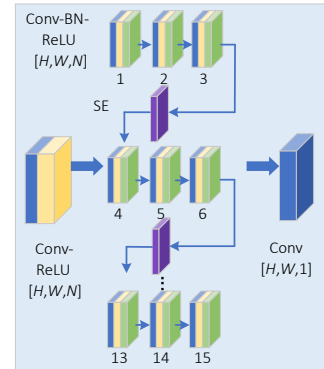


面向复合噪声的结构光三维重建多模态去噪方法

张立平¹, 胡晓峰^{1*}, 潘飞文², 郭 斌^{1,2}, 罗 戡¹

¹中国计量大学计量测试与仪器学院, 浙江 杭州 310018;

²杭州沃镭智能科技股份有限公司, 浙江 杭州 310018



摘要: 针对结构光三维成像中复合噪声干扰下表面细节难以准确还原的问题, 提出一种融合多模态输入与结构保持机制的深度去噪方法。采用双通道残差卷积神经网络, 以含噪条纹图与串联滤波图为联合输入, 引入多尺度特征提取与通道注意力模块, 结合均方误差、结构相似性指数和拉普拉斯边缘项的混合损失函数, 实现噪声抑制与边缘保持协同优化。搭建双目结构光实验平台, 对石膏像等三类样本结合多频相移与外差法进行训练验证。结果表明, 去噪后图像域的峰值信噪比提升约 7 dB, 结构相似性指数提高至 0.96, 均方根误差降低约 61.6%; 在三维重建结果层面, 点到面均方根误差降低约 47.1%, 点云密度提升 43.6%。该方法能够在抑制多源噪声的同时有效保持结构与边缘特征, 显著提升三维重建质量。

关键词: 结构光测量; 条纹去噪; 多模态输入; 深度学习; 三维点云重建

The English title and abstract appear at the end of the paper.

DOI: 10.12086/oe.2026.250272 | CSTR: 32245.14.oe.2026.250272 | 中图分类号: TH741 | 文献标识码: A

张立平, 胡晓峰, 潘飞文, 等. 面向复合噪声的结构光三维重建多模态去噪方法 [J]. 光电工程, 2026, 53(4): 250272

Zhang L P, Hu X F, Pan F W, et al. Multi-modal denoising method for structured light 3D reconstruction under composite noise[J]. Opto-Electron Eng, 2026, 53(4): 250272

1 引言

随着工业自动化与高精度检测需求的不断增长, 结构光三维重建技术因其高精度、高效率和非接触测量特性, 在电子元器件检测、机械零件测量等领域得到广泛应用^[1]。该技术基于条纹图像的相位信息, 通过相位解包裹与几何映射实现物体表面形貌的三维重建^[2]。然而, 在实际工业环境中, 条纹图像易受多源噪声干扰, 导致相位信息失真、解包裹失败及重建精度下降, 限制了其应用的稳定性与鲁棒性^[3]。

结构光三维重建步骤通常包括: 投影编码条纹、相

机同步采集变形条纹序列、提取并解包裹相位^[4], 结合相机标定与“逆相机”模型^[5-6]实现相位值到三维坐标的精确映射, 最终生成三维点云。实际测量中, 噪声来源主要包括环境噪声、设备噪声及物体表面特性噪声, 这些干扰会在相位处理过程中逐步累积, 引发相位跳变、不连续及局部失真^[7]。因此, 在相位解包裹之前进行有效的图像去噪, 是保证测量精度的关键。

现有去噪方法包括空域滤波(高斯、中值、均值)与频域滤波(傅里叶、小波、带通)^[8-10], 此类方法实现简单, 在低噪声条件下能够有效抑制随机噪声, 但在复杂或强噪声场景下往往导致条纹对比度显著下降, 细节纹理模

收稿日期: 2025-09-10; 修回日期: 2026-01-12; 录用日期: 2026-01-16; 网络出版日期: 2026-04-25

基金项目: 国家自然科学基金资助项目(U24A20134, 52475576); 浙江省科技计划项目“尖兵”研发攻关计划基金资助项目(2023C01061)

*通信作者: 胡晓峰, 12B0202094@cjl.edu.cn

版权所有©2026 中国科学院光电技术研究所

糊, 且对高频边缘信息存在不可逆的损失。此外, 这类方法通常依赖固定的滤波核参数, 难以自适应应对多源噪声的变化。近年来, 基于深度学习的图像复原方法得到广泛关注, 其中卷积神经网络 (convolutional neural network, CNN) 在去噪任务中表现出显著优势^[11]。例如, 可训练非线性反应扩散模型 (trainable nonlinear reaction diffusion, TNRD)^[12] 和基于残差学习的去噪卷积神经网络 (denoising convolutional neural network, DnCNN)^[13] 在自然图像去噪中取得了优异效果。随后, 研究者提出了更复杂的改进网络, 例如, 引入多尺度特征融合与残差密集连接的残差密集网络 (residual dense network, RDN)^[14], 以及结合深度先验的即插即用去噪先验框架 (plug-and-play, PnP)^[15], 在低照度和高噪声环境下表现出更强的细节保持与鲁棒性。然而, 这些方法多针对自然图像设计, 直接应用于结构光条纹图时仍面临纹理特征与相位信息易受破坏的问题。

针对结构光成像场景的特殊噪声特性, 近年来也出现了一些专门的去噪策略。例如, 基于条纹方向先验的自适应正弦/余弦滤波方法^[16] 可在均匀纹理与中低反射样本上有效抑制方向性噪声, 但在强反射区域易产生条纹形变与边缘模糊, 对跨场景的适应性有限; 基于卷积神经网络去噪正则化的条纹图修复模型^[17] 能够修复缺失条纹、提升连通性并减少断裂, 然而该方法在复杂背景与高反差区域容易出现过平滑, 细节对比度下降; 深度自编码器 (DAE) 去噪模型^[18] 采用迭代去噪与相位-高度标定框架, 可有效抑制随机噪声, 但作为自编码器类重构, 存在全局平滑倾向的风险, 细节与边缘保持能力受限; 基于深度学习的单帧条纹相位检索方法^[19] 利用端到端网络实现条纹图像的自适应去噪与相位提取, 然而在多源噪声叠加和严重非线性失真场景下, 网络的域外泛化能力不足, 仍存在精度下降的风险; 面向非视距定位的 U-Net 光斑检测方法^[20] 可在低信噪比与遮挡条件下稳定提取 LED 反射光斑, 并结合几何约束实现三维位置估计, 但其成像模式与条纹相位结构差异较大, 难以直接用于结构光条纹噪声抑制; 基于 swin-UNet 的两步相位解包裹方法^[21] 能够在散斑噪声下增强包裹相位的平滑性并提升相位展开精度, 但其训练数据以仿真为主, 对真实复合噪声的泛化能力有限; 基于 UMNNet 的数字全息相位解包裹方法^[22] 可通过多尺度卷积提高全息相位恢复精度, 并增强对椒盐噪声的抑制能力, 然而该方法面向透射式全息成像, 对反射式结构光条纹的噪声适应性

不足。由此可见, 现有方法在应对复杂工业环境下的条纹图像去噪任务时仍存在不足, 有必要进一步探索更具鲁棒性和泛化能力的深度学习方法。

针对上述问题, 本文提出一种融合多模态输入与结构保持机制的深度去噪模型。方法基于改进的双通道残差卷积神经网络 (DnCNN), 将原始含噪条纹图与两级串联滤波结果拼接输入, 以兼顾细节保留与噪声抑制。网络中嵌入通道注意力机制 (SE), 自适应强化与相位连续性相关的关键信号, 抑制冗余特征; 并设计融合亮度、结构与边缘约束的混合损失函数, 引导模型在抑噪同时保持条纹方向性与边界锐利度。与仅依赖编码优化或系统改进不同, 该方法作为前端图像预处理, 可与多频相移和外差解包裹等算法互补, 显式降低噪声对相位误差传播的累积效应, 从而提升三维重建的精度与稳定性。

2 结构光三维测量系统与噪声建模分析

2.1 双目结构光系统三维测量原理

本文采用双目结构光三维测量系统, 由结构光投影仪、左右工业相机、非远心 FA 工业镜头、图像采集卡、计算机及运动控制平台组成, 其整体结构如图 1 所示^[23]。其中, 同一投影仪向被测物体投射同一条纹序列, 左右相机同步采集并在统一标定坐标系下完成三维重建与融合, 属于典型的单投影仪-双相机的双目结构光系统。该系统通过联合标定建立统一几何关系, 左右相机重建结果可在同一参考坐标系下直接融合, 无需额外配准。

系统基于相位测量轮廓术 (phase measuring profilometry, PMP) 原理^[24], 通过投射多组预先设计的条纹图案, 并采用双目相机同步采集物体表面的变形条纹图像。为兼顾测量精度与抗噪声性能, 本文采用三频四步相移编码方式: 在三个空间频率 $\{f_s\}_{s=1}^3$ 下分别投射四幅相移正弦条纹, 移相步进为 $k\pi/2$, $k=0, 1, 2, 3$, 两台相机同步采集 12 幅变形条纹序列。在以下相位建模与三维重建过程中, (x, y) 表示相机图像平面上的像素坐标, 条纹强度、包裹相位及相位展开等计算均在对应像素位置 (x, y) 处进行。对于任意频率 f_s 下的相移序列, 首先利用四步相移法计算包裹相位 $\phi_w^{(s)}(x, y)$, 由于包裹相位限制在 $(-\pi, \pi]$, 需进一步解包裹。本文采用多频外差策略, 由低频条纹相位逐级引导中频与高频相位展开, 实现由低频到高频的逐级相位恢复, 得到像素级的绝对相位 $\phi_a(x, y)$ 。为将相位映射为投影仪像素坐标, 记投影端对

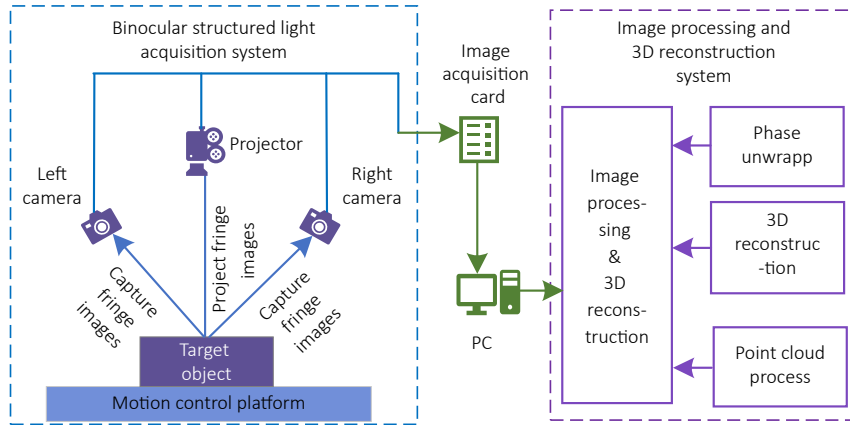


图 1 系统总体结构图

Fig. 1 Schematic diagram of the overall system architecture

应的等效空间周期 $\lambda_p = 1/f_s$ 。根据式 (1) 中的相位-像素映射

$$u_p(x, y) = \frac{\lambda_p}{2\pi} \phi_a(x, y) + u_{p0}, \quad v_p(x, y) = v_{p0}, \quad (1)$$

式中: (u_{p0}, v_{p0}) 为投影仪等效成像模型的主点坐标 (由标定得到)。以左相机坐标系为世界坐标系 ($R_1 = I, T_1 = \mathbf{0}$), 由相机内参 K_c 与投影仪等效内参 K_p , 以及投影仪相对于左相机的外参 ($R_{p \rightarrow 1}, T_{p \rightarrow 1}$) 组成投影矩阵为:

$$P_c = K_c [R_1 T_1], \quad P_p = K_p [R_{p \rightarrow 1} T_{p \rightarrow 1}]. \quad (2)$$

随后, 在前述相位提取与相位展开的基础上, 利用式 (1) 将像素级绝对相位换算为投影仪像素坐标 (u_p, v_p) , 并结合式 (2) 构建的相机-投影仪投影模型, 将左相机去畸变后的像素坐标 (u_c, v_c) 与 (u_p, v_p) 联立, 采用线性三角化方法求解物点在左相机坐标系下的三维坐标 $[X, Y, Z]^T$ 。上述各步骤对应的数学表达以公式组形式统一列于式 (3) 中。

$$\begin{cases} I_k(x, y) = I_0(x, y) + I_m(x, y) \cos \left[\phi_w(x, y) + \frac{k\pi}{2} \right] \\ k = 0, 1, 2, 3 \\ \phi_w(x, y) = \arctan \left[\frac{I_3(x, y) - I_1(x, y)}{I_0(x, y) - I_2(x, y)} \right] \\ \phi_a(x, y) = \text{unwrap} \left[\phi_w^{(s)}(x, y) \right], \quad s = 1, 2, 3 \\ \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} P_{c11} - u_c P_{c31} & P_{c12} - u_c P_{c32} & P_{c13} - u_c P_{c33} \\ P_{c21} - v_c P_{c31} & P_{c22} - v_c P_{c32} & P_{c23} - v_c P_{c33} \\ P_{p11} - u_p P_{p31} & P_{p12} - u_p P_{p32} & P_{p13} - u_p P_{p33} \end{bmatrix}^{-1} \cdot \begin{bmatrix} u_c P_{c34} - P_{c14} \\ v_c P_{c34} - P_{c24} \\ u_p P_{p34} - P_{p14} \end{bmatrix} \end{cases}, \quad (3)$$

式中: $I_k(x, y)$ 为第 k 幅相移条纹的灰度值; $I_0(x, y)$ 为背景光强; $I_m(x, y)$ 为调制度。本文在推导中选取左相

机作为参考坐标系以统一符号体系与结果表达, 因此式 (1)~(3) 以投影仪与左相机的几何约束为例给出相位-像素映射与三角化求解过程。需要说明的是, 在同一联合标定参数下, 右相机同步采集相同条纹序列并完成相位求解与展开, 其对应的投影模型与三角化求解可按相同思路建立, 得到右相机坐标系下的三维重建结果。随后, 利用双目标定获得的右相机相对于左相机的外参 ($R_{r \rightarrow 1}, T_{r \rightarrow 1}$), 将右相机重建结果变换至左相机参考坐标系下, 与左相机结果在同一坐标系中进行融合, 从而实现双目结构光系统的统一重建。图 2 展示了系统整体流程, 包括条纹投射、相位计算、解包裹以及三维重建步骤。

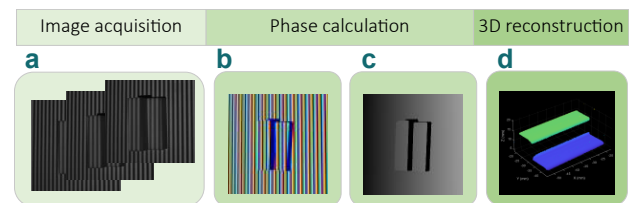


图 2 结构光三维测量系统工作流程示意图。(a) 条纹图像;

(b) 包裹相位图; (c) 解包裹相位图; (d) 三维表面点云图

Fig. 2 Schematic of workflow of the structured light 3D measurement system.

(a) Fringe images; (b) Wrapped phase map; (c) Unwrapped phase map; (d) 3D surface point cloud

2.2 成像误差与噪声来源分析及建模

在实际结构光三维测量中, 条纹图像在成像与传输过程中不可避免地受到系统性成像误差与多源随机噪声的影响, 导致条纹对比度降低并影响相位提取与解包裹的稳定性。为便于描述前端校正流程与后端噪声建模,

本文将其划分为两类: 1) 系统性成像误差 (来源稳定、可标定/校正), 包括相机与投影仪的几何畸变与外参误差、光学成像中的像面不均匀 (暗电流、照明不均匀)、背景/多次反射引入的低频结构化偏置等。此类项不具备零均值的随机统计特性, 本质上属于结构化的偏置/增益, 可通过系统标定、畸变校正等在重建前进行前端补偿。2) 随机性噪声 (时空不确定、需建模抑制), 包括光子/读出噪声 (近似高斯分布)、脉冲/极值噪声 (椒盐噪声)、散粒/斑点噪声, 以及环境光波动等^[25]。

在此基础上, 本文采用常见的观测模型, 把观测条纹图像表示为理想无噪条纹图像经系统映射后, 叠加缓变的乘性扰动与随机噪声。其中, 高斯噪声可建模为均值为 0、标准差为 σ_g 的白噪声^[26]; 椒盐噪声以概率 p 将像素值替换为最小 (0) 或最大 (255) 灰度^[27]; 散粒噪声属于与信号强度相关的乘性扰动^[28]。由此得到复合观测模型可表示为:

$$\begin{cases} N_{\text{Gauss}} \sim \mathcal{N}(0, \sigma_g^2) \\ N_{\text{sp}} = \begin{cases} 0, & p/2 \\ 255, & p/2 \\ I(x, y), & 1-p \end{cases} \\ I_{\text{noisy}}(x, y) = \mathcal{D}(I_{\text{clean}}(x, y))(1 + \varepsilon(x, y)) \\ \quad + N_{\text{Gauss}}(x, y; \sigma_g) + N_{\text{sp}}(x, y; p) \end{cases}, \quad (4)$$

式中: $I_{\text{noisy}}(x, y)$ 为观测条纹图像; $I_{\text{clean}}(x, y)$ 为理想无噪条纹图像; $\mathcal{D}(\cdot)$ 表示系统确定性的几何/光度映射; $\varepsilon(x, y)$ 为与信号相关的局部乘性扰动; N_{Gauss} 为高斯噪声项; N_{sp} 为椒盐噪声项; p 为椒盐噪声概率。系统性成像误差可通过标定与平场/背景校正在前端预补偿, 随机噪声则需依赖后续的深度去噪网络进行建模与抑制。

需要说明的是, 本文所建立的复合观测模型主要针对静态测量条件下的空间复合噪声。实验平台中投影仪与两台相机均通过刚性支架固定安装, 整体置于光学隔振平台, 被测 PCB 样本在采集周期内保持静止, 整个系统工作在外触发同步模式。两台相机曝光时间均设置为 $T_{\text{exp}} = 10 \text{ ms}$, 而实验室照明由 50 Hz 交流市电驱动, 其主导亮度闪烁频率约为 $f \approx 100 \text{ Hz}$ (闪烁周期 $T_f = 1/f \approx 10 \text{ ms}$), 二者处于同一量级。为描述环境光的闪烁, 可将环境光瞬时亮度写为

$$I_{\text{env}}(t) = I_0 + \Delta I \sin(2\pi f t), \quad (5)$$

式中: $I_{\text{env}}(t)$ 表示环境光亮度; I_0 为环境光的直流平均亮度; ΔI 为环境光闪烁幅度。则单帧曝光得到的环境光平均亮度为

$$\bar{I}_{\text{env}} = \frac{1}{T_{\text{exp}}} \int_{t_0}^{t_0+T_{\text{exp}}} I_{\text{env}}(t) dt = I_0 + \frac{\Delta I}{2\pi f T_{\text{exp}}} \{ \cos(2\pi f t_0) - \cos[2\pi f(t_0 + T_{\text{exp}})] \}, \quad (6)$$

式中: t_0 为单帧曝光的起始时刻。当 $T_{\text{exp}} \approx T_f$ 时, 式 (6) 中括号项趋近于 0, 可得 $\bar{I}_{\text{env}} \approx I_0$, 即闪烁的时变分量在曝光积分过程中被平均为近似恒定的直流亮度, 随后该直流分量会在平场/背景校正中被整体扣除。此外, 实验中相机自动曝光与自动增益均关闭 (增益 0 dB), 触发延迟为 0 μs 、线路防抖为 50 μs , 远小于 10 ms 的曝光窗口, 确保触发动作不会造成可见的帧间亮度波动。结合隔振平台与静态样本条件, 可以认为环境光闪烁、设备微振动等时变非平稳噪声在本实验配置下已被显著抑制, 成像误差主要由高斯噪声、椒盐噪声和散粒噪声等空间复合随机噪声主导。因此, 在本研究实验条件下, 公式 (4) 所建模的复合空间噪声可用于描述实际观测过程中的主要误差来源。在后续的噪声仿真实验及鲁棒性分析中, 均采用该模型对复合噪声进行建模。

3 多模态深度去噪模型设计

3.1 网络总体结构

本文提出的多模态深度去噪网络基于改进的去噪卷积神经网络 (DnCNN) 框架^[13], 其整体结构如图 3 所示。网络以双通道融合图像张量 $X_{\text{input}}(x, y)$ 为输入, 训练目标为学习实际含噪图 $I_{\text{noisy}}(x, y)$ 与理想无噪图 $I_{\text{clean}}(x, y)$ 之间的残差映射, 输出噪声预测图 $\hat{R}(x, y)$, 并通过残差反演得到去噪结果图 $\hat{I}_{\text{denoised}}$, 其数学形式如式 (7) 所示

$$\begin{aligned} \hat{R}(x, y) &= F_{\theta}(X_{\text{input}}(x, y)), X_{\text{input}}(x, y) \in \mathbb{R}^{H \times W \times 2}, \\ \hat{I}_{\text{denoised}}(x, y) &= I_{\text{noisy}}(x, y) - \hat{R}(x, y), \end{aligned} \quad (7)$$

式中: $F_{\theta}(\cdot)$ 为网络映射函数; θ 为参数集; H 为图像高度; W 为图像宽度。网络主体采用残差学习结构, 通过多层卷积提取特征, 并在末端重构干净图像。为增强特征表示能力, 引入通道注意力机制 (SE), 对不同通道特征分配自适应权重, 强化关键信息并抑制冗余特征。

网络主要由三部分构成: 1) 特征提取模块。由多层卷积 (Conv)、批归一化 (BN) 与修正线性激活单元 (ReLU) 构成, 用于捕捉并标准化局部纹理与结构信息; 2) 特征增强模块。SE 模块实现通道级特征的压缩与激励; 3) 重构模块。卷积层将增强后的特征映射为残差图, 并通过与输入图像相减得到最终去噪结果。

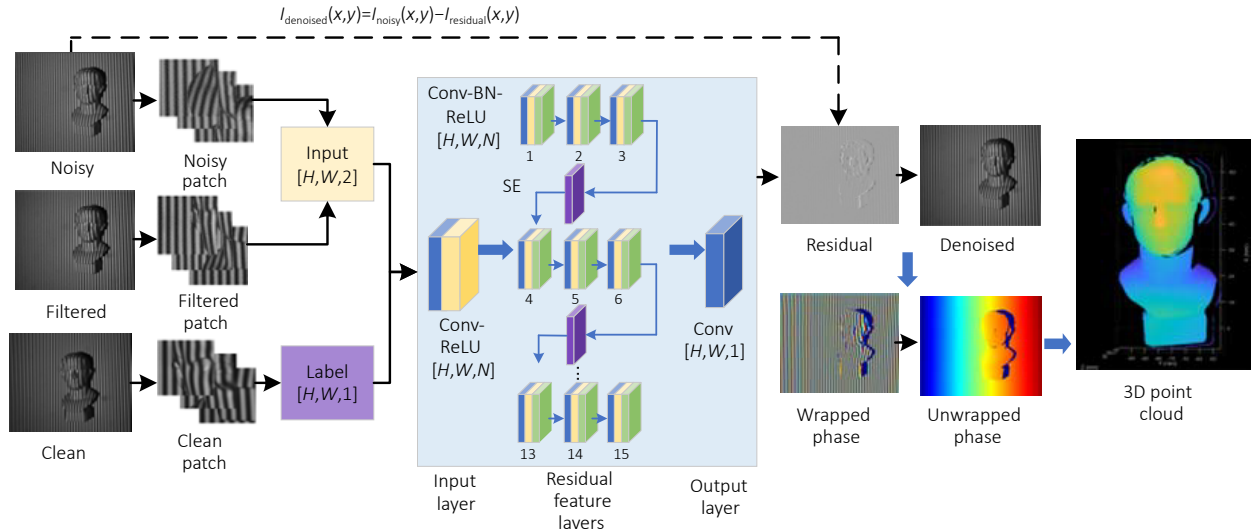


图 3 多模态 DnCNN 网络架构示意图

Fig. 3 Schematic of the multimodal DnCNN network architecture

3.2 双通道输入设计

第二节已表明结构光条纹图像常受复合噪声干扰。若仅以原始含噪图 (Noisy) 作为输入, 噪声分布复杂且高反射区域伪影严重, 模型容易过拟合局部噪声特征; 若仅依赖传统滤波结果 (Filtered), 虽然全局噪声被削弱, 但高频细节和边缘特征往往被平滑甚至丢失。二者各有优缺点^[29]。

为此, 本文提出双通道输入策略: 将 Noisy 与 Filtered 在通道维度拼接为联合输入, 使网络在学习过程中同时接收“细节保留”与“噪声抑制”的互补信息。其中, Noisy 保证纹理、条纹方向和高频结构的完整性; Filtered 则提供去除脉冲与高斯噪声后的平滑先验。通过特征融合, 网络能够自适应提取关键结构并抑制冗余干扰, 从而增强对复合噪声特性的建模鲁棒性。

具体而言, 设计了两级串联滤波模块作为预处理分支, 生成 Filtered 作为第二模态输入, 其结构所示如图 4 所示, 该模块由极值检测、选择性中值滤波 (selective median filter, SMF) 以及二维高斯平滑滤波三部分组成, 分别针对极端脉冲噪声与高频感光噪声, 实现差异化抑制。其中, SMF 相对传统中值滤波的改进在于, 先经极值检测判定脉冲噪声像素, 仅对被判定像素执行中值替换, 其余像素保持原值, 从而在抑制脉冲噪声的同时尽量减少对条纹主结构的整体平滑与边缘模糊。

输入含噪条纹图像 $I_{\text{noisy}}(x, y)$, 其包含结构信号与复合噪声项。

$$I_{\text{noisy}}(x, y) = 0 \text{ or } I_{\text{noisy}}(x, y) = 255, \quad (8)$$

第一阶段为极值检测, 若像素满足以下条件则判定为极端噪声点, 并采用 SMF 替换为 $I_{\text{SMF}}(x, y)$, 即

$$I_{\text{SMF}}(x, y) = \begin{cases} \text{median} \{I_{\text{noisy}}(\Omega_{x,y})\}, & \text{if } I_{\text{noisy}}(x, y) \in \mathcal{N}_{\text{extreme}} \\ I_{\text{noisy}}(x, y), & \text{otherwise} \end{cases}, \quad (9)$$

式中: $\Omega_{x,y}$ 为以 (x, y) 为中心的局部邻域窗口; $\mathcal{N}_{\text{extreme}}$ 为极值噪声点集合。该方法可在避免平滑非噪声区域的同时, 精准去除强脉冲噪声。

第二阶段针对高频随机噪声, 对 SMF 输出图像执行二维高斯卷积平滑, 得到滤波输出图像 $I_{\text{filtered}}(x, y)$, 式中: G_{ω} 为二维高斯核; 参数 ω 为核的标准差, 控制平滑强度, 卷积运算与核函数定义为

$$\begin{cases} I_{\text{filtered}}(x, y) = (G_{\omega} * I_{\text{SMF}})(x, y) \\ G_{\omega}(u, v) = \frac{1}{2\pi\omega^2} \exp\left(-\frac{u^2 + v^2}{2\omega^2}\right) \end{cases}. \quad (10)$$

该操作能够有效抑制随机噪声, 增强灰度场的局部连续性, 同时尽可能保持条纹主结构的光滑一致性。

通过上述两级滤波处理, 得到的串联滤波条纹图像 $I_{\text{filtered}}(x, y)$ 具有更高的条纹清晰度与背景稳定性。最终, 将 Noisy 与 Filtered 沿通道维度拼接为双通道张量输入 $X_{\text{input}}(x, y)$ 可以表示为

$$X_{\text{input}}(x, y) = \text{cat}(I_{\text{noisy}}(x, y), I_{\text{filtered}}(x, y)) \in \mathbb{R}^{H \times W \times 2}, \quad (11)$$

式中: $\text{cat}(\cdot)$ 表示沿通道维度的拼接算子; H, W 分别为图像高度与宽度 (单位: 像素)。该设计保证了输入既保

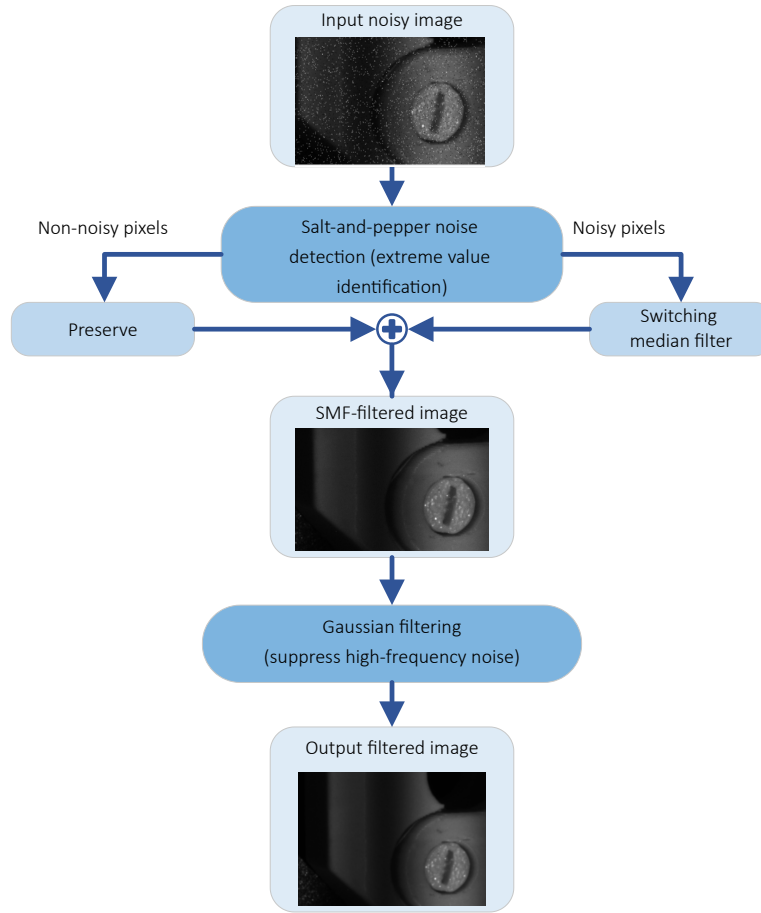


图 4 两级串联滤波模块结构图

Fig. 4 Architecture of the two-stage cascaded filtering module

留高频细节, 又抑制复合噪声干扰, 为后续网络的特征学习提供稳定而鲁棒的支撑。

3.3 通道注意力增强机制设计

在结构光图像去噪中, 基础 DnCNN 仅依赖空间卷积提取特征, 缺乏对通道维度的显式建模, 难以刻画多模态输入之间的互补与依赖关系。统一卷积易造成通道响应失衡, 导致有效结构被削弱、噪声残留增强、边缘模糊, 进而影响相位提取与三维重建的稳定性。为此, 本文在主干网络中引入轻量级通道注意力机制——SE 模块^[30], 通过压缩、激励和重标定三个阶段, 显式建模通道间的全局依赖关系^[31]。图 5 给出了 SE 模块的结构和作用机制, 设某一卷积模块的输出特征图为 $F \in \mathbb{R}^{H \times W \times C}$, 其中 C 为通道数, 第 c 个通道特征图记为 $F_c \in \mathbb{R}^{H \times W}$ 。

具体而言, SE 模块首先对每个通道进行通道汇聚(全局平均池化), 得到通道描述向量 $Z \in \mathbb{R}^C$ 为

$$Z_c = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W F_c(i, j), c = 1, \dots, C, Z = [Z_1, Z_2, \dots, Z_C], \quad (12)$$

用于刻画全局响应强度, 随后, 将第 c 个通道的全局响应强度 Z_c 输入由两层全连接(等效 1×1 卷积)构成的瓶颈映射, 先经 ReLU 再经 Sigmoid 激活得到通道权重向量 $s \in \mathbb{R}^C$ 为

$$s = \vartheta(W_2 \cdot \delta(W_1 \cdot Z)), W_1 \in \mathbb{R}^{\frac{C}{r} \times C}, W_2 \in \mathbb{R}^{C \times \frac{C}{r}}, \quad (13)$$

式中: r 表示压缩率, 控制瓶颈通道数 C/r , r 越大压缩越强、计算量更小, 但通道交互能力可能下降, r 越小则相反, 参考通用做法并结合本任务的计算预算, 我们在小范围内尝试 $r=\{8, 16, 32\}$ 等典型取值, 观察其对训练稳定性和指标表现的影响, 结果表明 $r=16$ 在精度与开销之间更为稳妥, 故本文取 $r=16$; $\delta(\cdot)$ 为 ReLU 激活函数; $\vartheta(\cdot)$ 为 Sigmoid 激活函数。最后, 利用通道权重对输入特征做逐通道重标定, 按权重对各通道 F_c 进行缩放, 实现结构增强与噪声抑制, 得到输出 $\tilde{F}_c \in \mathbb{R}^{H \times W}$ 为

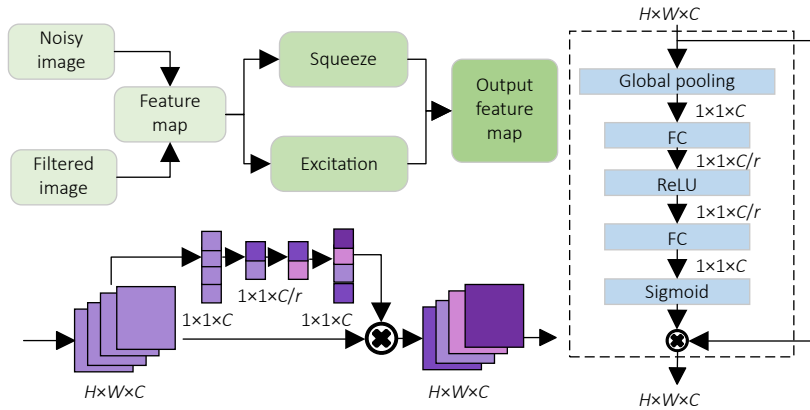


图 5 SE 模块结构与作用机制示意图

Fig. 5 Schematic of the structure and functional mechanism of the SE module

$$\begin{cases} \tilde{F}_c = s_c \cdot F_c, c = 1, \dots, C \\ \tilde{\mathbf{F}} = [\tilde{F}_1, \dots, \tilde{F}_C]^T \in \mathbb{R}^{H \times W \times C} \end{cases} \quad (14)$$

式中: $\tilde{F}_c$ 表示经 SE 重新标定后的输出第 c 通道特征图, $\tilde{\mathbf{F}}$ 表示经 SE 模块重新标定后的输出特征张量。通过该机制, SE 模块能够自适应突出条纹方向性和边缘显著区域, 提高去噪结果的结构保真与细节还原能力。为保持计算轻量化, 本文采用集成式策略, 将 SE 模块插入每个残差块的卷积层之后, 构成由 Conv、BN、ReLU 与 SE 组成的增强型卷积块。

3.4 多目标驱动的混合损失函数设计

传统图像去噪多以像素级均方误差 (MSE) 作为唯一优化目标, 虽然计算简单、收敛稳定, 但在结构光条纹图场景下往往伴随条纹模糊与边缘位置偏移。为此, 本文设计了融合亮度、结构与边缘约束的混合损失函数, 以多目标联合优化的方式提升去噪与结构保持能力^[32], 如图 6 所示。

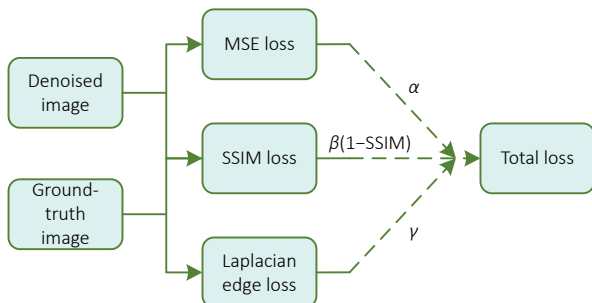


图 6 混合损失函数结构示意图

Fig. 6 Schematic of the hybrid loss function architecture

该损失函数包括三部分。

1) 亮度域损失 L_{MSE} 表示为

$$L_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N (\hat{I}_{\text{denoised},i} - I_{\text{clean},i})^2. \quad (15)$$

保证去噪图 $\hat{I}_{\text{denoised}}$ 与参考图 I_{clean} 在像素灰度 N 上的一致性, 提升峰值信噪比 (PSNR)。

2) 结构域损失 L_{SSIM} 表示为

$$\begin{cases} \text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \\ L_{\text{SSIM}} = 1 - \text{SSIM}(\hat{I}_{\text{denoised}}, I_{\text{clean}}) \end{cases} \quad (16)$$

式中: μ 、 σ 、 σ_{xy} 分别表示局部均值、方差与协方差; C_1 、 C_2 为稳定常数。强调局部结构、对比度与纹理一致性, 增强条纹连续性和感知稳定性。

3) 边缘域损失 L_{Edge} 表示为

$$L_{\text{Edge}} = \frac{1}{N} \sum_{i=1}^N \|\Delta(\hat{I}_{\text{denoised},i}) - \Delta(I_{\text{clean},i})\|_1. \quad (17)$$

通过二阶拉普拉斯算子 $\Delta(\cdot)$ 提取高频分量, 对比预测与真实边缘差异, 强化细节与突变区域的保持能力; 最终混合损失 L_{total} 表示为

$$L_{\text{total}} = \alpha \cdot L_{\text{MSE}} + \beta \cdot (1 - L_{\text{SSIM}}) + \gamma \cdot L_{\text{Edge}}, \quad (18)$$

式中: α 、 β 、 γ 为超参数, 分别对应 L_{MSE} 、 L_{SSIM} 、 L_{Edge} , 为确定混合损失中各分量的权重, 本文在合理范围内对 $\alpha \in [0.4, 0.8]$ 、 $\beta \in [0.1, 0.3]$ 、 $\gamma \in [0.05, 0.25]$ 进行小范围调节, 并在验证阶段结合 PSNR、SSIM 及边缘保持等指标观察其影响, 最优参数组合通过在验证集上的网格搜索 (grid search) 确定。综合考虑训练稳定性与三类指标表现, 最终采用 $\alpha=0.6$ 、 $\beta=0.2$ 、 $\gamma=0.2$, 该设置在三类指标间取得较好权衡, 以 MSE 为主导保证收敛与亮度保真, 同时结合 SSIM 与 Laplacian Edge 约束增强纹理与轮廓细节的保持能力。

3.5 数据集构建与增强

为使网络充分学习结构光条纹图像中的多源噪声特性与基本相位模式，本文以多种规则几何体 (标准阶梯块、球体、椎体) 作为主要训练对象，规则几何体能够覆盖边缘突变、曲率连续与尖锐顶点等典型结构特征，有助于网络在训练阶段获得更具代表性的条纹与相位模式。构建由三种模态 (Noisy、Filtered 与多帧平均 Clean 图^[33]) 组成的训练集，三类图像像像素级配准，分别用于原始输入、结构补充及监督标签。

此外，为增强网络的局部结构感知与尺度泛化能力，采用多尺度滑窗裁剪如图 7 与分阶段训练机制结合如表 1。同时利用随机滑窗、水平/垂直翻转及旋转等空间增强方式，结合梯度幅值引导的高频区域采样策略，重

点覆盖条纹边缘与纹理过渡区域^[34]，从而提升模型在结构细节保持与复合噪声抑制方面的鲁棒性。

4 实验验证与性能评估

4.1 实验方案设计

本研究搭建了一套双目结构光测量平台如图 8 所示，用于采集条纹图像并完成网络训练与三维重建验证。系统硬件包括：结构光投影仪 (DLPLCR4500EVM，912 px × 1140 px，120 Hz)、2 台工业相机 (MV-CH120-11CM，4096 px × 3000 px) 图像采集卡 (SOL2MEVCLF) 和运动控制平台 (步进电机与传送带，5~100 mm/s)。深度学习网络在 Windows10 平台运行，配置 i7-10750CPU 和 NVIDIA GTX1660Ti 显卡。网络训练和三维重建主要采用 MATLAB R2024b 和 CloudCompare 实现。

实验流程主要包括七个环节：系统标定^[35-36]、图案设计、图像采集、相位提取、相位解包裹^[37]、相位匹配与三维重建^[38]。标定结果如表 2 所示。

其中，内参记为 $K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$ ， f_x, f_y 为有效焦距； c_x, c_y 为主点坐标；外参 R 为旋转矩阵； T 为平移矩阵，

本文将左相机坐标系定义为世界坐标系，故左相机外参取 $R_l = I, T_l = \mathbf{0}$ ；右相机外参表示为 $(R_{r \rightarrow l}, T_{r \rightarrow l})$ ；投影仪外参表示为 $(R_{p \rightarrow l}, T_{p \rightarrow l})$ ； e_{rms} 为重投影误差反映标定拟合精度；畸变系数为径向 (k_1, k_2, k_3) 与切向 (p_1, p_2) 的去畸变参数。

4.2 图像去噪结果与质量评价

为验证所提方法在不同结构复杂度下的鲁棒性与泛化性，选取了三类典型样本如图 9 所示。标准台阶块几何边界与高度分布清晰，适于评估结构还原能力。此外，引入与训练类别不同的全新样本：石膏像，表面曲率复

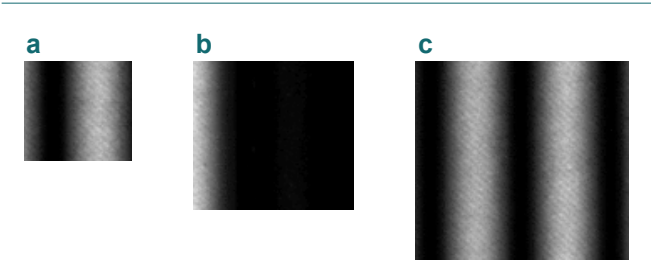


图 7 不同尺度 Patch 样本。(a) 64 px × 64 px; (b) 96 px × 96 px; (c)128 px × 128 px

Fig. 7 Patch samples at different spatial scales. (a) 64 px × 64 px; (b) 96 px × 96 px; (c) 128 px × 128 px

表 1 多尺度 patch 设置与训练分配策略

Table 1 Multi-scale patch configuration and training assignment strategy

Patch size /px	Sliding window stride /px	Training stage (Epoch)	Learning rate (Adam)
64×64	32	Epoch 1–20	5e-4
96×96	48	Epoch 21–35	3e-4
128×128	64	Epoch 36–50	1e-4、1e-5

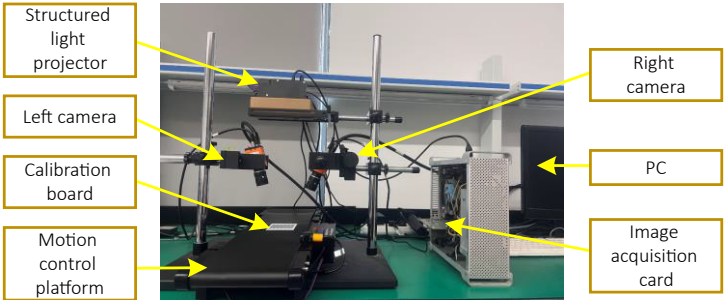


图 8 双目结构光图像采集系统实景图

Fig. 8 Photograph of the binocular structured light image acquisition system

杂、细节丰富，用于评估模型在真实曲面上的泛化能力；PCB 电路板，包含高反射元件与细密纹理，用于考察模型在复合噪声环境与材料特性差异下的适应性。整体采集共获得 1440 幅配准图像，其中 1152 幅用于训练，288 幅用于测试。

三类对象在相同采集条件下获得的含噪条纹图 (Noisy)、串联滤波结果 (Filtered) 及网络去噪结果 (Denoised) 进行对比，如图 10 所示。质量评估以 Clean 图为参考，采用峰值信噪比 (PSNR)、结构相似性指数 (SSIM) 及均方根误差 (RMSE) 三项指标^[39]。各样本经

表 2 双目结构光系统标定结果
Table 2 Calibration results for the binocular structured-light system

Sensor	Intrinsics $(f_x, f_y) / \text{px}; (c_x, c_y) / \text{px}$	Extrinsics $R / -; T / \text{mm}$	Reprojection errors $e_{\text{rms}} / \text{px}$	Distortion coefficients / -Radial $(k_1, k_2, k_3); \text{Tangential } (p_1, p_2) \times 10^{-3}$
Left camera	(3637.76, 3637.73); (1231.40, 999.93)	$R_1 = I \quad T_1 = 0$	0.070	(-0.0956, 0.2725, -0.9394); (-1.300, -1.100)
Right camera	(3628.19, 3629.24); (1197.52, 1012.36)	$R_{r \rightarrow l} = \begin{bmatrix} 0.9258 & -0.0274 & -0.3768 \\ -0.0151 & 0.9938 & -0.1096 \\ 0.3775 & 0.1072 & 0.9197 \end{bmatrix}$ $T_{r \rightarrow l} = \begin{bmatrix} 67.4552 \\ -58.2165 \\ 187.9625 \end{bmatrix}$	0.070	(-0.1155, 0.7025, -4.0291); (-0.1853, -0.5346)
Projector	(1113.33, 2224.74); (449.20, 1158.62)	$R_{p \rightarrow l} = \begin{bmatrix} 0.9649 & 0.0450 & -0.2586 \\ -0.0260 & 0.9967 & 0.0766 \\ 0.2612 & -0.0672 & 0.9629 \end{bmatrix}$ $T_{p \rightarrow l} = \begin{bmatrix} -76.9133 \\ -64.9604 \\ 173.2837 \end{bmatrix}$	0.045	(0.0273, 0.0048, -0.5186); (-2.800, -0.700)



图 9 典型测试样本示意图。(a) 标准台阶块；(b) 石膏像；(c) 印刷电路板
Fig. 9 Representative test samples in the experiments. (a) standard step block; (b) plaster bust; (c) PCB

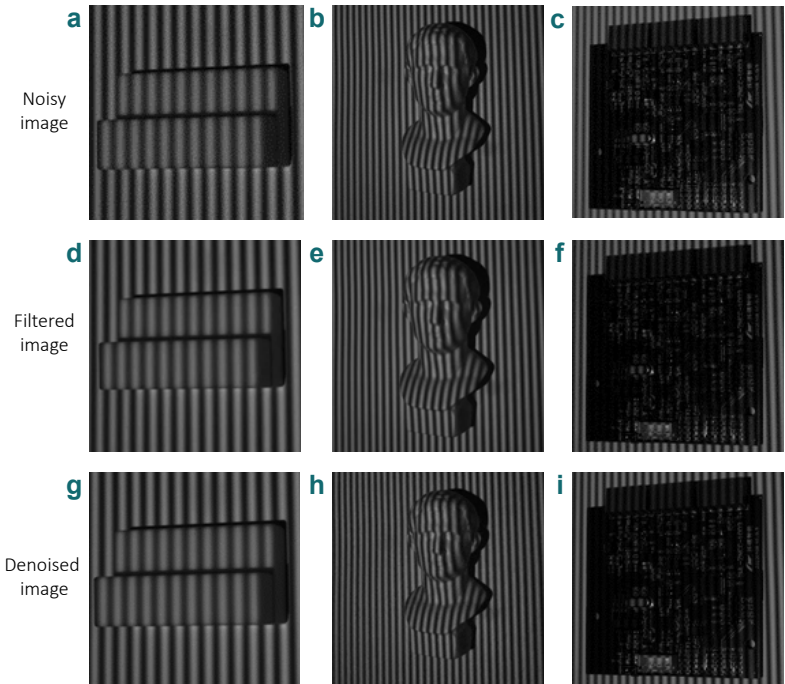


图 10 典型样本在各阶段图像处理下的可视化结果。(a,d,g) 标准台阶块；(b,e,h) 石膏像；(c,f,i) 印刷电路板
Fig. 10 Visual comparison of typical samples at different image processing stages. (a,d,g) Standard step block; (b,e,h) Plaster bust; (c,f,i) PCB

DnCNN 模型去噪前后的指标统计结果见表 3。

具体而言，标准台阶块样本 SSIM 提升约 8.4%，RMSE 降低约 62.6%，说明模型在规则边界场景中表现优异；石膏雕塑样本 SSIM 提升约 8.3%，RMSE 降幅约 61.6%，表明模型在曲率连续的平滑表面上能够有效抑制噪声并保持形貌完整性；PCB 样本 SSIM 提升约 8.6%，PSNR 提高约 19.5%，体现了模型在高频纹理与复合噪声条件下的细节恢复能力与干扰抑制能力。综上，所提方法在三类典型样本中均表现出针对性优势，为相位提取与三维重建提供了高质量输入。

4.3 消融实验与方法对比

为验证各模块对模型性能提升的贡献，以石膏像为对象，开展了消融实验，包括：1) 输入形式 (单通道 Noisy/双通道 Noisy 与 Filtered 拼接)；2) 注意力机制 (有/无 SE 模块)；3) 损失函数组合 (MSE、SSIM 与 Laplacian 的混合损失函数)。实验结果如表 4 所示。

结果表明，双通道输入较单通道显著提升去噪效果，

平均 PSNR 提高约 1.05 dB，说明多模态融合增强了特征鲁棒性；SE 模块在 PSNR 和 SSIM 上均有稳定增益，尤其在边缘细节恢复方面表现突出；混合损失函数优于单一损失，在三项指标上均达到最佳。与原始 DnCNN 相比，本文改进模型的平均 PSNR 提升约 1.79 dB，SSIM 增加 0.0185，RMSE 降低 27.5%，均优于原始模型，验证了本文设计多模态输入、结构保持机制与混合损失函数对复合噪声抑制的有效性。

为进一步验证本文改进模型的去噪效果，选取了两类常见的传统去噪方法 NLM^[40] 与 BM3D^[41] 和两种代表性的深度学习去噪方法 BM3D-Net^[42] 与 DIVA^[43]，以石膏像为对象进行对比实验。实验结果如图 11 和表 5 所示。

如表 5 所示，传统方法在一定程度上改善了图像质量，但整体劣于深度学习方法，说明神经网络方法在结构光条纹图像的去噪效果和细节保持方面更具优势。其次对比两种神经网络方法，从图 11 直观视觉效果来看，本文所提方法在几种方法中纹理保持最好、噪声残留最

表 3 去噪前后图像质量评价指标对比
Table 3 Comparison of image quality metrics before and after denoising

Sample type	Processing method	PSNR/dB	SSIM	RMSE
Standard step block	Noisy	35.74	0.8909	0.0163
	DnCNN	43.11	0.9654	0.0061
Plaster bust	Noisy	35.73	0.8892	0.0164
	DnCNN	42.90	0.9627	0.0063
PCB	Noisy	35.71	0.8881	0.0166
	DnCNN	42.69	0.9643	0.0062

表 4 消融实验与原始模型对比结果 (数据集：石膏像样本)
Table 4 Ablation study and comparison with the original DnCNN model (Dataset: plaster bust)

Experiment setting	Input	SE module	Loss function	PSNR/dB	SSIM	RMSE
Original DnCNN	Single-channel (Noisy)	No	MSE	41.11	0.9442	0.0087
Improved-1	Dual-channel (Noisy+Filtered)	NO	MSE	42.16	0.9552	0.0072
Improved-2	Dual-channel (Noisy+Filtered)	YES	MSE	42.85	0.9632	0.0071
Improved-3 (final)	Dual-channel (Noisy+Filtered)	YES	MSE,SSIM and Edge	42.90	0.9627	0.0063

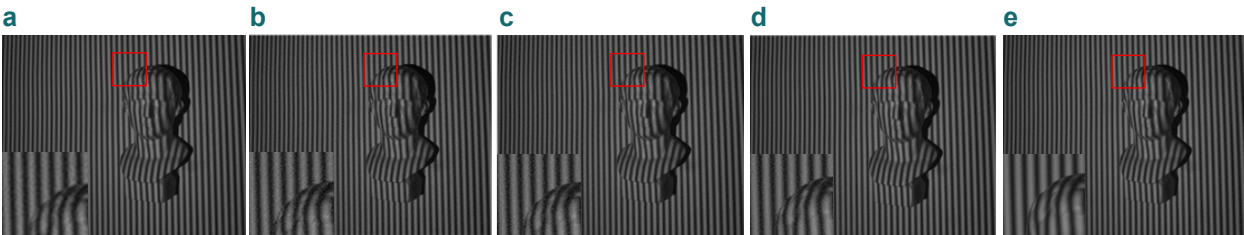


图 11 石膏像去噪效果对比。(a) NLM；(b) BM3D；(c) BM3D-Net；(d) DIVA；(e) 本方法
Fig. 11 Denoising performance comparison of the plaster bust. (a) NLM; (b) BM3D; (c) BM3D-Net; (d) DIVA; (e) Our method

表 5 去噪评价指标结果表 (数据集: 石膏像样本)
Table 5 Quantitative evaluation results of denoising performance (Dataset: plaster bust)

Method	PSNR/dB	SSIM	RMSE
NLM	39.4	0.9253	0.0097
BM3D	40.5	0.9361	0.0090
BM3D-Net	40.9	0.9445	0.0087
DIVA	41.8	0.9540	0.0078
Our method	42.9	0.9627	0.0063

少。从定量指标来看, 本文方法在三项指标上均取得最优结果, 优势较为明显。相比 BM3D-Net, PSNR 提升约 4.89%, SSIM 提升约 1.93%, RMSE 降低 27.59%; 相比 DIVA, PSNR 提升约 2.63%, SSIM 提升约 0.91%, RMSE 降低 19.23%。

4.4 噪声强度变化下的抗噪上限分析

为了进一步验证所提多模态 DnCNN 网络在不同噪声强度下的鲁棒性与抗噪上限, 本节以石膏像为对象, 在 Clean 条纹图上基于式 (4) 构造了一组由弱噪声到强噪声的复合噪声序列。具体而言, 参照图像去噪领域的通用设置^[44], 以高斯噪声标准差设为 $\sigma_g = 5, 10, 15, 20$ 四个典型噪声等级, 用于覆盖图像质量从轻微受损到明显退化的主要区间。在复合噪声模型中, 椒盐噪声密度 ρ_{sp} 随 σ_g 同步递增 (分别为 0.01、0.03、0.05、0.08), 以增强极值噪声成分并模拟更为恶劣的复合噪声条件。

实验过程如下: 1) 以石膏 Clean 条纹图作为参考图像; 2) 依据公式 (4) 在 Clean 图上叠加不同标准差的高斯噪声以及不同密度的椒盐噪声, 构造由弱到强噪声的一组复合观测图像; 3) 分别用原始 DnCNN 模型 (Original DnCNN) 和本文提出的多模态 DnCNN 网络 (Proposed DnCNN) 对上述含噪图进行去噪处理, 得到两

组结果, 同时将含噪图像 (Noisy) 作为基准参照; 4) 以 Clean 图为参考, 计算各噪声强度下 Noisy、Original DnCNN 与 Proposed DnCNN 的平均 PSNR、SSIM 与 RMSE 指标, 并绘制“图像质量指标-噪声强度”的变化曲线, 如图 12 所示。

可以看到, 随着噪声强度的增强, Noisy 曲线的 PSNR、SSIM 快速退化, RMSE 大幅上升, 说明噪声显著破坏条纹结构。在相同噪声条件下, Original DnCNN 和 Proposed DnCNN 两种网络均显著优于 Noisy, 验证了深度网络对复合噪声的抑制能力。更为重要的是, 在整个噪声范围内, 所提 Proposed DnCNN 网络的 PSNR 与 SSIM 始终高于原始 DnCNN 模型, 且随噪声增强的下降幅度更小, RMSE 增长更缓, 表现出更平滑的退化曲线。例如, 在最强噪声档 $\sigma_g = 20$ 下, Noisy 的 PSNR/SSIM 仅约为 15.7 dB/0.4057, 而 Original DnCNN 和 Proposed DnCNN 分别可达到约 25.6 dB/0.7514 与 27.6 dB/0.8032, 说明所提方法在远高于实际系统噪声水平的强噪声条件下仍能保持较好的稳定性和鲁棒性。

结合前文 4.3 节在实际采集条件下的结果可知, 真实系统中观测条纹图像的 PSNR 约为 35 dB, 而本节构造的复合噪声序列使 Noisy 的 PSNR 降至 28~15 dB 区间, 噪声水平明显高于实际采集场景。因此, 可以认为本节

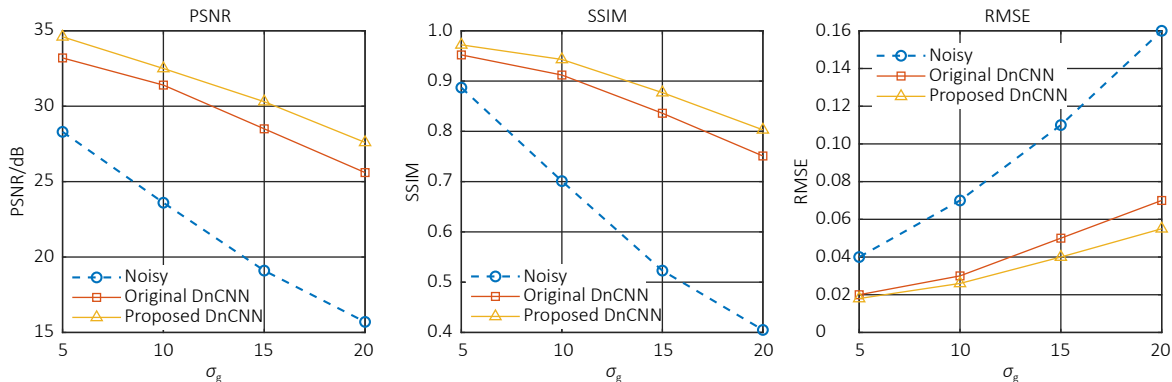


图 12 不同噪声等级下三种方法的图像质量指标变化曲线
Fig. 12 Image quality metrics under different noise levels

构造的噪声扫描实验已经覆盖了从实际系统噪声水平到更强噪声环境的宽噪声区间。在该区间内, 所提多模态 DnCNN 网络亦展现出良好的抗噪上限与稳定性能, 为后续相位解包裹与三维重建提供了更加可靠的输入保障, 并具备在高噪声复杂工业环境中的潜在应用价值。

4.5 相位解包裹与三维点云重建结果分析

基于三频四步相移与多频外差法^[37], 以 PCB 为对象, 对去噪前 (Noisy) 及网络去噪后 (Denoised) 的条纹图进行相位提取与三维重建。图 13(a,b,d,e) 展示了包裹相位、解包裹相位 (绝对相位) 的结果对比, 其中 Noisy 相位图存在大量相位跳变与条纹断裂, 尤其集中于边缘及高反射区域, 解包裹连续性受损; 而 Denoised 相位图条纹结构更清晰, 连续性显著增强, 在边缘与细节区域保持了更高的鲁棒性, 误差传播得到了有效抑制。为直观评估去噪对相位连续性的影响, 本文在 PCB 样本中选取一条位于高反射/复杂走线区域的列方向相位剖面, 其具体位置在图 13(a) 中以黄色直线标出, 并在局部像素区间内进行放大, 如图 13(c) 所示。可以看出, Noisy 曲线沿程存在明显的高频抖动和局部跳变, 个别区域出现突兀的相位落差; 而 Denoised 曲线整体更加平滑, 趋

势更接近理想基线, 并有效消除了噪声引起的不连续点, 表明去噪过程未破坏真实结构的整体形态。相位梯度分布同样呈现一致的变化, 如图 13(f) 所示, Denoised 直方图在零点附近更加集中, 高梯度离群值的数量明显减少, 说明局部畸变得到抑制, 条纹的连续性和局部几何保持性均有所改善。这些结果验证了所提出的去噪方法在抑制噪声扰动的同时能够维持相位场的真实结构, 为后续三维重建精度的提升奠定了可靠基础。

结合表 2 的系统标定参数, 通过三角测量将绝对相位转换为深度坐标, 生成三维表面点云如图 14 所示。在 xyz 三维视角下, Noisy 点云存在大量离散高反射伪点与异常沉点; Denoised 点云整体结构连续平滑, 杂散点显著减少, 空间一致性更优。在 xy 平面视角下, Noisy 点云存在边缘断裂与结构模糊现象, 局部轮廓扩散明显且干扰点较多; Denoised 点云轮廓完整度显著提升, 噪点区域明显减少。

量化评价采用点到面 (C2M) 距离作为几何误差度量, 比较去噪前后重建结果在点云密度、均方根误差 (RMSE)、95% 分位误差 (P95) 及最大误差等指标上的差异。结合已有研究在高反射或结构复杂表面条件下, 工业级结构光系统的重建精度范围^[45-48], 综合结果如表 6

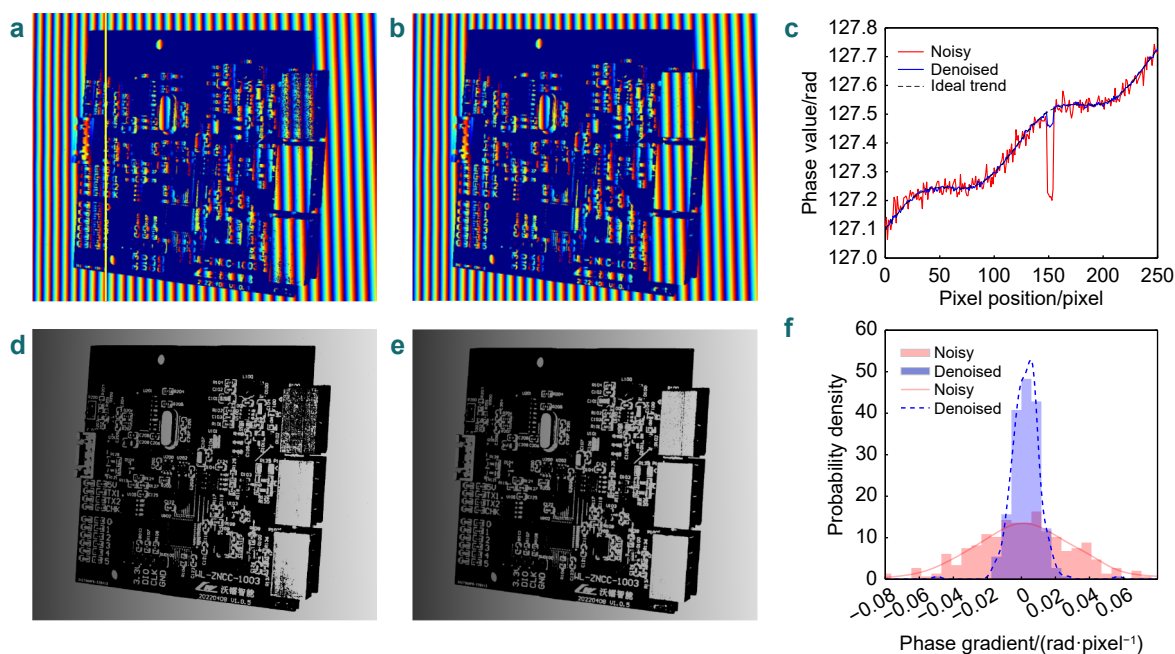


图 13 去噪前后 PCB 的相位提取结果对比。(a) Noisy 包裹相位图; (b) Denoised 包裹相位图; (c) 沿图 (a) 黄线位置截取的列方向相位剖面对比; (d) Noisy 绝对相位图; (e) Denoised 绝对相位图; (f) 相位梯度分布统计图

Fig. 13 Comparison of phase extraction results for PCB before and after denoising. (a) Noisy wrapped phase map; (b) Denoised wrapped phase map; (c) Column-wise phase profile comparison map extracted along the yellow line in (a); (d) Noisy absolute phase map; (e) Denoised absolute phase map; (f) Phase gradient distribution map

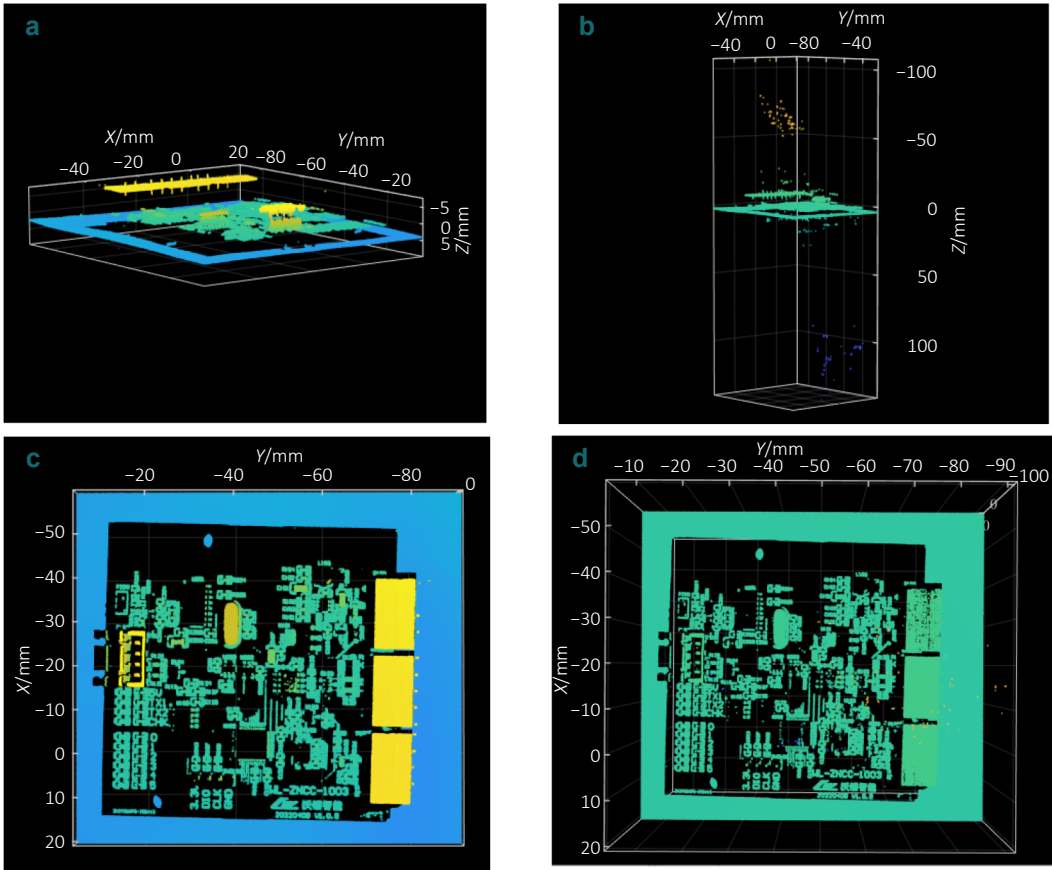


图 14 去噪前后 PCB 三维表面形貌对比。(a) Noisy 三维点云图；(b) Denoised 三维点云图；(c) Noisy 三维点云 XY 方向俯视图；(d) Denoised 三维点云 XY 方向俯视图

Fig. 14 Comparison of 3D surface morphology of the PCB before and after denoising. (a) Noisy 3D point cloud; (b) Denoised 3D point cloud; (c) Noisy 3D point cloud in XY top view; (d) Denoised 3D point cloud in XY top view

表 6 点云质量评估指标对比

Table 6 Comparison of point cloud quality evaluation metrics

Indicator	RMSE/mm	95% error/mm	Maximum error/mm	Point cloud density/(pts/mm ²)
Noisy Point Cloud	0.7140	0.2376	13.65	180.22
Denoised Point Cloud	0.3774	0.1369	11.37	258.74
Representative Metrics	0.20-0.60	0.20-0.80	8-15	100-500

所示。绝对误差定义为 $|E| = |Z_{\text{method}} - Z_{\text{ref}}|(\text{mm})$ ，其中 Z_{method} 为 Noisy 或 Denoised 的重建结果； Z_{ref} 为多帧均值重建面 (同条件采集 24 帧、像素级配准后在置信度掩膜内做截断均值/中值融合得到)，在本实验中作为高信噪比的参考几何，用于近似真实表面形貌。所有统计仅在统一 ROI 的有效像素上进行，百分比按有效像素数归一化。考虑到极少数饱和/强反射像素可能造成尾部偏置，结果解读以 P95 与 RMSE 为主，同时给出最大误差以完整体现尾部行为；连通域相关统计 (见表 7、表 8) 采用 8-连通，阈值 $t=2/5\text{ mm}$ 与图 15 保持一致。

在 PCB 样本上，Denoised 点云密度较 Noisy 增长

表 7 高误差像素占比 $P(E>t)$

Table 7 Proportion of high-error pixels $P(E>t)$

Method	$P(E>2\text{ mm})/\%$	$P(E>5\text{ mm})/\%$
Noisy	0.72	0.22
Denoised	0.33	0.07

约 43.6%，说明噪声抑制并非通过简单剔除散乱点实现，而是恢复了更多满足几何约束的有效点，曲面连续性与完整度得到增强；RMSE 降低约 47.1%，95% 分位误差降低约 42.4%，表明绝大多数点的重建精度稳定且细节保真度较高，处于复杂场景较优水平。最大误差降低约

表 8 高误差连通域的空间统计 ($E>2\text{ mm}$ 与 $E>5\text{ mm}$)
Table 8 Spatial statistics of high-error connected components ($E>2\text{ mm}$ and $E>5\text{ mm}$)

Method	Thr	Rsingle/%	A50/px	A90/px	$\bar{d}_{\text{NND}}$ /px
Noisy	>2	28	3	9	3.6
Denoised	>2	66	1	2	6.9
Noisy	>5	38	2	6	4.2
Denoised	>5	78	1	2	7.2

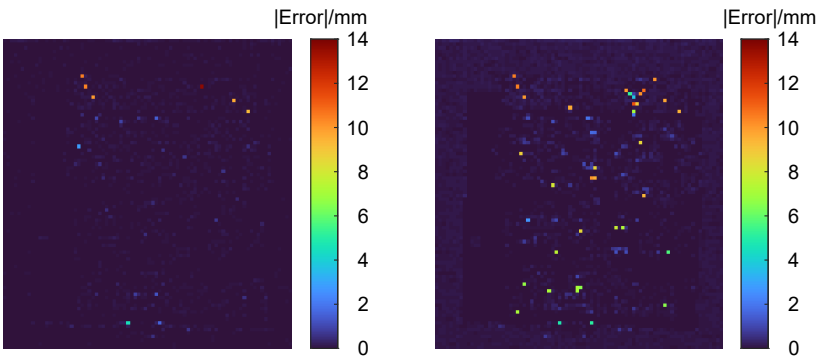


图 15 三维重建误差的空间分布热力图对比。(a) Denoised; (b) Noisy
Fig. 15 Comparison of spatial error distributions in 3D reconstruction. (a) Denoised; (b) Noisy

16.7%，去噪后仍存在约 11.37 mm 的极端值，但该类误差主要出现在 PCB 边缘及强反射元件附近，与场景最大高度变化处的局部饱和/遮挡有关。可视为少量局部重建失败点而非全局系统性偏差 (如图 15 所示呈稀疏星点分布)，在实际精密测量中通常通过 ROI 限定或掩膜剔除对评估结果的影响。剔除这类边缘/强反射引起的少量异常后，整体误差分布仍以 RMSE 与 P95 的下降为主。

在整体指标之外，进一步从空间分布的角度分析高误差像素的形态与连通性，以区分“片状聚集”的系统性失真与“稀疏孤立”的少量伪影。为此，除统计高误差像素占比 $P(E>t)$ 之外，还引入单像素连通域占比 Rsingle、连通域面积中位数/90 分位 (A50/A90) 以及平均最近邻距离 $\bar{d}_{\text{NND}}$ 等指标，用于刻画高误差连通域的尺度和分布，这在工程点云质量评估中能够补充反映伪影收敛程度与空间均匀性。

如图 15 所示的误差热力图直观地给出了去噪前后的差异：Noisy 点云在 PCB 边缘及强反射元件 (如金属端子、焊点) 周围存在成片/带状的高误差区，表明原始重建在高亮、高坡度区域易产生大面积失真；而 Denoised 点云在大部分主体区域呈均一且较低的误差水平，高亮区域数量显著减少，残余高误差主要呈局部零散小斑点分布。定量统计与此一致：表 7 显示，在统一

阈值下，高误差像素占比由 Noisy 的 0.72%/0.22% ($E>2/5\text{ mm}$) 降至 Denoised 的 0.33%/0.07%，尾部误差规模显著收敛；表 8 进一步表明，在 $E>2\text{ mm}$ 与 $E>5\text{ mm}$ 两档下，Denoised 的单像素连通域占比 (Rsingle) 提升至 66%/78%，连通域面积中位/90 分位 (A50/A90) 压缩至 1/3~1/2，平均最近邻距离 ($\bar{d}_{\text{NND}}$) 增至 6.9/7.2 px，高误差由“片状聚集”转为“稀疏孤立小斑点”。结合全局 C2M 指标，可以认为本文方法在显著抑制误差长尾、均化空间分布的同时，未引入新的结构性伪影。去噪后高误差像素总体数量明显减少，高误差区域由片状/带状结构收敛为局部零散小斑点，空间分布更加均匀。

在全球 C2M 指标及高误差连通性分析的基础上，我们进一步关心的关键问题是：去噪是否以牺牲边缘几何结构为代价。考虑到结构光在高度突变区域更易产生相位伪影和局部大误差，我们在参考面 Z_{ref} 上计算梯度幅值 $|\nabla Z_{\text{ref}}|$ ，并按分位阈值将像素划分为平坦区与高梯度边缘区。具体地，取 $|\nabla Z_{\text{ref}}|$ 的 Top10%~15% 作为边缘区，其余区域为平坦区，采用 8 连通掩膜，仅在统一 ROI 的有效像素上进行统计。

表 9 所示的区域化误差分析进一步从定量层面说明了去噪对结构和边缘几何的影响。平坦区与高梯度边缘区的 RMSE 与 P95 在去噪后均有所下降，表明整体与局

部区域的重建误差均得到不同程度改善。同时，边界精度 (DBE-Accuracy) 由 0.929 提升至 0.982，边界完整性 (DBE-Completeness) 由 0.463 提升至 0.964，法向角 (Normal angle)P95 由 5.47°降至 2.89°，法向角平均误差 (Mean) 由 4.04°下降至 1.41°。这些指标共同反映了局部表面坡度偏差减少，边缘位置及其几何形态与参考面更加一致。

这表明所提方法在压制噪声的同时保持了关键结构与边缘细节。结合图 15 与表 7、8 中阈值域与连通域形态由“片状聚集”收敛为“稀疏孤点”的观察，可确认本文方法在显著抑噪的同时未造成边缘圆滑化或结构削弱，关键几何细节得到有效保持。

为了在视觉层面验证上述指标，本文进一步选取 PCB 表面三个具有代表性的结构区域进行放大对比，

如图 16 所示。这些区域分别覆盖了典型的焊盘阵列 (ROI-1)、丝印字符笔画及其邻近狭窄高度带 (ROI-2)，以及走线与焊点混杂处的复杂边缘结构 (ROI-3)，均属于结构光重建中最易受到噪声干扰的敏感部位，因此能够有效反映去噪过程是否破坏真实几何。

如图 16 所示，Noisy 重建在三个区域中存在明显的高频散点噪声、局部轮廓断裂、走线边缘抖动以及少量离群点等典型失真现象，导致真实几何形态在视觉上呈现破碎与不连续。相比之下，Denoised 重建在相同位置的表现明显改进，焊盘边界、细小结构以及线路折弯等几何特征得以更好地保留，未出现边缘圆滑化或几何结构丢失的现象。同时，这些区域的整体几何轮廓、边界位置以及高度变化仍与原始重建保持一致，表明去噪并未“抹平”真实细节，而是有效地去除了噪声成分。可以

表 9 结构与边缘保持的分区量化评估
Table 9 Region-wise evaluation of structural/edge preservation

Region	Metric	Pre-denoising (Noisy)	Post-denoising (Denoised)
Flat region (~85% pixels)	95% error/ mm	0.225	0.130
	RMSE/ mm	0.680	0.360
	95% error/ mm	0.310	0.170
	RMSE/ mm	0.920	0.480
Edge region (Top10%–15%)	DBE-Accuracy	0.929	0.982
	DBE-Completeness	0.463	0.964
	Normal angle P95/(°)	5.47	2.89
	Normal angle Mean/(°)	4.04	1.41

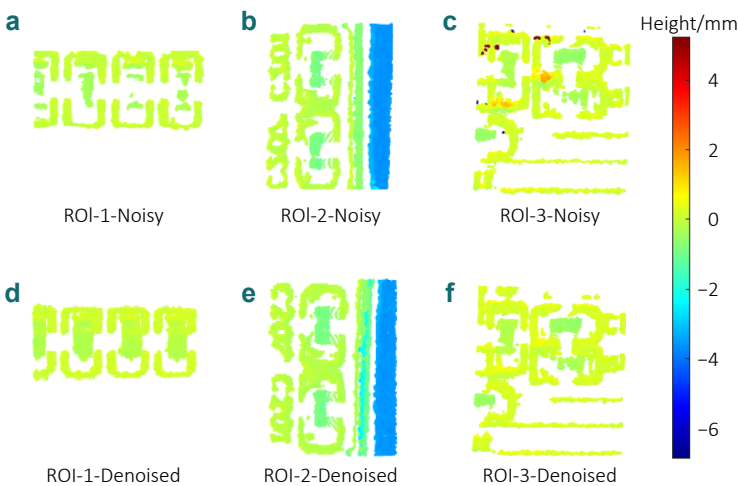


图 16 去噪前后局部点云结构对比图。(a) ROI-1 (去噪前); (b) ROI-2 (去噪前); (c) ROI-3 (去噪前); (d) ROI-1 (去噪后); (e) ROI-2 (去噪后); (f) ROI-3 (去噪后)

Fig. 16 Comparison of local point-cloud structures before and after denoising. (a) ROI-1 (Noisy); (b) ROI-2 (Noisy); (c) ROI-3 (Noisy); (d) ROI-1 (Denoised); (e) ROI-2 (Denoised); (f) ROI-3 (Denoised)

看出, 这些局部结构的视觉改善与表 9 中边缘区域 RMSE/P95 的下降及法向角误差的减小趋势保持一致, 进一步验证了去噪后几何细节得到了有效保留。

综上, 所提出的多模态 DnCNN 去噪方法, 在全局精度、误差稳健性以及局部几何细节保持方面均表现出显著优势。不仅提升了相位图的连续性与稳定性, 同时有效增强了三维点云的密度、精度及结构还原质量。为高精度结构光测量及缺陷检测提供了稳健、可信的技术支撑, 体现了深度学习去噪在结构光三维重建领域的工程实用价值与进一步的应用潜力。

5 结 论

本文提出了一种融合多模态输入与结构保持机制的结构光条纹图像深度去噪方法。该方法基于改进的 DnCNN 框架, 采用原始含噪条纹图与串联滤波图的双通道输入, 并引入 SE 通道注意力机制和混合损失函数, 实现了对复合噪声的有效抑制与结构细节的自适应保持。

针对结构光条纹图像的成像特性, 本文构建了包含规则几何体、石膏雕塑和 PCB 电路板在内的真实数据集, 并在多场景下开展训练与验证。实验结果表明, 所提方法在 PSNR、SSIM 与 RMSE 等图像域指标上均取得较为明显的提升, 在纹理保持、条纹连续性和边缘结构方面表现稳定, 具有良好的跨场景泛化能力。消融实验进一步验证了三项关键设计的有效性: 双通道输入增强了模型对复合噪声的鲁棒性; SE 通道注意力模块突出条纹方向性与边缘区域特征; 混合损失函数在保证收敛稳定的同时提高了结构保真度。

在三维重建实验中, 结合三频四步相移与多频外差法, 并根据系统标定参数通过三角测量完成相位提取与三维重建。对比去噪前后的三维点云结果, 所提方法显著改善了点云的几何连续性与空间一致性, 有效减少了边界断裂、局部起伏及强反射伪点。C2M 距离、RMSE、95% 分位误差等量化指标均大幅下降, 且未出现结构圆滑化或边缘削弱现象, 表明去噪过程保持了关键几何特征。

未来工作将面向动态场景, 探索结合运动补偿、多物理先验及时空一致性约束的深度去噪模型, 以进一步提升方法在在线测量与复杂工业环境中的泛化能力与工程实用价值。

作者贡献

张立平: 研究资料的收集、分析, 论文写作; 胡晓峰, 潘飞文, 郭斌, 罗戟: 指导论文选题、论文修改。

利益冲突

所有作者声明无利益冲突

参考文献

1. Zuo C, Qian J M, Feng S J, et al. Deep learning in optical metrology: a review[J]. *Light Sci Appl*, 2022, 11(1): 39.
2. Zuo C, Feng S J, Huang L, et al. Phase shifting algorithms for fringe projection profilometry: a review[J]. *Opt Lasers Eng*, 2018, 109: 23–59.
3. Nguyen H, Wang Z Y. Accurate 3D shape reconstruction from single structured-light image via fringe-to-fringe network[J]. *Photonics*, 2021, 8(11): 459.
4. Zhang S. Absolute phase retrieval methods for digital fringe projection profilometry: a review[J]. *Opt Lasers Eng*, 2018, 107: 28–37.
5. Zhang Z. A flexible new technique for camera calibration[J]. *IEEE Trans Pattern Anal Mach Intell*, 2000, 22(11): 1330–1334.
6. Wijenayake U, Baek S H, Park S Y. An error correcting 3D scanning technique using dual pseudorandom arrays[C]//*Proceedings of the 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*, 2012: 517–523. <https://doi.org/10.1109/3DIMPVT.2012.68>.
7. Ren M Y, Cui J, Cai S A, et al. Speckle suppression in dynamic structured light via single-element interference[J]. *Opt Express*, 2025, 33(5): 11452–11461.
8. Wang J, Chen L, Wu Q Y. Wavefront acquirement of single interferogram using the virtual grating moiré fringe method with Gaussian filter in the spatial domain[J]. *Chin J Lasers*, 2012, 39(5): 0508005.
王军, 陈磊, 吴泉英. 利用空域滤波虚光栅叠栅法提取干涉图波面[J]. *中国激光*, 2012, 39(5): 0508005.
9. Dabov K, Foi A, Katkovnik V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering[J]. *IEEE Trans Image Process*, 2007, 16(8): 2080–2095.
10. Wang Y H, Li J R, Sun J F, et al. Frequency domain filtering for phase fringe patterns of digital speckle pattern interferometry[J]. *Chin Opt*, 2014, 7(3): 389–395.
王永红, 李骏睿, 孙建飞, 等. 散斑干涉相位条纹图的频域滤波处理[J]. *中国光学*, 2014, 7(3): 389–395.
11. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. *Proc IEEE*, 1998, 86(11): 2278–2324.
12. Chen Y J, Pock T. Trainable nonlinear reaction diffusion: a flexible framework for fast and effective image restoration[J]. *IEEE Trans Pattern Anal Mach Intell*, 2017, 39(6): 1256–1272.
13. Zhang K, Zuo W M, Chen Y J, et al. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising[J]. *IEEE Trans Image Process*, 2017, 26(7): 3142–3155.
14. Zhang Y L, Tian Y P, Kong Y, et al. Residual dense network for image super-resolution[C]//*Proceedings of the 2018 IEEE/CVF*

- Conference on Computer Vision and Pattern Recognition, 2018: 2472–2481. <https://doi.org/10.1109/CVPR.2018.00262>.
15. Zhang K, Li Y W, Zuo W M, et al. Plug-and-play image restoration with deep denoiser prior[J]. *IEEE Trans Pattern Anal Mach Intell*, 2022, **44**(10): 6360–6376.
 16. Jiang H Y, Dai M L, Su Z L, et al. An adaptive sine/cosine filtering algorithm based on speckle phase fringe orientation[J]. *Acta Optica Sinica*, 2017, **37**(9): 0910001.
蒋汉阳, 戴美玲, 苏志龙, 等. 基于散斑相位条纹方向的自适应正弦/余弦滤波[J]. *光学学报*, 2017, **37**(9): 0910001.
 17. Peng G Z, Chen W J. Fringe pattern inpainting based on convolutional neural network denoising regularization[J]. *Acta Optica Sinica*, 2020, **40**(18): 1810002.
彭广泽, 陈文静. 基于卷积神经网络去噪正则化的条纹图修复[J]. *光学学报*, 2020, **40**(18): 1810002.
 18. Li C C, Mao J J, Zhang X N. Neural network-based method for structured light phase-shifted fringe denoising and phase-to-height calibrations[J]. *Meas Sci Technol*, 2025, **36**(1): 015047.
 19. Zhang X, Cheng P, You Z S, et al. Phase retrieval from single-shot square wave fringe based on image denoising using deep learning[J]. *Appl Opt*, 2024, **63**(4): 1160–1169.
 20. Huang W J, Lin B J, Ding Y Q, et al. Non-line-of-sight visible light positioning system based on deep learning[J]. *Study Opt Commun*, 2024, (6): 230091.
黄伟杰, 林邦姜, 丁永棋, 等. 基于深度学习的非视距可见光定位系统[J]. *光通信研究*, 2024, (6): 230091.
 21. Liao H Z, Kong Y, Zhang H, et al. Two-step phase unwrapping based on swin-UNet-denoise and least squares method[J]. *Laser Technol*, 2024, **48**(5): 752–758.
廖后章, 孔勇, 张贺, 等. 基于 swin-UNet-denoise 和最小二乘法的两步相位解包裹[J]. *激光技术*, 2024, **48**(5): 752–758.
 22. Chen C R, Wang H Y, Zhao B Q, et al. Digital holographic phase unwrapping based on UMnet[J]. *Laser Technol*, 2023, **47**(1): 73–79.
陈翠茹, 王华英, 赵宝群, 等. 基于 UMnet 的数字全息相位解包裹[J]. *激光技术*, 2023, **47**(1): 73–79.
 23. Chen C, Gao N, Zhang Z H. Simple calibration method for dual-camera structured light system[J]. *J Eur Opt Soc Rapid Publ*, 2018, **14**(1): 23.
 24. Lai S S, Liu Y K, Yu X, et al. A high-accuracy nonlinear phase error compensation method[J]. *Opto-Electron Eng*, 2021, **48**(4): 200296.
赖姗姗, 刘元坤, 于馨, 等. 一种高精度的非线性相位误差校正方法[J]. *光电工程*, 2021, **48**(4): 200296.
 25. Chatterjee A, Govindu V M. Noise in structured-light stereo depth cameras: modeling and its applications[Z]. arXiv: 1505.01936, 2015. <https://doi.org/10.48550/arXiv.1505.01936>.
 26. Fan L, Wang Y J, Zhang H X, et al. High-accuracy 3D contour measurement by using the quaternion wavelet transform image denoising technique[J]. *Electronics*, 2022, **11**(12): 1807.
 27. Boyat A K, Joshi B K. A review paper: noise models in digital image processing[J]. *Sign Image Process Int J*, 2015, **6**(2): 63–75.
 28. Fu B, Zhao X Y, Song C M, et al. A salt and pepper noise image denoising method based on the generative classification[J]. *Multimed Tools Appl*, 2019, **78**(9): 12043–12053.
 29. Wu W C, Lv G N, Duan Y Y, et al. Dual convolutional neural network with attention for image blind denoising[J]. *Multimed Syst*, 2024, **30**(5): 263.
 30. Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]// *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: 7132–7141. <https://doi.org/10.1109/CVPR.2018.00745>.
 31. Li H Y, Wu H Y, Chen X W, et al. Towards boosting the channel attention in real image denoising: sub-band pyramid attention[Z]. arXiv: 2012.12481, 2020. <https://doi.org/10.48550/arXiv.2012.12481>.
 32. Malyugina A, Anantrasirichai N, Bull D. Wavelet-based topological loss for low-light image denoising[J]. *Sensors*, 2025, **25**(7): 2047.
 33. Zhang Y, Li D S, Law K L, et al. IDR: self-supervised image denoising via iterative data refinement[C]// *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022: 2088–2097. <https://doi.org/10.1109/CVPR52688.2022.00214>.
 34. Liu J M, Wu C H, Wang Y Z, et al. Learning raw image denoising with Bayer pattern unification and Bayer preserving augmentation[C]// *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2019: 2070–2077. <https://doi.org/10.1109/CVPRW.2019.00259>.
 35. Zhang J R, Zhang Y J, Chen B. Out-of-focus projector calibration method with distortion correction on the projection plane in the structured light three-dimensional measurement system[J]. *Sensors*, 2017, **17**(12): 2963.
 36. Huang B Y, Ozdemir S, Tang Y, et al. A single-shot-per-pose camera-projector calibration system for imperfect planar targets[C]// *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 2018: 15–20. <https://doi.org/10.1109/ISMAR-Adjunct.2018.00023>.
 37. Deng S C, Zhou L B, He X K. Improved three-frequency four-step phase-shifted surface structured light 3D reconstruction algorithm[J]. *Mod Mach Tool Automat Manuf Techn*, 2024, (8): 13–16, 21.
邓仕超, 周炉保, 何新凯. 改进的三频四步相移面结构光三维重建算法[J]. *组合机床与自动化加工技术*, 2024, (8): 13–16, 21.
 38. Liu X R. Three-dimensional profile measurement system based on phase measuring profilometry[D]. Nanjing: Southeast University, 2015: 32–54. <https://doi.org/10.7666/d.Y2921534>.
刘欣冉. 基于相位测量轮廓术的三维轮廓测量系统研究[D]. 南京: 东南大学, 2015: 32–54. <https://doi.org/10.7666/d.Y2921534>.
 39. Tian H N, Li S M. Objective evaluation method for image quality based on edge structure similarity[J]. *Acta Photonica Sinica*, 2013, **42**(1): 110–114.
田浩南, 李素梅. 基于边缘的 SSIM 图像质量客观评价方法[J]. *光子学报*, 2013, **42**(1): 110–114.
 40. Buades A, Coll B, Morel J M. A non-local algorithm for image denoising[C]// *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005: 60–65. <https://doi.org/10.1109/CVPR.2005.38>.
 41. Shao G F, Gao Y B, Zuo J Y, et al. An improved BM3D method for eDNA microarray image denoising[C]// *Proceedings of the 13th International Conference on Computer Science & Education (ICCSE)*, 2018: 1–6. <https://doi.org/10.1109/ICCSE.2018.8468760>.
 42. Yang D, Sun J. BM3D-Net: a convolutional neural network for transform-domain collaborative filtering[J]. *IEEE Sig Process Lett*,

- 2018, **25**(1): 55–59.
43. Dutta S, Basarab A, Georgeot B, et al. DIVA: deep unfolded network from quantum interactive patches for image restoration[J]. *Pattern Recogn*, 2024, **155**: 110676.
 44. Singh P, Diwakar M, Gupta R, et al. A method noise-based convolutional neural network technique for CT image denoising[J]. *Electronics*, 2022, **11**(21): 3535.
 45. Polo M E, Cuartero A, Felicísimo Á M. Study of uncertainty and repeatability in structured-light 3D scanners[Z]. arXiv: 1910.13199, 2019. <https://doi.org/10.48550/arXiv.1910.13199>.
 46. Chen M J, Li Q, Shimasaki K, et al. A novel dynamic light-section 3D reconstruction method for wide-range sensing[J]. *Sensors*, 2024, **24**(12): 3793.
 47. Zhang D Y, Lingamanaik S N, Chung H. Image-based 3D reconstruction for rail profile measurement[J]. *Proc Inst Mech Eng Part F J Rail Rap Trans*, 2023, **237**(3): 309–321.
 48. Gu F F, Song Z, Zhao Z L. Single-shot structured light sensor for 3D dense and dynamic reconstruction[J]. *Sensors*, 2020, **20**(4): 1094.

作者简介



张立平 (2000-), 女, 硕士研究生, 主要从事机器视觉、三维重建方面的研究。

E-mail: 17395713083@163.com



【通信作者】胡晓峰 (1986-), 男, 硕士, 高级实验师, 硕士生导师, 主要从事智能检测、精密测量技术研究。

E-mail: 12B0202094@cjlu.edu.cn

Multi-modal denoising method for structured light 3D reconstruction under composite noise

Zhang Liping¹, Hu Xiaofeng^{1*}, Pan Feiwen², Guo Bin^{1,2}, Luo Zai¹

¹ College of Metrology & Measurement Engineering, China Jiliang University, Hangzhou, Zhejiang 310018, China;

² Hangzhou Wolei Intelligent Technology Co., Ltd., Hangzhou, Zhejiang 310018, China

Abstract

Objective Structured light 3D reconstruction is widely used in high-precision metrology and industrial inspection because of its non-contact and high-resolution characteristics. In practical industrial environments, however, captured fringe patterns are often corrupted by composite noise originating from sensor electronics, surface reflectance variations, and ambient interference. Such noise degrades phase quality, introduces unwrapping errors, and limits the achievable reconstruction accuracy. To address these issues, this paper proposes a multi-modal deep denoising method that integrates a dual-channel input, a squeeze-and-excitation channel attention mechanism, and a hybrid loss function combining MSE, SSIM, and Laplacian edge constraints. The objective is to effectively suppress multi-source noise while preserving fringe continuity and edge structures, thereby enhancing phase retrieval stability, reducing unwrapping failures, and ultimately improving both image-domain quality and 3D reconstruction accuracy under complex industrial conditions.

Methods The proposed method employs a dual-channel residual convolutional neural network based on the DnCNN architecture. The input is constructed by concatenating the original noisy fringe pattern with a serially filtered image along the channel dimension. The serial filtering module first applies extreme value detection and selective median filtering to remove impulse noise, followed by two-dimensional Gaussian smoothing to suppress high-frequency random noise. Multi-scale convolutional layers with batch normalization and ReLU activation are then introduced to extract hierarchical fringe features from the dual-channel input, capturing both fine textures and coarse structures while preserving edge information.

Subsequently, squeeze-and-excitation (SE) channel attention modules are embedded after convolutional layers in each residual block to adaptively emphasize informative features. Each SE module performs global average pooling to compress spatial information, followed by a bottleneck mapping with ReLU and Sigmoid activation to generate channel-wise weights, which are then used to recalibrate the feature maps. A hybrid loss function is designed to balance global noise suppression and local structure preservation, combining mean squared error for luminance fidelity, structural similarity index for contrast and texture consistency, and a Laplacian-based edge loss to maintain

Foundation item: Project supported by the National Natural Science Foundation of China (U24A20134, 52475576), and the Zhejiang Provincial Key Research and Development Program Fund (2023C01061)

*Correspondence: Hu X F, E-mail: 12B0202094@cjlu.edu.cn

high-frequency details and boundary sharpness.

Experiments are conducted on a binocular structured light system using three-frequency four-step phase shifting and heterodyne phase unwrapping. The left camera coordinate system is defined as the world coordinate system, and 3D coordinates are obtained via linear triangulation with calibrated camera and projector parameters. Training and validation data are collected from three representative objects: standard step blocks (rule geometries with sharp edges), plaster busts (smooth curved surfaces), and printed circuit boards (high-reflectivity components and fine textures). Reference images are obtained through multi-frame averaging.

Results and Discussions Quantitative evaluations in the image domain show that the proposed method achieves a PSNR of 43.11 dB / SSIM of 0.9654 on the standard step block, 42.90 dB / 0.9627 on the plaster bust, and 42.69 dB / 0.9643 on the PCB. Compared with noisy fringe patterns, the average PSNR increases by approximately 7 dB, SSIM improves to about 0.96, and RMSE decreases by around 61.6%. Ablation studies confirm the contribution of each component: dual-channel input improves PSNR by approximately 1.05 dB over single-channel input; the SE attention module provides consistent gains in both PSNR and SSIM, especially in edge detail recovery; and the hybrid loss function achieves the best overall performance. Comparative experiments against traditional methods (NLM, BM3D) and learning-based methods (BM3D-Net, DIVA) demonstrate that the proposed method achieves the highest PSNR (42.90 dB), highest SSIM (0.9627), and lowest RMSE (0.0063) on the plaster bust dataset. Noise robustness analysis under varying Gaussian noise levels ($\sigma_g = 5$ to 20) shows that even at the strongest noise level ($\sigma_g = 20$), the proposed method maintains a PSNR of 27.6 dB and an SSIM of 0.8032, substantially outperforming the original DnCNN (25.6 dB, 0.7514) and noisy inputs (15.7 dB, 0.4057), thereby demonstrating a high noise tolerance.

In the phase domain, denoised fringe patterns produce wrapped and absolute phase maps with significantly fewer phase jumps and fringe discontinuities, particularly in edge regions and high-reflectivity areas. Phase gradient analysis shows that the denoised histogram is more concentrated near zero, with a marked reduction in high-gradient outliers. In the 3D reconstruction domain, the point-to-plane RMSE decreases by 47.1%, the 95% quantile error decreases by 42.4%, the maximum error decreases by 16.7%, and point cloud density increases by 43.6%. Connectivity analysis reveals that high-error regions transition from large clustered patches to sparse isolated spots, indicating effective error tail suppression and spatial homogenization without introducing new artifacts. Region-wise error analysis further shows that RMSE and 95% quantile error decrease in both flat and edge regions, while edge metrics such as boundary accuracy, boundary completeness, and normal angle errors all improve. Consequently, the proposed method preserves key structural and edge details while suppressing noise, confirming that no edge rounding or structural weakening occurs.

Conclusions This paper proposes a deep denoising method for structured light fringe patterns that integrates multi-modal input and structure-preserving mechanisms. Based on an improved DnCNN framework, the method employs a dual-channel input combining the original noisy fringe pattern with a serially filtered image, introduces a squeeze-and-excitation channel attention mechanism, and adopts a hybrid loss function. Experiments on standard step blocks, plaster busts, and printed circuit boards demonstrate significant improvements in PSNR, SSIM, and RMSE, while maintaining texture preservation and fringe continuity. In the 3D reconstruction domain, point cloud density, accuracy, and structural fidelity are substantially enhanced. The proposed method provides a reliable preprocessing strategy for high-precision industrial measurement and defect detection.

Keywords: structured light measurement; fringe denoising; multi-modal input; deep learning; 3D point cloud reconstruction

Zhang L P, Hu X F, Pan F W, et al. Multi-modal denoising method for structured light 3D reconstruction under composite noise[J]. *Opto-Electron Eng*, 2026, 53(4): 250272; DOI: [10.12086/oe.2026.250272](https://doi.org/10.12086/oe.2026.250272)

