

引用格式: ZHANG Boxia, FU Shuai, GAO Nan, et al. Modulation-Guided and Digital Twin-Driven 3D Measurement Method for Highly Reflective Surfaces[J]. Acta Photonica Sinica, 2026, 55(8):0812002

张伯夏,付帅,高楠,等. 基于调制引导与数字孪生驱动的高反光表面三维测量方法[J]. 光子学报, 2026, 55(8):0812002

基于调制引导与数字孪生驱动的高反光表面三维测量方法

张伯夏,付帅,高楠,孟召宗,张宗华

(河北工业大学 机械工程学院,天津 300130)

摘 要:针对结构光三维测量中金属表面高反射导致的条纹饱和、调制信息缺失及相位解算失真问题,提出一种基于数字孪生数据驱动的调制引导注意力网络的三维重建方法。首先,构建了面向高反光场景的数字孪生条纹投影系统,生成具有多样反射特性的高动态范围虚拟数据集,解决了真实场景下数据采集成本高、标注困难的局限性。其次,设计一种融合多尺度残差编码、全局特征感知及调制引导模块的调制引导注意力网络,将条纹梯度与调制幅值作为空间先验信息引入注意力机制,引导网络自适应地滤除高反光区域的有效干扰,并对缺失的条纹信息进行精确特征重构,从而仅利用单幅受干扰的条纹图像即可稳定预测出理想的正弦与余弦分量。实验结果表明,在 $198\text{mm} \times 152\text{mm}$ 的有效视场下,所提方法在处理高反光表面时效果良好,虚拟测试集平均绝对误差为 0.0083 mm ;在实际金属台阶测量中,最大绝对误差为 0.032 mm 。与现有的深度学习网络相比,该方法在重建精度上具有显著优势,且在大幅提升测量效率的前提下达到了与传统多重曝光法相当的精度水平,为高反射工业零件的高速、高精度三维测量提供了有效方案。

关键词:条纹投影;数字孪生;三维重建;结构光;相位提取

中图分类号: TH741

文献标识码: A

doi: 10.3788/gzxb20265508.0812002

0 引言

条纹投影轮廓术(Fringe Projection Profilometry, FPP)因具有非接触、高精度、高点云密度及系统成本低等优点,在工业检测、质量控制及缺陷识别等场景中得到广泛应用^[1-2]。然而,在采用传统的条纹投影技术对高反光物体进行测量时,反射区域的亮度过高,超出了相机传感器的动态范围,使得这些区域的灰度值达到饱和,从而影响了三维(Three-Dimensional, 3D)重建结果^[3-4]。

针对高动态范围表面引起的条纹退化问题,已有大量研究开展。SONG Zhan等^[5]提出了一种基于二值条纹的高动态范围结构光方法,通过计算相机响应函数并在梯度域压缩动态范围,成功恢复了高反光物体表面的三维形貌。CHEN Chao等^[6]提出了一种自适应条纹投影强度的调节策略。该策略首先投影一组最大光强的相移条纹和一幅均匀的背景图像,以此计算出被测表面的反射率分布,随后逐像素地调整后续投影条纹的最大强度。实验结果表明,该自适应调节机制不仅有效抑制了镜面反射带来的过曝现象,还大大简化了测量流程,提高了三维形貌重建的效率与准确性。BUDIANTO等^[7]提出了一种针对高光区域的单帧条纹图像修复技术。该研究通过高斯混合模型定位盲区,并创新性地引入了缺失条纹的几何走向来引导迭代正则化过程,成功填补了高光区域的条纹信息,提高了高动态范围物体三维形貌恢复的鲁棒性。韩鹤翔等^[8]提出了一种基于三通道并行的条纹投影强度预测与图像融合方法。该方法利用互补金属氧化物半导

基金项目:国家自然科学基金(U2341275),河北省教育厅科学研究项目(JZX2024021)

第一作者:张伯夏(1999—),男,硕士生,主要研究方向为光学三维测量。Email: 3058356577@qq.com

通讯作者:张宗华(1974—),男,教授,博士,主要研究方向为光学检测、三维数字成像和造型等。Email: zhzhang@hebut.edu.cn

收稿日期: 2026-04-16; **录用日期:** 2026-05-25

<http://www.photon.ac.cn>

体相机在多光通道下对投影条纹的不同响应,通过增加一张预投影图像来预测三个通道的最佳投影光强比例,从而分离并合成高动态范围条纹图。上述方法在一定程度上改善了高反射表面的测量问题,但多依赖多曝光采集、投影调节或成像策略,系统复杂度较高且测量效率受限。此外,其他学者提出使用偏振滤光片技术^[9-10]、多相机视角融合技术^[11]以及基于颜色不变量的测量方法^[12]处理这类问题。然而,此类方案通常需要引入额外的硬件辅助,使得系统架构更为臃肿,且同样难以摆脱对繁复算法或多帧图像采集的依赖,进一步制约了其在高效测量场景中的应用。

近年来,深度学习技术被广泛应用于高动态范围物体的三维测量中,以解决条纹过曝或信噪比低的问题。WANG Hao等^[13]提出了一种基于注意力引导的端到端相位计算网络,利用注意力模块引导特征选择,通过深度神经网络直接从单幅多曝光调制的条纹图中精确提取基频的正弦和余弦分量,进而恢复出绝对相位。DANG Guoqing等^[14]提出了一种深度学习辅助的复合偏振条纹投影方法,通过投影仪通道特性生成复合偏振条纹并结合网络求解相位,减少了偏振片的物理旋转操作。SHEN Mengmeng等^[15]构建了基于改进型U-Net的相位解调模型,结合二值条纹建立输入与理想相位之间的映射关系,以从饱和及低对比度条纹中提取相位信息。XI Dejun等^[16]则提出了一种基于卷积神经网络的条纹融合方法,将高、低两种曝光时间下的条纹图像直接映射为多步相移条纹,省去了最佳曝光时间的计算过程,进一步提升了测量效率。李默晶等^[17]构建了一种包含相位预测与同名点匹配的双分支网络,针对网络在同时拟合过曝区域与非过曝区域时容易产生的梯度冲突与性能震荡问题引入了Pareto最优化,并提出了一种动态更新任务优先级的多任务权值搜索策略。然而,上述基于深度学习的方法仍面临两方面的挑战:一方面,在处理复杂高反光物体时,其相位恢复的稳定性和最终的三维重建精度依然难以满足高精度的测量需求;另一方面,目前的研究中获取数据集的常规手段包括算法合成与实景采集。前者由于虚拟合成技术很难完美复现真实物理空间中的复杂光影变化和非线性材质反射,从而造成生成数据在真实度与多样性上存在局限;后者需要获取海量且带有高精度标注的高动态范围条纹数据,成本高昂,难以大规模覆盖多样化的工业场景。

综上所述,复杂反射条件下的条纹信息缺失与高精度三维重建问题仍未得到充分解决。一方面,现有深度学习模型在复杂反射条件下的鲁棒性仍有待提升。尽管已有研究引入了注意力机制来引导特征选择,但这些方法大多依赖纯数据驱动的隐式映射,缺乏对结构光条纹物理退化过程的显式建模。在面对极端高反光导致的物理信息缺失时,权重的生成往往缺乏可解释性与鲁棒性;另一方面,受限于上述数据集获取的困境,现有方法难以系统模拟复杂反射退化规律,从而限制了模型的跨场景泛化能力。为解决上述问题,本文基于三维图形引擎Blender构建真实FPP系统的数字孪生模型,生成具有可控性与可扩展性的高动态范围的条纹数据集。在此基础上,提出一种用于提取单幅条纹图像中正弦与余弦分量的调制引导注意力网络,该网络采用编码-解码结构,并引入多尺度特征建模与条纹梯度及局部极差协同表征的特征引导机制,利用条纹梯度与局部极差显式建模条纹波形的退化程度,增强神经网络对于高反光区域条纹信息重建能力,从而提升高动态场景下相位恢复的稳定性和精度,最后通过系统实验验证所提方法在工业零部件三维测量中的有效性。

1 测量原理与方法

1.1 三维测量流程

本文的三维测量方法整体流程包括网络训练与实际测量两个部分。首先,利用数字孪生系统生成的大规模虚拟条纹数据集对网络进行训练,使其学习从退化条纹中预测理想正弦与余弦分量的映射关系。在实际测量应用中,将不同频率对应的单次曝光采集的被测物体条纹图像逐个输入网络,直接输出对应频率的十二步相移图像对应的正弦分量与余弦分量。在此基础上,通过反正切函数求取包裹相位,并利用最佳三频方法^[18]进行相位展开获取连续的绝对相位,最终结合逐像素拟合法^[19]得到的系统标定参数计算出目标物体的三维点云数据,完成完整的三维形貌重建。整体的三维测量流程如图1所示。

1.2 条纹投影轮廓术

条纹投影轮廓术作为一种典型的非接触式光学三维测量方法,其系统通常由计算机、投影仪和相机组成。测量过程中,计算机生成标准正弦条纹图案并由投影仪投射至被测物体表面,相机从固定视角采集经

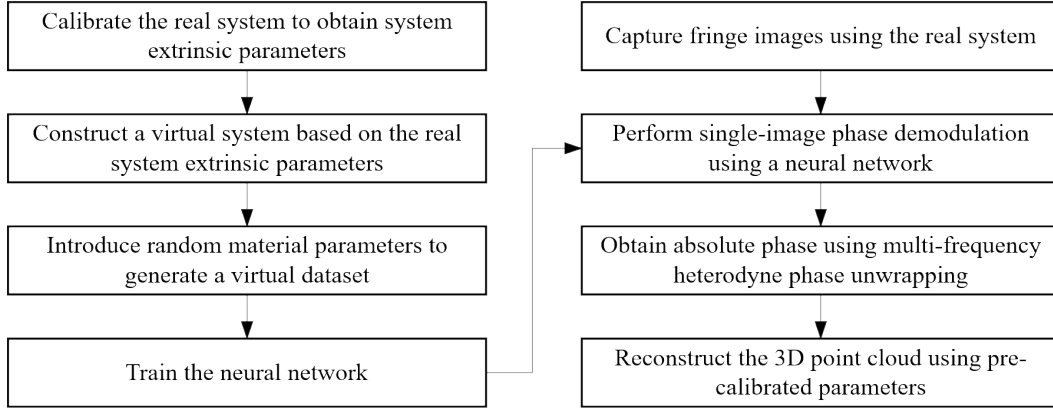


图1 三维测量流程图

Fig. 1 3D measurement process diagram

物体表面调制后的变形条纹图像。

针对FPP系统的相位解调,相移法是目前广泛应用的方法。通过投影多幅相移图像进行相位计算,能够有效保证测量精度。在 N 步相移法中,假设系统满足线性成像模型,在忽略高阶非线性和环境光影响的情况下,相机采集到的第 n 幅相移条纹图像在像素位置处的灰度强度可表示为:

$$I_n(x, y) = I_a(x, y) + I_b(x, y) \cos[\varphi(x, y) - 2\pi n/N], n \in [0, N-1] \quad (1)$$

其中, $I_a(x, y)$ 为背景强度, $I_b(x, y)$ 为调制幅度, N 为相移步数, $\varphi(x, y)$ 为包裹相位, 其取值范围为 $(-\pi, \pi]$ 。

对采集的相移图像进行解算以获得正弦与余弦分量,进而利用反正切函数提取包裹相位:

$$\varphi(x, y) = \tan^{-1} \left(\frac{\sum_{n=0}^{N-1} I_n(x, y) \sin(2\pi n/N)}{\sum_{n=0}^{N-1} I_n(x, y) \cos(2\pi n/N)} \right) = \tan^{-1} \left(\frac{S(x, y)}{C(x, y)} \right) \quad (2)$$

得到包裹相位后,利用相位展开算法进行相位展开获得绝对相位,最后结合系统标定参数获取物体表面的三维点云。

当被测物体表面存在高反光区域时,其局部反射率会急剧增大,呈现强烈的镜面反射特性。在曝光时间固定的情况下,相机传感器接收到的光强极易超过其动态范围的上限,导致相机响应不再满足线性假设,从而发生饱和截断现象。这种饱和截断破坏了条纹光强分布的正弦特性,在频谱中引入了大量的高次谐波成分。在相移算法中,这些高次谐波无法被完全滤除,会耦合到相位主值的解算中,产生周期性的相位误差。

1.3 网络结构

为实现对高反光条纹图像的精确相位恢复,文中提出了一种基于多尺度特征的调制引导注意力网络。该网络整体采用经典U-net网络的编码-解码结构,由多路径残差模块编码器(Residual Multi-scale Block, RMSB)、全局特征感知模块(Global Feature Perception Module, GFPM)以及带有调制引导模块(Modulation-Guided Module, MGM)的解码器三部分构成。该网络仅需输入单通道图像,即可有效输出该条纹图像中解调相位的正弦项和余弦项。其整体结构如图2所示。

作为网络的前端主干,编码器采用多级下采样结构逐层提取特征。每一级通过多路径卷积分支提取多尺度特征,不同于传统单分支卷积,该阶段将3个分别包含1个、2个和3个级联 3×3 卷积的并行分支与残差连接相融合,不仅逐步扩大了感受野,还保证了训练过程中的梯度稳定性,使网络在高反射区域具备更强的鲁棒性。

在编码器最深处引入全局特征感知模块,以解决深层特征历经多次下采样后易偏向局部统计特性、缺乏全局上下文约束的问题。其结构如图3所示。该模块首先通过多尺度空洞卷积结构进行特征融合,包含 1×1 卷积分支、全局平均池化分支,以及膨胀率为6和12的空洞分支,在不降低分辨率的前提下扩大有效感知范围;接着,结合头数为8的多头自注意力模块进行全局关系建模。通过计算空间位置间的相关性权重,建立长距离依赖关系。这既保证了条纹结构在大范围内的一致性,又为后续解码提供了稳定的语义基础。

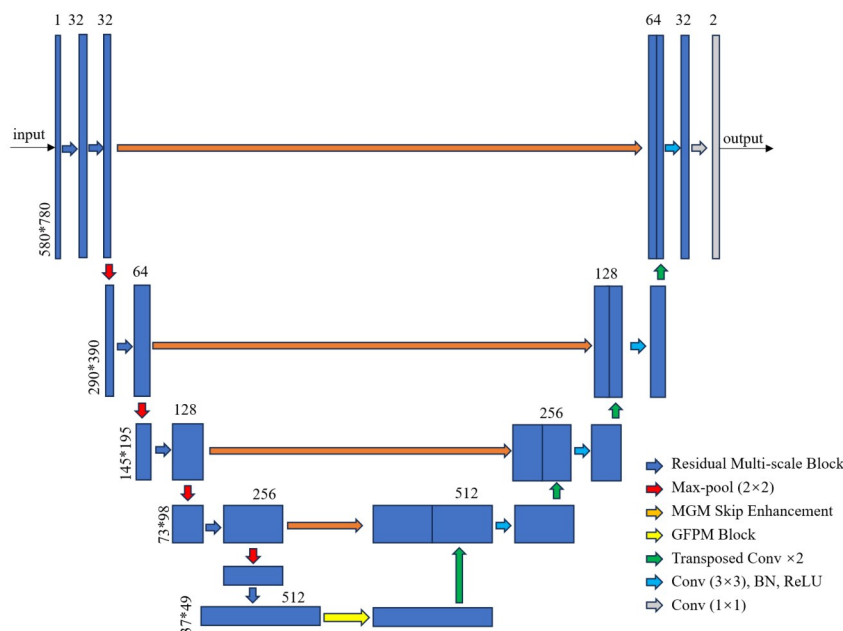


图2 调制引导注意力网络整体结构

Fig. 2 Overall architecture of the modulation-guided attention network

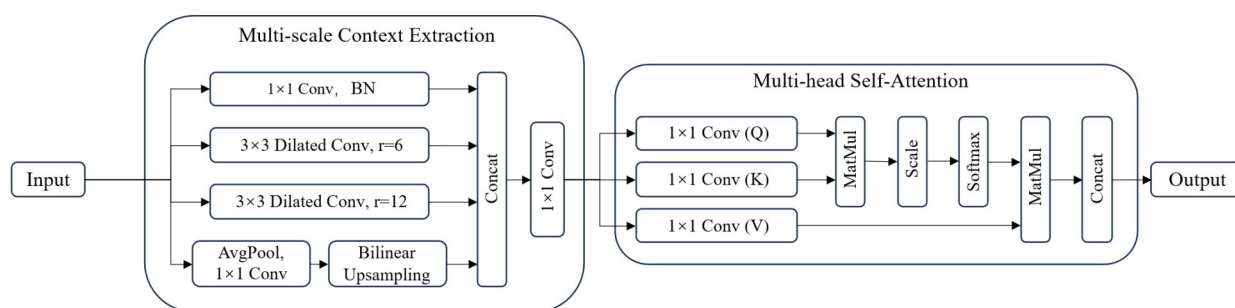


图3 全局特征感知模块结构示意图

Fig. 3 Architecture of the global feature perception module

在解码阶段,网络通过逐级上采样来恢复图像的空间分辨率,并在每一级上采样后,通过跳跃连接融合编码器中对应的浅层特征。然而,浅层特征通常包含较多的高频细节,极易受到高反射噪声和局部异常的干扰。为此,本文在跳跃连接路径中引入了调制引导模块,其结构如图4所示。调制引导模块旨在通过对条纹物理退化现象的定量评估,直接从输入特征中提取空间梯度与调制幅值,进而生成用于特征筛选的空间注意力权重图。具体而言,在强反射引起的过曝区域,相机传感器发生饱和,导致正弦条纹波形被严重截断。这种物理退化使得条纹的交变分量丢失,条纹的振幅会急剧衰减,其原有的规则正弦结构也会发生畸变甚至丢失。为了捕捉这种物理退化特征,计算输入特征沿通道维度的均值 X_m ,随后利用局部极大值池化与局部极小值池化的差值来估算条纹的动态范围。在结构光三维测量物理模型中,该动态范围等效于条纹的调制度,是衡量条纹信噪比与相位可靠性的核心指标。为此,提取出调制度特征图:

$$D = \text{Maxpool}(X_m) - \text{Minpool}(X_m) \quad (3)$$

同时,利用具有固定权重的 Sobel 算子提取 X_m 在水平方向的空间变化信息,计算出表示结构清晰度的梯度幅值:

$$G = \sum_i |\text{Conv}_{\text{Sobel}}(X_m)| \quad (4)$$

为了生成最终的先验掩膜,首先将通道均值 X_m 、梯度幅值 G 与调制振幅 D 沿通道维度进行拼接。随后,利用一个 7×7 的卷积层提取融合特征,并通过 Sigmoid 函数 σ 将其归一化,生成取值在 0 到 1 之间的空间注

意力权重图。该过程可表示为:

$$M = \sigma(\text{Conv}_{7 \times 7}([X_m, G, D])) \quad (5)$$

在空间注意力权重图 M 中,接近 0 的值指示该区域存在严重的过曝高反光退化或者为背景区域,而接近 1 的值则表明该区域拥有完整的条纹结构。这一机制将传统条纹分析中‘以调制度作为可靠性引导’的物理先验融入网络之中。借助这一空间引导分布,MGM 能够引导网络在特征传递过程中,自适应地抑制并滤除高反光区域的有效干扰。这一机制有效阻断了不可靠信息向解码阶段的渗透,从而指导解码网络更加精准地重构出缺失的条纹与相位信息。

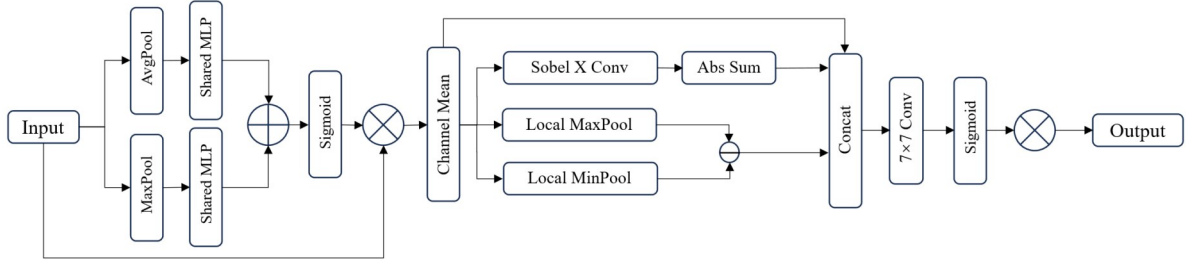


图 4 调制引导模块结构示意图

Fig. 4 Architecture of the modulation-guided module

最后,经过物理先验引导的浅层特征与上采样后的深层特征融合,并采用残差卷积模块进一步整合以抑制冗余响应。已有研究表明^[13],在神经网络任务中,输出对应相位的余弦分量与正弦分量测量精度往往优于预测深度图或预测绝对相位图,因此网络输出被设计为两个通道,分别对应相位的余弦分量与正弦分量,在解码末端,通过 1×1 卷积将多通道特征映射至目标输出空间,生成双通道结果,为后续的物理量计算提供高精度的估计。

为保证余弦和正弦分量的数值精度,损失函数被定义为预测值与真实标签之间的逐像素回归误差。本文采用 L1 损失作为损失函数,其具体形式为:

$$L(\theta) = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H (|S(x, y) - S_{gt}(x, y)| + |C(x, y) - C_{gt}(x, y)|) \quad (6)$$

其中, W 与 H 表示输出特征图的空间分辨率, θ 表示网络参数。

1.4 数据集生成

为了解决真实结构光条纹数据采集成本高、真值标注困难以及场景覆盖有限的问题,本文提出了一种基于数字孪生的虚拟条纹数据集构建方法。该方法的核心在于物理测量系统的精确虚拟化。首先,通过对真实条纹投影系统进行标定获取相机与投影仪的内外参数,并在三维图形引擎 Blender 中重建了如图 5 所示对应的虚拟系统,严格保证了虚拟空间与真实物理空间在几何成像关系上的一致性。随后,通过在虚拟环境中引入多样化的三维模型并随机调整其空间位姿,利用虚拟投影仪投射正弦条纹,并由虚拟相机同步采集受物体形貌调制的变形条纹图像,以此完成训练数据的采集过程。

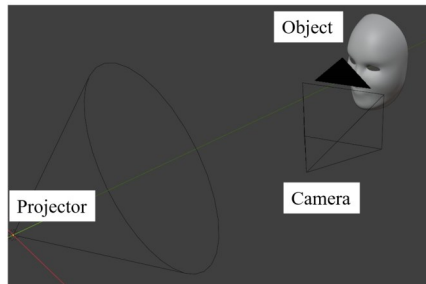


图 5 虚拟测量系统结构

Fig. 5 Architecture of the virtual 3D measurement system

为增强数据多样性与模型泛化能力,在数据生成过程中对物体材质、光照条件等参数进行随机扰动,具体参数如表1所示。针对高反光表面标签获取问题,本文使用高反光材质生成输入条纹图像以模拟真实测量,随后将同一模型替换为漫反射材质,并通过调节投影光强使其达到无过曝状态下的最高亮度,以最大化条纹信噪比,进而通过四步相移法计算出高精度的正弦与余弦分量作为理想标签。最终构建的虚拟数据集为深度学习模型训练提供了稳定且多场景的数据支撑。

表1 数据集生成过程中的随机参数与范围

Table 1 Random parameters and ranges for dataset generation

Parameter	Range
Fringe frequency	[50–90]
Power of projector/W	[2,30]
Metallicity	[0,0.8]
Roughness	[0.2,1.0]

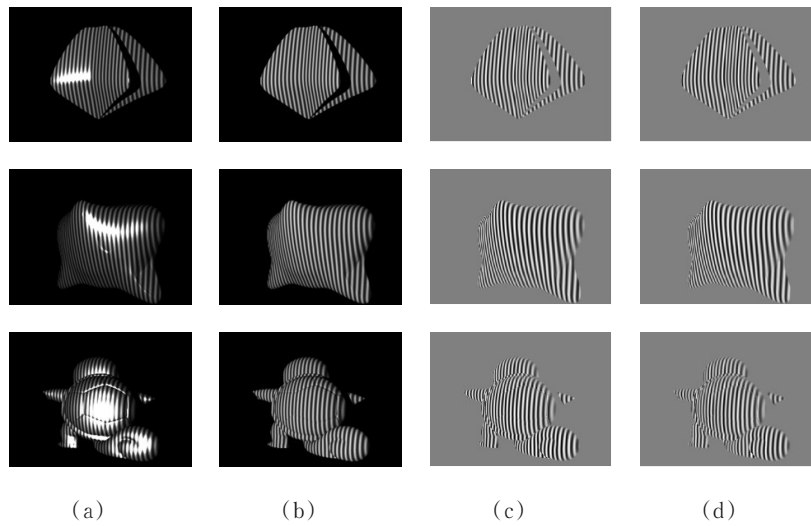


图6 数据集中的部分数据。(a)退化条纹;(b)理想条纹;(c)正弦项标签;(d)余弦项标签

Fig. 6 Partial data from the dataset. (a) Degraded fringe pattern; (b) Ideal fringe pattern; (c) Sine component label; (d) Co-sine component label

图6展示了在不同参数组合条件下渲染得到的条纹图像及其对应标签。其中,图6(a)展示了在高反光材质条件下的退化条纹图像,条纹的高反射区域由于过曝而出现明显的失真;图6(b)展示了漫反射材质条件下生成的清晰、均匀的正常条纹图像;图6(c)与(d)为正弦项标签和余弦项标签,表示相应四步理想条纹图像计算得到的正弦分量与余弦分量。这些图像与标签对准确表征了退化条纹与理想相位分量之间的映射关系,为后续模型训练和验证提供了必要的数据支持。

2 实验分析

本文基于Python语言与PyTorch深度学习框架实现了所提出的网络模型。实验在配备i7-8086K处理器和RTX2080Ti显卡的工作站环境下完成。构建的虚拟数据集共包含2 000张条纹图像,并按照8:2的比例划分为训练集和测试集。网络训练过程中采用Adam优化器,批大小为15,初始学习率设为0.001,并在训练过程中逐步衰减。模型共训练200个迭代轮次,训练过程中使用的损失函数如式(6)所示。

2.1 虚拟系统实验

文中选取测试集中如图7所示的两个具有代表性的场景进行实验评估本文方法。首先对相位恢复效果进行测试,结果如图8所示。图8(a1)–(a3)与(d1)–(d3)分别展示了传统三频四步法与本文方法得到的正弦项结果及对比;图8(b1)–(b3)与(e1)–(e3)为对应的余弦项结果及对比;图8(c1)–(c3)与(f1)–(f3)为

对应的包裹相位结果及对比。需要特别说明的是,为便于在统一尺度下进行对比,图中展示的正弦项与余弦项均为剔除背景光且经调制度归一化后的相对分量,属于无量纲数值。从结果与曲线中可以看出,传统三频四步法在面临高反光区域时,其正弦项与余弦项均出现了明显的畸变,导致最终的包裹相位发生严重的跳变与误差;而本文方法能够有效抑制高反光带来的饱和效应,准确地恢复出平滑的正余弦分量,从而稳定且高精度地重建出完整的包裹相位分布。

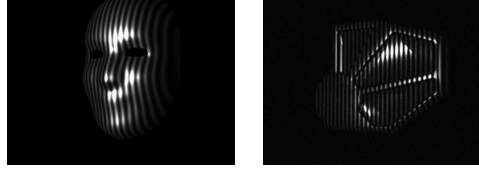


图7 两种典型高反光场景的原始条纹图。

Fig. 7 Original fringe patterns of two typical highly reflective scenes.

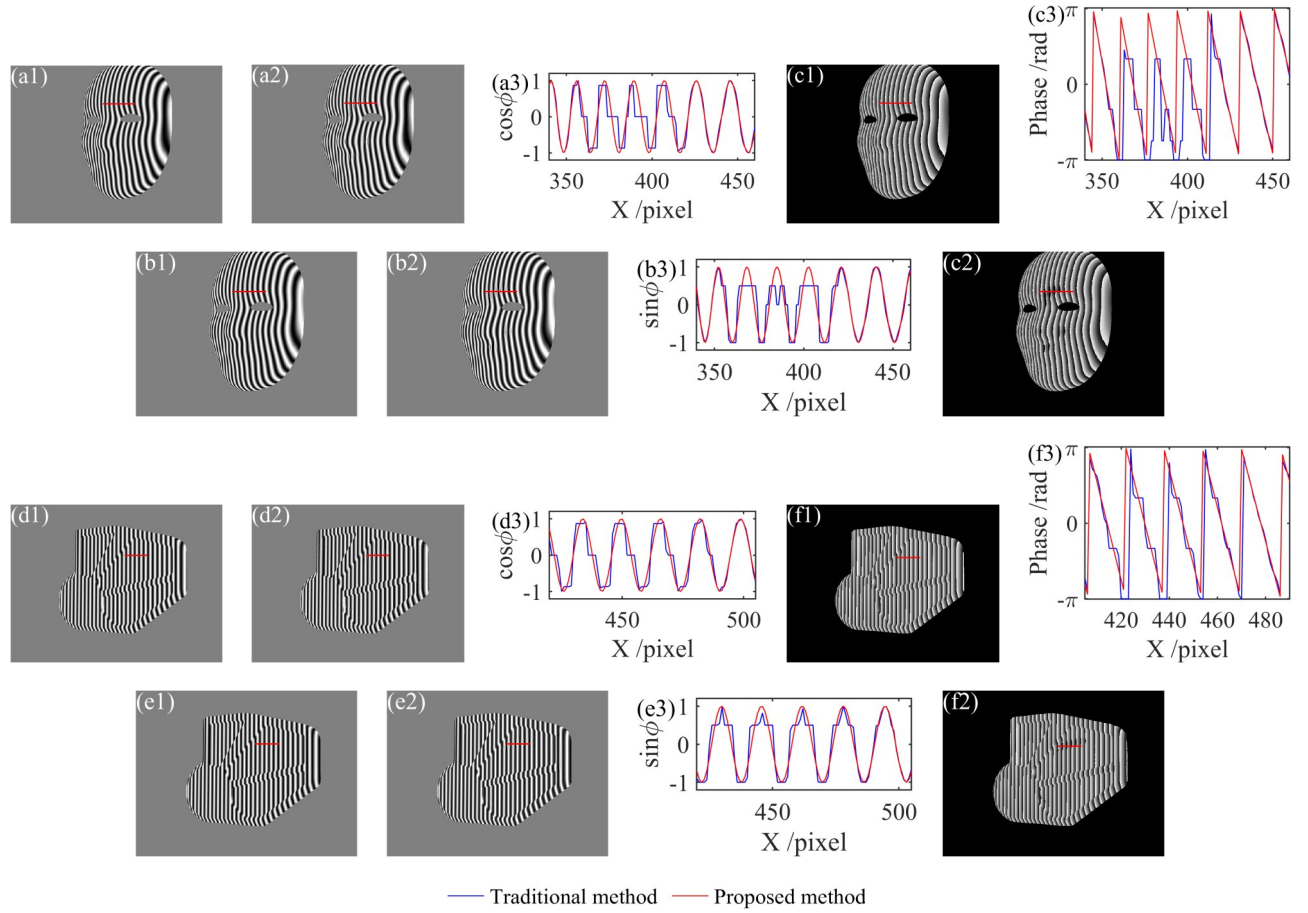


图8 不同高反光场景下余弦项、正弦项及截面包裹相位的恢复结果对比。(a1) - (a3)、(d1) - (d3)传统三频四步法与本文方法得到的余弦项结果及对比;(b1) - (b3)、(e1) - (e3)对应得到的正弦项结果及对比;(c1) - (c3)、(f1) - (f3)包裹相位结果及截面图对比。

Fig. 8 Comparison of recovered cosine components, sine components, and cross-sectional wrapped phases under different highly reflective scenes. (a1) - (a3) and (d1) - (d3) Recovered cosine components obtained by the traditional three-frequency four-step method and the proposed method; (b1) - (b3) and (e1) - (e3) Corresponding recovered sine components; (c1) - (c3) and (f1) - (f3) Comparison of the wrapped phases and their cross-sections.

为了验证所提方法的有效性,本文进一步对比了三维重建结果,如图9所示。可以明显看出,由于传统方法在局部高亮区域相位解算失败,导致图9(a)与(d)中的高反光区域存在大面积的点云缺失。相比之下,

图9(b)与(e)中本文方法的三维重建结果在点云完整度上有了显著提升,充分保留了复杂表面的形貌特征。图9(c)与(f)展示了本文方法重建结果与真值比较的绝对误差分布,其中面具场景与机械零件场景的整体平均绝对误差(Mean Absolute Error, MAE)分别为0.0054mm与0.0126mm。实验结果充分表明,所提方法在面临复杂高反光表面时具有出色的相位恢复能力与高精度的三维重建性能。

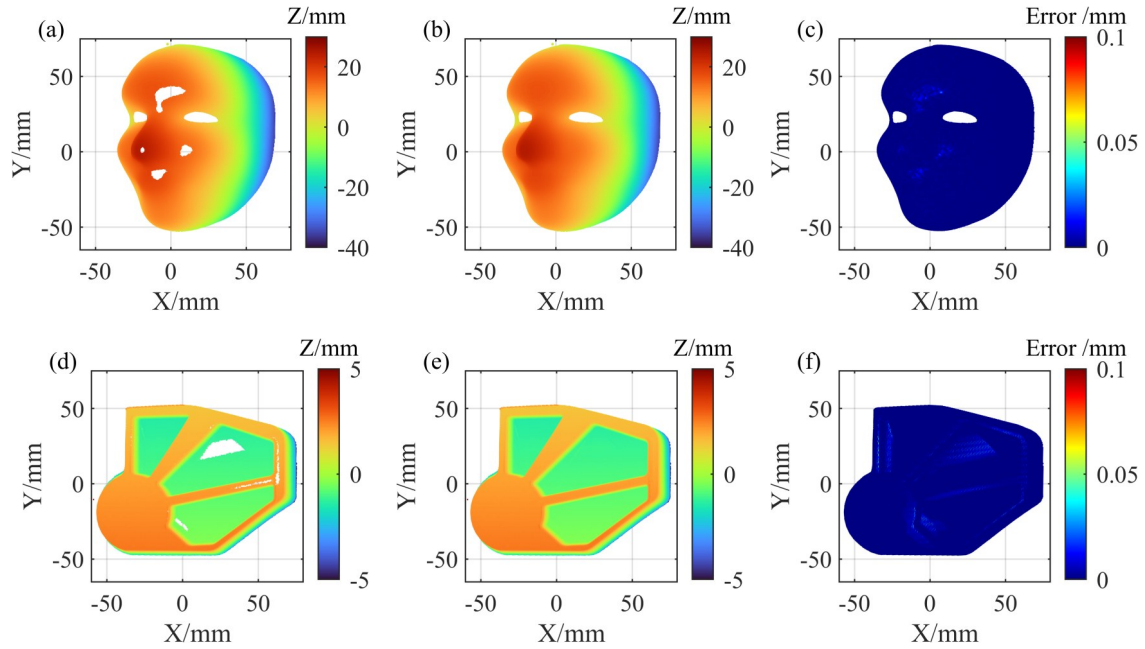


图9 不同场景下的三维重建结果。(a)(d)传统三频四步法重建结果;(b)(e)本文方法重建结果;(c)(f)本文方法的重建误差分布

Fig. 9 Comparison of 3D reconstruction results across different scenes. (a) (d) 3D reconstruction results using the traditional three-frequency four-step method; (b) (e) 3D reconstruction results using the proposed method; (c) (f) 3D reconstruction error distributions of the proposed method.

为了验证所提模型中各关键模块对整体性能的贡献,以标准U-Net结构作为基础模型,对所提出的网络进行系统性的消融实验。首先,以基础编码-解码结构作为基线模型,并逐步引入多尺度残差编码模块(RMSB)、全局特征感知模块(GFPM)以及调制引导模块(MGM),构建不同版本的网络结构,随后将上述模块完整集成,形成最终的模型。表2给出了不同模型版本在测试集的100个场景下得到的三维重建评估结果,主要包括峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似性(Structural Similarity Index, SSIM)以及平均绝对误差(MAE)。其中,PSNR与SSIM用于评估网络预测的深度图与虚拟真值的相似程度;MAE则用于量化最终三维形貌重建的绝对几何误差。为了进一步定量验证调制引导模块对高反光区域退化条纹的处理能力,在全局评价指标的基础上额外引入了饱和区域平均绝对误差(Saturated Region Mean Absolute Error, SR-MAE)。该指标通过掩膜提取原条纹图像中灰度值为255的饱和像素区域,并单独计算该区域内的三维重建绝对误差。

从表2结果可以看出,基础U-Net模型在高反光退化场景下的预测性能相对有限。随着多尺度残差编码模块(RMSB)的引入,模型对条纹结构的分析能力得到明显提升,其PSNR提升至27.98 dB;进一步加入

表2 消融实验结果
Table 2 Results of ablation experiments

Model	MAE /mm	SSIM	PSNR/dB	SR-MAE /mm
U-Net	0.1339	0.8721	22.85	0.3541
U-Net + RMSB	0.0218	0.9236	27.98	0.1885
U-Net + RMSB + GFPM	0.0118	0.9865	34.56	0.0752
Proposed (Full)	0.0083	0.9943	38.42	0.0256

全局特征感知模块(GFPM)后,网络在大尺度退化区域的全局上下文建模能力增强,从而提升了三维重建的稳定性,使其SSIM达到0.9865。最后,当在跳跃连接中引入调制引导模块(MGM)并构成完整模型后,各项评价指标均达到最优水平。此时,PSNR与SSIM进一步提升,整体MAE降至0.0083 mm,这种整体性能的提升,直接得益于MGM对高反光区域的有效处理:该模块通过对条纹振幅和空间梯度的自适应引导,使得极易出现异常的饱和区域局部误差(SR-MAE)从0.0752 mm减少至0.0256 mm。这充分证明MGM成功阻断了局部高光噪声对全局特征的污染,通过弥补局部短板,实质性地驱动了三维测量整体性能的提升。上述结果充分表明,本文提出的网络模型能够有效解决高反光带来的局部失真问题,提升三维重建的精度。

为探究数据集规模对模型泛化能力及相位恢复精度的影响,本文在同等网络设置下,使用不同规模的数据集对网络进行重新训练并在验证集中评估网络性能。如图10所示,当数据量低于1000张时,网络因未充分学习多样性特征,在未见过的场景下误差较高;随着数据量增加,误差显著下降,并在达到2000张左右时趋于平稳收敛。这表明2000张的数据集规模已涵盖足够丰富的表面特征,使网络能够充分学习并稳定收敛。

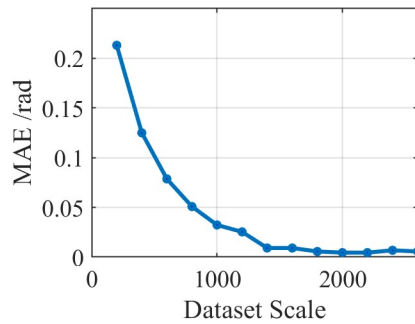


图10 不同数据集规模下的相位平均绝对误差曲线

Fig. 10 Phase mean absolute error curve under different dataset scales

2.2 实际系统实验

为验证所提方法在高反光场景下的有效性,本文搭建了光学三维测量实验系统,并基于标定所得的系统外参构建虚拟数据集并完成网络训练。该系统采用VisionFly6500投影仪,其分辨率为 1920×1080 。图像采集设备为eco415MVGE相机,分辨率为 780×580 ,配套使用焦距为8mm的定焦镜头,系统的有效视场为 $198\text{mm} \times 152\text{mm}$ 。所搭建的系统结构如图11所示。

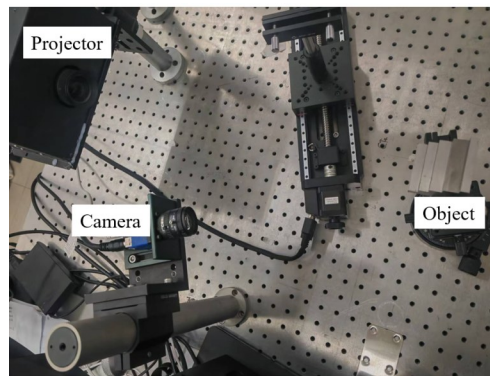


图11 三维测量系统结构

Fig. 11 Architecture of the 3D measurement system

为保证虚拟系统与真实系统在几何成像关系上的一致性,本文引入了虚实系统的标定参数验证环节。首先,在真实物理环境中使用 12×11 的标准棋盘格标定板(边长为15mm)采集多视角图像,并提取真实相机的内参矩阵与畸变系数。随后,通过模拟真实采集过程中的空间位姿变化,渲染输出同等分辨率的虚拟标定图像,并采用相同的算法流程求解虚拟相机的对应参数。

表 3 标定实验结果
Table 3 Results of calibration experiments

	Parameter	Real system	Virtual system
Camera	Focal length (f_x, f_y)/pixel	1009.49, 1009.59	1013.42, 1012.16
	Principal point (c_x, c_y)/pixel	395.93, 304.00	395.85, 304.74
	Distortion coefficients (k_1, k_2, p_1, p_2)	-0.1015, 0.1966 , 0.0025, 0.0000	-0.0986, 0.1876 , 0.0000, 0.0000
	Reprojection error/pixel	0.12, 0.16	0.08, 0.10
Projector	Focal length (f_x, f_y)/pixel	2958.79, 2959.21	2962.74, 2962.10
	Principal point (c_x, c_y)/pixel	1050.16, 554.25	1050.70, 552.14
	Distortion coefficients (k_1, k_2, p_1, p_2)	-0.0751, 0.3602, , 0.0036, 0.0000	-0.0735, 0.3614, , 0.0021, 0.0000
	Reprojection error/pixel	0.46, 0.45	0.24, 0.22
System parameters	Translation vector (T_x, T_y, T_z)/mm	94.78, 26.70, 287.62	93.07, 24.56, 290.68
	Rotation Euler angles(R_x, R_y, R_z)/°	1.60, -24.20, 0.04	2.06, -23.75, -0.51
	Baseline distance /mm	304.74	306.20

表 3 给出了真实系统与虚拟系统的标定结果。可以看出,两者在相机与投影仪的关键内参上具有较高一致性。以相机焦距参数为例,真实系统与虚拟系统的相对偏差均小于 0.4%,主点坐标偏差控制在 1 像素以内;对于投影仪参数,其焦距误差同样小于 0.2%,主点偏差约为 2pixel。在畸变参数方面,两套系统的径向畸变系数变化趋势保持一致,表明通过在虚拟场景中构建畸变节点模拟镜头的非线性成像过程,数字孪生模型能够较为准确地再现真实光学系统中的非理想成像特性。进一步分析系统外参可知,虚拟系统与真实系统的旋转欧拉角偏差均控制在 0.6° 以内,基线距离误差仅为 1.46mm。于此同时上述结果验证了数字孪生系统与真实物理系统在成像模型和空间几何关系上的高度一致性,能够对真实测量场景进行有效模拟与还原。

为验证本文方法对高反光金属工件的三维重建性能,实验分别采用 DC-UNet^[20]、HDRSL-Net^[21]、文中提出的方法以及作为参考方法的多重曝光法(Multiple exposure method, MEM)进行对比。为了保证对比的公平性,所有网络均在相同的数据集规模及训练参数下达到收敛。实验选取了三个典型高反光物体作为测量对象,如图 12 所示,图 12(a)-(c)展示了不同几何特征的高反光金属工件实物,包括法兰底座、金属圆盖与电机支架;图 12(d)-(f)则呈现了相机获取的不同零件的变形条纹图。

首先,本文对比了三种金属零件实物在不同深度学习方法下得到的正弦图与余弦图恢复效果。结果如图 13 所示。图 13(a1)-(a3)、(b1)-(b3)与(c1)-(c3)分别展示了三种金属零件实物采用 DC-UNet、HDRSL-Net 与本文方法得到的余弦项结果;图 13(a5)-(a7)、(b5)-(b7)与(c5)-(c7)则分别为对应方法得到的正弦项结果。图 13(a4)、(b4)、(c4)和图 13(a8)、(b8)、(c8)分别提取了不同方法以及参考方法在红色实线位置的余弦项与正弦项截面数据进行对比。从截面对比曲线中可以看出,在处理具有复杂反射特性的金属表面时,尽管各方法在整体走势上均基本实现了正余弦波形的恢复,但 DC-UNet 与 HDRSL-Net 方法重建的曲线依然存在一定的噪声波动与局部失真。相比之下,本文方法能够更有效地抑制金属表面高反光的干扰,呈现出更为平滑且逼近参考方法的波形。

图 14 展示了不同方法在高反光金属工件上的包裹相位提取效果。其中图 14(a1)-(a3)展示了法兰底座分别采用 DC-UNet、HDRSL-Net 与本文方法恢复得到的包裹相位结果;图 14(b1)-(b3)和图 14(c1)-(c3)分别给出了金属圆盖和电机支架在不同方法下的对应包裹相位分布。图 14(a4)、(b4)和(c4)提取了红色实线位置的截面包裹相位曲线。从结果可以看出,现有方法虽然能够在一定程度上恢复高反光区域的相位信息,但仍存在不同程度的相位偏移与波动。相比之下,本文方法恢复得到的包裹相位与参考趋势更加一致,曲线连续性和平滑性更优,能够有效抑制高反光干扰带来的局部畸变,提升复杂真实场景下的相位恢复精度与稳定性。

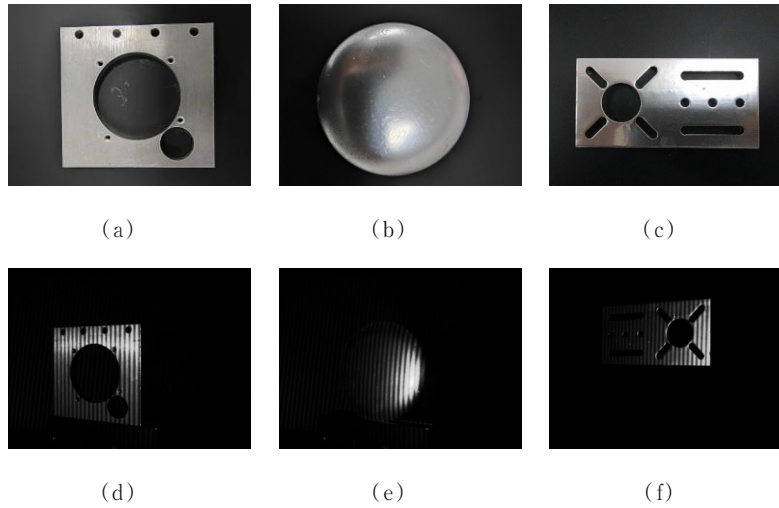


图12 物体实物图与变形条纹图。(a)-(c)三种金属零件实物图;(d)-(f)对应的变形条纹图

Fig. 12 Photographs and deformed fringe patterns of the objects. (a)-(c) Physical images of the three metal parts; (d)-(f) Corresponding deformed fringe patterns

在完成包裹相位恢复效果验证后,进一步基于重建得到的绝对相位对真实高反光金属工件进行三维测量,以评估不同方法在复杂场景下的三维重建性能。图15展示了法兰底座、金属圆盖与电机支架三类工件的重建结果。其中,图15(a)-(c)、(d)-(f)以及(g)-(i)分别对应法兰底座、金属圆盖与电机支架在DC-UNet、HDRSL-Net以及本文方法下的三维重建结果。

受金属表面强烈的镜面反射影响,高亮区域条纹调制度显著降低。DC-UNet在处理高亮区域时出现了明显的失真;HDRSL-Net虽在一定程度上抑制了反光干扰,但在复杂反射环境下的整体重建精度仍难以满足精密测量需求。相比之下,本文网络通过全局特征感知模块确保条纹在大范围内的结构一致性,并利用动态范围估计与梯度信息组合的调制信息自适应地抑制噪声干扰,从而提升了整体相位提取精度。实验结果显示,本文方法重构出的金属表面较为平滑,优于其他对比网络。

图15(j)-(l)展示了各工件在红色虚线位置处的深度剖面误差分布,其中参考真值由多重曝光法获取。由图可见,受高亮反光导致的调制度退化影响,各网络在过曝区域均出现了不同程度的误差提升。其中,DC-UNet在强反射核心区的误差曲线呈现剧烈的随机跳变,局部误差峰值显著升高;HDRSL-Net虽在一定程度上平抑了误差波动,但在反光剧烈区域仍存在明显的残余误差。相比之下,本文方法的误差曲线在零基准线附近保持了较小的波动范围,与参考值的一致性较高。上述实验结果表明,本文提出的调制引导注意力网络通过深度融合全局语义与局部调制信息,能够有效补偿高反光造成的物理信息缺失,在真实物理环境下展现出良好的泛化能力与测量精度。

为进一步量化评估各算法在高反光场景下的全局重建完整度与精度,本文引入了“有效点占比”评价指标。在此,有效点被定义为在工件目标区域内,既未发生数据缺失,又与参考真值深度误差小于0.1mm的重建点。表4展示了三种网络在不同金属工件上的有效点占比对比,受高光导致的条纹信息丢失与严重畸变影响,DC-UNet与HDRSL-Net在三个金属工件上的有效点占比平均仅为74.9%与87.02%。相比之下,本文方法凭借对高光区域浅层特征的有效加权与提取,大幅减少了缺失点与超差无效点,其法兰底座、金属圆盖与电机支架的有效点占比分别达到了99.32%、98.93%与99.67%。这一指标证明了本文网络在真实物理环境下对复杂高反光表面三维测量的全局完整性与可靠性。

为了进一步定量评估所提方法在处理复杂高反光表面时的三维重建精度,本文选取一个标准金属台阶作为测试目标,如图16所示。图16(a)为金属台阶的实物样图,图16(b)为其对应的变形条纹图。可以观察到,受金属材质影响,该样本表面呈现出典型的高反光特性,根据所处位置的不同,部分阶梯区域出现了强烈的局部高光与条纹饱和现象,而部分区域则光强较小。为了提供绝对可靠的评价基准,本实验预先采用高精度三坐标测量机(Coordinate Measuring Machine, CMM)对该台阶进行了测量,并将获取的数据作为三

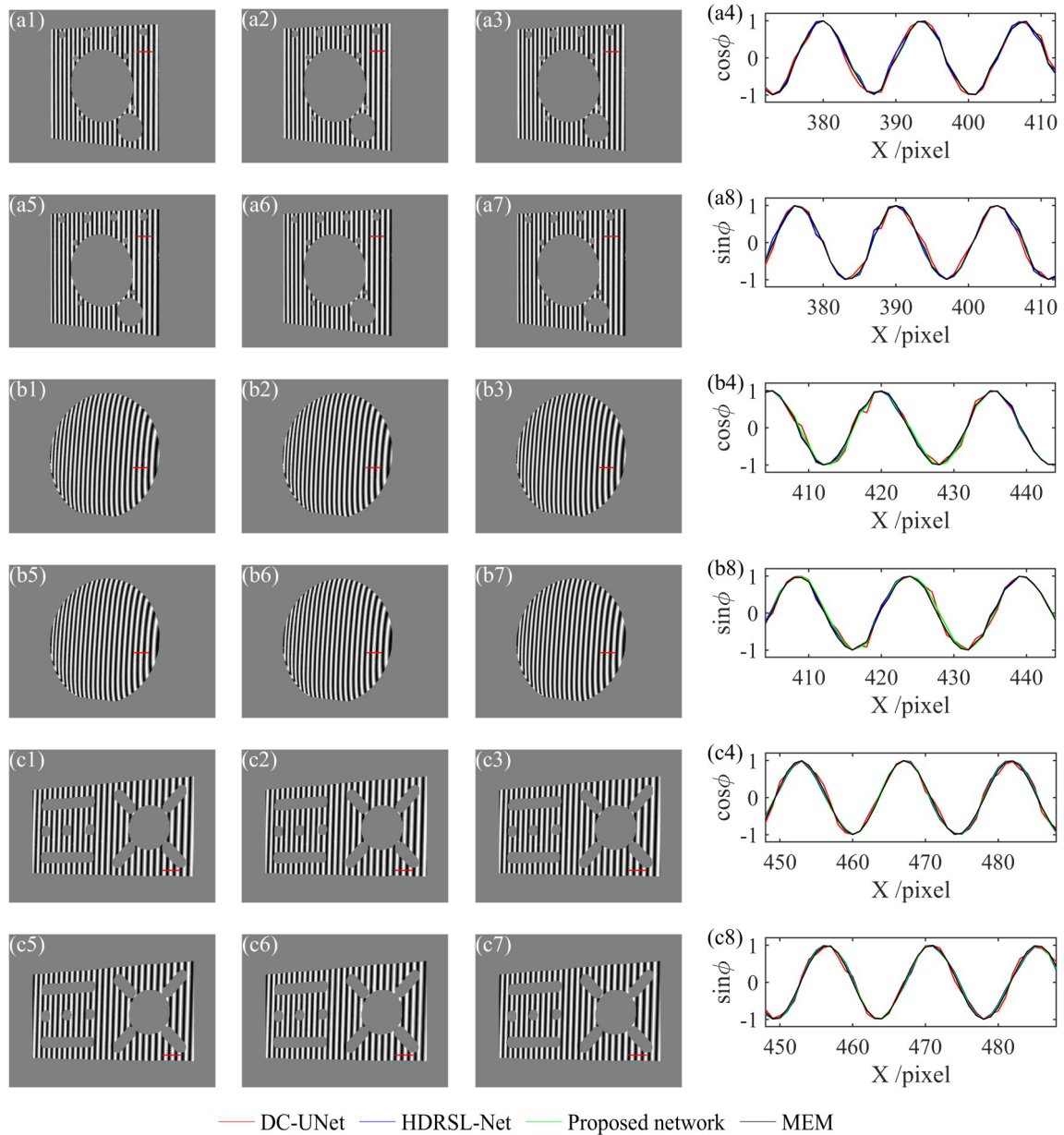


图13 不同方法重建的正弦与余弦项对比。(a1) - (a3)、(b1) - (b3)、(c1) - (c3)三种金属零件经DC-UNet、HDRSL-Net与本文方法重建的余弦项结果;(a5) - (a7)、(b5) - (b7)、(c5) - (c7)对应的正弦项结果;(a4)、(b4)、(c4)与(a8)、(b8)、(c8)分别为不同方法在红线处的余弦项与正弦项截面对比曲线。

Fig. 13 Comparison of reconstructed sine and cosine components by different methods. (a1) - (a3), (b1) - (b3), and (c1) - (c3) Cosine components of three metallic parts reconstructed by DC-UNet, HDRSL-Net, and the proposed method, respectively; (a5) - (a7), (b5) - (b7), and (c5) - (c7) Corresponding sine components; (a4), (b4), (c4) and (a8), (b8), (c8) Cross-sectional comparison plots of cosine and sine components along the red lines, respectively.

维形貌的真实值。基于此基准数据,本文随后将所提方法与其他现有的神经网络模型进行了定量的对比分析。各方法对金属台阶的三维重建结果如图17所示。从整体形貌上看,各方法均能基本恢复出台阶的物理结构。

为了进行定量分析,本文将重建获取的点云数据拟合为独立的台阶平面,并将相邻平面间的平均距离作为测量高度。以三坐标测量机获取的物理高度作为绝对真值,各方法的测量结果与高度误差详见表5。由表5可知,针对具有强反光特性的金属表面,本文方法的深度测量最大绝对误差为0.032 mm,优于其他深度学习模型,整体误差水平与多重曝光法基本持平。这一定量结果充分论证了本文方法的核心优势:能够在大幅减少图像采集数量的前提下实现与多重曝光法精度相当的高精度三维重建。

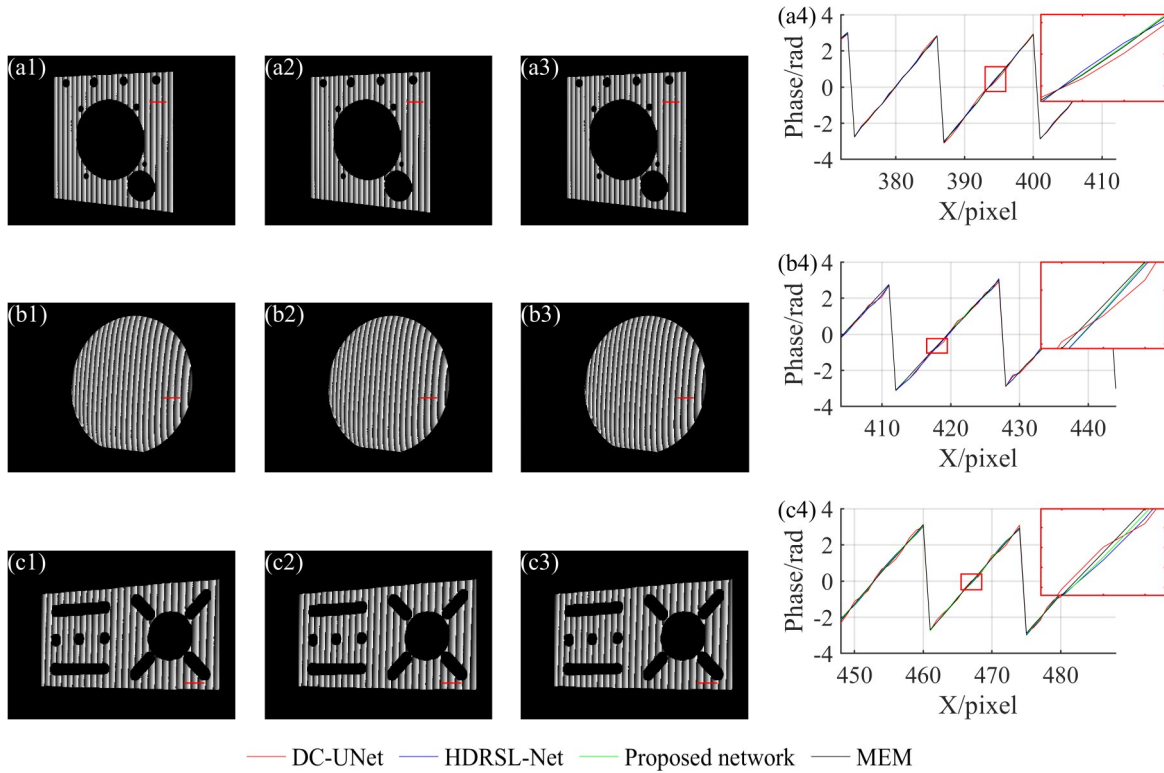


图 14 不同方法对金属工件的包裹相位对比。(a1)–(a3)法兰底座分别采用 DC-UNet、HDRSL-Net 与本文方法的包裹相位结果；(b1)–(b3)金属圆盖的对应包裹相位结果；(c1)–(c3)电机支架的对应包裹相位结果；(a4)、(b4)、(c4)为不同方法在红线处的包裹相位对比图

Fig. 14 Comparison of wrapped phases for metallic parts using different methods. (a1)–(a3) Wrapped phase results for the flange base using DC-UNet, HDRSL-Net, and the proposed method, respectively; (b1)–(b3) corresponding wrapped phase results for the metal cap; (c1)–(c3) corresponding wrapped phase results for the motor bracket; (a4), (b4), and (c4) comparisons of cross-sectional wrapped phases at the red lines.

在验证了针对高反光金属台阶的高精度重建能力后,为进一步评估本文方法在实际高速三维测量场景下的可行性与计算效率,本文以上述金属台阶的测量过程为例,对各网络模型的参数量及单次推理耗时进行了定量统计,结果如表 6 所示。

本文网络得益于物理模型与轻量化算子的深度结合,其内部包含幅度分支的特征提取机制有效避免了冗余的深层通道堆叠,使得计算资源高度集中于核心物理特征的解算上,其单次推理耗时仅为 27.76ms。相比之下,HDRSL-Net 为了同时实现高光校正与相位解算,采用了庞大的双 U-Net 串行级联架构,导致其参数量高达 76.32M。这种长链路的串行计算严重阻碍了硬件的并行效率,使其单次推理耗时攀升至 101.15ms。另一方面,DC-UNet 的参数量为 40.84M,尽管体量适中,但由于其扩张残差网络模块内部堆叠了大量 512 通道的高维密集卷积,带来了极高的显存访问成本,单次推理耗时达到了 34.85ms。

综上所述,本文所提方法不仅能在复杂的高反光表面场景下实现高精度的相位恢复与三维重建,更凭借其极致的轻量化设计与高效推理性能,有效克服了传统多重曝光方法拍摄图像多,耗时冗长的瓶颈,充分证明了其在高速在线三维测量任务中的应用潜力。

3 结论

针对结构光三维测量中金属零件等高反射表面易产生条纹过曝光、调制信息缺失及相位解算不稳定的问题,本文提出一种基于数字孪生数据驱动的调制引导注意力网络的三维重建方法。首先,本文构建了面向高反光场景的数字孪生条纹投影系统,对投影–成像物理过程进行建模,生成了具有多样反射特性的虚拟数据集,该方法不仅系统模拟了复杂反射退化规律,也为深度学习模型提供了精准的物理映射先验与稳定、无冗余的训练数据支撑,有效克服了真实场景下高动态条纹数据采集成本高、标注困难的问题。其次,

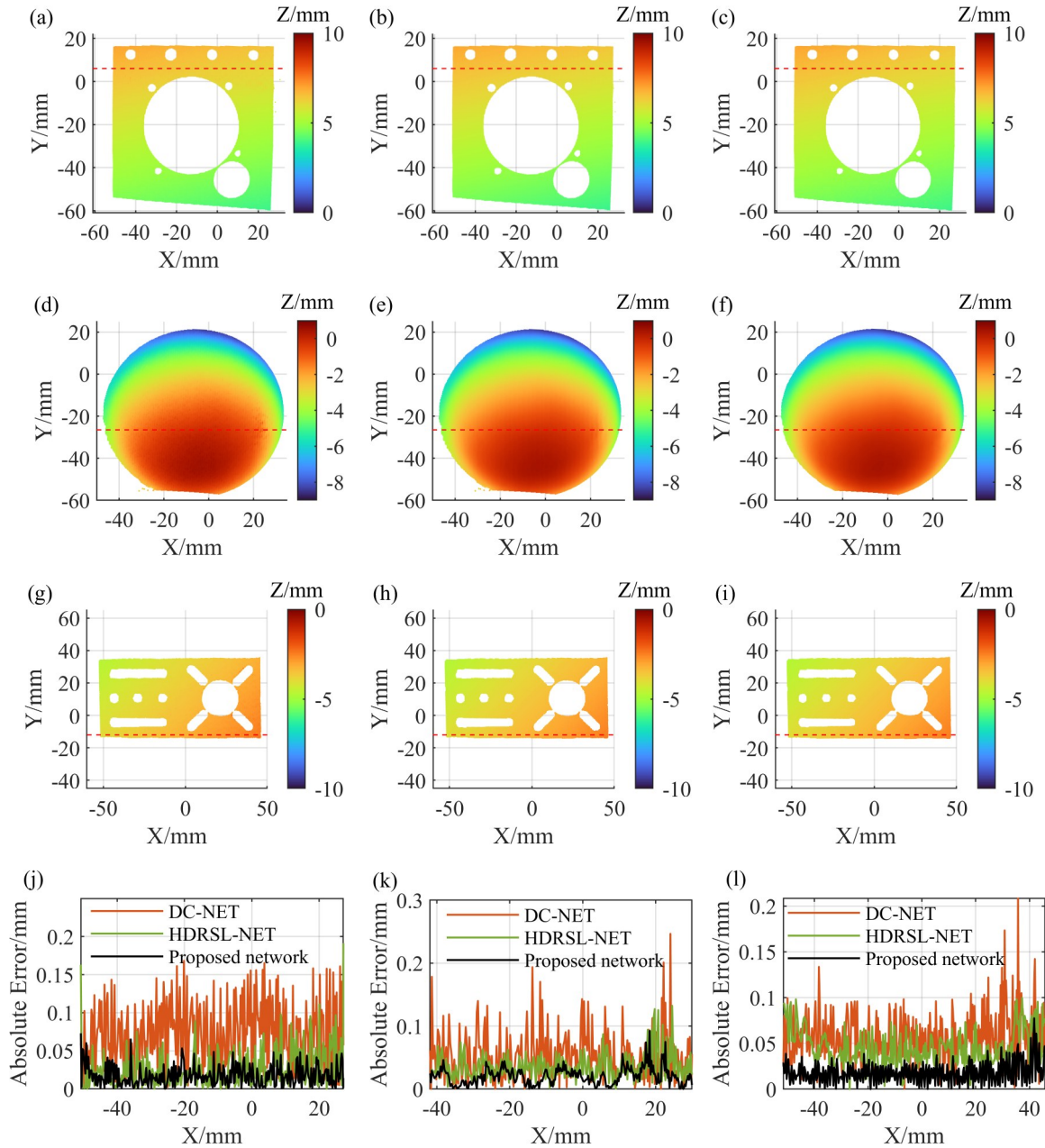


图 15 不同方法对金属工件的三维测量结果与误差分析。(a)-(c)法兰底座分别采用DC-UNet、HDRSL-Net与本文方法的重建结果;(d)-(f)金属圆盖的对应重建结果;(g)-(i)电机支架的对应重建结果;(j)-(l)不同方法在剖面线处的深度误差分布

Fig. 15 3D reconstruction results of metal workpieces using different methods. (a)-(c) Reconstruction results of the flange base using DC-UNet, HDRSL-Net, and the proposed method, respectively; (d)-(f) Corresponding reconstruction results of the metal cover; (g)-(i) Corresponding reconstruction results of the motor bracket; (j)-(l) Depth error distributions of each workpiece along the section line.

表 4 三种网络在不同金属工件上的有效点占比对比

Table 4 Comparison of valid point ratios of three networks on different metallic parts

Method	Flange base	Metal round cover	Motor bracket
DC-UNet	72.09%	75.87%	76.74%
HDRSL-Net	85.47%	87.21%	88.37%
proposed network	99.32%	98.93%	99.67%

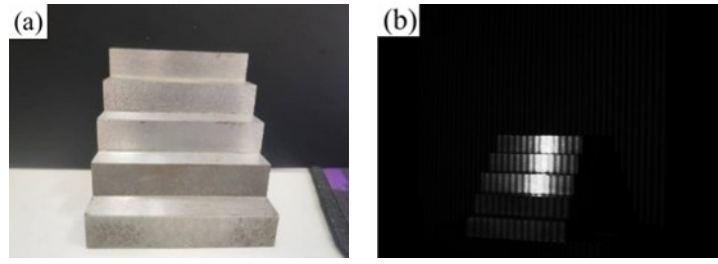


图 16 金属台阶。(a)实物样图;(b)变形条纹图

Fig. 16 Metallic step. (a) Physical sample; (b) Deformed fringe pattern.

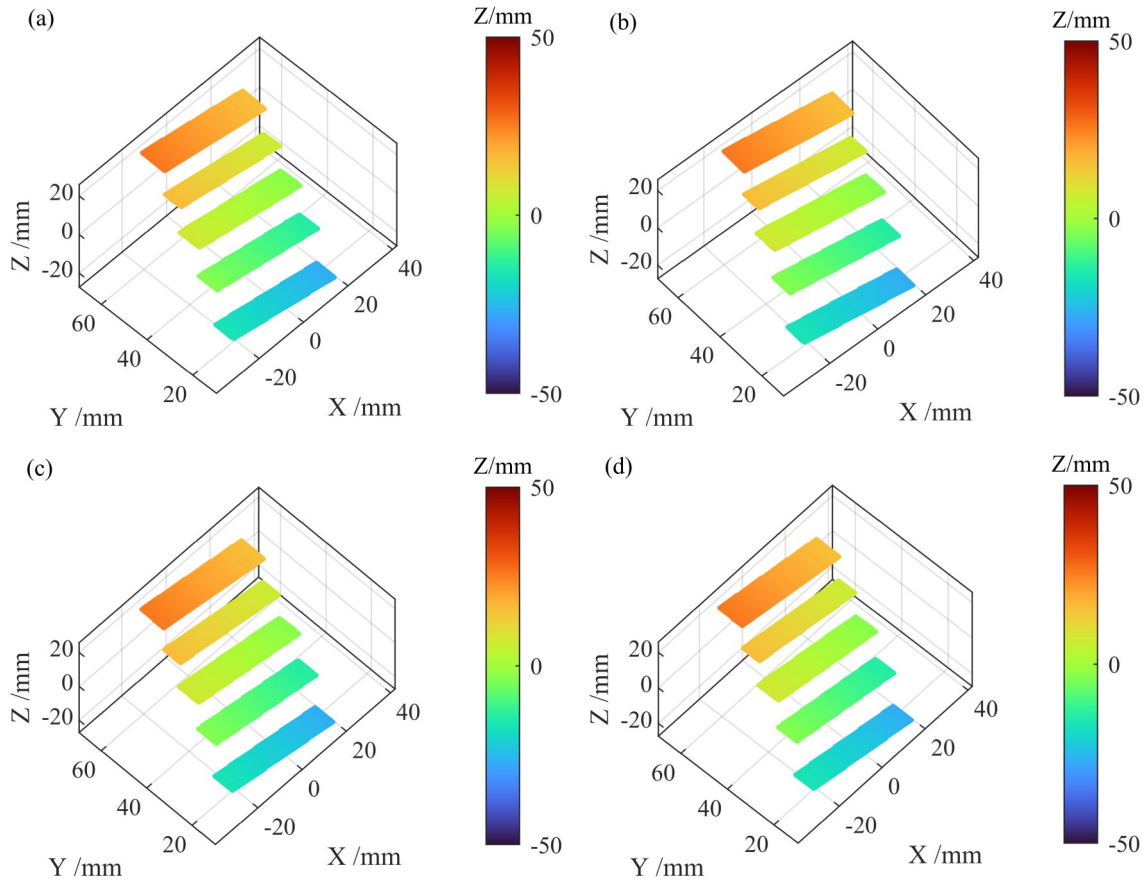


图 17 金属台阶测量结果。(a)多重曝光方法;(b)DC-UNet;(c)HDRSL-Net;(d)所提出的网络

Fig. 17 Measurement results of metallic step. (a) Multiple exposure method; (b) DC-UNet; (c) HDRSL-Net; (d) Proposed network.

所提出的网络整体采用编码-解码结构,创新性地结合了多尺度残差编码模块(RMSB)、全局特征感知模块(GFPM)以及调制引导模块(MGM)。该架构实现了对全局结构与局部特征的协同建模,将条纹的空间梯度与局部动态范围作为物理先验引入注意力机制,能够自适应抑制冗余响应并增强高反光区域的条纹信息,从而直接从单幅受高反光干扰的条纹图像中稳定预测出理想的正弦与余弦分量。系统性的验证实验表明,所提方法在处理复杂高反光表面时展现出了良好的鲁棒性与三维重建精度。在测试集消融实验中,基于多模块协同设计的模型的平均绝对误差仅为0.0083 mm。在对法兰底座、金属圆盖等实际零件的测量中,整体的三维重建更加平滑,高光区域点云完整度显著提升(有效点占比均达98.9%以上);标准金属台阶的定量实验显示,其最大绝对测量误差仅为0.032 mm,且单次推理耗时低至27.76 ms。该结果不仅在精度上优于现有深度学习网络,且在大幅减少图像采集数量的前提下,达到了与传统多重曝光法相当的测量水平,为高反射工业零件的高速、高精度三维测量提供了一种有效且实用的解决方案。

表5 金属台阶测量结果与误差
Table 5 Measurement results and errors of the metallic step

台阶面		1—2	2—3	3—4	4—5
CMM结果/mm		9.771	9.019	10.984	13.006
测量结果/mm	多重曝光方法	9.796	9.055	10.965	13.033
	DC—UNet	9.852	8.909	11.080	13.084
	HDRSL—Net	9.818	8.989	10.938	12.959
	本文方法	9.803	9.048	10.962	13.031
高度误差/mm	多重曝光方法	0.025	0.036	−0.019	0.027
	DC—UNet	0.081	−0.110	0.096	0.078
	HDRSL—Net	0.047	−0.030	−0.046	−0.037
	本文方法	0.032	0.029	−0.022	0.025

表6 不同网络模型的参数量与单次推理耗时对比
Table 6 Comparison of parameters and single-pass inference time among different network models

Network	Model size /M	Single—pass inference time /ms
DC—UNet	40.84	34.85
HDRSL—Net	76.32	101.15
Proposed network	33.45	27.76

参考文献

- [1] HUANG Yiyang, WANG Jie, SONG Yiping, et al. A novel defect detection method with eliminating dust for specular surfaces based on structured-light modulation analysis technique[J]. Optics and Laser Technology, 2021, 141: 107089.
- [2] LU Lilian, WU Zhoujie, ZHANG Qican, et al. High-efficiency dynamic three-dimensional shape measurement based on misaligned Gray-code light[J]. Optics and Lasers in Engineering, 2022, 150: 106873.
- [3] WANG Jianhua, YANG Yanxi. A new method for high dynamic range 3D measurement combining adaptive fringe projection and original-inverse fringe projection[J]. Optics and Lasers in Engineering, 2023, 163: 107490.
- [4] ZHANG Pan, ZHONG Kai, LI Zhongwei, et al. Hybrid-quality-guided phase fusion model for high dynamic range 3D surface measurement by structured light technology[J]. Optics Express, 2022, 30(9): 14600–14614.
- [5] SONG Zhan, JIANG Hualie, LIN Haibo, et al. A high dynamic range structured light means for the 3D measurement of specular surface[J]. Optics and Lasers in Engineering, 2017, 95: 8–16.
- [6] CHEN Chao, GAO Nan, WANG Xiangjun, et al. Adaptive projection intensity adjustment for avoiding saturation in three-dimensional shape measurement[J]. Optics Communications, 2018, 410: 694–702.
- [7] BUDIANTO, LUN D P K. Inpainting for fringe projection profilometry based on geometrically guided iterative regularization[J]. IEEE Transactions on Image Processing, 2015, 24(12): 5531–5541.
- [8] HAN Hexiang, GAO Nan, ZHANG Guofeng, et al. Three-dimensional morphology measurement of highly reflective objects based on three channel parallel processing[J]. Acta Photonica Sinica, 2024, 53(7): 0712003.
韩鹤翔, 高楠, 张国锋, 等. 基于三通道并行的高反光物体三维形貌测量[J]. 光子学报, 2024, 53(7): 0712003.
- [9] RIVIERE J, RESHETOUSKI I, FILIPI L, et al. Polarization imaging reflectometry in the wild[J]. ACM Transactions on Graphics, 2017, 36(6): 1–14.
- [10] SALAHIEH B, CHEN Z, RODRIGUEZ J J, et al. Multi-polarization fringe projection imaging for high dynamic range objects[J]. Optics Express, 2014, 22(8): 10064–10071.
- [11] FENG Shijie, CHEN Qian, ZUO Chao, et al. Fast three dimensional measurements for dynamic scenes with shiny surfaces[J]. Optics Communications, 2017, 382: 18–27.
- [12] BENVENISTE R, UNSALAN C. A color invariant for line stripe-based range scanners[J]. Computer Journal, 2011, 54(5): 738–753.
- [13] WANG Hao, LU Ziyu, HUANG Ziyang, et al. A high-accuracy and reliable end-to-end phase calculation network and its demonstration in high dynamic range 3D reconstruction[J]. Nanomanufacturing and Metrology, 2025, 8: 5.
- [14] DANG Guoqing, ZHU Zhenmin, CHENG Xiaoyan, et al. Deep-learning-assisted composite polarized fringe projection profilometry for three-dimensional measurement of high-dynamic-range objects[J]. Applied Optics, 2025, 64(1): 221–232.
- [15] SHEN Mengmeng, HE Liangbo, ZHANG Haihua, et al. Deep learning based measurement accuracy improvement of high dynamic range objects in fringe projection profilometry[J]. Optics Express, 2024, 32(20): 35689–35702.

- [16] XI Dejun, HOU Lei, WU Fei, et al. Deep learning-based inpainting of high dynamic range fringe pattern for high-speed 3D measurement of industrial metal parts[J]. *Advanced Engineering Informatics*, 2024, 60: 102428.
- [17] LI Mojing, SUN Changku. A multi-task learning method for 3D reconstruction of highly reflective surfaces[J]. *Acta Photonica Sinica*, 2024, 53(10): 1012001.
李默晶, 孙长库. 用于高反光表面三维重建的多任务学习方法[J]. *光子学报*, 2024, 53(10): 1012001.
- [18] ZHANG Zonghua, TOWERS CE, TOWERS DP. Time efficient color fringe projection system for 3D shape and color using optimum 3-frequency[J]. *Optics Express*, 2006, 14(14): 6444-6455.
- [19] ZHANG Zonghua, HUANG Shujun, MENG Shasha, et al. A simple, flexible and automatic 3D calibration method for a phase calculation-based fringe projection imaging system[J]. *Optics Express*, 2013, 21(10): 12218-12227.
- [20] LI Wenguo, LIU Tao, TAI Manli, et al. Three-dimensional measurement for specular reflection surface based on deep learning and phase measuring profilometry[J]. *Optik*, 2022, 271: 169983.
- [21] WANG Hao, ZHANG Chaobo, QIAN Xiang, et al. HDRSL Net for accurate high dynamic range imaging-based structured light 3D reconstruction[J]. *IEEE Transactions on Image Processing*, 2025, 34: 5486-5499.

Modulation-Guided and Digital Twin-Driven 3D Measurement Method for Highly Reflective Surfaces

ZHANG Boxia, FU Shuai, GAO Nan, MENG Zhaozong, ZHANG Zonghua
(School of Mechanical Engineering, Hebei University of Technology, Tianjin 300130, China)

Abstract: Objective The objective of this study is to address the critical challenges encountered in structured light three-dimensional (3D) measurement of highly reflective metallic surfaces, specifically the issues of fringe pattern saturation, loss of modulation information, and subsequent phase unwrapping distortion. To overcome these limitations, this paper proposes a novel 3D reconstruction method based on a modulation-guided attention network driven by digital twin data. The primary goal is to achieve robust, high-precision phase recovery and 3D shape reconstruction under high dynamic range (HDR) conditions using only a single degraded fringe image per spatial frequency. This approach is designed to eliminate the heavy dependence on traditional multi-exposure acquisition techniques and complex phase-shifting sequences while strictly maintaining optimal measurement accuracy.

Key words: Fringe projection profilometry; Digital twin; 3D reconstruction; Structured light; Phase retrieval

OCIS Codes: 120.5050; 120.2830; 100.6890; 110.5086

CSTR: 32255.14.gzxb20265508.0812002