

# 基于 MultiResHNet 的结构光三维重建技术

杨丽婷\*, 刘孝良, 储修祥, 周路

浙江农林大学光机电工程学院, 浙江 杭州 311300

**摘要** 随着深度学习和结构光条纹投影三维成像技术的发展, 直接从单幅条纹图中恢复物体的三维形状的研究近年来受到了多个领域的关注。提出改进的全局引导路径网络 MultiResHNet, 实现对单幅条纹图的 3D 形状重建, 将现有结构光学三维成像方案与深度卷积神经网络结合, 对仿真数据和实验数据分别进行了验证。实验结果表明, 所提方法预测的 3D 形状比已有的 U-Net 神经网络预测的 3D 形状更加准确, 误差更小, 精度更高。实验结果证明了所提技术的有效性和鲁棒性, 为后续的 3D 形状重建技术的提高提供了科学依据, 具有一定的参考和应用价值。

**关键词** 条纹图; 结构光; 卷积神经网络; 3D 形状; 机器视觉

中图分类号 O439 文献标志码 A

DOI: 10.3788/LOP223203

## Structured Light Three-Dimensional Reconstruction Technology Based on MultiResHNet

Yang Liting\*, Liu Xiaoliang, Chu Xiuxiang, Zhou Lu

School of Optical Mechanical and Electrical Engineering, Zhejiang A&F University,  
Hangzhou 311300, Zhejiang, China

**Abstract** Recently, advancement in deep learning and three-dimension (3D) imaging technology based on structured light fringe projection on the recovery of the 3D shape of objects from a single fringe image has attracted considerable attention. In this paper, MultiResHNet, an improved global guided path network, is proposed for the 3D shape reconstruction of a single fringe pattern. Herein, the simulation data and experimental data are verified by combining the existing structural optics 3D imaging scheme with the deep convolutional neural network. The experimental results show that the proposed method accurately predicts the 3D shape with lesser errors compared with the existing U-Net neural network. Therefore, our experiments prove the effectiveness and robustness of the proposed method, providing a scientific basis for the improvement of subsequent 3D shape reconstruction with a certain reference value and an application value.

**Key words** fringe pattern; structured light; convolutional neural network; 3D shape; machine vision

## 1 引言

近年来, 计算机技术、数字成像技术、激光技术的快速发展推动了三维成像技术的逐渐成熟<sup>[1-3]</sup>, 其中非接触式光学三维重建技术由于具有方便快捷、测量范围广、精度高和不损伤被测物体等优势, 在各个领域有着越来越多的应用需求<sup>[4]</sup>。

传统的结构光测量方法面临着一些问题, 如测量过程繁琐, 不能快速三维成像。但是, 随着深度学习的快速发展, 深度学习神经网络在图像分类、自然语言处理、语音识别、医疗影像处理等方面取得了重大进展<sup>[5]</sup>。由于神经网络有着强大的特征提取能力, 学者

们借助神经网络充分提取条纹、散斑等图像的特征, 进而恢复出物体的三维形状。最近几年, 基于深度学习的三维成像和形状重建方法层出不穷, 成为结构光三维测量技术在效率、精度等方面取得的新突破<sup>[6]</sup>。学者们利用神经网络取代了现有结构光学三维成像方法的一些实验步骤, 证明了该方法的可行性。例如 Spoorthi 等<sup>[7]</sup>用相位神经网络(PhaseNet)实现相位展开, 将原始包裹相位结合神经网络输出的条纹级次, 计算最终的展开相位。改进的神经网络 PhaseNet 2.0<sup>[8]</sup>可以直接将含噪声的包裹相位映射到去噪的绝对相位。Wang 等<sup>[9]</sup>利用具有 U-Net 结构的神经网络, 该网络可以省略计算条纹级次的步骤, 直接预测与包裹相

收稿日期: 2022-11-29; 修回日期: 2022-12-16; 录用日期: 2022-12-22; 网络首发日期: 2023-01-05

通信作者: \*542655072@qq.com

位对应的展开相位。

深度学习与传统的三维成像方法相比,可以实现端到端学习,神经网络可以一次性学习原始图像特征信息和所需参数信息之间的映射关系,相比分步学习分案,减小了人为干预产生的误差,提高了实验效率。许多学者如 Nguyen 等<sup>[10]</sup>将条纹图直接映射到其对应三维形状的端到端神经网络,并对全卷积网络(FCN)、自动编码器网络(AEN)和 U-Net 的 3 种深度卷积神经网络进行了训练,显示了 3 个网络预测的最佳三维重建结果。Machineni 等<sup>[11]</sup>利用多分辨率相似度评估的卷积神经网络(CNN),直接重建相应变形条纹图像中物体的形状。在低信噪比、低条纹密度和高动态范围等多种困难条件下,他们提出的方法可以取得很好的结果。van der Jeught 等<sup>[12]</sup>通过模拟条纹图像和深度图像训练网络,直接从单幅条纹图像中提取物体的高度信息。

对于三维形状重建,深度学习提供了一种可行且高效的方式,可以减轻整体计算负担。本文结合了生物学图像分割的神经网络 MultiResUNet<sup>[13]</sup>和全局引导路径<sup>[14]</sup>,提出了 MultiResHNet 神经网络,并将其应用在条纹投影三维形状重建中。所提网络模型通过检索和整合多个上下文层次的特征,对三维形状进行了更准确的重建,与基于相位提取和三角测量计算的最先进的 3D 形状重建技术相比,过程虽然简单,但仍然鲁棒,与使用 U-Net 进行三维形状重建的方法相比,可以提高预测的准确性。

## 2 算法基本原理

数字光栅投影的结构光三维测量系统由投影设备、被测物体、计算机及相机构成,如图 1 所示。基本思想是:计算机生成实验所需的结构光编码图案,通过投影设备,将编码图案依次投影到测试的物体表面,投射的编码图案受到被测物体表面高度的影响而发生形变,物体表面的三维坐标信息包含在调制的条纹图案

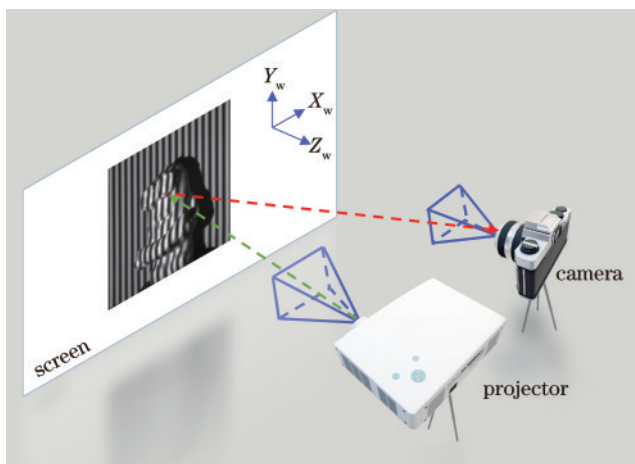


图 1 传统结构光三维测量系统

Fig. 1 Traditional structured light 3D measurement system

中;然后,相机拍摄每一幅被调制的条纹图案,并将这些图片传输到计算机上,提取相位,获得由相位-高度映射关系测量的物体空间的高度信息,从而对物体形貌进行三维重构。本文通过相移法获得被测物体的包裹相位,再结合格雷码编码得到物体的绝对相位,实现对被测物的三维形状重建<sup>[15-16]</sup>,结果作为网络训练数据集的输出标签。

### 2.1 基于相移法和格雷码编码的相位恢复

相移法的基本原理是在被测物体上投射一系列相移条纹图像,从拍摄到的物体变形的条纹图上获取被测物的相位主值。假设相移条纹图像为标准正弦分布,则其光强分布函数为

$$I_n(x, y) = A(x, y) + B(x, y) \times \cos\left[\varphi(x, y) + \frac{2\pi(n-1)}{N}\right], \quad (1)$$

式中: $N(N \geq 3)$ 为总的相移步数, $n$ 表示进行了第 $n$ 次相移, $n = 1, 2, \dots, N$ ;  $(x, y)$ 表示像素的空间坐标; $A(x, y)$ 表示背景光强, $B(x, y)$ 表示调制幅度, $\varphi(x, y)$ 表示相位主值, $A(x, y)$ 、 $B(x, y)$ 和 $\varphi(x, y)$ 是 3 个未知量。可以利用式(1), $N$ 步相移算法求得的相位主值 $\varphi(x, y)$ 为

$$\varphi(x, y) = \arctan\left[\frac{\sum_{n=1}^N I_n(x, y) \sin(2n\pi/N)}{\sum_{n=1}^N I_n(x, y) \cos(2n\pi/N)}\right]. \quad (2)$$

根据相移原理得到的相位主值 $\varphi(x, y)$ 是经过反正切运算的,以至于求解得到的值在 $(-\pi, \pi)$ 之间,呈周期分布,且相邻周期存在 $2\pi$ 相位跳变,该相位称为包裹相位,使得被测物体在图像中的每个像素点的相位分布失去了全场唯一性,需要对其进行相位展开得到连续的绝对相位值 $\phi(x, y)$ 。

利用格雷码对每个像素点的周期级次信息进行编码,如图 2 所示,将每个像素点的格雷码值转换为十进制数并将其映射到自然序列,得到该包裹相位的条纹级次,然后计算图像的绝对相位值,数学表达式为

$$\phi(x, y) = \varphi(x, y) + 2\pi k(x, y), \quad (3)$$

式中: $\phi(x, y)$ 为展开相位或去包裹相位; $k(x, y)$ 为光栅条纹的级次。利用展开后的绝对相位和高度之间的映射关系,以及基于系统标定得到的所需的各项系数,实现物体三维形状重建<sup>[17]</sup>。

### 2.2 网络结构

使用的 MultiResHNet 神经网络模型包括编码器路径、解码器路径和全局引导路径。网络的结构包括 3 个主要模块,即 Multi Block 模块、Res Path 模块和 Third 模块,如图 3 所示。其中,MultiRes 模块是主干,是 MultiResHNet 网络的核心结构。每个编码层对应一个解码层,编码器路径由卷积层和池化层组成,卷积

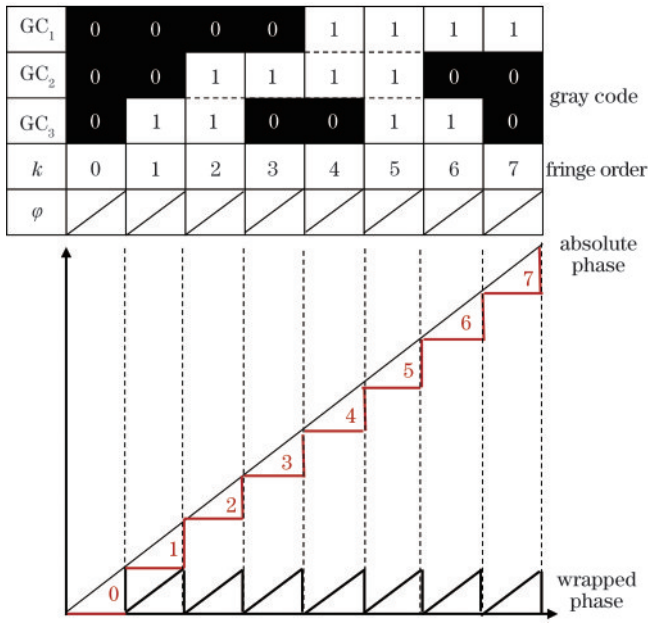


图 2 格雷码编码的相位展开原理

Fig. 2 Principle of phase unwrapping in Gray code

层负责获取图像局部特征,池化层对图像进行下采样并且将尺度不变特征传送到下一层。解码器对缩小后的特征图像进行上采样,然后对上采样后的图像进行卷积处理,完善物体的几何形状,弥补编码器中池化层将物体缩小造成的细节损失。本文在此基础上引入全局引导路径,通过从解码器路径中提取图像的附加信息来提高形状重建的精度。

为了增强网络的多分辨率分析能力,Multi Block 模块使用一系列更小、更轻便的  $3 \times 3$  卷积层,替换要求更高的  $5 \times 5$  和  $7 \times 7$  卷积层,如图 4(a)所示,第二和第三卷积模块的输出分别有效地逼近  $5 \times 5$  卷积和  $7 \times 7$  卷积运算,将三个卷积模块的输出连接在一起,从不同的尺度提取空间特征。Res Path 模块添加了残差连接,并引入了  $1 \times 1$  卷积层,如图 4(b)所示。沿着 4 个 Res Path 分别使用 4、3、2、1 个卷积块。Third 模块对 MultiRes 模块输出进行相应的  $1 \times 1$  卷积和信道深度为 1 的转置卷积,将感受野大小调整为 1,分辨率上采样为输入的大小,最后输出具有线性激活的  $1 \times 1$  卷积,如图 4(c)所示。

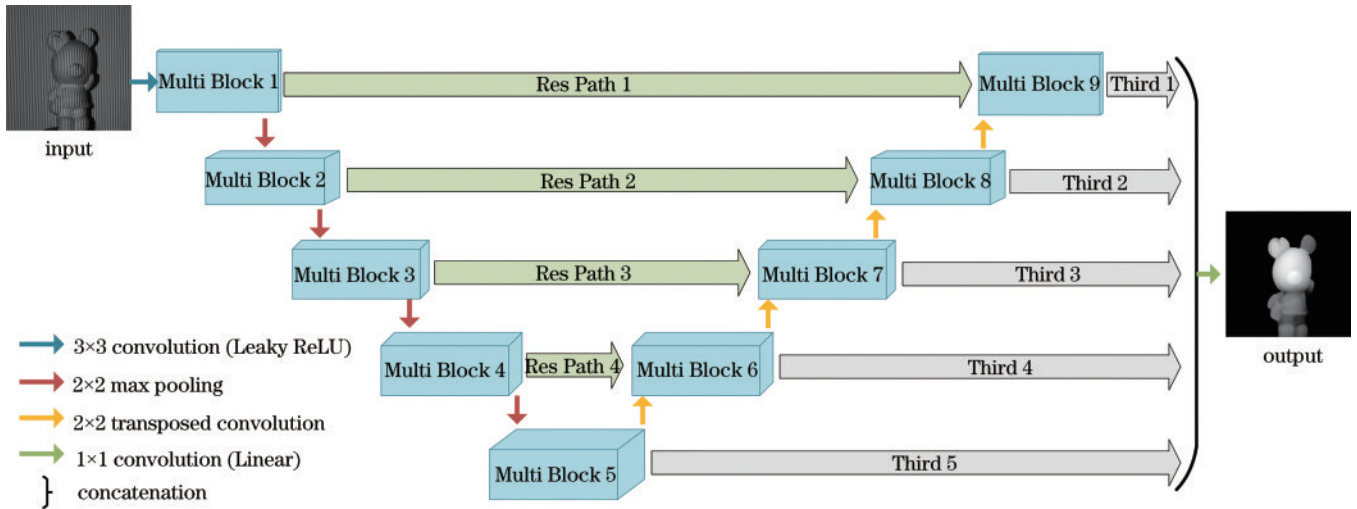


图 3 卷积神经网络结构

Fig. 3 Convolutional neural network structure

全局引导路径的网络结构如图 5 所示,从上到下依次为模块 Third 1 到 Third 5 的内部结构。由于模块 Third 1 与输入具有相同的分辨率,因此不需要再使用转置卷积进行上采样。全局引导路径可以对低分辨率到高分辨率的特征映射进行精确优化和训练。其次,对模块 Third 1~5 的输出进行进一步拼接和融合,进行最终预测,有助于提高输出的准确性。网络训练过程的总损失<sup>[18]</sup>为

$$L = w_c l_c + \sum_{n=1}^N w_n l_n, \quad (4)$$

式中:  $l_c$  表示路径中最后输出层的损失;  $l_n$  表示 Third 路径中其他输出层的损耗,  $n = 1, 2, \dots, N, N = 5$ 。  $w_c$  和  $w_n$  是每个损失相应的权重,在学习过程中使用反向

传播算法最小化学习损失,其中每个隐藏层<sup>[19]</sup>定义为

$$\mathbf{h} = g(\mathbf{W}^T \mathbf{x} + \mathbf{b}), \quad (5)$$

式中:  $\mathbf{x}$  是输入的向量,  $\mathbf{h}$  是输出向量;  $g$  是激活函数。矩阵形式的权重参数  $\mathbf{W}$  和向量形式的偏重参数  $\mathbf{b}$  通过训练过程得到优化。实验中采用 Leaky ReLU 非线性激活函数,表达式为

$$\mathbf{h}' = g(\mathbf{W}^T \mathbf{x} + \mathbf{b}) = \begin{cases} \mathbf{W}^T \mathbf{x} + \mathbf{b}, & \mathbf{W}^T \mathbf{x} + \mathbf{b} > \mathbf{0} \\ \alpha(\mathbf{W}^T \mathbf{x} + \mathbf{b}), & \mathbf{W}^T \mathbf{x} + \mathbf{b} \leq \mathbf{0} \end{cases}, \quad (6)$$

式中:  $\alpha$  为负斜率系数。

全局引导路径网络使用 Adam 优化器<sup>[20]</sup>控制网络的学习速率,损失函数使用均方误差(MSE)。采用两种常见的回调机制,即 Keras 库中的 LambdaCallback 和 ModelCheckpoint,来监控和保存学习的中间结果。

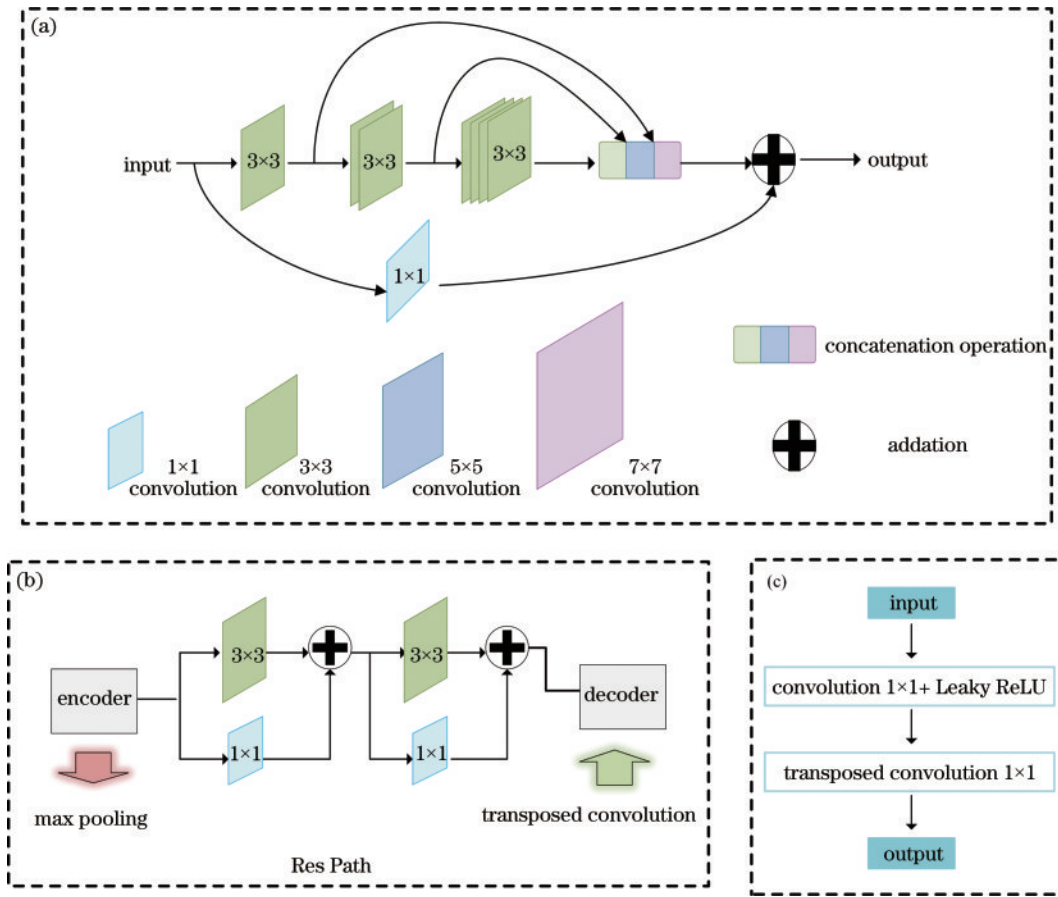


图 4 神经网络的三个模块。(a)Multi Block 模块;(b)Res Path 模块;(c)Third 模块

Fig. 4 Three modules of a neural network. (a) Multi Block module; (b) Res Path module; (c) Third module

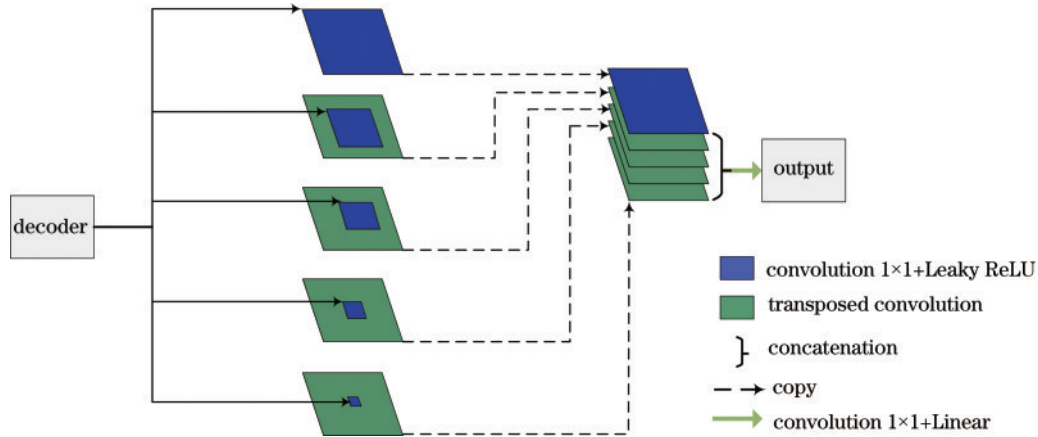


图 5 全局引导路径网络结构

Fig. 5 Global guidance path network structure

LambdaCallback 帮助保存带有更新参数的 3D 形状重建, ModelCheckpoint 设置为保存最佳优化模型。均方根误差(RMSE)用于监控优化过程, 模型是否被更新是基于验证数据集的 RMSE 的。

$$E_{\text{RMSE}} = \sqrt{\frac{1}{m} \sum_{i=1}^m [Y_i - f(x_i)]^2}, \quad (7)$$

式中:  $m$  是有效点的数量;  $Y_i$  指真实值;  $f(x_i)$  是预测值。

此外, 本研究还包括两个评价指标, 平均绝对误差 (MAE) 和结构相似性指数 (SSIM), 用于性能评估和预测结果与 3D 地面实况标签之间的比较。这两个指标的计算公式分别为

$$E_{\text{MAE}} = \frac{1}{m} \sum_{i=1}^m |Y_i - f(x_i)|, \quad (8)$$

$$I_{\text{SSIM}}(x, y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (9)$$

式中:  $\mu_x$  是  $x$  的平均值,  $\mu_y$  是  $y$  的平均值,  $\sigma_x^2$  是  $x$  的方差,  $\sigma_y^2$  是  $y$  的方差,  $\sigma_{xy}$  是  $x$  和  $y$  的协方差;  $c_1 = (k_1 L)^2$  和  $c_2 = (k_2 L)^2$  是用来维持稳定的常数;  $L$  是像素值的动态范围;  $k_1 = 0.01, k_2 = 0.03$ 。

### 3 实验与结果

通过模拟数据和实验数据验证了所提方法的有效性。在搭载 Nvidia GeForce RTX 2080 Ti 显卡、英特尔至强银牌 4110 处理器、16 GB 内存的服务器上进行网络训练和测试, 实验的网络架构是基于 Python 3.6.13 的 Keras 框架版本的, 采用了 Nvidia 的 cuDNN 深度神经网络库来加快训练过程。将投影仪和相机任意排列, 形成广义的布局。相机到感兴趣的场景的距离大约是 45 cm。在实验中, 使用一些大小大约为

4 cm 的小石膏雕塑作为对象, 其大小和表面性质适合于生成可靠和准确的三维数据集。

#### 3.1 仿真模拟

##### 3.1.1 数据集

为了制作训练网络所需要的数据集, 设计了随机表面地图生成器, 该生成器可以生成大小不同的随机矩阵, 再通过插值方法生成指定尺寸的模拟高度图。其中将模拟高度限制在 0~30 mm, 使用到的插值方法有线性插值、三次样条插值和三次 Hermite 插值等。仿真获得的数据集包括条纹投影式(1)生成的正弦条纹经物体调制后的变形条纹图及对应的深度图。实验共生成了 2500 对图片, 分别使用 2250 对数据作为训练集和 250 对数据作为验证集。数据集的部分模拟投影条纹图及对应的深度图如图 6 所示。

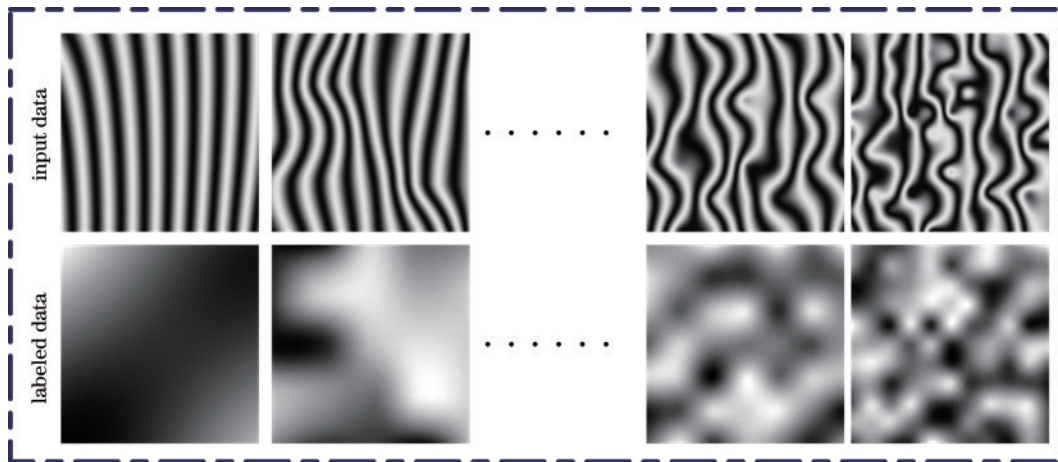


图6 训练数据集

Fig. 6 Training dataset

##### 3.1.2 三维测量结果

使用模拟数据集训练 U-Net 和 MultiResHNet 神经网络, 优化器设置为 Adam, 初始学习率为  $10^{-4}$ , 将 MSE 作为损失函数。为了得到更好的训练结果, 分析训练集、验证集的损失下降趋势, 调整批大小、训练次数和初始学习率。最终, 设定一次训练选取的样本数为 8, 训练总次数为 200, U-Net 和 MultiResHNet 两种方法的训练时间分别约为 1 h 和 2 h。网络输入的灰度图片大小是  $128 \times 128$ , 训练完成后, 使用 30 个测试样本对两种网络模型进行测试, 一次性输出对所有样本的测试结果。以 RMSE 作为衡量模型性能的标准, 训练过程中的训练集和验证集的 MSE 和 RMSE 如表 1 所示。

表 1 两种网络的模拟误差分析

Table 1 Simulation error analysis of the two networks

Parameter		U-Net	MultiResHNet
Training	MSE	$0.30 \times 10^{-1}$	$0.19 \times 10^{-1}$
	RMSE	$4.33 \times 10^{-2}$	$3.44 \times 10^{-2}$
Validation	MSE	$0.46 \times 10^{-1}$	$0.32 \times 10^{-1}$
	RMSE	$5.16 \times 10^{-2}$	$4.37 \times 10^{-2}$

图 7 是两个网络的测试结果。图 7(a) 和图 7(f) 分别是两个样本的真实深度图, 图 7(b) 和图 7(g) 是 U-Net 预测的两个样本的深度图, 图 7(c) 和图 7(h) 是两个样本真实深度图和 U-Net 预测的深度图之间的测试误差, 图 7(d) 和图 7(i) 是 MultiResHNet 预测的两个样本的深度图, 图 7(e) 和图 7(j) 是两个样本真实深度图和 MultiResHNet 预测的深度图之间的测试误差。从误差图可以看出, MultiResHNet 测试的整体误差小于 U-Net, 验证了所提方法的有效性。

由于在实际环境中采集的输入条纹图存在噪声, 在不改变训练方法的情况下, 对输入模拟条纹图添加噪声水平为 2% 的高斯噪声, 然后进一步测试了两种网络模型的性能, 训练时间分别为 2 h 和 3 h。加噪声数据集的部分模拟投影条纹图及其对应的深度图如图 8 所示, U-Net 和 MultiResHNet 对加噪声的测试样本进行处理得到的 RMSE 分别为 0.1414 和 0.1313。

图 9 是测试样本的三维形状图, 可以看出, 对于具有 2% 的高斯噪声的物体, MultiResHNet 和 U-Net 依然可以很好地恢复物体的三维形状, 进一步说明了所

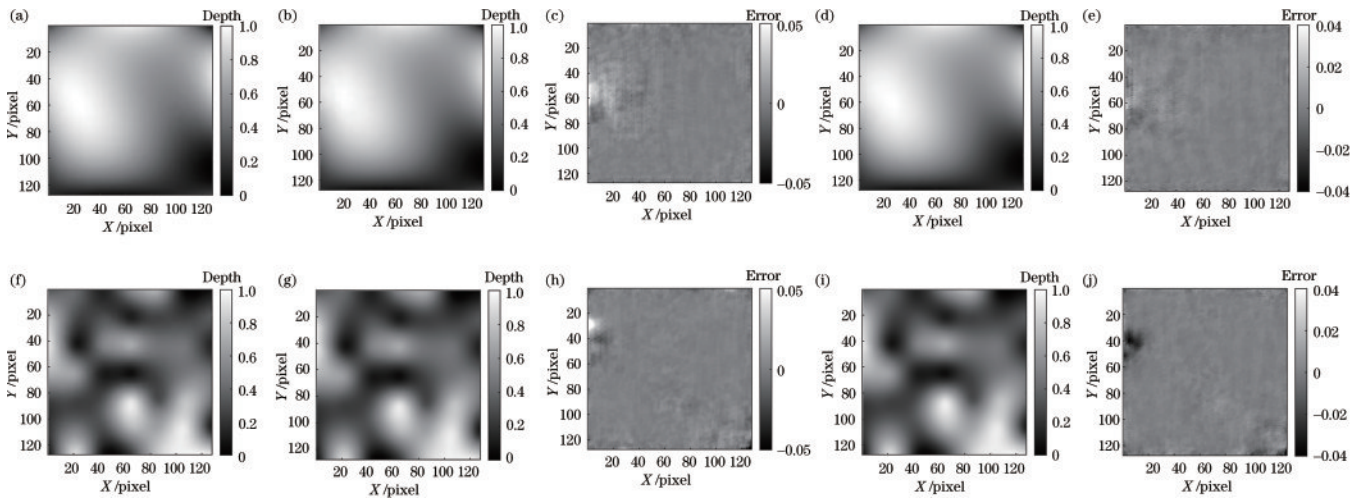


图7 物体的深度图和误差图。(a)(f)深度图;(b)(g)U-Net 预测的形貌图;(d)(i)MultiResHNet 预测的形貌图;(c)(h)U-Net 测量的误差图;(e)(j)MultiResHNet 测量的误差图

Fig. 7 Depth and error images of an object. (a)(f) Depth map; (b)(g) predicted 3D shape images predicted by U-Net; (d)(i) predicted 3D shape images predicted by MultiResHNet; (c)(h) error images measured by U-Net; (e)(j) error images measured by MultiResHNet

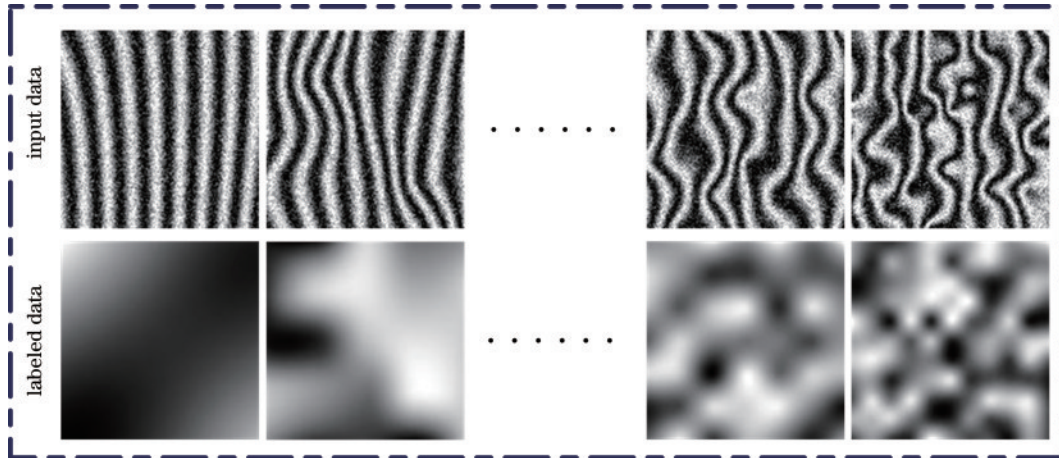


图8 加噪声的训练数据集

Fig. 8 Noised training dataset

提方法的可行性。其中图 9(a)是传统方法获得的物体实际三维形状,图 9(b)和图 9(d)分别对应 U-Net 和 MultiResHNet 预测的物体三维形状,图 9(c)和图 9(e)分别是网络预测值和地面真实值之间的误差。

### 3.2 实验验证

#### 3.2.1 数据集获取

在真实实验数据集上对所提算法进行可行性和网络性能的测试。图 10 列举了训练和验证时所使用的部分网络输入和输出的图像。实验数据是在真实场景下拍摄的小石膏雕塑的变形条纹图,共采集了 1120 组数据,数据中包含不同被测物或者相同被测物的不同姿态,每个样本有对应的传统三维成像方法获得的深度图。其中条纹的频率使用 48,获得了更多被测物体的特征信息,有助于神经网络的充分学习。

#### 3.2.2 训练、分析和评价

实验训练次数为 200,最小批量大小为 2,初始学习率设定为  $1 \times 10^{-4}$ 。在实验数据集上对 U-Net 和 MultiResHNet 神经网络进行了性能评估,两种方法的训练时间分别约为 5 h 和 9 h。完成训练后,在学习过程中获得的 RMSE 如图 11 所示。测试相同的数据集,实验结果表明,MultiResHNet 在验证集上的 RMSE 比 U-Net 的 RMSE 小,验证了 MultiResHNet 的可行性。

为了进一步评估两种网络模型在单帧图像三维重建方面的优势,对网络模型又增加了两种评价方法,即 MAE 和 SSIM。表 2 显示了用于单次 3D 形状重建的两神经网络模型的性能指标。从表 2 可知,MultiResHNet 模型总体要优于 U-Net,预测效果比 U-Net 好。

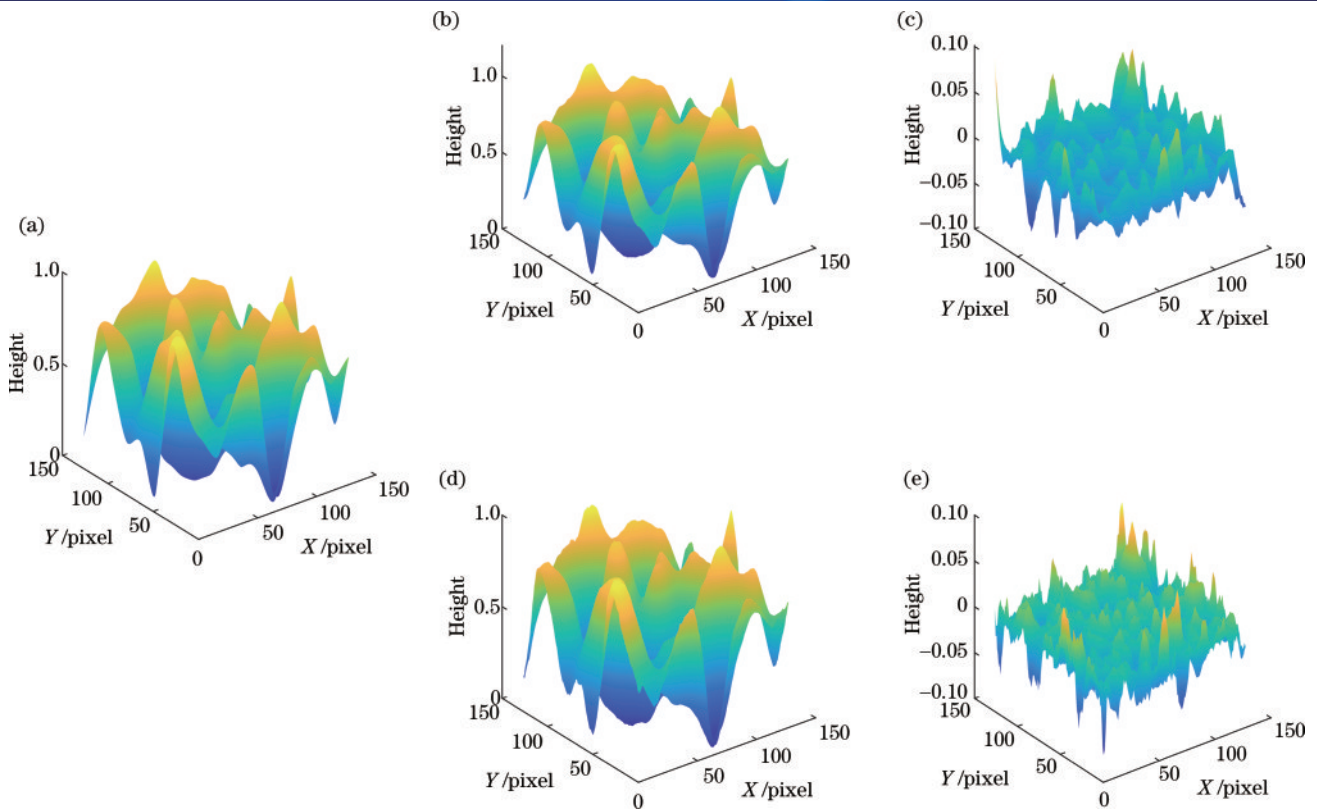


图9 物体的三维展示和误差图。(a)条纹图对应的三维形貌图;(b)U-Net预测的形貌图;(d)MultiResHNet预测的形貌图;  
(c)U-Net测量的误差图;(e)MultiResHNet测量的误差图

Fig. 9 3D effect and error images of an object. (a) 3D shape image corresponding to a fringe pattern; (b) 3D shape image predicted by U-Net; (d) 3D shape image predicted by MultiResHNet; (c) error image measured by U-Net; (e) error image measured by MultiResHNet

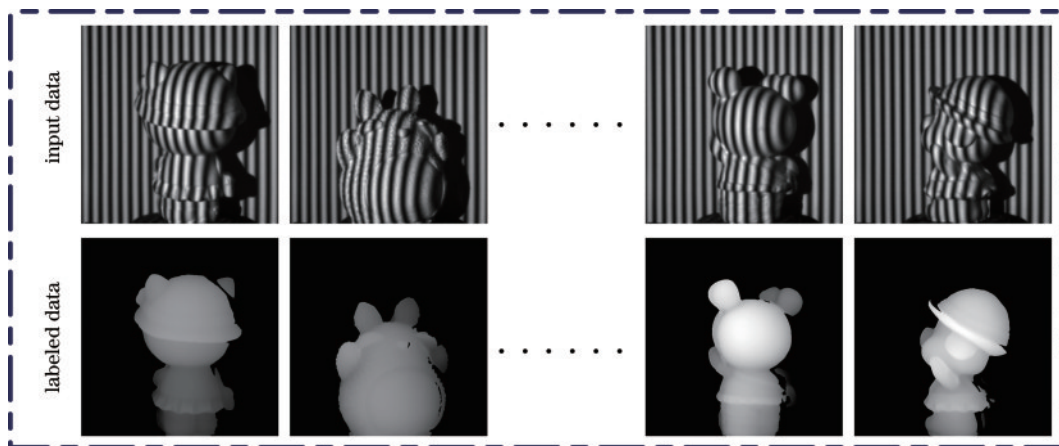


图10 神经网络训练数据集

Fig. 10 Neural network training dataset

选用测试数据中具有代表性的两组数据,对两种网络模型的测试结果与标签数据做误差分析。图12展示了地面真实3D数据和用U-Net与MultiResHNet测试的3D数据重建的视觉比较。网络输入如图12(a)和图12(e)所示,图12(b)和图12(f)是物体的实际三维图,图12(c)、(g)和图12(d)、(h)分别是U-Net和MultiResHNet测试的物体三维图。从图12可

以看出,MultiResHNet预测的深度比U-Net更加准确,并且重建的3D形状整体更加平滑,具有较多的细节。

图13为两个网络对三个物体的测试结果的误差分析。图13(a)是三个物体的网络输入图,图13(b)是真实值和U-Net预测值之间的误差图,图上的值是每个物体对应的平均误差。图13(c)是真实值和

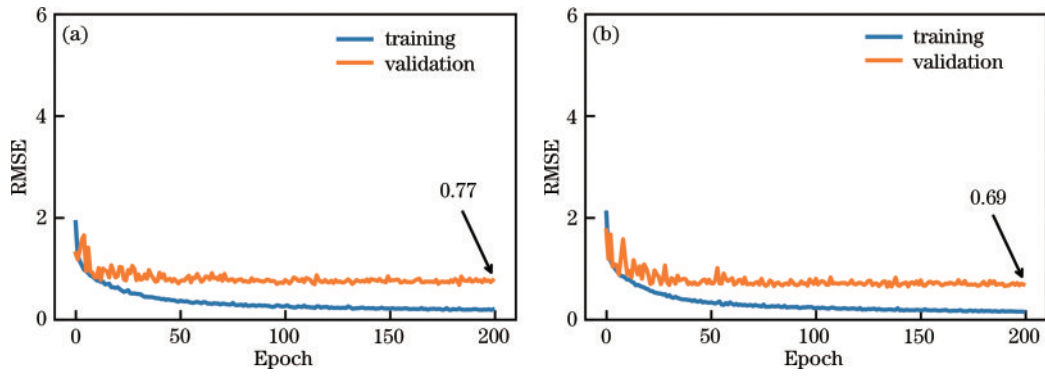


图 11 神经网络的 RMSE。(a)U-Net;(b)MultiResHnet

Fig. 11 RMSE of neural networks. (a) U-Net; (b) MultiResHNet

表 2 两种网络的实验指标分析

Table 2 Experimental index analysis of the two networks

Network	RMSE	MAE	SSIM
U-Net	0.77	$3.985 \times 10^{-1}$	0.98163
MultiResHNet	0.69	$3.528 \times 10^{-1}$	0.98764

MultiResHNet 预测值之间的误差图。平均误差的计

算方法是:将每个有效像素位置的深度误差相加,得到误差之和,然后除以有效像素数。对于第二个被测物体,MultiResHNet 三维重建的误差是 0.0132 mm,U-Net 的重建误差是 0.2000 mm。从结果可以看出,MultiResHNet 三维重建的误差要比 U-Net 低一些。实验结果表明了所提方法的有效性。

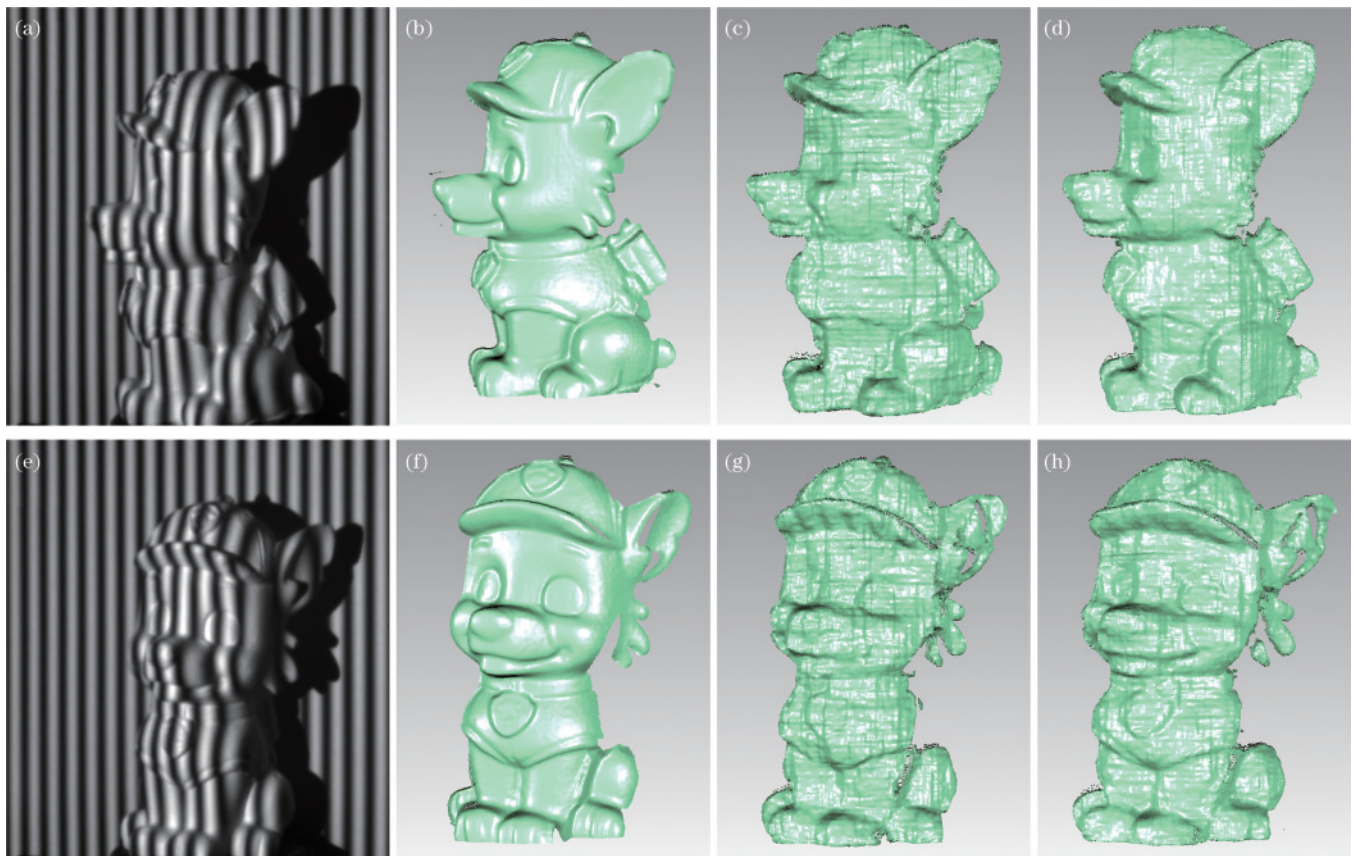


图 12 物体三维重建的视觉展示。(a)(e)网络输入的条纹图;(b)(f)条纹图对应的三维形貌图;(c)(g)U-Net 预测的形貌图;(d)(h)MultiResHNet 预测的形貌图

Fig. 12 Visual representation of three-dimensional reconstruction of objects. (a)(e) Fringe patterns of network input; (b)(f) 3D shape images corresponding to fringe patterns; (c)(g) 3D shape images predicted by U-Net; (d)(h) 3D shape images predicted by MultiResHNet



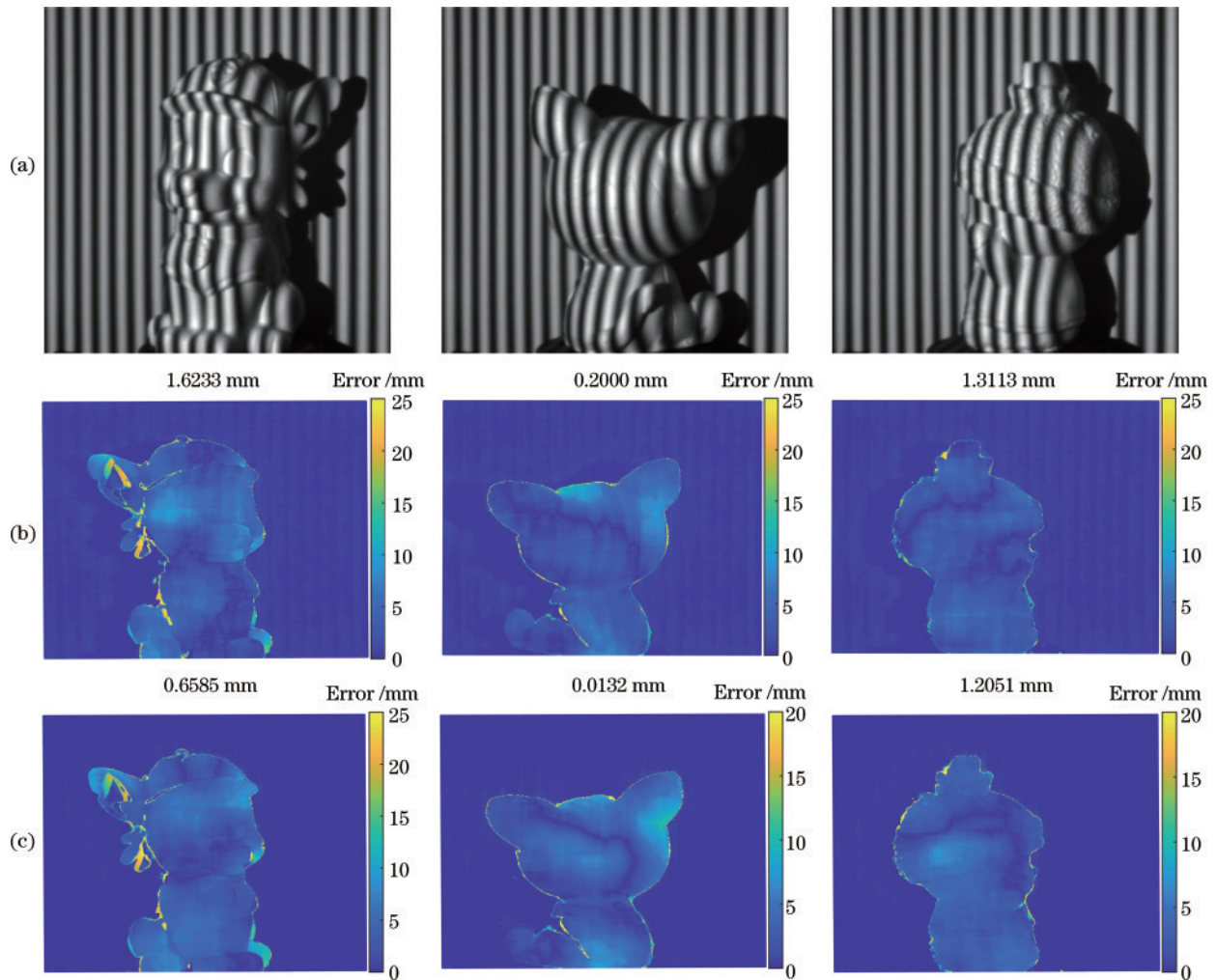


图 13 多个物体的误差结果。(a)网络输入的条纹图;(b)U-Net 预测误差图;(c)MultiResHNet 预测误差图

Fig. 13 Error results for multiple objects. (a) Fringe patterns of network input; (b) error images predicted by U-Net; (c) error images predicted by MultiResHNet

## 4 结 论

提出了一种用于单次条纹图 3D 成像的 MultiResHNet 深度神经网络模型,使用相移法和格雷码方法制作数据集,测量小型物体的三维形状,实验验证了所提网络的可行性和有效性。与 U-Net 训练结果相比,MultiResHNet 效果更加好,测量精度也更高,对表面较平坦的物体的测试精度可以达 0.01 mm,使三维形状测量技术能更好地应用于科学研究和实际生活中。为获得更高的三维重建测量精度,后续可以从多方面进行改进,包括对网络输入图像预处理的优化、网络模型的改进、训练数据集的优化和增强、网络模型泛化能力的提高,使网络可以应用在更加复杂的场景和物体上。

### 参 考 文 献

- [1] Sansoni G, Trebeschi M, Docchio F. State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation[J].
- [2] Vázquez-Arellano M, Griepentrog H W, Reiser D, et al. 3-D imaging systems for agricultural applications-a review [J]. Sensors, 2016, 16(5): 618.
- [3] Curless B. From the range scans to 3D models[J]. ACM Siggraph Computer Graphics, 1999, 33(4): 38-41.
- [4] Song Z. High-speed 3D shape measurement with structured light methods: a review[J]. Optics and Lasers in Engineering, 2018, 106: 119-131.
- [5] Schmidhuber J. Deep learning in neural networks: an overview[J]. Neural Networks, 2015, 61: 85-117.
- [6] 冯世杰, 左超, 尹维, 等. 深度学习技术在条纹投影三维成像中的应用[J]. 红外与激光工程, 2020, 49(3): 0303018.
- [7] Spoorthi G E, Gorthi S, Gorthi R K S S. PhaseNet: a deep convolutional neural network for two-dimensional phase unwrapping[J]. IEEE Signal Processing Letters, 2019, 26(1): 54-58.
- [8] Spoorthi G E, Gorthi R K S, Gorthi S. PhaseNet 2.0:

- phase unwrapping of noisy data based on the deep learning approach[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4862-4872.
- [9] Wang K Q, Li Y, Qian K M, et al. One-step robust deep learning phase unwrapping[J]. *Optics Express*, 2019, 27(10): 15100-15115.
- [10] Nguyen H, Wang Y Z, Wang Z Y. Single-shot 3D shape reconstruction using structured light and deep convolutional neural networks[J]. *Sensors*, 2020, 20(13): 3718.
- [11] Machineni R C, Spoorthi G E, Vengala K S, et al. End-to-end deep learning-based fringe projection framework for 3D profiling of objects[J]. *Computer Vision and Image Understanding*, 2020, 199: 103023.
- [12] van der Jeught S, Dirckx J J. Deep neural networks for single-shot structured light profilometry[J]. *Optics Express*, 2019, 27(12): 17091-17101.
- [13] Ibtehaz N, Rahman M S. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation[J]. *Neural Networks*, 2020, 121: 74-87.
- [14] Nguyen H, Ly K L, Tran T, et al. hNet: single-shot 3D shape reconstruction using structured light and h-shaped global guidance network[J]. *Results in Optics*, 2021, 4: 100104.
- [15] Zhang Z H. A review of single-shot 3D shape measurement by phase calculation-based fringe projection techniques[J]. *Optics and Lasers in Engineering*, 2012, 50(8): 1097-1106.
- [16] Zhang Q C, Su X, Xiang L, et al. 3-D shape measurement based on complementary Gray-code light [J]. *Optics and Lasers in Engineering*, 2012, 50(4): 574-579.
- [17] Yu S, Zhang J, Yu X, et al. 3D measurement using a combined Gray code and dual-frequency phase-shifting approach[J]. *Optics Communications*, 2018, 413: 283-290.
- [18] Xie S N, Tu Z W. Holistically-nested edge detection [C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1395-1403.
- [19] Goodfellow I, Bengio Y, Courville A. Deep learning [M]. Massachusetts: The MIT Press, 2016.
- [20] Kingma D P, Ba J. Adam: a method for stochastic optimisation[EB/OL]. (2014-12-22) [2022-08-06]. <https://arxiv.org/abs/1412.6980>.