引用格式：LIU Guijie, LIU Wulang, LI Mingfeng, et al. A Phase Recovery Method for Fringe Projection Profilometry Based on Multi-task Networks[J]. Acta Photonica Sinica, 2026, 55(3):0311001
刘规杰,刘伍浪,李明枫,等.一种基于多任务网络的条纹投影轮廓术相位恢复方法[J].光子学报,2026,55(3):0311001

一种基于多任务网络的条纹投影轮廓术相位恢复方法

刘规杰[1],刘伍浪[1],李明枫[1],黄朝光[2],黄文彬[3],马建平[1],李文杰[1,3]
(1 桂林电子科技大学 机电工程学院 广西制造系统与先进制造技术重点实验室,桂林 541000)
(2 桂林凯文彼德科技有限公司,桂林 541001)
(3 苏州大学 数字成像与显示教育部工程研究中心,苏州 215006)

摘　要:基于一种多任务网络,仅用单幅条纹图像实现包裹相位计算和条纹级次获取,继而直接获得展开相位。所设计网络模型以UNet为基础架构,引入残差模块,增强特征提取能力与训练稳定性。同时,所提模型结合信息收集与分发机制,实现多分支预测任务,直接输出包裹相位的分子项、分母项以及条纹级次,避免多个网络模型使用带来的繁琐、耗时问题。同时,充分利用包裹相位分布信息,构建条纹级次一致性连通域,以提高条纹级次分类的稳定性。实验结果表明,所提方法可实现精确、快速的相位信息恢复,并具有较好的泛化能力。

关键词:条纹投影轮廓术;多任务网络;信息收集与分发机制;相位恢复;连通域分割
中图分类号:O436　　**文献标识码**:A　　**doi**:10.3788/gzxb20265503.0311001

0　引言

三维重建技术近年来在多个领域得到了广泛应用,涵盖了医学成像、机器人导航、虚拟现实、3D动画建模、产品在线检测等[1-3]。在众多的三维重建技术中,条纹投影轮廓术(Fringe Projection Profilometry, FPP)通过向物体表面投射编码图案,实现三角测量中同名点对的准确定位,继而实现三维重构,具有抗干扰性强、非接触、精度高等优点[4]。目前,FPP广泛应用于工业检测、逆向工程、文物扫描和医学诊断等领域[5-7]。

尽管条纹投影轮廓术(FPP)已在多个应用领域展现出良好性能,但在实际工程应用中仍存在诸多待解决的问题[8]。首先,传统相移轮廓术通常需要投射多幅光栅图案,这会导致重建时间较长,难以满足实时性要求[9]。其次,傅里叶变换轮廓术虽可基于单帧图像提取相位信息,但在复杂表面或高噪声环境下重建精度难以保障[10]。近年来,研究者将深度学习与相位提取及相位展开过程相结合,在显著减少所需投影图案数量的同时获得很高的重建精度[11]。NGUYEN A H等[12]系统化研究了投影不同条纹频率图案在深度学习框架下单帧2D到3D图像转换精度的影响。FENG S等[13]提出了一种融合深度学习与传统相移法物理模型的相位提取方法,能够从单幅条纹图像中恢复高质量的包裹相位。为解决传统深度学习方法在光学计量中依赖大量数据且忽视物理规律的问题,YIN W等[14]提出了一种基于物理信息的深度学习方法。LI Y等[15]针对传统单帧相位测量方法的频谱混溶问题,将深度学习网络与FPP的物理模型相结合,实现从较大的表面不连续性或孤立的对象的复杂场景中获取可靠相位。SHI J等[16]则采用深度学习对条纹图像进行增强,从而有效提升了基于傅里叶变换的相位解算精度。然而,这类方法主要聚焦于包裹相位的高质量恢复,尚无法

基金项目:国家自然科学基金地区基金(52565058),广西自然科学基金面上项目(2025GXNSFAA069774),广西制造系统与先进制造技术重点实验室主任课题(AD25069080),苏州大学数字成像与显示教育部工程研究中心(SDGC2319)
第一作者:刘规杰,2917347458@qq.com
通讯作者:李文杰,li-wenjie@163.com
收稿日期:2025-11-04;**录用日期**:2025-12-31

实现绝对相位的直接提取。为此，QIAN J等[17]将三种频率的条纹图案分别编码至彩色图像的RGB通道，并设计五路径卷积神经网络预测包裹相位的分子项、分母项及粗略绝对相位。然而，该方法在面对彩色物体时容易受到物体本身颜色干扰，且复合条纹图案的解码过程复杂，影响系统鲁棒性。在随后的研究中，LI J等[18]通过将条纹序信息嵌入高频条纹图像的相位移分量中，并采用两个轻量级卷积神经网络，分别预测高频包裹相位的分子和分母以及条纹阶数。NGUYEN A H等[19]提出了一种单输入双输出网络结构，通过一条编码路径和两条解码路径将单幅条纹图像映射为多步相移图像与粗相位图。TAN H等[20]则将深度学习与双目条纹投影相结合，设计了一个五通道编码网络，将物体条纹图、参考平面条纹图与条纹级次信息编码为五通道数据输入，并采用双解码路径分别预测包裹相位与条纹阶数。虽然上述方法可以通过网络预测后恢复绝对相位，但多网络协同方案或双解码结构往往会显著增加模型复杂度与计算成本。

综上所述，目前基于深度学习的绝对相位恢复方法主要包括两类：一类是分别预测三个不同频率包裹相位的分子与分母项，并结合三频外差法或数论法计算绝对相位；另一类是采用多网络或双解码器的网络结构，分别预测高频包裹相位的分子分母项和条纹阶次，从而获得绝对相位。前者由于包裹相位误差在多频外差计算过程中累积，导致最终得到的绝对相位误差较大且稳定性较差；后者由于多网络或双解码器的设计，导致计算复杂度高，推理效率低。

针对当前方法存在的误差累积与模型复杂度高等问题，本文提出了一种基于多任务网络(UNet with Information Gather and Distribute Mechanism，GD-UNet)的单幅条纹图像绝对相位恢复方法。该方法在单个网络结构中实现了包裹相位与条纹阶次的同时预测，实现绝对相位直接恢复。所设计的GD-UNet网络以经典UNet架构为基础，引入残差模块以增强特征提取能力与训练过程的稳定性，并融合信息收集与分发模块以支持多分支预测任务。针对条纹边缘区域易出现包裹相位与条纹阶次像素点不对应的问题，本文充分利用包裹相位信息进一步引入连通域分割方法对初始预测条纹阶次进行修正。

1 基本原理

如图1所示，整体流程包含以下关键步骤：首先，向不同姿态的目标物体投影三组不同频率的条纹图像以构建训练数据集，并利用3频12步相移获得包裹相位的分子分母项以及条纹级次分布图，用于构建网络训练标签；随后，将3频投影图像融合为一幅RGB图像作为网络输入，并以最高频包裹相位的分子分母项、条纹级次作为预测目标；再通过解码网络输出的分子分母项和条纹级次信息，获得绝对相位；最终结合预先标定的投影仪-相机系统参数，实现目标物体的高精度三维重建。

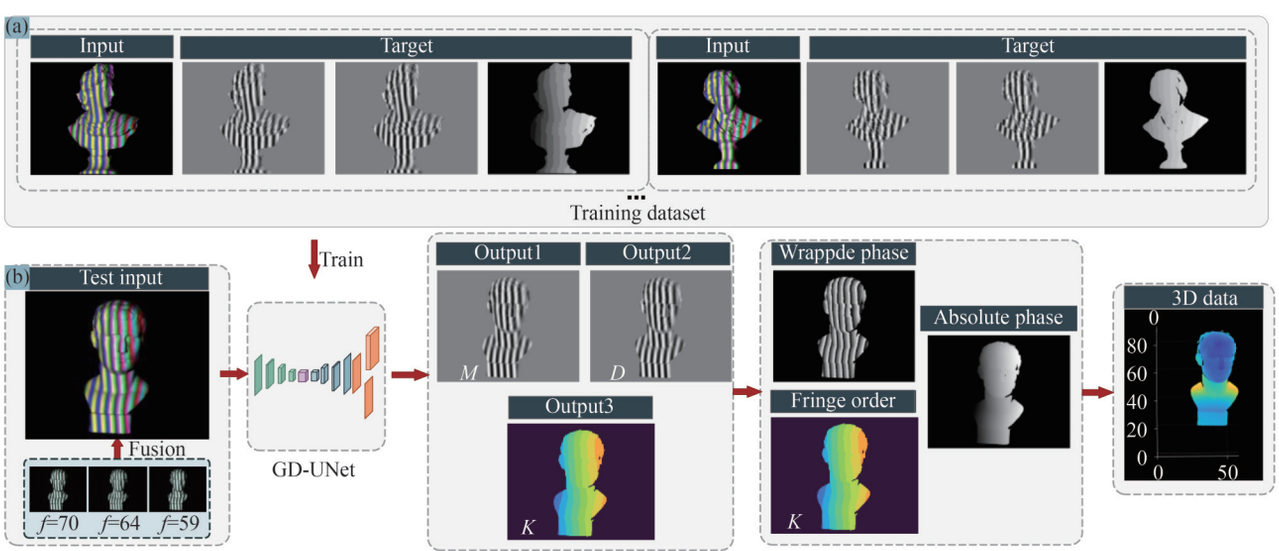

图1 基于深度学习条纹投影轮廓术的流程图。(a)部分网络训练数据集；(b)测试数据及训练模型预测结果

Fig.1 Flowchart of the deep learning-based fringe projection profilometry technique. (a) Part of network training datasets; (b) Test data and prediction results obtained by the training model

1.1 数据集生成

为了提升特征提取效率并降低计算负担，本文在目标区域周围裁剪，使用800×640像素的子图像作为数据集。数据集构建流程如图2所示。本文采用3频12步相移条纹投影方式，依据三频外差法[22]选择的条纹频率为70、64、59（如图2(a)所示），频率之间的关系是为了经过外差操作后获得频率为1的外差相位，继而消除相位歧义性，从而可以获得每个频率的条纹级次。为增强模型泛化能力，数据集涵盖了不同材质、纹理复杂度及几何形状的物体，并对每个物体进行了多角度、多姿态的图像采集。网络的输入为三幅不同频率的相移灰度图分别赋值至RGB图像的R、G、B通道后的融合图像，如图2(b)所示。网络的输出包括三个不同分支，分别为预测包裹相位的分子项M、分母项D（如图2(d)所示），以及对应的条纹级次K（如图2(e)所示）。

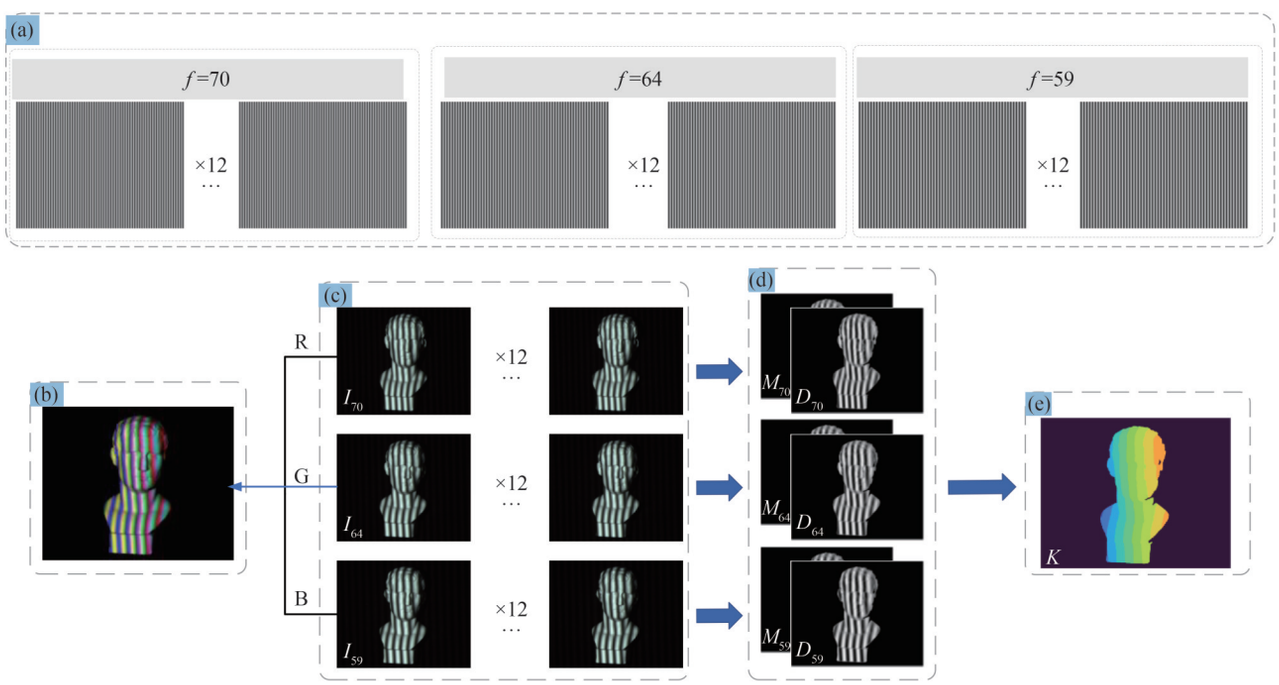

图2 数据集生成方法及流程。(a)投影仪投影的条纹图像；(b)三种条纹频率融合的RGB图像；(c)采集的n步相移条纹图像；(d)包裹相位的分子分母项；(e)条纹阶次

Fig.2 Dataset generation methodology and process. (a) Fringes projected by the projector; (b) Three-frequency fringe-fused RGB image. (c) Captured n-step phase-shifted fringe patterns; (d) Numerator and denominator terms of the wrapped phase; (e) Fringe order

采用3频率12步相移的条纹模式构建高质量标签数据，条纹图像可表示为

$$I_n(x,y)=A(x,y)+B(x,y)\cos\left[\phi(x,y)+\frac{2\pi n}{12}\right] \tag{1}$$

包裹相位为

$$\varphi(x,y)=\arctan\frac{\sum_{n=1}^{12}I_n(x,y)\sin\left(\frac{2\pi(n-1)}{12}\right)}{\sum_{n=1}^{12}I_n(x,y)\cos\left(\frac{2\pi(n-1)}{12}\right)}=\arctan\frac{M(x,y)}{D(x,y)} \tag{2}$$

式中，$M(x,y)$，$D(x,y)$是包裹相位的分子项和分母项。基于三个频率获得的包裹相位，利用三频外差法即可获得条纹级次[22]。为了增强网络的学习能力，通过调制函数和掩码函数设置适当的调制阈值（Thr），以屏蔽训练数据图中的无效点。

$$B(x,y)=\frac{2}{N}\sqrt{M(x,y)^2+D(x,y)^2} \tag{3}$$

$$\mathrm{Mask}(x,y)=\begin{cases}B(x,y), & B(x,y)\geqslant \mathrm{Thr}\\ 0, & B(x,y)<\mathrm{Thr}\end{cases} \tag{4}$$

1.2 网络模型

在GD-UNet的任务中，包裹相位预测被建模为图像到图像回归任务，属于典型的像素回归任务，而条纹级次预测则被建模为多类别分类任务。在像素回归问题和多类别分类问题中，已有多种卷积神经网络(CNN)架构被广泛应用。其中，UNet[21]以其典型的U形结构在众多视觉任务中表现优异。该结构由编码器(下采样路径)和解码器(上采样路径)组成，编码器通过卷积与池化操作逐步提取图像的高层特征，解码器则通过上采样与卷积操作恢复图像的空间分辨率，实现精细的输出预测。尽管UNet在特征提取和空间信息恢复方面表现优异，随着网络加深，仍存在梯度消失和特征退化的问题，限制了深层特征的有效利用。为了改善这些局限性，本文在UNet网络中引入残差模块，通过短路连接加强特征流动，提升网络的训练稳定性与特征表达能力。在此基础上，进一步提出GD-UNet，通过设计信息整合与分发模块，优化特征在多任务分支之间的分配策略，显著提升了单幅条纹图像相位恢复的准确性与鲁棒性。

GD-UNet整体框架如图3所示，主要由四个核心模块构成：下采样模块(Downsampling Module，DownM)、上采样模块(Upsampling Module，UPM)、残差模块(Residual Module，ResM)以及信息整合与分发模块(GDM)。如图3(b)所示，ResM采用双分支结构：一条分支由单层3×3卷积构成，另一分支由多个不同尺度和深度的卷积核组合而成，以提取多尺度特征信息。两条分支的输出通过逐元素相加的方式进行融合，随后接入Leaky_ReLU激活函数以增强非线性表达能力。

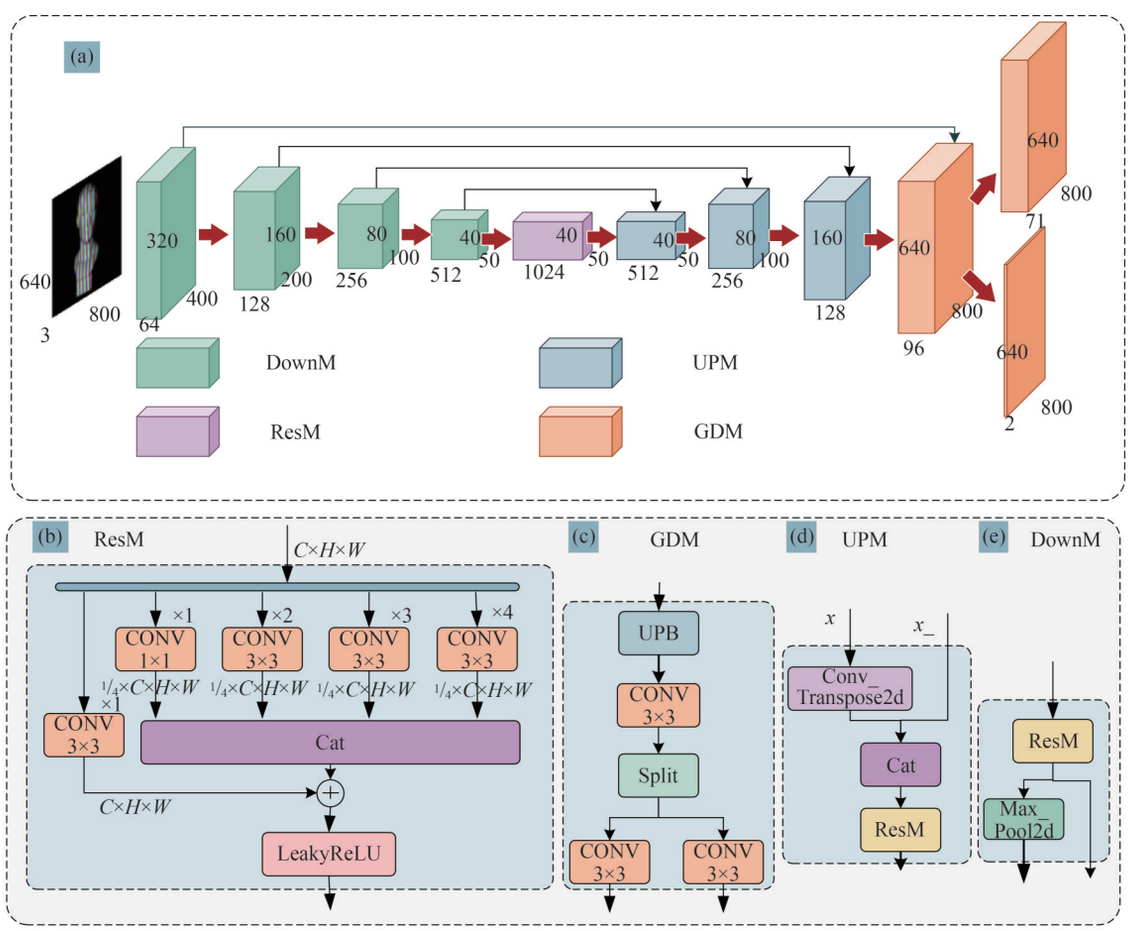

图3 GD-UNet的网络结构。(a)GD-UNet的整体网络架构；(b)残差模块；(c)信息整合与分发模块；(d)上采样模块；(e)下采样模块

Fig.3 Network structure of GD-UNet. (a) Overall network architecture of GD-UNet; (b) Residual module; (c) Gather-and-distribute mechanism module; (d) Upsampling module; (e) Downsampling module

在编码阶段，每个DownM由一个ResM和一个最大池化层组成，如图3(e)所示，其中，ResM负责提取特征信息，池化操作则用于压缩空间维度；同时，ResM的输出被保留并用于与解码阶段的上采样模块建立横向连接，以加强上下文信息的传递。整个下采样重复四次，实现从低层到高层语义特征的逐步提取，其中每次下采样均使特征通道数加倍、空间分辨率减半。

在解码阶段，UPM(如图3(d)所示)首先通过转置卷积对特征图进行上采样，并通过横向连接拼接来自编码器的相应横向特征。最终输出的特征图传递给GDM，其结构如图3(c)所示。该模块首先通过横向连接与上采样操作收集多尺度特征信息，并利用3×3卷积进一步提炼与增强特征表达。随后按照通道维度将特征划分，并分别分配至两个子任务分支。不同于传统使用全连接层的方法，本文采用全卷积结构，大大降低了计算复杂度，提高了推理效率。

对于包裹相位的分子与分母项的预测，通常采用平均绝对误差(L1损失)或均方误差(L2损失)作为网络的损失函数。然而，这些方法往往忽略了$M(x,y)$和$D(x,y)$图像中所蕴含的几何结构和空间结构信息。为了更全面地评估预测结果的质量，本文同时采用Smooth L1损失和结构相似性指数(SSIM)损失作为网络的损失函数，具体表达式为

$$\mathrm{Loss}_{\mathrm{ND}}=\lambda_1\mathcal{L}_{\mathrm{SSIM}}(x,y)+\lambda_2\mathrm{Smooth}_{L1}(d) \tag{5}$$

$$L_{\mathrm{SSIM}}(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \tag{6}$$

$$\mathrm{Smooth}_{L1}(d)=\begin{cases}0.5d^2 & \text{if } |d|<1\\ |d|-0.5 & \text{otherwise}\end{cases} \tag{7}$$

式中，λ_1、λ_2是常数，表示Smooth L1和SSIM损失在$\mathrm{Loss}_{\mathrm{ND}}$的重要程度；$\mu_x$和$\mu_y$分别是图像$x$和$y$的均值，表示亮度；$\sigma_x^2$和$\sigma_y^2$分别是图像$x$和$y$的标准差，表示对比度；$\sigma_{xy}$是图像$x$和$y$的协方差，表示结构相似性；$C_1$和$C_2$是常数，用于避免分母为零的情况；$d$是预测值与真实值之间的差异。

为明确损失函数中的权重选择，本文通过在合理范围内对λ_1与λ_2的多组组合进行实验，并通过验证集上的相位误差与损失平滑性等指标综合评估其效果。当λ_1与λ_2均设为0.5时，模型既能准确恢复相位数值，又能保持损失的平滑性，从而兼顾相位精度与结构完整性。

对于条纹级次预测，采用多分类交叉熵损失

$$\mathrm{CrossEntropyLoss}\left(k,\hat{k}\right)=-\sum_{i=1}^{C}k_i\log\left(\hat{k}_i\right) \tag{8}$$

式中，k_i是真实标签的one-hot编码；$\hat{k}_i$是模型预测的第i类的概率。

1.3 条纹级次修正

在实际应用中，条纹阶次的预测会受到各种干扰因素的影响，产生偶然性的误差。如图4中红色圆圈标记区域所示，该区域的包裹相位对应的真实条纹级次为32，而网络预测结果为33。为了提高条纹级次预测的稳定性，本文采用基于包裹相位的连通域分割技术[23]，如图5所示。

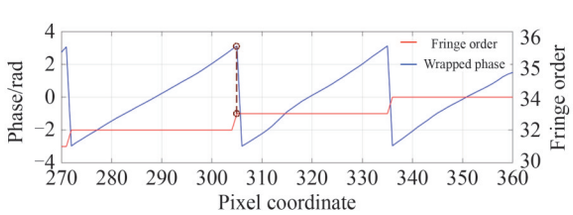

图4 条纹级次与包裹相位的对应关系

Fig.4 The correspondence between the fringe order and the wrapped phase

首先，对包裹相位进行二值化操作，得到二值图像掩码mask

$$\begin{cases}\mathrm{mask1}(x,y)=\varphi(x,y)>0\\ \mathrm{mask2}(x,y)=\varphi(x,y)<0\end{cases} \tag{9}$$

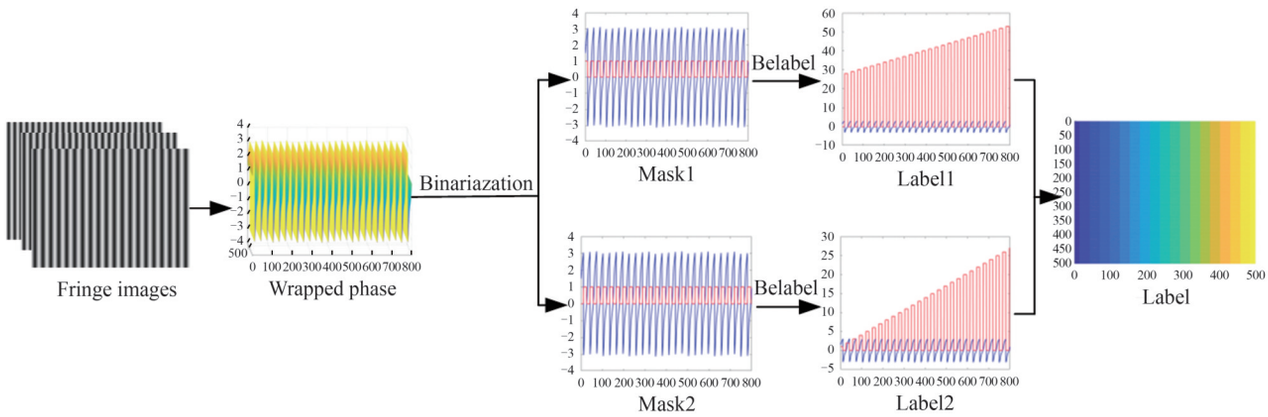

图5 基于包裹相位的连通域分割方法

Fig.5 Connected domain segmentation method based on the wrapped phase

随后,利用bwlabel函数对mask中的白色区域进行连通域分割,分别获得不同的连通区域标签Label1和Label2。根据这些标签可构建出完整的连通域分割图Label,从而提取出多个相互独立的连通区域。理论上,在同一连通域内的所有像素应该具有相同的条纹阶次。由于预测误差通常集中在连通域的边缘区域,因此采用多数投票策略来纠正错误的边缘阶次预测。具体而言,通过选择该域中频次最高的值来确定每个连通域的最终条纹阶数,从而实现条纹阶次的有效校正。

2 实验验证与分析

2.1 实验平台

为验证本文提出的方法,构建了基于单目结构光的三维重建实验平台,主要由维视工业相机(MV-EM120M)与德州仪器工业投影仪(DLP4500)组成。本文采集了包含500组多场景数据的数据集,其中400组用于训练、50组用于验证、50组用于测试,涵盖了不同纹理复杂度及材质类型(如石膏、金属等),充分验证模型在多样环境下的适应性与泛化能力(如图6所示)。

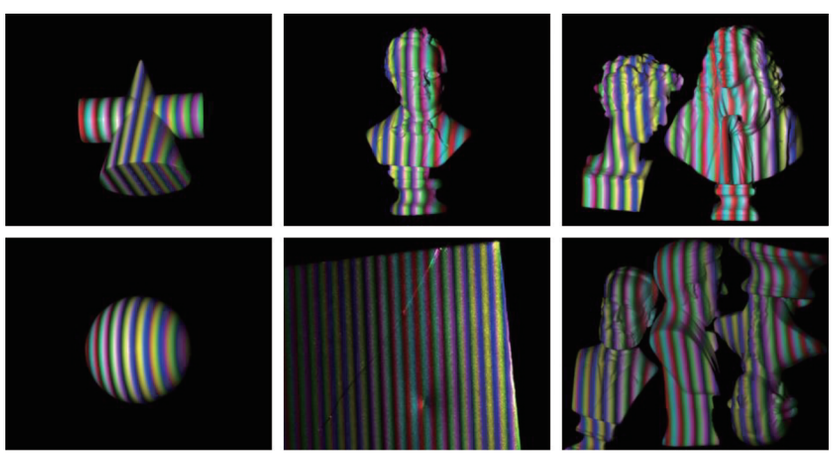

图6 数据集部分样本

Fig.6 Partial sample of the dataset

实验配置包括16核Intel(R) Xeon(R) Platinum 8474C处理器、NVIDIA RTX 4090D显卡(24 GB显存)及120 GB内存,操作系统为Ubuntu 20.04.3 LTS。训练过程中设置迭代周期为200个epoch,优化器采用Adam,batch_size设为3,并引入正则化项(weight_decay=10^{-5})以提升模型泛化性能。

2.2 结果分析与评估

2.2.1 条纹级次修正必要性验证

本文引入准确率(Accuracy, Acc)作为条纹阶次预测的评价指标,其定义为在剔除背景区域后,条纹阶

次预测正确的像素数占物体区域总像素数的比例。本文以空白背景区域为测试对象，对条纹级次校正算法的有效性进行验证，如图7所示。其中，图7(a)为输入图像；图7(b)为初始计算得到的绝对相位，其中局部区域存在显著的2π整数倍跳变误差，表明部分点的条纹级次预测存在错误；图7(c)展示了某行条纹级次修正前后的绝对相位局部放大对比图，能够直观地看出校正前后的差异；图7(d)为矫正后的绝对相位。结果表明，所提出的条纹阶次修正方法可有效消除跳变误差。

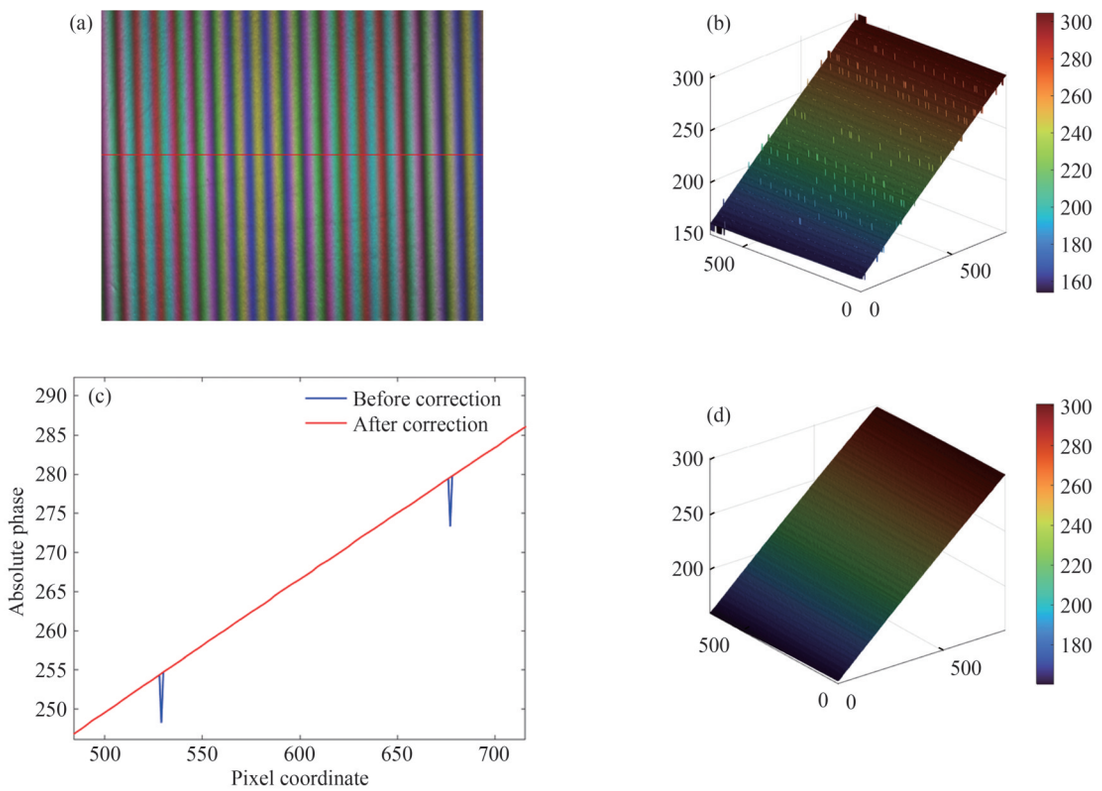

图7　条纹级次修正示意图。(a)输入图像；(b)初始绝对相位图；(c)初始绝对相位与校正后绝对相位的局部放大比较图；(d)矫正后的绝对相位

Fig.7　Fringe order correction. (a) Input image; (b) Initial absolute phase map; (c) Local magnification comparison of initial and corrected absolute phase; (d) corrected absolute phase map

2.2.2　GD-UNet与傅里叶变换方法的对比分析

为验证所提出方法的有效性，实验选取了网络从未见过的四个不同场景进行测量，如图8所示。如图9所示，在不同测试场景下，所提出的网络均能够准确预测包裹相位的分子项M、分母项D以及条纹级次k，体现了其良好的通用性与实用潜力。

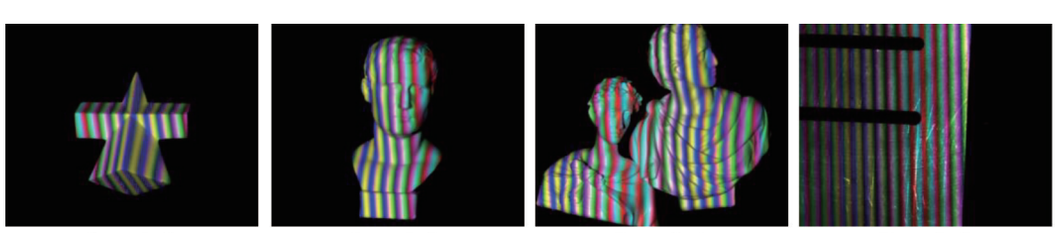

图8　输入网的四种不同测试场景

Fig.8　Network test input in four different scenarios

为了评估所提出的方法的有效性，本文将其与传统傅里叶变换方法进行对比。以12步相移法计算得到的包裹相位和多频外差法生成的绝对相位作为真值，相位误差如图10所示。实验结果表明，GD-UNet在相位提取和解包裹方面都优于FTP，能够在复杂表面以及不同材质的物体上获得更准确的绝对相位。

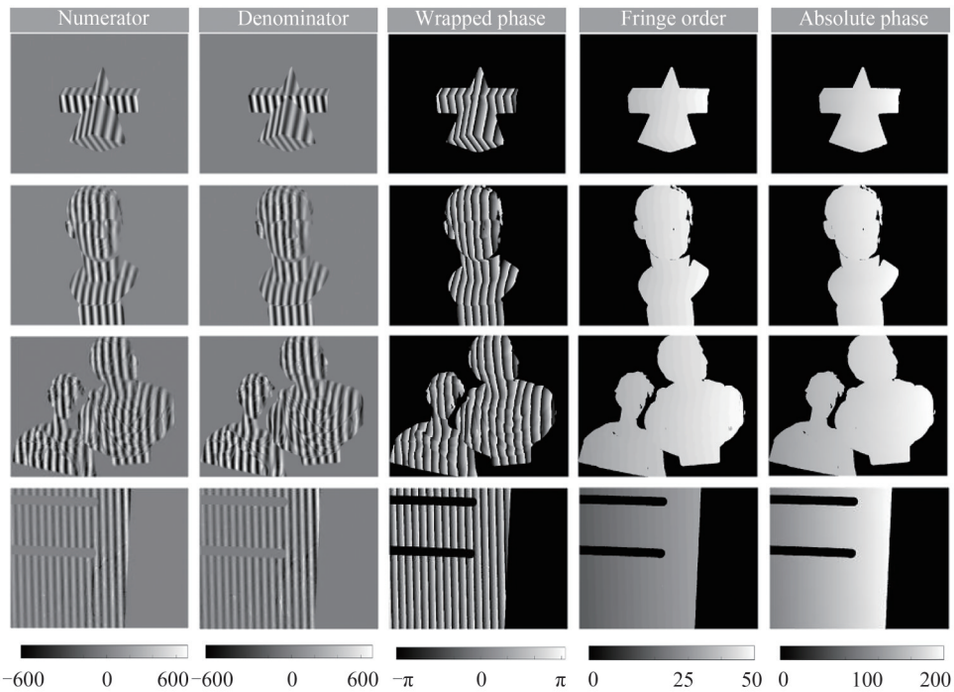

图 9 测试场景下GD-UNet的预测结果，依次为包裹相位分子项M、分母项D、包裹相位φ、条纹阶次K以及绝对相位Φ

Fig.9 The prediction results of GD-UNet in test scenario, showing the numerator M results, denominator D results, wrapped phase φ results, fringe order K results and absolute phase Φ maps respectively

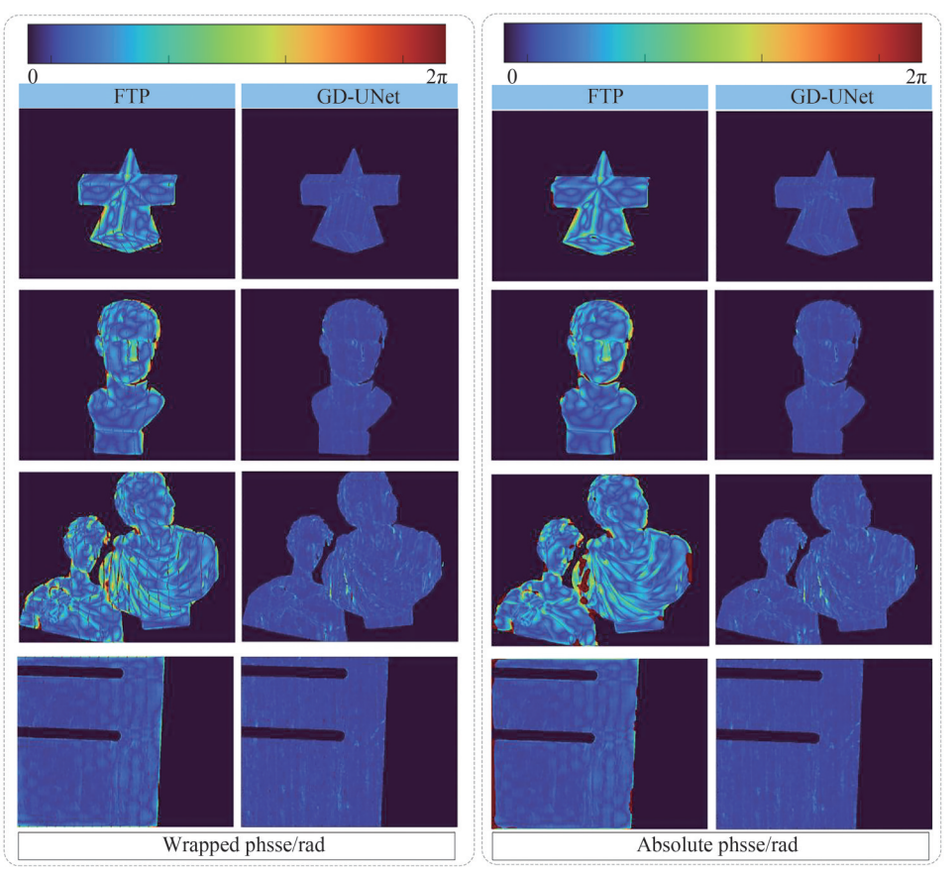

图 10 相位误差对比结果。前两列分别为FTP方法与GD-UNet的包裹相位误差，后两列为FTP方法与GD-UNet的绝对相位误差

Fig.10 Phase error results comparison. The first two columns are the wrapped phase error calculated by FTP method and GD-UNet. The last two columns are the absolute phase error by FTP method and GD-UNet

为了定量评估该方法的性能，本文采用均方根误差（RMSE）作为定量评价参数，结果如表1所示。在大多数测试场景中，GD-UNet的绝对相位误差有效控制在0.04 rad以内，即使面对表面极其复杂的物体，仍能保持较高的恢复精度。此外，所提方法的包裹相位误差的均方根误差与绝对相位误差的均方根误差相同，表明了条纹阶数预测的鲁棒性和准确性。本文方法的三维重建结果如图11所示。

表1　传统傅里叶变换法与GD-UNet方法的RMSE

Table 1　RMSE comparisons for FTP and proposed GD-UNet methods.

RMSE/rad	Scenario 1		Scenario 2		Scenario 3		Scenario 4	
	FFT	Our	FFT	Our	FFT	Our	FFT	Our
Wrapped phase	0.276 8	0.031 9	0.310 4	0.037 8	0.513 1	0.067 2	0.165 2	0.028 9
Absolute phase	0.276 4	0.032 0	0.416 1	0.037 8	0.658 3	0.067 0	0.230 4	0.029 0

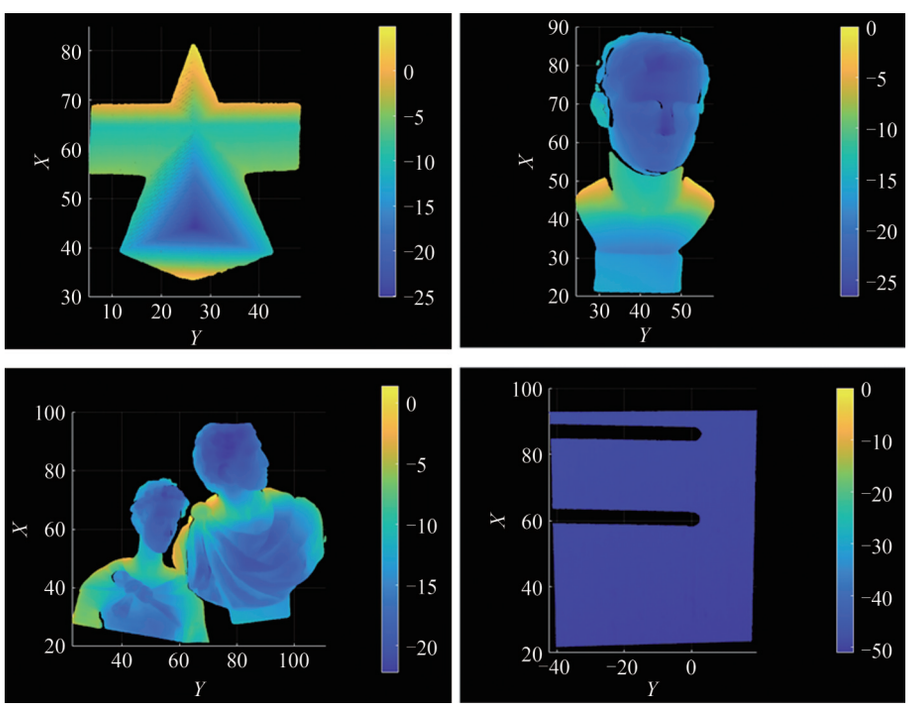

图11　使用GD-UNet方法获得的三维重建的点云图（单位：mm）

Fig.11　The point cloud map of 3D reconstructions obtained using the GD-UNet method (unit: mm)

2.2.3　GD-UNet方法与其他基于深度学习展开方法对比分析

为了体现本文所提出的方法在绝对相位计算方面的优越性，根据文献[24]、[18]、[20]所提出的方法，本文构建了三种深度学习方法F1、F2和F3进行对比，如图12所示。为了确保网络超参数的合理性和训练稳定性，本文在训练过程中引入了数据增强（包括平移、旋转和仿射扭曲）以增加样本多样性并降低过拟合风险；同时采用早停机制，当训练误差或验证误差不再下降时自动终止训练，从而进一步保证训练过程的稳定性。在网络设计上，采用相同的模块化设计（ResM、UpM、DownM），降低网络对超参数变化的敏感性。为进一步验证超参数配置的合理性，本文还进行了多组对照实验，比较了不同学习率、batch size、优化器及损失权重组合的性能，以保证模型稳定收敛并达到较优性能。

实验结果如图14和表2所示，与F1、F2和F3相比较，GD-UNet在精度上表现更优。不同方法的模型复杂度和推理速度如表3所示。尽管F2和F3的精度与GD-UNet较为接近，但其网络结构更为复杂：其中F2采用了两个独立的UNet，导致参数量几乎翻倍；F3采用双解码路径，参数量亦显著高于GD-UNet。相比之下，F1的参数量与本文方法接近，但精度存在明显不足。综上所述，GD-UNet在保持较低参数量的同时，实现了更高的相位恢复精度与更优的计算效率，充分体现了所提方法的有效性与优势。

为了进一步说明包裹相位误差累积对绝对相位展开的影响，本文构建了一种基于深度学习的单帧条纹解包裹方法（记为F3），如图13所示。不同方法所得到的包裹相位均方根误差（RMSE），如表4所示。由表4

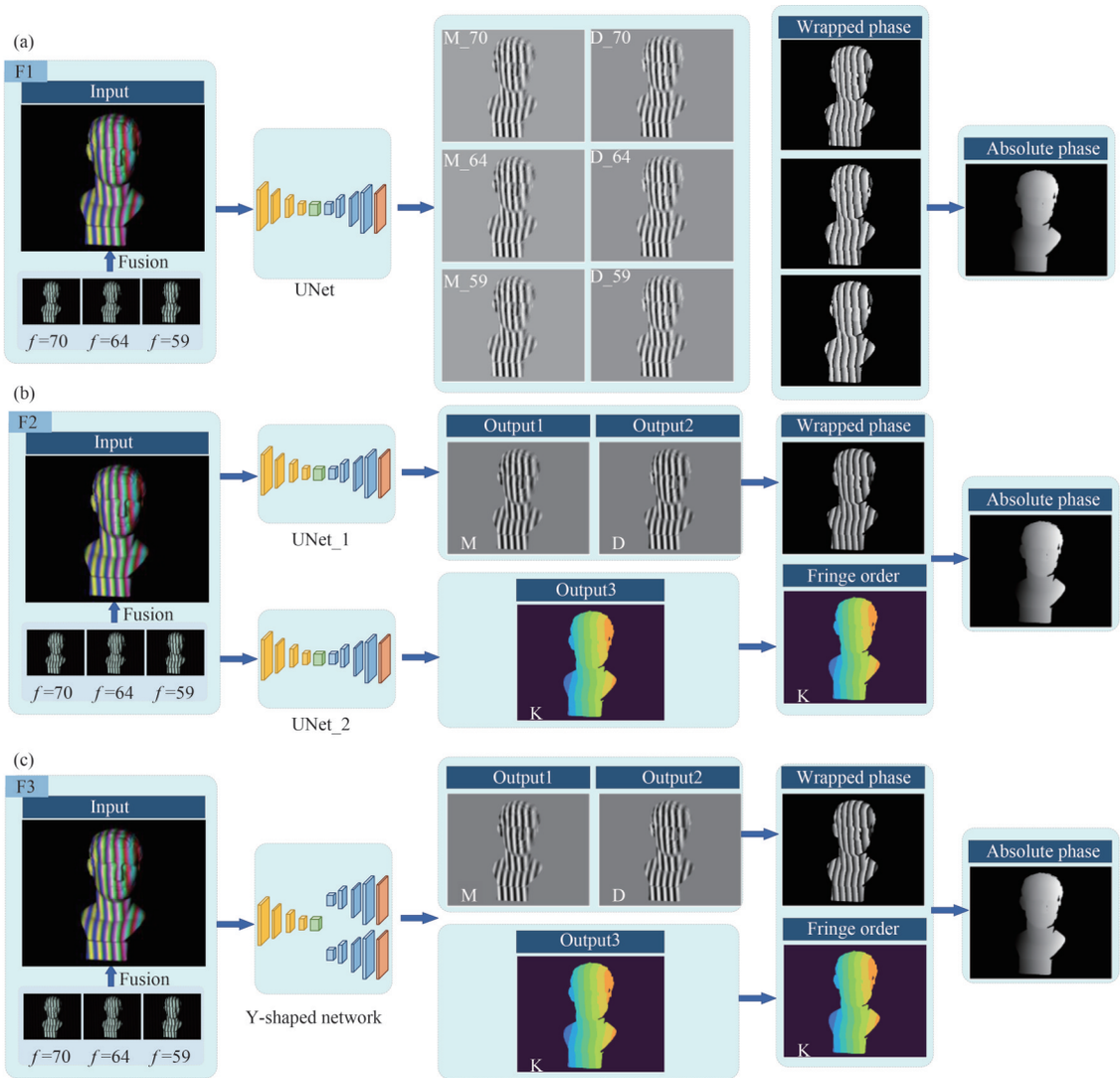

图 12 基于深度学习进行绝对相位展开的不同方案。(a)方法 F1:以三频合成图像为输入,网络输出三频包裹相位的分子项 M 和分母项 D;(b)方法 F2:以三频合成图像为输入,采用两个独立网络分别输出高频包裹相位的分子项 M、分母项 D 和条纹阶次 K;(c)方法 F3:以三频合成图像为输入,采用双解码路径的 Y 型网络同时输出高频包裹相位的分子项 M、分母项 D 和条纹阶次 K

Fig.12 Different deep learning-based methods for absolute phase unwrapping. (a) Method F1: The three-frequency synthesized image is used as input, and the network outputs the numerator M and denominator D terms of the wrapped phase at all three frequencies; (b) Method F2: The three-frequency synthesized image is used as input, and two separate networks are employed to output the numerator M and denominator terms D of high frequency wrapped phase and the fringe order K, respectively; (c) Method F3: The three-frequency synthesized image is used as input, and a dual-decoder network, Y-shaped network, is employed to simultaneously output the numerator M and denominator terms D of high frequency wrapped phase and the fringe order K, respectively

表 2 绝对相位误差 RMSE

Table 2 RMSE of absolute phase

Method	Scenario 1	Scenario 2	Scenario 3
F1	0.070 5	0.084 3	0.122 1
F2	0.056 1	0.075 6	0.039 3
F3	0.056 2	0.077 8	0.044 8
Our	0.048 3	0.068 7	0.039 0

的数据可见,F1、F2、F3 和 GD-UNet 方法的包裹相位误差比较接近,均显著优于 F4。实验结果表明,尽管网络的训练目标仅为高频包裹相位的分子与分母项,但低频条纹图在特征提取阶段提供了丰富的上下文信

表 3　不同方法模型的参数量和 FPS 比较

Table 3　Comparison of model complexity and inference speed for different methods

Method	Number of parameters (M)	FPS
F1	26.12	31.43
F2	26.12×2	15.79
F3	40.50	18.02
Our	26.40	25.93

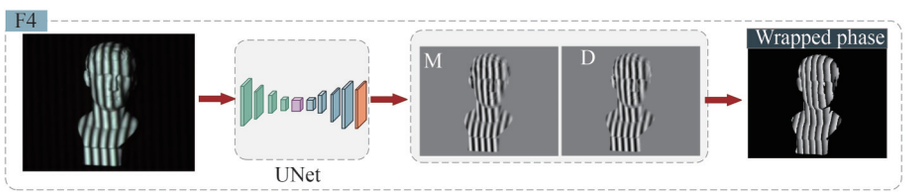

图 13　基于深度学习的单帧条纹解包裹

Fig.13　Single-frame fringe wrapped phase extraction based on deep learning

表 4　包裹相位 RMSE

Table 4　RMSE of wrapped phase.

Method	Scenario 1	Scenario 2	Scenario 3
F1	0.055 2	0.07 29	0.039 3
F2	0.048 0	0.070 8	0.042 9
F3	0.056 1	0.079 1	0.044 7
F4	0.087 6	0.117 4	0.058 6
Our	0.044 7	0.069 0	0.039 0

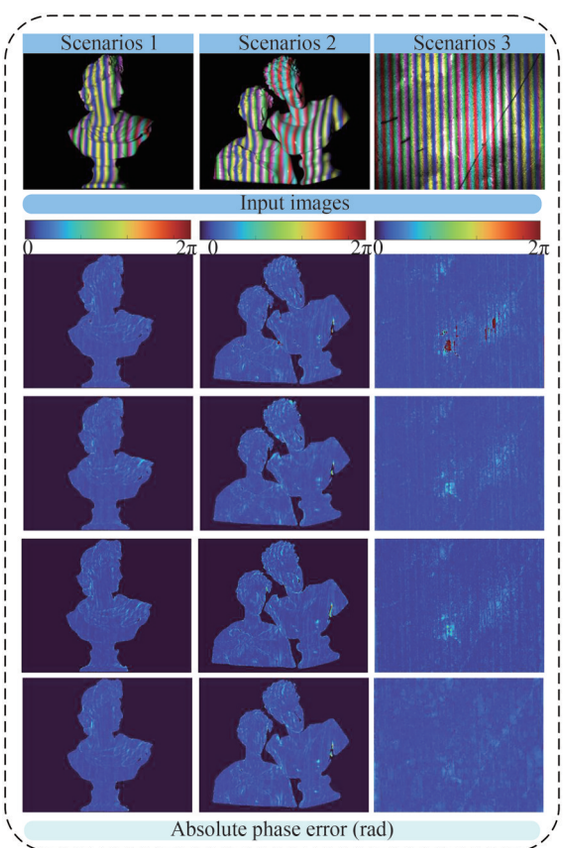

图 14　不同深度学习方法的绝对相位误差对比结果。第一行为输入图像,最后四行分别为 F1、F2、F3 及所提 GD-UNet 方法的绝对相位误差

Fig.14　Comparison of absolute phase errors among different deep learning-based methods. The first row is input images. Last four rows are the absolute phase errors by F1, F2, F3 and proposed GD-UNet methods

息，增强了网络对高频特征的建模能力，从而显著提升了包裹相位的预测精度。

2.3 阶梯块重建实验

为了系统评估GD-UNet方法的测量精度，本研究采用阶梯块作为标准测试样本进行三维重建实验验证。如图15所示，阶梯块由四个不同高度的阶梯面组成，其分别为3 mm、5 mm、10 mm和15 mm。重建的阶梯块三维点云数据，如图15(c)所示。5 mm阶梯面的拟合结果，如图15(d)所示，其中散点为原始点云数据，红色平面为拟合得到的基准平面。测量结果，如表5所示，其高度误差小于0.05 mm，验证了GD-UNet方法在毫米级尺度测量中的高精度特性。

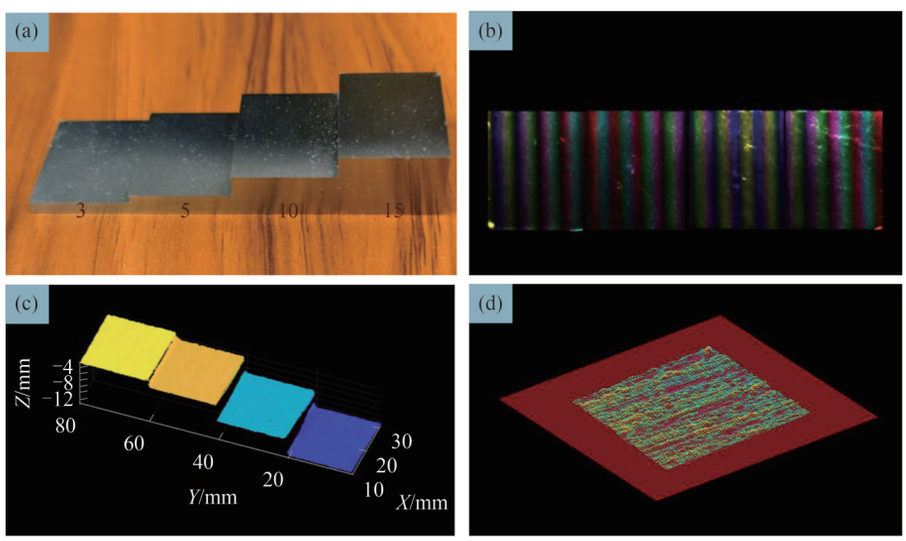

图15 阶梯块三维重构。(a)阶梯块；(b)采集的条纹图像；(c)重建的三维点云；(d)5 mm阶梯面的拟合结果

Fig.15 3D reconstruction of step block. (a) Physical step block specimen; (b) Captured fringe image; (c) Reconstructed 3D point cloud; (d) Plane fitting result for 5 mm step surface

表5 阶梯块平面间测量值(单位：mm)

Table 5 Measured height differences between adjacent step block planes (unit: mm)

Step height pair	3～5 mm	5～10 mm	10～15 mm
Measured height difference	1.976 3	5.032	4.958 1
Ideal height	2	5	5
Error	0.023 7	0.032	0.041 9

3 结论

本文提出了一种基于GD-Unet网络单帧条纹图像相位提取方法，只需要一幅融合条纹图像，一个网络架构就可以同时预测包裹相位的分子和分母项、条纹阶次，具有精度和效率两方面的优势。综合实验评估表明，该方法在复杂曲面重建、强噪声干扰以及多材质物体测量等挑战性场景下均展现出优异的测量精度和较强的鲁棒性。

本文旨在通过单个网络模型，实现对表面存在较大不连续性或包含孤立物体的复杂场景中的单幅条纹图像进行稳定且高精度的相位恢复。然而，在物体表面存在高反光等干扰因素时，所提出的方法在绝对相位的精确映射方面仍面临一定挑战，端到端的深度学习网络难以稳定地恢复高质量的绝对相位信息。未来的研究将致力于探索更先进的网络结构，并尝试将更符合条纹投影物理特性的模型融入深度学习框架中，以提升网络对复杂表面反射特性的适应能力。通过优化模型结构与参数设计，有望在减少网络复杂度或采用纯端到端方式的前提下，有效应对高反光等挑战，实现更高速、更高鲁棒性以及更高精度的三维形状测量。

参考文献

[1] HE K, LIU X, LIU J, et al. A multitask learning-based neural network for defect detection on textured surfaces under weak supervision[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-14.

[2] YU Jiajie, ZHOU Jianping, XUE Ruilei, et al. Weld surface quality detection based on structured light and illumination

model[J]. Chinese Journal of Lasers, 2022, 49(16): 170-178.
余佳杰,周建平,薛瑞雷,等.基于结构光视觉和光照模型的焊缝表面质量检测[J].中国激光,2022,49(16):170-178.
[3] WU Fupei, PENG Junlong, YE Weilin, et al. A printed circuit board three-dimensional reconstruction method based on fringe phase characteristics[J]. Acta Optica Sinica, 2024, 44(14): 152-161.
吴福培,彭俊龙,叶玮琳,等.一种基于条纹相位特征的印刷电路板三维重建方法[J].光学学报,2024,44(14):152-161.
[4] GENG J. Structured-light 3D surface imaging: a tutorial[J]. Advances in Optics and Photonics, 2011, 3(2): 128-160.
[5] 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.
[6] ZHANG Wei, CHEN Yuchong, YAO Pengcheng, et al. Three-dimensional measurement of complex textured objects based on bidirectional fringe projection[J]. Acta Optica Sinica, 2024, 44(9): 101-112.
张巍,陈玉翀,姚鹏程,等.基于双向条纹投影的复杂纹理物体三维测量[J].光学学报,2024,44(9):101-112.
[7] XIA R, ZHAO J, ZHANG T, et al. Detection method of manufacturing defects on aircraft surface based on fringe projection[J]. Optik, 2020, 208: 164332.
[8] GUO Wenbo, ZHANG Qican, WU Zhoujie. Real-time three dimensional imaging technique based on phase-shift fringe analysis: a review[J]. Laser & Optoelectronics Progress, 2021, 58(8): 0800001.
郭文博,张启灿,吴周杰.基于相移条纹分析的实时三维成像技术发展综述[J].激光与光电子学进展,2021,58(8):0800001.
[9] VAN DER JEUGHT S, DIRCKX J J J. Real-time structured light profilometry: a review[J]. Optics and Lasers in Engineering, 2016, 87: 18-31.
[10] WU Z, ZHANG Q. High-speed 3D topography measurement based on fringe projection: a review[J]. Laser & Optoelectronics Progress, 2023, 60(8): 0811001.
[11] LI Z, WANG J, JI Y, et al. Fringe projection profilometry based on deep learning phase demodulation combined with temporal phase unwrapping[J]. Applied Physics B, 2024, 130(12): 213.
[12] NGUYEN A H, SUN B, LI C Q, et al. Different structured-light patterns in single-shot 2D-to-3D image conversion using deep learning[J]. Applied Optics, 2022, 61(34): 10105-10115.
[13] FENG S, CHEN Q, GU G, et al. Fringe pattern analysis using deep learning[J]. Advanced Photonics, 2019, 1(2): 025001.
[14] YIN W, CHE Y, LI X, et al. Physics-informed deep learning for fringe pattern analysis[J]. Opto-Electronic Advances, 2024, 7(1): 230034.
[15] LI Y, QIAN J, FENG S, et al. Deep-learning-enabled dual-frequency composite fringe projection profilometry for single-shot absolute 3D shape measurement[J]. Opto-Electronic Advances, 2022, 5(5): 210021.
[16] SHI J, ZHU X, WANG H, et al. Label enhanced and patch based deep learning for phase retrieval from single frame fringe pattern in fringe projection 3D measurement[J]. Optics Express, 2019, 27(20): 28929-28943.
[17] QIAN J, FENG S, LI Y, et al. Single-shot absolute 3D shape measurement with deep-learning-based color fringe projection profilometry[J]. Optics Letters, 2020, 45(7): 1842-1845.
[18] LI J, ZHANG K, LUO L, et al. Deep-learning-enabled single-shot fringe projection profilometry based on inner shifting-phase encoding[J]. Optics Express, 2025, 33(23): 49530-49550.
[19] NGUYEN A H, LY K L, LAM V K, et al. Generalized fringe-to-phase framework for single-shot 3D reconstruction integrating structured light with deep learning[J]. Sensors, 2023, 23(9): 4209.
[20] TAN H, XU Y, ZHANG C, et al. A Y-shaped network based single-shot absolute phase recovery method for fringe projection profilometry[J]. Measurement Science and Technology, 2023, 35(3): 035203.
[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, 2015: 234-241.
[22] ZUO C, HUANG L, ZHANG M, et al. Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review[J]. Optics and Lasers in Engineering, 2016, 85: 84-103.
[23] LI W, SUN H, LI F, et al. A complementary binary code based phase unwrapping method[J]. Optoelectronics Letters, 2024, 20(4): 228-233.
[24] YU H, CHEN X, ZHANG Z, et al. Dynamic 3-D measurement based on fringe-to-fringe transformation using deep learning[J]. Optics Express, 2020, 28(7): 9405-9418.

A Phase Recovery Method for Fringe Projection Profilometry Based on Multi-task Networks

LIU Guijie[1], LIU Wulang[1], LI Mingfeng[1], HUANG Chaoguang[2], HUANG Wenbin[3], MA Jianpin[1], LI Wenjie[1,3]

(1 *Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology, School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology, Guilin* 541004, *China*)

(2 *Guilin Kevin Peter Technology Co., Ltd., Guilin* 541001, *China*)

(3 *Engineering Research Center of Digital Imaging and Display, Ministry of Education, Soochow University, Suzhou* 215006, *China*)

Abstract: 3D reconstruction techniques have been extensively applied across a wide range of fields in recent years, including medical imaging, robotic navigation, virtual and augmented reality, 3D animation modeling, and online product inspection. Among these techniques, Fringe Projection Profilometry (FPP) has attracted significant attention owing to its non-contact measurement capability, full-field acquisition, and high spatial resolution. These advantages have led to its widespread adoption in industrial inspection, cultural heritage preservation, biomedical applications, and reverse engineering. Within an FPP system, phase recovery constitutes a fundamental and indispensable step, as both the accuracy and computational efficiency of phase estimation directly determine the quality of the reconstructed 3D surface and the overall system performance. Consequently, the development of fast, accurate, and robust phase recovery methods remains a central research topic in fringe projection profilometry.

Traditional phase retrieval techniques mainly include Fourier Transform Profilometry (FTP) and multi-step Phase-Shifting Profilometry (PSP). Multi-step phase-shifting methods achieve high phase accuracy by projecting and capturing multiple phase-shifted fringe patterns; however, their reliance on multiple frames significantly restricts their applicability in dynamic scenes or high-speed measurement scenarios. In contrast, Fourier transform-based methods can extract phase information from a single fringe image, offering improved measurement efficiency. Nevertheless, their performance tends to degrade considerably when dealing with complex surface geometries, depth discontinuities, severe noise, or strong surface reflectivity, resulting in reduced accuracy and robustness. In recent years, researchers have increasingly integrated deep learning with phase extraction and phase unwrapping processes, achieving high reconstruction accuracy while significantly reducing the number of required projection patterns. Compared with conventional analytical approaches, deep learning-based methods exhibit superior capability in handling noise, nonlinear distortions, and surface discontinuities. Despite these advantages, existing deep learning-based absolute phase recovery methods still suffer from several limitations. Existing deep learning-based absolute phase recovery methods mainly fall into two categories. The first predicts the numerator and denominator terms of wrapped phase at three different frequencies separately, then computes the absolute phase using multi-frequency or number-theoretic methods. The second employs either a dual-network or dual-decoding architecture to separately predict the numerator and denominator terms of the high-frequency wrapped phase along with the fringe order, thereby obtaining the absolute phase. The former suffers from error accumulation during multi-frequency unwrapping, leading to significant inaccuracies and poor stability. The latter incurs high computational complexity and low inference efficiency due to the multi-network or dual-decoder design.

Aiming to address the issues of error accumulation and high model complexity inherent in existing methods, this paper proposes a novel multi-task phase recovery framework based on GD-UNet (UNet with an Information Gather-and-Distribute Mechanism). The proposed method enables simultaneous prediction of the wrapped phase and fringe order within a single network, thereby allowing direct recovery of the absolute phase. Built upon the classical UNet architecture, the proposed model incorporates residual modules to enhance feature extraction capability and improve training stability. In addition, by integrating an information gather-and-distribute mechanism, the network supports multi-task learning and directly outputs the numerator and denominator of the wrapped phase as well as the corresponding fringe order.

This unified design eliminates the need for multiple networks, effectively reducing computational complexity and inference time. Furthermore, to enhance the robustness of fringe order prediction-particularly in challenging regions such as object boundaries, sharp depth discontinuities, and highly reflective surfaces-a fringe order correction strategy based on Connected Domain Segmentation (CDS) of the wrapped phase is introduced. The proposed CDS-based correction method exploits the spatial continuity of the wrapped phase, under the assumption that all pixels within the same connected domain theoretically share an identical fringe order. Since prediction errors tend to occur more frequently near domain boundaries, the final fringe order for each connected region is determined through majority voting, thereby effectively suppressing local misclassifications. This strategy significantly improves the stability and accuracy of fringe order estimation without introducing additional computational burden. Extensive experiments are conducted to evaluate the performance of the proposed method under various conditions, including different surface materials, complex geometries, and discontinuous scenes. Both quantitative and qualitative comparisons with state-of-the-art methods demonstrate that the proposed GD-UNet-based framework achieves superior phase recovery accuracy while maintaining lower model complexity and faster inference speed.

The experimental results indicate that the proposed approach effectively mitigates error accumulation, enhances robustness against noise and surface reflectivity, and exhibits strong generalization capability across diverse measurement scenarios. In conclusion, this paper aims to achieve stable and high-precision phase recovery from a single fringe image in complex scenes containing large surface discontinuities or isolated objects through a unified network model. A single-frame phase extraction method is presented, in which only one fringe image is required, and a single network simultaneously predicts the numerator and denominator of the wrapped phase as well as the fringe order map. The proposed approach demonstrates clear advantages in both accuracy and efficiency. Comprehensive experimental evaluations confirm that the method achieves excellent measurement accuracy and strong robustness in challenging scenarios, including complex surface reconstruction, high noise interference, and multi-material object measurement.

Key words: Fringe projection profilometry; GD-UNet network; Information Gather and Distribute Mechanism; Phase recovery; Connected domain segmentation

OCIS Codes: 150.1708; 150.3040; 150.3045

CSTR: 32255.14.gzxb20265503.0311001

Foundation item: National Natural Science Foundation of China Regional Fund (52565058), Guangxi Natural Science Foundation (2025GXNSFAA069774), Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology (AD25069080), Engineering Research Center of Digital Imaging and Display, Ministry of Edu-cation, Soochow University (SDGC2319)