中国科学 : 信息科学 2026 年 第 56 卷 第 10 期: 2663–2677
SCIENTIA SINICA Informationis
面向 6G 的通感算可重构无线网络专题 • 论文

基于夹持天线的通感一体化波束成形: 多尺度注意力深度学习方法

张栖瀚[1], 盛冯浩[2], 张海洋[2,3], 杨龙祥[2*]

1. 南京邮电大学波特兰学院, 南京 210012
2. 南京邮电大学通信与信息工程学院, 南京 210003
3. 东南大学移动通信全国重点实验室, 南京 210096
* 通信作者. E-mail: yanglx@njupt.edu.cn

收稿日期: 2026–04–30; 修回日期: 2026–08–07; 接受日期: 2026–09–04; 网络出版日期: 2026–10–10

国家自然科学基金 (批准号: 62471246)、东南大学移动通信全国重点实验室开放课题 (批准号: 2025D09) 和四川省自然科学基金 (批准号: 2025ZNSFSC0514) 资助项目

摘要 通感一体化 (ISAC) 被认为是第六代移动通信的关键使能技术之一, 而夹持天线系统 (PASS) 凭借其天线位置沿波导灵活可调的独特优势, 为通感联合波束成形提供了传统固定阵列所不具备的额外空间自由度. 本文研究 PASS-ISAC 系统中给定天线部署下的发射波束成形设计, 在满足通信信干噪比和感知波束图增益的约束下最小化发射功率. 针对该问题的非凸性和实时求解需求, 提出了一种多尺度注意力波束成形网络 MSA-BFNet-ISAC. 该网络通过引入通信信道跨域融合、多尺度空间特征提取、感知目标角度编码和导向向量编码的四路联合特征处理架构, 结合优化引导预训练与无监督微调协同的两阶段混合训练策略, 在通信约束、感知约束和功率目标之间实现高效优化. 仿真实验表明, 所提方法在推理速度提升约两个数量级的情况下, 性能接近迭代优化算法; 与固定位置天线阵列及不同 PASS 布局的对比实验进一步验证了 PASS 架构的增益, 系统性消融实验则量化了网络各模块的贡献与复杂度开销, 验证了 PASS 在 ISAC 场景下的应用潜力.

关键词 夹持天线系统, 通感一体化, 波束成形优化, 深度学习, 多尺度注意力网络

1 引言

通感一体化被认为是未来第六代移动通信 (6G) 的关键使能技术 [1,2]. 在过去, 通信与雷达各自独立运行, 通感一体化 (ISAC) 将二者整合在同一硬件平台上, 共享频谱和波形, 同时完成高速数据传输和高精度目标感知 [3], 从而能够缓解频谱资源的紧张同时降低硬件成本. 在 ISAC 系统设计中, 如何通过优化发射端波束成形矢量来实现通信和感知性能的折中是一个关键的问题 [4,5]. 例如, Liu 等人 [4] 率先提出了双功能雷达 – 通信系统的最优波形设计框架, Stoica 等人 [6] 在波形多样性约束下分析了感知波束图设计的理论界, 此后文献 [3,5] 将多用户 MIMO 通信与 MIMO 雷达的波束成形设计纳入统一的优化框架, 借助半定松弛 [7] 等凸优化工具进行求解. 这些方法虽然取得了较好的理论结果, 但计算复杂度随天线规模呈指数增长, 在大规模天线和实时性要求较高的场景中, 实际部署仍有困难.

引用格式: 张栖瀚, 盛冯浩, 张海洋, 等. 基于夹持天线的通感一体化波束成形: 多尺度注意力深度学习方法. 中国科学: 信息科学, 2026, 56: 2663–2677, doi: 10.1360/SSI-2026-0201
Zhang Q H, Sheng F H, Zhang H Y, et al. Integrated sensing and communication beamforming for pinching antenna surface systems: a multi-scale attention deep learning approach. Sci Sin Inform, 2026, 56: 2663–2677, doi: 10.1360/SSI-2026-0201

© 2026《中国科学》杂志社

天线架构层面的创新也在为 ISAC 系统打开新的设计空间. 近年来, 以可移动天线 (movable antenna, MA) [8] 和六维可移动天线 (six-dimensional movable antenna, 6DMA) [9] 为代表的柔性可重构天线架构, 通过灵活调整天线的空间位置和姿态来主动改善信道条件, 已成为 6G 物理层的重要研究方向. 夹持天线系统 (PASS) 是 Ding 等人 [10] 近期提出的新架构, 其核心思路是在介质波导表面通过局部夹持机制动态激活天线辐射点. 与传统均匀线阵或平面阵不同, PASS 中的天线单元可以沿波导灵活调整位置, 信号先在波导内部传播, 再从各激活位置辐射到自由空间. 天线位置可沿波导灵活调整, 为波束成形带来了额外的空间自由度. 与 MA/6DMA 通常在孔径内波长量级范围移动或旋转天线不同, PASS 可以在数米乃至数十米的尺度上重构天线位置, 从而直接缩短天线与用户的距离; 多个天线共享同一波导射频源, 硬件复杂度因此大幅降低; 加之波导内传播损耗远低于自由空间, PASS 架构尤其适合亚太赫兹乃至太赫兹频段的部署 [10,11]. 另外, 文献 [10] 还从理论上证明了 PASS 相比传统固定阵列在创建视距链路和降低路径损耗方面的独特优势, 并分析了多天线共用单根波导时的性能.

PASS 与 ISAC 的结合有望带来比传统固定阵列更优的性能 – 功率折中, 但相关研究仍刚起步. Zhang 等人 [12] 将 PASS 引入 ISAC 场景, 提出了基于交替优化的联合波束成形方案; Li 等人 [13] 建立了多波导 PASS-ISAC 系统的克拉美 – 罗下界最小化框架; 文献 [14] 从克拉美 – 罗下界的角度揭示了 PASS 天线灵活部署对均匀定位精度的提升. 这些工作揭示了 PASS-ISAC 系统的潜力, 但均依赖传统迭代优化方法, 计算开销较大且在信道快速变化时难以做到实时响应.

深度学习近年来在无线通信物理层设计中已有不少成功应用 [15,16]. 文献 [15] 较早验证了用深度神经网络直接学习从信道到波束成形映射的可行性; Xia 等人 [16] 利用卷积神经网络 (CNN) 结构和上下行对偶等领域知识, 设计了面向多输入单输出系统 (MISO) 下行波束成形优化的深度学习框架. 在 ISAC 场景下, 文献 [17] 利用图神经网络建模多用户间的干扰关系. 在物理层优化之外, 数据驱动方法在更上层的无线资源管理中同样表现出色, 例如 Zhang 等人 [18] 基于深度强化学习联合优化了车联网中的信息年龄与能耗, 展示了学习类方法应对复杂无线优化问题的应用潜力. 同时, Transformer 架构 [19] 依靠自注意力机制捕获长程依赖, 在波束成形设计等物理层优化任务中展现出了强大的建模能力 [20]. 然而, 面向 PASS-ISAC 系统的数据驱动波束成形方法目前尚未研究.

综合以上分析, 本文将深度学习方法引入 PASS-ISAC 波束成形设计, 所做的主要工作和贡献概括如下.

(1) 在 PASS 通信信道模型的基础上, 引入包含波导相位响应的感知导向向量, 建立了联合信干噪比 (SINR) 约束和感知波束图增益约束的 PASS-ISAC 系统发射功率最小化问题.

(2) 提出了 MSA-BFNet-ISAC 网络, 其核心是通信信道跨域融合、多尺度空间特征提取、目标角度正弦编码和导向向量编码的四路联合特征处理架构, 并以 Transformer 编码器、门控残差单元和压缩激励模块构成后续的深度特征处理, 通过两阶段混合训练策略实现端到端波束成形映射.

(3) 设计了优化引导预训练与无监督微调协同的两阶段混合训练策略, 以迭代优化器的解作为监督标签引导网络快速学会波束成形结构, 再通过物理层无监督损失进一步压缩功率, 实现了对非凸双约束问题的高效训练.

(4) 在多组参数扫描实验中, 从 SINR 阈值、感知增益阈值和天线数量三个维度, 系统对比了所提方法与半定松弛 (SDR) 迭代优化、ZF-ISAC 和时分 ISAC 三种基准, 验证了所提优化设计的有效性. 同时, 通过发射波束图分析、与固定位置天线阵列及不同 PASS 布局的对比实验, 以及网络各模块与角度编码阶数的消融实验, 分别验证了感知约束的实际效果与 PASS 架构的性能增益, 并量化了网络各模块的贡献与复杂度开销.

本文后续内容安排如下: 第 2 节建立 PASS-ISAC 系统模型; 第 3 节给出问题建模; 第 4 节详述 MSA-BFNet-ISAC 网络设计; 第 5 节报告仿真结果; 第 6 节总结全文.

2　系统模型

2.1　PASS 发射端架构

考虑如图 1 所示的 PASS-ISAC 下行传输场景. 发射端配有 N 根介质波导, 每根波导上配置 L 个夹持天线

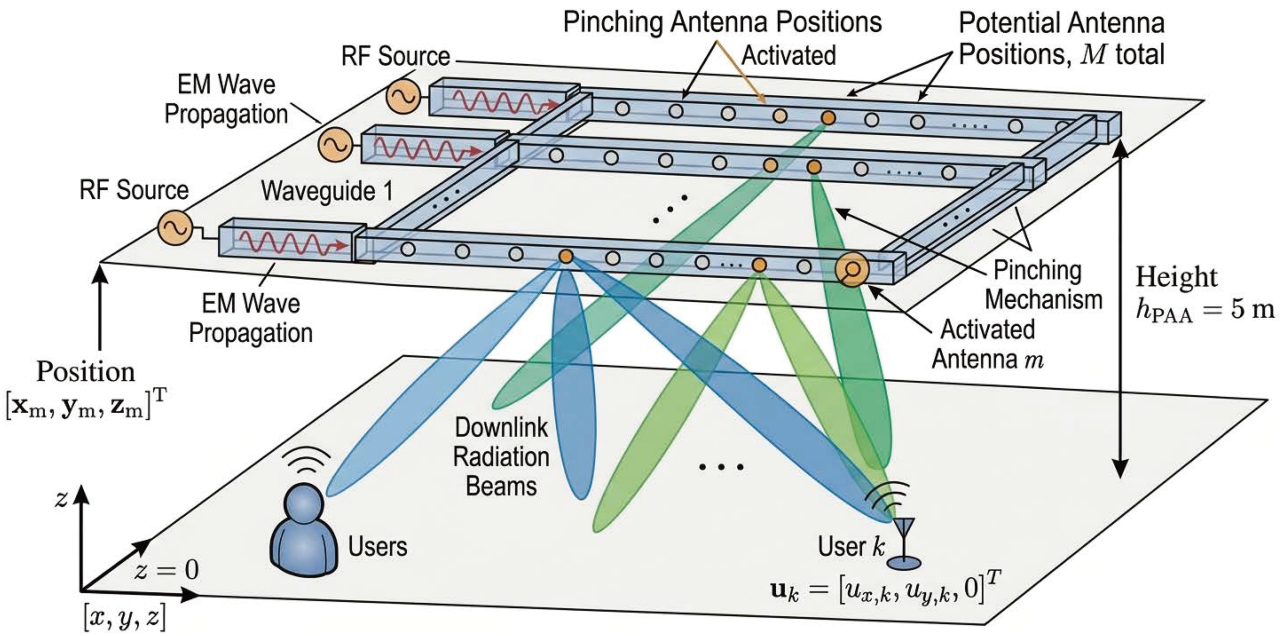

图 1　(网络版彩图) PASS-ISAC 系统架构示意图. PASS 发射端由多根介质波导组成, 每根波导上可灵活配置多个可编程天线, 信号经波导传播后辐射至地面用户和感知目标.

Figure 1　(Color online) PASS-ISAC system architecture. The PASS transmitter consists of multiple dielectric waveguides, each with flexibly configured pinching antennas. Signals propagate through waveguides and radiate toward ground users and sensing targets.

(PA), 天线总数 $M = N \times L$. 系统工作于亚太赫兹频段, 载波频率 $f_c = 150$ GHz, 对应波长 $\lambda = c/f_c = 2$ mm. 天线阵列部署在距地面 $h_{\mathrm{PAA}} = 5$ m 的天花板上, 下行同时服务 K 个单天线通信用户和探测 T 个感知目标, 感知性能以发射波束图增益衡量 [4,6]. 本文假设夹持天线位置在部署阶段按均匀间距给定, 并在波束成形时间尺度上保持固定, 优化变量为发射波束成形矩阵; 天线位置与波束成形的联合优化留待后续工作.

第 m 个 PA 位于三维坐标 $\mathbf{p}_m = (x_m, y_m, h_{\mathrm{PAA}})$. 馈入波导的信号传播至第 m 个 PA 时产生相位偏移

$$g_m = \exp\left(-\mathrm{j}\kappa\, n_{\mathrm{eff}}\, d_m^{\mathrm{wg}}\right), \tag{1}$$

其中 $\kappa = 2\pi/\lambda$ 是自由空间波数, $n_{\mathrm{eff}} = 1.4$ 是波导有效折射率, d_m^{wg} 是馈源点沿波导到第 m 个 PA 的距离. 传统阵列中每个天线都有独立的射频链路, PASS 则是所有天线共享同一波导馈源, 信号经波导色散到达各 PA, 因此 g_m 包含了一层确定性的相位旋转 [10].

2.2　通信信道模型

第 k 个用户位于 $\mathbf{u}_k = (u_{k,x}, u_{k,y}, 0)$. PASS 工作于亚太赫兹频段, 在室内场景中视距分量占主导 [21], 因此本文采用自由空间信道模型. 所有 PA 到用户 k 的有效信道向量可表示为波导相位与自由空间传播的级联:

$$\mathbf{h}_k = \mathbf{G}_{\mathrm{diag}} \cdot \boldsymbol{\alpha}_k \in \mathbb{C}^{M\times 1}, \tag{2}$$

$\mathbf{G}_{\mathrm{diag}} = \mathrm{diag}(g_1, g_2, \ldots, g_M)$ 是波导相位对角矩阵, 自由空间信道的第 m 个分量为

$$[\boldsymbol{\alpha}_k]_m = \frac{\sqrt{\beta}}{d_{m,k}} \exp(-\mathrm{j}\kappa\, d_{m,k}), \tag{3}$$

$d_{m,k} = \|\mathbf{p}_m - \mathbf{u}_k\|$ 为 PA m 到用户 k 的三维欧氏距离, $\beta = (\lambda/(4\pi))^2$ 是 Friis 公式对应的功率域路径损耗参考系数.

2.3　感知导向向量模型

对于阵列法线方向偏角 θ_t 处的远场感知目标, 发射导向向量定义为

$$\mathbf{a}(\theta_t) = \left[\exp\left(-\mathrm{j}\kappa(x_1\cos\theta_t + y_1\sin\theta_t)\right), \ldots, \exp\left(-\mathrm{j}\kappa(x_M\cos\theta_t + y_M\sin\theta_t)\right)\right]^{\mathrm{T}}. \tag{4}$$

叠加上波导相位, 有效导向向量为

$$\tilde{\mathbf{a}}(\theta_t) = \mathbf{G}_{\mathrm{diag}} \cdot \mathbf{a}(\theta_t). \tag{5}$$

式 (2) 的通信信道体现的是近场特性 —— 每个 PA 到同一用户的距离不同, 信道分量同时含有幅度和相位的变化; 而式 (4) 的导向向量基于远场平面波假设, 只涉及各天线间的相位差. 这种近场 – 远场的双重结构是 PASS-ISAC 建模的独特之处, 也是后续网络设计中需要对两类输入分别编码再融合的物理根源.

2.4　信号模型与性能度量

发射信号由线性波束成形生成:

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{w}_k s_k = \mathbf{W}\mathbf{s}, \tag{6}$$

$\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_K] \in \mathbb{C}^{M\times K}$ 为波束成形矩阵, $s_k \sim \mathcal{CN}(0,1)$ 为各用户的独立数据符号, 总发射功率 $P_{\mathrm{tx}} = \|\mathbf{W}\|_F^2$.

用户 k 的 SINR 为

$$\mathrm{SINR}_k = \frac{|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k|^2}{\sum_{j\neq k}|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_j|^2 + \sigma^2}. \tag{7}$$

感知目标方向 θ_t 的波束图增益 [6] 定义为

$$P(\theta_t) = \tilde{\mathbf{a}}^{\mathrm{H}}(\theta_t)\mathbf{W}\mathbf{W}^{\mathrm{H}}\tilde{\mathbf{a}}(\theta_t) = \|\mathbf{W}^{\mathrm{H}}\tilde{\mathbf{a}}(\theta_t)\|^2. \tag{8}$$

式 (8) 的物理意义是发射信号在方向 θ_t 上的辐射功率, 即发射波束图在该方向的取值 [6].

3　问题建模

在上述建模基础上, 本文通过优化 PASS-ISAC 系统的发射波束成形矢量, 在满足通信用户信干噪比和感知波束图增益需求的基础上最小化发射功率. 因此, PASS-ISAC 系统的功率最小化问题可形式化为

$$\begin{aligned} \mathcal{P}: \quad & \min_{\mathbf{W}\in\mathbb{C}^{M\times K}} \ \|\mathbf{W}\|_F^2 \\ & \text{s.t.} \qquad \mathrm{SINR}_k \geqslant \gamma_{\min}, \quad k = 1, \ldots, K, \\ & \qquad\quad\ \ P(\theta_t) \geqslant \Gamma_s, \quad t = 1, \ldots, T, \end{aligned} \tag{9}$$

其中, $\gamma_{\min}$ 是各用户的最低 SINR 门限, Γ_s 是各感知目标方向的最低波束图增益.

问题 $\mathcal{P}$ 的求解是一个非凸优化问题, 数学上很难进行求解, 现有处理该类非凸优化问题的方法是半定松弛 (SDR). 加入感知约束后, 可行域是纯通信问题可行域的子集, 因而在两问题均可行时, 最优发射功率不低于纯通信问题的最优值, 二者之差表征满足感知约束所需的附加功率.

本文的比对基准采用标准的提升变量技巧 [7], 令 $\mathbf{X}_k = \mathbf{w}_k\mathbf{w}_k^{\mathrm{H}} \succeq 0$ 并放松秩一约束, 问题 $\mathcal{P}$ 松弛为如下凸 SDP:

$$\begin{aligned} & \min_{\{\mathbf{X}_k\}_{k=1}^K} \ \sum_{k=1}^{K} \mathrm{tr}(\mathbf{X}_k) \\ & \text{s.t.} \qquad \mathrm{tr}(\mathbf{H}_k\mathbf{X}_k) \geqslant \gamma_{\min}\left(\sum_{j\neq k}\mathrm{tr}(\mathbf{H}_k\mathbf{X}_j) + \sigma^2\right), \ \forall k, \end{aligned} \tag{10}$$

$$\sum_{k=1}^{K}\mathrm{tr}(\mathbf{A}_t\mathbf{X}_k)\geqslant\Gamma_s,\quad\forall t,$$
$$\mathbf{X}_k\succeq 0,\quad\forall k,$$

其中, $\mathbf{H}_k=\mathbf{h}_k\mathbf{h}_k^{\mathrm{H}}$, $\mathbf{A}_t=\tilde{\mathbf{a}}(\theta_t)\tilde{\mathbf{a}}^{\mathrm{H}}(\theta_t)$.

上述 SDP 问题是凸优化问题, 可以通过内点法直接求解. 若所得最优解恰好满足秩一条件, 则松弛紧致, 可由特征值分解恢复 $\mathbf{w}_k$; 否则需要高斯随机化做近似恢复 [7]. SDR 的求解复杂度为 $\mathcal{O}(M^{6.5})$ [22]. 因此, 当 M 较大时难以实时运行, 这正是本文引入数据驱动方法的出发点.

4　MSA-BFNet-ISAC 网络设计

4.1　整体架构

MSA-BFNet-ISAC 是一个端到端的映射网络, 输入通信信道 $\mathbf{H}\in\mathbb{C}^{K\times M}$、感知导向向量 $\tilde{\mathbf{A}}\in\mathbb{C}^{T\times M}$ 和目标角度 $\boldsymbol{\theta}\in\mathbb{R}^T$, 直接输出波束成形矩阵 $\mathbf{W}\in\mathbb{C}^{M\times K}$, 整体结构如图 2 所示. 网络的设计出发点在于: 通信信道和感知信息在物理上具有不同的结构特征, 应分别提取特征后再融合, 而非简单拼接为一个长向量.

本文提出的四路并行特征处理是指网络内部四条并行的特征变换支路. 通常情况下, 网络的原始输入只有通信信道与感知目标信息两类: (1) 跨域特征融合与多尺度特征提取两条支路共享同一通信信道输入 $\mathbf{H}$, 分别从实部 – 虚部耦合关系与多尺度空间相关性两个互补角度对其进行差异化编码; (2) 目标角度编码与导向向量编码两条支路均源自感知目标信息. 其中, 前者仅使用标量角度 $\boldsymbol{\theta}$ 以显式提供目标方向的几何先验, 后者使用由 PA 位置与波导相位共同决定的有效导向向量 $\tilde{\mathbf{A}}$ 以刻画特定阵列布局下的响应结构. 四条支路对共享输入的差异化处理提高了特征的多样性, 其各自的贡献将在 5.3 小节的消融实验中定量评估. 各组成部分介绍如下.

4.2　网络结构设计

4.2.1　通信信道特征提取

通信信道特征提取分为跨域特征融合和多尺度特征提取两个环节. 在跨域特征融合环节, 复数信道 $\mathbf{H}$ 的实部与虚部分别包含不同的物理信息. 跨域融合模块先用两个独立的 MLP 分支分别编码实部和虚部, 再用一个可学习的门控向量 $\mathbf{g}$ 自适应地决定融合比例:

$$\begin{aligned}\mathbf{f}_{\mathrm{real}}&=\mathrm{MLP}_r(\mathrm{Re}(\mathbf{H})),\mathbf{f}_{\mathrm{imag}}=\mathrm{MLP}_i(\mathrm{Im}(\mathbf{H})),\\ \mathbf{g}&=\varsigma\left(\mathbf{W}_g[\mathbf{f}_{\mathrm{real}};\mathbf{f}_{\mathrm{imag}}]\right),\\ \mathbf{f}_{\mathrm{cross}}&=\mathrm{MLP}_f([\mathbf{f}_{\mathrm{real}};\mathbf{f}_{\mathrm{imag}}])\odot\mathbf{g}+\mathbf{f}_{\mathrm{real}}\odot(\mathbf{1}-\mathbf{g}),\end{aligned}\tag{11}$$

其中, $\varsigma(\cdot)$ 为 Sigmoid 激活函数, $\odot$ 为逐元素乘积. 网络可以根据输入信道的统计特性动态调节实部与虚部的贡献比例, 而非以固定权重合并.

在此基础上, 多尺度特征提取器采用三支路 MLP 金字塔结构 [23], 分别由 1∼3 层 MLP 构成, 对应捕获局部天线关联、中等范围空间相关和全局阵列模式三种尺度的特征. 三路输出拼接后经一个融合层压缩为 $D_h=256$ 维表征 $\mathbf{f}_{\mathrm{multi}}$.

4.2.2　感知特征编码

感知特征编码包含目标角度编码和导向向量编码两部分. 在目标角度编码方面, 目标角度 $\boldsymbol{\theta}=[\theta_1,\ldots,\theta_T]$ 通过正弦谐波编码转化为高维表示:

$$\mathbf{e}_{\mathrm{angle}}=\left[\sin(k\theta_t),\cos(k\theta_t)\right]_{k=1,\ldots,4;\,t=1,\ldots,T},\tag{12}$$

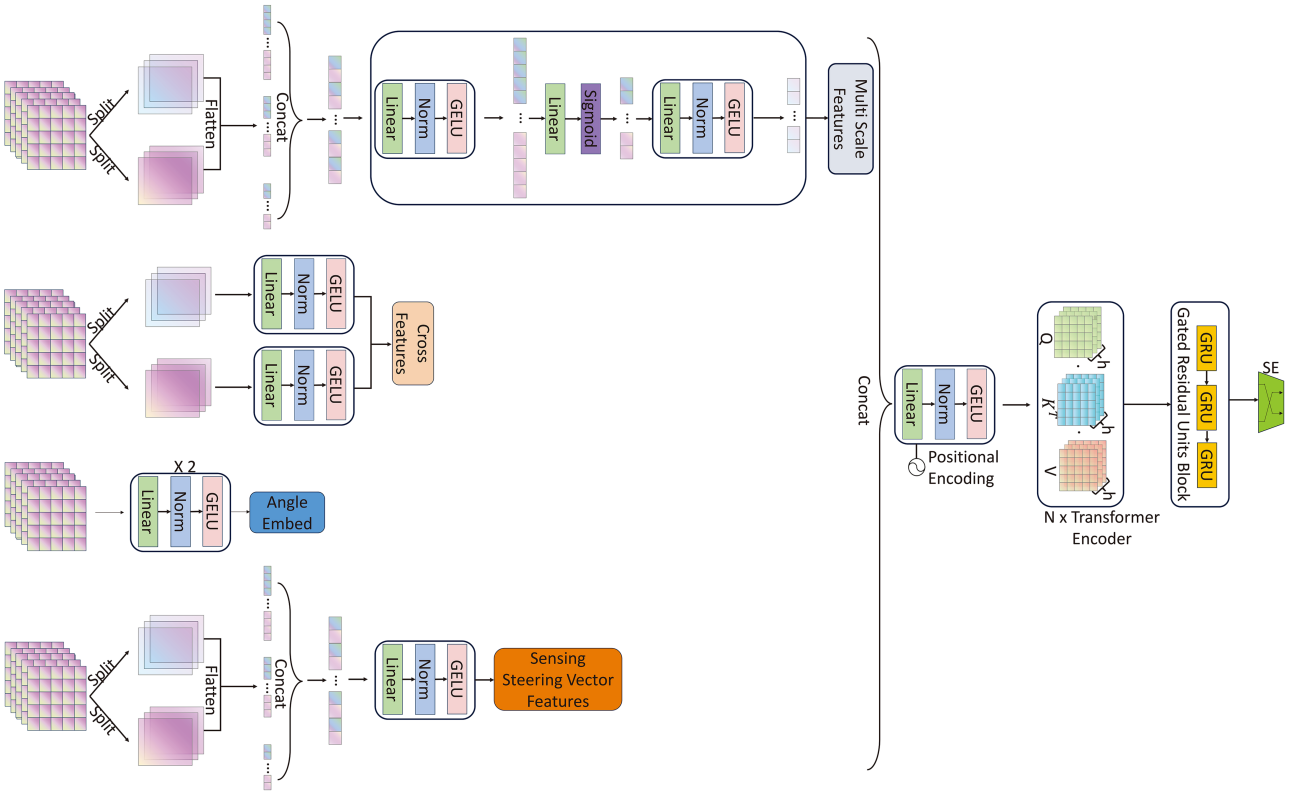

图 2　(网络版彩图) MSA-BFNet-ISAC 网络整体架构. 网络包含四条并行特征提取路径: 信道跨域融合、多尺度特征提取、目标角度编码和导向向量编码, 融合后经 Transformer 编码器、门控残差单元和 SE 注意力模块处理, 最终由波束成形输出头生成 W.

Figure 2　(Color online) Overall architecture of MSA-BFNet-ISAC. The network comprises four parallel feature extraction pathways: cross-domain channel fusion, multi-scale feature extraction, target-angle encoding, and steering-vector encoding. The fused features are processed by a Transformer encoder, gated residual units, and an SE attention module, followed by a beamforming output head that generates $\mathbf{W}$.

再经两层 MLP 映射到 D_h 维, 得到 $\mathbf{f}_{\text{angle}}$. 正弦编码天然保留了角度的周期性和相邻角度的连续性, 相比直接输入原始角度值, 该编码在小样本下收敛更快.

式 (12) 中 θ_t 以弧度为单位, 取值范围为 $[-\pi/2, \pi/2]$. 该编码虽然在形式上借鉴了 Transformer 的正弦位置编码 [19], 但存在两点本质区别, 特此说明. 其一, Transformer 位置编码的输入是离散的整数位置索引, 并采用几何级数间隔的多频率以唯一表示任意长序列中的位置; 而此处输入为连续实值角度, 故采用整数倍频率 k 的谐波形式, 以保证编码函数关于角度连续且以 2π 为周期. 其二, 由于 $\theta_t \in [-\pi/2, \pi/2]$ 不超过基频分量 $(k = 1)$ 的单个周期, $(\sin\theta_t, \cos\theta_t)$ 对本身已足以唯一确定 θ_t, 编码的唯一性不依赖多频率组合; 更高阶分量 $k = 2, 3, 4$ 的作用是以更高的角频率放大相邻角度之间的表示差异, 增强网络对小角度差的分辨能力. 关于频率阶数的选取: 阶数越高编码维度越大、对角度差异越敏感, 但也会引入冗余特征并增加首层计算量.

在导向向量编码方面, 有效导向向量 $\tilde{\mathbf{A}} \in \mathbb{C}^{T\times M}$ 的实部与虚部拼接后, 由独立的编码器网络处理:

$$\mathbf{f}_{\text{sens}} = \text{MLP}_s\big([\text{Re}(\tilde{\mathbf{A}}); \text{Im}(\tilde{\mathbf{A}})]\big) \in \mathbb{R}^{D_h}. \tag{13}$$

单独设置这条路径是因为: 仅靠角度编码无法完整反映 PASS 中导向向量的结构 —— 各 PA 位置不等间距, 同一角度在不同天线布局下对应的导向向量差异显著.

表 1　两阶段混合训练策略.

Table 1　Two-stage hybrid training strategy.

Stage	Epochs	Loss function	Description
1	1∼800	$\mathcal{L}_{\text{sup}}$	Supervised pre-training using iterative-optimizer solutions as labels
2	801∼3300	$\mathcal{L}_{\text{power}}+\lambda_1\mathcal{L}_{\text{SINR}}+\lambda_2\mathcal{L}_{\text{sensing}}$ $+\lambda_3(\mathcal{L}^2_{\text{SINR}}+\mathcal{L}^2_{\text{sensing}})$	Unsupervised fine-tuning with power minimization as the primary objective

4.2.3　四路融合与深度处理

四条路径的输出拼接后经线性层统一到 D_h 维:

$$\mathbf{f}_{\text{fused}}=\text{Linear}\big([\mathbf{f}_{\text{cross}};\mathbf{f}_{\text{multi}};\mathbf{f}_{\text{angle}};\mathbf{f}_{\text{sens}}]\big). \tag{14}$$

融合后的特征依次经过三个处理阶段. 首先是 2 层 4 头、隐层维度 256 的 Transformer 编码器 [19], 自注意力机制负责捕获特征间的全局依赖, 每层含多头自注意力子层和前馈网络子层, 各子层配有残差连接与层归一化.

其次是 3 层门控残差单元, 每层的计算过程为

$$\text{output}=\text{LN}\big(\mathbf{x}+\varsigma(\mathbf{W}_g\mathbf{h})\odot\text{FC}(\mathbf{h})\big), \tag{15}$$

其中, $\mathbf{h}=\text{GELU}(\text{FC}(\text{LN}(\mathbf{x})))$ 为中间特征. 门控向量 $\varsigma(\mathbf{W}_g\mathbf{h})$ 值域为 $[0,1]$, 逐元素控制残差分量的通过强度, 避免信息在多层堆叠中无差别地传递或衰减 [24].

最后是压缩激励注意力模块 [25], 结构如图 3 所示. 该模块先对特征向量做全局平均池化, 再经降维 – 激活 – 升维-Sigmoid 的瓶颈结构生成逐通道的注意力权重, 对各特征维度做自适应校准. 其结构紧凑, 参数量仅为主网络的 1/16 左右, 在视觉和信号处理任务中均能有效提升特征的判别力 [25]. 上述各深度处理模块以及四路特征支路的独立贡献, 将在 5.3 小节通过逐一移除或替换的消融实验加以量化.

4.2.4　波束成形输出头

经上述处理后得到 D_h 维特征, 由 3 层 MLP 映射为 $M\times K\times 2$ 维实值输出, 分别对应 $\mathbf{W}$ 的实部和虚部:

$$\mathbf{W}=\big(\mathbf{W}_{\text{real}}+\text{j}\,\mathbf{W}_{\text{imag}}\big)\times|s_{\text{power}}|, \tag{16}$$

s_{power} 是可学习的标量缩放参数, 初始化为 0.1, 作用是让训练初期的输出功率处于合理量级, 避免梯度爆炸或消失.

4.3　两阶段混合训练

本文采用优化引导预训练结合无监督微调的两阶段混合训练策略, 其核心思路是: 先以传统迭代优化的解作为监督标签让网络快速学会波束成形的基本结构, 再以物理层指标作为无监督目标进一步压缩功率, 参数见表 1.

4.3.1　阶段 1: 优化引导预训练

对训练集中的每个信道实现, 首先调用最小均方误差 (MMSE) 初始化结合感知补偿与序列最小二乘规划 (SLSQP) [26] 精调的迭代优化器求解问题 $\mathcal{P}$, 得到参考波束成形矩阵 $\mathbf{W}^{(b)}_{\text{ref}}$. 预训练阶段以均方误差作为损失:

$$\mathcal{L}_{\text{sup}}=\frac{1}{B}\sum_{b=1}^{B}\|\mathbf{W}^{(b)}-\mathbf{W}^{(b)}_{\text{ref}}\|_F^2, \tag{17}$$

该阶段使网络输出迅速逼近迭代优化器的解, 为后续微调提供高质量初始化. 实验中预训练轮数为 800, 优化器采用 AdamW [27], 其中学习率为 5×10^{-4}, 权重衰减为 10^{-5}.

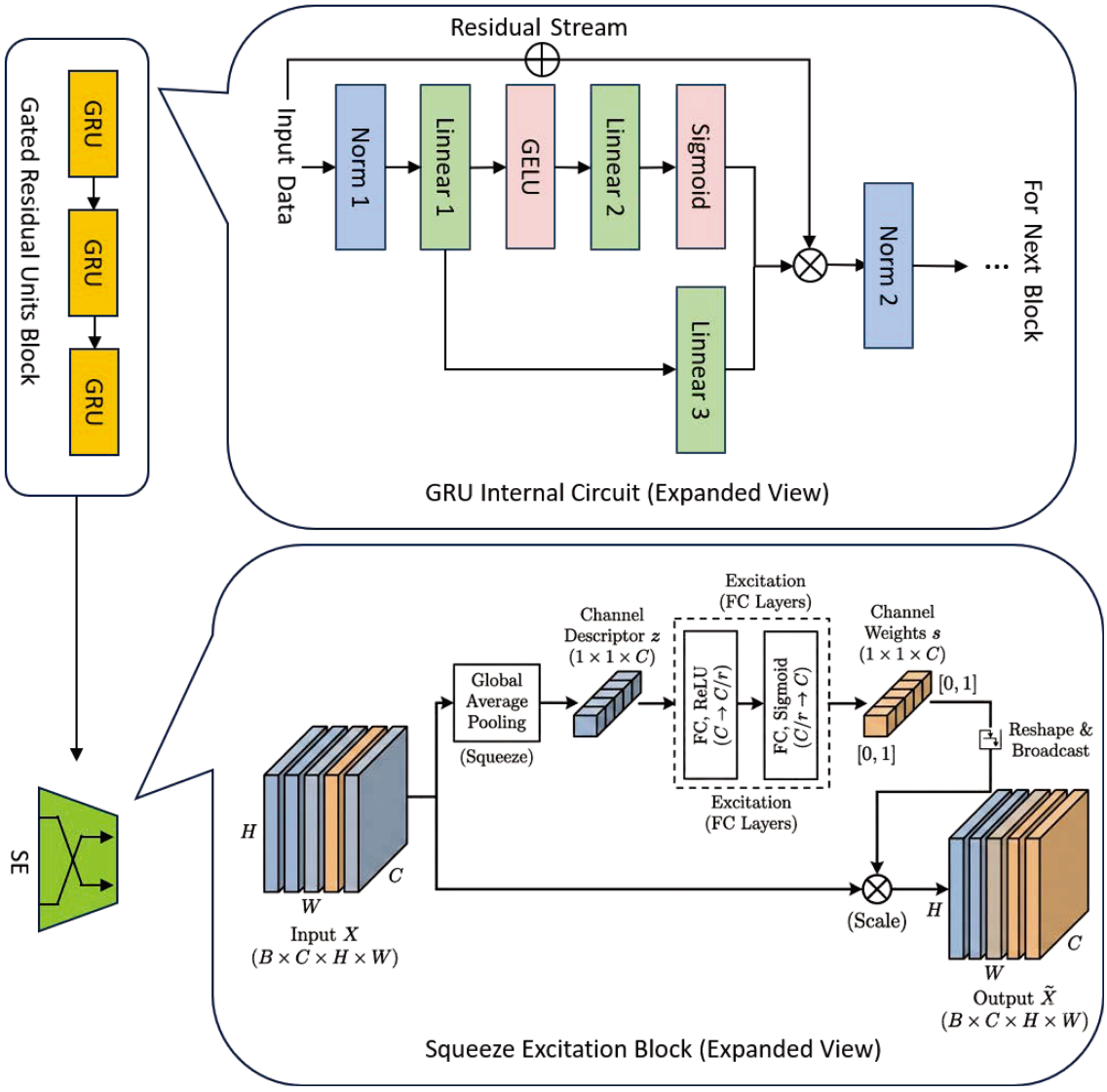

图 3　(网络版彩图) 压缩激励注意力模块结构. 通过全局池化和瓶颈全连接层生成逐通道注意力权重.

Figure 3　(Color online) Structure of the squeeze-and-excitation attention module. Channel-wise attention weights are generated through global pooling and bottleneck fully connected layers.

4.3.2　阶段 2: 无监督微调

在预训练基础上, 切换为物理层驱动的无监督损失:

$$\mathcal{L} = \mathcal{L}_{\text{power}} + \lambda_1 \mathcal{L}_{\text{SINR}} + \lambda_2 \mathcal{L}_{\text{sensing}} + \lambda_3 \left(\mathcal{L}_{\text{SINR}}^2 + \mathcal{L}_{\text{sensing}}^2\right), \tag{18}$$

各项为

$$\begin{aligned} \mathcal{L}_{\text{power}} &= \frac{1}{B}\sum_{b=1}^{B} 10\log_{10}\left(\|\mathbf{W}^{(b)}\|_F^2/\sigma^2\right), \\ \mathcal{L}_{\text{SINR}} &= \frac{1}{B}\sum_{b=1}^{B} \max\left(0,\, \gamma_{\min}^{\text{dB}} - \text{SINR}_{\min}^{(b),\text{dB}}\right), \\ \mathcal{L}_{\text{sensing}} &= \frac{1}{B}\sum_{b=1}^{B} \max\left(0,\, \Gamma_s^{\text{dB}} - P_{\min}^{(b),\text{dB}}\right). \end{aligned} \tag{19}$$

式 (19) 中的样本级最小指标构造如下: 对批次内第 b 个样本, 先按式 (7) 计算其全部 K 个用户的 SINR 并取最小值 $\mathrm{SINR}_{\min}^{(b)} = \min_{k=1,\dots,K} \mathrm{SINR}_k^{(b)}$, 再按式 (8) 计算其全部 T 个目标方向的波束图增益并取最小值 $P_{\min}^{(b)} = \min_{t=1,\dots,T} P^{(b)}(\theta_t)$; 上标 dB 表示相应量的分贝值. 以最小值汇聚多用户、多目标约束, 等价于对每个样本仅惩罚违反最严重的一条约束, 且当且仅当损失项为零时该样本的全部 $K+T$ 条约束均得到满足. $\mathcal{L}_{\mathrm{power}}$ 以噪声功率 σ^2 为参考将发射功率化为无量纲的相对 dB 值, 它与以 1 mW 为参考的绝对功率仅相差常数 $10\log_{10}(\sigma^2/1\,\mathrm{mW})$, 因而不改变梯度方向与最优解, 仅用于使损失各项处于相近的数值量级. SINR 和感知惩罚项均采用铰链形式, 只在约束违反时产生非零梯度. 微调阶段固定 $\lambda_1 = \lambda_2 = 100$, $\lambda_3 = 20$. 式 (18) 中的一次铰链项在约束违反时提供恒定幅值的梯度, 二次项则提供与违反量成正比的梯度, 二者结合使违反较大的样本得到更强的修正力度, 从而加速约束收敛. 训练时将 SINR 目标设为 $1.05\gamma_{\min}$, 为实际推理预留 5% 的余量. 微调轮数为 2500, 学习率 2×10^{-4}.

两阶段策略的优势在于: 预训练阶段为网络提供了接近可行的初始权重, 避免了从随机初始化出发时约束满足困难、功率居高不下的问题; 微调阶段的物理层损失则可以在保持高约束满足率的前提下进一步压缩功率, 使最终性能超越迭代优化基准. 优化器采用 AdamW [27], 学习率调度采用余弦退火重启 [28], 梯度范数裁剪阈值 1.0, 批量大小 32. 网络中所有 MLP 层和 Transformer 编码器均配置 Dropout 率 0.05 以防止过拟合.

4.4　复杂度分析

MSA-BFNet-ISAC 的可训练参数量由两部分构成: 一是与系统规模相关的输入/输出投影层, 约 $\mathcal{O}\big((K+T)MD_h\big)$; 二是与系统规模无关的 Transformer 编码器、门控残差单元等稠密层, 约 $\mathcal{O}(D_h^2)$. 在 $D_h = 256$, $M = 20$, $K = T = 2$ 的默认配置下, 参数总量约为 5.3×10^6. 设 Transformer 编码器层数为 L_{T}、输入 token 序列长度为 n_s, 则单样本前向推理的计算复杂度为

$$\mathcal{O}\Big(\underbrace{(K+T)MD_h}_{\text{四路特征提取}} + \underbrace{L_{\mathrm{T}}\big(n_s^2 D_h + n_s D_h^2\big)}_{\text{Transformer 编码器}} + \underbrace{D_h^2}_{\text{门控残差与 SE 模块}} + \underbrace{MKD_h}_{\text{波束成形输出头}}\Big), \tag{20}$$

其中第一项来自四路特征提取中与输入维度 MK (通信信道支路)、TM (导向向量支路) 和 T (角度支路) 成正比的首层 MLP 投影, 第二项为标准 Transformer 的复杂度, 即自注意力子层 $\mathcal{O}(n_s^2 D_h)$ 与前馈子层 $\mathcal{O}(n_s D_h^2)$ [19]. 在本文实现中, 四路特征融合后的 D_h 维全局表征作为单个 token 输入 Transformer, n_s 为不随系统规模增长的常数, 因此式 (20) 随系统参数 M, K, T 仅线性增长, 整体由 $\mathcal{O}(D_h^2)$ 量级的稠密层运算主导. 在 GPU 上单样本推理时间约为 2∼5 ms.

5　仿真结果与分析

5.1　仿真设置

系统默认参数见表 2. 本文所用的数据假定, 用户均匀分布于服务区域, 两个感知目标分别位于 $-30°$ 和 $45°$ 方向. 所有神经网络实验在 PyTorch 框架下完成, GPU 型号为 NVIDIA RTX 系列.

5.2　基准方案

为公平评估, 本文实现了三种基准方案加以对比. 第一种为 ISAC-SDR 迭代优化基准, 先用 MMSE 波束成形初始化并进行感知增益补偿, 再以 SLSQP [26] 在 SINR 与感知双约束下精调功率, 代表传统数值优化方法的性能上界. 第二种为 ZF-ISAC, 先用迫零 (ZF) 波束成形 [29] 消除用户间干扰并满足 SINR 约束, 再对功率进行均匀缩放以满足感知增益约束; 该方案虽有闭式解, 但通信与感知分步设计, 感知所需的功率缩放导致 SINR 明显高于所需门限, 功率效率较低. 第三种为时分 ISAC (TD-ISAC), 通信与感知各占 50% 时间, 分别独立设计波束成形, 等效通信速率减半, 总功率约为联合设计的 2 倍.

表 2　仿真默认参数.

Table 2　Default simulation parameters.

Parameter	Symbol	Value
Carrier frequency	f_c	150 GHz
Number of users/waveguides	$K = N$	2
Antennas per waveguide	L	10
Antenna height	h_{PAA}	5 m
Waveguide effective refractive index	n_{eff}	1.4
Waveguide length/width	S_x/S_y	10/30 m
Minimum SINR	$\gamma_{\min}$	20 dB
Sensing target angles	θ_1, θ_2	$-30°$, $45°$
Sensing gain threshold	Γ_s	10 dB
Noise power	σ^2	-80 dBm

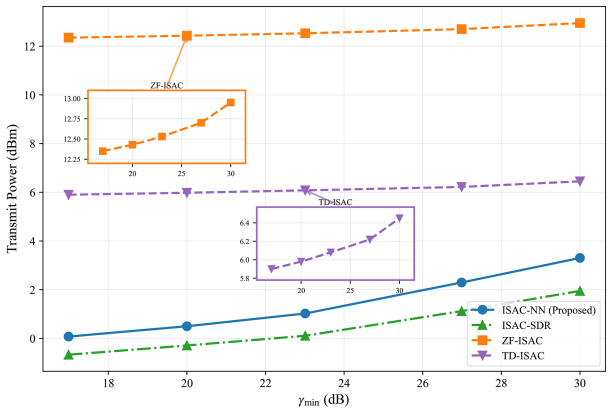

图 4　(网络版彩图) 发射功率随 SINR 阈值 $\gamma_{\min}$ 的变化. $\Gamma_s = 10$ dB, $L = 10$, $S_x = 10$ m.

Figure 4　(Color online) Transmit power vs. SINR threshold $\gamma_{\min}$. $\Gamma_s = 10$ dB, $L = 10$, $S_x = 10$ m.

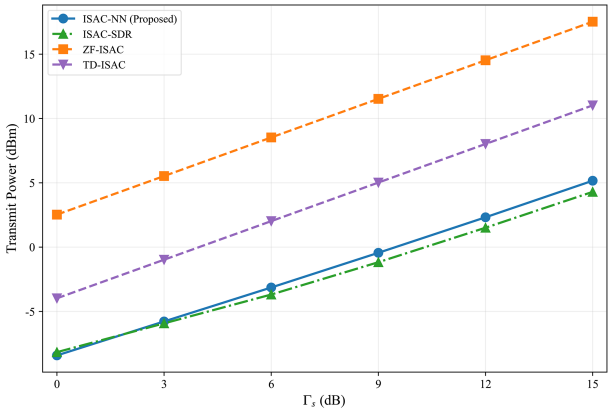

图 5　(网络版彩图) 发射功率随感知增益阈值 Γ_s 的变化. $\gamma_{\min} = 20$ dB.

Figure 5　(Color online) Transmit power vs. sensing threshold Γ_s. $\gamma_{\min} = 20$ dB.

后续图中, “ISAC-NN” 指本文所提 MSA-BFNet-ISAC 方法, “ISAC-SDR” 指上述 MMSE 初始化结合 SLSQP 精调的迭代优化基准, “ZF-ISAC” 为迫零加功率缩放的启发式基线, “TD-ISAC” 为时分基线.

5.3　实验结果与讨论

图 4, 考虑在 $\Gamma_s = 10$ dB, $L = 10$, $S_x = 10$ m 的参数设置下, 研究了发射功率随 SINR 阈值的变化. 从图 4 可见, 本文提出的算法, 明显优于 ZF-ISAC 和 TD-ISAC 算法. 例如在 $\gamma_{\min} = 20$ dB 的时候, 发射功率比 TD-ISAC 小 5.7 dBm, 比 ZF-ISAC 小 12.2 dBm. 同时从图 4 可以看出, 所提方法所需的发射功率接近于最优的 ISAC-SDR 方法. 值得指出的是, 所提方法在推理时间仅需 3 ms, 显著快于 ISAC-SDR 方法.

接下来分析感知约束对功率的影响. 固定 $\gamma_{\min} = 20$ dB, 扫描 Γ_s 从 0 到 15 dB, 结果如图 5 所示. 从仿真中可以看出, ISAC-NN 的功率从 $\Gamma_s = 0$ dB 时的 -8.42 dBm 单调递增至 15 dB 时的 5.16 dBm, 约 13.6 dB 的增幅直观地反映了通信 – 感知之间的功率折中. 在全部 Γ_s 取值下, ISAC-NN 与 ISAC-SDR 的差距保持在不到 1 dB, 且均显著低于 TD-ISAC 约 4∼6 dB.

为了直观检验感知约束的实际效果, 图 6 给出了默认配置 $\gamma_{\min} = 20$ dB, $\Gamma_s = 10$ dB 下一个典型信道实现的发射波束图. 从仿真中可以看出, ISAC-NN 与 ISAC-SDR 均在 $-30°$ 和 $45°$ 两个目标方向形成满足约束的

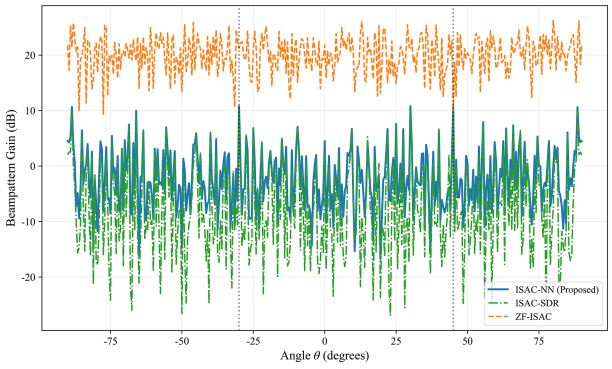

图 6　(网络版彩图) 默认配置下典型信道实现的发射波束图. $\gamma_{\min} = 20$ dB, $\Gamma_s = 10$ dB, 竖直虚线标记两个感知目标方向 $-30°$ 与 $45°$.

Figure 6 (Color online) Transmit beampattern for a representative channel realization under the default configuration. $\gamma_{\min} = 20$ dB, $\Gamma_s = 10$ dB. Vertical dotted lines mark the two sensing target directions $-30°$ and $45°$.

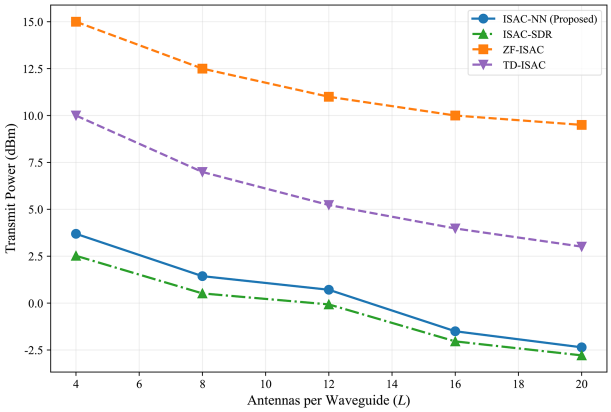

图 7　(网络版彩图) 发射功率随每波导天线数 L 的变化. $\gamma_{\min} = 20$ dB, $\Gamma_s = 10$ dB.

Figure 7 (Color online) Transmit power vs. antennas per waveguide L. $\gamma_{\min} = 20$ dB, $\Gamma_s = 10$ dB.

增益峰: ISAC-NN 的增益约为 10.8 dB, 以约 0.8 dB 的裕量满足 $\Gamma_s = 10$ dB 的约束; ISAC-SDR 的增益约为 10.0 dB, 恰好压在约束阈值上. 二者的平均旁瓣电平分别约为 0.7 和 -1.1 dB, 比目标方向增益低约 10 dB; 尽管阵元间距远大于半波长导致少量栅瓣, 但并不影响通信与感知约束的满足. 相比之下, ZF-ISAC 由于采用分步设计, 只能通过整体功率缩放使最差目标方向 (45°) 达到 10 dB 阈值, 其全角度平均辐射电平高达约 21.2 dB, 这直观解释了其约 12.5 dBm 的高发射功率 (见图 4). 该实验从波束形态上验证了波束图增益约束能有效将辐射能量聚焦于感知目标方向, 而联合设计则避免了分步设计带来的全向能量浪费.

图 7 研究了天线规模对发射功率的影响, 考虑每波导天线数 L 从 4 扫描至 20, 固定 $\gamma_{\min} = 20$ dB, $\Gamma_s = 10$ dB. 从图 7 可见, 本文提出的 ISAC-NN 算法在所有天线规模下均明显优于时间分割策略 TD-ISAC; 例如在 $L = 4$ 时, ISAC-NN 所需功率为 3.65 dBm, 比 TD-ISAC 低约 4∼6 dB; 当 $L = 20$ 时, ISAC-NN 功率进一步降至 -2.38 dBm, 降幅达 6.03 dB, 充分体现了更多天线带来的波束成形与空间分集增益. 同时, 从图 7 可以看出, ISAC-NN 的发射功率始终接近于基于 SDR 的最优波束成形方法, 两者差距在全部 L 取值下仅为 0.41∼1.13 dB, 验证了其近似最优性. 值得指出, 所提 ISAC-NN 方法的推理时间显著快于 ISAC-SDR (约 3 ms 对比约 500 ms), 这得益于神经网络的一次性前向传播, 避免了迭代求解的高计算开销.

图 8 定量评估了 PASS 架构相对传统固定位置阵列的增益, 以及不同 PASS 布局的影响, 考虑与上文相同的用户分布、约束及 ISAC-NN 求解器, 比较了 10 m 展宽 PASS, 30 m 展宽 PASS 和中心部署的 $\lambda/2$ 间距固定 ULA. 从图 8 可见, 默认 10 m 展宽 PASS 在全部天线规模下的发射功率均不高于固定 ULA, 且优势随 L 增大而扩大: $L = 4$ 时功率低约 0.2 dB, $L = 20$ 时低约 0.9 dB, 这得益于 PASS 天线沿波导大尺度分布带来的路径损耗降低与用户信道去相关. 同时可以看出, PASS 配置下 ISAC-NN 求解器的成功率始终保持在 98%∼100%, 而固定 ULA 因多用户信道高度相关导致问题条件变差, 成功率从 86% 降至 62%∼72%, 表明阵列架构是鲁棒性的关键. 两种 PASS 布局相比, $L = 4$ 时 30 m 展宽功率高出约 1.5 dB, 其余规模下差距仅 0.1∼0.9 dB, 说明部署几何是重要的设计维度. 值得指出, 采用相同 ISAC-NN 求解器时, 上述所有架构的推理时间均维持在毫秒量级, 远优于传统迭代优化方法, 为实时通感调度提供了基础.

最后, 为量化网络各组成模块的贡献并检验 4.2 小节的设计选择, 表 3 给出了系统性消融实验结果. 所有变体均在与主实验完全相同的数据、SLSQP 参考标签和两阶段训练策略下重新训练, 并按相同协议评估. 消融对象包括: 将跨域融合替换为对实虚部拼接向量的普通 MLP 编码、将多尺度提取器替换为单尺度 MLP、移除

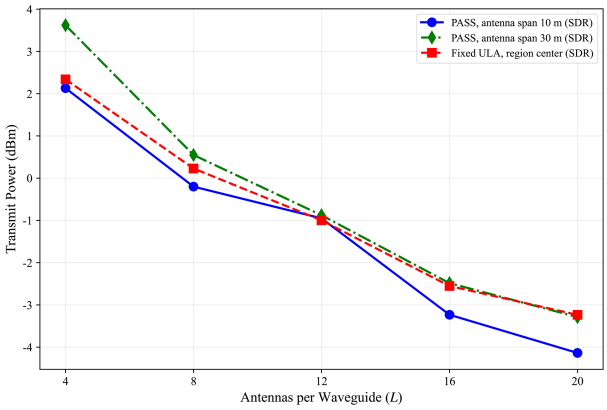

图 8　(网络版彩图) PASS 与传统固定位置 ULA 及不同 PASS 天线展宽布局的发射功率对比. $\gamma_{\min} = 20$ dB, $\Gamma_s = 10$ dB, 三种架构均采用 5.2 小节的迭代优化器求解, 功率为求解成功样本的中位数.

Figure 8　(Color online) Transmit power comparison between PASS, a conventional fixed-position ULA, and PASS layouts with different antenna spans. $\gamma_{\min} = 20$ dB, $\Gamma_s = 10$ dB. All architectures are solved by the same iterative optimizer of Subsection 5.2, and each power is the median over successfully solved samples.

表 3　网络模块与角度编码阶数消融实验 (默认配置: $\gamma_{\min} = 20$ dB, $\Gamma_s = 10$ dB, $L = 10$).

Table 3　Ablation study on network modules and angle-encoding order under the default configuration. $\gamma_{\min} = 20$ dB, $\Gamma_s = 10$ dB, $L = 10$.

Network variant	Transmit power (dBm)	SINR satisfaction (%)	Sensing satisfaction (%)	Parameters ($\times 10^6$)	Inference latency (ms)
Full model (4th-order angle encoding)	−0.19	100	100	5.28	3.29
w/o cross-domain fusion	−0.24	100	100	5.02	3.08
w/o multi-scale feature extraction	−0.12	100	100	4.99	2.90
w/o Transformer encoder	−0.19	100	100	3.70	2.53
w/o SE module	−0.14	100	100	5.27	3.34
Plain MLP (parameter-matched)	−0.27	100	100	5.58	0.89
2nd-order angle encoding	−0.26	100	100	5.28	3.24
8th-order angle encoding	−0.22	100	100	5.28	3.61

Transformer 编码器、移除 SE 模块、将整个网络替换为参数量匹配的纯 MLP, 以及将角度编码阶数由 4 改为 2 和 8. 表中同时报告中位发射功率、两类约束的满足率、参数量与实测 GPU 单样本推理时延.

从表 3 可以得到以下观察. 其一, 在 $K = N = 2$, $M = 20$ 的默认配置下, 各变体均能达到 100% 的约束满足率, 最终功率差异在 0.15 dB 量级. 这表明在该规模下, 框架的约束满足能力与功率性能主要由优化引导预训练与无监督微调的两阶段混合训练策略保证, 网络容量存在一定冗余, 单一模块的移除可由其余部分补偿. 其二, 各模块对复杂度的影响差异明显: 移除 Transformer 编码器可减少约 30% 的参数量并降低约 23% 的推理时延, 而 SE 模块的参数与时延开销几乎可以忽略; 结合各变体的功率表现, 这些数据为算力受限场景下的模型裁剪提供了量化依据. 其三, 角度编码阶数在 2, 4, 8 之间的功率差异不超过 0.05 dB, 与 4.2.2 小节的分析一致, 在本文目标角度固定的设置下, 网络性能对该超参数不敏感, 取 4 阶是编码维度与表达能力之间的折中默认值. 需要说明的是, 上述消融在小规模默认配置与固定信道集合上进行, 多支路结构化设计在更大系统规模和更多样输入分布下的收益仍需进一步验证, 这也是后续工作的方向之一.

为了深入理解 MSA-BFNet-ISAC 的学习机制, 本文从权重分布、奇异值谱、特征嵌入和输出相关性四个维度对训练后的网络进行内在特性分析, 结果如图 9 所示. 图 9(a) 展示了网络 9 个模块的权重值分布. 可以看到, 跨域融合和多尺度提取器的权重分布相对集中且近似零均值, 表明这两个通信特征提取模块的权重已经过充分优化; 角度编码器的权重分布最为分散, 这与其需要将低维角度信息映射到高维空间的功能一致; 感知编码器的权

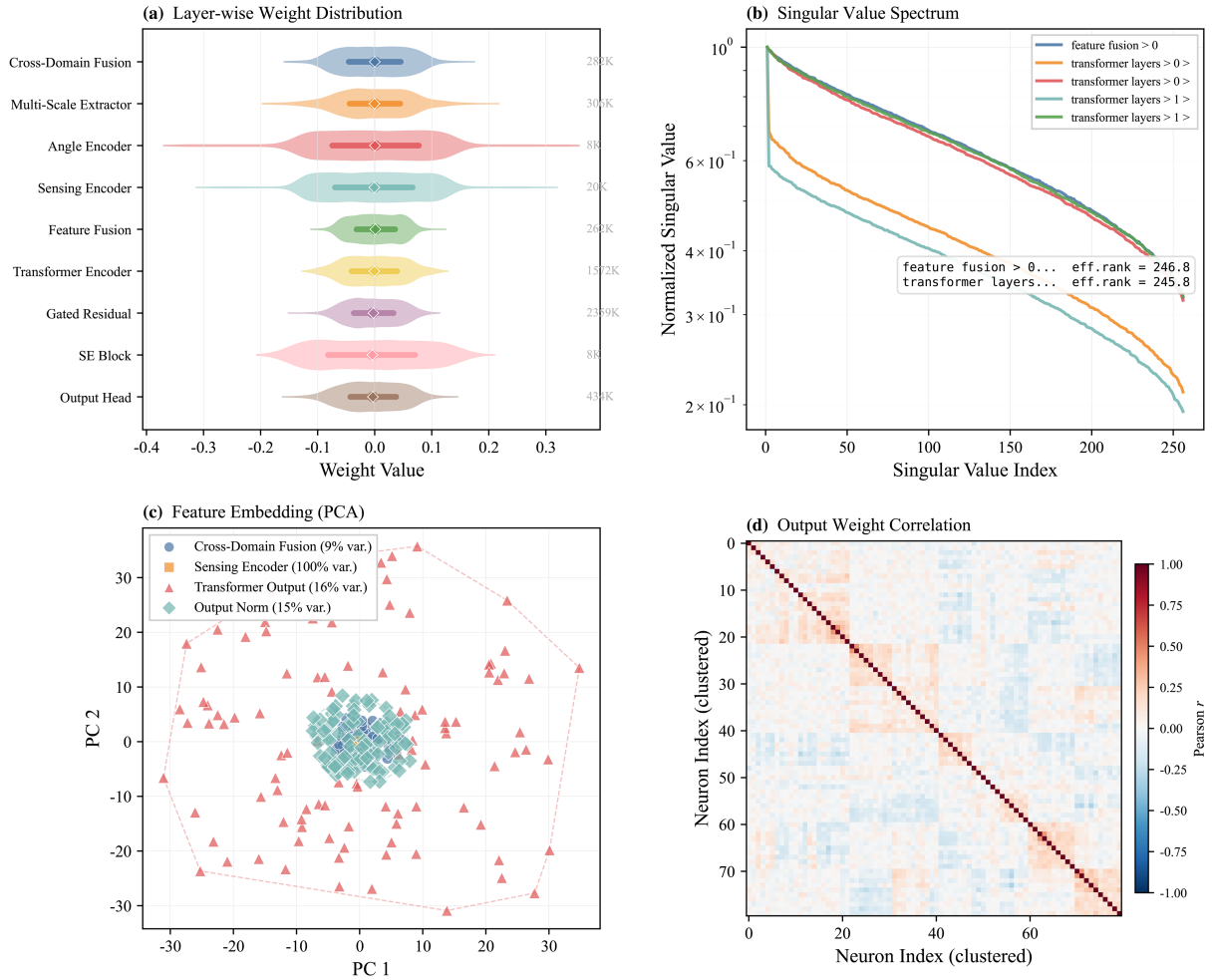

图 9　(网络版彩图) MSA-BFNet-ISAC 网络内在特性分析. (a) 各模块权重分布的小提琴图, 菱形表示均值, 粗线段为四分位距; (b) 关键权重矩阵的归一化奇异值谱及有效秩; (c) 四路特征的 PCA 投影, 括号内为前两个主成分解释的方差百分比; (d) 波束成形输出头权重矩阵的层级聚类 Pearson 相关热力图.

Figure 9　(Color online) Intrinsic analysis of trained MSA-BFNet-ISAC. (a) Violin plots of weight distributions for each module; diamonds indicate means, and thick line segments indicate interquartile ranges; (b) normalized singular value spectra of key weight matrices and their effective ranks; (c) PCA projections of four-pathway features; the percentages in parentheses indicate the variance explained by the first two principal components; (d) hierarchically clustered Pearson correlation heatmap of the beamforming output-head weight matrix.

重分布略宽于通信模块, 反映了导向向量编码的复杂性; 四路特征融合层的权重相对集中, 说明融合操作主要起到维度压缩和信息整合的作用. 图 9(b) 的奇异值谱揭示了各权重矩阵的有效自由度. 特征融合层和 Transformer 编码器的有效秩均在 245 左右, 接近隐层维度 256, 说明这些模块充分利用了全部参数空间, 信息在网络中的传递没有发生严重的维度退化. 图 9(c) 的 PCA 投影展示了不同处理阶段的特征几何结构. 跨域融合和感知编码器的特征点呈现不同的聚类形态, 验证了四路并行处理架构对通信和感知特征的差异化编码能力; Transformer 输出和最终归一化层的特征分布更为紧凑, 表明深层处理逐步将异质特征整合为面向波束成形输出的统一表示. 图 9(d) 的输出头权重相关热力图经层级聚类后, 呈现出明显的对角块结构, 表明网络自发地将输出神经元组织为若干功能组 —— 每组负责生成波束成形矩阵的特定分量, 这种结构化特性有助于网络同时满足通信和感知的双重约束. 综合以上实验来看, ISAC-NN 的发射功率与 ISAC-SDR 迭代优化基准的差距稳定在 0.4~1.4 dB 之间, 显著低于 TD-ISAC 3~6 dB, 且约束满足率保持在 98% 以上. 随 Γ_s 从 0 dB 增至 15 dB, ISAC-NN 功率单调递增约 13.6 dB, 说明网络能根据感知需求自适应地调整功率分配; 天线数从 $L = 4$ 增至 $L = 20$ 时功率稳

步下降约 6 dB. 在推理速度方面, ISAC-NN 的单样本前向推理约 3 ms, 而 ISAC-SDR 的逐样本迭代优化约需 500 ms, 速度差距约两个数量级, 在信道快速变化的场景下 ISAC-NN 具有明显的实时部署优势.

6　结论

本文针对 PASS-ISAC 系统中通信 SINR 约束与感知波束图增益约束下的功率最小化问题, 提出了 MSA-BFNet-ISAC 深度学习求解框架. 该框架将 PASS 特有的波导相位传播效应纳入通信信道和感知导向向量的建模, 区分了近场通信信道与远场感知导向向量的结构差异, 并以跨域融合、多尺度特征提取、角度编码和导向向量编码四路特征融合为核心, 结合 Transformer 编码器、门控残差单元和 SE 注意力完成端到端波束成形映射, 通过优化引导预训练与无监督微调协同的两阶段混合训练策略实现对非凸双约束问题的高效求解. 仿真实验表明, ISAC-NN 的推理速度与 ISAC-SDR 迭代优化基准相比提升约两个数量级, 在信道快速变化的实时场景下具有显著优势. 发射波束图分析、与固定位置阵列及不同 PASS 布局的对比实验, 以及网络模块与角度编码阶数的消融实验, 进一步验证了感知约束的实际效果与 PASS 部署几何的架构增益, 并量化了网络各组成部分的贡献.

未来值得进一步研究的方向有: 首先, 本文将夹持天线位置视为部署阶段给定的系统参数, 仅优化给定部署下的波束成形矩阵, 尚未将针对每个信道实现的天线位置自适应选择或调整纳入联合优化框架; 其次, 感知性能采用发射波束图增益这一能量覆盖型指标, 尚未直接约束角度估计 CRB、检测概率等终端感知指标; 再次, 扩展至更多用户和 OFDM 宽带场景, 研究用户定位误差与信道估计误差下的鲁棒智能波束成形, 以及探索将两阶段训练扩展至在线自适应学习以应对动态信道环境.

参考文献

1 Liu F, Cui Y, Masouros C, et al. Integrated sensing and communications: toward dual-functional wireless networks for 6G and beyond. IEEE J Sel Areas Commun, 2022, 40: 1728–1767

2 Zhang J A, Rahman M L, Wu K, et al. Enabling joint communication and radar sensing in mobile networks-a survey. IEEE Commun Surv Tutorials, 2022, 24: 306–345

3 Liu X, Huang T, Shlezinger N, et al. Joint transmit beamforming for multiuser MIMO communications and MIMO radar. IEEE Trans Signal Process, 2020, 68: 3929–3944

4 Liu F, Zhou L, Masouros C, et al. Toward dual-functional radar-communication systems: optimal waveform design. IEEE Trans Signal Process, 2018, 66: 4264–4279

5 Li B, Petropulu A P. Joint transmit designs for coexistence of MIMO wireless communications and sparse sensing radars in clutter. IEEE Trans Aerosp Electron Syst, 2017, 53: 2846–2864

6 Stoica P, Li J, Xie Y. On probing signal design for MIMO radar. IEEE Trans Signal Process, 2007, 55: 4151–4161

7 Luo Z Q, Ma W K, So A M C, et al. Semidefinite relaxation of quadratic optimization problems. IEEE Signal Process Mag, 2010, 27: 20–34

8 Zhu L, Ma W, Mei W, et al. A tutorial on movable antennas for wireless networks. IEEE Commun Surv Tutorials, 2026, 28: 3002–3054

9 Shao X, Mei W, You C, et al. A tutorial on six-dimensional movable antenna for 6G networks: synergizing positionable and rotatable antennas. IEEE Commun Surv Tutorials, 2026, 28: 3666–3709

10 Ding Z, Schober R, Vincent Poor H. Flexible-antenna systems: a pinching-antenna perspective. IEEE Trans Commun, 2025, 73: 9236–9253

11 Xu Y, Ding Z, Schober R, et al. Pinching-antenna systems with in-waveguide attenuation: performance analysis and algorithm design. ArXiv:2506.23966

12 Zhang Z, Wang Z, Mu X, et al. Integrated sensing and communications for pinching-antenna systems (PASS). ArXiv: 2504.07709

13 Li H, Zhong R, Lei J, et al. Pinching antenna systems for integrated sensing and communications. ArXiv: 2508.19540

14 Ding Z. Pinching-antenna assisted ISAC: a CRLB perspective. ArXiv: 2504.05792

15 Sun H, Chen X, Shi Q, et al. Learning to optimize: training deep neural networks for interference management. IEEE Trans Signal Process, 2018, 66: 5438–5453

16 Xia W, Zheng G, Zhu Y, et al. A deep learning framework for optimization of miso downlink beamforming. IEEE Trans Commun, 2020, 68: 1866–1880

17 Liu X, Fischione C. Coordinated beamforming for multi-cell ISAC using graph neural networks. IEEE Trans Wireless Commun, 2025, 25: 5876–5889
18 Zhang Z, Wu Q, Fan P, et al. DRL-based optimization for AoI and energy consumption in C-V2X enabled IoV. IEEE Trans Green Commun Netw, 2025, 9: 2144–2159
19 Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. In: Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), 2017. 5998–6008
20 Li Y, Liu Y F. HPE transformer: learning to optimize multi-group multicast beamforming under nonconvex QoS constraints. IEEE Trans Commun, 2024, 72: 5581–5594
21 Rappaport T S, Xing Y, Kanhere O, et al. Wireless communications and applications above 100 GHz: opportunities and challenges for 6G and beyond. IEEE Access, 2019, 7: 78729–78757
22 Boyd S, Vandenberghe L. Convex Optimization. Cambridge: Cambridge University Press, 2004
23 Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection. In: Proceedings of IEEE CVPR, 2017. 936–944
24 Dauphin Y N, Fan A, Auli M, et al. Language modeling with gated convolutional networks. In: Proceedings of ICML, 2017. 933–941
25 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: Proceedings of IEEE CVPR, 2018. 7132–7141
26 Kraft D. A Software Package for Sequential Quadratic Programming. Technical Report DFVLR-FB 88-28, DLR. 1988
27 Loshchilov I, Hutter F. Decoupled weight decay regularization. ArXiv: 1711.05101
28 Loshchilov I, Hutter F. SGDR: stochastic gradient descent with warm restarts. ArXiv: 1608.03983
29 Spencer Q H, Swindlehurst A L, Haardt M. Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. IEEE Trans Signal Process, 2004, 52: 461–471

Integrated sensing and communication beamforming for pinching antenna surface systems: a multi-scale attention deep learning approach

Qihan ZHANG[1], Fenghao SHENG[2], Haiyang ZHANG[2,3] & Longxiang YANG[2*]

1. *Portland Institute, Nanjing University of Posts and Telecommunications, Nanjing 210012, China*
2. *School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China*
3. *National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China*
* Corresponding author. E-mail: yanglx@njupt.edu.cn

Abstract Integrated sensing and communication (ISAC) is widely recognized as a key enabling technology for sixth-generation mobile communications. Pinching antenna surface systems (PASS), which allow flexible antenna placement along dielectric waveguides, offer additional spatial degrees of freedom unavailable in conventional fixed arrays for joint sensing-communication beamforming. This paper addresses the transmit power minimization problem in PASS-ISAC systems under a given antenna deployment, subject to joint communication signal-to-interference-plus-noise ratio (SINR) and sensing beampattern gain constraints. To handle the non-convexity and real-time requirements, we propose MSA-BFNet-ISAC, a multi-scale attention beamforming network whose core is a four-pathway joint feature processing architecture comprising cross-domain channel fusion, multi-scale spatial feature extraction, sensing target angle encoding, and steering vector encoding. A two-stage hybrid training strategy—optimization-guided pre-training followed by unsupervised fine-tuning—is further designed to efficiently optimize across communication constraints, sensing constraints, and the power objective. Simulations at 150 GHz sub-THz band show that the proposed method achieves transmit power within 1 dB of the iterative optimization benchmark while providing two-orders-of-magnitude inference speedup, significantly outperforms time-division ISAC, and steadily reduces power as the antenna count grows; comparisons with conventional fixed-position arrays and different PASS layouts further validate the architectural gain of PASS, and systematic ablation studies quantify the contribution and complexity overhead of each network component, confirming the potential of PASS in ISAC scenarios.

Keywords pinching antenna surface system, integrated sensing and communication, beamforming optimization, deep learning, multi-scale attention network