arXiv:2601.02572v1 [physics.optics] 5 Jan 2026

基于深度学习的单帧复合条纹投影轮廓术与像素级不确定性量化

XIAN开军 KONG,[1,*] QING康 BAO,[2] TIBEBEYALEW,[1] GERARDOADESSO[3] 和 SAMANTAPIANO,[1,*]

[1]诺丁汉大学工程学院制造计量团队，英国诺丁汉[2]西安交通大学机械工程学院，中国陕西西安[3]诺丁汉大学量子非平衡系统数学与理论物理中心及数学科学学院，英国诺丁汉[*]
xiangjun.kong@nottingham.ac.uk **samanta.piano@nottingham.ac.uk*

摘要：随着先进制造对高速三维测量的需求日益增长，光学计量算法必须在动态条件下实现高精度和高鲁棒性。条纹投影轮廓术（FPP）具有高精度，但包裹相位存在 2π 歧义，因此传统的绝对相位恢复通常依赖多幅编码图案，牺牲了时间分辨率。基于深度学习的复合条纹投影轮廓术（CFPP）为从复合条纹中进行单帧相位恢复展现出良好前景，但其可解释性有限，因此在缺少真实值的情况下，难以评估重建可靠性或追溯误差来源。为此，我们提出 HSURE-CFPP（面向 CFPP 的异方差快照集成不确定性感知比值估计）。HSURE-CFPP 利用异方差快照集成网络预测用于包裹相位计算的分子与分母“比值”，从单幅复合条纹实现超高速三维成像，并生成像素级不确定性图，用于置信度评估和不可靠区域识别。具体而言，异方差似然联合估计像素级噪声方差，以表征数据不确定性；快照集成则通过快照间的离散程度量化模型不确定性，从而得到总预测不确定性，作为一种可解释的可靠性度量。静态和动态场景下的实验表明，HSURE-CFPP 能够以高速度实现高精度重建，且预测的不确定性与重建误差高度相关，为基于深度学习的条纹投影轮廓术提供了可部署的质量评估机制。

1. 引言

受制造业在线缺陷检测等高速应用的推动，市场对实时三维测量的需求日益增长，要求在不牺牲精度的前提下提高吞吐量 [1, 2]。与此同时，在实际部署中，评估重建可靠性并追溯误差来源日益重要，尤其是在无法获取真实三维数据的情况下。在此背景下，条纹投影轮廓术（FPP）凭借非接触式操作、高精度和低成本，已成为广泛应用的三维形状测量技术 [3–5]。在 FPP 中，通过系统标定和条纹分析恢复三维形状，其中相位提取和相位展开是关键步骤 [4, 5]。相位展开是影响速度和精度的主要瓶颈，因为反正切运算会产生具有固有 2π 不连续性的包裹相位。常用方法包括空间相位展开（SPU）和时间相位展开（TPU） [6, 7]：TPU 通常精度更高，且对噪声和不连续性的鲁棒性更强，但需要额外投影，因此会降低采样效率 [4]。此外，传统 FPP 大多只报告确定性高度，而没有明确的不确定性模型，这限制了误差来源预算以及可追溯、风险感知的符合性判定。在计量学中，不确定度是测量结果不可或缺的一部分；对于 FPP，这种不确定度

会因表面相关成像效应和偶发的解包裹失败而呈现显著的空间变化，因此需要像素级不确定性图，以支持可靠性感知的三维重建与检测。

条纹投影轮廓术中有两种技术可用于快速三维测量，即傅里叶变换轮廓术（FTP）[8] 和相移轮廓术（PSP）[9]。由于 FTP 仅使用单幅条纹图案，通常采用空间相位展开（SPU），但对于表面彼此孤立、条纹级次无法传播且绝对相位存在歧义的物体，SPU 会失效。为提高相位解包裹精度，研究人员探索了利用复合条纹图案进行空间频率复用（FM）[10, 11], , 将多个具有不同空间载波频率的正弦波编码到单幅图案中。然而，在 FTP 中，陡峭的表面梯度和高频成分可能导致频谱泄漏，使基波分量与相邻频谱分量发生重叠；在复合条纹中尤其如此，从而降低带通滤波的效果。一种改进的复合条纹将两个具有正交载波频率的正弦图案复用 [12], , 以减少频谱重叠并实现更可靠的时间相位展开（TPU）。使用高速相机和投影仪时，FTP 可实现超高速测量；但系统成本高，而且通常需要在帧率与信噪比之间进行权衡。此外，在 FTP 中，由于不可避免的频谱混叠，频域滤波无法干净地分离出基波分量，即使采用先进滤波器（如正交椭圆滤波器）也是如此，从而降低了测量精度 [13]。

与 FTP 相比，相移轮廓术通常具有更高的精度，因为它采用多幅相移条纹，从而抑制随机噪声并减轻反射率效应 [5]。为实现稳健解码，相移轮廓术通常采用时间相位展开，但这会牺牲实时性能。随着硬件的改进，千赫兹级数字光处理（DLP）投影可减轻物体运动在相移轮廓术中引起的图像伪影 [14]。这适用于单比特二值离焦方法，该方法通过投影仪离焦形成正弦条纹，不过由此产生的低图像对比度可能降低相位精度。为避免多帧投影，研究人员探索了单帧复合图案相移轮廓术。例如，三片式 RGB 相机可捕获一种颜色编码图案，其中三个条纹分量被复用到颜色通道中 [15]。然而，颜色串扰和非线性传感器响应会降低相位提取的精度。此外，散斑嵌入式条纹方法将类散斑信号嵌入三幅正弦相移条纹中，以实现绝对相位恢复 [16]。尽管该方法速度快，但对噪声敏感，而且空间分辨率与测量精度之间存在固有权衡。

近年来，基于数据驱动的深度学习方法，尤其是卷积神经网络，被引入条纹投影轮廓术，以缓解相位模糊和噪声问题 [1, 17–20]。与传统方法（如 FTP）相比，这些方法能够从单频条纹图案中更准确地恢复绝对相位 [1, 19, 20]。然而，直接根据单频条纹回归包裹相位、条纹级次或解包裹相位，所提供的物理约束有限，会增加解码难度，并可能降低稳定性和精度。此外，对预测的包裹相位进行空间相位展开（SPU）可能会进一步降低精度。因此，为了在保持相位解调精度的同时实现实时测量，研究者开始探索基于深度学习的复合条纹投影轮廓术（CFPP），因为复合条纹比单频条纹提供了更丰富的线索 [21–26]。其中，卷积神经网络辅助的彩色 CFPP 可以减轻通道间串扰 [25, 26], , 但在彩色表面上并不可靠，因为光谱反射率会扭曲逐通道条纹对比度。为避免对颜色的依赖，研究者采用频率复用 CFPP 来恢复绝对相位 [21–24]。然而，将多个高频分量复用到单幅图案中，会因调制传递函数（MTF）的衰减而降低条纹对比度；当各分量频率间隔较小时，还会加剧光谱串扰。已有研究提出贝叶斯卷积神经网络，可从单频条纹图案中恢复相位，同时提供用于量化置信度的不确定性图，从而提升可解释性并支持可靠性评估 [27]。然而，这种不确定性分析仅适用于单频推断，且从单幅

条纹中估计的绝对相位仍不可靠。采用蒙特卡罗（MC）方法的贝叶斯卷积神经网络需要进行多次随机前向传播，这会增加延迟和计算量，并使不确定性对丢弃率等超参数较为敏感。

为提升实时 CFPP 的精度，同时提供具备不确定性感知能力且可解释的像素级预测，我们提出了 HSURE-CFPP（用于 CFPP 的异方差快照集成不确定性比值估计）。此处，“比值”指用于计算包裹相位的逐像素分子—分母表示（N, D）。HSURE-CFPP 采用异方差快照集成 U-Net（HSU-Net），联合预测均值（N, D）及对应的像素级不确定性图。据我们所知，这是首次展示采用异方差似然的快照集成卷积神经网络，能够从复合条纹图案中进行单次相位解调，同时估计两个输出的像素级不确定性。实验表明，HSURE-CFPP 能够准确重建瞬态场景，且生成的不确定性图与重建误差相关，因此无需真实参考数据即可进行置信度评估。此外，与贝叶斯卷积神经网络替代方案相比，HSURE-CFPP 的推理成本更低，部署也更简单。

2. 结果

在 HSURE-CFPP 中，将一种在频谱域设计的复合图案（称为分层频率复合条纹 HFCF）投射到物体上。随后，使用经过异方差似然训练的 HSU-Net 对采集到的条纹图案进行处理，并通过集成来自不同训练轮次的检查点进行部署，从而实现实时重建并给出逐像素不确定性估计。HFCF 嵌入了高、中、低频调制，以减轻频谱混叠并实现单帧 TPU，从而生成高频绝对相位图。由于高频分量具有更高的深度灵敏度，因此能够保持空间分辨率。输出的不确定性包括数据不确定性和模型不确定性。数据不确定性也称为偶然不确定性，反映了测量噪声以及条纹图案中固有的模糊性。模型不确定性也称为认知不确定性，源于训练数据有限或泛化误差；对于分布外（OOD）或罕见表面（即在材料、几何形状或外观方面与训练分布显著不同的样本），模型不确定性往往较大。这两项合并后得到总不确定性，并通过一阶误差传播计算相位不确定性，从而得到可与相位误差直接比较的估计值。

数据（偶然）不确定性编码在预测的逐像素对数方差中，而模型（认知）不确定性则通过不同快照预测结果的方差进行量化。图 1 展示了我们方法的工作流程。我们使用 $(\hat{\mu},\ \hat{\sigma}^2)$ 表示单个快照预测的均值和方差，使用 $(\bar{\mu},\ \bar{\sigma}^2)$ 表示对应的集成平均估计值。HSU-Net 在相同的 HFCF 输入数据集上训练三次，每次使用不同的相位展开真实值目标，从而得到三个训练完成的网络实例。对于每个 HFCF 输入，训练后的网络会输出四个量：各调制频率下包裹相位的分子均值 $\bar{\mu}_N$ 和分母均值$\bar{\mu}_D$，以及用于表征逐像素噪声的对应对数方差 $\log\hat{\sigma}_N^2$ 和 $\log\hat{\sigma}_D^2$。其中，N 和 D 分别表示包裹相位的分子和分母。对于频率为 $f_{\text{mod的第}\,i}$个调制条纹（$i = 1,\ 2,\ 3$ 分别对应高、中、低频），将 atan2 应用于预测的 $\bar{\mu}_N$ 和 $\bar{\mu}_D$即可得到包裹相位。随后，利用这三个包裹相位确定条纹级次，并通过 TPU 展开相位。最后，结合标定参数，通过三角测量得到三维点云。图 1 下方展示了整体重建流程。不确定性在方差域中进行量化。对逐像素对数方差取指数即可得到数据方差，而模型方差则根据快照集成中各快照预测结果的方差估计。将两者在方差域中相加，得到 N 和 D的总预测方差，随后通过以下方式将其传播至包裹相位

在 $\bar{\mu}_N$ 和 $\bar{\mu}_D$处通过 atan2 进行一阶误差传播，得到包裹相位的方差 $\hat{\sigma}_\phi$。因此，对每一对 N 和 D 所得的不确定性估计能够更细致、全面地评估预测可靠性，从而有助于定位并判定不同频率及网络输出通道中的潜在误差来源。

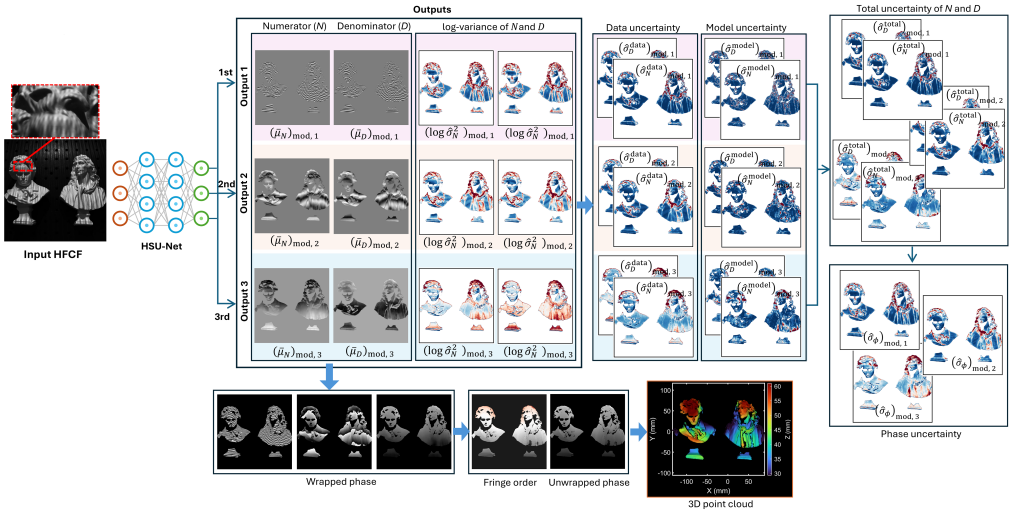

图 1。HSURE-CFPP方法的工作流程。将一个HFCF输入训练好的HSU-Net，以输出四个量，包括分子均值 $\bar{\mu}_N$、分母均值 $\bar{\mu}_D$ 以及相应的对数方差 $\log\hat{\sigma}_N^2$ 和 $\log\hat{\sigma}_D^2$。利用这些均值计算包裹相位、执行相位展开并重建三维点云。利用这些对数方差估计数据σ^{data}和模型σ^{model}不确定性，随后将二者结合以获得总不确定性σ^{total}和相位不确定性 σ_ϕ。

*2.1. HSU-Net CFPP*的原理

2.1.1. HFCF的生成

在HFCF中，三个具有不同空间频率且相位仅沿垂直方向 y^p 变化的调制条纹图案被编码为一个高频载波条纹图案。具体而言，将每个垂直调制项分别与沿 x^p方向具有不同空间频率的水平载波相乘，最后将三个乘积叠加，形成一个复合条纹图案。投影仪坐标（x^p， y^p）中的第 i个投影调制条纹可写为

$$I^p_{\text{mod},i}(x^p, y^p) = a + b\,\cos\left(2\pi f^p_{\text{mod},i}\, y^p\right), \quad i \in \{1,2,3\}, \tag{1}$$

其中， a和 b是投影常数，用于将 I_{mod_i}偏移为非负值。 y^p 轴为相位方向，对应垂直方向。调制频率呈层级关系，$f_{\text{mod},\,1} > f_{\text{mod},\,2} > f_{\text{mod},\,3}$；在我们的实现中，$f_{\text{mod},\,3}$ 对应于投影仪分辨率上的一个周期。单帧 HFCF 表示为

$$I^p_{\text{HFCF}}(x^p, y^p) = \sum_{i=1}^{3} I^p_{\text{mod},i}(x^p, y^p)\,\cos\left(2\pi f^p_{\text{car},i}\, x^p\right), \tag{2}$$

其中，载波频率 $f^p_{\text{car},\,i}$ 沿 x^p 方向排列，并按从高到低的顺序排列为 $f_{\text{car},\,1} > f_{\text{car},\,2} > f_{\text{car},\,3}$。为便于说明，图2(a)展示了 HFCF 的构造

由三个调制条纹和三个载波条纹构成；生成的合成条纹如图2(b)所示。图2(c)展示了 HFCF 的二维傅里叶频谱。 $F(f_{\text{mod},0})$ 为零频项，对应于条纹的背景强度。三对一阶频谱分量分别记为 $F(f^{p}_{\text{调制},1})$、$F(f^{p}_{\text{调制},2})$ 和 $F(f^{p}_{\text{调制},3})$，它们源自三个调制条纹分量，并包含相位信息。$f^{p}_{\text{载波},i}$ 和 $f^{p}_{\text{调制},i}$经过精心设计，使这些一阶谱瓣沿水平频率轴（通过不同的载波频率）和垂直频率轴（通过不同的调制频率）彼此充分分离。这可避免频谱混叠，并使后续的基于网络的学习和特征提取更加稳定可靠。考虑到投影仪和相机的强度容量，需要将投射条纹的强度归一化至 $[0, 255]$，以将条纹灰度映射到投影-相机系统 [4] 的有效动态范围内，方法是

$$(I^{p}_{\text{HFCF}})_{\text{norm}} = \frac{I^{p}_{\text{HFCF}} - I_{\min}}{I_{\max} - I_{\min}} \times 255 \tag{3}$$

其中，$(I^{p}_{\text{HFCF}})_{\text{norm}}$ 为归一化后的复合条纹，本文后续将其表示为 I^{p}_{HFCF}；$I_{\min}$ 和 $I_{\max}$ 分别为原始 HFCF 的最小强度值和最大强度值。

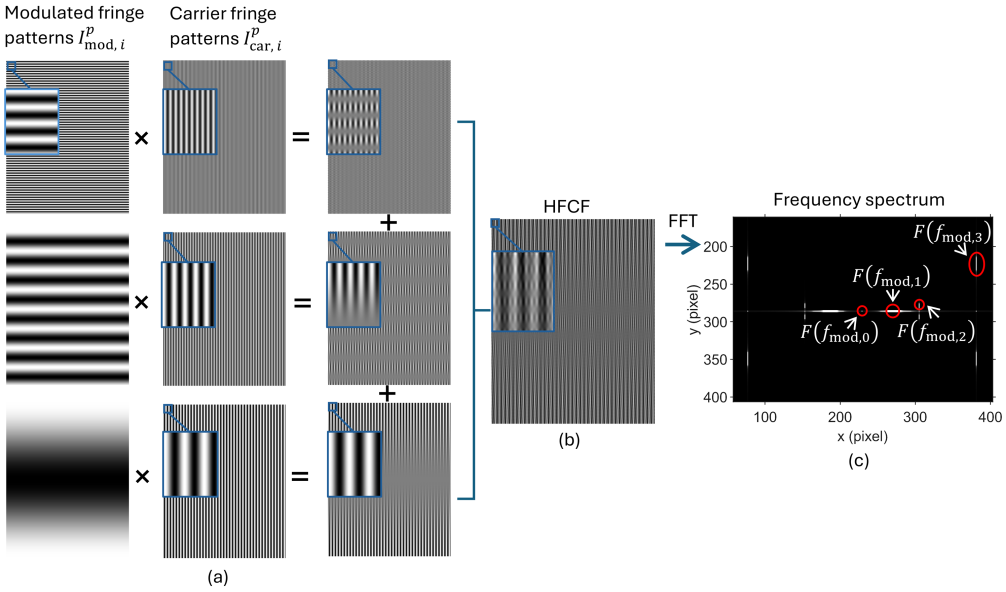

图 2. HFCF 的构造及其傅里叶频谱。(a) 将三个具有不同垂直频率的调制条纹分别与三个具有不同水平频率的载波条纹相乘，再将得到的三个条纹图案相加，得到 (b) 中的复合条纹图案（即 HFCF）。(a) 和 (b) 中放大的区域展示了条纹图案的结构。使用三个调制条纹是为了满足稳健相位展开的 TPU 准则。这些条纹图案的尺寸由投影仪的规格决定。(c) HFCF 的二维傅里叶频谱，显示了各调制分量的频率分布。

使用 DLP投影仪将 HFCF 照射到被测物体后，采集图像可表示为

$$I^{c}_{\text{HFCF}}(x^{c}, y^{c}) = R(x^{c}, y^{c})[I^{p}_{\text{HFCF}}(x^{p}, y^{p}) + \beta_1(x^{c}, y^{c})] + \beta_2(x^{c}, y^{c}) \tag{4}$$

其中，$(x^c,\ y^c)$ 是相机的像素坐标，$R(x^c,\ y^c)$ 表示测量区域的反射率变化。$\beta_1(x^c,\ y^c)$ 对物体表面的背景照明进行建模，而 $\beta_2(x^c,\ y^c)$ 则表示相机处的加性环境光和传感器噪声。变形调制条纹 $I^c{}_{\mathrm{mod},\ i}(x^c,\ y^c)$ 为

$$\begin{aligned} I^c_{\mathrm{mod},i}(x^c,y^c) &= R(x^c,y^c)[I^p_{\mathrm{mod},i}(x^p,y^p)+\beta_1(x^c,y^c)]+\beta_2(x^c,y^c), \\ &= A(x^c,y^c)+B(x^c,y^c)\cos\big(\phi(x^c,y^c)\big) \end{aligned} \tag{5}$$

其中，A^p 和 B^p 分别表示 $I^c{}_{\mathrm{mod},\ i}(x^c,\ y^c)$的平均强度和调制幅度。$\phi(x^c,\ y^c)$是编码高度信息且需要计算的畸变相位。载波条纹引入空间频移，使调制分量彼此分离，并与傅里叶域中的零频率分离，从而使 HSU-Net 能够提取具有信息量的条纹内容。调制频率和载波频率经过联合设计。三个调制条纹图案用于 TPU，而最高调制频率$f_{\mathrm{mod},\ 1}$处的解包裹相位则用于表面重建，因为其对微小高度变化具有更高的敏感性。频率比根据投影仪分辨率及其在空间频率谱中的位置进行选择，如图2(c)所示，从而确保在噪声、饱和、非均匀反射率和遮挡条件下仍能可靠地展开相位。

2.1.2. HSU-CFPP重建原理

我们引入了一种 HSU-Net（详见第4节），直接从捕获的 HFCF图案中回归包裹相位的正弦和余弦分量，随后利用这些分量进行相位解调。根据公式（5），第 i个调制频率 ϕ_i $f^c{}_{\mathrm{mod},\ i}$ 处的包裹相位可通过其正弦通道和余弦通道经四象限反正切恢复：

$$\phi_i(x^c,y^c)=\operatorname{atan2}\big(N_i(x^c,y^c),\ D_i(x^c,y^c)\big), \tag{6}$$

其中，$N_i(x^c,y^c)$ 和 $D_i(x^c,y^c)$ 分别表示 $\phi_i(x^c,y^c)$ 的正弦通道和余弦通道。在理想的无噪声情况下，它们分别与包裹相位的真实正弦值和余弦值成正比，即 $N_i(x^c,y^c)\ \propto \sin\big(\phi_i(x^c,y^c)\big)$ 和 $D_i(x^c,y^c)\ \propto \cos\big(\phi_i(x^c,y^c)\big)$。

HSU-Net $f_{\boldsymbol{\theta}}$ 以捕获的 HFCF 图案 I^c_{HFCF} 作为输入，并针对每个像素 (x^c,y^c) 预测这两个通道的概率估计。具体而言，网络输出每个通道的均值和对数方差：

$$\begin{bmatrix} \bar{\mu}_{N_i}(x^c,y^c) \\ \bar{\mu}_{D_i}(x^c,y^c) \\ \log\hat{\sigma}^2_{N_i}(x^c,y^c) \\ \log\hat{\sigma}^2_{D_i}(x^c,y^c) \end{bmatrix} = f_{\boldsymbol{\theta}}\big(I^c_{\mathrm{HFCF}}\big)\big|_{(x^c,y^c)}, \tag{7}$$

其中，$\boldsymbol{\theta}$ 表示在训练数据上优化的网络可学习参数。μ_N 和 μ_D 分别表示分子和分母的预测均值。$\log\hat{\sigma}^2_N$和$\log\hat{\sigma}^2_D$ 是异方差高斯似然的逐像素对数方差参数，不参与重建。随后，利用预测的 $\bar{\mu}_{N_i}$ 和 $\bar{\mu}_{D_i}$ 计算每个调制频率下的包裹相位，再进行三频率时间相位展开。借助较低频率，对最高频率的包裹相位 ϕ_3 进行展开。一般而言，绝对相位 $\Phi_i(x,\ y)$与其对应的包裹相位 $\phi_i(x,\ y)$之间相差一个未知的 2π整数倍：

$$\Phi_i(x,y)=\phi_i(x,y)+2\pi\,k_i(x,y),\qquad k_i(x,y)\in\mathbb{Z}. \tag{8}$$

时间相位展开按频率从低到高依次确定整数映射 k_i。假设最低频率在整个视场内均未发生包裹：

$$\Phi_1 = \phi_1, \qquad k_1 = 0. \tag{9}$$

对于 $i = 2,\ 3$，中频和高频对应的整数映射 k_i 计算如下：

$$k_i = \mathrm{round}\left(\frac{f_i}{f_{i-1}} \frac{\Phi_{i-1}}{2\pi} - \frac{\phi_i}{2\pi} \right) \tag{10}$$

最终的绝对相位为 $\Phi_3(x,\ y)$，通过投影仪-相机标定将其转换为三维点云。为了训练 HSU-Net，我们使用多步相移轮廓术为输入准备对应的真实值目标，以获得高精度的 N_i 和 D_i。随后进行逐像素监督，促使预测分量在统计意义上与真实值标签相匹配；具体做法是对每个像素使用异方差损失（具体损失函数及权重见第3.5节）。

2.1.3. HFCF 中的频率设计

尽管将复合条纹与深度学习相结合，相较于单条纹方法能够同时提升速度和精度，但在单帧中嵌入多个频率也会带来风险。较大的高度变化和器件非线性可能引发光谱串扰，光学系统会衰减高频分量，且还存在采样限制。这些因素可能降低训练稳定性。精心设计载波频率和调制频率可以缓解这些问题。载波频率应彼此充分分离，以便在频域中分离调制分量；同时，在系统限制允许的范围内应尽可能设得较高，从而将这些分量移离零频率。调制频率则应根据相位展开策略进行选择。本研究采用三频时间相位展开方案，因此选用低、中、高三种载波。与采用三个间隔很近的高频率的外差法相比， [7], 这种分层频率设计可减少频谱重叠，并减轻高频衰减的影响，因为它只使用一条高频条纹。为追求更高的高度分辨率，条纹空间频率不能任意提高。最大可用空间频率受以下因素共同限制：相机和投影仪的奈奎斯特采样限制、系统调制传递函数所规定的最低对比度要求，以及 [20]中定义的最大可测相位梯度。这四项限制共同构成：

$$f^{\mathrm{p}}_{\mathrm{min}} = \max\left\{ 2,\ 2\,\frac{\Delta x_{\mathrm{c}}}{\Delta x_{\mathrm{p}}},\ \frac{1}{\Delta x_{\mathrm{p}}\, f_{\mathrm{MTF}}},\ \frac{3\,|\nabla\phi|_{\max}}{2\pi} \right\}. \tag{11}$$

其中， f^{p} 表示沿条纹方向每个条纹周期所包含的投影仪像素数， $f^{\mathrm{p}}{}_{\mathrm{min}}$ 是理论上确保无混叠且具有足够 MTF 对比度的最小值。 Δx^{c} 和Δx^{p} 分别表示相机和投影仪在物体上的有效采样间距。f_{MTF} 是确保条纹具有足够对比度的调制传递函数限制频率。 $|\nabla\phi|_{\max}$ 是最大可测相位梯度（更多详情请参阅补充文件）。采用这种频率选择策略，我们减少了频谱混叠并保留了对比度，从而能够测量表面梯度较陡的物体，进而有利于深度学习多频解调的训练。

2.1.4. 不确定性建模与估计

不确定性分析对于评估模型预测的可靠性至关重要。贝叶斯神经网络已被用于估计单频条纹的逐像素输出不确定性；然而，相位模糊会使相位展开变得脆弱，而近似的蒙特卡洛 dropout 可能会系统性地

误估不确定性，同时带来较高的推理成本。我们采用包含多个不同调制频率下相位分量的 HFCF 来解决相位模糊问题，从而使独立光谱观测的数量显著超过式（1）中的未知量。在实际应用中，传感器噪声、频谱混叠、表面反射率变化和条纹衰减仍可能导致不可靠的预测，因此恢复的相位可能以依赖数据的方式偏离真实值。为追踪和分析此类误差，我们采用 HSU-Net 从 HFCF 中恢复包裹相位分量 N 和 D，同时通过异方差建模和快照集成策略生成逐像素的模型不确定性和数据不确定性。

给定复合条纹 I_{HFCF}，HSU-Net 会为每个像素输出预测的 N 和 D，以及用于参数化异方差（偶然）噪声的逐输出对数方差参数。具体而言，网络会生成四个逐像素量：$\hat{\mu}_N$、$\hat{\mu}_D$、$\log\hat{\sigma}_N^2$、$\log\hat{\sigma}_D^2$。对于每个像素 $s \in S$，$(\hat{\mu}_{N,s}, \hat{\mu}_{D,s})$ 被视为 N 和 D的点估计，而 $(\hat{\sigma}^2_{N,s}, \hat{\sigma}^2_{D,s})$（其中 $\hat{\sigma}^2 = \exp(\log\hat{\sigma}^2)$）表示异方差（偶然）噪声方差。对于 $v \in \{N, D\}$，假设其服从条件高斯似然 $v_s^* \sim \mathrm{N}(\hat{\mu}_{v,s}, \hat{\sigma}^2_{v,s})$，并且给定网络权重和输入后，$N$和 D 被视为条件独立。逐像素负对数似然为

$$\mathcal{L}_{\mathrm{NLL},s} = \frac{(N_s^* - \hat{\mu}_{N,s})^2}{2\,\hat{\sigma}^2_{N,s}} + \frac{1}{2}\log\hat{\sigma}^2_{N,s} + \frac{(D_s^* - \hat{\mu}_{D,s})^2}{2\,\hat{\sigma}^2_{D,s}} + \frac{1}{2}\log\hat{\sigma}^2_{D,s}. \tag{12}$$

训练过程中，逐像素负对数似然在每幅图像的所有有效像素上计算并取平均，然后再对各个小批次取平均，因此该目标反映的是图像层面的性能，而非孤立像素的性能。负对数似然通过推动预测的逐像素方差与经验残差的幅值相匹配，促进不确定性校准。当预测方差被低估且出现较大残差时，逆方差加权会赋予更大的惩罚。当预测方差被高估、数值偏大时，对数方差项会对这种膨胀施加惩罚。因此，在预测准确的位置往往会产生较小的方差（例如 -7×10^{-3}），而在数据本身噪声较大的位置则会产生较大的方差（例如 -2×10^{-3}）。随后，对一种兼顾均值精度与方差学习的混合总损失进行优化：

$$\mathcal{L} = \alpha\,\mathcal{L}_{\mathrm{MSE}} + \beta\,\mathcal{L}_{\mathrm{NLL}}. \tag{13}$$

其中，$\mathrm{L}_{\mathrm{MSE}}$ 表示两个输出$\hat{\mu}_N$ 和$\hat{\mu}_D$的均方误差。其计算方式为 as

$$\mathcal{L}_{\mathrm{MSE}} = \frac{1}{|S|}\sum_{s\in S}\Big[(\hat{\mu}_{N,s} - N_s^*)^2 + (\hat{\mu}_{D,s} - D_s^*)^2\Big]. \tag{14}$$

系数 β 在训练早期轮次从 0 逐渐升至目标值（我们固定 α = 1.0 和 β = 1.0，并在前 30 个训练轮次中将 β 从 0 线性增加到 0.04），这样当 β 较小时，优化会侧重于 $\mathrm{L}_{\mathrm{MSE}}$，以稳定均值（$\hat{\mu}_N, \hat{\mu}_D$）；随着 β 增大，$\mathrm{L}_{\mathrm{NLL}}$ 会使预测方差与残差幅度相匹配，并抑制无意义的方差膨胀。我们通过快照集成量化认知不确定性，而无需学习权重后验分布。与贝叶斯神经网络不同，我们不学习权重后验分布或各层的丢弃率；丢弃法仅用作标准正则化器，并在测试阶段禁用。采用热重启的余弦退火会生成 T个确定性检查点 $\{\boldsymbol{\theta}(t)\}_{t=1}^{T}$。对于测试图像 x^*，检查点 t会得出 $\hat{\mu}_v^{(t)},_s = \mu_v(x^*, s;\ \boldsymbol{\theta}(t))$ 和 $\hat{\sigma}^2,_v{}^{(t)},_s = \sigma_v^2(x^*, s;\ \boldsymbol{\theta}(t))$。集成均值为：

$$\bar{\mu}_{v,s} = \frac{1}{T}\sum_{t=1}^{T}\hat{\mu}_{v,s}^{(t)}. \tag{15}$$

数据不确定性通过对各检查点预测的异方差方差取平均得到：

$$\bar{\sigma}_{v,s}^{2,\mathrm{data}} = \frac{1}{T}\sum_{t=1}^{T}\hat{\sigma}_{v,s}^{2,(t)}. \tag{16}$$

模型不确定性通过快照均值围绕集成均值的离散程度进行度量：

$$\bar{\sigma}_{v,s}^{2,\mathrm{model}} = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\mu}_{v,s}^{(t)} - \bar{\mu}_{v,s}\right)^2. \tag{17}$$

总预测方差通过对数据不确定性项和模型不确定性项求和得到：

$$\bar{\sigma}_{v,s}^{2,\mathrm{total}} = \bar{\sigma}_{v,s}^{2,\mathrm{data}} + \bar{\sigma}_{v,s}^{2,\mathrm{model}}. \tag{18}$$

其中，$\bar{\sigma}^{2,\mathrm{data}}_{v,s}$ 表示网络预测的固有数据噪声，而 $\bar{\sigma}^{2,\mathrm{model}}_{v,s}$ 度量不同检查点之间的离散程度，通常会随着数据覆盖度或模型容量的提升而减小。$\bar{\sigma}^{2,\mathrm{total}}_{v,s}$ 将这两个来源聚合为单一的可靠性度量。当任一分量较大时，该度量便会增大；我们将其传播至相位，并将其可视化为置信度图。为了将 N 和 D 的不确定性传播至包裹相位 ϕ，需围绕集成均值估计进行一阶误差传播。设$N_s = \bar{\mu}_{N,s}$ 和$D_s = \bar{\mu}_{D,s}$，则在（N_s，D_s）处的雅可比矩阵为

$$\frac{\partial \phi_s}{\partial N_s} = \frac{D_s}{N_s^2 + D_s^2}, \qquad \frac{\partial \phi_s}{\partial D_s} = -\frac{N_s}{N_s^2 + D_s^2}. \tag{19}$$

假设 N_s 和 D_s的扰动相互独立，则相位方差为：

$$\begin{aligned}\bar{\sigma}_{\phi,s} &= \left(\frac{\partial \phi_s}{\partial N_s}\right)^2 \bar{\sigma}_{N,s}^{2,\mathrm{total}} + \left(\frac{\partial \phi_s}{\partial D_s}\right)^2 \bar{\sigma}_{D,s}^{2,\mathrm{total}} \\ &= \frac{D_s^2\,\bar{\sigma}_{N,s}^{2,\mathrm{total}} + N_s^2\,\bar{\sigma}_{D,s}^{2,\mathrm{total}}}{(N_s^2 + D_s^2 + \varepsilon)^2}.\end{aligned} \tag{20}$$

在实际应用中，我们会在分母中加入一个极小的 ε（例如 -8×10^{-3}），以确保数值安全性，并屏蔽 $N_s^2+D_s^2$ 极小的像素。为便于可视化，标准差报告为相应方差的平方根。所有组合与传播均在方差域中进行。随机性（数据）不确定性的概率基础得以保留，而通过确定性检查点的快照集成近似表示模型平均所对应的认识（模型）不确定性。随后生成相位域置信度图（$\hat{\sigma}_\phi$），可用于质量评估和重建加权。图3展示了如何使用HSU-Net的输出进行相位展开和不确定性估计。

*2.2. HSU-Net*架构

HSU-Net 的构想受到深度学习技术近期在 FPP 中成功应用的启发，例如卷积神经网络 [17]。考虑到 HSU-Net 的目标是完成逐像素的图像到图像变换任务，图像分割网络（如基于 CNN 的 U-Net、ResNet 和 U-Net衍生网络）是较为理想的选择 [28, 29]。借鉴以往研究，我们综合考虑其功能性与实用性，采用标准的编码器–解码器U-Net作为骨干网络。如图 4 所示，基线U-Net 的输入为单通道灰度HFCF。编码器包含四个卷积块，每个卷积块后接一个步幅为 2 的 2×2 最大池化层。编码器的通道宽度为 {1, 64, 128, 256, 512}。每个卷积块由两个连续的 3×3 卷积组成（步幅为 1，

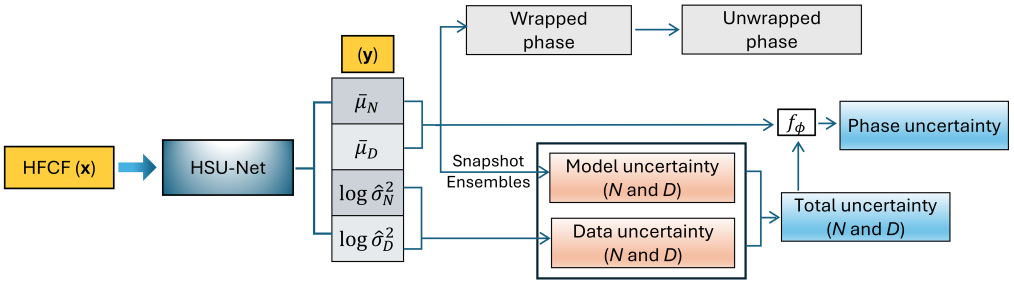

图 3. 使用 HSU-Net 预测结果的不确定性估计流程。将 HFCF 输入 HSU-Net，以预测均值 $\bar{\mu}_N$ 和 $\bar{\mu}_D$ 以及相应的对数方差，并据此计算数据不确定性。快照集成提供多个 $\hat{\mu}_N$ 和 $\hat{\mu}_D$ 的估计值，据此计算模型不确定性。随后，将数据不确定性与模型不确定性相结合，得到总不确定性和相位不确定性。

填充为 1），之后接批归一化和 ReLU 激活。为减轻过拟合，在两个卷积之间插入一个概率为 $p = 0.1$ 的 Dropout2d 层。瓶颈层采用相同的卷积块结构，通道数为 1024。解码器与编码器结构相对应。特征图通过步幅为 2 的 2×2 转置卷积进行上采样，并通过跳跃连接与尺寸相对应的编码器特征进行拼接。每次拼接后均接一个与编码器中相同的卷积块。解码器的通道宽度为 $\{1024, 512, 256, 128, 64\}$。最后，通过一个 1×1 卷积将特征映射到所需数量的输出通道。该架构通过跳跃连接保留高分辨率空间细节，同时利用较深层中较大的感受野来保证全局一致性 [29]。除非另有说明，我们采用零填充，以使每个卷积块内的特征图尺寸保持不变。

在基础编码器–解码器U-Net的基础上，我们附加一个异方差回归头，以联合预测 (N, D) 的逐像素均值和对数方差。在推理阶段，我们采用包含 $T = 4$ 个确定性检查点的快照集成 $\{\boldsymbol{\theta}_{(t)}\}_{t=1}^{T}$；检查点 t 生成 $(\hat{\mu}_N^{(t)}, \hat{\mu}_D^{(t)})$ 和 $(\log \hat{\sigma}^{2}{}_{,N}^{(t)}, \log \hat{\sigma}^{2}{}_{,D}^{(t)})$，最终的均值估计 $(\bar{\mu}_N, \bar{\mu}_D)$ 则通过对各检查点的均值取平均获得。随后，按照第 2.1.4 节所述，将各检查点的均值和方差进一步结合，以构建不确定性图。训练通过最小化由 L_{MSE} 和 L_{NLL} 的加权和构成的混合目标函数来完成，其权重分别为 α 和 β；这在保持 HSU-Net 点估计性能的同时，实现了异方差方差学习，并在空间变化噪声下获得了更好的不确定性校准。对于相位展开，我们使用相同的 HFCF 输入训练三个 HSU-Net 模型，但为其提供特定于频率的真实值目标（高频、中频和低频），从而得到三组学习到的网络参数 $\{\boldsymbol{\theta}_{\text{high}}, \boldsymbol{\theta}_{\text{mid}}, \boldsymbol{\theta}_{\text{low}}\}$，用于在每个调制频率下预测 $(\hat{\mu}_N, \hat{\mu}_D)$。

2.3. 数据集构建

为了将数据集输入 HSU-Net 进行网络训练，我们构建了一个包含输入数据和真实值的高质量数据集。输入数据由从各种物体采集的单帧条纹图案组成，相应的真实值为 N_i 和 D_i，它们是通过传统的多步相移轮廓术方法，利用相同的测量装置从同一物体的多帧条纹图案中计算得到的。这些多步相移轮廓术有助于计算包裹相位的高质量分子和分母，并进一步推导出绝对相位和三维点云，作为验证我们模型预测结果的基准。

在相移轮廓术中，对于第 i 个条纹频率，第 m 个相移变形条纹的强度

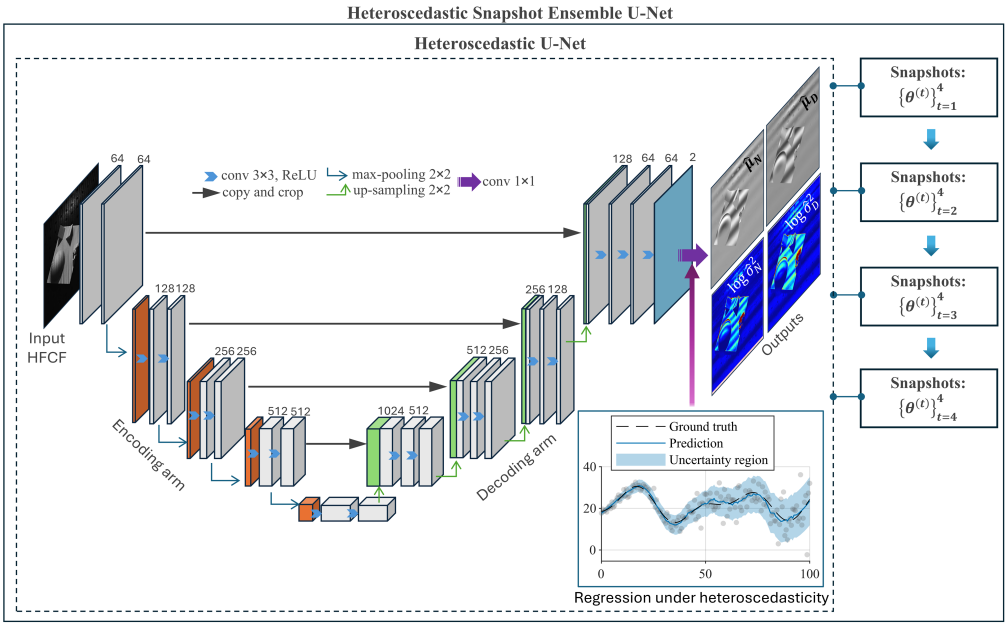

图 4. 我们的 HSU-Net 架构以 U-Net 骨干网络为基础，接收 HFCF 输入，并通过带有跳跃连接的编码器–解码器，利用异方差回归头联合预测 $(\hat{\mu}_N, \hat{\mu}_D)$ 和 $(\log\hat{\sigma}_N^2, \log\hat{\sigma}_D^2)$，从而在回归 (N, D) 的同时实现逐像素偶然不确定性估计；快照集成则汇总多个检查点，以提高稳健性并量化认知不确定性。

$I_{m,\ i}(x^c,\ y^c)$ 在像素 $(x^c,\ y^c)$ 处可建模为

$$I_{m,i}(x^c, y^c) = R(x^c, y^c)\left(A(x^c, y^c) + B(x^c, y^c)\cos\left[\phi_m(x^c, y^c) - \frac{2\pi m}{M-1}\right]\right). \tag{21}$$

其中，$\phi_i(x^c, y^c)$ 表示第 i 个条纹频率 f_i 下的包裹相位，$i \in \{1, 2, 3\}$。下标 m 表示相移条纹的索引，$m \in \{0, 1, 2, \ldots, M-1\}$。在我们的框架中，用于数据集生成的 PSP 程序所采用的条纹频率，与网络要恢复的 HFCF 条纹的空间调制频率 $f_{\mathrm{mod},i}$ 相同。匹配这组频率可确保监督信号与输入 HFCF 图案对应于同一底层相位场。此外，一旦涉及相位包裹，改变条纹频率就会改变相位的采样方式以及 2π 不连续点的位置。因此，通常无法通过对相位进行简单的线性缩放，将在一个频率下计算的标签转换为另一频率下的标签。对于第 i 个条纹频率，像素 (x^c, y^c) 处的包裹相位可由 M 步相移条纹计算得出：

$$\begin{aligned} N_i^{\mathrm{gt}}(x^c, y^c) &= \sum_{m=0}^{M-1} I_{m,i}(x^c, y^c)\ \sin\left(\frac{2\pi m}{M-1}\right), \\ D_i^{\mathrm{gt}}(x^c, y^c) &= \sum_{m=0}^{M-1} I_{m,i}(x^c, y^c)\ \cos\left(\frac{2\pi m}{M-1}\right), \end{aligned} \tag{22}$$

基于这两个分量，可得到像素 $(x^c,\ y^c)$ 处的包裹相位：

$$\phi_i(x^c, y^c) = \operatorname{atan2}\left(N_i^{\mathrm{gt}}(x^c, y^c),\ D_i^{\mathrm{gt}}(x^c, y^c)\right). \tag{23}$$

其中， $N_i^{gt}(x^c,\ y^c)$ 和 $D_i^{gt}(x^c,\ y^c)$ 分别作为包裹相位余弦通道和正弦通道的真实值标签。随后按照第 2.3 节所述，通过时间相位展开（TPU）计算绝对相位。较大的相移次数 M 能够有效抑制各种相位误差，例如相移误差和系统非线性误差。为在保持测量效率的同时降低相位计算误差，我们在三个调制频率下采用不同相移次数的相移轮廓术（PSP）： $M_1 = 9$、 $M_2 = 6$ 和 $M_3 = 6$ 分别对应高频、中频和低频条纹。由于用于三维重建的绝对相位由高频条纹得出，因此高频采用的相移次数多于中频和低频，以增强包裹相位估计的鲁棒性。总计，测量系统将对每个目标物体投射并同步采集 21 幅条纹图案和一幅单独的 HFCF，以构建训练数据集。监督式深度学习网络通常需要大量且多样的数据集，以确保训练可靠。因此，数据集中包含的不同尺寸和几何形状的被测物体会被随机放置并改变朝向，再由相机采集，以增加数据集的样本数量和多样性。此外，为提高模型的泛化能力，还会将若干单个物体随机组合并放置在一起，使其相互连接或彼此分离，以扩充数据集。最终生成的数据集包含 1000 个数据样本，并按 8:1:1 的比例划分为训练数据集、验证数据集和测试数据集。训练数据集和验证数据集用于确定网络权重，并在模型收敛时停止训练。测试数据集用于评估训练后模型的性能，且不会出现在其他数据集中。我们的数据集未采用图像旋转和图像翻转等数据集增强技术。因为这些操作违反标定的几何约束，相当于改变相机和投影仪的位置，导致训练不稳定并降低模型的泛化能力。采集图像仅采用裁剪操作，以定位感兴趣区域（ROI）并提高训练效率。

2.4. 实验

我们进行了全面的定量实验，以验证所提出的 HSU-Net CFPP 方法在动态三维形状测量与重建中的有效性和鲁棒性。对 HSU-Net 模型进行超参数调优后，我们使用测试数据集中的不同物体样本对所提方法进行测试和分析。通过将预测结果与真实值进行比较，并采用包裹相位和重建高度图的定量误差指标，对所提方法进行评估。随后，对用于描述模型不确定性和数据不确定性的不确定性图进行量化。最后，根据实验分析得出与 HSU-Net-CFPP 相关的结论。

2.4.1. 实验系统

测量系统由一台分辨率为912×1140 的 DLP投影仪（LightCrafter 4500）和一台分辨率为 5472×3648 的单目机器视觉相机（Basler acA5472-17uc）组成，相机配备一支 16 mm 焦距镜头（Basler C23-1616-2M-S）。相机的最高帧率为 17 fps，投影仪在投射 8位灰度图案时的帧率为 120 fps。光圈设置为 f/11（镜头最小光圈：f/16），曝光时间为 50 ms，帧周期设为 110 ms，相机快门模式为全局复位释放。我们的测量系统提供约 8.73 cm^2的测量视场，空间分辨率约为 $50\,\mu$m。为满足GPU内存限制，每个原始帧（5472×3648 像素）均从中心裁剪为 3648×3648 像素的方形感兴趣区域，裁去左右两侧各 912 像素。随后使用 OpenCV 的基于区域的重采样（INTER_AREA）将该感兴趣区域降采样至 912×912 像素，在保留精细细节的同时有效抑制混叠和莫尔伪影。预处理后，该感兴趣区域对应的成像

面积约为 5.29cm^2，并生成紧凑的网络输入，从而降低计算成本。实验所用的计算机组件包括一台联想 ThinkStation P3 Tower，配备 Intel® CoreTM i9-13900K 处理器、128 GB 内存和 NVIDIA GeForce RTX 4070 GPU。网络使用 PyTorch 2.5.1（Python 3.9.18）实现并训练，并通过官方 NVIDIA CUDA 构建版本（cu118）使用 GPU 加速。

2.4.2. HSU-Net 实现细节

为实现稳定收敛，我们对超参数进行了调优，重点调整学习率、训练轮次、批量大小和快照数量。训练时最小化一个复合目标函数，其中结合了均方误差（权重 1.0）和负对数似然（权重 0.04）；为稳定方差学习，负对数似然项在前 30 个训练轮次中线性预热。优化采用 AdamW （初始学习率 5×10^{-4}，权重衰减 1×10^{-4}）。学习率按带热重启的余弦退火策略调整（周期为 20 个训练轮次），并在第 $20, 40, 60, 80$ 个训练轮次保存快照，以构建用于模型不确定性估计的集成。每个卷积块在两个卷积层之间采用 ReLU 激活，并加入 0.1 丢弃法，以抑制过拟合。训练最多进行 80 个训练轮次，小批量大小为 4；若验证性能连续 25 个训练轮次未提升，则通过早停终止优化。为提高稳定性和吞吐量，将梯度裁剪至l_2-范数 1.0，并启用带动态损失缩放的自动混合精度（float16）。此配置可加快收敛，提高对异方差像素级噪声的鲁棒性，并产生校准良好的预测不确定性。图 6 绘制了三个频率特定模型的训练损失和验证损失；这些模型分别针对高、中、低调制频率，对（N_i, D_i）进行回归。在所有情况下，总损失（均方误差 +负对数似然）和均方误差项均随训练轮次增加而下降并趋于平稳，表明模型已收敛。由于计算异方差高斯负对数似然时未包含其常数归一化项，因此总损失可能为负。当残差较小且模型预测的方差较低时，即使均方误差分量仍为正，负对数似然项也可能为负。使用相同输入、但针对不同目标（高、中、低频（N_i, D_i））训练三个频率特定模型，总计耗时 42 小时。推理每张图像需要 1.94 秒，以生成（N_i, D_i）及对应的不确定性图。对于时间要求严格的测量任务，可在采集过程中缓存捕获的条纹，并在离线状态下进行处理。

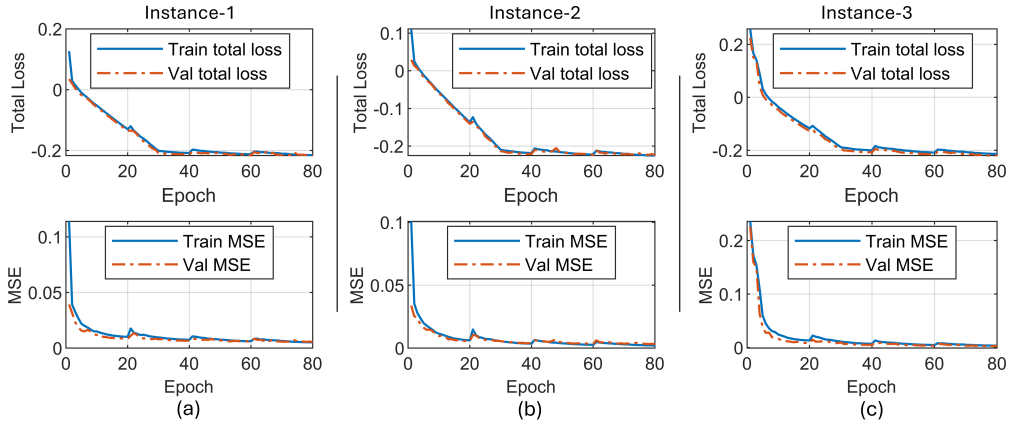

图 5. 总损失（均方误差和负对数似然的加权和）及均方误差的训练和验证曲线显示，三个频率特定模型均稳定收敛。图 (a)–(c) 展示了三个模型实例，它们以相同的 HFCF 输入进行训练，但使用不同的真实值目标（N_i^{gt}， D_i^{gt}），对应高、中、低调制频率。两项指标均随训练轮次增加而下降，并最终趋于平稳。由于高斯负对数似然未包含常数归一化项，因此总损失可能为负。

2.4.3. 定量分析

所提出的基于 HSU-Net 的 CFPP方法从单幅 HFCF图案中恢复包裹相位的正弦和余弦分量 N 和 D，随后将其用于相位展开和实时三维重建。由于三维形状由解包裹相位推导而来，因此 N 和 D 的预测精度在很大程度上决定了最终的重建精度。为评估所提方法，我们在训练过程中未见过的测试物体和场景上开展实验，并将重建结果与相应的真实数据进行比较。如图 6 所示，将两幅 HFCF图案输入训练好的模型进行推断，一幅对应单个物体，另一幅对应多个孤立物体；模型同时推断 N 和 D 及其对数方差，以评估分布外泛化。由于每幅 HFCF图案包含三组调制条纹，模型会输出三个频率各异的成对分量 $((\mu_N)_{\mathrm{mod},i}, (\mu_D)_{\mathrm{mod},i})$，以及相应的对数方差参数 $((\log \sigma_N^2)_{\mathrm{mod},i}, (\log \sigma_D^2)_{\mathrm{mod},i})$。成对分量 $(\mu_N)_{\mathrm{mod},i}$ 和 $(\mu_D)_{\mathrm{mod},i}$ 用于计算包裹相位和条纹级次，以进行相位展开；对数方差参数则用于构建不确定性图。在条纹分析中，包括包裹相位计算和条纹级次判定，我们使用图 6 所示的复合掩膜来去除离群点。该复合掩膜由两者的交集定义：根据相移条纹得到的幅值掩膜，以及依据物体有效高度范围设定的高度掩膜。

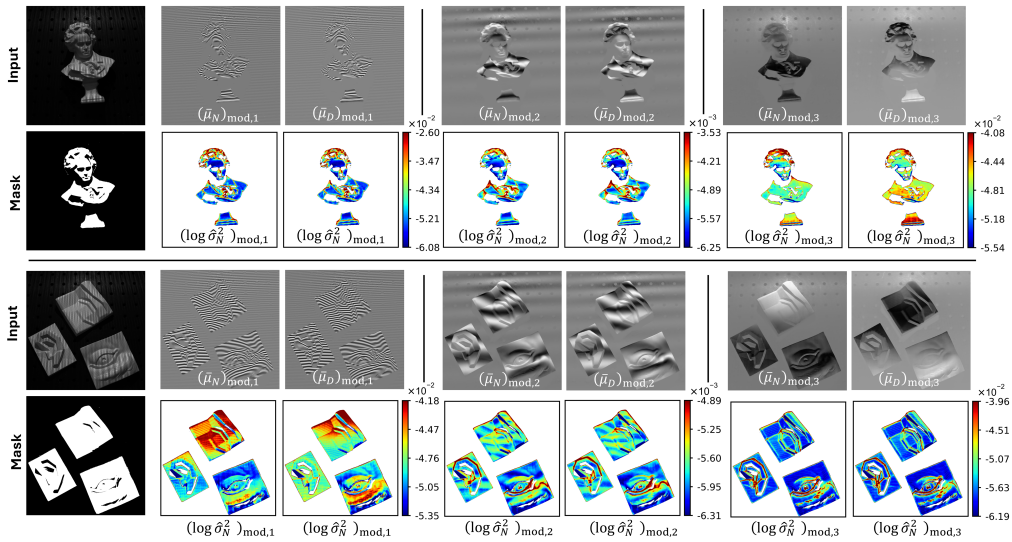

图 6. 测量了两种不同类型的物体场景，以评估 HSU-Net 的可靠性和泛化能力。将两个场景的 HFCF 作为训练后模型的输入；推理得到三个调制频率下 N 和 D 的均值及其对数方差。根据条纹幅值和物体高度范围定义的复合掩膜用于剔除离群值。

为进一步突出我们方法的优势，我们还将 FTP方法应用于同一物体，所得点云如图 7(a) 所示。为与 HSU-Net 进行公平比较，我们训练了一个内部增强型 U-Net基线，该基线采用批归一化和丢弃法，并启用带梯度缩放（GradScaler）的自动混合精度（AMP） [30]。我们使用与 HSU-Net 相同的数据集、超参数以及编码器–解码器通道宽度。损失函数为 L = 均方误差(μ_N, N) + 均方误差(μ_D, D)。使用增强型 U-Net 预测结果重建的三维点云如图 7(b) 所示，采用我们的 HSU-Net 获得的三维点云如图 7(c) 所示，而图 7(d) 中的真实点云则通过九步PSP方法获得。在这四个点云中，FTP 的质量明显低于真实

值；精细表面细节缺失，并出现波纹状伪影，我们将其归因于高频谱损失和频谱混叠。增强型 U-Net 可结合我们的 HFCF 输入，为相位展开提供一种替代方案，但无法提供不确定性估计。由于其假设噪声为同方差噪声，且缺乏逐像素不确定性建模，其精度仍低于 HSU-Net，这降低了其在低信噪比和高频区域中的鲁棒性。对于图 6 中的两种物体场景，FTP 的高度平均绝对误差分别为 1.37 mm 和 0.95 mm，增强型 U-Net 的分别为 0.22 mm 和 0.11 mm，而 HSU-Net 的分别为 0.12 mm 和 0.03 mm。包含多个孤立物体的场景误差低于单物体场景。我们认为，这是因为单个物体具有更复杂的几何形状，而孤立物体的表面更加平滑。这表明，表面复杂度是限制重建精度的主要因素之一，这一趋势也得到了后续不确定性分析的验证。

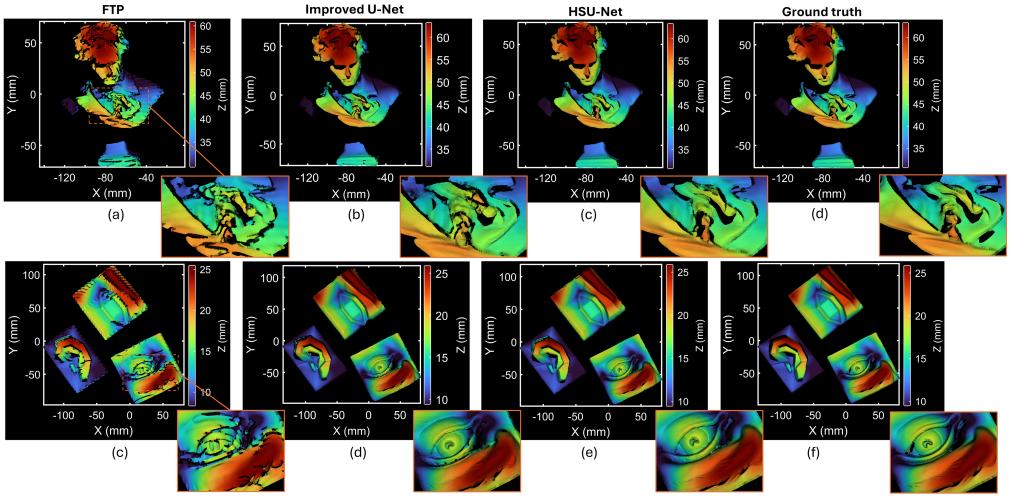

图 7. 使用不同方法获得的图 6 所示两个物体的三维点云，包括 FTP（a）、改进型 U-Net（b）和我们的方法（c），并与通过传统九步 PSP 获得的真实点云（d）进行比较。在这些方法中，FTP 的准确率最低；与真实值相比，改进型 U-Net 的表现略逊于我们的方法。

2.4.4. 不确定性估计

不确定性图提供了一种经过校准的逐像素可靠性表示，能够指示可能出现误差的位置并明确误差来源，从而提高三维重建的鲁棒性和准确性。为计算不确定性，我们使用模型输出 $\bar{\mu}_N$ 和 $\bar{\mu}_D$ 及其对应的对数方差，如图 6 所示。根据公式（15）–（17），我们得到数据不确定性（$\hat{\sigma}_N^{\text{data}}$）$_{\text{mod},\,i}$ 和（$\hat{\sigma}_D^{\text{data}}$）$_{\text{mod},\,i}$，以及模型不确定性（$\hat{\sigma}_N^{\text{model}}$）$_{\text{mod},\,i}$ 和（$\hat{\sigma}_D^{\text{data}}$）$_{\text{mod},\,i}$，如图 8(b) 所示。随后，根据公式（18）–（20）计算总不确定性（$\hat{\sigma}_N^{\text{total}}$）$_{\text{mod},\,i}$ 和（$\hat{\sigma}_D^{\text{total}}$）$_{\text{mod},\,i}$（图 8(a)），以及相位不确定性（$\hat{\sigma}_\phi$）$_{\text{mod},\,i}$（图 8(c)）。

图 9 展示了包含孤立物体的场景的不确定性图，以及对应的包裹相位误差。利用模型预测，我们计算了三个调制频率下的包裹相位，并将其与真实值进行比较，从而得到图 9(b) 所示的包裹相位误差图。我们利用这些误差图验证不确定性图能否识别预测误差可能增大的区域。在图 9(c) 和图 9(d) 中，包裹相位误差相对较大的区域与总不确定性较高的区域相吻合。图 9(e) 所示的相位不确定性也呈现出相同趋势。总之，总不确定性汇总了

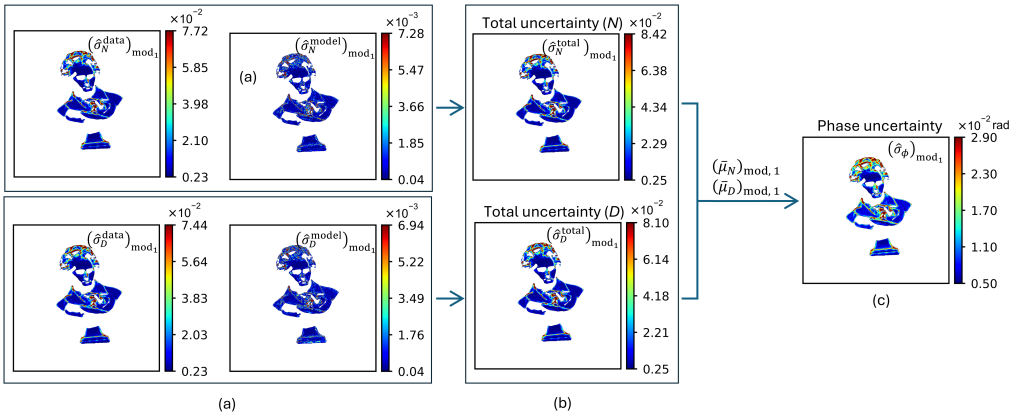

图 8. HSU-Net 不确定性图的可视化。将数据分量和模型分量结合起来，得到 N 和 D 的总不确定性，并进一步传播以获得相位不确定性。(a) N 和 D的数据不确定性图和模型不确定性图。(b) 结合数据不确定性和模型不确定性得到的总不确定性图。(c) 根据总不确定性和预测均值 $\bar{\mu}_N$ 和 $\bar{\mu}_D$计算得到的相位不确定性图。

误差来源，并且在统计上保持校准。仅依赖数据不确定性可能会导致对分布外区域或欠拟合区域过度自信，而仅依赖模型不确定性则无法捕捉低信噪比下的物理噪声。相位不确定性将风险投射到最终任务变量——解包裹相位 ϕ——上，因此可直接用作质量度量。数据不确定性和模型不确定性仍有助于识别误差来源并指导改进。数据不确定性反映了测量噪声以及条纹图案中的固有歧义，例如低对比度、饱和或阴影。模型不确定性源于训练覆盖范围有限或泛化误差，在分布外或较少观测到的表面条件下通常较高，可用于识别因模型知识不足而导致预测准确性较低的区域。因此，在实际应用中，应结合总不确定性和相位不确定性，同时分析数据不确定性和模型不确定性，以得出更可靠的结论。

图 9(b) 报告的高、中、低频包裹相位平均绝对误差分别为 0.055 rad、0.025 rad 和 0.032 rad，解包裹相位误差为 0.055 rad。我们认为，高频包裹相位误差与最终解包裹相位误差接近，是因为低频和中频的误差较小；这些误差在解包裹过程中不会传播，因为这些频率主要用于确定条纹级次。此外，中频包裹相位的相位误差小于高频和低频包裹相位的相位误差。在所有不确定性图中也能观察到这一现象。这主要是因为高频条纹受到系统 MTF 的衰减，导致调制度降低；而低频条纹则更容易受到环境照明、非均匀照明和频谱泄漏的影响 [8]。与此一致，估计的不确定性与整体预测误差呈正相关。

此外，若将这一实验置于质量控制场景中，它将展示 HSURE-Net 如何帮助人们做出更好的决策。例如，在 Additive Manufacturing 中，基于深度学习的条纹投影轮廓术因速度快、精度高而被用于表面缺陷检测。人们可能面临这样的风险：由于预测失败而将零件的某个区域错误地判定为缺陷，或者漏检缺陷。解决方法是查看估计的不确定性图，它们能够反映预测和模型可能存在的问题。这能提高深度学习模型的可靠性，并为人们识别缺陷提供参考指标，而不是盲目相信预测。为评估动态测量性能

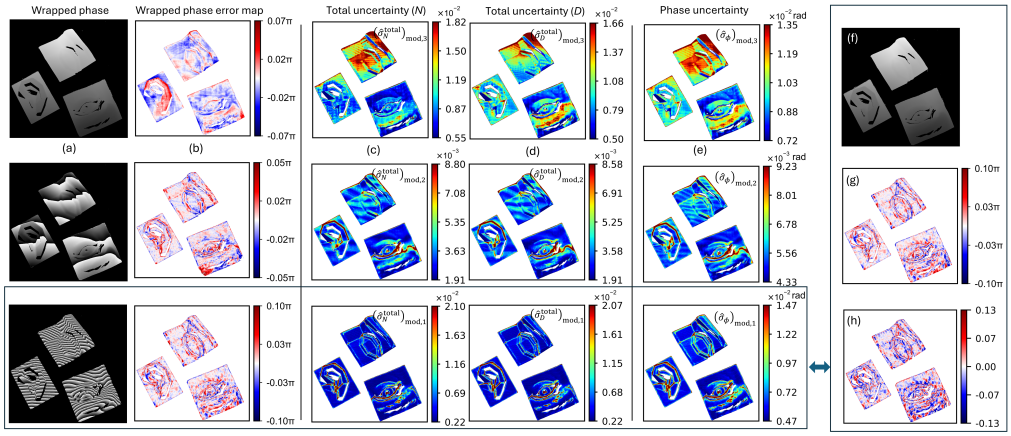

图 9. 总不确定性和相位不确定性与包裹相位误差之间的关联。不确定性图显示了由预测的 N 和 D 得到的包裹相位可能出现较大误差的区域。(a) 三种调制频率 f_{mod_1}、f_{mod_2} 和 f_{mod_3}下的包裹相位值。(b) (a) 中各图相对于真实值的包裹相位误差。(c) 和 (d) N 和 D的总不确定性图。(e) 相位不确定性图。

对 HSURE-Net 的动态测量性能进行评估时，投影仪持续投射图案，同时相机采集运动物体的图像。两个物体在同一视场内同时进行测量，其中一个作为场景内对照：一个物体保持静止，另一个物体则相对于静止物体进行旋转和平移。相机以 17 fps 运行，帧周期设为 67 ms。共采集 27 帧，总采集时间为 1809 ms。选取在 $\tau_k (k = 1、2、\ldots\ldots、6)$ 时刻采集的六帧（图 10(a)），相应的重建结果和相位不确定性图分别如图 10(b) 和图 10(c) 所示。如图 10(b) 所示，位于视场底部的静止物体（“嘴唇石膏模型”）在不同时刻的重建结果几乎相同，平均高度平均绝对误差为 0.035 mm。对于旋转的物体（“眼睛石膏模型”），即使在持续运动状态下，HSURE-Net 仍能生成准确的高分辨率重建结果，平均高度平均绝对误差为 0.039 mm。由于物体旋转时受到手指遮挡，采集图像中出现阴影，导致 τ_4 处的重建结果部分缺失。图 10(c) 中的相位不确定性图标示了更可能存在重建误差的区域。物体旋转时，同一表面区域出现在图像中的不同空间位置，其不确定性会略有变化，这主要是因为条纹质量随位置而变化。尽管如此，不确定性仍主要与表面几何复杂度相关。值得注意的是，当物体运动速度相对于相机帧率过快时，可能会产生运动模糊伪影，从而降低重建精度。

3. 结论

我们提出了一种实时三维测量方法，称为 HSU-CFPP，该方法基于异方差快照集成深度神经网络。它能够实现高精度单帧三维重建，同时输出逐像素不确定性估计，包括数据不确定性、模型不确定性、二者的总不确定性以及由此得到的相位不确定性。基于空间频率调制原理，我们采用频率分层策略，通过载波条纹将调制条纹编码为 HFCF，从而提供多个频率的相位信息以进行相位展开，同时减轻频谱混叠。将 HFCF 输入 HSU-Net 网络后，该网络预测用于计算包裹相位的正弦和余弦分量（即分子和分母）

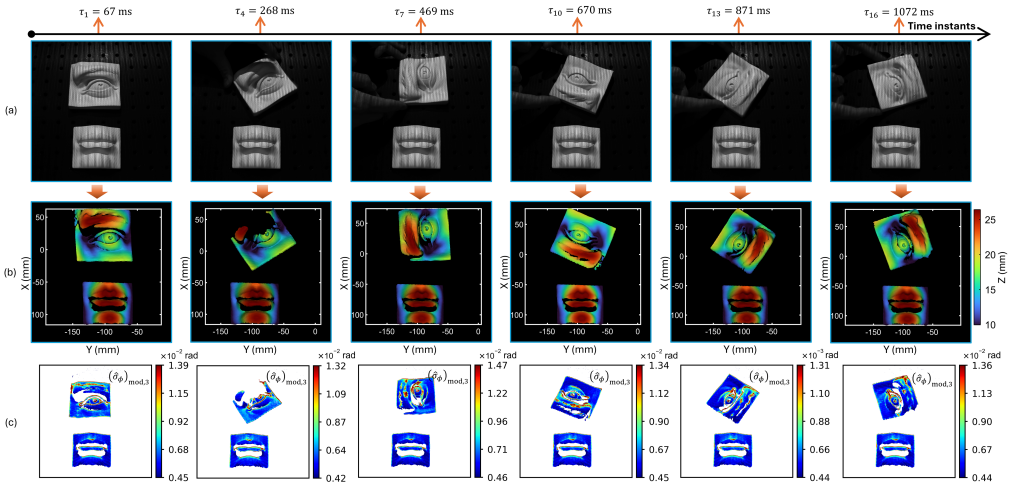

图 10. 单一视场内 HSURE-Net 的静态—动态验证：一个物体保持静止，另一个物体相对于它旋转。除三维重建结果外，还给出了相应的相位不确定性图。(a) 在时间点 $\tau_k (k = 1, 2、\ldots\ldots、6)$ 捕获的 HFCF帧。(b) 各时间点的三维重建结果。(c) 与相应重建结果的高度图相关联的相位不确定性图。

。随后，根据 HFCF 预测结果中不同频率的包裹相位恢复绝对相位，从而实现无歧义的单次相位解包裹。HSU-Net 通过预测分子和分母的逐像素对数方差来建模异方差性；数据不确定性由这些方差推导得出，模型不确定性则由快照预测的离散程度得出。这些分量分别反映测量噪声和模型知识的有限性。二者之和构成总不确定性，在异方差高斯似然下经实验证实具有良好的校准效果。通过一阶误差传播得到的相位不确定性，为置信度图和下游加权提供了相位质量的定量度量。实验表明，所提方法在连续表面和孤立离散物体上均具有较高精度，并揭示不确定性度量与相位误差、高度图误差均存在显著相关性：估计不确定性较高的区域测量误差也较大，并会传递至最终的三维重建。有效利用这些不确定性图，能够为深度学习输出和重建点云提供可量化的置信度，这对于实时条纹投影轮廓术测量中的质量控制至关重要，尤其是在无法获得真实值时。

此外，我们认为，所提出的 HSU-Net 在其他需要校准不确定性的逐像素回归任务中具有巨大潜力，例如光学干涉测量、雷达与遥感，以及医学和科学成像。鉴于不确定性分析在光学计量中的重要性，我们未来的工作将不再局限于网络层面的不确定性，而是研究基于深度学习的实时FPP系统的整体不确定度预算，包括测量环境、测量硬件、重建算法和模型输出。

参考文献

1. B. Wang, W. Chen, J. Qian, 等，“单次超分辨率条纹投影轮廓术（sssr-fpp）：利用深度学习实现每秒 100,000 帧的三维成像”，《Light: Science & Applications》**14**，70（2025）。2. B. Zhang、J. Ziegert、F. Farahi 和 A. Davies，“利用条纹投影对激光粉末床熔融进行原位表面形貌测量”，《增材制造》**12**，100–107（2016）。3. J. Geng，“结构光三维表面成像：教程”，《先进光子学》**3**，128–160（2011）。

4. C. Zuo、S. Feng、L. Huang 等，“条纹投影轮廓术的相移算法：综述”，《光学与激光工程》109，23–59（2018）。5. J. Salvi、S. Fernandez、T. Pribanic 和 X. Llado，“用于表面轮廓测量的结构光图案研究现状”，《模式识别》43，2666–2680（2010）。6. S. Zhang，“数字条纹投影轮廓术的绝对相位提取方法：综述”，《光学与激光工程》107，28–37（2018）。7. C. Zuo、L. Huang、M. Zhang 等，“条纹投影轮廓术的时间相位展开算法：比较综述”，《光学与激光工程》85，84–103（2016）。8. M. Takeda 和 K. Mutoh，“用于自动测量三维物体形状的傅里叶变换轮廓术”，《应用光学》22，3977–3982（1983）。9. S. Zhang、D. Van Der Weide 和 J. Oliver，“用于三维形状测量的超快速相移法”，《光学快报》18，9684–9689（2010）。10. M. Takeda、Q. Gu、M. Kinoshita 等，“频分复用傅里叶变换轮廓术：对具有较大高度不连续和/或表面隔离区域的物体进行单次三维形状测量”，《应用光学》36，5347–5354（1997）。11. H.-M. Yue、X.-Y. Su 和 Y.-Z. Liu，“基于复合结构光图案的傅里叶变换轮廓术”，《光学与激光技术》39，1170–1175（2007）。12. Y.-c. Wu、Y.-p. Cao、Z.-f. Huang 等，“改进的复合傅里叶变换轮廓术”，《光学与激光技术》44，2037–2042（2012）。13. L.-C. Chen 和 H. H. Hai，“采用新型正交椭圆带通滤波实现精确三维表面重建的傅里叶变换轮廓术”，《精密工程》38，512–524（2014）。14. Y. Gong 和 S. Zhang，“利用商用 DLP 投影仪实现超高速三维形状测量”，《光学快报》18，19743–19754（2010）。15. Z. Zhang、C. E. Towers 和 D. P. Towers，“采用最优三频选择进行三维形状与颜色测量的高效彩色条纹投影系统”，《光学快报》14，6444–6455（2006）。16. Y. Zhang、Z. Xiong 和 F. Wu，“从散斑嵌入条纹中实现无歧义三维测量”，《应用光学》52，7797–7805（2013）。17. C. Zuo、J. Qian、S. Feng 等，“光学计量中的深度学习：综述”，《Light: Science & Applications》11，39（2022）。18. S. Feng、Q. Chen、G. Gu 等，“利用深度学习进行条纹图案分析”，《先进光子学》1，025001–025001（2019）。19. W. Yin、Y. Che、X. Li 等，“用于条纹图案分析的物理信息引导深度学习”，《光电进展》7，230034–1（2023）。20. H. Yu、D. Zheng、J. Fu 等，“用于精确条纹投影轮廓术的深度学习条纹调制度增强方法”，《光学快报》28，21692–21703（2020）。21. Y. Li、J. Qian、S. Feng 等，“用于单次绝对三维形状测量的深度学习赋能双频复合条纹投影轮廓术”，《光电进展》5，210021–1（2025）。22. Y. Chen、J. Kang、L. Feng 等，“用于单次三维形状测量的基于深度学习的频分复用复合条纹投影轮廓测量技术”，《测量》233，114640（2024）。23. Y. Jiang、J. Qin、Y. Liu 等，“利用空间复合图案进行基于深度学习的单次条纹投影轮廓术”，《IEEE仪器与测量汇刊》73，1–14（2024）。24. Y. Li、J. Qian、S. Feng 等，“用于单次绝对三维形状测量的复合条纹投影深度学习轮廓术”，《光学快报》30，3424–3442（2022）。25. Y. Fu、Y. Huang、W. Xiao 等，“用于快速三维测量的基于深度学习的双目复合彩色条纹投影轮廓术”，《光学与激光工程》172，107866（2024）。26. B. Zhang、K. Jiang、J. Lin 等，“基于双内移相法的深度学习赋能单次高频彩色条纹投影轮廓术”，《测量》239，115462（2025）。27. S. Feng、C. Zuo、Y. Hu 等，“基于深度学习并带有不确定性估计的条纹图案分析”，《Optica》8，1507–1510（2021）。28. T. Falk、D. Mai、R. Bensch 等，“U-Net：用于细胞计数、检测和形态测量的深度学习”，《自然-方法》16，67–70（2019）。29. N. Siddique、P. Sidike、C. Elkin 和 V. Devabhaktuni，“U-Net 及其变体在医学图像分割中的理论与应用”，arXiv 预印本 arXiv:2011.01118（2020）。30. C. Garbin、X. Zhu 和 O. Marques，“丢弃法与批归一化：对其深度学习影响的实证研究”，《多媒体工具与应用》79，12777–12815（2020）。