

北斗短报文超低码率语音压缩与窄带通信传输技术研究

Research on Ultra-Low Bit Rate Speech Compression and Narrowband Communication Transmission Technology for Beidou Short Message System

高艺嘉,叶海纳,杨 杉,杜忠岩,谢梦楠,郑志扬(中国联通智能城市研究院,北京 100048)

Gao Yijia, Ye Haina, Yang Shan, Du Zhongyan, Xie Mengnan, Zheng Zhiyang (China Unicom Smart City Research Institute, Beijing 100048, China)

摘 要:

针对北斗短报文等窄带卫星通信单次传输容量受限与多模态数据传输需求之间的矛盾,提出一种融合传统参数化建模与轻量级神经网络重建的超低码率语音压缩算法,编码端采用线性预测编码参数化建模提取声道特征,通过残差矢量量化实现极低码率压缩;在解码端利用轻量级 ConvNeXt 神经网络完成语音重建。结合面向窄带卫星信道的智能分片与可靠传输机制,实现语音数据在窄带高时延信道下的可靠投递与有序重组,传输中有效保持语音可懂度和感知质量。该方案为窄带卫星多模态通信提供可行技术路径,拓展北斗短报文在应急通信、智慧海洋等场景的应用边界。

Abstract:

Aiming at the contradiction between the limited single transmission capacity of narrowband satellite communications such as Beidou short message and the demand for multimodal data transmission, an ultra-low bit rate speech compression algorithm integrating traditional parametric modeling and lightweight neural network reconstruction is proposed. At the encoding end, linear predictive coding (LPC) parametric modeling is used to extract vocal tract features, and residual vector quantization (RVQ) is employed to achieve extremely low bit rate compression. At the decoding end, a lightweight ConvNeXt neural network is utilized for speech reconstruction. Combined with an intelligent segmentation and reliable transmission mechanism tailored for narrowband satellite channels, speech data is reliably delivered and ordered reassembly over narrowband high-latency channels, while effectively maintaining speech intelligibility and perceptual quality during transmission. This scheme provides a feasible technical path for narrowband satellite multimodal communication and expands the application boundaries of Beidou short message in scenarios such as emergency communication and smart ocean.

Keywords:

Narrowband satellite communication; Beidou short message; Low bit rate speech compression; Hybrid speech codec

引用格式:高艺嘉,叶海纳,杨杉,等. 北斗短报文超低码率语音压缩与窄带通信传输技术研究[J]. 邮电设计技术,2026(7):31-36.

1 概述

在窄带卫星通信研究中,信道带宽的刚性约束与数据传输的可靠性保障始终是核心挑战,而富媒体信息在广域覆盖场景下的高效递送需求,则进一步放大了这一技术张力的复杂性。我国北斗短报文系统兼

关键词:

窄带卫星通信;北斗短报文;低码率语音压缩;混合式语音编解码

doi: 10.12045/j.issn.1007-3043.2026.07.006

文章编号:1007-3043(2026)07-0031-06

中图分类号:TN927

文献标识码:A

开放科学(资源服务)标识码(OSID):



具窄带卫星低功耗、广覆盖的物理层特性,聚焦应急通信^[1]、电力物联网^[2]、海洋渔业^[3]、林草畜牧、环境监测^[4]等垂直行业^[5],提供通导一体化双向短消息服务。随着应用场景从结构化文本向压缩语音、轻量化图片延伸,受限信道下多源异构数据的高效处理与可靠投递面临新的挑战^[6]。

语音通信的核心瓶颈在于极低信道速率与语音可懂度要求之间的矛盾,北斗短报文单次传输容量有

收稿日期:2026-06-02

限,须将语音压缩至极低码率,同时保持自然度和可懂度^[7]。

低码率语音编码是实现窄带卫星通信中语音业务传输的关键技术。近年来,国内外学者围绕极低码率下的语音质量与编码效率展开了大量研究。针对北斗短报文等窄带专用信道,敖振等人基于正弦激励线性预测(SEL P),提取线性预测系数、基音周期、清浊音等参数,研发了 600 bit/s SEL P 声码器,实现了语音编码与定位信息同传^[8];Liang Yutu 等人基于混合激励线性预测(MEL P),采用矢量量化压缩参数,并加入解码端后处理增强,在 1.2 kbit/s 低码率下实现了语音重建,通过对 LPC、基音周期、增益等多个参数联合量化,在极低码率下尽可能保留关键语音特征,为后续更先进的混合编码奠定了重要的技术基础^[9]。传统参数编码方法(如 LPC、SEL P)在 2.4 kbit/s 以下码率时^[10],合成语音自然度与抗噪性不足,且其通用性、重建自然度及对传输环境的鲁棒性在实际应用中面临诸多挑战。为突破这一瓶颈,混合式神经编解码器成为主流方向。Zheng Y Q 等人将传统参数编码器的粗量化比特流与神经声码器结合,在 1.1 kbit/s 下实现了优于 3 kbit/s 传统编码器的重建质量^[11];Zhen K 等人设计了可扩展的 CNN 自编码框架,支持单一模型灵活切换多种码率^[12]。与此同时,端到端神经编解码器在极低码率下展现出巨大潜力,如 Liu 等人利用大规模 Transformer 与有限标量量化,在 400~700 bit/s 下获得了顶尖质量^[13],但其模型规模大、计算资源要求高,难以在星载或终端侧直接部署。

综合来看,传统语音编码方法在计算复杂度方面

具有优势,但在低码率条件下会导致语音质量下降^[14];而基于深度学习的神经语音编解码方法虽能显著提升语音的自然度和流畅度,但模型规模和计算量大,难以直接部署于资源受限平台^[15]。现有研究在超低码率与低复杂度、高可懂度、自然度及鲁棒性之间仍存在权衡困境。对于窄带卫星非实时语音通信,须在保证语音编码压缩效率的前提下,提升低码率语音重建质量,同时控制模型规模和计算复杂度^[16]。

为此,本文提出一种融合传统参数化建模与轻量级神经网络重建的低码率语音压缩算法,结合轻量化调度与智能重组机制,实现窄带卫星信道下语音数据的高可靠传输与精准还原,为北斗短报文在应急通信等场景中的应用提供可行的技术路径。

2 轻量化超低码率语音压缩算法

2.1 混合式语音编解码系统架构

北斗短报文通信信道窄,单次可传输的字节数有限,无法直接发送原始语音波形。为此,结合低码率语音压缩编码技术,在发送端对语音进行分析、处理和多级量化,并将量化编码后的语音参数封装成短报文数据包,通过北斗终端经卫星链路发出;卫星接收后转发至地面监控中心,经北斗短报文网关及北斗短报文平台,最终传输至接收方终端。接收端将收到的压缩参数进行语音解码,经参数反量化、重构、特征增强、声学特征恢复等处理,最终输出语音信号。混合式语音编解码系统如图 1 所示。

本方案采用一种融合传统语音参数化建模与轻量级神经网络重建的混合式语音编码与解码架构。

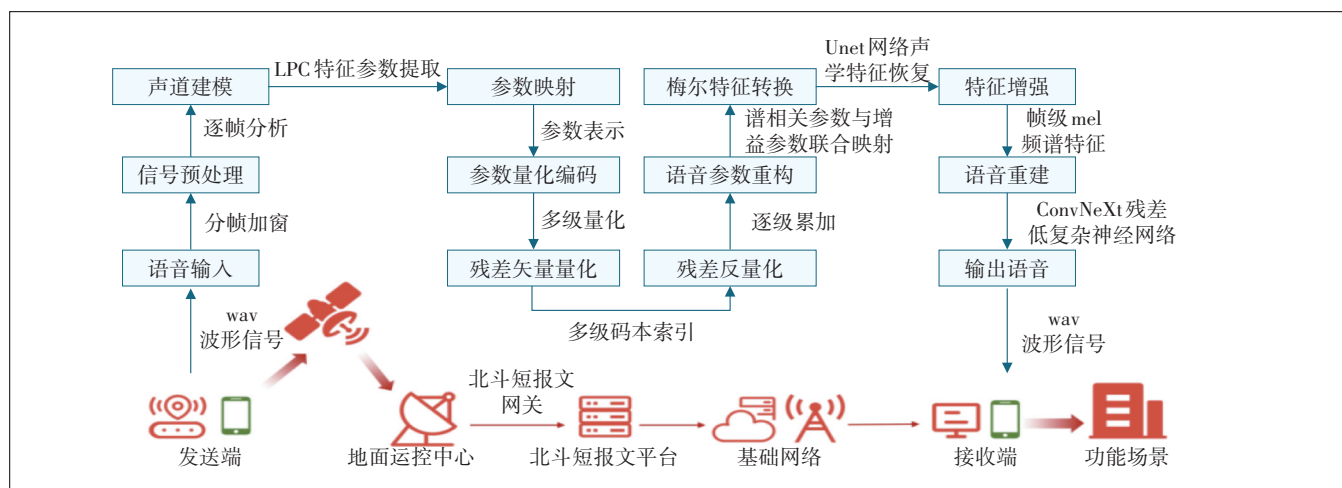


图 1 混合式语音编解码系统

该架构遵循“参数分析—低比特率编码—神经辅助重建”的设计思想。在编码端,通过传统参数化建模对语音信号进行逐帧分析,提取能够表征声道特性和语音结构的参数表示,并采用残差矢量量化的编码方式对关键参数进行编码压缩,有效且显著地降低传输数据量。在解码端,系统对接收到的索引参数进行反量化与重构,在此基础上引入轻量级神经网络对重构参数得到的声学特征进行细节修复与重建,以补偿传统参数化建模和量化过程在低码率条件下带来的信息损失。

2.2 LPC 参数化建模及残差矢量量化的编码算法

编码端对输入语音信号进行分析处理,将其转换为适于低比特率传输或存储的离散参数索引信息。

编码端通过传统参数化建模对语音信号进行逐帧分析,基于线性预测编码(LPC)等参数化建模提取能够表征声道特性和语音结构信息的参数,优先保留基频轨迹、谱包络/主要共振峰等决定语音可懂度与特征信息的要素,其核心功能包括语音分析、参数提取、参数变换以及参数量化编码等步骤(见图2)。最后,采用残差矢量量化的编码方式将连续参数映射为稳定、可控的码本索引,对关键参数进行编码压缩。通过对输入语音进行LPC特征提取与残差矢量量化,有效降低传输数据量,实现极低码率下的高鲁棒性表示。

2.2.1 语音分析与LPC特征提取

基于LPC的基本假设,用 P 阶滤波器近似声道模型,得到:

$$X(z) = \frac{G}{1 + \sum_{i=1}^P a_i z^{-i}} E(z) \quad (1)$$

其中, $X(z)$ 表示语音信号的 z 变换, a_k 表示预测系

数,描述了信道的共振峰。该线性预测系数能够有效刻画语音的短时谱包络特性,是语音声道特性的紧凑表示。

编码端接收时域语音信号 $x(n)$,并对所接收的语音信号进行分帧与加窗处理,得到语音帧序列:

$$x_m(n) = x(n+mH) \cdot w(n), n=0, 1, 2, \dots, N-1 \quad (2)$$

其中, n 为帧长, m 表示帧索引, H 为帧移, $w(n)$ 为窗函数。

对每一帧采用线性预测分析,建立LPC模型:

$$x_m(n) = -\sum_{i=1}^P a_i^{(m)} x[n-k] + e_m[n] \quad (3)$$

其中, $a_k^{(m)}$ 为第 m 帧对应的线性预测系数, P 为预测阶数, $e_m[n]$ 为残差激励信号。

在获得线性预测系数后,编码端将线性预测系数变换为谱相关参数,以提高参数表示的稳定性与量化鲁棒性:

$$I_m = T(a_m) \quad (4)$$

其中, $a_m = [a_1^{(m)}, \dots, a_P^{(m)}]$, T 为从预测系数到谱相关参数的变换函数。此外,对每帧语音计算能量增益,与谱相关参数组合,得到该帧的参数表示:

$$z_m = \{I_m, g_m\} \quad (5)$$

其中, I_m 为该语音帧的参数向量, g_m 为对应的增益参数。

2.2.2 基于RVQ的参数量化编码

采用残差矢量量化(RVQ)方法对语音参数进行离散化编码,通过多个量化阶段对输入向量进行逐级逼近。首先对原始参数向量进行第一阶段量化,得到初始量化结果;随后对量化残差进行进一步量化,直至达到预设的量化级数。

对输入参数向量 z_m ,定义初始残差为:

$$r_m^{(0)} = z_m \quad (6)$$

设共有 L 级残差量化,第 l 级码本为:

$$C^{(l)} = \{c_k^l\}_{k=1}^{K_l} \quad (7)$$

其中, K_l 为第 l 级码字数。

在第 l 级量化阶段,通过在第 l 个子码本中根据最小平方欧式距离准则:

$$d(a, b) = \|a - b\|_2^2 \quad (8)$$

搜索与当前残差最匹配的码字:

$$i_m^{(l)} = \operatorname{argmin}_{i \in \{1, \dots, K_l\}} d(r_m^{(l-1)}, c_i^l) \quad (9)$$

并将对应码字作为该级量化结果:

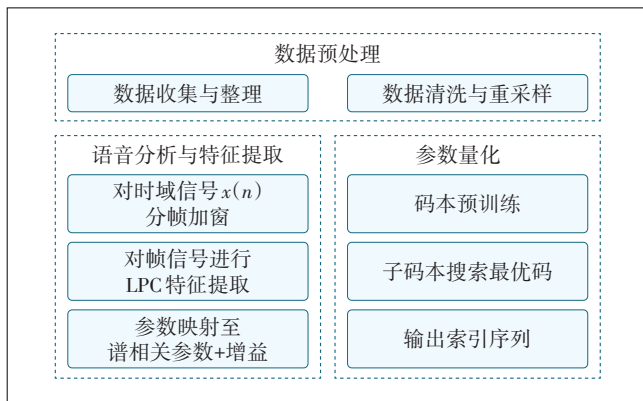


图2 轻量化超低码率语音压缩算法编码端设计

$$\hat{r}_m^{(l)} = c_{i_m}^{(l)} \quad (10)$$

随后更新残差:

$$r_m^{(l)} = r_m^{(l-1)} - \hat{r}_m^{(l)} \quad (11)$$

上述过程依次进行,直至完成第 l 级量化。第 m 帧语音参数的编码结果由多级码本索引组成:

$$i_m = \{i_m^{(1)}, i_m^{(2)}, \dots, i_m^{(l)}\} \quad (12)$$

索引序列 i_m 作为编码端输出,用于后续的传输或解码处理。

2.3 轻量级神经网络重构的解码算法

解码端用于对编码端输出的参数索引进行解析与恢复,并在此基础上生成重建语音信号。系统首先对接收到的索引参数进行反量化与重构,在此基础上引入轻量级神经网络对重构参数得到的声学特征进行细节修复与重建,以补偿传统参数化建模和量化过程在低码率条件下带来的信息损失。解码端包括参数反量化、声学特征构建、特征增强以及语音重建等处理步骤(见图3)。该算法通过将传统模型的可解释性与神经网络的非线性表征能力有机结合,有效补足高频细节与自然度损失,从而在低比特率条件下最大限度地恢复原始语音的声学特性和感知质量。

2.3.1 参数反量化与残差重构

解码端首先接收编码端传输的参数索引序列,然后根据索引 i_m ,在各级子码本中查找对应的码字,并对其逐级累加,以重构量化后的语音参数向量:

$$\hat{z}_m^{(l)} = \sum_{l=1}^L c_{i_m}^{(l)} \quad (13)$$

通过上述残差反量化过程,解码端获得用于语音重建的参数估计值:

$$z_m = \{I_m, g_m\} \quad (14)$$

其中, I_m 表示重构后的谱相关参数, g_m 表示对应的增益参数。

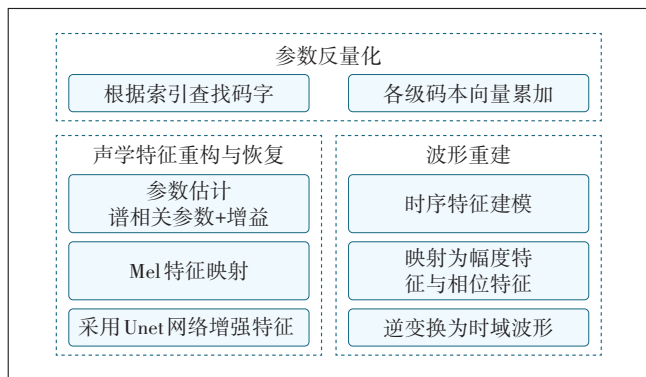


图3 轻量化超低码率语音压缩算法解码端设计

2.3.2 声学特征构建与语音重建

在获得参数重构结果后,系统将谱相关参数与增益参数联合映射至感知相关的特征空间——Mel频谱特征空间:

$$M_m = \gamma(z_m) \quad (15)$$

其中, z_m 表示谱相关参数与增益参数的组合, γ 表示从谱相关参数到Mel特征空间的变换操作, M_m 表示Mel谱。为补偿参数量化和参数化建模过程中引入的误差,解码端引入声学特征恢复模块,该模块利用Unet网络对构建的声学特征进行进一步处理,有效恢复了语音细节信息,提升重建语音的感知自然度。

解码端利用修复后的帧级梅尔频谱,通过多层堆叠的ConvNeXt残差模块构成的低复杂度神经网络,重建出时域语音信号。

3 基于智能分片重组的窄带语音消息可靠传输机制

在窄带通信环境下,语音消息传输常面临带宽受限、信道质量波动大、丢包率高等问题,严重影响消息的完整性。为此,本文提出一种基于智能分片重组的窄带语音消息可靠传输机制。

其核心为“动态分包—智能重组”:发送端将长报文拆分为带有总包数、分包编号的多个短报文,接收端依据协议头部缓存重组。针对卫星信道固有的丢包问题,本文引入可靠传输控制机制,包括超时重传、选择重传请求等,并利用内存数据库对分包数据进行临时缓存管理。

3.1 数据动态拆分

数据动态拆分主要实现初始化与阈值匹配、数据拆分判断与封装。根据通信卡等级或预设分级确定拆分阈值^[17],计算原始数据总字节数与当前有效阈值确定分包数量。将原始数据拆分为多个数据包进行传输,且阈值预留标识字节,以保障后续数据包的标识封装空间。通过分片标识与顺序控制,为每个分片的数据包头添加包序号、总包数及CRC32校验码等字段,以支持接收端正确排序、组装与校验。

3.2 数据智能重组

接收端依据“包头标识+包序号”对分片数据进行排序,当检测到缺失包时触发重传机制,确保重组成功率达100%。数据智能重组流程划分为数据接收、数据整合、分类处理3个阶段。其中,在分类处理阶段,

根据完整数据类型执行不同的处理逻辑。

3.3 数据重传与校验

数据重传和校验是窄带卫星通信网络可靠传输中的核心机制之一^[18]。本文采用CRC校验算法验证数据完整性,并通过分包回执机制触发重传,同时将重传队列优先级设置为高于新任务优先级。

在数据发送结束后,发送端发送回执申请,接收端将接收的数据通过回执确认信息将接收信息返回到发送端。当发送端接收到的回执确认信息标识有

数据丢失或在预设的重传超时时间内未收到接收端返回的对应确认帧时,触发数据补发重传,将未收到的分包数据进行补发。系统会将补发重传的数据插入到传输前置队列,确保优先发送。

3.4 窄带语音消息传输机制

窄带语音消息的可靠传输是通过分包和重传机制来确保大数据量的完整到达^[19]。窄带语音消息处理及传输流程如图4所示。

采集端采集音频,经自定义压缩算法压缩后,再

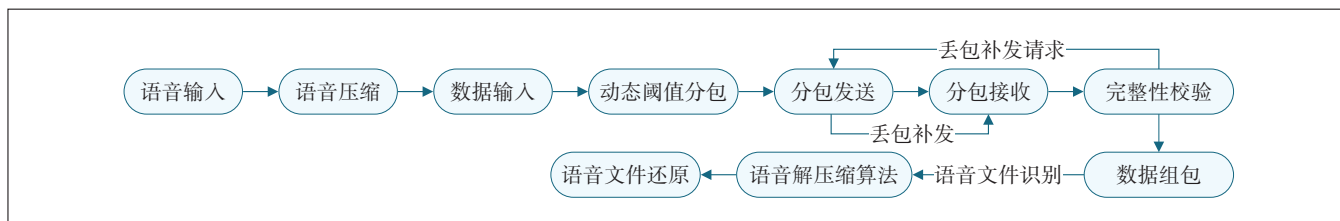


图4 窄带语音消息处理及传输流程

根据压缩后的数据包大小以及北斗信道最大应用层协议数据单元长度切分为固定长度的数据片,并将每片封装为独立短报文帧后传输到传输队列,经北斗射频模块发送至卫星。接收端在接收到分包数据后,逐帧接收并校验物理层,根据分片序列号与总包数判断丢包情况;若超时未收齐,则生成含缺失序列号的反馈报文反向请求重传。收齐全部分片后,按序列号排序拼接并确认为语音文件,调用语音解压缩算法进行解压,生成语音文件重组数据。

4 仿真试验分析

为验证语音编码压缩算法在实际语音数据上的重建效果,对同一语音样本在编码与解码前后的信号进行对比分析。测试数据来源于实际场景,采集端通过麦克风采集时长为16 s的音频(单声道,采样率8 kHz,采样位深16 bit),对采集的语音信号进行预处理,并完成特征提取与量化编码,同时根据原始WAV文件大小与编码后索引数据大小计算压缩比与压缩率:

$$\text{Compression Ratio} = \frac{\text{originalSize}}{\text{compressedSize}} \quad (16)$$

随后进入解码阶段,对编码索引进行反量化与波形重建,生成最终输出语音文件,以实录数据为例,图5所示为原始音频的时域波形图和语谱图,图6所示为编码压缩重建后音频的时域波形图和语谱图。

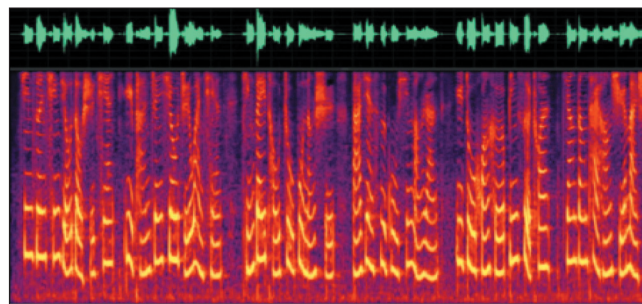


图5 原始传输音频

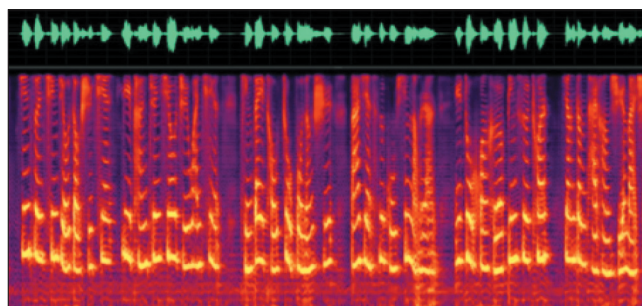


图6 编码压缩重建后的音频

从频谱分析结果可以看出,编码压缩重建后的语音在整体时频结构上与原始音频保持高度一致。重建语音能够有效保留基频轨迹和主要共振峰结构,中低频能量分布稳定,从而保证了良好的语音可懂度。尽管在高频细节上存在一定程度的平滑和能量衰减现象,但该现象符合低比特率语音编码的典型特征,

并未引入明显的结构性失真。整体而言,该编码方案在压缩效率与重建质量之间达成了较为合理的平衡。

5 结论与展望

本文针对北斗短报文等窄带卫星通信单次传输容量极小、难以直接传输语音这一核心瓶颈,提出了一套融合传统参数化建模(LPC特征提取与残差矢量化)与轻量级神经网络重建(采用ConvNeXt残差网络进行梅尔谱增强)的超低码率语音压缩算法,并结合动态分包、智能重组与选择重传的可靠传输机制,解决了在极低信道速率下保持语音可懂度与可靠投递大数据量的问题,为北斗短报文从纯文本向多模态通信演进提供了可行的技术路径。

本文研究成果已完成了可在Windows平台、Linux平台、Android平台上直接调用的算法库文件,接口统一、功能可靠,随时可以嵌入到各类后端或终端软件中使用,实现了从算法模型到工程部署的完整闭环。并在应急通信、能源电力、工业物联网、水利生态等多个项目中发挥关键价值:在青海玉树高原无移动网络覆盖区域,实现常态数据上报、控制指令上传下达;在水利水库项目中,以北斗短报文技术为核心,凭借其“全天时、无盲区”通信优势,实现雨水工情自动测报、汛期紧急沟通和异地协同调度,水库自动监测覆盖率从不足50%跃升至90%以上,监测效率提升85%,监测准确率达99%,保障了水库安全运行。

随着通导融合的深度演进与人工智能技术的持续赋能,窄带卫星通信必将在应急通信、行业应用、消费电子等领域发挥更加重要的作用,面向未来万物互联、空天地一体化的通信需求,仍有许多关键技术值得进一步探讨研究,如文本、语音、图像等多模态数据的联合编码与协同传输机制、构建适应多样化业务需求的弹性传输框架、根据信道质量实时调整压缩比与传输策略等。

参考文献:

- [1] 申翔,吴培仁. 基于北斗卫星导航系统的海上应急搜救系统[J]. 指挥控制与仿真,2018,40(6):43-49.
- [2] 李静,龙强,臧志斌,等. 北斗卫星导航系统在电力行业的应用研究[J]. 电力信息与通信技术,2022,20(10):87-97.
- [3] 姜斌,徐振华,葛振营,等. 基于北斗短报文的海洋观测实时通信系统的设计与研发[J]. 海洋科学,2023,47(8):68-74.
- [4] 王春芳,陈永涛,李春来,等. 基于北斗卫星的预警信息发布技术研究及实现[J]. 应用气象学报,2014(3):375-384.

- [5] 刘艳,刘晓莉,石书祝,等. 北斗短报文电网状态监测及应急通信系统[J]. 自动化与仪表,2015,30(1):43-47.
- [6] 李志炜,傅军,刘李欢,等. 基于北斗短报文的远程图像通信系统设计[J]. 导航定位与授时,2019,6(6):119-123.
- [7] 王晶,徐亮,陈晓娇,等. 基于神经网络的低码率语音编码技术研究综述[J]. 信号处理,2024,40(12):2261-2280.
- [8] 敖振,李凤,马婧,等. 基于北斗导航星座的语音与定位同传通信系统[J]. 西北工业大学学报,2020,38(5):1010-1017.
- [9] Liang Y T, Zhang J, Shi Z, et al. Research on coding and enhancement algorithm for very low-rate speech compression of beidou short message [C]//2022 4th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP). Hangzhou: IEEE, 2022:1105-1110.
- [10] 张怀峰,张浩. 一种应用于北斗三号RSMC的低码率语音传输方法:CN202311218774.7[P]. 2023-09-21.
- [11] Zheng Y Q, Xiao L, Tu W P, et al. CQNV: a combination of coarsely quantized bitstream and neural vocoder for low rate speech coding [C]//Interspeech 2023. Dublin: Interspeech, 2023:171-175.
- [12] Zhen K, Sung J, Lee M S, et al. Scalable and efficient neural speech coding: a hybrid design [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2022, 30: 12-25.
- [13] Parker J D, Smirnov A, Pons J, et al. Scaling transformers for low-bitrate high-quality speech coding [C]//The Thirteenth International Conference on Learning Representations. Rio de Janeiro: ICLR, 2025.
- [14] 江文斌. 面向复杂场景的极低码率语音编解码和语音增强关键技术研究[D]. 上海:上海交通大学,2018.
- [15] Chen X P, Wang R, Khalilian-gourtani A, et al. A neural speech decoding framework leveraging deep learning and speech synthesis [J]. Nature Machine Intelligence, 2024, 6(4):467-480.
- [16] 王敬凯,秦董洪,白凤波,等. 语音识别与大语言模型融合技术研究综述[J]. 计算机工程与应用,2025,61(6):53-63.
- [17] 李政霖,介婧,柴佳辉,等. 面向低算力设备的改进轻量化语音识别模型[J]. 计算机工程与设计,2025,46(10):2969-2977.
- [18] 王振业,江勋林,李明. 基于北斗的装备保障信息长报文可靠通信研究[J]. 计算机与数字工程,2016,44(1):88-94.
- [19] 李文金,苏凯雄. 基于存储管理的北斗报文传输协议设计与应用[J]. 微型机与应用,2015,34(24):63-65.

作者简介:

高艺嘉,工程师,硕士,主要从事5G+北斗通导融合、窄带卫星通信、产品研发及行业应用工作;叶海纳,正高级工程师,博士,主要从事5G+北斗通导一体化方向技术攻关、产品研发和应用示范工作;杨杉,正高级工程师,博士,主要从事5G+北斗通信导航融合技术研究、产品研发及应用示范工作;杜忠岩,正高级工程师,硕士,主要从事空天地一体化网络、5G+北斗通导融合、产品研发和应用示范工作;谢梦楠,高级工程师,硕士,主要从事通信网络、产品研发及行业应用工作;郑志扬,高级工程师,学士,主要从事新一代移动通信、AI智能体应用研究工作。