科学通报 2026 年 第 71 卷 第 30 期: 6972 ~ 6983
电磁超材料与人工智能双向赋能专题 论 文
《中国科学》杂志社 SCIENCE CHINA PRESS

基于多模态感知的双频共口径智能超表面通信系统

夏得校†, 韩家奇†, 冯强†, 穆亚洁, 朱利豪, 刘海霞, 薛皓*, 史琰, 李龙*

西安电子科技大学超高速电路设计与电磁兼容教育部重点实验室, 西安 710071

† 同等贡献

* 联系人, E-mail: xuehao@xidian.edu.cn; lilong@mail.xidian.edu.cn

2025-11-27 收稿, 2026-02-20 修回, 2026-02-24 接受, 2026-02-25 网络版发表

国家自然科学基金“信息超材料”卓越研究群体项目(62288101)、国家重点研发计划(2023YFB3811503)、国家资助博士后研究人员计划C档(GZC20250387)、国家自然科学基金青年科学基金(62401443)和中国博士后科学基金(2024M752522)资助

摘要 为满足6G对超高数据速率与极低时延的迫切需求, 毫米波通信成为关键支撑技术, 但其易受阻塞、波束管理困难等挑战制约了在动态场景中的应用. 可重构智能超表面(reconfigurable intelligent metasurfaces, RIS)可动态调控无线信道以克服毫米波通信的遮挡和衰减, 然而, 现有超表面系统严重依赖高开销、高复杂度的信道状态信息(channel state information, CSI)估计. 为此, 本文提出一种基于人工智能(artificial intelligent, AI)的“智能感知-无线通信一体化”创新架构, 通过对AI模型进行训练和验证, 实现对智能终端的快速识别, 促进从“反应式”到“感知驱动式”通信的范式转变. 本研究核心包括: (1) 设计了一款X波段与Ka波段的双频共口径智能超表面, 并通过子阵拼接的方式, 构建超大规模阵列以实现双频的精准波束调控; (2) 提出一种视觉与电磁信息融合的多模态定位方法, 通过前期数据集的训练使其能够对无人机进行感知, 融合YOLOv8的目标识别优势与射频信号的精确测距能力, 克服单一传感器的局限性; (3) 基于AI感知直接驱动超表面波束赋形, 规避传统CSI估计的瓶颈. 通过面向无人机毫米波通信场景的实验验证表明, 该架构能实现对终端的精确定位、毫米波通信信号增强等, 为未来低空经济中的毫米波可靠通信和无线传能提供了有效的技术路径.

关键词 双频共口径, 可重构智能超表面, 毫米波无线通信, 多模态感知, 低空经济

为支撑虚拟现实(virtual reality, VR)[1]、自动驾驶[2]、数字孪生[3]等前沿应用, 未来第六代移动通信(6G)系统对数据传输速率与网络时延提出了极致要求[4~6]. 在此背景下, 拥有丰富频谱资源的毫米波频段, 已成为国际电信联盟(International Telecommunication Union, ITU)认可的关键发展方向[7,8]. 然而, 毫米波信号固有的高路径损耗与弱穿透能力, 导致其在复杂环境下的覆盖性能显著下降. 根据3GPP的高频信号穿透损耗模型, 毫米波甚至难以穿透混凝土等常规建筑结构[9]. 此外, 毫米波通信对波束对准精度极为敏感[10], 这使得其能够为高速移动的非静止终端维持稳定链接, 成为一项严峻挑战.

可重构智能超表面(reconfigurable intelligent metasurfaces, RIS)作为一种低功耗、低成本的技术[11~13], 通过独立控制大规模人工超表面单元, 实现对电磁波相位[14,15]、幅度[16,17]、极化[18,19]等特性的现场可编程调控, 从而动态塑造无线信道[20,21], 以克服信号阻塞、增强接收功率并实现精准的波束赋形. 此外, 由于毫米波波长短, 易于实现超大规模阵列的集成, 这使得超表面在毫米波频段可发挥近场能量聚焦[22]、高精度感知[23]和点对点通信[24]等方面的显著优势. 然而, 现有基于超表面的通信系统大多依赖通过信道估计获取信道

引用格式: 夏得校, 韩家奇, 冯强, 等. 基于多模态感知的双频共口径智能超表面通信系统. 科学通报, 2026, 71: 6972–6983
Xia D, Han J, Feng Q, et al. Multimodal sensing-based dual-band co-aperture intelligent metasurface communication system (in Chinese). Chin Sci Bull, 2026, 71: 6972–6983, doi: 10.1360/CSB-2025-5838

© 2026《中国科学》杂志社

状态信息(channel state information, CSI)[25~27]. 频繁的信道估计需要发送大量导频信号, 将占用宝贵的频谱资源. 此外, 针对大规模超表面的信道估计会产生高维度信道矩阵, 导致估计过程计算复杂度极高[28~30].

基于上述瓶颈, 我们开展了“感知-通信一体化”的智能超表面架构研究, 从传统的“反应式通信”转向“感知驱动式通信”, 即通过目标进行独立感知, 从而指导超表面编码以实现毫米波信号增强和盲区覆盖. 我们研究设计了X波段和Ka波段的双频共口径智能超表面, 并通过类似LED大屏的可拼接方式, 构建了超大规模阵列, 其具有可拓展性, 可用于精确的波束调控. 其中, X波段用于对目标的无线感知, Ka波段则用于毫米波无线通信, 实现了多功能一体化的系统架构. 此外, 对于精确的目标定位, 本研究采用了人工智能(artificial intelligent, AI)视觉和电磁感知的双模态融合的定位方法. 该方法能够克服单一摄像头传感器在远距离测距上的不足[31,32], 和超表面系统的单通道缺陷, 通过融合计算机视觉的目标测向和多频连续波雷达测距, 实现“1+1>2”的效果.

本研究面向未来6G毫米波通信和低空经济的应用场景, 采用YOLOv8机器视觉框架对无人机终端进行识别[33~35], 通过AI技术和多功能超表面的双向赋能, 实现了双频智能超表面的自适应毫米波通信增强架构. 该架构通过感知与通信的一体化设计, 利用AI感知信息直接驱动超表面配置, 规避了传统CSI估计的开销, 从而实现无人机等终端设备的毫米波通信增强. 此外, 多模态融合的目标定位方法, 一方面充分发挥AI机器视觉的快速识别, 另一方面利用电磁波进行远距离测距[36], 能够克服传统短基线双目摄像头在目标测距上精度不足的问题[37,38]. 本文基于多模态感知的双频共口径智能超表面通信系统研究将为未来毫米波通信网络建设和低空经济发展提供重要的技术支撑, 并且为AI多模态融合的精确目标感知提供可行的技术路径.

1 大规模双频共口径智能超表面设计与验证

本研究面向未来毫米波通信的发展趋势, 结合X波段在雷达探测领域的独特优势, 创新性地提出了X/Ka双频段共口径的智能超表面设计方案. 双频的共口径设计能够为毫米波携能通信(simultaneous wireless information and power transfer, SWIPT)[22]、自适应波束扫描[39]、辐射散射调控[40]等提供高集成硬件平台, 具有显著的应用价值. 当前, 双频可重构超表面的研究在国内外仍相对较少, 尤其是针对大频率间隔(如X/Ka波段)的双频设计更具挑战性, 由于高频(Ka波段)的半波长尺寸对低频(X波段)而言属于深亚波长, 这使得对低频的相位调控变得尤为困难. 此外, 面向未来“通信-感知-传能一体化”的需求[41,42], 我们提出了一种可拼接的大规模阵列设计方案, 其思路类似于大型LED显示屏的装配方式, 充分体现了超表面设计的模块化和扩展性. 这一方式不仅能够实现更精确的波束调控, 也为未来超大规模阵列的应用奠定了架构基础.

1.1 1-bit双频超表面单元设计

在现有的双频共口径超表面设计中, 多采用变容二极管调控[43], 通过改变电容值使其工作在不同谐振频率下. 相比之下, 基于PIN二极管的方案在阵列控制和馈电方式上更加灵活便利. 此外, 本研究希望单元结构本身具备多频点谐振的效果, 当PIN二极管状态切换时, 单元能够在双频实现1-bit响应. 其难点在于低频段的设计, 为了在一个有限单元尺寸内实现0°/180°的切换, 最直接的方式是延长感应电流的路径, 并通过缝隙加载或者绕线的方式, 间接影响等效电容和电感, 从而降低结构的谐振频率. 如图1(a)所示, 是我们所设计的双频1-bit超表面单元, 它由三层金属层构成, 在其顶部加载了一个交趾电容结构, 一方面增加低频的感应电流路径, 另一方面平衡贴片与地平面之间的感抗, 提升低频带宽. 该单元仅用到了一个PIN二极管(MADP000907, MACOM, Inc.), 相比于加载多个PIN二极管的设计方案[44], 其控制复杂度低, 并且损耗更小, 对于毫米波频段来说优势显著. 图中给出了该结构的一些主要尺寸参数, 其中单元的周期尺寸为5.35 mm、厚度为3.3 mm, 整体非常紧凑.

如图1(b)所示, 是该单元在低频段的反射幅相响应. 可以看出, 在11.5~13 GHz范围内, 有良好的1-bit相位调控精度, 并且插损低于1.2 dB. 以180°±20°的调相标准来判别, 低频的带宽为1.5 GHz左右, 这对于一个亚波长单元来说性能不错, 这也得益于对交趾电容的加载和优化. 在本文中, 低频段将用于对目标的探测, 并通过AI机器视觉的赋能, 对目标进行准确的定位. 该频段除了雷达探测的重要价值, 对X波段的散射调控也具有重要意义, 未来可以用于电磁隐身方面的研究. 此外, 如图1(c)所示, 对于高频响应, 工作频段能够覆盖26~29 GHz, 该频段可以用于5G NR标准中的n258 (24.25~27.5 GHz)、n257 (26.5~29.5 GHz)、n261

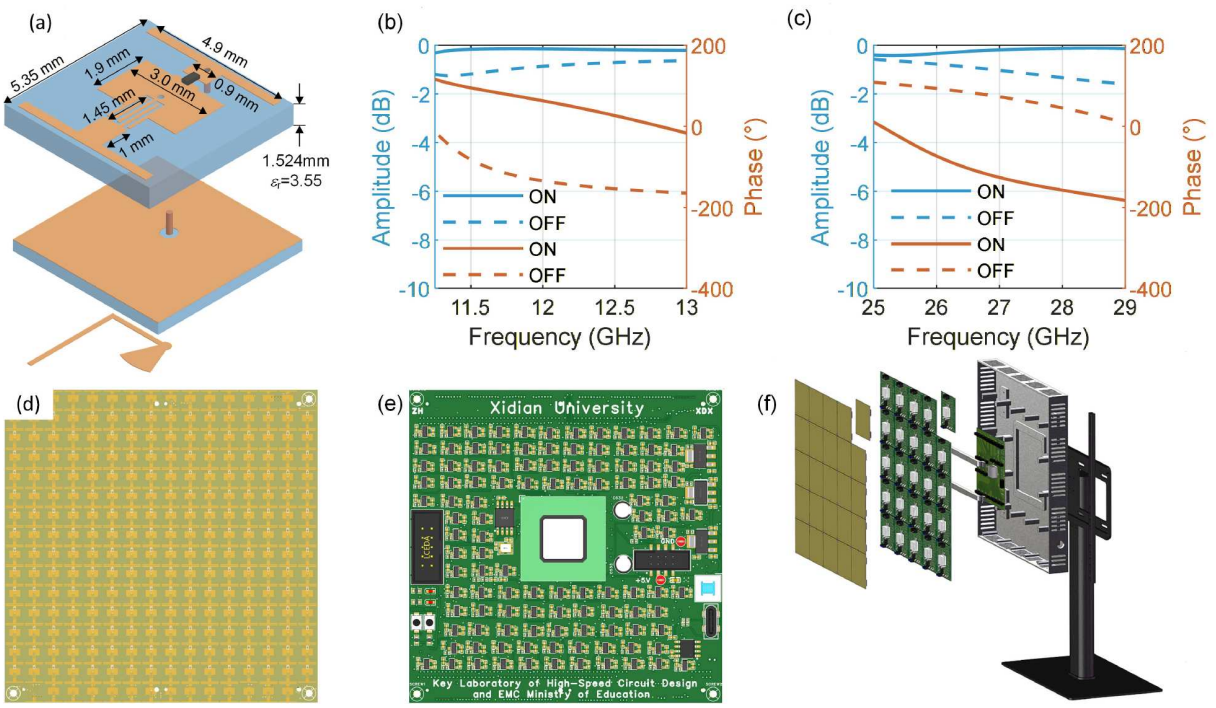


图 1 (a) 双频共口径超表面单元示意图. (b) 超表面单元在X波段的反射幅度和相位. (c) 超表面单元在Ka波段的反射幅度和相位. (d) 16×16双频共口径超表面子阵. (e) 256端口的FPGA控制板. (f) 基于子阵拼接的超大规模阵列爆炸图, 超表面单元数>6000

Figure 1 (a) Diagram of the dual-frequency co-aperture metasurface unit. (b) Reflection amplitude and phase of the metasurface unit in the X-band. (c) Reflection amplitude and phase of the metasurface unit in the Ka-band. (d) 16×16 dual-frequency co-aperture metasurface sub-array. (e) 256-port FPGA control board. (f) Exploded view of the ultra-large-scale array based on sub-array splicing, with the number of metasurface units >6000

(27.5~28.35 GHz)频段, 因此在未来毫米波通信系统中具有重要价值.

1.2 基于可拼接技术的大规模智能超表面设计

上一节中给出了双频共口径智能超表面的单元设计, 为了进一步设计大规模阵列, 我们采用模块化可拼接的方式, 该方式借鉴了大型LED显示屏的原理架构, 使得阵列规模更容易定制化并且便于扩展. 如图1(d)所示, 将16×16个单元设计成一个子阵, 该子阵的尺寸为85.8×85.8 mm^2, 电口径约为8个高频波长. 在该子阵的背面, 通过板对板连接器与FPGA控制子板相连接, 并且控制子板的尺寸与子阵的尺寸完全一致, 从而确保多个子阵的无缝拼接. 由于需要在如此小尺寸的控制板上对256个单元进行独立控制, 设计过程需要进行小型化和高密度布局, 使其能够对子阵上的所有单元进行独立控制, 如图1(e)所示. 该子阵及其控制子板将作为复用的子模块, 通过拼接的方式进行大规模阵列的设计.

如图1(f)所示, 本文通过25块子阵的拼接, 构建了一个80×80个单元的超大规模阵列, 单元数超过6000个. 其中, 采用SPI协议对每个子阵进行控制, 利用底部的主控板将编码指令以并行的方式进行下发, 整体的阵列切换时间可以达到微秒级, 能够支持实时的波束跟踪. 类似于LED大屏, 超表面单元相当于屏幕中的一个像素, 并且从像素到模块、再到大规模阵列的互联层级, 可以为大规模智能超表面的环境布置提供巨大的应用空间. 如果构建一个足够大的智能超表面阵列, 比如与楼宇墙面融合为“智能墙”, 它将能够为高精度电磁感知、点对点近场通信、生命健康监护、低空无人机无线能量传输等提供多功能一体化的平台. 此外, 结合人工智能的感知和决策能力, 将充分发挥超大规模阵列和大数据模型的优势, 从而更好地服务于未来6G的万物智联.

1.3 双频大规模智能超表面阵列测试与验证

为验证所设计的双频大规模智能超表面的近场聚焦调控和远场波束扫描能力, 本研究在微波暗室中开展了相关实验测试. 图2(a)展示了平面近场暗室的实验

场景, 超表面被放置于距离扫描平面2.5 m处, 通过波导探头在指定区域内扫描以获取近场能量分布. 在上位机的控制下, 所设计的超表面被动态配置为在z = 2.5 m的平面上, 分别于预设位置(x, y, z)=(0.1, 0, 2.5) m和(0, –0.1, 2.5) m处实现Ka波段的近场能量聚焦. 测量结果如图2(b)和(c)所示, 可见形成的波束焦点清晰, 位置与设计预期吻合, 证明了其优异的波束切换能力. 值得注意的是, 即使在远离超表面2.5 m的距离上, 能量焦斑依然显著且轮廓分明, 这得益于所设计的大规模阵列. 该阵列的电尺寸达40个波长, 根据$2D^2/\lambda$近场距离公式计算, 其近场范围可延伸至35.2 m. 因此, 在2.5 m处的距离完全处于近场区域内, 使得超表面能够充分利用其大口径优势, 实现高效的毫米波点对点能量聚焦, 从而为提升近场通信质量和无线传能效率提供有效途径.

对于远场方向图和波束扫描测试, 考虑到大规模阵列的电尺寸较大, 我们采用紧缩场暗室对其进行测试, 如图2(d)所示. 对于X波段, 它将用于对目标的无线定位, 因此对其远场波束扫描性能进行了测试和评估, 测试频率为11.5 GHz. 如图2(e)所示, 在–60°~60°扫描范围内, 方向图的波束指向准确, 最大辐射增益为20.1 dBi, 最佳副瓣电平抑制优于13 dB, 可以在无线探测与跟踪任务中发挥重要作用. 对于Ka波段, 我们给出了27 GHz下的波束扫描结果, 如图2(f)所示, 得益于大阵列口径, 该频段在增益和副瓣抑制方面具备明显优势. 其中, 最佳副瓣电平抑制达18 dB, 最大辐射增益为34.12 dBi. 在未来远距离的毫米波通信和无线能量传输中, 基于可拼接的技术, 双频RIS阵列的规模可以进一步拓展, 以满足工程应用需求.

2 多模态融合的智能超表面目标感知

2.1 基于视觉的目标测向

为了充分发挥双频智能超表面的多功能集成优势, 在本研究中, 我们分别利用X波段进行目标定位, 并根据目标位置信息计算阵列编码; 利用Ka波段进行毫米波通信补盲, 调控波束对准目标所在的区域, 提升无线通信质量. 由于超表面的单通道限制, 难以实现精确的目标测向, 为此本研究采用多模态融合的方式对目标进行定位: 基于机器视觉的目标测向基础上, 进一步融合X波段的多频连续波测距[45], 并最终实现高精度定位. 获取了目标位置后, 通过上位机计算阵列编码, 并将阵列编码进行传输和映射, 从而将毫米波通信信号聚焦到无人机所在的位置, 实现信号增强, 如图3(a)所示. 现有的双目摄像头定位方法, 在远距离精确测距场景下需要很长的物理基线, 这在实际应用中往往很难实现. 而我们选择的多模态融合定位方案, 通过电磁波传播直接测距, 突破了基线限制, 能够实现稳定可靠的中远距离目标定位. 此外, 该方案不仅能避免复杂的立体匹配计算、降低系统算力负担, 还能通过多普勒频移获取目标的潜在径向速度信息, 实现了“1+1>2”的综合优势.

为实现视觉和无线的融合定位, 本文引入了AI视觉算法-YOLOv8目标检测模型, 它能够从光学摄像头采集的原始图像中, 快速而准确地识别并定位出无人机的边界框. 相较于早期版本, YOLOv8在保持高推理速度的同时, 显著提升了检测精度, 有效克服了传统图像处理方法在应对衬度变化、部分形变或微小瑕疵时的局限性. 在YOLOv8模型的训练中, 我们采用了开源的无人机数据集, 该数据集包含四旋翼无人机图像共计10000张. 该数据集按7:3的比例划分为训练集和测试集, 并在NVIDIA A100计算平台上完成模型训练, 整个过程耗时约3小时. 进一步地, 训练完成的YOLOv8将三维空间中的无人机在二维图像中进行识别, 从而确认无人机所在的俯仰角θ和方位角φ. 基于针孔相机模型, 无人机的坐标(X, Y, Z)到图像中(u, v)的投影关系由如下关系给出:

$$s\begin{bmatrix}u\\v\\1\end{bmatrix}=\begin{bmatrix}f_x & 0 & c_x\\0 & f_y & c_y\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}X\\Y\\Z\end{bmatrix}, \tag{1}$$

其中, s是一个任意的比例因子, f_x和f_y分别表示摄像头在x轴和y轴的焦距, c_x和c_y是图像中心的坐标, 它们可以通过前期的相机标定进行校准. 进一步, 对(X, Y, Z)进行归一化, 转换到摄像头后方一个虚拟的Z = 1的平面上, 使得转换后的坐标为$x=(u-c_x)/f_x$, $y=(v-c_y)/f_y$, z = 1, 由此可以计算无人机所在俯仰角$\theta=\arctan(y/(1+x^2)^{1/2})$和方位角$\varphi=\arctan(x)$. 上述俯仰角定义为目标所在方位与水平面的夹角, 方位角定义为目标所在方位在水平面上的投影与法向的夹角. 在获取到无人机所在的方位后, 调控超表面的阵列编码, 使得X波段的波束对准无人机所在方位, 从而提升无线探测的信噪比. 需要强调的是, 如果考虑相机镜头带来的畸变, 需要进一步对畸变进行校准和补偿, 在式(1)中呈现的是一个简化的投

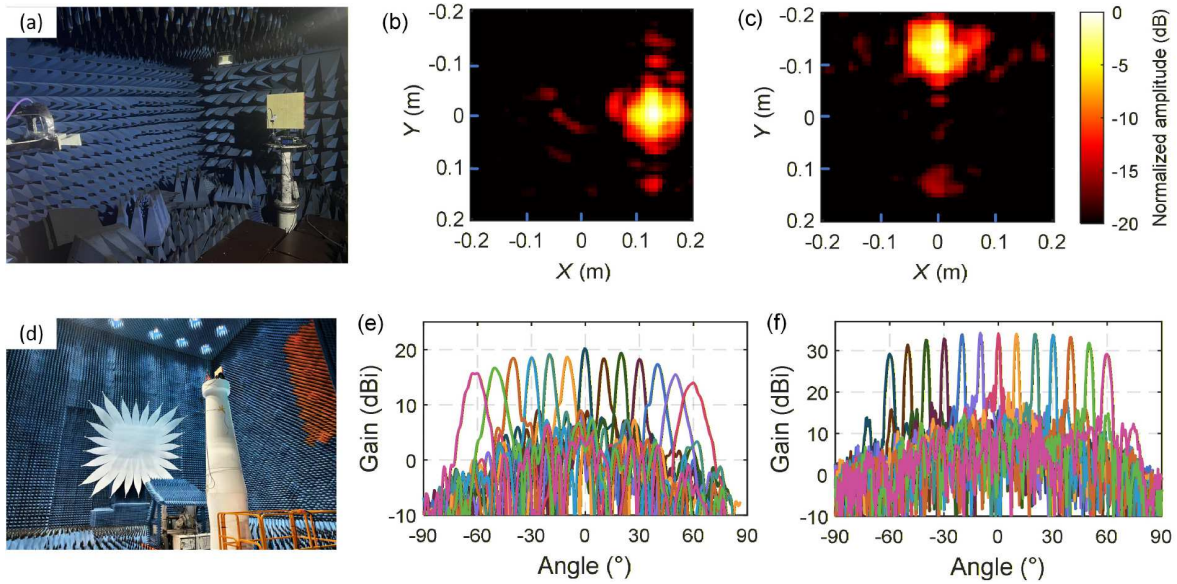


图 2 (a) 平面近场测试场景图. (b)和(c)分别为焦点位置在(0.1, 0, 2.5) m和(0, –0.1, 2.5) m处的Ka波段近场能量分布. (d) 紧缩场暗室测试场景图. (e) X波段波束扫描方向图. (f) Ka波段波束扫描方向图

Figure 2 (a) Planar near-field test scenario. (b) and (c) show the Ka-band near-field energy distribution with the focal points located at (0.1, 0, 2.5) m and (0, –0.1, 2.5) m, respectively. (d) Test scenario in a compact range anechoic chamber. (e) X-band beam scanning patterns. (f) Ka-band beam scanning patterns

影关系, 适用于畸变很小或对精度要求不高的场景.

2.2 基于多频连续波的目标测距

对于我们所设计的超大规模智能超表面阵列, 其近场范围足够大, 非常适合用于近场聚焦的无线通信和无线传能, 因此终端所在位置的距离信息需要进一步精确获取. 考虑在X波段下利用多频连续波的方式, 对目标所在位置的距离R进行定位. 整个过程可以描述为, 超表面向无人机依次辐射多个频率为f_i的连续波, 其中f_i=11 GHz + i×4 MHz(i = 0, 1, 2, 3, 4), 而在无人机上我们放置了一个有源标签, 它能对入射的电磁波进行周期性的时间调制, 其产生的谐波进一步反射回超表面, 通过环形器后最终由接收通道进行时域采样. 有源标签如图3(b)所示, 它由一个Vivaldi天线和1-bit反射式调相单元组成, 其中调相单元对接收的信号进行周期性的相位调制. 为此, 采用了NE555芯片产生2 kHz的方波, 对调相单元中的PIN二极管的通断状态进行调控, 这个过程将产生非线性的调制谐波, 该谐波将作为信标并最终反馈到超表面的接收端. 其中, 有源主动发射信号能提供更高的信噪比, 显著提升对抗多径衰落和干扰的鲁棒性; 并且调制标签的设计允许我们嵌入专用的测距序列, 从而在算法层面实现最优的时延估计精度.

在超表面上, 通用软件无线电平台(Universal Software Radio Peripheral, USRP)和X波段变频模块用于产生和接收载波频率为f_i的信号. 对于谐波信号的接收和提取, 在通过USRP的下变频和数字采样之后, 利用数字下变频(digital downconverter, DDC)的方式进行采集. DDC的核心步骤是: 首先, 以目标谐波的频率作为本振, 对原始信号进行数字混频, 将其下变频至基带; 随后, 利用一个高性能的窄带IIR低通滤波器, 滤除混频产生的高频分量以及其他非目标谐波, 从而在时域中高保真地提取出该阶谐波信号. 该过程在文献[46]中进行了详细的阐述, 它能够精确地提取出任意阶谐波的时域信号. 最终, 在通过USRP和有源标签的非同步校准后[17], 超表面接收端采集的相位可以表示为$\phi_i=(4\pi f_i R/c+\phi_{\mathrm{H}}+\phi_{\mathrm{n}})\mathrm{mod}\ 2\pi$, 其中, ϕ_{n}表示由噪声引起的随机误差, $\phi_{\mathrm{H}}=\phi_{\mathrm{tx}}+\phi_{\mathrm{rx}}+\phi_{\mathrm{tag}}$表示由发射链路、接收链路和标签天线引入的固定相位偏移, 它可以通过系统校准的方式进行补偿, 最终可以简化为$\phi_i=(4\pi f_i R/c)\mathrm{mod}\ 2\pi$. 因此, 距离$R$可以由如下给出:

$$R=\frac{c\phi_i}{4\pi f_i}+\frac{cl_i}{2f_i}, \tag{2}$$

其中, l_i表示一个非负整数. 由上式可知, 由于相位的周

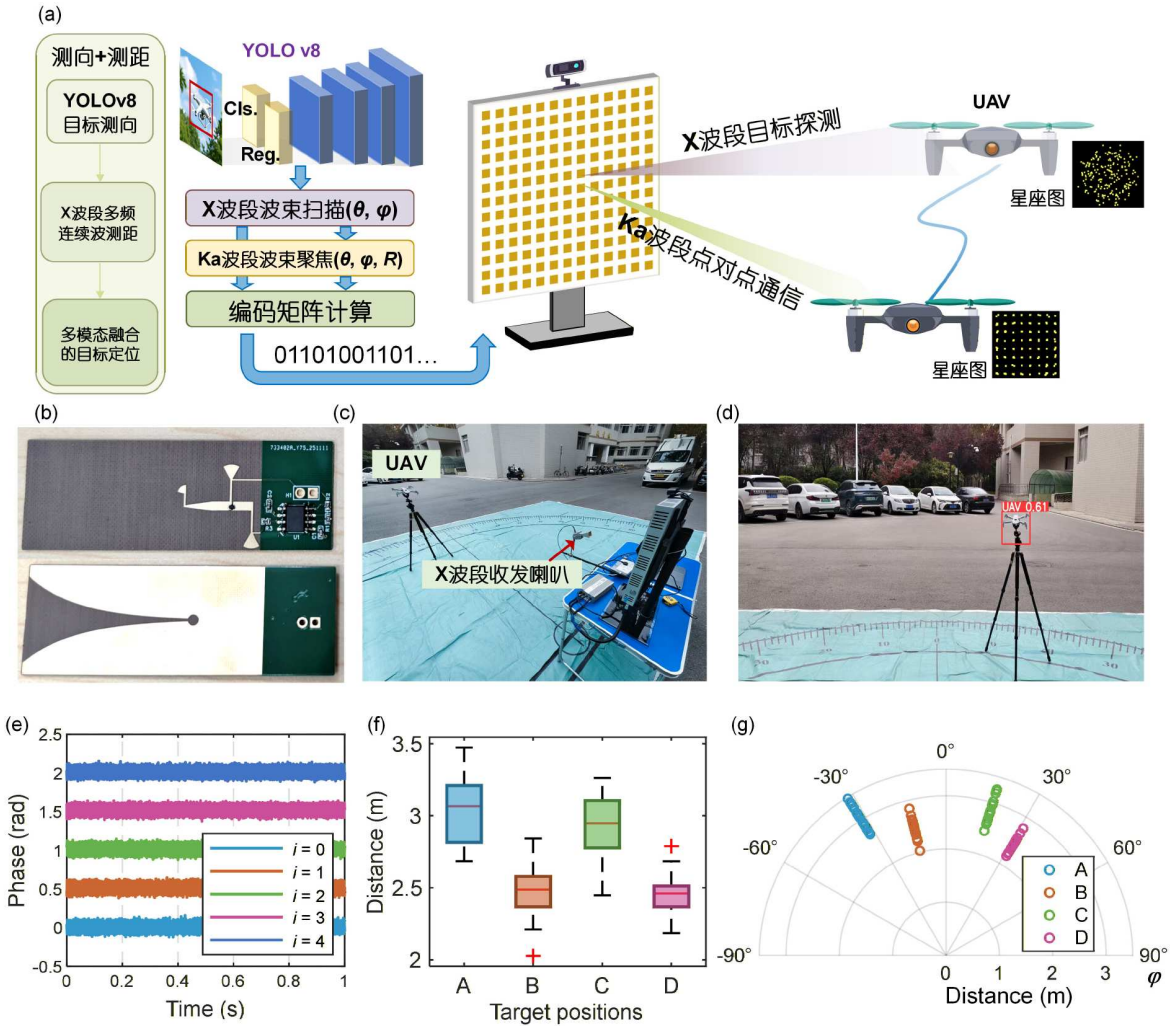


图 3 (a) 基于摄像头和电磁波的多模态目标感知, 及其自适应波束调控. 通过识别无人机所在的位置, 将毫米波通信波束进行聚焦, 从而提升通信质量和信号覆盖能力. (b) 加工的有源标签的正面和反面, 通过1-bit移相器对入射波进行时间调制. (c) 无人机定位测试场景. (d) 将无人机架设在$(\theta, \varphi, R) = (0°, 15°, 3\ \text{m})$处, 通过计算机视觉对其进行测向, 并通过多频连续波实现测距. (e) 在不同载波下, 接收的有源标签返回的信号相位. (f) 在A、B、C和D四个不同位置处, 通过多频连续波测距的统计结果. (g) 在4个不同位置下, 依次进行20次目标定位的结果, 极坐标给出了$\theta = 0°$时的二维平面

Figure 3 (a) Multimodal target perception based on camera and RF identification, and its adaptive beam control. By identifying the position of the drone, the millimeter-wave communication beam is focused, thereby enhancing communication quality and signal coverage. (b) Top and bottom views of the fabricated active tag, with time modulation of the incident wave via a 1-bit phase shifter. (c) Test scenario of UAV positioning. (d) The UAV is positioned at $(\theta, \varphi, R) = (0°, 15°, 3\ \text{m})$. Direction finding is performed via computer vision, and ranging is achieved through multi-frequency continuous wave. (e) Signal phase feedback by the active tag at different carrier frequencies. (f) Statistical results of multi-frequency continuous wave ranging at four different positions A, B, C, and D. (g) Target positioning resulted from 20 consecutive trials at each of the 4 different positions, and polar coordinates provide the 2-D plane when $\theta = 0°$

期性, 仅依靠某一个载波频率进行测距, 会出现距离模糊. 因此, 需要采用跳频的方式, 通过多个频率的连续波进行测距, 并根据不同的工作频率f_i, 存在一组上述关于R的方程组. 在无模糊距离中(对于4 MHz调频间隔, 最大无模糊距离为37.5 m), 有唯一的R与之对应. 对于方程组中R的求解, 我们采用了基于最大似然的中国余数定理(Chinese Remainder Theorem, CRT)方法[47], 找到满足最小均方根误差的结果, 从而估计目标所在的距离. 此外, 在多模态数据融合方面, 由于采用了YOLOv8与多频连续波测距融合的目标感知方法, 并且视觉与无线信号的产生频率不同, 不同模态之间的数据同步至关重要. 这里, 我们以后到达的数据时间为基准

(距离信息), 对先到达的另一模态数据进行线性插值(方位信息), 从而估计出同一时刻的对应状态, 消除异构模态间的固有时延.

2.3 目标感知测试

为了对上述所提的AI多模态融合感知方法进行评估, 我们先将无人机固定在(θ, φ, R) = (0°, 15°, 3 m), 分别采用摄像头和无线定位的方式对目标的方位和距离进行估计, 测试场景图如图3(c)所示. 通过基于AI的计算机视觉, YOLOv8实现了对图像中无人机的精准识别, 如图3(d)所示, 这里以图中检测框中心所在的二维坐标为结果, 估计出目标的俯仰角θ和方位角φ, 其结果为(θ, φ) = (0.1°, 16.6°). 值得注意的是, 融合了AI视觉的目标感知, 对于单通道的超表面来说至关重要, 它不仅克服了通道数不足的缺陷, 也能避免复杂环境中无线定位鲁棒性差的问题. 在获取了目标方位后, 我们调控阵列的编码, 将X波段的远场波束对准目标所在的方位, 从而利用电磁波进行无线测距. 随后, 通过USRP采集不同载波下返回的谐波相位, 图3(e)中给出了1 s时间间隔内不同载波下的相位曲线, 其中i表示载波序号. 由于噪声和多径效应的影响, 图中的相位曲线围绕着均值来回波动, 但总体趋于平稳, 据此计算的目标距离的均值为R=2.93 m. 从结果来看, 我们所提的多模态融合目标感知方法是可行的, 该方法所获取的无人机坐标信息为(θ, φ, R) = (0.1°, 16.6°, 2.93 m), 与真实位置之间的误差为厘米级, 精确的位置信息将用于指导阵列在Ka波段进行编码, 以实现毫米波近场聚焦的无线通信和无线供能.

此外, 为了对所提出的目标感知方法进行评估, 我们将无人机依次架设到4个不同的位置, 分别为(0°, –30°, 3 m)、(0°, –15°, 2.5 m)、(0°, 15°, 3 m)和(0°, 30°, 2.5 m), 并将4个位置依次标记为A, B, C和D. 实验中, 对每个位置均进行20次的距离估计, 其统计结果以箱线图的形式进行呈现. 从图3(f)中可以看出, A点所在位置的误差波动较大, 其中位数为3.075, 50%样本的误差波动在0.39 m以内; 除此之外, 其余位置的定位精度和误差波动略有改善, 它们的中位数也非常接近目标所在的距离. 误差的主要来源是超表面收发通道之间的耦合和多径效应, 整体来看, 目标定位的误差在亚分米级, 能够对目标位置有不错的估计精度. 进一步, 为了直观呈现多模态融合的目标定位效果, 将所测结果以散点图的形式绘制在极坐标中, 如图3(g)所示, 给出了θ=0°时的二维平面. 可以看出, 图中散点图所在的方位角与预期一致, 说明基于AI的计算机视觉可以精确地获取目标方位. 进一步地, 配合亚分米级的无线测距后, 无人机所在的位置可以被最终确认, 图中多次测试的结果是以目标真实位置为中心呈现合理波动. 对于测距精度的进一步改善, 未来可拓宽多频连续波的扫频带宽, 以提升无线测距的距离分辨率, 并改善多径效应的影响. 此外, 当视觉因天气、光照失效时, 多频连续波测距仍可独立工作, 提供目标的距离信息. 虽然失去了方位, 但结合历史轨迹或简单运动模型, 系统仍可维持粗略的目标跟踪与存在性判断, 这远比纯视觉系统完全失效更稳健.

3 自适应毫米波通信补盲实验

基于多模态目标感知, 本文进一步开展了自适应通信补盲实验. 针对Ka波段毫米波通信, 在估计出终端所在的位置之后, 本实验利用大规模智能超表面的灵活波束调控能力, 将入射的通信波束聚焦到无人机所在的位置, 实现点对点无线通信增强. 如图4(a)所示, 以超表面中心为原点, 将Ka波段馈源架设于坐标(X, Y, Z) = (–2, 1.2, 2.5) m处, 通过矢量信号发生器产生中心频率为27 GHz的64QAM信号, 该信号入射到超表面后, 超表面动态调整编码从而将波束聚焦到无人机. 无人机端则配置超宽带Vivaldi天线作为接收天线, 并借助矢量信号分析仪(Ceyear 4051F)对通信信号质量进行评估, 包括信噪比和星座图. 其中, 信噪比的计算依据其定义: 信号功率与噪声功率的比值. 具体而言, 从频谱数据中分别提取信号频带内的总积分功率和邻近无信号频段内的噪声积分功率, 两者相除并取对数得到信噪比(单位为dB).

如图4(b)所示, 在超表面前方架设了X波段喇叭天线, 作为无线测距的收发端口, 再结合摄像头的目标测向功能, 实现对目标位置的精确感知. 为验证通信补盲和增强效果, 实验中将无人机分别置于两个不同位置: (1, 1.2, 2.8) m和(–0.5, 1.2, 3) m, 它们分别对应图4(c)与(d). 通过多模态融合的实时目标位置估计, 计算机可快速计算阵列编码, 使得Ka波段馈源入射的通信信号在无人机位置处聚焦, 从而提升接收信号质量. 在对图4(c)中无人机位置进行精确估计后, 图4(e)展示了上位机计算的超表面编码矩阵(左下角), 以及在该编码下测试的接收信号星座图, 可见星座点收敛良好, 并且误差向量幅度(error vector magnitude, EVM)最大值为

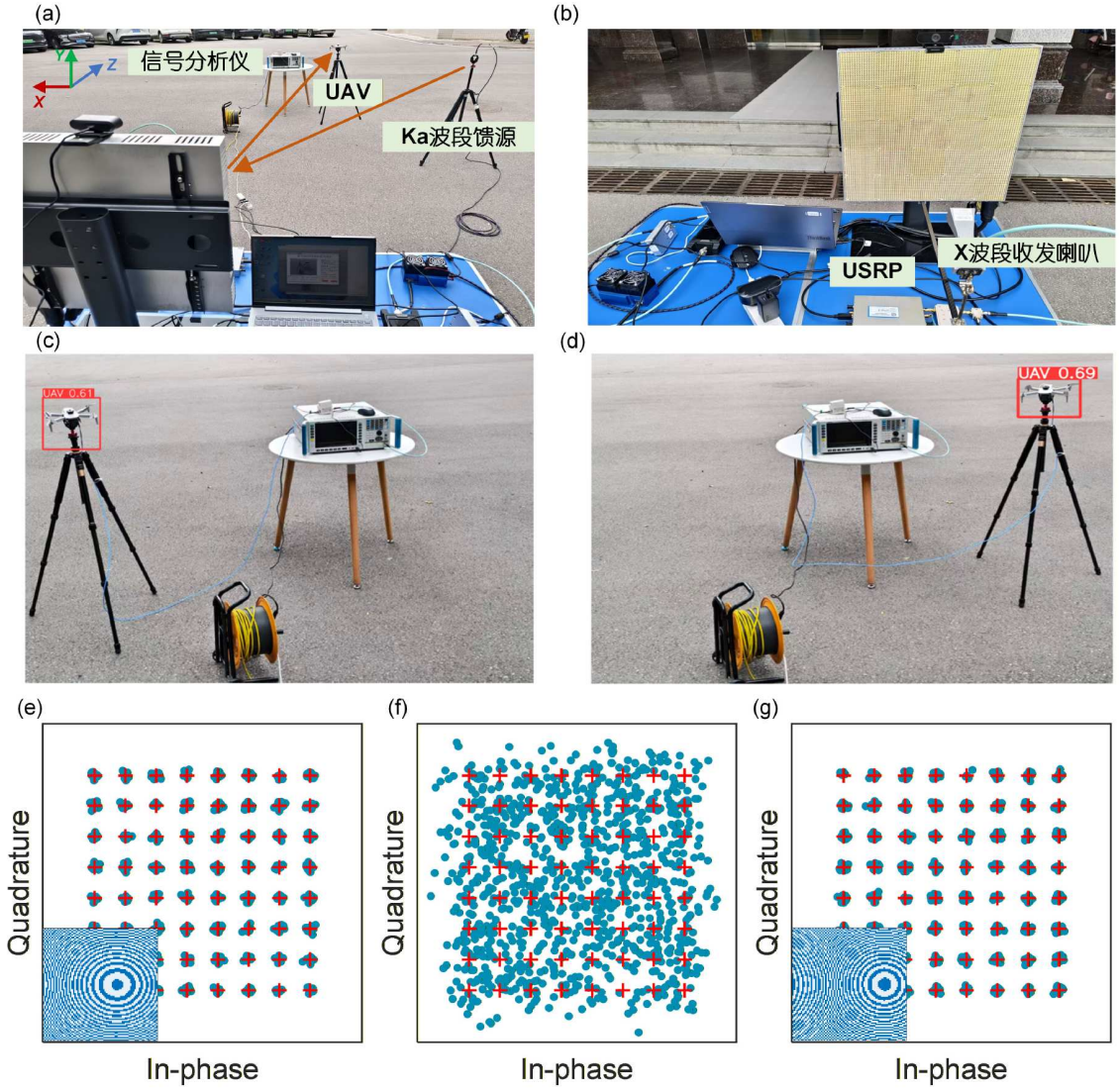


图 4 (a) 自适应波束聚焦的毫米波通信补盲实验场景图, 其中Ka波段信号从远区入射. (b) 超表面系统的主视图, X波段收发喇叭用于对无人机进行无线定位. (c)和(d)分别为当无人机分别架设于(X, Y, Z) = (1, 1.2, 2.8) m和(–0.5, 1.2, 3) m处, 计算机视觉对无人机的识别结果. (e) 当波束焦点对准无人机所在的(1, 1.2, 2.8) m处时, 无人机接收信号的星座图. (f) 当超表面断电时, 无人机接收信号的星座图. (g) 当波束焦点对准无人机所在的(–0.5, 1.2, 3) m处时, 无人机接收信号的星座图

Figure 4 (a) Experimental scenario for mmWave communication blind zone coverage using adaptive beam focusing, with a Ka-band signal incident from the farther region. (b) Front view of the metasurface system, where an X-band horn antenna is used for wireless localization of the UAV. (c) and (d) show the computer vision-based detection results of the UAV when it is positioned at (X, Y, Z) = (1, 1.2, 2.8) m and (–0.5, 1.2, 3) m, respectively. (e) Constellation diagram of the signal received by the UAV when the beam is focused on its location at (1, 1.2, 2.8) m. (f) Constellation diagram of the signal received by the UAV when the metasurface is powered off. (g) Constellation diagram of the signal received by the UAV when the beam is focused on its location at (–0.5, 1.2, 3) m

2.13%. 作为对比, 图4(f)给出了在图4(c)场景中, 超表面断电时的接收信号星座图, 其性能恶化明显, 接收端不能对通信信号进行正常的解调. 上述对比结果表明, 通过超表面的动态波束调控, 毫米波通信的信号质量可以得到明显的改善, 验证了其在通信增强和补盲中的有效性.

类似地, 如图4(g)所示, 当无人机位置发生改变时, 感知系统能够引导超表面进行重新编码, 并将波束再次对准无人机所在的位置. 在波束对准后, 接收信号质量可以得到显著改善, 星座图清晰可辨, 且系统误码率低于可正确解调的阈值. 无人机位置变化下, 通信质量从差到好的转变, 充分体现了智能超表面对无线信道

的灵活重构. 并且得益于超大规模阵列的口径优势, 通信信号的信噪比改善超过23 dB, 这将显著提升信号的覆盖能力. 此外, 针对低空经济无人机续航能源问题, 采用本文提出的双频共口径智能超表面系统, 将能为移动目标的自适应定位和无线传能提供平台基础, 实现无线能量和信息的协同高效传输.

4 总结

面对AI技术和移动互联的高速发展, 如何开展AI与无线技术的双向赋能, 实现多功能集成、环境自适应、精确目标感知、信息与能量协同传输等, 是未来电子信息的发展必经之路. 在本文的研究中, 通过双频共口径的大规模智能超表面设计, 实现了电子信息平台的多功能一体化, 为“感知-通信一体化”提供了硬件基础. 此外, 本文提出了一种AI视觉与电磁感知融合的目标定位方法, 结合YOLOv8视觉识别的快速定向优势与X波段射频信号的精确测距能力, 克服了单一传感器的局限, 从而实现不同感知模态的优势互补. 最终, 基于上述感知信息直接驱动超表面产生定向波束, 构建了完整的“感知-决策-调控-通信”闭环系统. 实验结果表明, 所提架构能够对无人机等智能终端进行毫米波通信信号增强, 通信链路信噪比获得超过23 dB的显著提升. 因此, 本研究工作不仅为低空毫米波可靠通信提供了一种有效的解决方案, 其提出的多模态融合技术路径, 也为未来6G网络实现智能、高效的无线资源调控提供了重要的理论基础与实践依据. 未来, 本研究将继续沿着多模态融合和高精度的方向深入研究, 将采用分布式传感网络和更大带宽的定位信号, 提升在城市低空复杂场景和高性能感知需求的实时性和可靠性, 为点对点无线通信、室内导航定位、生命健康监测等提供精确的位置信息.

参考文献

1 Berrezueta-Guzman S, Wagner S. Immersive multiplayer VR: unreal engine’s strengths, limitations, and future prospects. IEEE Access, 2025, 13: 85597–85612
2 Li Y, Zhou F, Yuan L, et al. A novel knowledge graph driven automatic modulation classification framework for 6G wireless communications. IEEE Trans Wireless Commun, 2025, 24: 2373–2388
3 Tao Y, Wu J, Lin X, et al. DRL-driven digital twin function virtualization for adaptive service response in 6G networks. IEEE Netw Lett, 2023, 5: 125–129
4 Ishteyaq I, Muzaffar K, Shafi N, et al. Unleashing the power of tomorrow: exploration of next frontier with 6G networks and cutting edge technologies. IEEE Access, 2024, 12: 29445–29463
5 Jagannath A, Jagannath J, Melodia T. Redefining wireless communication for 6G: signal processing meets deep learning with deep unfolding. IEEE Trans Artif Intell, 2021, 2: 528–536
6 Priyadarshani R, Ki-Hong R, Ata Y ı, et al. Jamming intrusions in extreme bandwidth communication: a comprehensive overview. IEEE Open J Commun Soc, 2025, 6: 7125–7149
7 Verdecia-Peña R, Martinez-de-Rioja D, Alonso J I, et al. mmWave dual-coverage RIS-aided 6G wireless communications: experimental demonstration. IEEE Wireless Commun Lett, 2025, 14: 3630–3634
8 Hong W, Jiang Z H, Yu C, et al. The role of millimeter-wave technologies in 5G/6G wireless communications. IEEE J Microw, 2021, 1: 101–122
9 Neil C T, Shafi M, Smith P J, et al. Impact of microwave and mmWave channel models on 5G systems performance. IEEE Trans Antennas Propagat, 2017, 65: 6505–6520
10 Biason A, Zorzi M. Multicast via point to multipoint transmissions in directional 5G mmWave communications. IEEE Commun Mag, 2019, 57: 88–94
11 Xia D, Guan L, Liu H, et al. MetaBreath: multitarget respiration detection based on space-time-coding digital metasurface. IEEE Trans Microwave Theor Techn, 2024, 72: 1433–1443
12 Mu Y, Xia D, Han J, et al. Low in-band RCS reconfigurable metasurface with LP feed for CP transmission, LP reflection, and simultaneous transmission-reflection modes’ switching. IEEE Trans Antennas Propagat, 2024, 72: 7749–7762
13 Mu Y, Xia D, Han J, et al. Time-space-coding radiation-stealth metasurface with amplitude-phase co-modulation. Adv Funct Mater, 2024, 34: 2407802
14 Lin M, Xu M, Wan X, et al. Single sensor to estimate DOA with programmable metasurface. IEEE Internet Things J, 2021, 8: 10187–10197
15 Han J, Li L, Ma X, et al. Adaptively smart wireless power transfer using 2-bit programmable metasurface. IEEE Trans Ind Electron, 2022, 69: 8524–8534

16 Wang X, Han J Q, Li G X, et al. High-performance cost efficient simultaneous wireless information and power transfers deploying jointly modulated amplifying programmable metasurface. Nat Commun, 2023, 14: 6002
17 Xia D, Wang X, Han J, et al. Accurate 2-D DoA estimation based on active metasurface with nonuniformly periodic time modulation. IEEE Trans Microwave Theor Techn, 2023, 71: 3424–3435
18 Chang M, Mu Y, Han J, et al. Tailless information–energy metasurface. Adv Mater, 2024, 36: 2313697
19 Khan M I, Hu B, Chen Y, et al. Multiband efficient asymmetric transmission with polarization conversion using chiral metasurface. Antennas Wirel Propag Lett, 2020, 19: 1137–1141
20 Najafi M, Jamali V, Schober R, et al. Physics-based modeling and scalable optimization of large intelligent reflecting surfaces. IEEE Trans Commun, 2021, 69: 2673–2691
21 Hua M, Wu Q, Poor H V. Power-efficient passive beamforming and resource allocation for IRS-aided WPCNs. IEEE Trans Commun, 2022, 70: 3250–3265
22 Liu H X, Li Y C, Cheng F J, et al. Holographic tensor metasurface for simultaneous wireless powers and information transmissions using polarization diversity. Adv Funct Mater, 2024, 34: 2307806
23 Han J, Li L, Tian S, et al. Millimeter-wave imaging using 1-bit programmable metasurface: simulation model, design, and experiment. IEEE J Emerg Sel Top Circuits Syst, 2020, 10: 52–61
24 Yu Q, Dai L. Near-field wideband beamforming for RIS based on fresnel zones. IEEE Trans Commun, 2025, 73: 5126–5138
25 Xu M, Zhang S, Ma J, et al. Deep learning-based time-varying channel estimation for RIS assisted communication. IEEE Commun Lett, 2022, 26: 94–98
26 Liu Y, Zhang S, Gao F, et al. Cascaded channel estimation for RIS assisted mmWave MIMO transmissions. IEEE Wireless Commun Lett, 2021, 10: 2065–2069
27 Guan X, Wu Q, Zhang R. Anchor-assisted channel estimation for intelligent reflecting surface aided multiuser communication. IEEE Trans Wireless Commun, 2022, 21: 3764–3778
28 Peng Z, Li T, Pan C, et al. RIS-aided D2D communications relying on statistical CSI with imperfect hardware. IEEE Commun Lett, 2022, 26: 473–477
29 Ren H, Pan C, Wang L, et al. Long-term CSI-based design for RIS-aided multiuser MISO systems exploiting deep reinforcement learning. IEEE Commun Lett, 2022, 26: 567–571
30 Zhou G, Pan C, Ren H, et al. Channel estimation for RIS-aided multiuser millimeter-wave systems. IEEE Trans Signal Process, 2022, 70: 1478–1492
31 Sid-Ahmed M A, Boraie M T. Dual camera calibration for 3-D machine vision metrology. IEEE Trans Instrum Meas, 1990, 39: 512–516
32 Dong J, Sun L, Chen L, et al. Multieye visual fusion encoderless control with permanent magnet synchronous machines. IEEE Sens J, 2025, 25: 40901–40912
33 Zhou G, Liu X, Bi H. Recognition of UAVs in infrared images based on YOLOv8. IEEE Access, 2025, 13: 1534–1545
34 Hussain M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. Machines, 2023, 11: 677
35 Di J, Xi K, Niu H, et al. Enhanced YOLOv8 framework for precise small object detection in UAV imagery. IEEE Access, 2025, 13: 157811–157827
36 Kong S H, Cho S, Kim T, et al. Stepped-frequency binary offset carrier modulation for global navigation satellite systems. IEEE Trans Aerosp Electron Syst, 2023, 59: 2733–2751
37 Maheepala M, Kouzani A Z, Joordens M A. Light-based indoor positioning systems: a review. IEEE Sens J, 2020, 20: 3971–3995
38 Alam M S, Mohamed F B, Selamat A, et al. A review of recurrent neural network based camera localization for indoor environments. IEEE Access, 2023, 11: 43985–44009
39 Xia D X, Han J Q, Mu Y J, et al. Adaptive wireless-powered network based on CNN near-field positioning by a dual-band metasurface. Nat Commun, 2024, 15: 10358
40 Mu Y, Han J, Xue H, et al. Electromagnetic all-in-one radiation-scattering reconfigurable intelligent metasurface. Natl Sci Rev, 2025, nwaf470
41 Liu H, Wu X, Xue H, et al. Multitarget simultaneous wireless information and power transfer (SWIPT) scheme based on anisotropy-metasurface field synthesis. IEEE Trans Antennas Propagat, 2024, 72: 8466–8475
42 Mu Y, Xia D, Han J, et al. Time-space-coding radiation-stealth metasurface with amplitude-phase co-modulation. Adv Funct Mater, 2024, 34: 2407802
43 Lin H, Yu W, Tang R, et al. A dual-band reconfigurable intelligent metasurface with beam steering. J Phys D-Appl Phys, 2022, 55: 245002
44 Saifullah Y, Chen Q, Yang G M, et al. Dual-band multi-bit programmable reflective metasurface unit cell: design and experiment. Opt Express, 2021, 29: 2658–2668

45 Schroth C A, Eckrich C, Kakouche I, et al. Emergency response person localization and vital sign estimation using a semi-autonomous robot mounted SFCW radar. IEEE Trans Biomed Eng, 2024, 71: 1756–1769
46 Guan L, Yang X, Zhao N, et al. Domain-independent gesture recognition using single-channel time-modulated array. IEEE Trans Antennas Propagat, 2024, 72: 3386–3399
47 Chen R, Huang X, Zhou Y, et al. UHF-RFID-based real-time vehicle localization in GPS-less environments. IEEE Trans Intell Transp Syst, 2022, 23: 9286–9293

Summary for “基于多模态感知的双频共口径智能超表面通信系统”

Multimodal sensing-based dual-band co-aperture intelligent metasurface communication system

Dexiao Xia†, Jiaqi Han†, Qiang Feng†, Yajie Mu, Lihao Zhu, Haixia Liu, Hao Xue*, Yan Shi & Long Li*

Key Laboratory of High-Speed Circuit Design and EMC, Ministry of Education, Xidian University, Xi'an 710071, China
† Equally contributed to this work
* Corresponding authors, E-mail: xuehao@xidian.edu.cn; lilong@mail.xidian.edu.cn

The impending era of sixth-generation (6G) wireless networks is defined by its demand for transformative performance metrics, including terabit-per-second data rates and sub-millisecond latency, which are essential for enabling advanced applications like virtual reality (VR), digital twin, and the low-altitude economy. To meet these stringent requirements, millimeter-wave (mmWave) communication has emerged as a cornerstone technology, offering vast amounts of underutilized spectrum. However, the inherent propagation characteristics of mmWave signals, notably their severe susceptibility to blockage by common obstacles and pronounced path loss, present significant challenges for reliable connectivity. These limitations are acutely manifested in dynamic scenarios, such as communications with unmanned aerial vehicles (UAVs), where traditional beam management strategies—reliant on high-overhead channel state information (CSI) estimation—struggle with the rapid channel variations. While reconfigurable intelligent metasurfaces (RIS) represent a promising solution by dynamically manipulating the wireless environment to mitigate obstructions, existing metasurface-assisted systems remain fundamentally constrained by their dependency on complex and latency-intensive CSI acquisition processes, creating a critical bottleneck for real-time operation.

To address this bottleneck, this paper proposes a novel integrated sensing and communication (ISAC) architecture that orchestrates a paradigm shift from a “reactive” model, which responds to channel changes after they occur, to a proactive, “sensing-driven” framework. The core contributions of this work are threefold. First, we introduce the design and implementation of a dual-frequency co-aperture metasurface. This innovative hardware platform operates concurrently in the X-band for accurate sensing and the Ka-band for mmWave communication. An ultra-large-scale array aperture is constructed using a sub-array splicing approach, which enables independent and highly precise beam manipulation for both sensing and data transmission functions on a single physical platform. Second, we develop a novel multimodal localization method that achieves data fusion between artificial intelligence (AI) visual and electromagnetic sensing modalities. This method strategically integrates the superior real-time target recognition and bounding box precision of the YOLOv8 computer vision model with the accurate range information derived from RF signal processing. This fusion effectively overcomes the inherent limitations of single-sensor systems, such as the visual model's sensitivity to lighting conditions and the RF system's lack of contextual awareness, resulting in potential for continuous 3D trajectory tracking of mobile terminals. The third and most significant contribution is the direct utilization of the acquired high-precision sensing information—specifically, the real-time spatial coordinates of the target—to analytically compute and configure the phase profile of the metasurface. This process entirely bypasses the conventional, iterative CSI estimation and feedback loop, thereby eliminating its associated latency and computational overhead.

The proposed sensing-driven communication architecture was validated through a comprehensive set of experiments emulating a challenging UAV mmWave communication link. The experimental results confirm that the system achieves two primary objectives simultaneously: AI-based precise localization of the UAV terminal and a significant enhancement of the received mmWave signal strength. By proactively steering the communication beam based on predictive sensing data, the system maintains a stable and high-quality link even in an NLoS scenario. The demonstrated architecture effectively decouples beam management from the traditional CSI dependency, resulting in a substantial reduction in link configuration latency and signaling overhead. Consequently, this work provides a concrete and effective technical pathway for realizing reliable, high-capacity, low-altitude mmWave communications and wireless power transfer. It offers a feasible solution that paves the way for future 6G applications demanding extreme reliability and ultralow latency in dynamic environments.

dual-frequency co-aperture, reconfigurable intelligent metasurface, mmWave wireless communication, multimodal sensing, low-altitude economy

doi: 10.1360/CSB-2025-5838