引用格式：WANG Linlin, LI Bingzhen, GUO Wanyi, et al. Data Augmentation and Normalization Optimization for Single-frame Fringe Projection Depth Estimation[J]. Acta Photonica Sinica, 2026, 55(2):0210001
王琳霖,厉秉桢,郭万意,等.单帧条纹投影深度估计的数据增强与归一化优化[J].光子学报,2026,55(2):0210001

单帧条纹投影深度估计的数据增强与归一化优化

王琳霖,厉秉桢,郭万意,高骞,王传云
(沈阳航空航天大学 人工智能学院, 沈阳 110136)

摘　要:针对单帧条纹投影深度估计中传统数据增强与归一化方法适配性不足的问题,提出了一种结合区间裁剪幂次归一化与结构化条纹遮挡增强的数据预处理策略。区间裁剪幂次归一化通过裁剪极端亮度值、幂次变换与分段映射优化像素分布,强化暗部细节表现的同时,保持整体亮度均衡与对比度稳定。结构化条纹遮挡增强则通过规则化遮挡与条纹一致性填补增加样本的多样性。通过多组实验验证,所提出的归一化与增强方法的有效性。进一步分析发现,所提出的数据增强与归一化方法在多种网络结构中均展现出稳定的性能提升,体现出良好的通用性与适配性,同时对不同结构的网络呈现出一定的差异化适应特征。该策略兼具普适性与差异化适配能力,为单帧条纹投影深度估计的性能提升提供了一种切实可行的优化思路。

关键词:条纹投影;深度估计;数据增强;归一化;数据预处理
中图分类号:TP391.41　　**文献标识码**:A　　**doi**:10.3788/gzxb20265502.0210001

0　引言

近年来,深度学习方法广泛应用于单帧条纹投影的深度估计任务中[1]。利用卷积神经网络,可以直接从单帧条纹图像中回归深度图或相位图[2],并在多种工业测量任务中取得了优异的表现[3-4]。然而,基于单帧条纹投影的深度估计方法在实际应用中仍面临诸多挑战。其中,最核心的问题是训练数据的获取困难与多样性不足[5]。高质量的条纹投影数据集通常需要昂贵的设备和精密的标定过程,且采集过程耗时耗力[6-7]。这导致现有数据集的规模相对有限,严重制约了深度学习模型在单帧条纹投影任务中的性能和泛化能力。

在训练数据有限的条件下,数据增强成为提升模型性能的重要方法[8]。传统的数据增强方法,如旋转、缩放、裁剪[9-10]等,在自然图像任务中能够有效增加数据多样性,但在条纹投影任务中却可能带来新的问题。由于条纹图像中的周期性结构直接承载了三维形貌信息,这些几何变换可能会破坏条纹的相位一致性与空间分布规律,从而影响深度估计精度。例如,随机裁剪可能改变条纹周期,旋转和缩放会引入非物理性的频率变化,这些都会导致网络学习到的特征与真实测量系统不匹配。因此,如何设计兼顾物理一致性与数据多样性的增强方法,是单帧条纹投影深度估计中亟需解决的问题。

除了数据增强,输入数据的归一化也是影响网络性能的关键环节。归一化的目的是将输入图像的像素值映射到合适的数值范围或分布,以加快收敛速度、稳定训练过程并提升模型性能。常见的归一化方法包括 Min-Max 归一化[11]、Z-Score 标准化[12]、分位数归一化[13]、Gamma 校正[14]等。然而,这些方法通常基于自然图像的统计分布假设,而条纹投影图像的亮度分布则受到光源非均匀性、相机响应及物体反射率等多因素影响,与自然图像的统计特性存在显著差异。一方面,投影系统的光源亮度分布、相机响应曲线以及被测物体的反射率会导致图像亮度分布不对称[15];另一方面,背景与条纹区域的亮度差异显著,且可能存在局部过曝或欠曝区域。在这种情况下,直接使用如 Z-Score 标准化等方法,可能会导致亮度分布被不均衡拉伸或

基金项目:国家自然科学基金(62376165),辽宁省教育厅基本科研项目(LJ212410143055)
第一作者:王琳霖, wlin_23@163.com
通讯作者:厉秉桢, libingzhen_0111@163.com
收稿日期:2025-09-01;**录用日期**:2025-10-17

压缩，进而削弱网络对有效条纹信息的利用能力。

综上，针对单帧条纹投影深度估计任务中训练数据有限、数据分布不均及传统数据增强与归一化方法适配性不足的问题，提出了一种结合结构化遮挡增强与区间裁剪幂次归一化的新型数据预处理策略。在归一化方面，提出了区间裁剪幂次归一化（Range-clipped Piecewise Contrast Normalization，RPCN）法，通过亮度直方图分析确定裁剪范围以抑制异常亮度值，再将数据线性映射至标准区间并施加幂次变换，以优化暗部细节与高亮区域的平衡。在数据增强方面，设计了结构化条纹遮挡增强（Structured Stripe Occlusion Augmentation，SSOA）方法，将输入条纹图像划分为规则的区域，并在训练过程中以固定概率随机选取区域进行条纹遮挡模拟，同时利用条纹周期性特征进行物理相近性填补与标签修正，从而提升模型在复杂场景下的精度。该策略能够有效改善输入数据分布与深度网络训练需求的匹配性，在保证物理一致性的同时提升了模型的泛化性能与预测精度。

1 基本原理

1.1 区间裁剪幂次归一化

围绕抑制异常极值对分布的干扰、保留并增强暗部有用条纹信号、将像素分布调整为更利于网络学习的动态范围的设计目标，提出了区间裁剪幂次归一化。

1.1.1 单帧条纹图像的亮度分布特性

为了验证单帧条纹图像的亮度分布特性，将训练集中每一张图像的所有像素值提取并汇总，统计其亮度分布直方图，如图1所示。可以明显观察到，在灰度分布的左端存在一个显著峰值，该峰主要对应原始图像中的阴影区域或低反射率表面，属于无效信息区。这部分像素不仅缺乏有效条纹结构，而且更易受噪声干扰。相比之下，有效条纹信息主要集中在[−0.9，0.5]的亮度范围内，承载了绝大多数与高度估计相关的细节。

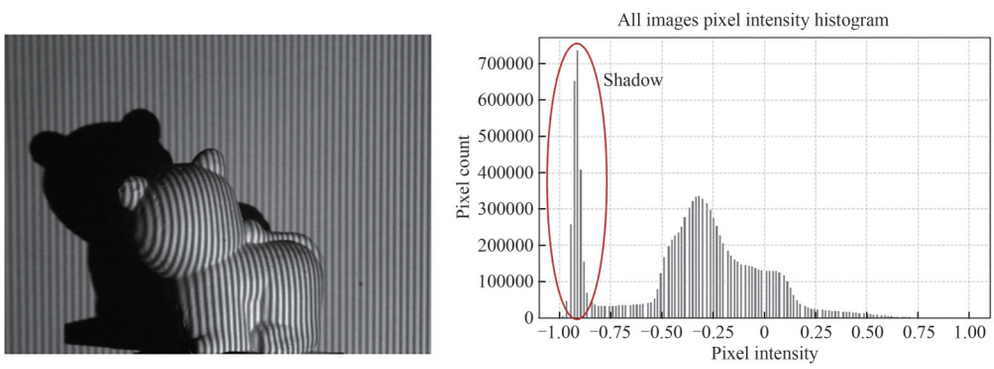


图1 原始条纹图像和数据集图像的像素亮度分布直方图

Fig.1 Original stripe image and the histogram of pixel brightness distribution in the dataset image

这种亮度分布特性表明，若直接使用全局归一化方法，极端暗部和少量高亮值会主导归一化区间，导致有效信息区的动态范围被压缩，从而影响网络对条纹细节的感知。为了进一步量化这一影响，首先分析常见归一化方法在此类亮度分布下的表现。

1.1.2 常见的归一化方法

在单帧条纹图像的亮度分布中，存在大量极暗阴影区域与少量高亮异常点，这种长尾分布特性使得传统归一化方法存在明显不足。下面分析了几种典型方法及其局限性。

（1）Min-Max 归一化

将像素 x 映射到[0，1]区间

$$x'=\frac{x-x_{\min}}{x_{\max}-x_{\min}} \tag{1}$$

式中，$x_{\min}$ 和 $x_{\max}$ 分别是图像像素的最小值和最大值。

在单帧条纹投影图像中，阴影区域的像素值导致 $x_{\min}$ 极小，而高反射噪点使得 $x_{\max}$ 偏大，原始图像中的承载有效信息亮度区间 $[-0.9,\ 0.5]$ 在经过 Min-Max 归一化后，被压缩至较窄范围 $[0.1,\ 0.8]$，如图 2 所示，这种压缩较低了有效条纹区域的对比度，削弱了图像中的条纹细节表现。

(2)Z-Score 标准化

利用均值 μ 和标准差 σ 进行归一化

$$x'=\frac{x-\mu}{\sigma} \tag{2}$$

式中，$\mu=\frac{1}{N}\sum_{i=1}^{N}x_i$，$\sigma=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}$。

由于阴影区域占比较大，μ 会显著偏低，σ 会因极值而被拉大，从而使有效像素区间的均值偏离 0，未能有效抑制无效信息的影响，可能在模型训练过程中，降低特征提取的稳定性与准确性。

(3)分位数归一化(Robust)

使用中位数 median 和四分位距 $\mathrm{IQR}=Q_3-Q_1$ 缩放

$$x'=\frac{x-\mathrm{median}}{\mathrm{IQR}} \tag{3}$$

式中，Q_1 和 Q_3 分别为第一四分位数和第三四分为数，该方法相比 Z-Score 标准化在一定程度上削弱了极值的影响，但由于阴影区域低光像素占比过高，整体分布特性未得到显著改善。

(4)Gamma 矫正

对像素进行 Min-Max 归一化后，通过幂次变换调整亮度分布

$$x'=\left(\frac{x-x_{\min}}{x_{\max}-x_{\min}}\right)^{\gamma} \tag{4}$$

式中，$\gamma<1$ 可增强暗部细节，$\gamma>1$ 可增强亮部细节。针对单帧条纹投影图像需要优先恢复暗部细节的特点，本研究选取 $\gamma=0.7$。然而，由于阴影区域分布的影响，虽然暗部细节得到提升，但阴影区像素值分布由原图像中的 $[-1,\ -0.9]$ 扩展至 $[0.1,\ 0.3]$，反而将有效像素信息区域的亮度范围压缩至 $[0.3,\ 0.8]$，导致整体对比度下降，如图 2 所示。

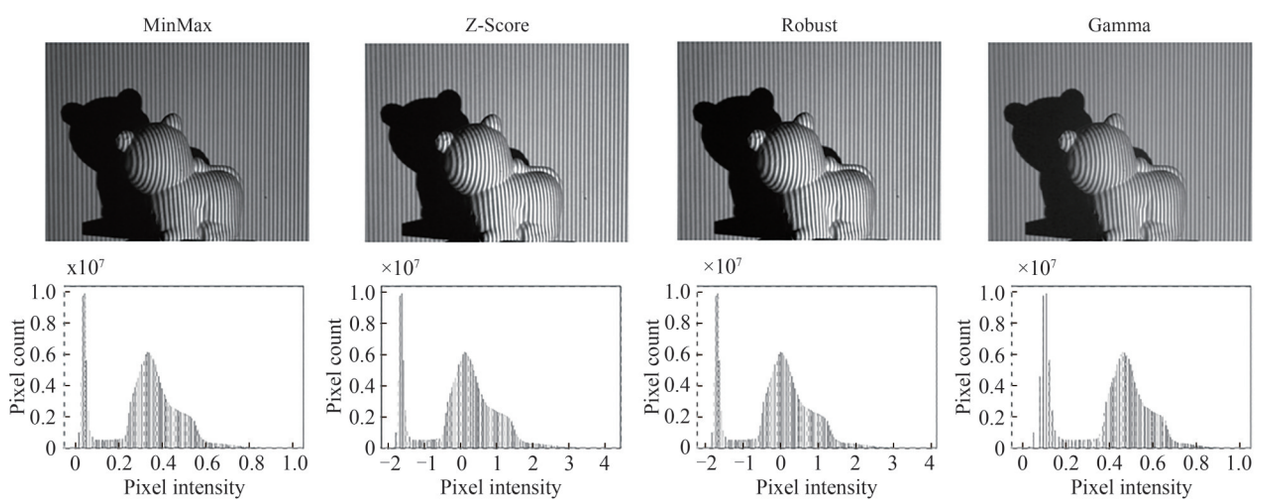


图 2　四种归一化方法后的图片示例和像素亮度分布直方图

Fig.2　Image examples and pixel brightness distribution histograms after four normalization methods

1.1.3　区间裁剪幂次归一化(RPCN)

为解决上述问题，提出区间裁剪幂次归一化方法(Range-clipped Piecewise Contrast Normalization，RPCN)。该方法先去除极值干扰，再针对性增强暗部条纹，在保持物理连续性的同时提升网络对条纹结构的感知能力。

设输入像素为 x，数据集的裁剪上下限为$[L, H]$，幂次参数为 $\gamma>0$。RPCN的映射为四步

(1)区间裁剪

$$x'=\mathrm{clip}(x, L, H)=\min\left(\max(x, L), H\right) \tag{5}$$

(2)线性映射到$[0, 1]$

$$x''=\frac{x'-L}{H-L} \tag{6}$$

(3)幂次变换

$$x'''=\left(x''\right)^{\gamma} \tag{7}$$

(4)分段区间映射

根据幂次变换后的直方图和任务需要，将$[0, 1]$范围进一步划分为若干子区间$[s_i, e_i]$，并将其线性映射到指定的目标区间$[S_i, E_i]$

$$R(x)=\begin{cases} x'''\cdot k_1+b_1, & x'''\in[s_n, e_n] \\ x'''\cdot k_2+b_2, & x'''\in[s_n, e_n] \\ \cdots & \\ x'''\cdot k_n+b_n, & x'''\in[s_n, e_n] \end{cases} \tag{8}$$

式中，$k_i=\dfrac{E_i-S_i}{e_i-s_i}$，$b_i=S_i-k_i\cdot s_i$

RPCN过程见图3。根据研究所使用的公开数据像素亮度分布特点，本文选取 $\gamma=0.5$，$L=-0.9$，$H=0.5$，该配置是基于对视觉动态范围与特征保真度的综合考量，经由多轮实验对比、参数微调与验证后得到的优化结果。其核心目标在于在强化暗部细节表现的同时，保持整体影像的亮度均衡与对比度稳定，从而提升

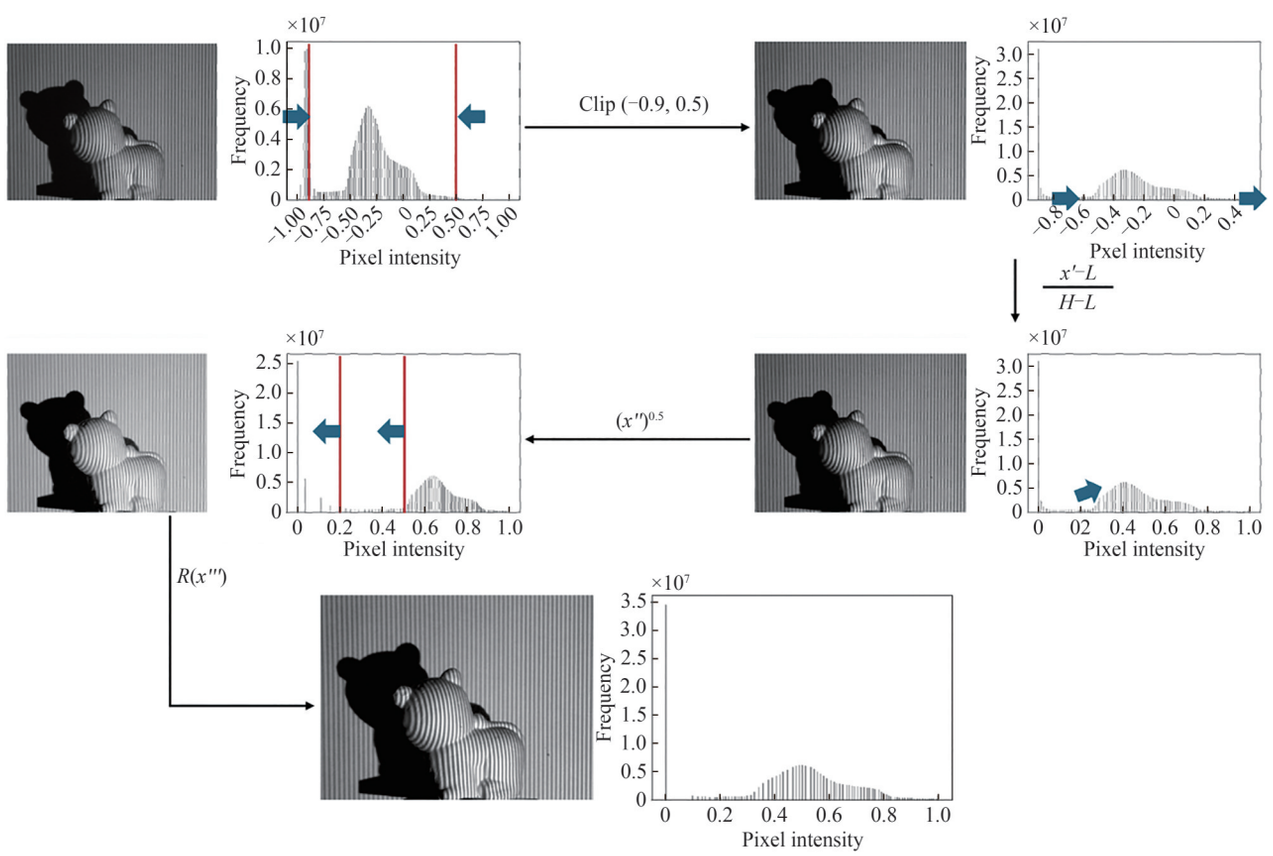


图3 RPCN过程图

Fig.3 RPCN process diagram

模型对低亮度区域的感知与重建能力。需要强调的是，具体参数的选取应结合不同数据集的亮度分布特性进行适当调整，以在视觉动态范围与特征保真度之间实现最优平衡。RPCN的最终输出 $y=\mathcal{R}(x)\in[0,\ 1]$ 可直接作为网络输入。经处理后，有效像素区间分布在$[0.1,\ 1]$，阴影区域与有效区域的灰度分布实现了明显分离，直方图峰值分布更加集中且互不重叠，从而显著提高了条纹图像的整体对比度。该优化不仅减少了无效阴影信息对特征提取的干扰，还为后续三维重建网络提供了更稳定、信息密度更高的输入基础。

1.2 结构化条纹遮挡增强

针对单帧条纹投影深度估计任务的特殊性，提出了一种结构化条纹遮挡增强方法，通过制造局部条纹变化并进行物理一致性填补，在不破坏条纹整体结构的情况下显著提升样本多样性和模型精度。

1.2.1 区域划分与遮挡概率

在结构化条纹遮挡增强（SSOA）方法中，首先需要对输入条纹图像进行规则化的空间划分，以保证遮挡操作在不同空间位置上均匀分布，同时为后续的概率遮挡策略提供离散化的区域单元。

设输入条纹图像为二维矩阵 $\boldsymbol{I}\in\mathbb{R}^{H\times W}$，其中 H 和 W 分别表示图像的高度和宽度。计算划分边界，$h_1=\lfloor H/3\rfloor$，$h_2=2\times\lfloor H/3\rfloor$，$w=\lfloor W/2\rfloor$，利用边界将图像划分为六个等大小的矩形区域$\{R_1,\ R_2,\cdots,R_6\}$，如图4所示。

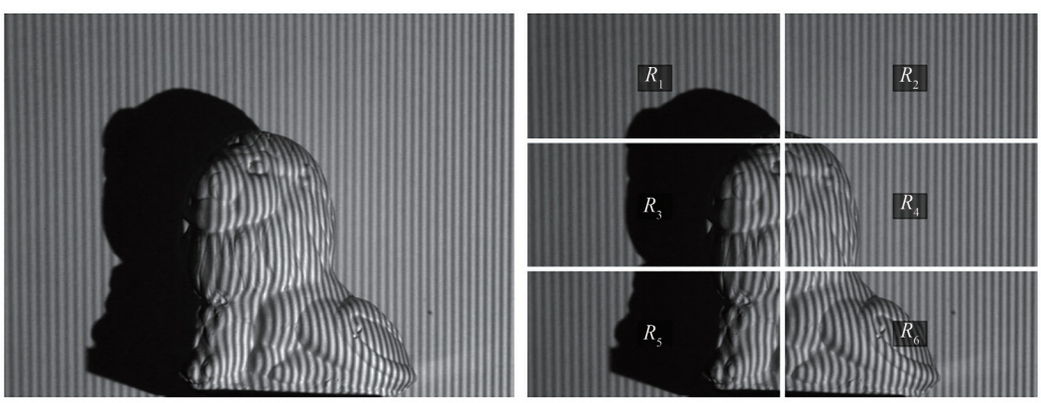


图4 单帧条纹投影图像的区域划分示意

Fig.4 Schematic diagram of region division for single frame stripe projection image

其中各区域定义为

$$\begin{cases}[\,R1=\mathrm{I}0{:}\,h1,\quad 0{:}\,w\,]\\ [\,R2=\mathrm{I}0{:}\,h1,\quad w{:}\,W\,]\\ [\,R3=\mathrm{I}h1{:}\,h2,\quad 0{:}\,w\,]\\ [\,R4=\mathrm{I}h1{:}\,h2,\quad w{:}\,W\,]\\ [\,R5=\mathrm{I}h2{:}\,H,\quad 0{:}\,w\,]\\ [\,R6=\mathrm{I}h2{:}\,H,\quad w{:}\,W\,]\end{cases}\tag{9}$$

为了模拟不同区域条纹背遮挡的情况，引入二值随机变量

$$O_i\sim\mathrm{Bernoulli}(p),\ i=1,2,\cdots,6\tag{10}$$

用于决定每个区域是否被遮挡，其中 p 为遮挡概率。在每轮训练开始前，对于每张训练图像中的第 i 个区域，若 $O_i=1$，则执行遮挡操作；若 $O_i=0$，则保持原状。在单张图像的6个区域中，期望被遮盖区域数量为

$$E[K]=\sum_{i=1}^{6}\mathbb{E}[O_i]=6p\tag{11}$$

该概率设计确保在训练过程中，遮挡区域数量保持稀疏且随机，即增强了数据的多样性，又避免了过度破坏条纹图像的特征结构。

1.2.2 条纹特征一致性填补

在结构化条纹遮挡增强(SSOA)中,对选中区域的遮挡并非简单地将像素值置零或用随机噪声替代,这样会导致条纹信息的物理不连续性,进而影响模型对条纹结构的学习。为此,本文提出基于条纹特征一致性的填补方法,效果示意图见图5。具体地,首先在条纹图像顶部的前 h_{top} 行提取该区域的条纹特征信息,记该参考区域为矩阵 $\boldsymbol{I}'\in\mathbb{R}^{h_{\text{top}}\times W}$,对待填补的遮挡区域 $R_{\text{occ}}{}^{h_r\times h_w}$ 内的每一列像素,根据对应列的参考条纹特征及其邻近列的信息进行重构。具体计算方法如下。

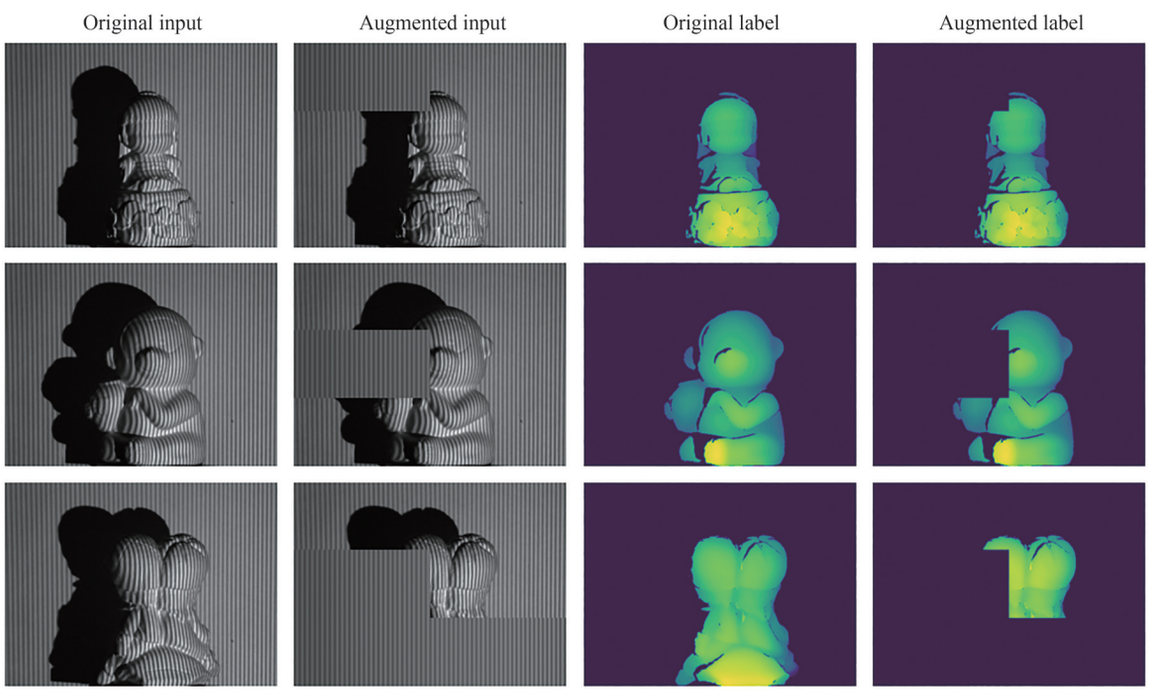


图5 结构化条纹遮挡增强(SSOA)效果示意图

Fig.5 SSOA effect demonstration

对于第 $j(1\leqslant j\leqslant W)$ 列,考虑其自身及左右相邻列的参考条纹像素值

$$S_j=\left\{I[:,1],\ 1\in[j-k,\ j+k],1\in[1,\ W]\right\},\ \ k=0,1,2,3 \tag{12}$$

然后对这些列的像素值进行均值计算,得到填补列像素均值

$$R_{\text{occ}}[:,j]=\frac{1}{|S_j|}\sum_{s\in S_j}\frac{1}{h_{\text{top}}}\sum_{i=1}^{h_{\text{top}}}s_i \tag{13}$$

式中,s_i 表示参考区域第 i 行对应列的像素值,$|S_j|$ 为实际参与计算的列数。该均值作为该列在遮挡区域内所有像素的填充值,保留了条纹结构的物理特性。同样地,为保证训练标签的一致性,遮挡区域的标签值也采用对应参考区域的均值填充

$$\text{label}_{\text{occ}}(i,j)=\frac{1}{h_{\text{top}}\times W}\sum_{m=1}^{h_{\text{top}}}\sum_{n=1}^{W}\text{label}(m,m),\ \ \forall(i,j)\in E_{\text{occ}} \tag{14}$$

2 实验设计与方案

2.1 数据集介绍

本文采用两类数据集。首先,使用公开的真实单帧条纹投影数据集[16],该数据集包含分辨率为480×640像素的单通道单帧条纹投影图像及其对应的深度标签,划分为训练集和测试集,其中训练集包含612对样本,测试集包含36对样本。其次,为进一步验证模型的鲁棒性,基于Blender自建了单帧条纹投影数据集,图像分辨率同样为480×640像素,包含单通道条纹图像及其对应的深度标签,并划分为训练集与测试集,其中训练集包含216对样本,测试集包含36对样本。

2.2 衡量指标

为全面评估深度估计性能,采用定量与定性相结合的指标,包含图像质量、几何误差。

(1) MAE

衡量预测值与真值之间的平均绝对偏差,取值范围[0, +∞),数值越小偏差越小。

$$\mathrm{MAE}=\frac{1}{N}\sum_{i=1}^{H}\sum_{j=1}^{W}\left|\hat{Y}_{ij}-Y_{ij}\right| \tag{15}$$

(2) RMSE

强调大误差的影响,对异常值更敏感, 取值范围[0, +∞),数值越小偏差越小。

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{H}\sum_{j=1}^{W}\left(\hat{Y}_{ij}-Y_{ij}\right)^2} \tag{16}$$

(3) SSIM

SSIM是一种综合考虑亮度、对比度和结构三个分量来综合衡量两幅图像相似度的指标。SSIM的曲直范围在[1, 1],值接近1表示图像质量较高。

$$\mathrm{SSIM}\left(\hat{Y},\ Y\right)=\frac{\left(2\mu_{\hat{Y}}\mu_Y+C_1\right)\left(2\sigma_{\hat{Y}Y}+C_2\right)}{\left(\mu_{\hat{Y}}^2+\mu_Y^2+C_1\right)\left(\sigma_{\hat{Y}}^2+\sigma_Y^2+C_2\right)} \tag{17}$$

式中,$\mu_{\hat{Y}}$、μ_Y 分别为预测图像于真值的均值,$\sigma_{\hat{Y}}^2$ 和 σ_Y^2 分别为预测图像和真值的方差,$\sigma_{\hat{Y}Y}$ 为二者的协方差,$C_1=(k_1L_{\max})^2$、$C_2=(k_2L_{\max})^2$ 为稳定常数,通常 $k_1=0.01$、$k_2=0.03$。

(4) PSNR

PSNR是基于图像的最大可能像素值和均方误差(MSE)之间的比率。单位为dB,越大表示预测图像与真值越接近。

$$\mathrm{PSNR}=10\cdot\log_{10}\left(\frac{L_{\max}^2}{\mathrm{MSE}}\right) \tag{18}$$

式中,$\mathrm{MSE}=\frac{1}{N}\sum_{i=1}^{H}\sum_{j=1}^{W}\left(\hat{Y}_{ij}-Y_{ij}\right)^2$,$L_{\max}$为图像像素的最大可能值。

2.3 实验方案

实验共分为四个阶段,前两个阶段采用以base=32为初始通道数的标准UNet,后两个阶段引入以base=32为初始通道数的UNet++,以及MobileNetV2为主干网络的DeepLabV3+。三者在结构上存在显著差异,UNet依托对称的编码-解码架构及跳跃连接实现多尺度特征融合;UNet++在此基础上通过密集连接进一步增强了特征传递与语义表达;而DeepLabV3+则借助空洞卷积与多尺度特征金字塔,在扩大感受野的同时保持空间分辨率。不同的特征提取策略为本研究比较优化方法的适应性与有效性提供了实验基础。

模型性能通过四项指标量化评估,包括峰值信噪比(PSNR)、结构相似性指数(SSIM)、均方误差(MSE)和平均绝对误差(MAE),以全面反映模型在重构精度、结构保持能力和误差控制方面的综合表现。

具体三个阶段分别为

(1)归一化策略对比

为评估不同归一化方法对模型训练与重建性能的影响,在输入端分别采用Min-Max归一化、Z-score标准化、分位数归一化、Gamma矫正以及本文提出的RPCN方法进行对比实验。同时,为验证RPCN中各子模块设计的合理性,设置消融实验,逐步启用区间裁剪、线性映射、幂次变换与分段区间映射,以分析各步骤的独立作用及整体协同效果。

(2)数据增强概率对比

引入SSOA条纹遮挡策略,概率p分别取$\frac{1}{3}$, $\frac{1}{6}$, $\frac{1}{9}$, $\frac{1}{12}$,并与不添加SSOA策略的UNet进行对比实验。该实验的目的在于评估SSOA对模型性能的提升效果,同时确定最优增强概率,为后续协同优化阶段的数据增强方案提供依据。

(3)归一化与数据增强协同优化

在验证RPCN与SSOA策略有效性的前提下,设计了四类对比方案:①不引入优化的基线模型;②仅加入SSOA模块;③仅采用RPCN方法;④同时结合SSOA与RPCN。实验首先在UNet框架下分别考察优化方法的独立作用与协同效应,随后扩展至UNet++与DeepLabV3+,以检验所提优化策略在不同网络结构中的通用性与稳定性。

(4)泛化与鲁棒性验证实验

本阶段基于Blender自建的单帧条纹投影数据集上进行了扩展实验。使用UNet、UNet++和DeepLabV3+三种模型,在相同训练策略与参数设定下进行性能评估。进一步验证RPCN和SSOA优化方法在不同数据分布下均具备良好的泛化效果与鲁棒性能,见图6。

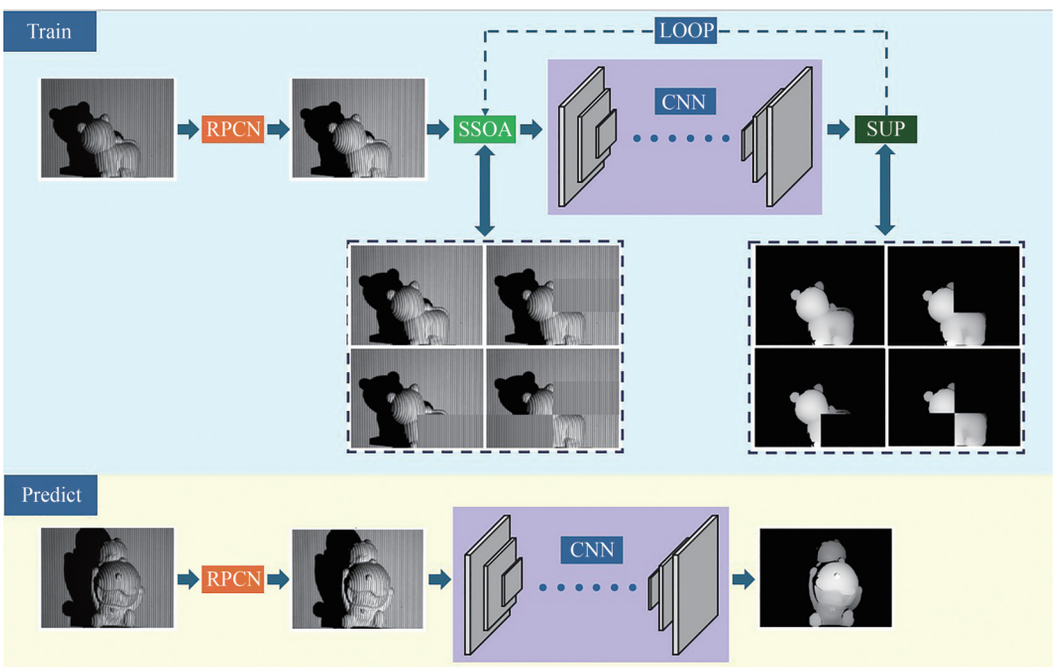


图6 添加RPCN和SSOA后的训练和预测过程

Fig.6 Training and prediction process after adding RPCN and SSOA

训练细节如下,实验数据来自分辨率为480×640的单帧条纹投影图像,训练时统一下采样至240×320,以平衡计算效率与特征表达能力。所有实验均在相同硬件环境下进行,训练轮数固定为100轮,优化器采用AdamW,损失函数为L1Loss。

3 实验结果与分析

3.1 归一化策略对比

在未进行归一化处理的情况下,模型在各项指标上的表现有限结果见表1和图7。

表1 归一化策略对比实验结果

Table 1 Comparison of normalization strategy experimental results

Model	Normalize	MAE	RMSE	SSIM	PSNR
UNet (base=32)	—	1.948 924	6.534 913	0.967 8	28.48 dB
	Min-Max	1.858 128	6.269 591	**0.971 2**	28.81 dB
	Z-score	2.014 634	6.553 381	0.965 9	28.48 dB
	Robust	1.838 592	6.240 499	0.971 0	28.83 dB
	Gamma	1.812 360	6.260 796	0.967 6	28.80 dB
	RPCN	**1.600 360**	**5.967 960**	0.971 0	**29.25 dB**

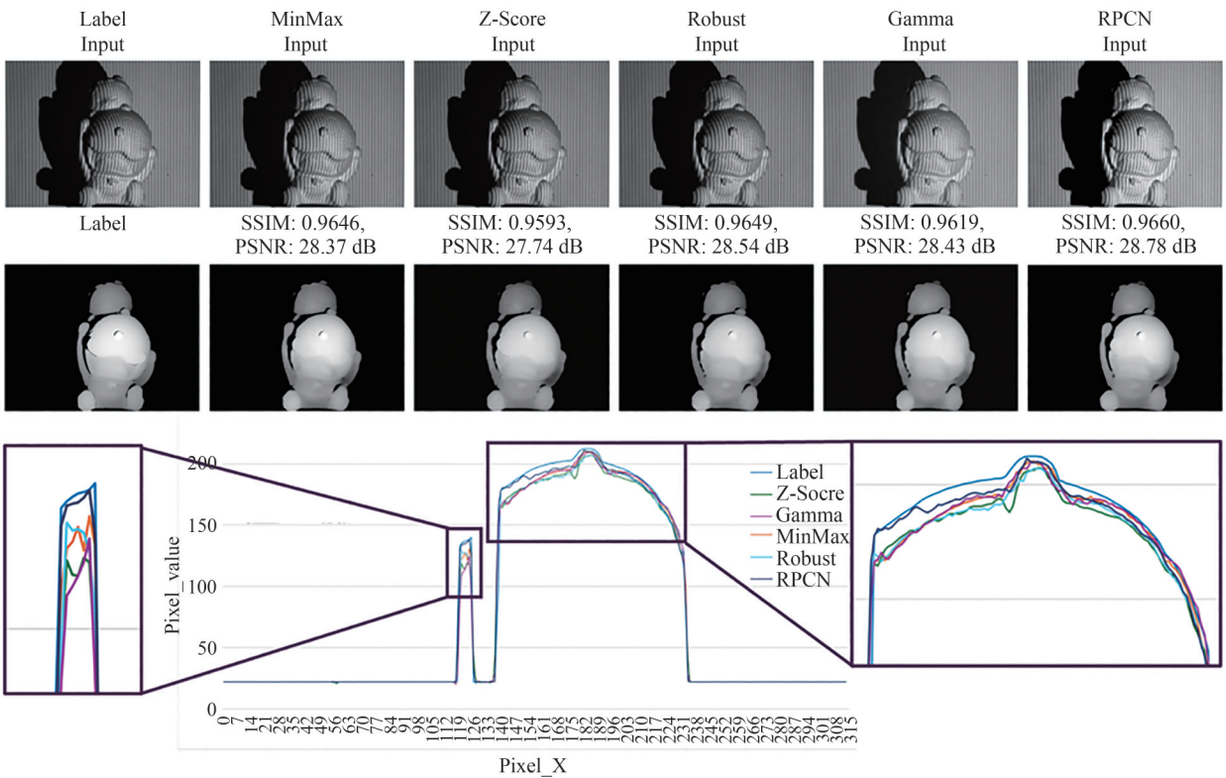


图7　不同归一化方法预测结果对比

Fig.7　Comparison of prediction results using different normalization methods

Min-Max归一化在整体上取得了提升，尤其在SSIM（0.971 2）和PSNR（28.81 dB）指标上取得了明显提升。该方法通过将数据映射到[0, 1]区间，提高了训练过程的稳定性，从而提升网络的重建精度。

Z-score标准化虽然在一些任务中常被用于加速训练，但在本实验中表现最差，其效果甚至略低于无归一化。这可能是由于条纹投影图像具有非对称分布和复杂光照背景，直接进行均值-方差标准化破坏了条纹相位分布的相对关系，导致网络对条纹特征的提取反而受损。

Robust归一化相较于Z-score有所提升，相较于无归一化结果各指标有一定增强，说明该方法对异常值的抑制作用有效，单难以充分利用条纹信号中的细微结构特征，因此对整体精度提升有限。

Gamma归一化在MAE（1.812 4）上表现出一定优势，这表明通过非线性映射改善图像亮度分布，能够一定程度上增强暗部区域细节，有助于网络学习到更清晰的条纹细节。

RPCN归一化方法在四项评价指标上均取得了显著提升，MAE降低至1.600 4，RMSE降至5.968 0，SSIM提升至0.971 0，PSNR提升至29.25 dB。与传统归一化方法相比，RPCN不仅有效降低了像素级误差，同时在结构相似性和峰值信噪比方面实现了同步优化。这一结果表明，RPCN能够更好地保留与增强图像暗部细节，提升整体对比度，从而促进网络对条纹信息的充分学习，最终显著提高重建精度。

为了进一步验证RPCN归一化各个子模块的有效性，本文以UNet作为基础模型，进行了分步消融实验。其中，Step1对应区间裁剪，Step2为线性映射至[0,1]，Step3为幂次变换，Step4为分段区间映射。从实验结果如表2所示，各步骤在不同指标上均体现出渐进式的性能提升规律。

区间裁剪在去除极端像素值的同时抑制异常亮度干扰，使模型在MAE与RMSE上均有所改善，限制输入动态范围有助于网络集中学习主要条纹区域的有效特征；线性映射至[0,1]提升了结构性指标，SSIM由0.967 8升至0.971 2，PSNR从28.48 dB升至28.81 dB，有效增强图像对比一致性，提高网络还原条纹相位结构的能力，是提升SSIM的关键环节；幂次变换通过非线性调整亮度分布强化暗部细节，有助于降低像素误差，实验中MAE从1.858 1降至1.812 4，RMSE下降至6.260 8，但局部纹理对比略有损失，导致SSIM轻微下降；分段区间映射实现多区间适应增强，弥补幂次变换带来的结构性损失，同时继续优化误差指标，达到最佳整体性能。

表 2 RPCN 归一化消融实验

Table 2 RPCN normalization ablation experiment

Model	Step1	Step2	Step3	Step4	MAE	RMSE	SSIM	PSNR
UNet (base=32)	×	×	×	×	1.948 924	6.534 913	0.967 8	28.48 dB
	×	√	×	×	1.858 128	6.269 591	**0.971 2**	28.81 dB
	×	√	√	×	1.812 360	6.260 796	0.967 6	28.80 dB
	×	√	√	√	1.751 553	6.259 093	0.970 4	28.78 dB
	√	×	×	×	1.885 312	6.354 612	0.968 8	28.68 dB
	√	√	×	×	1.755 622	6.137 174	0.971 1	29.05 dB
	√	√	√	×	1.712 098	5.981 053	0.970 3	29.21 dB
	√	√	√	√	**1.600 360**	**5.967 960**	0.971 0	**29.25 dB**

综合来看，RPCN 归一化的四步结构呈现出由线性到非线性、由整体到分段的层级优化特征。区间裁剪奠定稳定基础，线性映射确保全局一致性，幂次变换强化暗部细节，分段映射实现全区域适应优化，使 RPCN 在多项指标上均优于传统归一化方法，充分验证了其在条纹投影三维重建任务中的有效性。

3.2 SSOA 概率对比

选择不使用 SSOA 的 UNet(base=32)模型作为性能基准，见表 3 和图 8。相比基线结果有

当 $p=\frac{1}{3}$ 时，模型性能明显下降，MAE 升至 2.160 7，RMSE 增至 6.879 0，PSNR 降至 27.96 dB。这表明过高的遮挡概率会破坏原始条纹结构，使得网络难以有效学习条纹特征，甚至引入过度的数据噪声；

当 $p=\frac{1}{6}$ 时，模型表现显著优于基线，MAE 降至 1.589 5，RMSE 降至 5.894 6，SSIM 升至 0.971 1，PSNR 升至 29.35 dB，适度的遮挡增强有效提升模型对缺失区域的适应性，通过数据多样性显著改善重建精度；

当 $p=\frac{1}{9}$ 时，模型同样获得较优结果，MAE 降至 1.613 2，RMSE 降至 5.897 3，SSIM 升至 0.971 1，PSNR 升至 29.38 dB，该设置在结构相似性与信噪比方面实现了最佳平衡；

表 3 SSOA 概率对比实验结果

Table 3 Comparison experiment results of SSOA probability

Model	p	MAE	RMSE	SSIM	PSNR
UNet (base=32)	—	1.948 924	6.534 913	0.967 8	28.48 dB
	1/3	2.160 693	6.878 972	0.965 4	27.96 dB
	1/6	**1.589 545**	**5.894 574**	0.969 9	29.35 dB
	1/9	1.613 174	5.897 283	**0.971 1**	**29.38 dB**
	1/12	1.945 093	6.448 927	0.965 6	28.52 dB

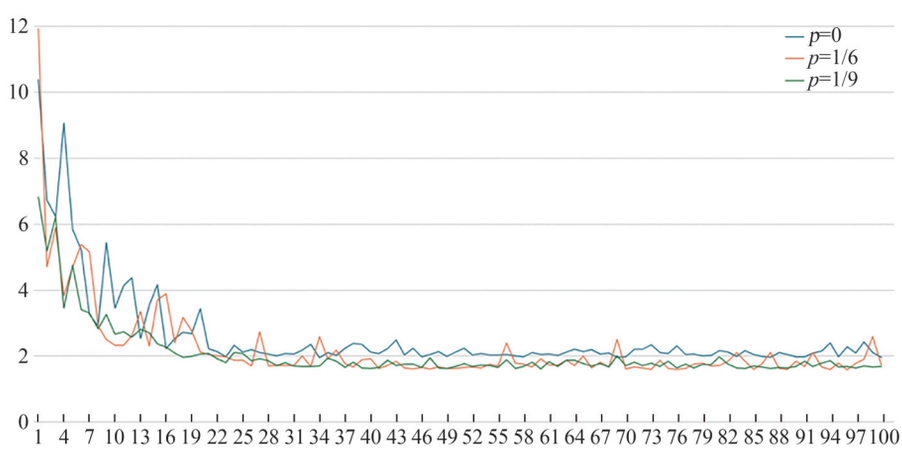


图 8 不同 SSOA 遮挡概率(1/6，1/9)与无增强条件下的 Test Loss 收敛曲线

Fig.8 Convergence curves of test loss under different SSOA occlusion probabilities (1/6, 1/9) and no enhancement conditions

当$p=\frac{1}{12}$时，增强效果几乎消失，与基线结果接近，表明过低的遮挡概率不足以为网络提供有效增强，难以发挥数据增强的作用。

综合分析可见，SSOA在合适的概率设置下能够显著增强模型性能。尤其是$p=\frac{1}{6}$与$p=\frac{1}{9}$两种情况均表现优异，前者在MAE和RMSE方面更具优势，后者在SSIM和PSNR方面略胜一筹。进一步结合Test Loss曲线的收敛趋势可知，当$p=\frac{1}{9}$时，模型在训练过程中表现出更稳定的下降特征，因而被选定为后续协同优化阶段的最优数据增强概率。

3.3 RPCN与SSOA协同优化

实验结果表明，不同优化策略在各项指标上均展现出差异化优势：在引入RPCN后，模型在MAE指标上取得最佳表现；采用SSOA时，则在RMSE与PSNR指标上表现更优；而当RPCN与SSOA联合使用时，则在SSIM指标上达到最优水平。总体来看，无论采用哪种优化方法，模型性能均显著优于基线UNet，以及针对条纹投影三维重建任务优化的DD-Inceptionv2-UNet，见表4，表5。

表4 协同优化UNet试验结果

Table 4 Collaborative optimization of UNet experimental results

Model	Optimize	MAE	RMSE	SSIM	PSNR
DD-Inceptionv2-UNet[17]	—	1.832 585	6.254 072	0.970 0	28.86 dB
UNet (base=32)	—	1.948 924	6.534 913	0.967 8	28.48 dB
UNet (base=32)	RPCN	**1.600 360**	5.967 960	0.971 0	29.25 dB
UNet (base=32)	SSOA	1.613 174	**5.897 283**	0.971 1	**29.38 dB**
UNet (base=32)	RPCN+SSOA	1.653 553	6.055 997	**0.971 8**	29.11 dB

表5 协同优化UNet++、DeepLabV3+试验结果

Table 5 Collaborative optimization of UNet and DeepLabV3+ experimental results

Model	optimize	MAE	RMSE	SSIM	PSNR
UNet++ (base=32)	—	1.770 640	5.642 854	0.977 2	29.73 dB
UNet++ (base=32)	RPCN	**1.508 020**	5.302 199	0.978 3	30.30 dB
UNet++ (base=32)	SSOA	1.525 771	5.288 515	0.977 5	30.31 dB
UNet++ (base=32)	RPCN+SSOA	1.573 725	**5.232 229**	**0.979 1**	**30.34 dB**
DeepLabV3+(MoblieNet)	—	2.415 629	8.326 917	0.945 1	26.27 dB
DeepLabV3+(MoblieNet)	RPCN	**2.319 025**	**8.175 125**	**0.945 7**	**26.41 dB**
DeepLabV3+(MoblieNet)	SSOA	2.603 039	8.471 635	0.942 5	26.12 dB
DeepLabV3+(MoblieNet)	RPCN+SSOA	2.455 791	8.353 243	0.941 6	26.22 dB

具体而言，RPCN通过裁剪极端亮度值、幂次变换与分段映射优化像素分布，有效抑制了过曝与欠曝区域的亮度失衡，强化了条纹细节与灰度层次，使模型在像素级误差MAE方面得到显著改善。相较之下，SSOA通过规则化遮挡与条纹一致性填补增强了数据多样性，提升了模型在复杂场景中的泛化能力与亮度重建稳定性，因此在整体误差RMSE与信噪比PSNR指标上表现更优。而两者结合时，前者提供像素分布优化，后者增强结构一致性，二者互为补充，使得模型在结构相似性SSIM上取得最优效果，体现了局部像素校正与全局结构平衡的协同优势。

在进一步的实验中，引入了UNet++和DeepLabV3+两种网络进行对比分析。对于结构更复杂且具备更强特征融合能力的UNet++，RPCN与SSOA的协同优化实现了最优的综合性能。相比之下，基于MobileNet的轻量化DeepLabV3+网络由于特征表达能力有限，在引入SSOA后可能增加数据复杂性，反而使模型更难以学习有效规律，导致拟合效果下降。

具体而言，MobileNetV2以深度可分离卷积与倒残差结构为核心设计理念，以显著减少参数量和计算复杂度。该结构在高效性与轻量化方面具有明显优势，但同时也削弱了网络在多尺度纹理细节与复杂结构特

征上的表达能力。当SSOA增强策略引入较强的条纹遮挡与强度扰动时，MobileNetV2主干在浅层特征提取阶段的通道表达受限，导致网络对局部结构和相位纹理的敏感度下降。此外，DeepLabV3+依赖空洞卷积和ASPP模块进行多尺度特征聚合，但其特征融合主要作用于高层语义空间，对低层细节变化的适应性较弱，因此在面对增强后高噪声或高遮挡样本时，模型的特征分辨能力进一步受限，出现了拟合困难与性能波动。

因此，对于UNet网络结构，不同优化策略在各项指标上各具优势，应根据具体任务需求灵活选择；对于UNet++，RPCN与SSOA的协同优化能够充分发挥其特征融合潜力，实现最优的综合性能；而对于轻量化网络或特征学习能力相对有限的模型，单独采用RPCN归一化更为适宜，它不仅能够保持训练过程的稳定性，还能在一定程度上提升整体性能。

3.4 泛化与鲁棒性验证实验

为进一步验证所提优化方法的鲁棒性与泛化能力，在原有公开数据集实验的基础上，增加了在Blender系统自建数据集上的泛化实验，实验结果如表6所示。

表6 泛化与鲁棒性验证实验结果

Table 6 Generalization and robustness verification experimental results

Model	Optimize	MAE	RMSE	SSIM	PSNR
UNet (base=32)	—	0.042 986	0.153 151	0.913 8	24.75 dB
UNet (base=32)	RPCN	0.036 642	0.138 912	0.934 7	25.67 dB
UNet (base=32)	SSOA	0.036 714	**0.131 191**	0.918 4	**26.09 dB**
UNet (base=32)	RPCN+SSOA	**0.033 676**	0.134 677	**0.952 2**	25.87 dB
UNet++(base=32)	—	0.033 432	0.135 403	0.932 8	25.84 dB
UNet++(base=32)	RPCN	0.031 920	0.131 490	0.961 2	26.19 dB
UNet++(base=32)	SSOA	0.029 648	0.127 220	0.962 5	26.35 dB
UNet++(base=32)	RPCN+SSOA	**0.029 076**	**0.125 657**	**0.967 8**	**26.48 dB**
DeepLabV3+(MoblieNet)	—	0.049 993	0.162 215	0.950 4	24.26 dB
DeepLabV3+(MoblieNet)	RPCN	**0.042 185**	**0.151 943**	**0.954 5**	**24.83 dB**
DeepLabV3+(MoblieNet)	SSOA	0.057 308	0.173 859	0.939 0	23.68 dB
DeepLabV3+(MoblieNet)	RPCN+SSOA	0.045 904	0.158 225	0.945 4	24.45 dB

各模型在不同优化策略下的性能表现与前述真实数据集实验结果相近，进一步验证了所提方法的稳定性与鲁棒性。值得注意的是，由于数据集中训练样本数量相对较少，SSOA数据增强策略在缓解过拟合与提升模型泛化能力方面表现出更加显著的效果，在多数指标上较未增强模型均有明显提升。而RPCN与SSOA的联合使用仍保持最优综合性能，表现出较强的结构一致性与误差控制能力，表明该归一化与增强策略在不同数据域下均具备良好的适应性与鲁棒性。

4 结论

针对单帧条纹投影深度估计中训练数据有限、传统预处理方法适配性不足的问题，提出了结合区间裁剪幂次归一化(RPCN)与结构化条纹遮挡增强(SSOA)的新型数据预处理策略，并通过多组实验系统验证了其有效性与通用性。RPCN能够通过异常值抑制、幂次映射和分段归一化，显著改善像素分布与条纹暗部细节，在UNet模型上将MAE降至1.600 4、PSNR提升至29.25 dB；SSOA则通过物理一致性的遮挡增强提升了样本多样性，在最优参数下使UNet的RMSE降至5.897 3、PSNR升至29.38 dB，并提升了收敛稳定性。进一步实验发现，不同网络结构对策略的适配性存在差异，在特征融合能力更强的UNet++上，RPCN与SSOA协同优化可实现最优平衡，MAE1.573 7，RMSE5.232 2，SSIM0.979 1，PSNR30.34 dB；而在轻量化的DeepLabV3+中，单独使用RPCN更能兼顾性能提升与稳定性。综上，本文提出的预处理策略在保证条纹物理一致性的前提下显著提升了模型精度，同时具备良好的通用性和差异化适配潜力。

参考文献

[1] LIU H, YAN N, SHAO B, et al. Deep learning in fringe projection: a review[J]. Neurocomputing, 2024, 581: 127493.

[2] NGUYEN H, WANG Y, WANG Z. Single-shot 3D shape reconstruction using structured light and deep convolutional neural networks[J]. Sensors, 2020, 20(13):3718.

[3] GSVIK K J, ROBBERSMYR K G, VADSETH T. Online dimensional control of rolled steel profiles using projected fringes[J]. The International Journal of Advanced Manufacturing Technology, 2020, 107(3): 1725-1730.

[4] WANG M, ZHANG Q, LI Q, et al. Research on morphology detection of metal additive manufacturing process based on fringe projection and binocular vision[J]. Applied Sciences, 2022, 12(18): 9232.

[5] MUMUNI A, MUMUNI F. Data augmentation: a comprehensive survey of modern approaches[J]. Array, 2022, 16: 100258.

[6] WU Z, ZHANG Q. High-speed 3D topography measurement based on fringe projection: a review [J]. Laser & Optoelectronics Progress, 2023, 60(8): 0811001.
吴周杰,张启灿.基于条纹投影的高速三维形貌测量技术发展综述[J].激光与光电子学进展,2023,60(8): 0811001.

[7] ZHANG S. Recent progresses on real-time 3D shape measurement using digital fringe projection techniques[J]. Optics and Lasers in Engineering, 2010, 48(2): 149-158.

[8] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, 6(1): 1-48.

[9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25: 84-90.

[10] FALLAH E A OEL. Investigating the impact of different representations of data on neural network and regression[D]. College of Arts and Sciences, Universiti Utara Malaysia, 2008.

[11] YUN J, KIM H. Znorm: Z-score gradient normalization for deep neural networks[J]. arXiv preprint arXiv:2408.01215, 2024. https://arxiv.org/abs/2408.01215.

[12] ZHAO Y, WONG L, GOH W W B. How to do quantile normalization correctly for gene expression data analyses[J]. Scientific Reports, 2020, 10(1): 15534.

[13] WANG Y, LIU Z, LIU J, et al. Low-light image enhancement with illumination-aware gamma correction and complete image modelling network[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 13128-13137.

[14] WANG F, WANG C, GUAN Q. Single-shot fringe projection profilometry based on deep learning and computer graphics [J]. Optics Express, 2021, 29(6): 8024-8040.

[15] ZHANG S. High-speed 3D shape measurement with structured light methods: a review [J]. Optics and Lasers in Engineering, 2018, 106: 119-131.

[16] NGUYEN H, WANG Y, WANG Z. Single-shot 3D shape reconstruction using structured light and deep convolutional neural networks[J]. Sensors. 2020; 20(13):3718.

[17] WANG L, XUE W, WANG C, et al. Depth estimation from a single-shot fringe pattern based on DD-Inceptionv2-UNet [J]. Applied Optics, 2023, 62(34): 9144-9155.

Data Augmentation and Normalization Optimization for Single-frame Fringe Projection Depth Estimation

WANG Linlin, LI Bingzhen, GUO Wanyi, GAO Qian, WANG Chuanyun

(*School of Artificial Intelligence*, *Shenyang Aerospace University*, *Shenyang* 110136, *China*)

Abstract: This study aims to enhance the accuracy, robustness, and training stability of deep-learning-based single-frame fringe projection profilometry. It focuses on two practical limitations commonly encountered in existing methods: the incompatibility between conventional intensity normalization schemes and the highly skewed brightness distribution of fringe images, and the lack of physically consistent data augmentation strategies under limited labeled data conditions. The objective is to develop a preprocessing framework that preserves effective fringe information, improves model adaptability across different neural network architectures, and provides a model-aware optimization strategy without altering network structures or learning objectives.

To achieve this goal, a unified preprocessing framework integrating Range-clipped Piecewise Contrast Normalization (RPCN) and Structured Stripe Occlusion Augmentation (SSOA) is proposed. RPCN is specifically developed to address the intrinsic characteristics of fringe images, which typically exhibit a large proportion of low-intensity pixels and a small number of extreme outliers caused by saturation, specular reflection, or sensor noise. The method combines interval clipping to suppress extreme pixel values, power-law mapping to enhance contrast in dark fringe regions, and piecewise remapping to separate shadow-dominated regions from effective fringe regions. This method redistributes pixel intensities into a range more suitable for neural network learning while maintaining stripe continuity and contrast stability. In parallel, SSOA is introduced to improve data diversity in a physically meaningful manner. Instead of directly masking or zeroing pixels, SSOA generates structured occlusions whose missing regions are reconstructed using neighboring stripe patterns, thereby preserving spatial continuity and phase-related information. This augmentation strategy simulates realistic measurement disturbances such as partial occlusion, surface defects, and local signal loss while maintaining physical consistency. The proposed framework is evaluated on multiple network architectures with varying capacities, including UNet, UNet++, and DeepLabv3+ with a MobileNetV2 backbone, to assess its generality, adaptability, and sensitivity to network complexity.

Extensive experiments demonstrate that single-frame depth estimation performance is significantly improved across different network architectures by the proposed preprocessing strategies, while distinct effects are observed depending on model capacity. On the UNet architecture, the mean absolute error was reduced by 17.9% through the application of RPCN, indicating enhanced feature representation under limited normalization adaptability. When SSOA was applied independently, the root mean square error was reduced by 9.8%, and the peak signal-to-noise ratio was increased from 28.48 dB to 29.38 dB, reflecting improved robustness to structured disturbances. The joint application of RPCN and SSOA further increased the structural similarity index from 0.967 8 to 0.971 8. For the more powerful UNet++ architecture, the combined RPCN-SSOA framework yielded the best overall performance, with the structural similarity index increased from 0.977 2 to 0.979 1 and the peak signal-to-noise ratio improved from 29.73 dB to 30.34 dB. In contrast, for lightweight models such as DeepLabv3+ with a MobileNetV2 backbone, performance gains were constrained when SSOA was introduced, as the increased data complexity hindered effective learning, whereas a more favorable balance between accuracy improvement and training stability was achieved using RPCN alone. Additional evaluations across multiple test scenes further confirm that prediction stability is consistently improved under varying surface reflectance and illumination conditions. These results indicate that preprocessing strategies interact strongly with network capacity and architectural design, and that inappropriate augmentation may limit performance gains in compact models.

The proposed RPCN-SSOA preprocessing framework provides a practical and adaptable solution for improving deep-learning-based single-frame fringe projection profilometry. By aligning intensity normalization with fringe-specific brightness distributions and enforcing physical consistency in data augmentation, the framework enhances depth estimation accuracy and robustness without increasing model complexity. The results highlight the importance of model-aware preprocessing demonstrate that normalization and augmentation strategies should be tailored to network capacity rather than uniformly applied. This work offers a feasible preprocessing paradigm and practical guidance for fringe-based three-dimensional measurement systems in both research and engineering applications.

Key words: Fringe projection; Depth estimation; Data augmentation; Normalization; Data preprocessing

OCIS Codes: 100.1160; 100.2650; 110.6890

CSTR: 32255.14.gzxb20265502.0210001

Foundation item: National Natural Science Foundation of China (62376165), Basic Scientific Research Project of Liaoning Provincial Department of Education (LJ212410143055)