

文章编号 1004-924X(2026)09-1496-11

增强反射的被动非视域成像 U-NET 重构

于相之, 金伟其*, 裘 溯, 李 力

(北京理工大学 光电学院 光电成像技术与系统教育部重点实验室, 北京 100081)

摘要:非视域(Non-Line-of-Sight, NLOS)成像可通过中介面捕获障碍物后方目标的间接光信号以重建隐藏场景, 在安防、自动驾驶等领域应用广泛。针对被动非视域成像有效信号弱、噪声强及现有方法未充分利用中介面非朗伯反射特性导致重建精度不足的问题, 搭建了不同材质瓷砖中介面的被动 NLOS 成像实验系统, 并构建含 8 组超 80 万张投影图像的 NLOS-Passive-TILES-stl10 数据集; 通过提出增强反射的改进 U-NET 架构(ER-UNET)被动 NLOS 成像方法, 整合反射特征提取器、可微分 CLAHE、自适应实例归一化等模块, 将 RGB 转换为 $Y, R-Y, B-Y$ 三通道并强化 Y 通道权重, 经编码-解码提取多尺度特征并做 SVD 低秩分解, 以实现梯度感知 RGB 重建。实验结果表明, ER-UNET 在 NLOS-PASSIVE 测试集上平均 PSNR 达 15.14 dB, SSIM 为 0.50, 优于 NLOS-OT 与 C-GAN 算法, 在人脸、图标等测试数据上泛化能力优异, 为被动 NLOS 成像实用化提供了新思路。

关键词:非视域成像; 被动成像; U-NET; NLOS 数据集; 瓷砖; 中介面; 反射特性

中图分类号: TP391.4; TN249 **文献标识码:** A

doi: 10.37188/OPE.20263409.1496 **CSTR:** 32169.14.OPE.20263409.1496

Enhanced reflection U-Net reconstruction for passive Non-Line-of-Sight imaging

YU Xiangzhi, JIN Weiqi*, QIU Su, LI Li

(MOE Key Laboratory of Optoelectronic Imaging Technology and System, School of Opto-Electronics, Beijing Institute of Technology, Beijing 100081, China)

* Corresponding author, E-mail: jinwq@bit.edu.cn

Abstract: Non-line-of-sight (NLOS) imaging reconstructs hidden scenes by capturing indirect optical signals reflected from targets occluded by obstacles via intermediate surfaces, and has broad applications in security monitoring, autonomous driving, and related fields. Existing methods often suffer from weak effective signals, strong noise, and limited reconstruction accuracy due to insufficient exploitation of the non-Lambertian reflection characteristics of intermediate surfaces. To address these limitations, a passive NLOS imaging experimental system was constructed using tile intermediate surfaces of different materials, and the NLOS-Passive-TILES-stl10 dataset was established, comprising eight groups and more than 800 000 projected images. An enhanced-reflection improved U-Net architecture (ER-UNET) was proposed for passive NLOS imaging, incorporating a reflection feature extractor, differentiable CLAHE, adaptive instance normalization, and other modules. RGB images are transformed into $Y, R-Y,$ and $B-Y$

收稿日期: 2026-02-05; 修订日期: 2026-04-14.

基金项目: 国家自然科学基金面上项目 (No. 62171024)

channels, with an increased weight assigned to the Y channel. Multi-scale features are extracted through an encoder-decoder structure, and SVD-based low-rank decomposition is performed to enable gradient-aware RGB reconstruction. Experimental results demonstrate that ER-UNET achieves an average PSNR of 15.14 dB and an SSIM of 0.50 on the NLOS-Passive test set, outperforming NLOS-OT and C-GAN. The proposed method also exhibits strong generalization ability on test data such as faces and icons, providing a promising approach for the practical application of passive NLOS imaging.

Key words: NLOS imaging; passive imaging; U-NET; NLOS datasets; ceramic tiles; intermediate surfaces; reflection characteristics

1 引言

非视域(Non-Line-of-Sight, NLOS)成像技术通过中介面反射的间接光信号捕获物体,以重建被障碍物遮挡场景的信息,这突破了传统光学成像的视野限制^[1],可用于自动驾驶^[2]、光学隐蔽侦察^[3]及医学治疗^[4]等领域。根据是否使用可控光源照明,NLOS成像可分为主动和被动NLOS成像两类。其中,前者采用激光主动照射中介面,通过主动发射脉冲或调制光信号照射在景物上,并捕捉经中介面散射的景物反射时空信息,结合高精度探测与先进的重构算法,实现对视域外隐藏目标的三维成像与状态感知^[1,5-8];后者则无需主动光源照明,仅利用环境光进行成像^[9],具有更高的隐蔽性和安全性,在侦察、救援和自动驾驶等领域具有重要的应用价值。

然而,被动NLOS成像的探测照明不可控,各向同性漫反射导致像素间关联性较强,成像中光传输矩阵的条件数极大^[10],这导致从观测数据中难以获得理想的重建结果。为缓解这一问题,研究人员提出了多种方法,包括放置局部遮挡物^[11-12]、使用偏振器^[13]以及应用深度学习^[14-16]等。其中,基于深度学习的被动NLOS成像^[17-18]凭借深度神经网络强大的表征能力备受关注,这种能力可显著提升重建分辨率。但是,将深度学习应用于被动NLOS成像仍面临若干挑战。首先,现有较优的方法采用Intro-VAE^[14]作为基础网络结构,该架构虽在复杂场景的被动NLOS成像中被证实有效,但其三模块两阶段设计需要分两步进行训练,训练过程繁琐且耗时较长。其次,当前NLOS成像数据集大多为合成数据集,且模型假设较理想化,导致这些数据集无法用于实际光照条件^[19]。Chen等^[20]进行了非视域成像中介面的

BRDF散射特性研究,但未进行更加深入的数据集采集和非视域重建。当前仅有的实拍数据集为基于普通墙面和白板的被动非视域大规模数据集^[14],并不适用于非视域场景下常见的墙面瓷砖场景。

U型网络(U-NET)是一种经典的编码器-解码器架构,最初用于医学图像分割^[21],通过编码器提取图像的高级特征,解码器恢复图像的空间细节,并采用跳跃连接方法融合不同尺度的特征信息。由于在图像重建中的出色表现,U-NET也被应用于NLOS成像领域。Christopher等^[22]提出一种基于U-NET架构的卷积神经网络解决方案,通过光谱估计推导NLOS成像相关噪声模型,生成合成训练数据,再设计平移不变损失函数训练网络解决噪声相位恢复问题,最终实现1 m距离下的主动NLOS成像。Tingyi Yu等^[18]在U-NET架构上提出了一种被动NLOS成像的UR-NET网络结构,通过改进瓶颈层提升细节恢复能力。然而,目前的U-NET架构只能实现简单目标的重建,无法对复杂场景目标实现重建。

本文提出一种基于增强反射的改进U-NET架构(Enhanced Reflection UNET, ER-UNET)的被动NLOS成像方法,通过增强反射的特征提取机制等设计,利用普通相机即可完成复杂目标的NLOS重建,期望提升被动NLOS成像性能和实用性。

2 基本原理

与主动NLOS成像方法相比,被动NLOS成像仅采集反射光的光强信息,缺少主动方法所具备的时间、相位等信息,故目前的研究大多聚焦于二维重建或目标定位。如图1所示,被动

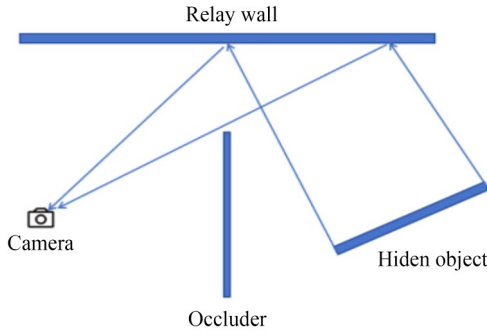


图1 被动NLOS成像系统原理

Fig. 1 Principle of passive NLOS imaging system

NLOS 成像技术是被动非视域场景下的二维图像重建,其目标是通过处理漫反射墙面上的投影图像还原目标场景图像。假设目标图像上的每个像素均为独立点光源,那么墙面上与之对应的实测投影可表示为:

$$I(p_y) = \iint_{p_f \in F} A(p_f, p_y) I(p_f) dp_f + n_{b+d}(p_y), \quad (1)$$

其中: $I(p_y)$ 为探测投影区域像素 p_y 处的光强, $I(p_f)$ 为隐藏源显示区域像素 p_f 处的光强, $A(p_f, p_y)$ 为点光源 P_f 到中继墙面区域 p_y 的光传输系数, F 代表整个屏幕的所有像素空间, $n_{b+d}(p_y)$ 为像素 p_y 处的噪声,由背景光 b 与探测器自身 d 共同产生。

该模型的离散化形式为:

$$y = Af + n_{b+d}, \quad (2)$$

其中: $f \in \mathbf{R}^{H_f \times W_f}$ 为向量化场景光强, $H_f \times W_f$ 为显示器像素数, $A \in \mathbf{R}^{H_y \times W_y \times H_f \times W_f}$ 为光传输矩阵, $y \in \mathbf{R}^{H_y \times W_y}$ 为向量化观测数据, $H_y \times W_y$ 为实测投影图像像素数, $n_{b+d} \in \mathbf{R}^{H_y \times W_y}$ 为向量化的噪声。

光传输矩阵 A 由墙面 w 的双向反射分布函数BRDF、隐藏物体与墙面的距离 r 及相机的位置 c 共同决定,记为 $A(w, r, c)$,则式(2)可改写为:

$$y = A(w, r, c) f + n_{b+d}. \quad (3)$$

即被动NLOS成像可看作去噪后求出 A^{-1} 的过程。

被动NLOS成像主要采用基于光场分析的方法和基于深度学习的方法。前者通过分析间接光场的统计特性来重建场景,计算复杂度高且重建精度有限。后者通过学习间接光场与隐藏场景之间的映射关系,实现高效准确的重建,但基于深度学习的方法存在泛化能力弱的问题,其重建性能很大程度上取决于训练集与测试集之间的相似度。因此,基于ER-UNET的被动NLOS成像模型既可在特定数据集上能实现高精度的重建,又能在大数据集上具备优异的泛化能力。

3 网络结构

如图2所示,ER-UNET是基于编码器-解码器结构的U-NET架构。其中,编码器部分通过

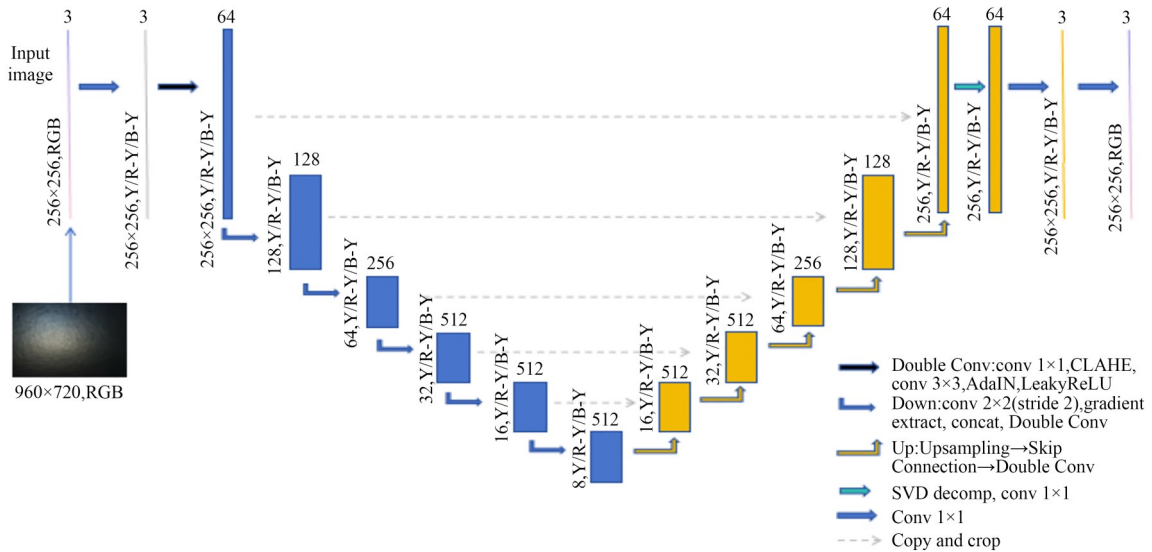


图2 ER-UNET模型架构

Fig. 2 ER-UNET model architecture

下采样操作提取图像的高级特征,解码器部分通过上采样操作恢复图像的空间细节,在编码器和解码器之间,通过跳跃连接融合不同尺度的特征信息。与传统U-NET不同,ER-UNET在编码器部分引入了反射特征提取器、可微分CLAHE和通道注意力机制,实现了架构创新。在解码器部分引入方向梯度特征提取和自适应实例归一化,有效利用了非朗伯表面的反射特性,实现了机制创新。ER-UNET流程如下:

(1)通过反射特征提取将RGB转换为Y,R-Y,B-Y通道,分离亮度与色度信息,聚焦非朗伯表面的亮度反射特征,剔除无关色彩干扰;

(2)利用通道注意力机制强化Y通道权重,让模型学习中介面反射的共性光传输规律,而非单一图像特征;

(3)方向梯度提取捕捉非朗伯反射的方向性,使模型适配不同目标的反射投影;

(4)通过编码-解码路径提取多尺度特征;

(5)对Y通道进行SVD低秩分解,仅保留前8个主奇异值以压缩有效模式数量,提升重建效率;

(6)基于光谱映射输出层,利用三独立卷积核实现梯度感知的RGB重建。

3.1 反射特征提取

反射特征提取器将RGB输入转换为Y,R-Y,B-Y三通道特征,模拟人眼视觉系统对亮度和色差的分离处理,即:

$$Y = 0.299R + 0.587G + 0.114B, \quad (4)$$

其中:Y通道代表亮度信息,R,G和B通道代表色彩信息。这种分离符合人类视觉感知特性,在非朗伯表面重建中能更好地保留亮度信息。

3.2 图像增强

限制对比度自适应直方图均衡化(Contrast Limited Adaptive Histogram Equalization, CLAHE)是一种常用的图像增强技术,能够增强局部对比度。传统的CLAHE算法是不可微分的,无法直接应用于深度学习框架^[23]。在深度学习框架中,可微分CLAHE通过基于滑动窗口的连续统计量计算替代传统CLAHE的离散直方图操作,从而实现梯度可传播。具体而言,计算以像素 (i,j) 为中心的局部邻域内均值和标准差,通过光滑的连续函数将局部统计量映射为增强因子,替代传统CLAHE中的直方图裁剪和重分配步

骤,使局部对比度调整过程完全可微分。该方法使网络能够自适应地放大低对比度区域的间接反射信号,提升网络对微弱反射特征的感知能力,同时将图像增强操作纳入可微计算图,保持梯度可端到端回传,实现增强参数与网络权重的联合优化。

可微分CLAHE算法通过局部均值和标准差计算增强因子,在保持梯度可传播的条件下,增强局部对比度的通道强度 I_c (如Y,R-Y,B-Y):

$$I_c(i,j) = \alpha(\mu_\Omega, \sigma_\Omega) \cdot (I(i,j) - \mu_\Omega) + \mu_\Omega, \quad (5)$$

其中: Ω 为以像素 (i,j) 为中心的局部邻域, μ_Ω 为局部邻域 Ω 内的像素均值, σ_Ω 为局部邻域 Ω 内的像素标准差, $\alpha(\mu_\Omega, \sigma_\Omega)$ 为连续增强因子,是关于 μ_Ω 和 σ_Ω 的光滑函数。

本文可微分CLAHE模块仅对Y通道执行可微分CLAHE增强,色彩差异通道不增强,增强因子为 $[0.5, 3.0]$,根据局部对比度动态调整:低对比度区域强增强(因子 ≈ 3.0),高对比度区域弱增强(因子 ≈ 0.5)。相较于传统CLAHE,该算法具有完全可微分、计算高效、无缝集成等优势。

3.3 自适应实例归一化

自适应实例归一化(Adaptive Instance Normalization, AdaIN)^[24]能够根据局部反射强度动态调整归一化参数,使用局部方差计算自适应权重。标准AdaIN通过将内容特征 x 的统计量(均值 $\mu(x)$ 和标准差 $\sigma(x)$)替换为风格特征 s 的统计量(均值 $\mu(s)$ 和标准差 $\sigma(s)$),实现风格迁移。输入为特征图 x ,输出为经过自适应实例归一化后的特征图 x' 。归一化计算公式为:

$$AdaIN(x) = \sigma(s) \cdot \frac{x - \mu(x)}{\sigma(x)} + \mu(s). \quad (6)$$

模型中未使用外部风格特征作为 s ,而是使用输入特征 x 自身的局部反射强度统计量动态调整归一化参数,即通过 $K \times K$ 滑动窗口计算特征图 x 在空间位置 (h, w) 处的局部均值与方差,可避免普通实例归一化对微表面反射突变的过度平滑,并保持训练稳定性。具体而言,归一化标准为:在高方差区域(局部方差均值较大),自适应因子 α 趋近于0,归一化强度减弱,保留微表面反射突变细节;在低方差区域(局部方差均值较小),自适应因子 α 趋近于1,归一化强度增强,提高训练稳定性。典型处理效果示例如图3所示,

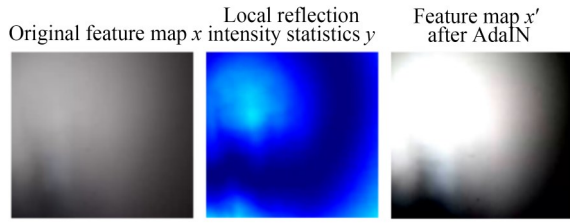


图3 归一化处理效果示意图

Fig. 3 Schematic diagram of effect of normalization processing

AdaIN 对高光细节实现了清晰保留,能够识别并保留微表面反射细节。

AdaIN 主要用于模型的 DoubleConv 块,作为优化双卷积序列的一部分,其输入为输入特征图,输出为归一化后的特征图,能适应局部光照变化,有助于利用非朗伯表面的反射特性。

3.4 通道注意力机制

通道注意力机制根据通道的重要性对特征进行加权,突出重要的特征信息。在非朗伯表面重建中,亮度信息通常比色度信息重要。因此,设计了如图4所示的通道注意力机制流程图,强化 Y 通道的权重 (≥ 0.6),减弱色彩差异通道的权重 (≤ 0.2),同时利用自适应平均和最大两种池化操作获取通道维度的统计信息,分别捕捉全局平均响应和最大响应特征,形成并行的特征提取分支;两个分支输出通过共享的 MLP 结构进行处理,结构包含两个 1×1 卷积层,实现通道压缩和非线性转换,中间使用 RE-LU 激活函数增强模型的表达能力;经 MLP 处理的两个分支结果相加并通过 Sigmoid 激活函数生成通道注意力权重,实现对不同通道特征的动态加权。

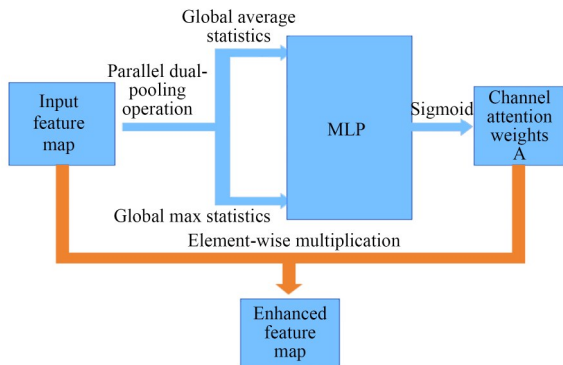


图4 通道注意力机制流程图

Fig. 4 Flow chart of channel attention mechanism

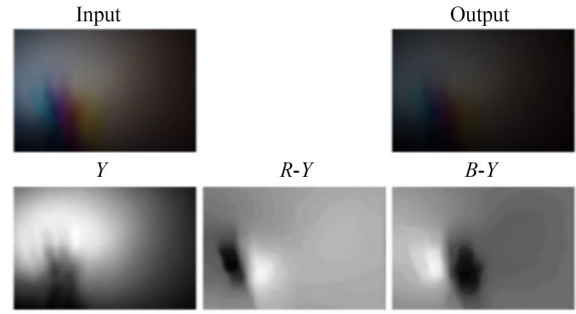


图5 通道注意力机制处理示意图

Fig. 5 Schematic diagram of channel attention mechanism processing

图5为一幅彩色图像的通道注意力机制处理示意图,模型聚焦非朗伯反射核心亮度信息,输出图像反射纹理更清晰且抑制了色彩信息的影响,明确优势为解决传统注意力无偏向问题,精准提取NLOS所需的有效反射信号,通过增强 Y 通道的权重,使模型更关注亮度信息,提高反射特征的捕捉能力。

3.5 方向特征梯度提取

非朗伯表面的反射特性具有明显的方向性,方向梯度能有效捕捉这些特征。基于 Sobel 算子提取 0° , 45° , 90° 方向的梯度信息,并加入下采样模块,计算梯度幅值和方向熵,作为额外的特征输入到解码器中。

模块从输入特征中提取 Y 通道完成梯度计算,定义 0° (水平)、 45° (对角线)、 90° (垂直) 3 个方向的固定权重 Sobel 算子:

$$\text{Sobel}_{0^\circ} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \text{Sobel}_{45^\circ} = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix},$$

$$\text{Sobel}_{90^\circ} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \quad (7)$$

对输入特征图 $X \in \mathbb{R}^{B \times C \times H \times W}$, 提取 Y 通道特征 Y:

$$Y = \begin{cases} X[:, 0:1, :, :], & C > 1 \\ X, & C = 1 \end{cases}, \quad (8)$$

其中: $Y \in \mathbb{R}^{B \times 1 \times H \times W}$; B, C, H, W 分别为批次大小、通道数、特征图高与宽。对 Y 通道执行二维卷积得到各方向梯度响应 $grad_\theta$, 基于各方向梯度计算梯度幅值,表征梯度的整体强度特征:

$$\text{mag} = \sqrt{\text{grad}_{0^\circ}^2 + \text{grad}_{45^\circ}^2 + \text{grad}_{90^\circ}^2}. \quad (9)$$

进一步计算方向熵以量化梯度方向的分布

均匀性,将梯度幅值与方向熵在通道维度拼接,得到最终方向梯度特征,随后将原始下采样特征与方向梯度特征拼接,增加特征维度。

3.6 损失函数设计

为了平衡像素级准确性、特征相似性和平滑性,采用多损失函数组合:

$$L = \lambda_1 L_{L1} + \lambda_2 L_p + \lambda_3 L_{L2}, \quad (10)$$

其中: $\lambda_1, \lambda_2, \lambda_3$ 为权重, L_{L1} 表示 L_1 损失,促进像素级准确性, L_p 表示感知损失,确保重建图像在特征空间中的相似性, L_{L2} 表示 L_2 损失,促进平滑性。各损失的计算公式如下:

$$L_1 = \frac{1}{N} \sum_p |I(p) - \hat{I}(p)|, \quad (11)$$

$$L_2 = \frac{1}{N} \sum_p (I(p) - \hat{I}(p))^2, \quad (12)$$

$$L_p = \frac{1}{M} \sum_q |\phi(I)(q) - \phi(\hat{I})(q)|, \quad (13)$$

其中: I 为真实参考图像, \hat{I} 为网络重建输出图像, p 为像素索引, N 为总像素数, q 为特征图元素索引, M 为特征总元素数, ϕ 为预训练卷积神经网络。

L_1 损失用于衡量生成图像与真实图像在像素级别上的绝对差异,能够有效惩罚大偏差的像素值。感知损失旨在捕捉生成图像与真实图像在高层语义特征空间中的差异,而非仅关注像素级别的匹配。 L_2 损失用于衡量生成图像与真实图像在像素级别上的平方差异。

4 实验

4.1 实验设置

4.1.1 数据集

为检验ER-UNET网络在常见的瓷砖中介面场景下NLOS成像的图像重建性能,搭建了一个被动NLOS成像实验系统,如图6所示。以普通液晶显示器作为不同目标图像的投射端,再采用手机摄像头或SPAD相机采集通过不同瓷砖中介面散射投影的模糊图像,最后通过对采集图像的处理,重构目标图像。

以现有被动NLOS数据集NLOS-PASSIVE^[14]作为基础投影图像,构建以瓷砖中介面的真实被动NLOS成像数据集,命名为NLOS-Passive-TILES-stl10。数据集样本遵循“复杂分



图6 被动NLOS成像实验场景

Fig. 6 Experiment scene of passive NLOS imaging

布数据可提升模型泛化能力”的原则^[13],通过镜面、白色和黑色瓷砖中介面(如图7)采集了8组有效数据,累计投影图像超80万张(示例如图8所示),每组数据包含超过10万张图片,按照8:1:1的比例划分为训练集、验证集和测试集。其中,所用白色瓷砖的微面元均方根斜率约为0.1,折射率约为1.7,漫反射系数约为0.78,镜面反射系数约为0.2;黑色瓷砖的微面元均方根斜率约为0.1,折射率约为2.2,漫反射系数约为0.67,镜面反射系数约为0.3。



图7 镜子、白和黑瓷砖中介面

Fig. 7 Interface surface of mirror, white tile and black tile

由于数据量过大,在实际测试中随机选取500张进行测试。此数据集全面覆盖STL-10图像在不同材质瓷砖表面的投影信息,且在同组数据采集全过程中,严格保持成像设备的白平衡、焦距等关键参数固定不变,有效消除设备参数波动对数据一致性的干扰,保障数据集的可靠性与有效性。数据集特意纳入不同材质的瓷砖中介面成像样本,可支持算法在差异化光学环境下的性能对比分析。

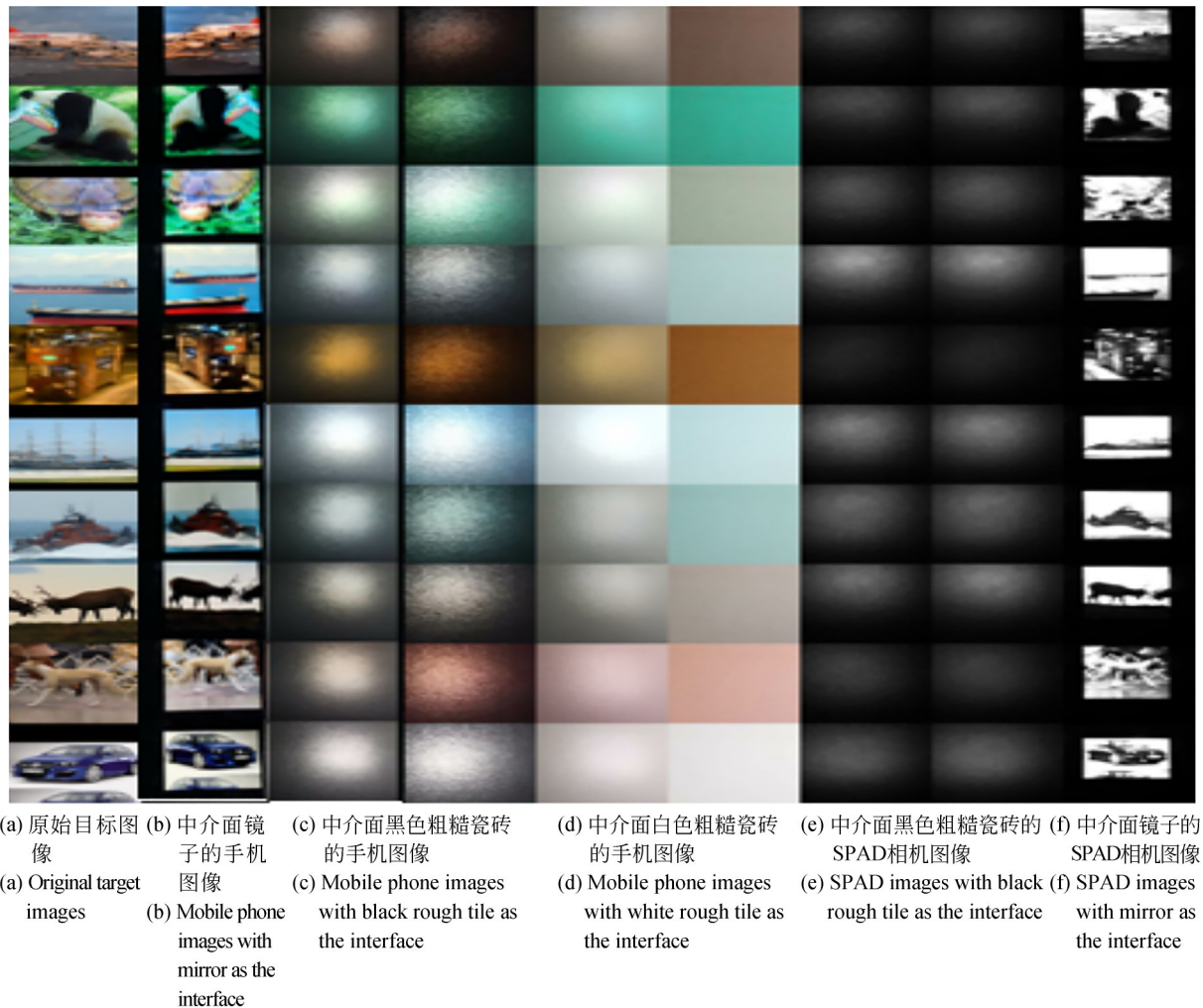


图 8 数据集采集图例

Fig. 8 Dataset collection legend

4.1.2 基线方法与评估指标

为了验证 ER-UNET 的性能,选择 NLOS-OT^[14]和 C-GAN 方法进行对比。其中,NLOS-OT 性能与重建效果优良的开源被动 NLOS 深度学习方法,通过流形嵌入与最优传输结合的两步训练实现;虽暂无开源的 C-GAN 架构用于被动 NLOS,但生成对抗网络在图像去噪方面的表现较好,故将其作为基线进行对比。

以测试集的平均峰值信噪比 PSNR 和结构相似度 SSIM 作为评价指标,评估不同方法的重构质量。评价指标的基础图像为投影图像。

通过实验分析,对于粗糙中介面的复杂场景的非视域重建,我们认为当 $PSNR \geq 12$ dB, $SSIM \geq 0.4$ 可认为重建有效,可较为清晰地分辨目标轮廓与关键纹理,满足目标识别应用需求。

4.1.3 训练细节

ER-UNET 与两种基线方法均基于 Pytorch 实现,训练在单块 GeForce RTX 3090 显卡上进行。采集的投影图像尺寸,利用手机摄像头采集的彩色数据集为 960×720 ,利用 SPAD 相机采集的单色数据集为 512×512 。所有输入与输出图像尺寸均调整为 256×256 。

4.2 复杂隐藏场景的重构

以往被动 NLOS 成像算法大多以数字或几何结构物件的简单场景进行重构,且重构图像较为模糊。近年来,NLOS-OT 方法^[13]可重构较为复杂的隐藏场景,但需要两步训练且细节重构效果较差。本文 ER-UNET 通过引入反射特征提取,以期降低重构和训练的难度,能有效应用于复杂的隐藏场景,提升重构质量。

表 1 不同方法在测试集上的平均定量结果对比

Tab. 1 Comparison of quantitative results from different methods on test set

Method	PSNR/dB	SSIM
NLOS-OT	14.08	0.50
C-GAN	14.48	0.48
ER-UNET	15.14	0.50

表 1 给出了 3 种方法在 NLOS-PASSIVE 测试集上的平均 PSNR 和 SSIM 值。可以看出, ER-UNET 在 PSNR 和 SSIM 指标上均优于其他对比方法,具有更好的重建质量。

图 9 和表 2 分别展示了 3 种方法在 4 个测试样本上的重建图像以及定量评价结果。可以看出, NLOS-OT 和 C-GAN 的重建结果有所改善,但仍然存在细节丢失; ER-UNET 重建图像细节

有所提高,对比度适中,平均 PSNR 更高, SSIM 与 NLOS-OT 持平,视觉效果更好,获得了更多的目标纹理和结构信息。



图 9 不同方法的图像重建结果

Fig. 9 Image reconstruction results for different methods

表 2 单张重建图片的定量对比

Tab. 2 Quantitative comparison of single reconstructed image

Method	C-GAN(PSNR/SSIM)	NLOS-OT(PSNR/SSIM)	ER-UNET(PSNR/SSIM)
INPUT1	11.10/0.35	15.56/0.31	15.17/0.37
INPUT2	12.87/0.37	12.03/0.42	12.59/0.34
INPUT3	14.71/0.51	18.06/0.48	17.59/0.53
INPUT4	15.57/0.25	15.24/0.15	15.21/0.25

4.3 不同中介面下的 NLOS 成像重构

为评估 ER-UNET 对不同中介面的稳定性,选取数据集 NLOS-Passive-TILES-stl10 中白色和黑色瓷砖中介面作为对比数据集,实验结果及其定量分析分别如图 10 和表 3 所示。可以看出,不同中介面下的重构效果存在差异,中介面特性对重建结果影响较大。在所有测试条件下, ER-UNET 均能实现较好的重建效果,证明 ER-UNET 对不同中介面具有良好的稳定性。

表 3 不同中介面重建指标

Tab. 3 Reconstruction index of different interface

Relay wall-method	PSNR	SSIM
White rough tile-ER-UNET	15.94	0.51
White rough tile-NLOS-OT	15.20	0.45
Black rough tile-ER-UNET	17.03	0.53

4.4 泛化能力评估

由于隐含着测试数据的分布与训练数据相似的假设,泛化能力成为评价学习方法的关键指标。在 NLOS 深度学习任务中,泛化能力通常指固定中介面的跨图像类别泛化,测试为训练集未见的图像类别,证明模型学习逆光传输通用知识。在被动 NLOS 任务中, ER-UNET 主要可学习三类知识:逆光传输过程、噪声分布和数据先验。若学习过多数据先验,会导致泛化能力下降;当训练数据集的分布足够广泛时,学习数据先验的难度会增加,迫使网络学习更多逆光传输相关知识,从而具备更强的泛化能力。

本文的泛化能力评估是在固定中介面条件下进行的跨图像类别泛化测试,旨在验证算法在特定光学环境下的有效性和稳定性。关于跨中介面泛化能力,由于不同中介面的 BRDF 参数差异会从根本上改变光传输矩阵的结构,当前算法

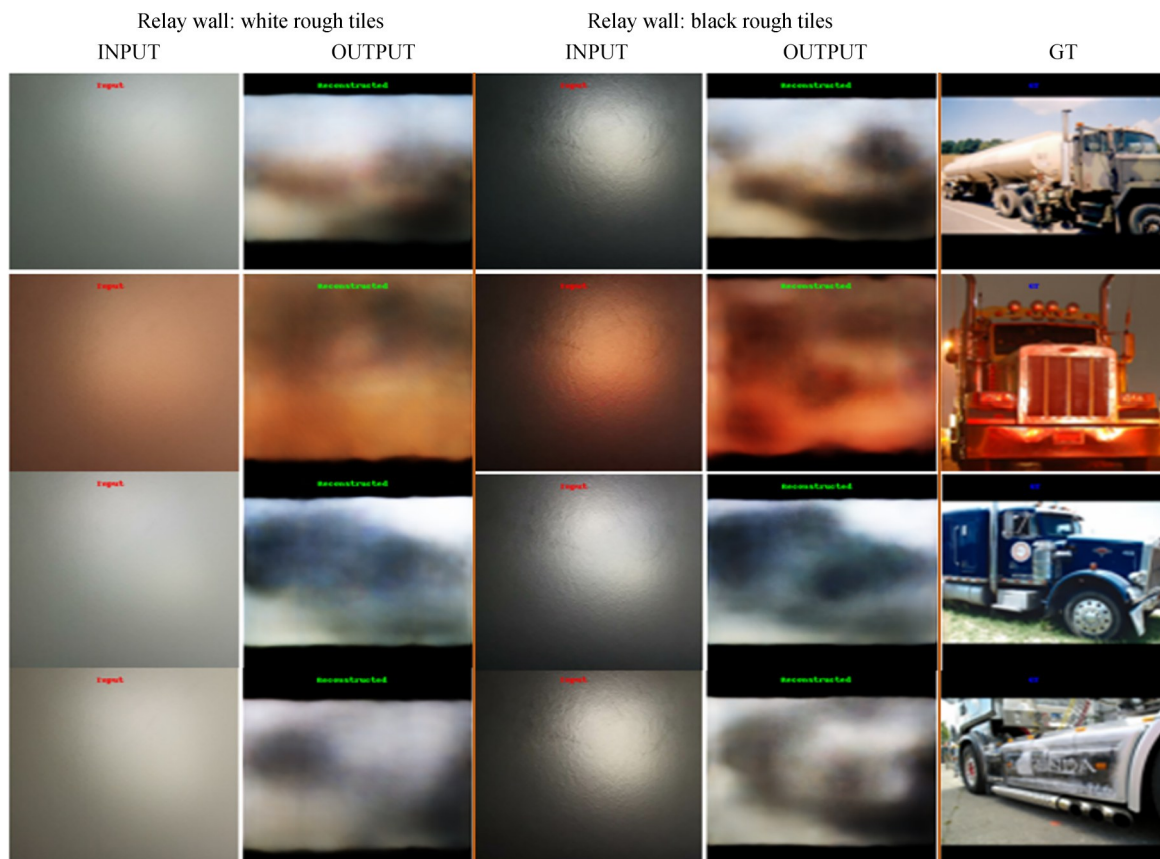


图 10 不同中介面下 ER-UNET 的重建结果

Fig. 10 Reconstruction results of ER-UNET under different interface planes

暂无法实现跨中介面泛化,未来需要通过扩充包含多种中介面及其对应 BRDF 参数的大规模数据集来进一步研究。

为测试网络泛化性,通过在 stl10 上的训练,在未见的人脸以及图标数据类别上测试评估模型泛化能力。实验结果如图 11 所示,ER-UNET

在测试数据上仍能取得良好的重构效果。表明当训练数据集选择适当时,ER-UNET 能够学习到类似逆光传输矩阵的映射关系,具备强化泛化能力。

5 结 论

为了提升被动 NLOS 成像在目标场景图像重建的性能,本文提出了一种基于增强反射的被动 NLOS 成像方法 ER-UNET。针对瓷砖非朗伯表面的反射特性,ER-UNET 通过集成反射特征提取器、可微分 CLAHE、自适应实例归一化与通道注意力机制和方向梯度特征提取等核心模块,区别于传统 U-NET 与现有 NLOS 网络,针对瓷砖非朗伯表面反射特性,设计了反射感知的编码-解码路径,突破传统 U-NET 无法利用中介面反射特性的局限,实现了对瓷砖非朗伯反射特性的有效利用。在 NLOS 重建中,引入 SVD 低秩分解+梯度感知 RGB 重建,提升了重建效率。实验表明,ER-

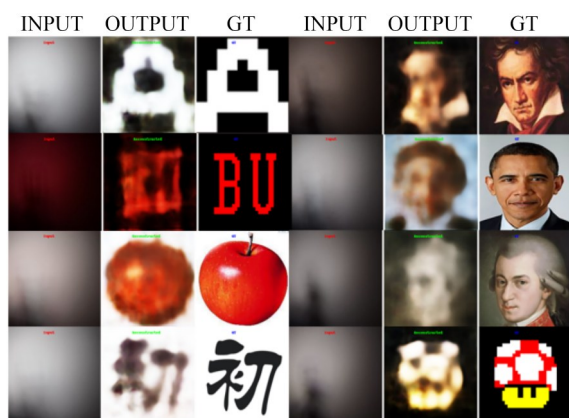


图 11 ER-UNET 的泛化性检验

Fig. 11 Generalization test of ER-UNET

UNET的平均PSNR和SSIM指标达到了15.14 dB和0.50,优于NLOS-OT,C-GAN等对比方法,能够更准确地重建非视域场景。

ER-UNET为被动NLOS成像技术提供了新的发展思路,有望推动在安全监控、自动驾驶、医疗诊断等领域的实际应用,进一步将优化模型结构,提高其重构性能和适用性,推进NLOS成像技术的发展和應用。此外,当前研究仅针对特定中介面条件下的非视域成像进行了算法验证,

跨中介面泛化能力有待未来通过扩充包含多种中介面及其对应BRDF参数的数据集来进一步优化,从而使网络能够学习到不依赖特定中介面特性的通用光传输映射。

作者贡献声明:

于相之:算法研究,实验设计,论文撰写;
金伟其:方法指导,论文修改与审核;
裘溯、李力:实验资源提供与指导。

参考文献:

- [1] VELTEN A, WILLWACHER T, GUPTA O, *et al.* Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging[J]. *Nature Communications*, 2012, 3: 745.
- [2] YANG Y, DU K, WU X D. Iterative optimization dehazing algorithm based on improved linear transformation [J]. *Laser & Optoelectronics Progress*, 2021, 58(6): 0610011.
- [3] KATZ O, HEIDMANN P, FINK M, *et al.* Non-invasive single-shot imaging through scattering layers and around corners *via* speckle correlations[J]. *Nature Photonics*, 2014, 8(10): 784-790.
- [4] LUESIA P, CRESPO M, JARABO A, *et al.* Non-line-of-sight imaging in the presence of scattering media using phasor fields [J]. *Optics Letters*, 2022, 47(15): 3796.
- [5] KIRMANI A, HUTCHISON T, DAVIS J, *et al.* Looking around the corner using ultrafast transient imaging[J]. *International Journal of Computer Vision*, 2011, 95(1): 13-28.
- [6] O'TOOLE M, LINDELL D B, WETZSTEIN G. Confocal non-line-of-sight imaging based on the light-cone transform [J]. *Nature*, 2018, 555(7696): 338-341.
- [7] LINDELL D B, WETZSTEIN G, O' TOOLE M. Wave-based non-line-of-sight imaging using fast *f-k* migration[J]. *ACM Transactions on Graphics*, 2019, 38(4): 1-13.
- [8] 韩圆,李有治,纪珍辰,等.阵列飞行时间实时非视域成像技术[J].*光学精密工程*,2025,33(3):377-388.
HAN Y, LI Y ZH, JI ZH CH, *et al.* Technologies of real-time non-line-of-sight imaging based on time-of-flight array [J]. *Optics and Precision Engineering*, 2025, 33(3): 377-388. (in Chinese)
- [9] BATARSEH M, SUKHOV S, SHEN Z, *et al.* Passive sensing around the corner using spatial coherence [J]. *Nature Communications*, 2018, 9: 3629.
- [10] LI W W, ZHOU Y J, *et al.* Turning rough surfaces into non-line-of-sight cameras [J]. *Optica*, 2025, 12(5): 626.
- [11] SAUNDERS C, MURRAY-BRUCE J, GOYAL V K. Computational periscopy with an ordinary digital camera [J]. *Nature*, 2019, 565(7740): 472-475.
- [12] YEDIDIA A B, BARADAD M, THRAMPOULIDIS C, *et al.* Using unknown occluders to recover hidden scenes[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 12223-12231.
- [13] TANAKA K, MUKAIGAWA Y, KADAMBI A. Polarized non-line-of-sight imaging [C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 13-19, 2020, Seattle, WA, USA. IEEE, 2020: 2133-2142.
- [14] GENG R X, HU Y, LU Z, *et al.* Passive non-line-of-sight imaging using optimal transport [J]. *IEEE Transactions on Image Processing*, 2022, 31: 110-124.
- [15] 刘玉红,杨恒.联合图像层级特征的压缩感知迭代重构[J].*光学精密工程*,2024,32(14):2311-2324.
LIU Y H, YANG H. Iterative reconstruction of compressive sensing combining image hierarchical feature [J]. *Optics and Precision Engineering*, 2024, 32(14): 2311-2324. (in Chinese)
- [16] 石敦攀,徐伟,朴永杰,等.基于Mamba-2的视

- 频快照压缩成像重构方法[J]. 液晶与显示, 2025, 40(6): 881-894.
- SHI D P, XU W, PIAO Y J, *et al.* Reconstruction method of video snapshot compressive imaging based on Mamba-2[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(6): 881-894. (in Chinese)
- [17] SHARMA P, AITTALA M, SCHECHNER Y Y, *et al.* What you can learn by staring at a blank wall[C]. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 10-17, 2021, Montreal, QC, Canada. IEEE, 2022: 2310-2319.
- [18] 于亭义, 乔木, 刘红林, 等. 基于深度学习的非视域成像[J]. 光学学报, 2019, 39(7): 0711002.
- YU T Y, QIAO M, LIU H L, *et al.* Non-line-of-sight imaging through deep learning[J]. *Acta Optica Sinica*, 2019, 39(7): 0711002. (in Chinese)
- [19] PHONG B T. Illumination for computer generated pictures[J]. *Communications of the ACM*, 1975, 18(6): 311-317.
- [20] 陈海鹏, 金伟其, 李力, 等. 时间门控 SPAD 阵列非视域成像的中介面 BRDF 散射特性研究[J]. 红外技术, 2024, 46(11): 1225-1234.
- CHEN H P, JIN W Q, LI L, *et al.* Study on BRDF scattering characteristics of relay wall in non-line-of-sight imaging based on time-gated SPAD array[J]. *Infrared Technology*, 2024, 46(11): 1225-1234. (in Chinese)
- [21] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015*. Cham: Springer, 2015: 234-241.
- [22] METZLER C A, HEIDE F, RANGARAJAN P, *et al.* Deep-inverse correlography: towards real-time high-resolution non-line-of-sight imaging[J]. *Optica*, 2020, 7(1): 63.
- [23] PIZER S M, AMBURN E P, AUSTIN J D, *et al.* Adaptive histogram equalization and its variations[J]. *Computer Vision, Graphics, and Image Processing*, 1987, 39(3): 355-368.
- [24] HUANG X, BELONGIE S. Arbitrary style transfer in real-time with adaptive instance normalization [EB/OL]. 2017: *arXiv*: 1703.06868. <https://arxiv.org/abs/1703.06868>.

作者简介:



于相之(2000—),男,四川成都人,硕士研究生,2023年于北京理工大学获得学士学位,主要从事计算机视觉和微光探测等研究。E-mail: slintonyu@163.com

通讯作者:



金伟其(1961—),男,上海人,博士,教授,博士生导师,1982年于北京工业学院获得学士学位,1990年于北京理工大学获得博士学位,主要从事夜视与红外技术、光电图像处理及光电检测与仪器等方面的研究。E-mail: jin-wq@bit.edu.cn