RIS辅助的单视角三维重建无线模态感知

Yuhan Wang[1], Haobo Zhang[1], Qingyu Liu[1],Hongliang Zhang[2],Lingyang Song[1,2,3]1中国北京大学深圳研究生院电子与计算机工程学院2中国北京大学电子学院3中国湘潭湘潭大学湖南省先进感知与信息技术研究院

arXiv:2602.02148v1 [eess.SP] 2 Feb 2026

摘要——**全模态感知对于各种现实世界感知应用至关重要，因为它能够恢复复杂环境中部分遮挡物体的完整形状。在各种全模态感知范式中，无线模态外感知凭借其环境鲁棒性、隐私保护和低成本等优势，成为一种潜在解决方案。然而，由于空间分辨率较低，无线系统获取的感知数据较为稀疏，不利于形状重建；在存在遮挡的复杂环境中，这一问题会进一步加剧。为解决这一问题，我们提出一种由可重构智能表面（RIS）辅助的无线模态外感知方案，利用大规模RIS提升空间分辨率，并创建能够绕过障碍物的反射路径。我们还采用生成式学习模型，根据从RIS视点采集的感知数据重建完整形状。在此类系统中，优化RIS相移颇具挑战，因为RIS相移与模态外感知精度之间的关系十分复杂，且其闭式表达式未知。为应对这一挑战，我们开发了一种误差预测模型，用于学习RIS相移到模态外感知精度的映射关系，并据此优化RIS相移。在基准数据集上的实验结果表明，在RIS配置数量相同的情况下，与传统方案相比，我们的方法至少可将重建误差降低56.73%**。

索引词—**全模态感知、可重构智能表面、波束成形**。

I. 引言

三维重建对于从自动驾驶到工业机器人等各种现实应用至关重要。例如，自动驾驶车辆会预测被遮挡行人的完整形状，以准确估计其位置和行为；机械臂则需要推断堆叠货物的完整几何形状，以实现高效抓取 [1]。为解决遮挡导致的形状重建精度下降问题，近期提出了全模态感知技术。该技术利用物理世界的形状连续性、对称性和结构先验信息，根据可见部分推断被遮挡部分 [2], ，从而实现完整形状的重建。

在各种全模态感知范式中，无线模态外感知是一种潜在的解决方案。具体而言，无线模态外感知通过分析目标散射的无线信号来重建完整形状。与现有的基于视觉的模态感知方法相比 [3],，无线模态外感知能够抵御光照和天气变化。它还不会记录面部细节等敏感数据，因此能够保护隐私。与基于激光雷达的方法相比 [4], ，它能够适应恶劣天气条件，例如雨雾天气，而且可以利用现有的无线系统（如WiFi），从而大幅降低硬件成本 [7]。

尽管无线模态外感知具有潜在优势，但数据稀疏性问题限制了其性能。与摄像机和激光雷达相比，无线信号的空间分辨率较低，导致目标的感知数据稀疏。在全模态感知场景中，这一问题会进一步加剧：由于遮挡，部分信号传播路径受到阻断，从而加重数据稀疏性，使完整形状的重建更加困难。

本文提出一种可重构智能表面（RIS）辅助的无线模态外感知方案，以提升无线模态外感知的性能。所提出的系统由一对发射和接收天线、一个RIS以及待重建的目标物体组成。部署的RIS由大量亚波长单元构成，各单元具有可调相移，可用于波束赋形 [5]。凭借较大的孔径，RIS能够提升无线感知系统的空间分辨率，从而提高重建精度 [6]。此外，RIS还可创建能够绕过障碍物的反射路径，从而部分缓解环境中的遮挡 [5]。此外，系统采用生成模型，根据从RIS视点采集的感知数据重建完整形状。

所提出的RIS辅助全模态感知方案与现有的RIS辅助方案有显著区别。具体而言，文献 [5] 的作者提出了一种RIS辅助的单视点三维重建方法，其中RIS通过定制反射波束来获取目标形状信息。然而，由于自遮挡，该方法在重建物体背面时面临挑战。为缓解这一问题，文献 [8], 设计了一种多视角方法：由移动发射机发送感知信号，从不同方向感知目标；RIS则将目标散射信号反射至接收机，以实现形状重建。尽管如此，该方法会增加系统复杂度，而且在复杂或动态场景中可能无法完全消除遮挡。与这些方法不同，我们的方案利用生成模型推断被遮挡部分，从而解决遮挡问题，同时保持较低的系统复杂度。

RIS辅助无线全模态感知的主要挑战在于优化RIS相移配置，因为RIS相移与全模态感知精度之间的关系复杂，且其闭式表达式未知。为应对这一挑战，我们提出了一种深度

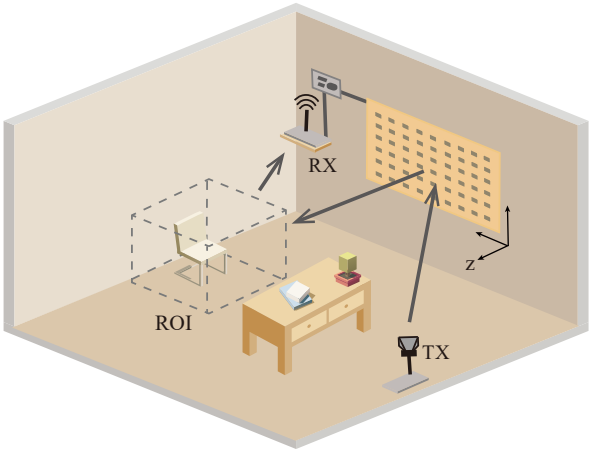


图1. RIS辅助无线非模态感知系统的系统模型。

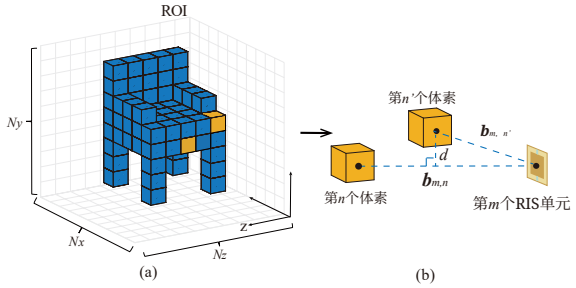


图2. 目标物体及体素间遮挡的示意图。

神经网络（DNN）模型，用于近似RIS相移与全模态感知重建误差之间的关系，从而能够基于这一关系有效优化RIS相移配置。在基准数据集上的实验结果表明，在RIS配置数量相同的情况下，与基线方法相比，我们的方法将重建误差至少降低了56.73%。

II. 系统模型

A. 场景描述

如图 1 所示，我们考虑一种由 RIS 辅助的无线全模态感知系统，该系统包含一个单天线发射机（Tx）、一个单天线接收机（Rx）、感兴趣区域（ROI）内的目标物体，以及一个垂直部署且具有 M 个可重构单元的 RIS。Tx 与 ROI 之间存在其他物体，会遮挡目标物体。为充分利用 RIS 波束成形，Tx 将波束指向 RIS，并且仅使用经 RIS 反射的链路进行感知。为实现全模态感知，Tx 首先向 RIS 发射信号，由 RIS 将信号反射至 ROI 内的物体。随后，Rx 接收散射信号，并利用这些信号重建物体的完整三维形状。

B. 物体模型

如图 2(a) 所示，感兴趣区域被离散化为 $N = N_x \times N_y \times N_z$ 个体素（每个体素的尺寸为 $\xi_x \times \xi_y \times \xi_z$）。第 n 个体素的散射系数记为ω_n，整个感兴趣区域的散射系数可表示为 $\boldsymbol{\omega} = [\omega_1, \omega_2, ..., \omega_N]^T$。此外，令 $\boldsymbol{\chi} = [\chi_1, \chi_2, ..., \chi_N]^T$表示物体的形状，其中 χ_n 表示第 n 个体素是否被物体占据。给定 ω_n， χ_n 的值可由下式确定

$$\chi_n = 1 \text{ if } |\omega_n| \geq \tau_\omega, \text{otherwise } 0. \tag{1}$$

其中， τ_ω 表示用于确定体素是否被占据的阈值。

从 RIS 和 Rx 的角度来看，被占据体素可能会阻挡从 RIS/Rx 到其他体素的路径。令 $v_m^{RIS}{}_{,n}$ 和 v_n^{Rx}分别表示从第 m 个 RIS 单元到第 n 个体素，以及从第 n 个体素到 Rx 的路径是否被遮挡（1：未遮挡，0：被遮挡）。这些值取决于其他非空体素的位置。具体而言，若要阻挡从第 m 个 RIS 单元到第 n 个体素的路径，即 $v_m^{RIS}{}_{,n} = 0$，非空体素（记为第 n' 个体素）的位置必须同时满足以下条件 [9]:

1) n'号体素与从第 m个 RIS 单元到第 n个体素的路径之间的距离 d 小于阈值 τ_d，即 $d = |\boldsymbol{b}_{m,n'} \times \boldsymbol{b}_{m,n}| / |\boldsymbol{b}_{m,n}| < \tau_d$，如图 2(b) 所示。2) $\boldsymbol{b}_{m,n'}$ 与$\boldsymbol{b}_{m,n}$ 之间的夹角为锐角，即$\boldsymbol{b}_{m,n'} \cdot \boldsymbol{b}_{m,n} > 0$。3) $\boldsymbol{b}_{m,n'}$ 在$\boldsymbol{b}_{m,n}$方向上的投影小于$\boldsymbol{b}_{m,n}$的长度，即$|\boldsymbol{b}_{m,n'} \cdot \boldsymbol{b}_{m,n}| < |\boldsymbol{b}_{m,n}|^2$。

v_n^{Rx}的值可以用类似的方法推导得出。

如果从 RIS 或接收机到某个体素的所有路径都被阻断，导致该体素对感知系统不可见，则称该体素被完全遮挡 。令体素遮挡向量 $\boldsymbol{v} = [v_1, v_2, ..., v_N]^T$表示体素可见性（$v_n = 0$：不可见，$v_n = 1$：可见）。路径遮挡矩阵 $\boldsymbol{V} \in \mathbb{C}^{M \times N}$的第（$m$，$n$）个元素为$V_{m,n} = v_m^{RIS}{}_{,n} v_n^{Rx}$，用于确定从第 m个 RIS 单元经由第 n个体素到接收机的路径是否被阻断。于是， v_n的值为:

$$v_n = 0 \text{ if } \sum_{m=1}^{M} V_{m,n} = 0, \text{otherwise } 1. \tag{2}$$

然后，物体形状 $\boldsymbol{\chi}$ 可分为可见部分$\boldsymbol{\chi_v} = \boldsymbol{\chi} \odot \boldsymbol{v}$和不可见部分 $\boldsymbol{\chi_i} = \boldsymbol{\chi} \odot (\boldsymbol{i} - \boldsymbol{v})$，其中$\boldsymbol{i} = [1, 1, ..., 1]^T \in \mathbb{R}^{N \times 1}$和 $\boldsymbol{\chi} = \boldsymbol{\chi_v} + \boldsymbol{\chi_i}$。类似地，可见部分和不可见部分的散射系数分别表示为 $\boldsymbol{\omega_v} = \boldsymbol{\omega} \odot \boldsymbol{v}$和 $\boldsymbol{\omega_i} = \boldsymbol{\omega} \odot (\boldsymbol{i} - \boldsymbol{v})$。

C. 基于*RIS*的测量模型

RIS包含 M 个RIS单元（尺寸为 $\xi_s \times \xi_s$），每个单元都集成了PIN二极管，可根据施加的偏置电压切换状态。通过控制PIN二极管的状态，每个单元对入射信号施加的相移可以进行重构。假设每个单元可提供 2^b 种离散相移，则第 m个RIS单元的相移 ϕ_m表示为 $\phi_m \in \mathcal{F} = \left\{ \underbrace{0, \frac{2\pi}{2^b}, ..., \frac{2\pi(2^b-1)}{2^b}}_{2^b} \right\}$。

发射机-RIS-感兴趣区域-接收机路径的信道系数 h_{ROI} 可建模为所有 RIS 单元和感兴趣区域体素散射的所有路径之和 [8]:

$$h_{ROI}(\boldsymbol{\omega},\boldsymbol{q},\boldsymbol{V}) = \sum_{m=1}^{M}\sum_{n=1}^{N}\frac{\lambda\sqrt{GK_m\xi_s\xi_sG_IF_m^AF_{m,n}^D}}{(4\pi)^{\frac{3}{2}}d_{t,m}^{\alpha}d_{m,n}^{\alpha}d_{n,r}^{\alpha}}e^{-j\frac{2\pi}{\lambda}(d_{t,m}+d_{m,n}+d_{n,r})} \times e^{-j\phi_m}V_{m,n}\omega_n, \tag{3}$$

其中，$\boldsymbol{q}=[e^{-j\phi_1},e^{-j\phi_2},...,e^{-j\phi_M}]^T$为 RIS 相移配置。$\lambda$ 为波长。G和 G_I 分别对应发射机和 RIS 单元的天线增益。K_m 表示发射机天线在指向第 m 个 RIS 单元方向上的归一化功率增益。F_m^A 和$F_{m,\,n}^D$ 分别为第 m 个 RIS 单元针对来自发射机的入射信号和发往第 n 个感兴趣区域体素的离射信号的归一化功率辐射方向图。$d_{t,\,m}$、$d_{m,\,n}$和 $d_{n,\,r}$ 分别表示发射机与第 m 个 RIS 单元之间、第 m 个 RIS 单元与第 n 个感兴趣区域体素之间，以及第 n 个感兴趣区域体素与接收机之间的距离。α 为路径损耗指数。

由于发射机-RIS-感兴趣区域-接收机路径包含物体形状信息，因此首先使用现有方法根据接收信号测量其信道增益 [5], [6], [8] ，然后利用该增益重建物体形状。h_{ROI} 的测量模型可表示为 $r=h_{ROI}+z$，其中$z\sim\mathcal{N}(0,\ \varepsilon_z)$ 为测量噪声 [8]。根据式（3），由于不可见体素（满足 $\sum_{m=1}^{M}V_{m,\,n}=0$）不会影响$h_{ROI}$，测量可简化为:

$$r=\boldsymbol{q}^T(\boldsymbol{H}_r\odot\boldsymbol{V})\boldsymbol{\omega_v}+z, \tag{4}$$

其中，$\boldsymbol{H}_r\in\mathbb{C}^{M\times N}$ 表示发射机-RIS-感兴趣区域-接收机路径的自由空间信道系数。

D. 感知协议

所提出的RIS辅助全模态感知协议分为三个阶段:

1) 感兴趣区域获取：采用现有定位技术 [8]估计感兴趣区域的位置（x_{ROI}，y_{ROI}，z_{ROI}）和大小（l_x，l_y，l_z）。具体而言，RIS引导反射波束扫描整个空间，接收机分析来自不同方向的回波信号强度和时延，以获取感兴趣区域的位置和大小。

2）测量：K 通过第三节所述方法对RIS配置进行优化。随后，RIS采用 K优化后的相位配置对感兴趣区域进行精确扫描。对于每种配置 k，接收机记录 r_k 关于 h_{ROI}的测量值。

3）全模态感知：基于 K 测量值，采用第四节详细介绍的全模态感知算法重建物体的完整三维形状。

III. RIS CONFIGURATION OPTIMIZATION FOR AMODEL SENSING

A. 问题 *formulation*

本节优化 RIS 相移配置，以最小化可见和不可见形状的全模态感知重建误差。优化问题表述如下:

$$\min_{\boldsymbol{q}_1,\boldsymbol{q}_2,...,\boldsymbol{q}_K}\quad a \tag{5a}$$

$$\text{s.t. } \phi_{k,m}\in\mathcal{F},\forall k\in[1,K],\forall m\in[1,M], \tag{5b}$$

其中，a表示全模态感知重建误差。$\boldsymbol{q}_k=[e^{-j\phi_{k,1}},e^{-j\phi_{k,2}},...,e^{-j\phi_{k,M}}]^T$表示第 k 个 RIS 相移配置，其中$\phi_{k,\ m}$ 为第 m 个 RIS 单元的相位。约束 (5 b) 要求每个 RIS 单元的相移都从可用集合$\mathcal{F}$中选取。

B. 基于学习的*RIS*配置优化算法

由于全模态感知重建误差的闭式表达式未知，难以直接求解RIS 配置优化问题。此外，现有方法采用的信道相关性最小化指标 [7], [8] 由于以下两个原因无法应用于全模态感知：1）该指标仅关注可见形状的重建误差，而忽略了不可见部分；2）计算信道相关性时忽略了体素遮挡。为此，我们训练了一个基于监督学习的误差预测模型，以学习RIS配置到全模态感知重建误差的映射关系，并据此优化RIS配置。

1）模型架构：所提出的误差预测模型能够估计给定 RIS配置下的全模态感知重建误差。基于该预测结果，可通过梯度下降优化 RIS配置，以最小化误差。

具体而言，该模型采用深度神经网络来预测给定感兴趣区域和 RIS配置下的全模态感知重建误差 a 。该模型由一个 7 层全连接神经网络组成，各层维度为（$L+1$）-32-64-128-64-32-1。深度神经网络的输入是从 RIS 测量信道中提取的 $L+1$ 相关性特征。RIS配置 $\boldsymbol{Q}=[\boldsymbol{q}_1,\boldsymbol{q}_2,...,\boldsymbol{q}_K]^T\in\mathbb{C}^{K\times M}$和自由空间信道系数 $\boldsymbol{H}_r\in\mathbb{C}^{M\times N}$ 不会直接输入网络，因为它们的维度很高且配置空间很大，会显著增加训练复杂度。因此，我们提取复合信道矩阵 $\boldsymbol{QH}_r$ 的全局和 L 局部相关性，并使用低维输入训练模型，以降低训练复杂度，其中我们有 $L<<K$，M，N。全局相关性定义为 [7]

$$c_0=\|R(\boldsymbol{QH}_r)-\boldsymbol{I}_N\|_F/(N^2-N), \tag{6}$$

其中，$R(\boldsymbol{A})$表示矩阵 $\boldsymbol{A}$的归一化列向相关矩阵，其第(i,j)个元素由 $R(\boldsymbol{A})_{i,j}=|\boldsymbol{A}(:\ i)^T\boldsymbol{A}(:\ j)|/\|\boldsymbol{A}(:\ i)\|_2\|\boldsymbol{A}(:\ j)\|_2$给出，表示 $\boldsymbol{A}$的第 i列与第 j列之间的相关性。为计算局部相关性，将感兴趣区域划分为 L 个连续的立方子区域，每个子区域包含集合 $\mathcal{N}_l$中的 N_l个相邻体素。对于第 l个子区域，局部相关性由 $c_l=\|R(\boldsymbol{QH}_r(:\ \mathcal{N}_l))\ -\boldsymbol{I}_{N_l}\|_F/(N_l^2-N_l)$ [7]计算。这 $L+1$ 个相关性描述了全局与局部感兴趣区域中RIS至体素信道响应的可区分性 [5],从而促进从输入到重建误差的映射。

2) 模型训练：本部分重点介绍如何训练误差预测模型。首先，我们模拟具有随机分布感兴趣区域和 RIS配置的多种设置，以构建训练数据集。对于每个

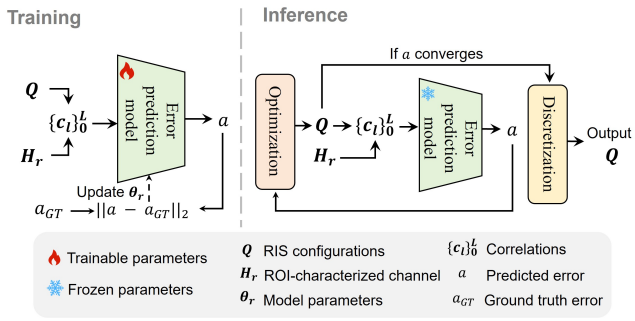


图 3. 所提出的用于 RIS 配置优化模型的训练与推理。

设置，计算相关性 $\{c_l\}_0^L$。为生成相应的真实标签，我们在相同设置下对一组ShapeNet [10]物体进行测量并应用非模态感知模型，然后计算这些物体的平均重建误差，以评估该设置对任意物体的整体重建性能。误差预测模型最终经过训练，以最小化预测误差与真实误差之间的均方误差（均方误差），如图3的训练相位所示。

3) 模型推理：如图 3 的推理阶段所示，从随机初始化的 $\boldsymbol{Q}$开始，$\boldsymbol{Q}$ 根据预训练误差预测模型预测的误差进行迭代优化。在每次迭代中，首先由误差预测模型评估当前的 $\boldsymbol{Q}$，以计算误差。然后，计算预测误差相对于 $\boldsymbol{Q}$的梯度，并通过梯度下降优化 $\boldsymbol{Q}$。该迭代过程持续进行，直至预测误差收敛。收敛后，将优化后的 $\boldsymbol{Q}$的每个元素离散化为最接近的离散值，以满足式 (5b) 中的约束。

IV. RIS辅助三维重建

A. 问题 *formulation*

本节旨在重建感兴趣区域（ROI）内目标物体的三维形状。由于根据式 (4)，测量值 $\{r_k\}_{k\in[1,K]}$只能捕获可见部分，因此我们提出依次求解以下两个问题。

1) 可见形状重建问题：在此问题中，我们根据测量值$\{r_k\}_{k\in[1,K]}$ 重建可见形状，以尽可能减小重建的可见形状 $\tilde{\boldsymbol{\chi}}_{\boldsymbol{v}}$ 与实际可见形状 $\boldsymbol{\chi}_{\boldsymbol{v}}$之间的误差。该问题可表示为

$$\min_{\tilde{\boldsymbol{\chi}}_{\boldsymbol{v}}} \|\tilde{\boldsymbol{\chi}}_{\boldsymbol{v}} - \boldsymbol{\chi}_{\boldsymbol{v}}\|_1, \ \text{s.t.}\ \tilde{\chi}_{v,n} \in \{0,1\}, \tilde{\boldsymbol{\chi}}_{\boldsymbol{v}} = \tilde{\boldsymbol{\chi}}_{\boldsymbol{v}} \odot \boldsymbol{v}. \quad (7)$$

约束条件规定，$\tilde{\boldsymbol{\chi}}_{\boldsymbol{v}}$中第 n个元素$\tilde{\chi}_{v,\ n}$ 必须为二进制值（0 或 1），且对应于不可见体素的元素必须为 0。

2) 完整形状重建问题：在该问题中，我们通过非模态感知恢复完整形状，以最小化重建形状与实际形状之间的误差，即分别为 $\tilde{\boldsymbol{\chi}}$ 和 $\boldsymbol{\chi}$。该问题可表示为

$$\min_{\tilde{\boldsymbol{\chi}}} \|\tilde{\boldsymbol{\chi}} - \boldsymbol{\chi}\|_1, \ \text{s.t.}\ \tilde{\chi}_n \in \{0,1\}, \tilde{\boldsymbol{\chi}}_{\boldsymbol{v}} = \tilde{\boldsymbol{\chi}} \cap \tilde{\boldsymbol{\chi}}_{\boldsymbol{v}}. \quad (8)$$

该约束确保 $\tilde{\boldsymbol{\chi}}$ 中的第 n 个元素 $\tilde{\chi}_n$ 仅取二进制值，并基于恢复出的可见部分重建完整形状。

B. 可见形状重建算法设计

通过压缩感知（CS）技术 [5]，可以根据测量值 $\{r_k\}_{k\in[1,K]}$ 重建$\boldsymbol{\omega_v}$，从而恢复可见形状 $\boldsymbol{\chi_v}$。首先，将问题（7）重新表述为一个 CS 问题。具体而言，CS 重建方程可表示为

$$\boldsymbol{r} = \boldsymbol{Q}(\boldsymbol{H}_r \odot \boldsymbol{V})\boldsymbol{\omega_v} + \boldsymbol{z}, \quad (9)$$

其中 $\boldsymbol{r} = [r_1, r_2, ..., r_K]^T$。根据散射体的稀疏性，我们可以通过最小化 $\boldsymbol{\omega_v}$ [5]的 l_1范数来求解 $\boldsymbol{\omega_v}$。此外，为了控制原目标函数的值，我们添加一个约束：重建的 $\boldsymbol{Q}$（$\boldsymbol{H}_r \odot \boldsymbol{V}$）$\boldsymbol{\omega_v}$ 与测量值$\boldsymbol{r}$ 之间的误差不得超过预定义阈值 ε_z，其中 ε_z 是式（9）中 $\boldsymbol{z}$ 的方差 [5]。因此，问题（7）转化为以下形式 [5]:

$$\min_{\boldsymbol{\chi_v}, \boldsymbol{\omega_v}} \|\boldsymbol{\omega_v}\|_1, \quad (10a)$$

$$\text{s.t.}\ \|\boldsymbol{r} - \boldsymbol{Q}(\boldsymbol{H}_r \odot \boldsymbol{V})\boldsymbol{\omega_v}\|_2 \le \varepsilon_z, \quad (10b)$$

$$\chi_{v,n} = 1 \ \text{if} \ |\omega_{v,n}| \ge \tau_\omega, \text{otherwise}\ 0, \quad (10c)$$

$$\boldsymbol{\omega_v} = \boldsymbol{\omega_v} \odot \boldsymbol{v}, \quad (10d)$$

与传统压缩感知问题不同，(10)在压缩感知测量矩阵$\boldsymbol{Q}$ ($\boldsymbol{H}_r \odot \boldsymbol{V}$)中包含未知遮挡变量 $\boldsymbol{V}$，因此诸如 GAMP [11] 等传统压缩感知方法并不适用。为此，我们提出一种遮挡感知GAMP算法，该算法迭代求解$\boldsymbol{\omega_v}$，并基于第II-B节所述的体素遮挡关系更新 $\boldsymbol{V}$，从而实现对可见形状 $\boldsymbol{\chi_v}$的准确恢复。

令 $\boldsymbol{\omega_v}^{(i)}$和 $\boldsymbol{V}^{(i)}$ 分别表示第 i次迭代中 $\boldsymbol{\omega_v}$ 和 $\boldsymbol{V}$ 的解。初始化时，我们假设所有路径均无遮挡，即 $\boldsymbol{V}^{(0)}$ 为全1。随后，我们迭代求解 $\boldsymbol{\omega_v}$ 和 $\boldsymbol{V}$。具体而言，每次迭代包含以下两个步骤：

步骤1： 给定 $\boldsymbol{V}^{(i-1)}$，执行GAMP算法 [11] 以得到$\hat{\boldsymbol{\omega}}_{\boldsymbol{v}}^{(i)}$，并通过 $\boldsymbol{\omega_v}^{(i)} = \hat{\boldsymbol{\omega}}_{\boldsymbol{v}}^{(i)} \odot \boldsymbol{v}^{(i-1)}$ 计算可见散射系数，以满足约束(10d)。此处，$\boldsymbol{v}^{(i-1)}$ 根据(2)由 $\boldsymbol{V}^{(i-1)}$ 计算得到。

步骤 2： 基于 $\boldsymbol{\omega_v}^{(i)}$，通过检查路径是否被其他占用体素阻挡来更新 $\boldsymbol{V}^{(i)}$，如第 II-B 节所示。

这一迭代过程持续进行，直到$\|\boldsymbol{V}^{(i)} - \boldsymbol{V}^{(i-1)}\|_F \le \varepsilon_v$。最后，通过约束条件 (10c) 恢复可见形状 $\boldsymbol{\chi_v}$。可见形状重建过程总结于算法1中。

C. 形状补全算法设计

形状补全问题可以视为一个条件生成任务，其目标是根据可见形状恢复完整三维形状。为解决这一问题，我们设计并训练了一个条件扩散模型。

1) 模型架构：条件扩散模型通过在 T 个步骤中逐步对随机噪声形状 $\boldsymbol{\chi}_T$ 进行去噪，在以下内容的引导下生成完整形状 $\boldsymbol{\chi}_0$：

算法1： 可见形状重建算法

输入： 测量 $\boldsymbol{r}$、配置 $\boldsymbol{Q}$、信道 $\boldsymbol{H}_r$.

输出： 可见形状 $\boldsymbol{\chi_v}$。

初始化： $i=0$，$\boldsymbol{V}^{(0)}=\mathbf{1}_{M\times N}$。

重复

  $i=i+1$。

  执行 GAMP [11]以获得 $\hat{}\boldsymbol{\omega_v}^{(i)}$。

  $\boldsymbol{\omega_v}^{(i)}=\hat{}\boldsymbol{\omega_v}^{(i)}\odot\boldsymbol{v}^{(i-1)}$，其中 $\boldsymbol{v}^{(i-1)}$ 来自 (2)。

  对于 $n=1:\ N$ **执行**

    检查路径遮挡以确定 v_m^{RIS} ,n 以及 v_n^{Rx}.

  end

  根据 V_m更新 $\boldsymbol{V}^{(i)}$,$n=v_m^{RIS},\ n v_n^{Rx}$。

直到 $\|\boldsymbol{V}^{(i)}-\boldsymbol{V}^{(i-1)}\|_F\le\varepsilon_v$。

根据 (10c) 得出 $\boldsymbol{\chi_v}$ 。

可见形状 $\boldsymbol{\chi_v}$的条件。$\boldsymbol{\chi}_t$ 表示第 t 步的中间受损形状。该模型由两个分支组成：执行去噪操作的主分支，以及从条件输入 $\boldsymbol{\chi_v}$ 中提取特征并将其注入主分支、以引导去噪过程的控制分支。

主分支：该分支负责对输入的受损形状进行迭代去噪。它基于 Stable Diffusion 框架 [13]，，采用 U-Net 架构，由多个编码器块、一个中间块和多个解码器块组成，并通过跳跃连接连接相应的编码器块与解码器块。主分支的输入是在第 t 步的受损形状 $\boldsymbol{\chi}_t$，经过处理后生成去噪后的形状$\boldsymbol{\chi}_{t-1}$。

控制分支：该分支用于处理条件 $\boldsymbol{\chi_v}$，以引导去噪过程。它与主分支采用相同的结构，但使用独立的参数。输入 $\boldsymbol{\chi_v}$ 由其自身的编码器块和中间块处理，以提取多尺度空间特征。每个块处理后，提取的特征会通过投影层，以逐体素加法融合的方式注入主分支的解码器块中。这种融合策略确保可见形状与生成形状的特征之间实现空间对齐，从而提高形状补全准确率。

2）模型训练：按照去噪扩散概率模型（DDPM）框架 [12]，，模型训练包括前向过程和反向过程。给定一对可见形状和真实完整形状，完整形状在前向过程中被逐步扰乱为噪声，并在反向过程中以可见形状为条件进行去噪。这两个过程的具体细节如下。

在前向过程中，经过 $\boldsymbol{\chi}_0$ over T 个时间步，逐步向完整形状 $\boldsymbol{\chi}_T$ 添加高斯噪声，以生成其噪声版本；转移为 $q(\boldsymbol{\chi}_t|\boldsymbol{\chi}_{t-1})=\mathcal{N}(\sqrt{1-\beta_t}\boldsymbol{\chi}_{t-1},\ \beta_t\boldsymbol{I})$，其中 $\beta_t\in(0,\ 1)$ 控制每个时间步 t 的噪声水平，而 $\boldsymbol{\chi}_t$ 表示时间步 t 时得到的受损形状。在反向过程中，一个去噪——网络利用学习得到的参数 $\boldsymbol{\theta}$，在以 $\boldsymbol{\chi_v}$ 为条件的情况下逐步去除 $\boldsymbol{\chi}_T$ 中的噪声，从而恢复 $\boldsymbol{\chi}_0$；其转移为 $p_{\boldsymbol{\theta}}(\boldsymbol{\chi}_{t-1}|\boldsymbol{\chi}_t,\ \boldsymbol{\chi_v})=\mathcal{N}(\boldsymbol{\mu_\theta}(\boldsymbol{\chi}_t,\ t,\ \boldsymbol{\chi_v})\ \sigma_t\boldsymbol{I})$。其中，$\boldsymbol{\mu_\theta}$ 表示当前时间步 t 的网络预测均值，σ_t 表示方差。按照 DDPM [12],，预测均值 $\boldsymbol{\mu_\theta}$ 可以简化为预测反向过程中要去除的噪声，以近似前向过程中添加的噪声。因此，模型通过最小化预测噪声 $\boldsymbol{\epsilon_\theta}(\boldsymbol{\chi}_t,\ t,\ \boldsymbol{\chi_v})$ 与实际添加噪声 $\boldsymbol{\epsilon}$ 之间的均方误差进行训练。

V. 实验结果

本节对所提出的方案进行评估。RIS位于 x-y 平面内，中心位于原点；Tx和Rx分别位于（-2.12，0，2.12）m和（2，0，0）m处。感兴趣区域（ROI）的中心位于（0，0，2.5）m处，尺寸为 $3\times3\times3\text{m}^3$,，默认划分为 $N=10^3$个体素。Tx发射功率为35 dBm [5],，载波频率为 $f=3GHz$ [5]。噪声功率为 $\varepsilon_z^2=-104$ dBm [5]。RIS单元间距为0.05 m，每个单元具有2比特相位 [5]。被占据体素的散射系数为1，空体素的散射系数为0。遮挡阈值为 $\tau_w=0.5$ 和$\tau_d=\min\{\xi_x,\xi_y,\xi_z\}$ [9]。算法1的收敛阈值为 $\varepsilon_v=0.01\sqrt{MN}$。实验中，从ShapeNet [10]的10个类别中选取31,000个物体，生成93,000个目标，并将其置于不同的观察角度。为训练形状补全生成模型，我们将目标转换为分辨率为 32^3,的完整形状，从而支持将补全结果下采样至不同的 N。通过模拟RIS观测获得相应的可见形状。扩散模型采用$T=1000$ 个步骤，β_t 从0.0001线性递增至0.02。随后，基于训练好的形状补全生成模型，按照第三节B部分所述，训练用于RIS配置优化的误差预测模型。

在图4(a)中，我们将所提出的 RIS 辅助全模态感知方案与两种现有基线进行比较：

- GAMP-Vision：一种不使用 RIS 的全模态感知方案，其中物体有 20% 的部分被遮挡。该方案通过 GAMP [11] 估计可见形状，并使用基于视觉的算法 [3]补全完整形状。
- MetaSketch [5]：一种 RIS 辅助的单视角感知方案，通过压缩感知技术重建散射系数。

可以看出，所提出的方案在不同分辨率下的误差均最低；当 $N=10\times10\times10$时，其误差比 GAMP-Vision 低 65.54%，比 MetaSketch 低 56.73 %。这是因为在 GAMP-Vision 方案中，由于没有 RIS 的辅助，物体与收发器之间的无线链路会受到部分遮挡，而 MetaSketch 无法恢复物体的不可见部分。相比之下，我们的方法利用 RIS 绕过遮挡，并采用生成模型推断不可见部分，因此三维形状重建误差更低。

图4(b)展示了不同方案获得的物体形状重建效果示例。由于遮挡且缺少 RIS 辅助，GAMP-Vision 的结果存在严重误差，物体形状几乎无法

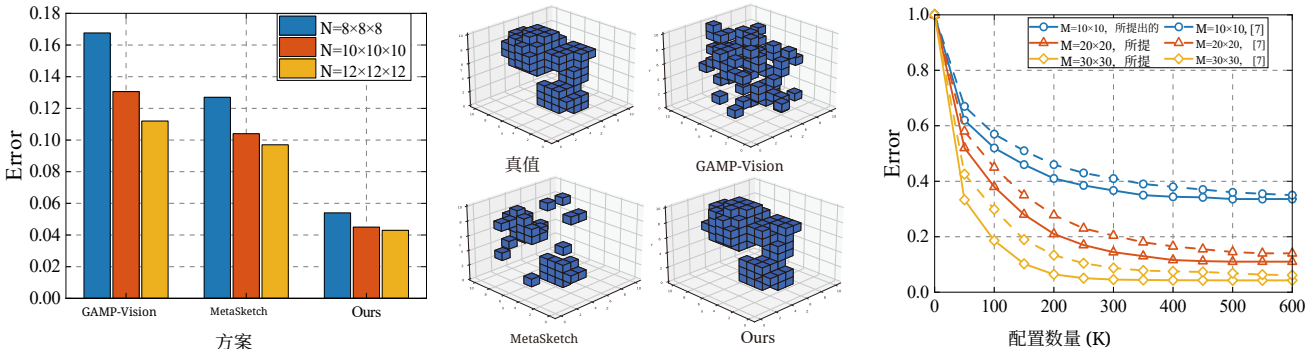


图4。(a) 所提方案与其他感知方案的重建误差。(b) 不同方案下灯的重建结果（N = 10×10×10）。(c) 我们的RIS配置优化方案与基于相关性最小化的方案的重建误差。

辨认。虽然 RIS辅助MetaSketch 能够重建可见部分，但无法恢复不可见部分，导致三维重建不完整且不准确。相比之下，我们提出的方法生成的重建结果与真值高度吻合，能够有效捕捉物体的可见部分和不可见部分。

图4(c)展示了重建误差随RIS配置数量的变化，以评估所提RIS配置优化方法的性能。作为对比，我们还给出了通过最小化信道相关性 [7] 来优化RIS配置所获得的结果，而非采用所提出的基于监督学习的方法。随着配置数量 K增加，重建误差逐渐减小并趋于收敛。RIS尺寸 M增大时，误差也会减小；随着 M继续增大，性能增益则逐渐变小。在所有设置下，我们的方法始终优于基于相关性最小化的方案。值得注意的是，在固定 M的情况下，与现有方案相比，我们的方法使用更少的配置即可达到更低的误差，这体现了该方法在优化时考虑可见和不可见形状的全模态感知重建误差所带来的效率优势。

六、结论

本文研究了一种RIS辅助全模态感知系统，该系统借助大规模RIS和生成式人工智能模型，恢复部分被遮挡物体的完整形状。鉴于RIS配置与全模态感知重建误差之间的关系复杂，且不存在闭式表达式，我们提出了一种基于学习的RIS配置优化方案，用于近似这一关系，并据此优化RIS相移配置。实验结果得出两个主要结论：

1）与现有单视角感知方案相比，所提出的RIS辅助的模态外感知方案可将重建误差至少降低56.73%。
2）随着感兴趣区域的分辨率和RIS配置数量增加，重建误差会降低。

致谢

本研究部分得到了美国国家科学基金会（项目编号：62371011）、北京市自然科学基金（项目编号：L243002）、广东省基础与应用基础研究基金（项目编号：2023B0303000019）、深圳市科技计划（项目编号：JCYJ20241202125910015）以及北京市杰出青年科学家计划（项目编号：JWZ020240102001）的资助。

参考文献

[1] Y. Yue 等，“GridNet-3D：一种基于点云的新型实时三维目标检测算法”，《中国电子学报》，第30卷，第5期，931–939页，2021年9月。[2] B. Nanay，“日常感知中非模态补全的重要性”，《*i-Perception*》，第9卷，第4期，文章编号2041669518788887，2018年7月。[3] R. Chu、E. Xie、S. Mo 等，“DiffComplete：基于扩散的生成式三维形状补全”，载于《神经信息处理系统进展》（*NeurIPS*），美国洛杉矶，2023年12月，第75951–75966页。[4] X. Yu 等，“AdaPoinTr：采用自适应几何感知变换器的多样化点云补全”，《*IEEE*模式分析与机器智能汇刊》，第45卷，第12期，14114–14130页，2023年12月。[5] J. Hu、H. Zhang、K. Bian 等，“MetaSketch：基于可重构智能表面的无线语义分割”，《*IEEE*无线通信汇刊》，第21卷，第8期，5916–5929页，2022年8月。[6] X. Tong 等，“利用信号和环境稀疏性的多用户通信与感知联合方法”，《*IEEE*信号处理专题精选期刊》，第15卷，第6期，1409–1422页，2021年11月。[7] S. Zhu、Z. Yu、Q. Guo、J. Ding、Q. Cheng和T. J. Cui，“RIS辅助的联合上行通信与成像：相位优化与贝叶斯回波解耦”，2023年，arxiv:2301.03817。[8] Y. Huang、J. Yang、C. K. Wen和S. Jin，“利用多视角图像相关性的RIS辅助单频三维成像”，《*IEEE*通信汇刊》，第72卷，第8期，5003–5018页，2024年8月。[9] X. Tong 等，“考虑遮挡效应的环境感知：一种多视角方法”，《*IEEE*信号处理汇刊》，第70卷，3598–3615页，2022年6月。[10] A. X. Chang 等，“ShapeNet：信息丰富的三维模型库”，*arXiv*预印本 *arXiv:1512.03012*，2015年。[11] S. Rangan，“采用随机线性混合进行估计的广义近似消息传递”，《*IEEE*国际信息论研讨会论文集》，2011年，第2168–2172页。[12] J. Ho、A. Jain、P. Abbeel，“去噪扩散概率模型”，《神经信息处理系统进展》，2020年，第6840–6851页。[13] L. Zhang、A. Rao、M. Agrawala，“为文本到图像扩散模型添加条件控制”，《*IEEE/CVF*计算机视觉国际会议论文集》，2023年，第3836–3847页。