

语义通信技术商用化之路



Road to Commercialization of Semantic Communication Technology

石光明/Shi Guangming^{1,2}, 杨旻曦/Yang Minxi^{1,3},
高大化/Gao Dahua^{1,3}, 肖泳/Xiao Yong^{2,3,4}

(1. 西安电子科技大学人工智能学院, 中国 西安 710126;

2. 鹏城实验室, 中国 深圳 518055;

3. 琶洲实验室(黄埔), 中国 广州 510555;

4. 华中科技大学, 中国 武汉 430074)

(1. School of Artificial Intelligence, Xidian University, Xi'an 710126, China;

2. Pengcheng Laboratory, Shenzhen 518055, China;

3. Pazhou Laboratory, Huangpu, Guangzhou 510555, China;

4. Huazhong University of Science and Technology, Wuhan 430074, China)

DOI: 10.12142/ZTETJ.202603010

网络出版地址: <https://link.cnki.net/urlid/34.1228.TN.20260626.1023.006>

网络出版日期: 2026-06-26

收稿日期: 2026-05-20

摘要: 随着人工智能技术的快速发展, 传统以比特传输为核心的通信范式正面临根本性变革。系统梳理了语义信息处理与通信从理论起源到技术演进的发展脉络, 阐明其本质内涵与核心优势, 分别从通信与智能两个维度, 剖析其商用化面临的关键挑战并研判发展前景。语义通信的商用化仍受制于理论体系不完善、全链路技术成熟度不足、标准化进程滞后三重瓶颈; 而智能化的持续推进, 特别是智能体技术的发展, 将有力推动语义通信网的落地部署。在工业物联网、应急通信、深海水声通信等具有明确语义任务和场景约束的垂直行业中, 语义通信技术有望在3~5年内实现局部试商用; 伴随6G技术和标准化的演进, 语义编解码和语义认知网络架构等有望在5~10年纳入全球6G标准框架; 而实现全网络语义内生智能化的大规模部署, 则预计需要10~20年。本研究可为学术界与产业界开展语义通信研究及应用部署提供参考。

关键词: 语义通信; 商用化; 语义通信网; 通信体系架构; 智能体通信; 下一代移动通信

Abstract: With the rapid development of artificial intelligence technology, the traditional communication paradigm centered on bit transmission is undergoing fundamental changes. This paper systematically reviews the development route of semantic information processing and communication from theoretical origins to technological evolution, and clarifies its essential connotation and core advantages. From the two dimensions of communication and intelligence, it analyzes the key challenges faced by its commercialization and forecasts its industrial prospects. The commercialization of semantic communication is still restricted by three major bottlenecks: imperfect theoretical system, insufficient maturity of full-link technologies, and lagging standardization process. The continuous advancement of intelligence, especially the development of agent technology, will effectively promote the practical deployment of semantic communication networks. In vertical industries with clear semantic tasks and scenario constraints, such as industrial Internet of Things, emergency communications, and deep-sea underwater acoustic communications, semantic communication technology is expected to achieve partial pilot commercialization within 3–5 years. With the evolution of 6G technologies and standardization, semantic encoding and decoding as well as semantic cognitive network architectures are expected to be incorporated into the global 6G standard framework in 5–10 years. In contrast, the large-scale deployment of full-network endogenous semantic intelligence is estimated to take 10–20 years. This work can provide references for academic and industrial communities to carry out research and application deployment of semantic communication.

Keywords: semantic communication; commercialization; semantic communication network; communication system architecture; agent communication; next-generation mobile communication

引用格式: 石光明, 杨旻曦, 高大化, 等. 语义通信技术商用化之路 [J]. 中兴通讯技术, 2026, 32(3): 51–59. DOI: 10.12142/ZTETJ.202603010

Citation: Shi G M, Yang M X, Gao D H, et al. Road to commercialization of semantic communication technology [J]. ZTE technology journal, 2026, 32(3): 51–59. DOI: 10.12142/ZTETJ.202603010

回顾人类文明的技术演进脉络, 工业革命解放了人类的体力, 互联网构建了全球化的信息通道, 而人工智能正在规模化地生产智力。当今智能时代, 信息技术的服务对象正从人本身扩展到智能机器。机器不仅是工具, 更成为智力的生产者和消费者。在这一背景下, 大约10年前, 少数先驱学者开始倡导语义通信技术, 主张通信系统应从传输比

特转向传递语义。经过近年来的快速发展, 语义通信已成为通信与人工智能交叉领域公认的研究热点^[1-2]。当前不论是高校、研究所, 还是华为、中兴通讯等设备厂商, 以及中国移动、中国电信、中国联通等通信运营商, 都在投入资源从事语义通信技术的研究。国际上, 著名电信企业和研究机构也在密集跟进, 例如: Nokia Bell Labs 已联合 KT 启动 6G 语

义通信研发^[3]；Samsung Research在2025年发布的6G白皮书中，将语义/任务导向通信列为关键方向^[4]；NTT DOCOMO在6G架构中纳入流式语义通信^[5]。不同于大数据、区块链、虚拟现实/增强现实（VR/AR）等先在国际形成热点再被中国追赶的技术轨迹，语义通信走出了一条反向路径：先在中国通信学术圈形成热点，再由国际学者跟进研究，最后企业参与研发落地。

一项信息技术从萌芽到商用，需要技术成熟度、市场迫切性、生态完整性与政策支持度多维度共同支撑，周期通常为10~20年。语义通信从萌芽到成为热点已近10年。那么语义通信何时以何种方式进入商用阶段？许多信息技术企业正在关注此事并作出各自预测，本文就此问题给出系统分析与判断。

1 通信目的与通信技术目标

约10万年前，人类最早靠手势、眼神、简单音节传递信息，逐渐形成口语，实现面对面的高效即时沟通；约3万年前，洞穴壁画、岩画、结绳记事用于记录事件、传递简单信息，形成最早的“符号编码”；公元前，中国烽火狼烟传递军情、非洲击鼓传信、希腊火炬接力实现远距离预警；1844年莫尔斯电报机与莫尔斯电码首次用电信号远距离即时传输文字，华盛顿至巴尔的摩电报线路的开通标志着电气通信时代开启。至此，用电子信号通信成为人类社会不可或缺的要害。无论是远古的洞穴壁画、结绳记事、烽火狼烟、击鼓传信，还是当今的电磁波信号，通信的目的始终是给接收方传递信息、传授知识、下达指令或布置任务。通信的本质是要“达意”^[1]——接收者要能收到发送者的本意，这个本意（信息、知识、指令、任务）被人类用各种模态的信号承载并能跨越时（隔时）空（远距）传播，人脑收到这些信号是要对其处理用于智能决策。当电子信息技术出现后，所有模态的信号都转变成了数字信号，数字通信成为当下隔空跨时的主流手段，占比达99%以上，几乎成为唯一通信模式。传统通信技术目标就定位于无损快速地把信源数据传递给接收方，使得接收方能够复现发送方的消息信号，也就是信号搬运；至于数据中承载的信息、知识、指令、任务等内容，并未成为传统通信技术的关注点。传统电子通信技术对比特保真的投入重于对内容保真的投入，因为它不知道信号内容在哪，只能坚守“信号保真保证内容保真”的信条。这导致传统网络在智能化时，如AI赋能的网络流量预测、切片优化等，只能基于外挂式的人工智能模型实现“黑盒”数据分析。而语义通信实现了网络数据流的语义化，使网络节点“理解”流量意图与

任务需求，最后可实现真正的内生智能调度与网络自愈。

2 语义通信的起源与发展

随着信息技术和人工智能的发展，移动机器人即将广泛融入人类社会的方方面面，人类社会活动需要传输的信息量越来越大。传统以频带和能耗资源换取通信容量性能的模式受到了挑战。特别是智能体之间的通信，需要一种低代价、高性能、直达意的通信模式。语义通信正是人工智能发展的产物^[1]。当智能终端需要理解视频内容、解读语音意图、判断图像情感时，传统的比特传输方式不得不将海量的语义信息转换为比特流进行传输，造成了严重的“语义鸿沟”——发送端编码的比特序列与接收端理解的语义内容之间存在巨大落差。

早在1988年，北京邮电大学钟义信教授在《信息科学原理》中系统提出“全信息”理论框架^[6]，2017年在*China Communications*上进一步发表语义信息理论^[7]，2018年在《智能系统学报》中构建“信息—知识—智能”转换机制^[8]，关联语义通信与通用人工智能。可以说钟义信先生是中国语义信息、语义通信研究的先驱。同期，北京邮电大学张平院士也在探讨语义通信的可行性，及其相较于传统通信的变革性。2017年12月，西安电子科技大学石光明教授在第S39次香山会议上首次讨论类脑语义信息技术，并于2018年发表文章指出“语义通信是智能时代的产物”^[11]。国际学者在2019—2020年也对语义通信进行了探索^[9-12]。在2020年之前，语义通信或语义信息并不为众多学者所熟知，也不完全被认可，因为很多基本概念尚未刻画清楚和准确。需要特别指出的是，张平院士在当时各类学术会议与报告中多次阐述语义通信的优势和前景，并首次提出“智简通信系统”^[13]。在他的倡导下，语义通信受到了众多学者的关注。据统计，语义通信领域的论文发表数量在2021—2026年间呈快速增长态势。语义通信已成为通信和人工智能领域的研究热点^[14-15]。此外，陆建华院士、张文军院士团队也针对语义通信开展了大量研究工作，取得了丰硕成果。截至目前，从基础理论（语义的数学刻画与度量^[16]、信号内容的表达^[17]），到核心技术（信源语义编解码方法^[2,18]、信源信道联合语义弹性编解码^[19]、语义知识构建^[20]、语义通信系统^[21]），以及语义通信网络^[22-26]，都已经有了相应的解决方案。

3 语义通信的本质与优势

语义通信是从“传信号比特”升级为“传语义比特”的新一代达意通信范式，针对信号中代表内容的重要比特进行

传输。其本质是以知识库为先验,降低信源熵,从而用很少的比特或特征表达信号中的信息内容。早期研究语义通信的学者普遍认可其两大优势。一是频带利用率高,或者说传输信息的密度高:其本质是基于语义知识库提取信源中的语义内容,然后用另一种简洁模态表达语义,在极高的数据压缩比下保留信源中的语义信息。它突破了传统通信传输所有比特的范式,提取并传输有效信息密度更高的语义符号,可在窄带条件下实现大容量传输。也就是说,它在信源压缩方面兼具内容保真与传输高效的优点。二是抗干扰能力强:传统抗干扰信道编码通过增加冗余比特实现可靠传输,但降低了频带资源利用率;而语义通信是利用学习得到的离线先验,在线智能纠正误码,其本质是用算力换取带宽^[13]。语义通信的信源压缩能力与信道抗误码能力,除了在海洋水声、战场等特殊场景中有利外,在当前5G带宽充裕、误码率较低的通信环境中,似乎不再具有明显优势——当前持这种看法的学者很多。然而,该观点仅局限于传统通信视角,存在一定片面性。

其实从智能的角度看,语义通信的一项突出能力——对信源语义内容的直接表达——尚未得到学界广泛重视,而正是这一能力使其成为网络内生智能的基础。具体而言,传统网络智能采用包括大模型等在内的人工智能技术,会产生巨大的传输和计算开销。而语义通信在信息源端就完成了“数据到特征/知识”的转变。由于传输的是直接可供认知和决策使用的语义流,网络节点不需要使用“先重构信号再提取语义”的传统方法,而是能直接根据接收到的语义信息进行任务协同与决策。这种“通信即计算”的新范式,大幅降低了整个智能系统的端到端时延与算力资源消耗。所以说,语义通信是智能时代的产物^[1]。

4 语义通信理论与技术成熟度

4.1 语义通信理论成熟度

语义数学刻画与语义通信的理论研究已从零散探索迈入体系化、可证明、可工程化阶段,核心进展集中在同义性语义信息论、语义数学公理/拓扑刻画、语义通信架构与编码定理、多模态与量子语义拓展四大方向,中国团队(北京邮电大学、西安电子科技大学、华为、中兴通讯等)处于全球引领地位。北京邮电大学张平院士团队构建语义熵、语义互信息、语义信道容量度量体系,该体系兼容香农理论;同时提出ComAI系统架构,支撑6G“原生智简”设计,并证明语义无损信源编码、语义信道编码、语义率失真编码定理,奠定了语义通信的理论极限。石光明教授领导西安电子科技

大学与鹏城实验室研发团队突破了信号-语义鸿沟^[17,27],对信号内容进行了刻画,提出从信号空间到语义空间的跨层映射数学模型^[16],突破了传统信息论仅处理语法信号的局限。石光明教授与华中科技大学肖泳教授合作,推导出任务引导的语义信息容量上界。中兴通讯研发团队发表了《基于熵函数的语义信息边界研究》^[28]。华为实验室提出了编码器速率与解码器准确率的权衡、语义熵的定义与计算等语义信息论基础问题。海南大学段玉聪团队量化刻画了语义对齐偏差与多维度语义歧义问题^[29]。

4.2 语义通信技术成熟度

为系统评估语义通信的技术成熟度,本文按端到端通信链路的功能分层与近年研究脉络,将语义通信划分为5个相互衔接的技术模块,并依次展开分析。

4.2.1 语义信源编码

语义信源编码是语义通信的基础,负责将原始数据转换为语义表示,按传输类型可分为文本、语音、图像3类。

文本语义编码最为成熟。Xie等提出DeepSC框架^[2],利用自然语言处理提取文本语义表示并完成传输,接收端通过深度网络重建文本,在语义准确率相当的情况下,将传输带宽压缩至传统方案的十分之一。

语音语义编码由Weng等于2021年率先开展研究^[18]。他们设计神经网络编码器提取语音高层语义特征,接收端通过语音合成恢复语义等价内容。实验表明,3 kbit/s低比特率下,其主观质量评分(MOS)可达到传统4G语音编码(AMR-WB, 23.85 kbit/s)的同等水平。2023年,该团队实现完整闭环系统,涵盖语音识别、语义提取、重建及语音合成四大环节^[21]。

图像语义编码是新兴方向。2022年,北京邮电大学张平院士团队Dai等提出DTSCC框架^[19],将语义通信拓展至图像领域,通过非线性变换前置模块实现端到端信源-信道联合优化,显著提升结构相似性指数(SSIM)指标。2023年,该团队进一步提出语义切片模型^[30],实现语义单元的按需传输与可扩展编码。

任务导向语义通信是重要扩展方向。2022年,Xie等在JSAC发表论文^[31],首次系统研究以任务完成为目标的语义通信,通过优化编解码器,丢弃与任务无关信息,进一步压缩传输比特率。

4.2.2 语义知识库与知识增强

语义知识库是关键使能技术,发送端与接收端共享先验知识库,仅传输缺失或不确定的语义内容^[20]。2023年,

Wang等在 *IEEE Communications Letters* 发表相关研究成果^[20]，在接收端引入知识图谱，以增强语义理解，提升语义一致性。

当前知识库研究面临两大挑战：一是跨领域通用知识库缺失，现有知识库多为专用知识库；二是静态知识与动态场景信息难以融合。此外，Masked VQ-VAE方案^[32]通过学习语义码本和掩码机制，增强语义码字抗信道噪声能力，保障恶劣信道下的语义恢复质量。

4.2.3 语义信道编码与联合编码

语义信道编码负责在语义层面对抗噪声，是语义通信实用化的关键。2025年，北京邮电大学研究团队在 *IEEE Transactions on Cognitive Communications and Networking* 发表论文^[33]，初步建立语义信道容量理论框架，证明其容量可显著高于传统信道容量。

联合信源—语义—信道编码是当前语义通信领域前沿研究热点。2026年，北京邮电大学研究团队在 *IEEE Transactions on Wireless Communications* 发表Token通信联合编码调制方案^[34]，验证了联合设计的可行性。2023年，Qin等提出通用化语义通信框架^[35]，将语义通信统一到联合编码框架下。在协议层面，2025年石光明团队Wang等提出无重传语义传输协议（SITP）^[36]，实现高可靠传输。目前语义信道编码仍处于起步阶段，与成熟比特层编码差距较大。

4.2.4 语义网络与资源管理

语义网络与资源管理将语义通信拓展至多用户场景，是支撑语义通信大规模商用的必要条件。现有研究已尝试将语义相似性引入用户分组、功率控制与带宽分配^[22]，以提升网络语义吞吐率与能效。此外，多任务场景用户体验质量（QoE）度量与资源分配方法也已得到初步验证^[22]。

在隐私保护方面，2024年刊发于 *IEEE Transactions on Wireless Communications* 的相关研究证实^[37]，语义层面可实现更精细的隐私保护，发送端可选择性传输任务相关语义信息，保留无关隐私。

4.2.5 生成式语义通信

生成式语义通信是新一代语义通信范式^[38-41]，发送端传递高度压缩的语义意图，接收端借助生成模型“再创造”语义等价内容。2024年，石光明团队在ICASSP上提出SG2SC框架^[38]，验证了生成式方法在图像语义通信中的优越性。2025年，石光明教授带领西安电子科技大学和鹏城实验室团队实现公里级水声图像视频通信^[40]。崔曙光团队在 *Engineering* 期刊上综述生成式语义通信（GSC）三大技术支柱，

搭建了对应的范式分类体系^[39]。

GSC对语义通信商用化影响深刻，主要表现在3个方面：一是提升接收端算力需求，约束终端部署；二是放弃逐比特精确性，使语义保真度的定义与验证成为新问题；三是在带宽极度受限的场景（如深空、应急通信）中具有“以算力换带宽”的独特优势。

4.3 语义通信系统成熟度

在频带资源稀缺的通信领域，如水下声通信、短波通信、星地通信，已有语义通信成功传输视频、图像和语音的报道。张平院士在多个学术报告中展示了相关研究进展，他的研究团队成功研制并展示了星地语义通信系统。在海洋水声通信领域，石光明教授团队研制的语义通信系统，实现了超百米级远距离多媒体信源语义无线传输。该语义通信系统所需要的算力不足200 TOPS（每秒万亿次运算）。在此之前，国际上未见有类似成果报道。

5 语义通信商用化挑战

尽管语义通信已在多类受限通信场景验证了技术可行性，但面向商业化落地仍存在诸多挑战。目前语义通信技术的推进多数集中在信源处理，在通信网络协议、信道编解码、发射波形等方面，语义赋能的相关研究成果仍较为匮乏。5G/6G的带宽很大，完全能够支撑目前的应用场景，加上语义通信与现有的网络协议的兼容性还没有完全建立，导致设备制造商和通信运营商对语义通信的落地迭代需求缺乏迫切性。这是当前产业推进的现实瓶颈。因此，语义通信从学术研究走向商用部署仍有很长的路要走。

5.1 兼容性挑战：与现有系统的融合

如果单从通信角度看语义通信，语义通信要实现商用，首先需要解决的是与现有通信系统的兼容问题。现实情况是：从成本和运营效益角度分析，无论是设备制造商还是通信运营商，都难以在短期内为语义通信研发专用设备、搭建全新通信网络。因此，语义通信必须能够与现有的4G/5G网络基础设施共存，甚至在某些场景下需要与传统的比特传输模式切换使用。然而，现有语义通信研究几乎未涉及与现有系统的融合方案。本文调研发现，截至目前，尚无文献系统研究语义层与比特层的共存机制、切换策略与协同工作模式。这一研究空白已成为语义通信商用化的重大障碍^[13]。

5.2 全链路语义化：最大的技术瓶颈

即使解决了兼容性问题，语义通信要实现端到端的语义传输，仍需解决全链路各环节的技术成熟度问题。

语义信源编码是当前成熟度最高的环节。文本^[2]、语音^[18,21]、图像语义编码^[19,30]均有较为完整的解决方案。但这些方案大多在理想实验室条件下完成验证,缺乏在真实复杂信道条件下的测试。

语义信道编码在2025年才刚刚起步^[33],目前尚未形成成熟的工程实现体系。传统的信道编码经过数十年的研究已经非常成熟(Turbo码、Polar码等均已标准化),而语义信道编码的编解码算法设计、性能分析、标准制定等均处于探索阶段。其中,石光明教授团队提出的SITP协议^[36]通过利用语义鲁棒性实现无重传高可靠传输,是协议层首个工程化尝试。

语义调制的相关研究目前几乎是空白^[34]。调制作为物理层的核心功能,负责将编码后的比特或符号映射到波形。语义调制需要将语义表示映射到电磁波形,而如何设计语义调制的星座图、检测算法、抗干扰性能分析,目前均缺乏系统性的研究。

语义网络协议研究仅有极少数初步探索工作^[22]。现有的网络协议栈(网络层、传输层等)均是为比特传输设计的,如果要支持语义网络,就需要在协议层面进行革命性的设计。这不仅是技术问题,还涉及标准制定、产业生态等非技术因素。

在基础支撑层面,语义知识库建设工作已取得进展。2026年5月16日,在中国通信学会现代语义通信专委会上,鹏城实验室联合合作团队发布了《语义知识库白皮书2.0》^[42]。这为语义通信的落地应用提供了重要支撑。

5.3 生成式语义通信的额外商用障碍

GSC在带来新能力的同时,也引入了传统语义通信尚未面临的商用障碍^[39]。

幻觉与可追溯性是GSC商用化面临的首要问题^[39]。生成模型可能在接收端“创造”出发送端并不存在的内容——这在扩展现实(XR)/元宇宙等娱乐场景中幻觉尚可接受,但在工业控制、医疗监测等安全关键场景中,任何非预期的内容生成均构成严重的系统可靠性隐患。现阶段,如何建立GSC的输出追责机制,目前尚无成熟方案。

推理时延与算力约束制约着GSC的实时性^[39]。扩散模型的去噪推理通常需要数十至数百步迭代运算,在当前终端算力水平下难以满足实时通信要求;即使在基站侧进行辅助推理,也面临网络时延叠加的问题。这意味着现阶段GSC更适合非实时或对时延不敏感的场景,如视频点播缓存、场景预渲染等。

感知一致性度量的缺失导致GSC质量评价体系不完

善^[39]。传统通信的峰值信噪比(PSNR)/SSIM度量在GSC场景中失效——高PSNR的生成图像可能存在语义错误,低PSNR的生成图像反而语义准确。目前学界尚未形成被广泛接受的GSC语义保真度量标准,这直接制约了GSC系统的公正评测和标准化推进工作。

内容安全、监管合规与隐私保护是GSC产业化不可避免的非技术壁垒^[37,39]。生成模型固有的内容生成能力可能被滥用于虚假信息传播,各国监管机构对AI生成内容的监管框架正在快速演进,GSC系统的商用合规路径尚不明朗。

6 语义通信商用化的动力

当前设备商和运营商对单纯降低带宽的“急迫感不强”,但对“网络智能化转型和新业务催生”具有较高的内生动力。随着千亿级异构智能体在网络中的大规模部署(大模型具身智能、无人集群),网络发展的核心痛点不再是速率不足,而是异构智能体之间“无法理解与无缝协作”,即语义鸿沟。语义通信作为“智能体网络网络层协议”的潜在颠覆者,其核心商用驱动力将不再是“省带宽”,而是“赋能网络具备连接、协同高阶智能的能力”。

目前加速语义通信商用的有利因素有3点:一是智能体的出现,催生了智能体集群之间的语义通信需求;二是发展海洋经济、开发深海的国家政策已经颁布,资源有限条件下的多媒体通信业务需求旺盛;三是国际电信组织正在推动语义通信相关国际标准的规划制定工作。

6.1 智能体集群之间的语义通信需求增加

智能体是AI技术发展的一个里程碑,它实现了智能从感知、认知、决策到执行的闭环。从单一智能体作业走向群体智能体协同作业是必然趋势。因此,智能体集群之间的语义通信需求快速攀升。智能体之间的协作交互重点是达意,通信的核心随之由“比特传输”转向“意义传递”,以突破传统句法通信在效率、理解与协作上的瓶颈。

一方面,全球智能体数量正迈向千亿级,通信主体从“人人互联”转向“智能体互联”。异构智能体(大模型Agent、工业机器人、无人机、车载单元等)跨厂商、跨平台、跨领域集群化部署,形成开放动态的智能体语义通信网络。智能体本身也从单一功能执行器进化为具备感知、推理、决策、自主学习能力的认知实体。此时交互目的不再是数据复制,而是意图理解、任务协同、知识共享。

另一方面,自动驾驶、智慧工厂、空天地一体化应急救援、数字孪生城市等场景,需要多智能体分布式协作(如无人机集群搜救、多车协同编队等)。这类业务的任务高度耦

合、动态变化,要求智能体实时对齐意图、共享上下文、同步决策逻辑。专业智能体(如医疗诊断、金融风控、工业控制)跨领域协作时,用户指令高度抽象,传统结构化数据交互难以传递抽象语义与隐性知识,必须依赖语义通信实现“共同理解”。

随着智能体全面融入各行各业,低成本跨主体信息交互成为刚需,将持续推动语义通信技术落地商用。为了保障信息安全和提高通信效率,语义通信网必然会向着“多传知识、少传数据”的方向发展^[43]。当单个智能体无力完成某项任务时,无须向协作方传输全量业务数据,转而由能力适配的智能体输出知识与技能,协助其完成本体能力迭代升级、独立完成任务。这样不仅减少了传输的数据量,而且还能有效保护数据隐私。

6.2 资源有限下的大容量通信需求正在增加

随着卫星直连的快速普及,其场景特性与传统通信的局限使得语义通信的需求持续激增。语义通信已成为卫星直连规模化、高质量发展的核心支撑,具体表现在3个方面:

一是卫星直连信道受限,传统通信难以适配。卫星直连面临链路时延长、多普勒频移明显、带宽稀缺且频谱资源紧张等问题。传统比特传输模式需全量传输数据,不仅容易造成资源浪费,在低信噪比环境下还易出现信号衰减、信息丢失,无法满足高效通信需求。

二是语义通信可破解卫星直连核心痛点。语义通信以“意义传递”为核心,可提取各类终端传输数据的核心语义,压缩冗余信息,大幅降低带宽消耗,同时通过语义保真技术确保关键信息准确传递,适配卫星直连的资源约束场景,提升传输效率与可靠性^[44]。

三是卫星直连场景拓展放大需求。随着卫星直连向应急救援、远洋航运、偏远地区通信等场景延伸,叠加低轨卫星星座规模化部署、多终端协同产生海量数据,语义通信可结合天上计算实现数据智能处理,同时打破异构终端互操作壁垒,支撑天地融合协同。

与此同时,中国实施深海经济开发战略,进一步催生和

放大语义通信刚需。深海开发依赖水下机器人、潜航器等智能体集群,声波作为主要通信载体,存在传播慢、衰减强、噪声高、多径干扰等问题,传统比特传输难以应对低带宽、高干扰场景,易出现指令误解、数据丢失,直接制约深海开发进程。语义通信可提取勘探数据、作业指令的核心语义,在有限带宽下实现高效传输,既能支撑精准勘探,也能保障水下智能体与水面平台的指令对齐,避免协同作业出现失误。随着深海经济向深远海推进,“十五五”时期对深海探测、开发能力的要求将进一步提升。语义通信结合边缘计算可破解传输低效难题,其抗衰落、高容错能力也能适配深海极端场景,未来需求将持续攀升。2025年11月,石光明领导的鹏城实验室与西安电子科技大学联合团队已经实现了水平和垂直千米级的水声多媒体信源通信^[40]。

6.3 标准化正在布局

标准化是技术商用的前提条件。国际标准化组织正在积极布局语义通信标准研制^[45]。中国学者和企业陆续在《中兴通讯技术》等国家级期刊发表面向6G标准化路径的语义通信架构与场景研究^[44,46]。2026年3月,鹏城实验室与北京邮电大学提交的技术提案正式写入3GPP TR 22.870规范^[47],多项6G语义通信场景用例提案在3GPP SA1福冈会议上获得表决通过,为后续语义通信关键技术的标准化推进打下了基础。此外,中国移动等通信运营商也在积极推动面向语义视频通话业务的端到端架构标准化工作。

7 商用前景与路径判断

尽管现阶段语义通信的发展面临诸多挑战,本文认为语义通信的商用前景仍然值得期待^[13,44]。本节将从商用条件成熟度、商用路径预测、关键使能因素3个方面展开具体分析。

7.1 商用条件的成熟度分析

如表1所示,评估语义通信的商用条件,需要从服务对象、效能收益和投资保护3个维度进行分析,针对不同应用

表1 语义通信典型应用场景的商用前景分析

应用场景	预期收益	算力需求	主要阻塞因素	最早落地窗口/年
工业物联网	带宽节省80%	低-中	知识库构建不充分	2026—2028
应急通信	极低带宽可用	中	鲁棒性不足	2026—2028
车联网	时延+带宽双降	中-高	端侧算力不足、标准缺失	2028—2030
扩展现实/元宇宙	沉浸感提升	高	端侧算力不足、GSC成熟度低	2029—2032
卫星/深空通信	高压缩比	中	信道建模不完善、验证不充分	2027—2030

GSC: 生成式语义通信

场景作出差异化判断。

从服务对象来看,智能终端的普及为语义通信提供了广阔的应用土壤。2026年,智能手机、智能音箱、智能汽车等终端已经具备强大的边缘AI能力,能够支持语义编码和解码的本地运行^[13]。然而,语义通信并非适用所有场景,对于实时性要求高、带宽资源丰富、终端能力充足的场景(如家庭Wi-Fi下的高清视频传输),传统比特传输仍然足够高效。语义通信更适合于带宽资源受限、终端能力有限的场景,如偏远地区的物联网数据传输、应急通信等。

从效能收益来看^[2,18,31],目前特定场景下的性能增益已经得到验证,但在更多场景下的泛化性能仍有待验证^[48]。本文认为,语义通信的性能优势具有场景依赖性:在文本摘要、语音情感识别等语义密集型任务中优势明显^[18],在高速率音频传输等语义稀疏型任务中优势并不明显。

从投资保护来看,当前运营商已经大规模建设了4G/5G网络基础设施,不可能在短期内对现有网络进行革命性替换。语义通信的商用需要考虑对现有投资的保护,设计渐进式的演进路径^[13]。

7.2 商用路径预测

基于上述分析,本文对语义通信的商用路径作出如下分阶段预测:

1) 近期(3~5年):特定垂直场景局部商用落地。在工业物联网(IIoT)、远程医疗、应急通信和深海无线通信等垂直场景中,语义通信有望实现局部部署^[44]。在这类场景中,带宽资源有限,传输需求以语义理解而非比特精确为目标,因此语义通信具有突出的应用优势^[21,40]。智能体之间的语义通信也将率先在这一阶段实现商用落地。然而,短期内语义通信不会大规模替代现有通信系统,仅作为传统通信技术的补充方案存在。

2) 中期(5~10年):6G标准框架嵌入语义通信特性。6G标准化工作将进入实质阶段^[47]。语义通信有望作为6G候选技术之一被纳入规范级标准化讨论^[33-34],但更可能是以部分特性(如任务导向的语义编码^[31,37])的形式嵌入6G标准,而非完整的全语义化网络框架。标准化工作将推动技术收敛,助力语义通信产业生态初步成型。

3) 远期(10~20年):全网范围规模化商用部署。随着核心技术持续成熟^[33-34]、标准体系逐步完善以及AI算力的大幅提升,语义通信有望实现更大规模的商用部署。然而,即便在这一阶段,语义通信也不太可能完全取代传统比特传输,该阶段更可能形成语义传输与比特传输并存的混合传输新格局^[1,35]。

7.3 关键使能因素

1) 大模型与多模态AI技术成熟^[13,43]。以GPT-4、Claude等为代表的大语言模型,以及多模态模型的快速发展,为语义理解与生成提供了强有力的技术支撑。随着AI模型能力的进一步提升^[14],语义编码器和解码器的性能有望持续改善。

2) 标准化体系稳步推进^[13,47]。第3代合作伙伴计划(3GPP)、国际电信联盟(ITU)等标准化组织逐步将语义通信纳入标准框架,旨在制定统一的技术规范和接口标准^[33]。标准化将使得不同厂商的设备能够互联互通,降低部署成本,促进产业生态形成。

3) 芯片算力持续提升。语义编码和解码涉及复杂的神经网络运算,对终端和基站的算力有较高要求。随着神经网络处理单元(NPU)、张量处理单元(TPU)等专用AI芯片的持续迭代,边缘设备的算力将不断提升^[31],语义通信的端侧部署将更具可行性。

4) 新型信道编码取得突破^[33-34]。语义信道编码是当前语义通信领域最薄弱的技术环节,需要在理论研究和工程应用两个方面取得突破。未来需要研究设计出复杂度可实现、性能可分析、工程可部署的语义信道编码方案^[36]。

综合来看,语义通信的商用化还需要一些时间积累。本文对语义通信在5~10年内进入商用阶段持乐观态度。从长期发展来看,语义通信具有重要战略价值,值得学术界和产业界持续深耕探索。

8 结论与展望

本文系统探讨了语义通信的商用化前景,并围绕三大核心问题展开分析。一是语义通信能否商用?答案是肯定的,但需循序渐进。语义通信在特定场景下已展现显著优势,可实现5~10倍带宽压缩,理论上具备重塑通信范式的潜力。然而,受技术成熟度、标准化进度等因素制约,其商用化将是渐进迭代过程,无法在短期内实现对现有体系的全面替代。二是何时能够商用?本文提出分阶段判断:近期(3~5年)将在工业物联网等垂直领域实现局部商用;中期(5~10年)有望融入6G标准体系;远期(10~20年)可能实现大规模部署。该判断基于当前技术发展与标准化节奏得出,后续可随关键技术突破与产业政策迭代进一步优化调整。三是商用价值何在?语义通信将推动通信从“比特传输”向“意义传递”跃迁,为带宽受限、语义密集、人机协作等场景带来革命性价值。

本文研究仍存在一定局限性:相关文献调研覆盖范围仍有待完善,对语义通信商用前景的阶段性的研判存在一定主观

倾向性,对产业政策、产业链协同、商业化运营机制等非技术关键影响因素的分析不够系统。为此,后续研究将重点关注语义通信与现有系统融合、通用语义知识库构建、语义信道编码、语义网络协议栈以及语义保真度量等,为语义通信的规模化商用落地提供理论支撑与技术保障。

参考文献

- [1] 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. DOI: 10.16451/j.cnki.issn1003-6059.201801008
- [2] 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. DOI: 10.1109/TSP.2021.3071210
- [3] Nokia Bell Labs. Nokia Bell Labs strengthens semantic communications research with KT collaboration [EB/OL]. (2025-03-03) [2026-04-15]. <https://www.nokia.com/blog/nokia-bell-labs-strengthens-semantic-communications-research-with-kt-collaboration/>
- [4] Samsung Research. 6G white paper: AI-native and sustainable communication [EB/OL]. [2026-04-15]. https://cdn.codeground.org/nsr/downloads/researchareas/Samsung_6G_Vision_2025.pdf
- [5] NTT DOCOMO. 5G evolution and 6G white paper v5.0 [EB/OL]. [2026-04-15]. https://www.docomo.ne.jp/english/binary/pdf/corporate/technology/whitepaper_6g/DOCOMO_6G_White_PaperEN_v5.0.pdf
- [6] 钟义信.信息科学原理[M].福州:福建人民出版社,1988
- [7] Zhong Y X. A theory of semantic information [J]. China communications, 2017, 14(1): 1-17. DOI: 10.1109/cc.2017.7839754
- [8] 钟义信.机制主义人工智能理论:一种通用的人工智能理论[J].智能系统学报,2018,13(1):2-18. DOI: 10.11992/tis.201711032
- [9] O' Shea T, Hoydis J. An introduction to deep learning for the physical layer [J]. IEEE transactions on cognitive communications and networking, 2017, 3(4): 563-575. DOI: 10.1109/TCCN.2017.2758370
- [10] 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). IEEE, 2018: 2326-2330. DOI: 10.1109/icassp.2018.8461983
- [11] Bourtsoulatz 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. DOI: 10.1109/TCCN.2019.2919300
- [12] Güler B, Yener A, Swami A. The semantic communication game [J]. IEEE transactions on cognitive communications and networking, 2018, 4(4): 787-802. DOI: 10.1109/TCCN.2018.2872596
- [13] 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. DOI: 10.1016/j.eng.2021.11.003
- [14] 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. DOI: 10.1109/COMST.2022.3223224
- [15] 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. DOI: 10.1109/MWC.101.2100269
- [16] 石光明,高大化.语义信息及其信号内容的数学刻画方法[J].中国科学:信息科学,2025,55(8):1950-1970. DOI: 10.1360/SSI-2025-0032
- [17] 石光明,高大化,杨雯曦,等.信号的语义刻画与度量[J].电子学报,2022,50(9):2068-2078. DOI: 10.12263/DZXB.20210238
- [18] 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. DOI: 10.1109/JSAC.2021.3087240
- [19] 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. DOI: 10.1109/jsac.2022.3180802
- [20] Wang B Y, Li R P, Zhu J H, et al. Knowledge enhanced semantic communication receiver [J]. IEEE communications letters, 2023, 27(7): 1794-1798. DOI: 10.1109/lcomm.2023.3274562
- [21] 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. DOI: 10.1109/TWC.2023.3240969
- [22] 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. DOI: 10.1109/TWC.2024.3386807
- [23] Shi G M, Xiao Y. An introduction to semantic communication and semantic-aware networking standardization for 6G [J]. GetMobile: mobile computing and communications, 2024, 28(3): 14-19. DOI: 10.1145/3701701.3701707
- [24] 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. DOI: 10.1109/MCOM.001.2001239
- [25] Xiao Y, Shi G M, Zhang P. Toward agentic AI networking in 6G: a generative foundation model-as-agent approach [J]. IEEE communications magazine, 2025, 63(9): 68-74. DOI: 10.1109/MCOM.001.2500005
- [26] Xiao Y, Sun Z J, Shi G M, et al. Imitation learning-based implicit semantic-aware communication networks: multi-layer representation and collaborative reasoning [J]. IEEE journal on selected areas in communications, 2023, 41(3): 639-658. DOI: 10.1109/JSAC.2022.3229419
- [27] 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. DOI: 10.11959/j.issn.2096-3750.2021.00209
- [28] 唐雪,许进,冯雨龙.基于熵函数的语义信息边界研究[J].中兴通讯技术,2025,31(6):61-69
- [29] 段玉聪.语义数学概论[M].西安电子科技大学出版社,2026
- [30] Dong C, Liang H T, Xu X D, et al. Semantic communication system based on semantic slice models propagation [J]. IEEE journal on selected areas in communications, 2023, 41(1): 202-213. DOI: 10.1109/jsac.2022.3221948
- [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. DOI: 10.1109/JSAC.2022.3191326
- [32] 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. DOI: 10.1109/TWC.2023.3265201
- [33] Ma S, Zhang C H, Qi H Y, et al. A theory for semantic channel coding with many-to-one source [J]. IEEE transactions on cognitive communications and networking, 2025, 11(6): 3618-

3627. DOI: 10.1109/tccn.2025.3544275
- [34] Ying J K, Qin Z J, Feng Y L, et al. Joint semantic-channel coding and modulation for token communications [J]. IEEE transactions on wireless communications, 2026, 25: 8179–8193. DOI: 10.1109/TWC.2025.3636174
- [35] 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. DOI: 10.1109/mwc.013.2200553
- [36] Wang Y H, Ma S, Wu Y L, et al. SITP: a high-reliability semantic information transport protocol without retransmission for semantic communication [PP/OL]. arXiv (2025–12–10) [2026–04–15]. <https://arxiv.org/abs/2512.09291>
- [37] Wang Y H, Guo S S, Deng Y Q, et al. Privacy-preserving task-oriented semantic communications against model inversion attacks [J]. IEEE transactions on wireless communications, 2024, 23(8): 10150–10165. DOI: 10.1109/TWC.2024.3369170
- [38] Yang M X, Gao D H, Xie F, et al. SG2SC: a generative semantic communication framework for scene understanding-oriented image transmission [C]//Proceedings of ICASSP 2024 – 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2024: 13486–13490. DOI: 10.1109/icassp48485.2024.10446594
- [39] Ren J K, Sun Y P, Du H Y, et al. Generative semantic communication: architectures, technologies, and applications [J]. Engineering, 2026, 56: 45–61. DOI: 10.1016/j.eng.2025.07.022
- [40] Shi G M. World-first kilometer-level underwater acoustic video transmission based on semantic communication [R]. 2025
- [41] Chang M K, Hsu C T, Yang G C. GenSC: generative semantic communication systems using BART-like model [J]. IEEE communications letters, 2024, 28(10): 2298–2302. DOI: 10.1109/LCOMM.2024.3450309
- [42] 鹏城实验室, 北京邮电大学, 琶洲实验室(黄埔), 等. 语义知识库白皮书 2.0 [R]. 2026
- [43] Shi G M, Lan N, Xie X M, et al. Rethinking intelligence in the era of large language models [J]. Journal of communications and information networks, 2026, 11(1): 39–50
- [44] 李东博, 王新宇, 尹志胜, 等. 星地一体化语义通信网络: 探索与展望 [J]. 中兴通讯技术, 2025, 31(5): 57–65. DOI: 10.12142/ZTETJ.202505009
- [45] Shi G M, Xiao Y. Requirements of semantic-aware networking in future networks [R]. 2023
- [46] 张黎明. 基于通信扩展定义的语义通信三层架构 [J]. 中兴通讯技术, 2024, 30(6): 92–99. DOI: 10.12142/ZTETJ.202406014
- [47] 3GPP TSG SA WG1, TR 22.870 (Release 20): Study on 6G use cases and service requirements [S]
- [48] Dai J C, Wang S X, Yang K, et al. Toward adaptive semantic communications: efficient data transmission via online learned

nonlinear transform source-channel coding [J]. IEEE journal on selected areas in communications, 2023, 41(8): 2609–2627. DOI: 10.1109/JSAC.2023.3288246

作者简介



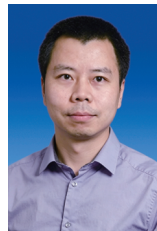
石光明, 鹏城实验室副主任, 西安电子科技大学教授、博士生导师, IEEE Fellow, IET Fellow, 中国电子学会会士, “长江学者”特聘教授, ACM中国杰出学者, IEEE CASS西安分会主席; 主要研究方向为语义通信、语义信号处理等; 获国家自然科学奖二等奖1项、省部级一等奖2项。



杨旻曦, 西安电子科技大学人工智能学院博士后; 主要研究方向为语义通信、生成式人工智能、图像视频压缩等。



高大化, 西安电子科技大学教授、博士生导师, 陕西省中青年科技创新领军人才, 陕西省青年科技新星; 主要研究方向为语义通信、智能信息处理、新体制智能成像、计算光谱成像等。



肖泳, 华中科技大学教授、IMT-2030 (6G) 网络智能组副组长、武汉市(中国-新加坡)智能与通信网络融合“一带一路”联合实验室主任; 主要研究方向为语义认知网络、智能体通信、边缘智能、通信网络博弈理论等。