

# 光学学报

## 基于卷积神经网络的单帧复合图像绝对相位恢复

李文健<sup>1,2</sup>, 盖绍彦<sup>1,2\*</sup>, 俞健<sup>1,2</sup>, 达飞鹏<sup>1,2,3\*\*</sup>

<sup>1</sup>东南大学自动化学院, 江苏 南京 210096;

<sup>2</sup>东南大学复杂工程系统测量与控制教育部重点实验室, 江苏 南京 210096;

<sup>3</sup>东南大学深圳研究院, 广东 深圳 518063

**摘要** 提出一种使用卷积神经网络从单帧复合图像中获得高质量绝对相位的方法。所提方法使用的复合图像为嵌入散斑的条纹图像, 卷积神经网络包括两个子网络, 分别利用复合图像中的条纹模式分量和散斑模式分量完成包裹相位的求解及展开。所提方法在相位展开的过程中使用预拍摄的平板复合图像及其条纹阶次作为辅助信息, 从而保证相位展开的准确性。实验结果表明, 所提方法可以利用单帧复合图像最大限度地减少投影图片的数量, 能够获得较高精度的绝对相位, 为实现高精度动态场景下的三维测量提供一种可行性的解决方案。

**关键词** 测量; 相位恢复; 条纹投影; 神经网络; 三维测量; 散斑相关

中图分类号 TP391

文献标志码 A

doi: 10.3788/AOS202141.2312001

### Absolute Phase Recovery of Single Frame Composite Image Based on Convolutional Neural Network

Li Wenjian<sup>1,2</sup>, Gai Shaoyan<sup>1,2\*</sup>, Yu Jian<sup>1,2</sup>, Da Feipeng<sup>1,2,3\*\*</sup>

<sup>1</sup>*School of Automation, Southeast University, Nanjing, Jiangsu 210096, China;*

<sup>2</sup>*Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education, Southeast University, Nanjing, Jiangsu 210096, China;*

<sup>3</sup>*Shenzhen Research Institute, Southeast University, Shenzhen, Guangdong 518063, China*

**Abstract** In this paper, a convolutional neural network is proposed to obtain high quality absolute phase from single frame composite images. The composite image used in the proposed method is the fringe image embedded with speckle. The convolutional neural network consists of two sub-networks, which use the fringe mode component and the speckle mode component in the composite image to solve and unfold the wrapping phase. In the process of phase unwrapping, the proposed method uses the pre-photographed composite image and its fringe order as auxiliary information to ensure the accuracy of phase unwrapping. Experimental results show that the proposed method can minimize the number of projected images by using single-frame composite images and obtain high precision absolute phase, which provides a feasible solution for 3D measurement in high precision dynamic scenes.

**Key words** measurement; phase recovery; fringe projection; neural network; three-dimensional measurement; speckle correlation

**OCIS codes** 120.5050; 120.2650; 100.5070

收稿日期: 2021-04-14; 修回日期: 2021-05-14; 录用日期: 2021-06-10

基金项目: 国家自然科学基金(51475092)、江苏省自然科学基金(BK20181269)、江苏省前沿引领技术基础研究专项(BK20192004C)、深圳市科技创新委员会(JCYJ20180306174455080)

通信作者: \*qxxyymm@163.com; \*\*dafp@seu.edu.cn

## 1 引言

三维测量技术在工业元件缺陷检测、医疗疾病诊断、人工智能和机器视觉等众多领域得到了广泛的研究及应用,其中结构光三维测量技术因其具有非接触、精度高、条纹设置灵活以及测量范围较大等优点而成为三维测量技术中的主流光学技术<sup>[1]</sup>。最常见的两种结构光三维测量技术分别为散斑投影测量技术和条纹投影测量技术。

条纹投影测量技术的基本处理过程主要分为包裹相位计算、相位展开和三维重建等步骤<sup>[2]</sup>。常见的包裹相位计算方法主要有相移(PS)法<sup>[3]</sup>和傅里叶变换(FT)法<sup>[4-5]</sup>。傅里叶变换法只需要一幅条纹图像就可以计算包裹相位,但是存在频谱混叠和频谱泄漏等问题,这会影响其计算精度,尤其是在被测物体表面形貌不连续的情况下精度较低。相移法具有解相位精度较高和抗干扰能力强的优点,因此得到了更广泛的应用。由上述方法求解得到的包裹相位还需要通过相位展开法来获得正确的绝对相位。常用的相位展开法主要有时间相位展开(TPU)法<sup>[6]</sup>和空间相位展开(SPU)法<sup>[7]</sup>。时间相位展开法需要将多张辅助图形(格雷码<sup>[8]</sup>或多波长条纹<sup>[9]</sup>等)投影到被测物体表面,而且每个像素的相位展开互不影响,因此当采用该方法来测量具有复杂表面形貌的被测物体时表现良好,抗噪声能力强。空间相位展开法仅需要对一幅图像根据相邻相位差进行逐行逐列的积分展开,这要求被测物体表面满足连续性的假设,但实际获得的包裹相位图由于欠采样而存在边缘跳变和局部阴影等问题,造成该方法的测量误差较大。

散斑投影测量技术可以通过单幅散斑图片来实现三维测量,但传统散斑投影测量技术存在测量速度慢和测量精度低等问题,而且双摄像机的引入会增加制作成本和系统的复杂度。因此,研究人员考虑能否综合利用条纹投影技术和散斑投影技术的优势进行结构光的投影测量。如笪健等<sup>[10]</sup>基于标准的三步相移法向正弦条纹图像中嵌入额外的随机数字散斑,并结合左右相机之间的极限约束实现对三幅图像的相位展开。Deng 等<sup>[11]</sup>同样将散斑嵌入条纹图像中,并专门创建了一个用于区分不同周期包裹相位的散斑匹配函数,采用六步相移结合循环投影的方法实现了三维测量。这些将散斑与条纹融合产生复合图像的方法,因为需要投影多帧图像或者需要双摄像机来实现散斑匹配,所以它们的应用场

景受到了一定的限制。

近些年,深度学习及神经网络取得了飞速发展,而且在人工智能、计算机视觉和自然语言处理等领域取得了令人瞩目的突破<sup>[12-14]</sup>。与此同时,越来越多的研究人员考虑能否利用深度学习和卷积神经网络(CNN)来实现条纹投影和散斑投影。对于散斑投影,钟锦鑫等<sup>[15]</sup>利用双目系统采集数据集,通过深度学习的方法来实现散斑图像的逐像素匹配,这提高了散斑图像匹配算法的精度。对于条纹投影,有一些研究人员将深度学习应用于计算包裹相位的过程,Feng 等<sup>[16]</sup>利用深度神经网络进行条纹分析,这大大提高了单条纹图的相位解调精度。Qiao 等<sup>[17]</sup>针对镜面物体提出了一种单镜头相位检索方法,其可以通过深度学习来重建镜面物体的相位分布。进一步,如何利用深度学习来实现相位展开成为了研究热点。Wang 等<sup>[18]</sup>利用了 CNN 专门学习相位展开,证明了复杂的非线性相位展开任务可以由单个深度神经网络直接完成。Budianto 等<sup>[19]</sup>将条纹阶次信息编码后嵌入条纹图中,利用 CNN 对条纹图中的条纹阶次信息进行解码以实现相位展开。紧接着便有研究人员直接利用深度学习从条纹信息到高度信息之间的映射,van der Jeught 等<sup>[20]</sup>提出了一种基于深度学习的从单个变形条纹图案中提取高度信息的方法,证明深度学习可以从条纹图中获得高精度的高度信息,但这种端到端的单阶段网络无法利用一些重要的中间信息。

如何在保证测量精度的同时提高测量速度成为结构光测量的重要研究方向,为此本文提出一种基于 CNN 从单帧融合散斑的条纹图中获得高精度绝对相位的方法。在该方法中,设计两个子网络并分别实现包裹相位的计算和展开。在包裹相位的计算过程中复合图像的条纹分量起主要作用,而在相位展开的过程中散斑分量较为重要。由于散斑图像的随机性,没有较好的方法能从一张复合图像中将条纹分量和散斑分量区分开来并分别利用。由于 CNN 具备提取不同维度特征的能力,所以设计两个子网络来提取复合图像中的不同信息。利用子网络一来提取复合图像中的相位信息以获得包裹相位,再利用复合图像中散斑信息的高分辨率特点,通过子网络二来区分包裹相位所属的条纹阶次以实现相位展开。同时在相位展开的过程中,利用辅助信息来提高条纹阶次精度。本文方法的优势在于投影图像较少,仅需单帧复合图像即可求解得到绝对相位。实验结果证明,所提出的基于 CNN 的方法可以通

过单幅复合图像来获得较高精度的绝对相位,而且可应用于动态测量场景。

## 2 传统方法获取绝对相位

本节介绍如何通过相移法结合格雷码法来获得绝对相位,其原理如图 1 所示,将绝对相位作为训练 CNN 的真值。以  $N$  步相移法为例,投影仪向被测物体表面投射  $N$  幅正弦周期相移图案,则相机采集到的条纹图案可表示为

$$I_i(x, y) = A(x, y) +$$

$$B(x, y) \cos[\varphi(x, y) + 2i\pi/N], \quad (1)$$

式中:  $(x, y)$  表示像素点坐标;  $I_i$  表示采集到的第  $i$  幅图像,  $i = 1, 2, \dots, N$ ;  $A(x, y)$  表示背景光强;  $B(x, y)$  表示调制强度;  $\varphi(x, y)$  表示包裹相位,可表示为

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

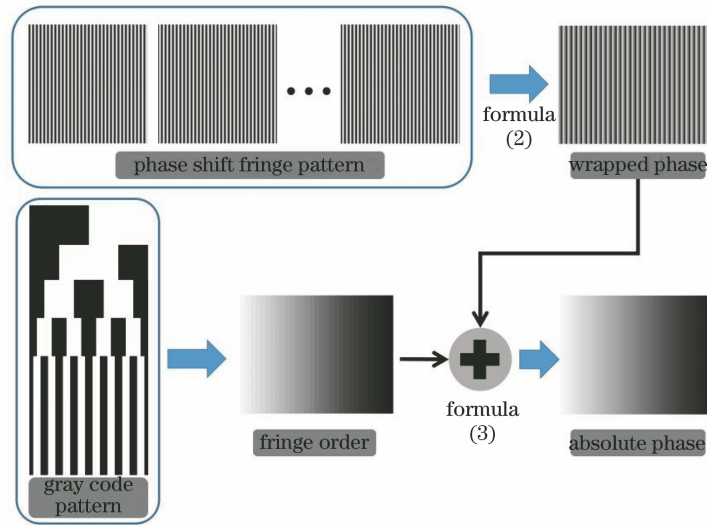


图 1 利用相移法结合格雷码法获得绝对相位的原理图

Fig. 1 Principle diagram of absolute phase is obtained by using phase shift method combined with gray code method

为了确定每个像素点的相位,相移法使用反正切函数来计算,将反正切函数的分子和分母分别记为  $M$  和  $D$ ,此时得到的相位通常被称为包裹相位,取值范围为  $-\pi \sim \pi$ ,相邻周期之间存在  $2\pi$  的相位跳变,对包裹相位进行相位展开可以获得绝对相位。绝对相位  $\phi(x, y)$  与包裹相位之间的关系可以描述为

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

式中:  $k(x, y)$  表示像素点所在的条纹阶次。使用格雷码法来标记条纹阶次  $k(x, y)$ ,再根据(3)式逐像素地展开相位。

## 3 所提方法

### 3.1 复合图像的设计

由于散斑具有高度的空间辨别度,保证了不同区域的散斑图像具有唯一性的特点,所以通过散斑图像可以辅助确定条纹阶次<sup>[21]</sup>。因此在本文方法中,将散斑图像嵌入条纹图像中可以得到复合条纹图像,经过 CNN 的训练可以实现从单帧复合条纹

图像中获得绝对相位。复合图像可表示为

$$I_c(x, y) = I(x, y) + I_s(x, y), \quad (4)$$

式中:  $I_c(x, y)$  表示复合图像;  $I(x, y)$  表示根据(1)式设计的正弦条纹图像,此时  $i = 1$ ;  $I_s(x, y)$  为随机散斑图像,可表示为

$$I_s(x, y) = \sum_{j=1}^S I_0 \exp \left[ -\frac{(x-x_j)^2 + (y-y_j)^2}{d^2} \right], \quad (5)$$

式中:  $S$  表示随机散斑的总个数;  $d$  表示散斑的最大直径;  $I_0$  表示散斑的光强,即散斑图像的灰度;  $x_j$  表示散斑中心位置的横坐标;  $y_j$  表示散斑中心位置的纵坐标。复合图像的分辨率为  $1280 \text{ pixel} \times 960 \text{ pixel}$ ,生成原理如图 2(a)所示。条纹图像  $I(x, y)$  中的  $A$  和  $B$  均设置为 100,即幅值为 200;条纹周期数设置为 32,实验过程中使用竖向条纹,每个条纹周期中的像素个数为 30。散斑图像  $I_s(x, y)$  中的  $S$  设置为 10000,  $d$  设置为 7,  $I_0$  设置为 255;  $x_j$  随机取  $1 \sim 960$  之间的整数,  $y_j$  随机取  $1 \sim 1280$  之间的整数。得到随机散斑图像后对其进行归一化处

理,再将每个像素灰度值乘以 55 可以得到幅值为 55 的散斑图像。最后在每个像素位置叠加条纹图

像与散斑图像的灰度值,得到的复合图像如图 2(b) 所示。

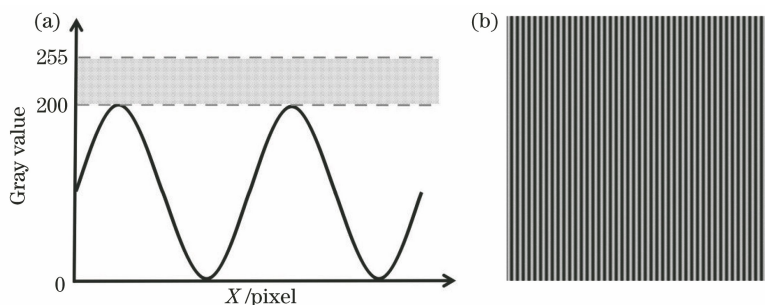


图 2 复合条纹图像的设计。(a)复合条纹图像的生成原理;(b)复合条纹图像

Fig. 2 Design of composite fringe pattern. (a) Generation principle of composite fringe pattern; (b) composite fringe pattern

### 3.2 所提方法的整体框架

本文提出了一种利用 CNN 从单帧复合图像中

获得绝对相位的方法,其原理如图 3 所示,其中  $I_w$  和  $k_w$  分别为参考平面复合图像及其条纹阶次。

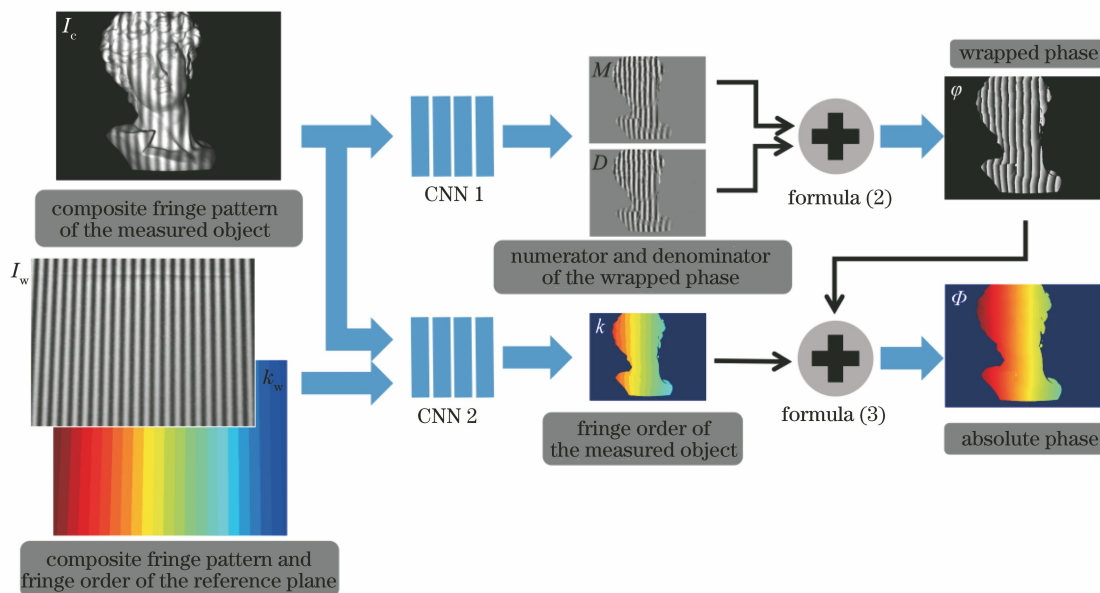


图 3 本文方法的原理图

Fig. 3 Principle of proposed method

构造两个子网络(CNN 1 和 CNN 2),分别用于训练得到包裹相位和条纹阶次。为了直接从复合图像中获得绝对相位,使用这种多阶段网络<sup>[22]</sup>而不使用一个单阶段网络的主要原因如下。1)包裹相位和条纹阶次是两种不同的特征,包裹相位更关心某一像素点的值,条纹阶次更关心某一周期内所有像素点的值,它们并不在同一尺度上,所以不使用一个网络来提取特征;2)单阶段网络虽然结构简单,但为了得到较好的预测结果不可避免地使网络参数的设计变得更加复杂,灵活性不够,而且相比于多阶段网络,单阶段网络无法利用一些重要的中间变量信息。所以,本文采用多阶段网络结构,训练两个子网络

CNN 1 和 CNN 2 可以分别获得包裹相位及条纹阶次,根据任务的不同可为子网络设计各自具有针对性的结构特点,灵活性更强。

如图 3 所示,CNN 1 的输入为由相机拍摄得到的单帧复合图像,输出为(2)式中的分子项和分母项,即  $M$  和  $D$ 。采用这种策略的原因是包裹相位存在  $2\pi$  的跳变,所以 CNN 难以精确预测跳变位置,而  $M$  和  $D$  不存在跳变,则利用  $M$  和  $D$  可获得高质量的包裹相位<sup>[16]</sup>,训练过程中真值由 8 步相移法得到。

对于 CNN 2,除了被测物体的单帧复合图像以外,一个参考平面的复合图像及其条纹阶次图也会

同时输入到 CNN 2 中, CNN 2 的输出为被测物体的条纹阶次图, 训练过程中真值由格雷码法得到。本文在测量距离(相机焦距)下拍摄表面平整的白色平板, 可以得到参考平面的复合图像, 再通过格雷码法得到参考平面的条纹阶次图。CNN 2 利用参考平面的复合条纹与被测物体的复合条纹之间的形变信息来调整参考平面的条纹阶次, 从而获得被测物体的条纹阶次。上述设计的目的是利用深度神经网络的优势来实现单帧复合图像的相位展开。值得一提的是, 虽然 CNN 2 有三个输入, 但这里辅助完成相位展开的参考平面信息是拍摄被测物体复合图像之前预先拍摄的, 且拍摄一次即可, 后续不需要重复拍摄, 所以在实际测量的过程中仍是利用单帧复合图像来预测得到绝对相位。最后利用 CNN 1 和 CNN 2 的输出可以得到高质量的绝对相位。

### 3.3 CNN 结构

本节将详细介绍 CNN 1 和 CNN 2 的网络结构。CNN 1 的网络结构如图 4 所示, 使用大小为 (640, 480, 1) 的张量作为网络的输入, 其中 640 pixel×480 pixel 为输入图像的分辨率, 1 为输

入数据的通道数, 同时将每层张量大小标注于图 4。网络主要由卷积层、剩余块网络、最大池化层、上采样层和融合层组成。对于每个卷积层, 卷积核的大小为  $3 \times 3$ , 卷积步长为 1, 每个卷积层所用的卷积核数目已在图 4 中标注出来, 在边界使用零填充以控制输出的空间大小与输入保持一致。除了最后一个卷积层使用线性激活函数以外, 其余的卷积层使用 ReLU<sup>[23]</sup> 作为激活函数。每个剩余网络块<sup>[24]</sup> 的结构如图 5 所示, 剩余网络块的使用可以解决网络退化的问题, 其中  $H$  和  $W$  分别为图像的高和宽。在整个网络结构中, 随着网络深度的增加, 利用最大池化层对图像进行下采样处理以提取更明显和精确的特征, 利用上采样层将图像大小恢复到与输入图像一致。对于 CNN 1, 输入复合图像中的条纹分量对解包裹相位精度的影响较大, 所以在主路径上提取条纹分量中的细节信息, 再利用最大池化层在支路上对特征信息进行下采样处理以减少散斑分量对最终结果的影响; 然后通过融合层来叠加主路径与支路所提取的特征信息; 最后通过一个线性激活的卷积层来得到预测输出。

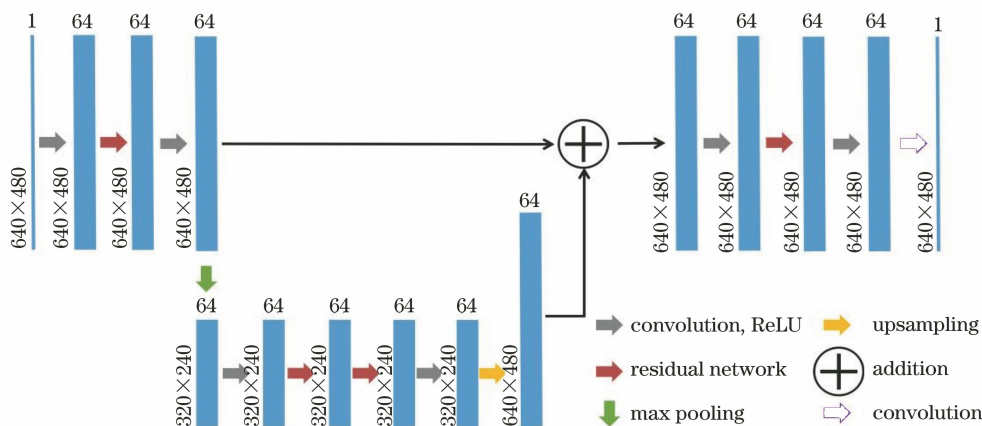


图 4 CNN 1 的结构

Fig. 4 Structure of CNN 1

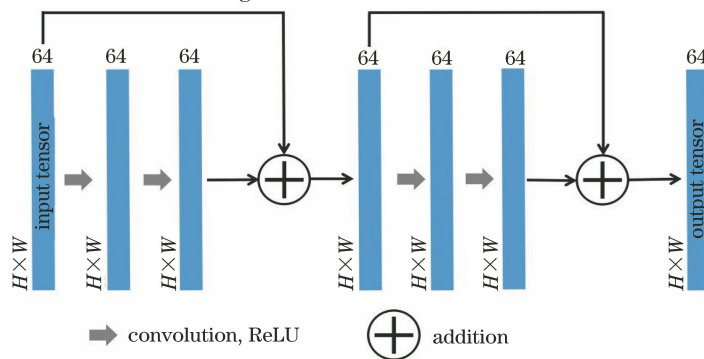


图 5 剩余块网络的结构

Fig. 5 Structure of residual block network

CNN 2 的网络结构如图 6 所示,以大小为  $(640, 480, 3)$  的张量作为输入,输入图像的分辨率仍为  $640 \text{ pixel} \times 480 \text{ pixel}$ 。与 CNN 1 不同的是,在 CNN 2 中输入数据的通道数为 3,因为除了被测物体复合图像外,CNN 2 额外加入一组通道数为 3 的输入数据,其由白色平板的复合图像及其条纹阶次图组成。CNN 2 同样主要由卷积层、剩余块网络、最大池化层、上采样层和融合层组成,其中卷积层参数和剩余块网络结构与 CNN 1 一样。不同于 CNN 1,CNN 2 提取细节信息是计算每个像素的包裹相位。对于相位展开而言,同一周期的连续多个像素的条纹阶次为同一个值,所

以 CNN 2 除了条纹阶次跳变处的细节信息,还需要根据散斑分量来提取大尺度的特征信息以表征条纹阶次。在 CNN 2 中设计了三条并行路径,在第一条路径中对输入图像不进行下采样处理以保证细节信息的有效获取;在第二条和第三条路径中对输入图像通过最大池化层分别进行下采样 2 倍和 4 倍的处理,可以得到分辨率分别为  $320 \text{ pixel} \times 240 \text{ pixel}$  和  $160 \text{ pixel} \times 120 \text{ pixel}$  的输入图像,再通过卷积层等来提取大尺度特征,接着通过融合层、卷积层及剩余块网络将不同尺度上的特征充分融合,最后通过一个线性激活的卷积层来得到与输入对应的条纹阶次信息。

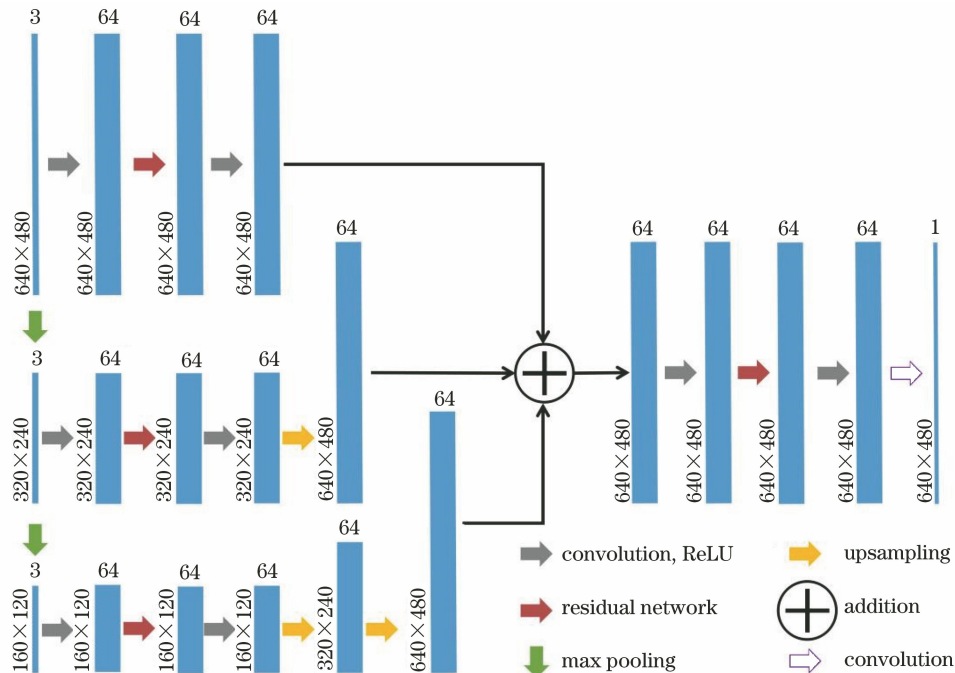


图 6 CNN 2 的结构

Fig. 6 Structure of CNN 2

## 4 实验与讨论

### 4.1 CNN 的训练阶段

为了验证所提方法的有效性,建立一个结构光投影三维测量系统,其包括一个摄像机和一个投影仪。在实验中使用具有 32 周期的相移条纹图案,被测物体与摄像机-投影仪对之间的距离约为 1.5 m (测量距离),摄像机的视场尺寸约为  $400 \text{ mm} \times 300 \text{ mm}$ 。为了收集训练数据集,对不同类型的石膏体进行任意组合和不同角度旋转,生成 350 个不同的场景图像,分辨率为  $640 \text{ pixel} \times 480 \text{ pixel}$ ,其中 300 组作为训练集,剩余 50 组作为测试集。每组数据图像包含 8 幅相移条纹图像、5 幅格雷码图像和 1

幅复合图像,共 14 幅图像。对于每组数据图像,通过 8 步相移法和格雷码法来获得包裹相位的分子项  $M$  和分母项  $D$  以及条纹阶次并作为真值。用于训练 CNN 的输入及真值如图 7 所示。同时由于 CNN 2 输入的需求,额外采集一组白色平板(图 3)作为参考平面,参考平面的摆放距离与被测物体的测量距离一致,摆放过程中使参考平面尽可能正对相机(摄像机光轴),保证在测量距离(约为 1.5 m)下占满相机视场即可。

为了使网络的学习过程更加容易以及收敛速度更快,在 CNN 1 中将复合图像进行归一化处理后再输入网络中,处理方式各像素点灰度值除以所有像素点最大灰度值与最小灰度值之差。在 CNN 2

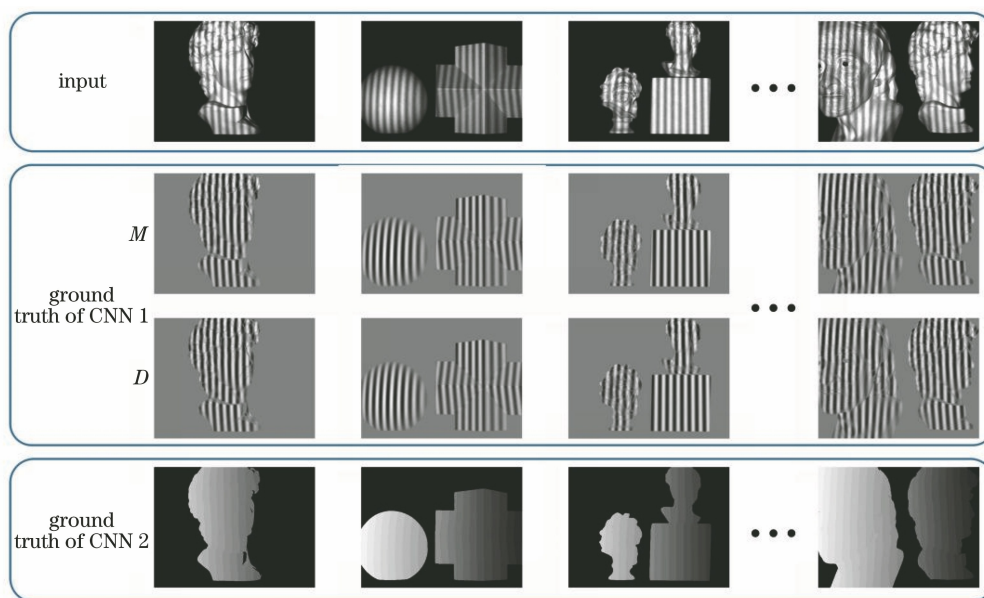


图 7 CNN 的部分训练数据

Fig. 7 Partial training data of CNN

中对复合图像先进行归一化处理再乘以条纹周期数 32, 而条纹阶次图不进行归一化处理。两个子网络都使用自适应矩估计 Adam<sup>[25]</sup> 作为优化器, 学习率初始化为 0.01, 每经过 100 轮的迭代学习率除以 10, 这样可以避免模型陷入局部最小值。由于两个子网络的目标任务不同, 所以使用不同的损失函数对网络参数进行调整。CNN 1 使用网络输出相对于真值的均方误差 (MSE) 作为损失函数, 可表示为

$$L_1 = \frac{1}{m} \sum_{o=1}^m (t_o - \hat{t}_o)^2, \quad (6)$$

式中:  $m$  表示像素个数之和;  $t_i$  表示真值;  $\hat{t}_i$  表示预测输出。同理, CNN 2 的损失函数如 (7) 式, 由于大误差对条纹阶次的影响较大, 所以在 (7) 式中给出了一个惩罚因子  $p_o$ 。当预测输出与真值之间的误差

( $t_o - \hat{t}_o$ ) 大于或等于 0.5 时, 惩罚因子  $p_o$  设置为 10, 这可以较好地抑制预测输出与真值之间的差异, 该过程可表示为

$$L_2 = \sqrt{\frac{\sum_{o=1}^m p_o \times (y_o - \hat{y}_o)^2}{m}}, \quad (7)$$

其中  $p_o$  满足

$$p_o = \begin{cases} 1, & \text{if } y_o - \hat{y}_o < 0.5 \\ 10, & \text{if } y_o - \hat{y}_o \geq 0.5 \end{cases}. \quad (8)$$

图 8 为 CNN 训练后的损失函数曲线, 横轴表示网络迭代的轮数, 纵轴表示训练集和验证集的损失值, 其中子网络 CNN 1 设置了 300 轮迭代, CNN 2 设置了 500 轮迭代。从图 8(a) 和图 8(b) 可以看到, CNN 1 的损失函数曲线在迭代大约 200

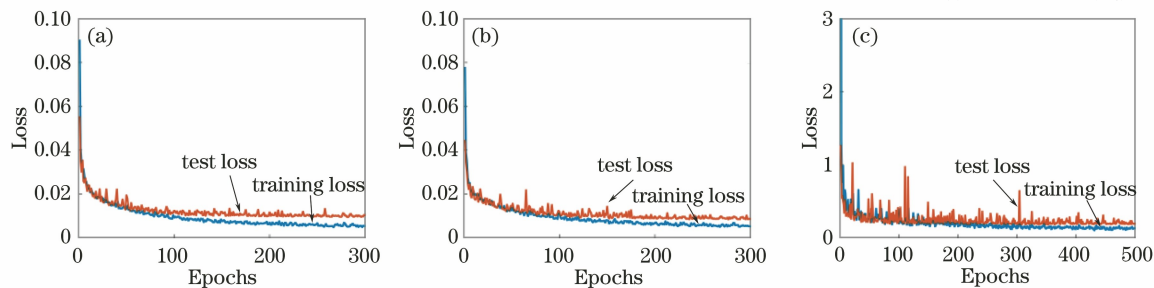


图 8 不同情况下 CNN 1 和 CNN 2 的损失函数曲线。(a) CNN 1 中分子项  $M$  的损失曲线; (b) CNN 1 中分母项  $D$  的损失曲线; (c) CNN 2 中阶次  $k$  的损失曲线

Fig. 8 Loss function curves of CNN 1 and CNN 2 under different conditions. (a) Loss curves of molecular term  $M$  in CNN 1; (b) loss curves of denominator  $D$  in CNN 1; (c) loss curves of order  $k$  in CNN 2

轮后收敛。从图 8(c) 可以看到, CNN 2 的损失函数曲线在迭代大约 350 轮后收敛。由于这两个网络的损失函数不同且归一化方式不同, 分子  $M$  和分母  $D$  均是进行归一化处理得到的, 值为  $0 \sim 1$ , 而条纹阶次未进行归一化处理, 值为  $0 \sim 31$ , 所以它们的输出不在同一尺度上。由于 CNN 2 应用了惩罚因子, 因此损失曲线会出现一些跳跃现象, 然而这些跳跃现象对最终结果的影响较小, 因

为网络本身会随着训练轮数的迭代最终趋于稳定。

### 4.2 实验结果与分析

4.1 节的损失函数曲线表明所提出的网络是有效的, 经过训练确定网络参数后, 对图 9(a) 的多个场景进行测量, 利用 CNN 1 的输出计算得到的包裹相位如图 9(b) 所示, 通过 CNN 2 得到的条纹阶次如图 9(c) 所示。

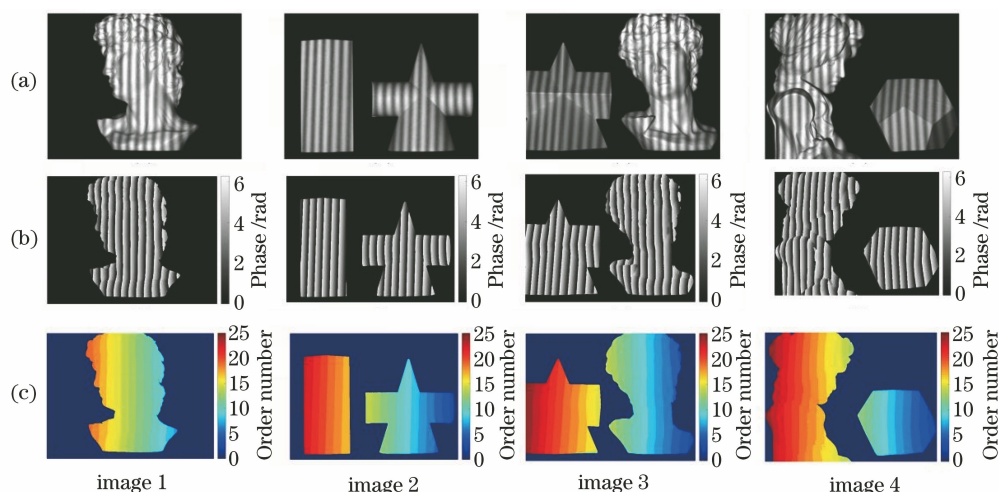


图 9 CNN 的输出结果。(a) 被测对象; (b) CNN 1 所得的包裹相位; (c) CNN 2 所得的条纹阶次

Fig. 9 Output results of CNN. (a) Measured object; (b) wrapping phase obtained by CNN 1; (c) fringe order obtained by CNN 2

将本文方法得到的结果与典型的加窗傅里叶变换(WFT)方法进行对比。由于没有额外的附加图像, 只能利用 WFT 方法来获得包裹相位, 因此针对本文方法与 WFT 方法之间的包裹相位进行比较, 结果如图 10 所示。以 8 步相移法得到的包裹相位作为真值, 图 10(a) 为本文方法所得的误差图, 图 10(b) 为 WFT 方法所得的误差图。同时, 计算每幅误差图的平均误差图, 计算方法为将每个有效

像素点的误差相加得到误差总和, 再除以有效像素点的个数得到平均误差, 单位为弧度(rad)。从图 10 可以看到, WFT 方法的误差明显大于本文方法, 尤其是在复杂纹理区域及边缘位置。对不同被测场景的误差范围进行进一步分析, 结果表明本文方法的平均误差约为  $0.07 \text{ rad}$ , 而 WFT 方法的误差范围为  $0.20 \sim 0.30 \text{ rad}$ , 证明无论被测物体表面的纹理复杂程度如何, 本文方法具有较高的精度和

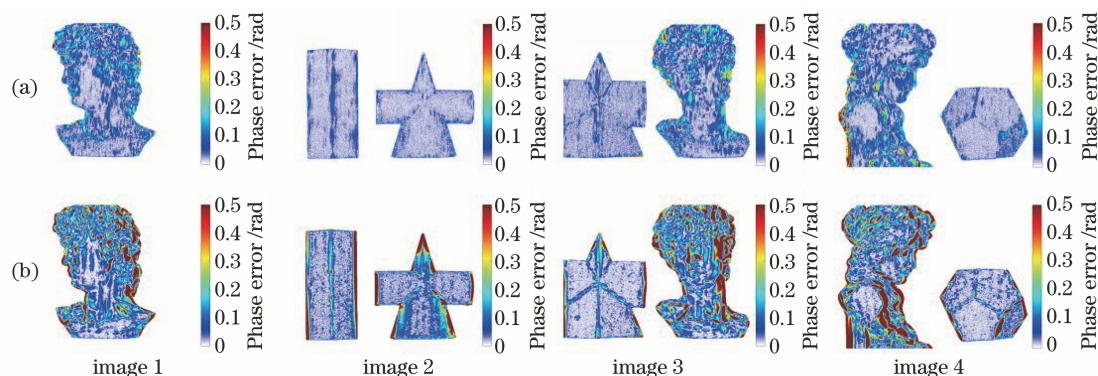


图 10 不同方法的包裹相位误差。(a) 本文方法的包裹相位误差; (b) WFT 方法的包裹相位误差

Fig. 10 Wrapped phase error diagram of different methods. (a) Wrapped phase error of proposed method; (b) wrapped phase error of WFT method



较强的鲁棒性。

为了验证本文方法在恢复绝对相位方面的表现,相移法结合格雷码法所得的绝对相位如图 11(a)所示,本文方法所得的绝对相位如图 11(b)所示。为了使结果的可视化程度更高,绝对相位图是由不同场景下的绝对相位减去参考平面绝对相位得到的。同时以真值为标准,本文方法所得的绝对相位误差如图 11(c)所示。从图 11 可以看到,本文方法具有较高的精度,大部分像素点的误差均小于 0.1 rad;同时也存在一些较大的位置误差,它们主要集中在边缘及遮挡位置,主要原因是这些位置跳

变剧烈或明亮度较低,可用信息较少。为了更清晰地显示误差结果,对图 11(c)中每幅误差图的所有有效像素点个数进行求和,不同误差范围的像素点个数及所占比例如图 12(a)~12(d)所示,其中标签格式为误差范围  $R$ ,像素点个数  $S$  及其占有效像素点总数的比例  $T$ 。从图 12 可以看到,本文方法具有较高的精度,对于不同被测场景基本都有 95% 的像素点误差小于 0.2 rad。从多组实验的结果可以看到,本文方法求解的绝对相位与传统相移法结合格雷码法存在一定的差距,误差主要是由被测物体的边界部分跳变和某些区域的遮挡问题导致的。

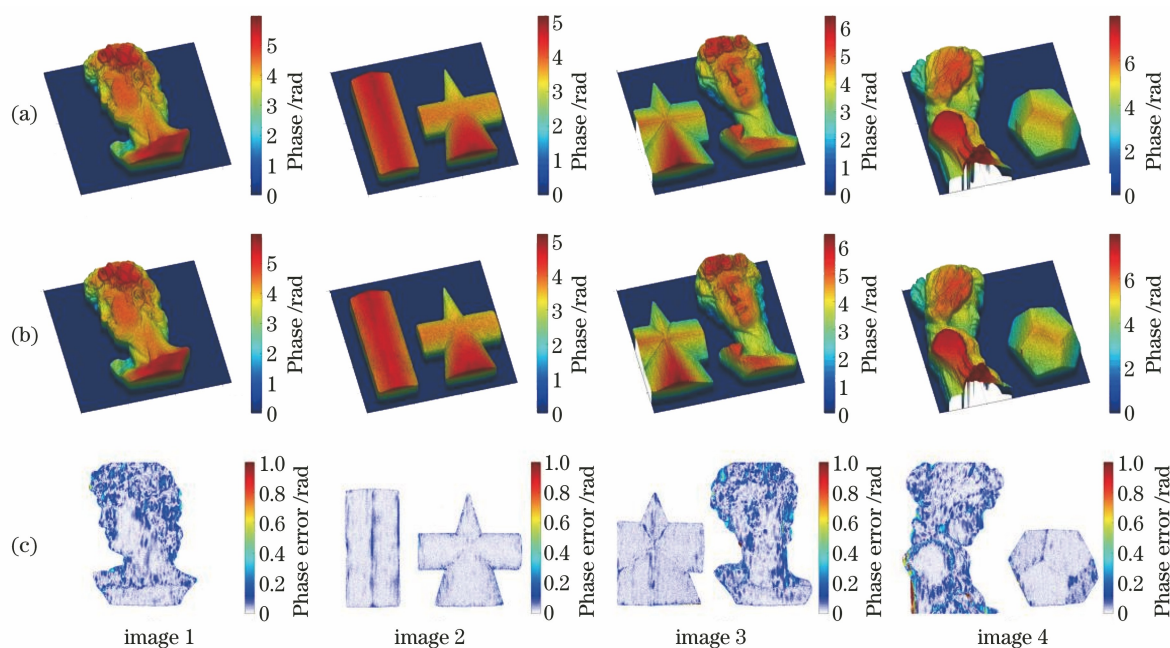


图 11 不同方法下的绝对相位误差。(a)相移法结合格雷码法所得的绝对相位;(b)本文方法所得的绝对相位;(c)本文方法相对于传统方法得到的绝对相位误差

Fig. 11 Absolute phase error under different methods. (a) Absolute phase obtained by phase shift method combined with gray code method; (b) absolute phase obtained by proposed method; (c) absolute phase error obtained by proposed method compared with traditional method

进一步通过实验来验证本文方法能否对快速场景进行绝对相位的分析,对图 13 的张开手掌过程及图 14 的撕纸过程进行绝对相位的恢复。整个绝对相位的恢复过程中主要耗时为 CNN 求解结果的时间,经过多次实验来统计单帧复合图像可以获得包裹相位及条纹阶次的平均时间,其中 CNN 1 获得包裹相位的时间约为 54 ms, CNN 2 获得条纹阶次的时间约为 60 ms,两者可同时进行,所以 CNN 获得绝对相位的时间约为 60 ms,整体求解速度可达 16 frame/s。后续工作可考虑采用 GPU 加速和并

行计算等技术来继续提高整个过程的求解速度。从所得结果可以看到,本文方法可以有效恢复动态场景中的绝对相位。由于本文方法仅需被测物体的单帧复合图像,所以绝对相位的恢复速度主要受到相机采集速度及计算机处理速度的制约,故而可以处理大部分快速场景下的动态测量任务,为动态三维测量提供了重要参考。总的来说,本文所提的单帧复合图像绝对相位恢复方法可以得到高质量的绝对相位且适用于动态场景,但边缘及遮挡区域的测量精度仍有上升空间,这也是今后研究中的重要方向。

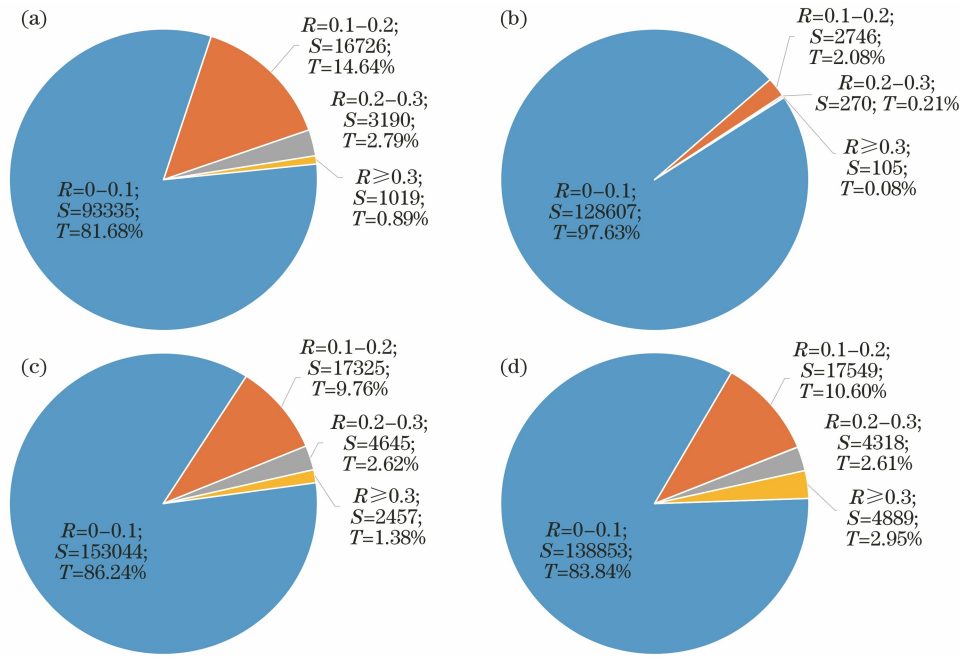


图 12 不同情况下的绝对相位误差统计图。(a)图 11(c) image 1;(b)图 11(c) image 2;(c)图 11(c) image 3;(d)图 11(c) image 4  
 Fig. 12 Statistical diagram of absolute phase error under different conditions. (a) Fig. 11(c) image 1;  
 (b) Fig. 11(c) image 2; (c) Fig. 11(c) image 3; (d) Fig. 11(c) image 4

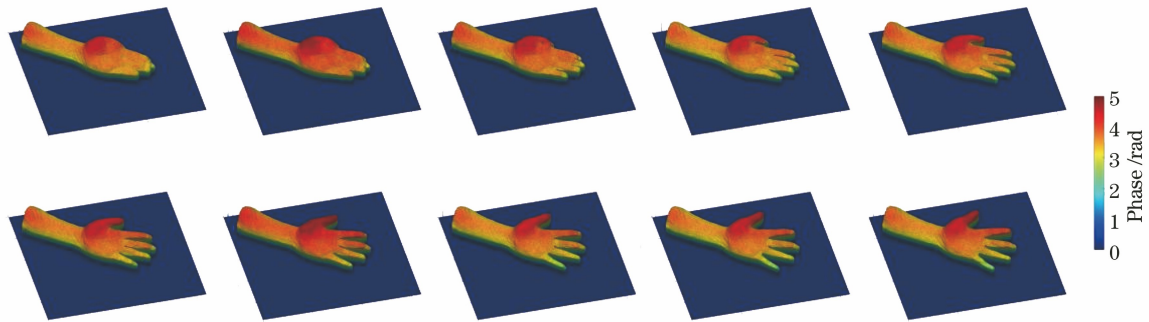


图 13 张开手掌过程的相位图  
 Fig. 13 Phase diagram of open hand process

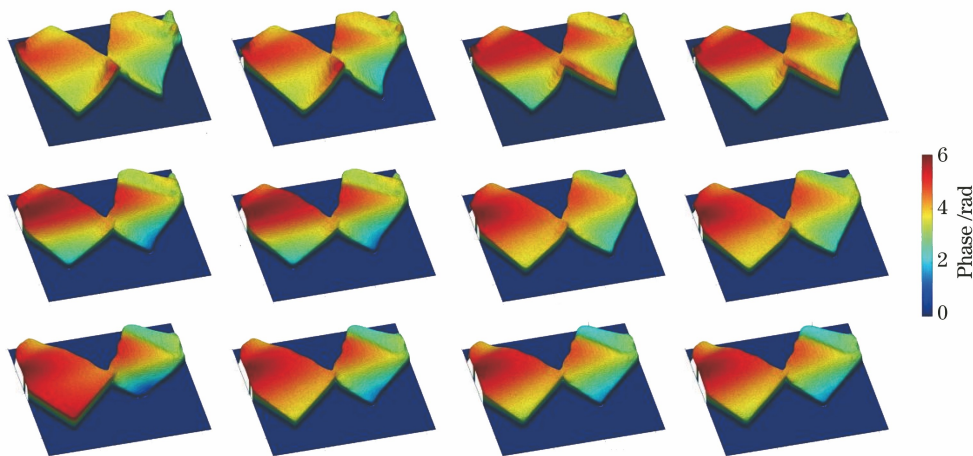


图 14 撕纸过程的相位图  
 Fig. 14 Phase diagram of paper tearing process

## 5 结 论

本文提出一种从单帧复合图像中快速获得高精度绝对相位的方法。在所提方法中,复合图像为嵌入散斑的条纹图像,以该复合图像为输入,利用 CNN 来获得包裹相位和条纹阶次。本文所涉及的 CNN 结构包含两个子网络(CNN 1 和 CNN 2),CNN 1 利用复合图像中的条纹模式来获得包裹相位,CNN 2 利用复合图像中散斑空间辨别度高的特点并结合辅助信息来获得被测物体的条纹阶次。辅助信息为预拍摄的平板复合图像及其条纹阶次,仅需要拍摄一次便可用于网络模型的训练,网络参数确定后在实际测量过程中无需重复拍摄,所以该辅助信息对测量速度没有影响。CNN 的特征提取能力可以保证其从单帧复合图像中提取出条纹分量和散斑分量,且能够实行不同尺度上的特征提取及融合,而这个过程传统方法实现起来较为困难。实验结果也证明了本文方法可以通过单帧复合图像来获得较高精度的绝对相位,由于只需要测量对象的单帧图像,所以该方法在实际测量精度要求较低的高速场景中具有重要的参考意义。但本文方法也存在一些问题,即在边缘及遮挡区域的精度还有待提高,同时该方法受制于真值所用方法精度的影响。未来,随着深度学习的进一步发展及多帧条纹投影测量方法的优化,本文方法可以进一步提升测量精度,最终实现高速、高精度的三维测量。针对后续的研究工作,考虑采用 GPU 加速和并行计算等技术来继续提高绝对相位的求解速度。此外,将在数据集中加入不同的表面发射率和颜色对象以增强 CNN 的预测能力,同时将每次所采集的数据集合并以逐步扩大数据集规模,并增加被测场景的多样性以提高网络模型的泛化能力,使其在更多的被测场景下得到较好的结果。

### 参 考 文 献

- [1] Salvi J, Fernandez S, Pribanic T, et al. A state of the art in structured light patterns for surface profilometry[J]. *Pattern Recognition*, 2010, 43(8): 2666-2680.
- [2] Song Z, Chung R. Determining both surface position and orientation in structured-light-based sensing[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(10): 1770-1780.
- [3] Zuo C, Feng S J, Huang L, et al. Phase shifting algorithms for fringe projection profilometry: a review[J]. *Optics and Lasers in Engineering*, 2018, 109: 23-59.
- [4] Su X Y, Zhang Q C. Dynamic 3-D shape measurement method: a review [J]. *Optics and Lasers in Engineering*, 2010, 48(2): 191-204.
- [5] Qian K M. Two-dimensional windowed Fourier transform for fringe pattern analysis: principles, applications and implementations [J]. *Optics and Lasers in Engineering*, 2007, 45(2): 304-317.
- [6] Zuo C, Huang L, Zhang M L, et al. Temporal phase unwrapping algorithms for fringe projection profilometry: a comparative review [J]. *Optics and Lasers in Engineering*, 2016, 85: 84-103.
- [7] Su X Y, Chen W J. Reliability-guided phase unwrapping algorithm: a review [J]. *Optics and Lasers in Engineering*, 2004, 42(3): 245-261.
- [8] Zheng D L, Da F P, Qian K M, et al. Phase-shifting profilometry combined with gray-code patterns projection: unwrapping error removal by an adaptive median filter [J]. *Optics Express*, 2017, 25(5): 4700-4713.
- [9] Cheng Y Y, Wyant J C. Multiple-wavelength phase-shifting interferometry[J]. *Applied Optics*, 1985, 24(6): 804-807.
- [10] Da J, Qu H M, Tao T Y, et al. Real-time three-dimensional measurement composite of epipolar constraint and speckle correlation [J]. *Acta Optica Sinica*, 2016, 36(10): 1012003.  
笮健, 屈惠明, 陶天阳, 等. 结合极线约束和散斑相关的实时三维测量方法 [J]. *光学学报*, 2016, 36(10): 1012003.
- [11] Deng H X, Ling X, Wang Y Y, et al. High-speed and high-accuracy fringe projection profilometry without phase unwrapping[J]. *Optics and Lasers in Engineering*, 2021, 140: 106518.
- [12] LeCun Y, Bengio Y, Hinton G. Deep learning [J]. *Nature*, 2015, 521(7553): 436-444.
- [13] Luo W J, Schwing A G, Urtasun R. Efficient deep learning for stereo matching [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 5695-5703.
- [14] Kuznietsov Y, Stückler J, Leibe B. Semi-supervised deep learning for monocular depth map prediction [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2215-2223.
- [15] Zhong J X, Yin W, Feng S J, et al. Speckle projection profilometry with deep learning [J].

- Infrared and Laser Engineering, 2020, 49 (6): 20200011.  
钟锦鑫, 尹维, 冯世杰, 等. 基于深度学习的散斑投影轮廓术 [J]. 红外与激光工程, 2020, 49 (6): 20200011.
- [16] Feng S J, Chen Q, Gu G H, et al. Fringe pattern analysis using deep learning [J]. Advanced Photonics, 2019, 1(2): 025001.
- [17] Qiao G, Huang Y Y, Song Y P, et al. A single-shot phase retrieval method for phase measuring deflectometry based on deep learning [J]. Optics Communications, 2020, 476: 126303.
- [18] 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.
- [19] Budianto, Law W, Lun D P K. Deep learning based period order detection in structured light three-dimensional scanning [C] // 2019 IEEE International Symposium on Circuits and Systems (ISCAS), May 26-29, 2019, Sapporo, Japan. New York: IEEE Press, 2019: 18815870.
- [20] van der Jeught S, Dirckx J J J. Deep neural networks for single shot structured light profilometry [J]. Optics Express, 2019, 27(12): 17091-17101.
- [21] Zhang Y Y, Xiong Z W, Wu F. Unambiguous 3D measurement from speckle-embedded fringe [J]. Applied Optics, 2013, 52(32): 7797-7805.
- [22] Zhang Z, Han B W, Yu H T, et al. Multi-stage deep learning based single-frame fringe projection 3D measurement method [J]. Infrared and Laser Engineering, 2020, 49(6): 20200023.  
张钊, 韩博文, 于浩天, 等. 多阶段深度学习单帧条纹投影三维测量方法 [J]. 红外与激光工程, 2020, 49(6): 20200023.
- [23] Hara K, Saito D, Shouno H. Analysis of function of rectified linear unit used in deep learning [C] // 2015 International Joint Conference on Neural Networks (IJCNN), July 12-17, 2015, Killarney, Ireland. New York: IEEE Press, 2015: 15503788.
- [24] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [25] Kingma D P, Ba J. Adam: a method for stochastic optimization [EB/OL]. (2014-12-22) [2021-04-10]. <https://arxiv.org/abs/1412.6980>.