 14(1): 148–152.
- [83] Mitrica I, Mercier E, Ruellan C, et al. Very low bitrate semantic compression of airplane cockpit screen content[J]. IEEE Transactions on Multimedia, 2019, 21(9): 2157–2170.
- [84] Velickovic P, Cucurull G, Casanova A, et al. Graph attention networks[J]. Stat, 2017, 1050(20): 4.
- [85] Wang L Y, Wu W, Zhou F H, et al. Adaptive resource allocation for semantic communication networks[J]. IEEE Transactions on Communications, 2024, 72(11): 6900–6916.
- [86] Hu Q Y, Zhang G Y, Qin Z J, et al. Robust semantic communications with masked VQ-VAE enabled codebook[J]. IEEE Transactions on Wireless Communications, 2023, 22(12): 8707–8722.
- [87] Su J W, Liu Z X, Xie Y A, et al. Semantic communication-based dynamic resource allocation in D2D vehicular networks[J]. IEEE Transactions on Vehicular Technology, 2023, 72(8): 10784–10796.
- [88] Ding G Y, Liu S L, Yuan J T, et al. Joint URLLC traffic scheduling and resource allocation for semantic communication systems[J]. IEEE Transactions on Wireless Communications, 2024, 23(7): 7278–7290.
- [89] Yan L, Qin Z J, Li C F, et al. QoE-based semantic-aware resource allocation for multi-task networks[J]. IEEE Transactions on Wireless Communications, 2024, 23(9): 11958–11971.
- [90] Mu X D, Liu Y W. Exploiting semantic communication for non-orthogonal multiple access[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(8): 2563–2576.
- [91] Zhang H J, Wang H Y, Li Y B, et al. DRL-driven dynamic resource allocation for task-oriented semantic communication[J]. IEEE Transactions on Communications, 2023, 71(7): 3992–4004.
- [92] Zhang H J, Wang H Y, Li Y B, et al. Toward intelligent resource allocation on task-oriented semantic communication[J]. IEEE Wireless Communications, 2023, 30(3): 70–77.

- [93] Chen X J, Xu J, Ni W, et al. Energy-efficient resource allocation for multi-user semantic communications: A deep reinforcement learning approach[J]. IEEE Wireless Communications Letters, 2025, 14(5): 1541–1545.
- [94] Wu H T, Shao Y L, Mikolajczyk K, et al. Channel-adaptive wireless image transmission with OFDM[J]. IEEE Wireless Communications Letters, 2022, 11(11): 2400–2404.
- [95] Tung T Y, Gündüz D. DeepWiVe: Deep-learning-aided wireless video transmission[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(9): 2570–2583.
- [96] Wang Y N, Chen M Z, Luo T, et al. Performance optimization for semantic communications: An attention-based reinforcement learning approach[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(9): 2598–2613.
- [97] Liu C H, Guo C L, Yang Y, et al. Adaptable semantic compression and resource allocation for task-oriented communications[J]. IEEE Transactions on Cognitive Communications and Networking, 2024, 10(3): 769–782.

Research on key technologies of semantic communication networks

LUO Liangjie, LIU Xiangnan, ZHANG Haijun*

School of Computer & Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

Abstract Traditional communication systems are unable to meet the urgent demand for transmission efficiency in intelligent communication scenarios. An overview of semantic communication is presented as a new communication paradigm, which can understand and transmit the essence of information in intelligent communication scenarios. The development prospects of semantic communication in future wireless communication technologies are also discussed. Firstly, based on the type of data, semantic communication is classified into three categories: single-modal semantic communication, cross-modal semantic communication, and multi-modal semantic communication. The current research status under each mode is combed. Furthermore, an in-depth analysis is conducted on performance evaluation metrics, transceiver system design, and resource management strategies within semantic communication networks. Related studies from recent years are analyzed with a focus on three enabling technologies and the challenges faced by various technologies and the key issues to be solved in future development have been proposed. Finally, the deep integration of semantic communication with future communication networks is explored and summarized. In the future, semantic communication urgently needs to establish a unified semantic performance evaluation system. Further research is required on cross-modal semantic fusion and semantic communication resource management strategies, and efforts should be made to continuously integrate semantic communication with the existing network architecture, in order to achieve the future communication vision of "semantic empowerment and intelligent interconnection".

Keywords semantic communication; multi-modal; semantic metrics; resource management ●



(责任编辑 王微)