photonics

MDPI

Article

Composite Deep Learning Framework for Single-Shot Absolute 3D Shape Measurement Based on 1-Bit Speckle-Embedded Pulse-Width Modulation Pattern

Zhen Ren, Zhaoyan Li * and Zhisheng Gao *

School of Computer and Software Engineering, Xihua University, Chengdu 610039, China; renzhen@stu.xhu.edu.cn
* Correspondence: lzy@xhu.edu.cn (Z.L.); gzs_xihua@mail.xhu.edu.cn (Z.G.)

Abstract

In fringe projection profilometry (FPP), how to measure objects quickly and accurately remains an important issue. Fourier transform profilometry (FTP) can recover the wrapped phase from a single frame of fringes, but it suffers from severe distortion when reconstructing complex surfaces. Phase-shift profilometry (PSP) can obtain high-precision three-dimensional (3D) reconstruction results by multi-frame fringe patterns, but the increase in projection time leads to a decrease in measurement efficiency. How to use the single-frame pattern for fast and high-quality 3D reconstruction remains a difficult problem. We propose a composite deep learning framework for absolute 3D shape measurement using single-frame 1-bit patterns. The framework adopts a 1-bit pattern named speckle embedded PWM (SPPWM), which is obtained by embedding speckle into pulse width modulation (PWM) and combines the advantages of speckle and phase-shift. The framework consists of three steps: (1) We propose two convolutional neural networks, SPSNet and SMDNet, to recover high-precision wrapped phase and the corresponding mask from SPPWM. (2) Apply the semi-global matching (SGM) algorithm to obtain a coarse disparity map at the integer level. (3) According to the constraint relationship between the precise wrapped phase map and the coarse disparity map, the coarse disparity map is refined using the precise wrapped phase map without phase unwrapping, and the sub-pixel disparity map is directly obtained. Experimental results show that our proposed SPSNet and SMDNet can obtain high-precision wrapped phase from a single frame pattern. The proposed method can achieve high-precision absolute 3D measurement with an accuracy of 41 μm using only a single-frame 1-bit SPPWM pattern, achieving a good balance between projection speed and reconstruction accuracy.

Keywords: phase retrieval; fringe projection profilometry; deep learning; single-shot

Check for updates

Received: 14 September 2025
Revised: 8 November 2026
Accepted: 18 November 2026
Published: 23 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

With the rapid development of optical metrology, 3D measurement has become increasingly important in our lives. Among them, optical measurement is widely used in industrial manufacturing, medical testing, cultural relics scanning, human–computer interaction, product testing, reverse engineering and other fields due to its non-contact and high-precision characteristics [1–7]. A typical FPP system consists of a projector and a camera, as shown in Figure 1. During the measurement process, the projector casts a series of fringes, one after another, onto the object being measured. A camera is used to capture the fringe patterns modulated by the surface of an object. Subsequently, the 3D shape

of the object is reconstructed through steps such as phase unwrapping, phase expansion, and phase-to-height mapping. Currently, the two most widely used FPPs are FTP [8,9] and PSP [10–12]. FTP can recover phase information using only a single-frame fringe pattern. However, this method is susceptible to spectrum aliasing [13], noise interference [14] and sudden changes on the surface of the object [15], which can lead to errors in phase unwrapping and a large amount of distortion in the reconstruction results. Unlike FTP, PSP recovers the phase information of an object by projecting multiple fringe patterns in the time domain, thus achieving high-precision measurement. In addition, PSP is highly robust to different lighting environments and surfaces of different materials. However, the measurement efficiency is inversely proportional to the number of projected fringes, especially when measuring moving objects, during which the accuracy of the phase-shift method will drop sharply [16,17]. Therefore, how to shorten the PSP projection time as much as possible without reducing the measurement accuracy has become the focus of research by many scholars.

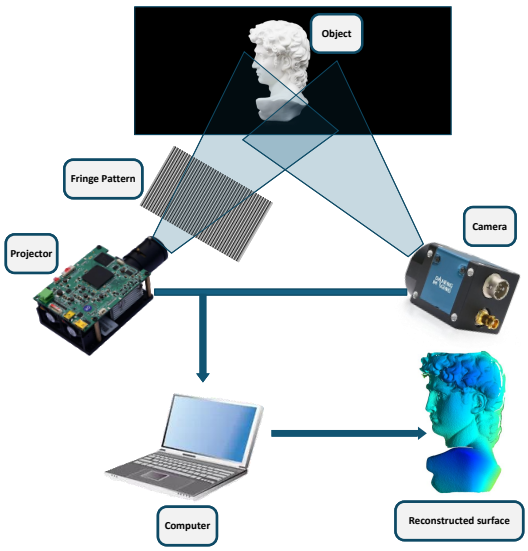


Figure 1. A typical FPP system.

The phase extracted by FPP through the inverse tangent function will be truncated to between $-\pi$ and π. In the phase jump region, a 2π jump may occur between adjacent pixels, resulting in phase ambiguity [18], which leads to the failure of stereo matching and the inability to find corresponding points. There are three methods of phase unwrapping: spatial phase unwrapping, temporal phase unwrapping, and stereo phase unwrapping. Spatial phase unwrapping relies on the relationship between adjacent phases in a single wrapped phase map to unwrap, but incorrect unwrapping often occurs in the case of complex surfaces [19]. Temporal phase unwrapping projects multiple auxiliary patterns in the time dimension to eliminate phase ambiguity. Commonly used temporal phase unwrapping methods include gray code phase unwrapping [20,21], multi-frequency heterodyne method [22], and number-theoretical phase unwrapping [23]. However, temporal phase unwrapping increases the number of projected fringes, thereby increasing the projection time.

Stereo phase unwrapping methods remove phase ambiguity by adding additional cameras and finding corresponding points in multiple cameras. To address the unreliability of high-frequency fringe patterns in stereo phase unwrapping, Li et al. [24] proposed a simple and effective deep anomaly detection strategy to improve the robustness of the unwrapping. Stereo phase unwrapping requires multiple cameras, which reduces the effective depth of measurement. The researchers addressed this issue by increasing the amount of information contained in each fringe. Yin et al. [25] embedded speckle into a

sinusoidal fringe pattern and obtained an accurate and distortion-free point cloud using only four fringe patterns. Lohry et al. [26] proposed a new measurement method based on stereo phase unwrapping. First, random patterns are embedded in sinusoidal fringes, and a coarse disparity map is obtained through the quality map. Then, the coarse disparity map is optimized by wrapped phase without the need for phase unwrapping. Lohry and Zhang [27] further used dithering technology to binarize the fringe pattern proposed in [26] to improve the projection speed. Li et al. [28] designed a fringe projection system embedded with speckle, which projects speckle and sinusoidal fringes onto the surface of an object in a superimposed manner through the cooperation between optical devices, thus achieving single-frame measurement. But this optical system requires the use of a lot of optical equipment. Song et al. [29] reconstructed 3D shapes using one frame of original image and one frame of fringe image. Gai et al. [30] projected a speckle pattern and a sinusoidal fringe pattern, bypassing phase unwrapping while increasing the number of projected fringes. These studies have shown that by carefully designing fringe patterns and reconstruction algorithms, high-quality 3D reconstruction can still be achieved while significantly reducing the number of projected patterns.

By reducing the number of projected fringes, the time required for projection can be significantly shortened. However, most methods adopt 8-bit patterns, which increases the projection time and introduces the nonlinear gamma effect of the projector. To solve this problem, researchers proposed the binary defocused projection technique [31]. The fringe pattern used in binary defocusing technology has a depth of only 1 bit, which far exceeds the fringe pattern with an 8-bit depth in projection speed, and the harmonics in the binary fringe can be removed by defocusing the projector. Researchers have proposed various binary fringe design methods, aiming to obtain better sinusoidal fringes after the projector is defocused. Yuan et al. [32] proposed a dual-focus optimization strategy, which solved the limitation of binary defocus that only the isolated front segment or rear segment can be used as the measurement depth range by adding a set of cylindrical lens groups. In order to solve the problem of reduced fringe contrast caused by high level defocus of the projector, Ayubi et al. [33] used a triangular wave to modulate a sine wave, resulting in a spatial version of PWM, which enabled the projector to measure at low defocus levels. Zheng et al. [34] proposed a set of complementary PWM patterns that cleverly eliminated high-frequency noise and still had a certain measurement capability in overexposed areas. Although binary defocusing technology greatly improves the single-frame projection speed, the information carried by each frame of binary pattern is usually less than that of an 8-bit pattern, so multiple frames still need to be projected to complete reliable 3D reconstruction.

In the past decade, with the development of deep learning, many researchers have combined deep learning with 3D reconstruction, which plays an increasingly important role in image denoising [35], phase recovery [36], and stereo matching [37,38]. Due to its powerful representation ability, researchers naturally think of using deep learning to reduce the number of projected fringes or reduce the time required to project each fringe. In order to resolve the contradiction between projection speed and reconstruction accuracy, Feng et al. [39] proposed a micro-deep learning profilometry method, which uses a neural network to extract phase information from three fringe patterns of different frequencies. Yin et al. [40] proposed a deep learning framework for 3D shape measurement based on fringe phase recovery and speckle correlation, which achieves 3D reconstruction by projecting two patterns. Zheng et al. [41] embedded speckle into PWM to obtain SP-PWM, and achieved high-precision 3D measurement using only three 1-bit patterns. These methods significantly reduce the number of projected fringes through deep learning, but experiments show that single-frame 3D measurement can be achieved through deep learning. Feng et al. [42] experimentally demonstrated that the use of neural networks can signifi-

cantly improve the accuracy of phase extraction from a single fringe pattern. Wan et al. [43] designed a deep network FRANET, which consists of three fringe analysis sub-networks and can achieve high-precision 3D measurement using single-frame sinusoidal patterns. Nguyen et al. [44] studied the differences in 3D reconstruction performance of six different single-frame structured lights. Yang et al. [45] proposed a single-shot phase extraction method based on a deep convolutional generative adversarial network to generate a phase map and quality mask from the input fringe pattern, thereby achieving single-frame 3D reconstruction. It is worth noting that these deep learning-based single-frame measurement methods all use 8-bit depth patterns instead of binary patterns that are projected faster.

In this work, inspired by binary defocus and deep learning, we propose a composite deep learning framework to achieve absolute 3D shape measurement using single-frame 1-bit SPPWM patterns. The SPPWM pattern adopted in this paper is obtained by embedding speckle in PWM. In our proposed method, the projector remains in focus at all times. PWM is chosen as the carrier because it employs one-dimensional modulation and is easy to generate; furthermore, it preserves more sinusoidal information during focusing. Square waves possess strong, fixed harmonics; furthermore, techniques such as dithering and random optimization coding—which involve two-dimensional modulation and must account for defocus—exhibit poor compatibility with our proposed method. The embedded speckle provides rich local random features, which significantly improves the discrimination of stereo matching algorithms such as SGM when calculating the matching cost. The SPPWM pattern also contains information about the sinusoidal fringe pattern, which provides the necessary information for deep learning to recover the wrapped phase from SPPWM. First, we use SPSNet and SMDNet to recover high-quality wrapped phase from SPPWM containing speckles and harmonics, and simultaneously obtain the object mask. In addition, the SPPWM images captured by the left and right cameras are input into the SGM algorithm to calculate the coarse disparity map. Finally, a simple and efficient disparity refinement algorithm based on wrapped phase and coarse disparity map is proposed, which can obtain high-precision sub-pixel disparity map without phase unwrapping. Experimental results show that our method can achieve high-quality absolute 3D measurement with an accuracy of 41 μm using only a single-frame SPPWM pattern.

This paper is organized as follows. Section 2 introduces the SPPWM pattern and our proposed composite deep learning framework. Section 3 experimentally validates our approach and compares it with other methods. Section 4 concludes this paper.

2. Principles

2.1. SPPWM Projection Pattern Description

The fringe pattern projected by N-step phase-shift profilometry is mathematically expressed as follows:

$$I_k^p(x,y) = 0.5 + 0.5\cos(2\pi f_s x + \delta_k) \tag{1}$$

where $I_k^p(x,y)$ $(k = 0, 1, 2, \cdots, N-1)$ represents the light intensity of the k-th phase-shift fringe at (x,y) in N-step phase-shift, f_s represents the spatial frequency of the fringe in the x direction. δ_k is the offset of the k-th phase-shift. We use the superscript p to denote the pattern projected by the projector. An ideal sinusoidal fringe pattern is shown in Figure 2a.

When a sinusoidal fringe pattern is projected, the fringes are modulated by the surface of the object, and the image captured by the camera contains the phase information of the object. The expression of the fringe pattern of N-step phase-shift captured by the camera is as follows:

$$I_k^c(x,y) = A^c(x,y) + B^c(x,y)\cos(\phi(x,y) + \delta_k) \tag{2}$$

where $I_k^c(x,y)(k = 0,1,2,\cdots,N-1)$ represents the light intensity of the k-th phase-shift fringe at (x,y) in the N-step phase-shift captured by the camera. $A^c(x,y)$ represents the background light intensity of the fringe at (x,y). $B^c(x,y)$ represents the modulation amplitude of the fringe at (x,y), reflecting the contrast of the fringe. $\phi(x,y)$ is the phase information of the solution. δ_k has the same meaning as the projection fringe. We use superscript c to represent the pattern captured by the camera.

For the measured object, the N-step phase-shift method uses the least squares method to solve the numerator (sine component) $M^c(x,y)$ and denominator (cosine component) $D^c(x,y)$ of the object's wrapped phase. Finally, the wrapped phase $\phi_w^c(x,y)$ is solved by the inverse tangent function:

$$M^c(\mathrm{x},\mathrm{y}) = \sum_{k=0}^{N-1} I_k^c(\mathrm{x},\mathrm{y})\sin(\delta_k) \tag{3}$$

$$D^c(\mathrm{x},\mathrm{y}) = \sum_{k=0}^{N-1} I_k^c(\mathrm{x},\mathrm{y})\cos(\delta_k) \tag{4}$$

$$\phi_w^c(\mathrm{x},\mathrm{y}) = \tan^{-1}\left(\frac{M^c(\mathrm{x},\mathrm{y})}{D^c(\mathrm{x},\mathrm{y})}\right) \tag{5}$$

In practical applications, in addition to the object being measured, there is a lot of background information that we do not need. In order to display the measured object more intuitively, we often calculate the effective object area mask to avoid the influence of background pixels:

$$B^c(x,y) = \frac{2}{N}\sqrt{M^c(\mathrm{x},\mathrm{y})^2 + D^c(\mathrm{x},\mathrm{y})^2} \tag{6}$$

$$\mathrm{Mask}^c(x,y) = \begin{cases} 1 & B(x,y) \geq thr \\ 0 & B(x,y) < thr \end{cases} \tag{7}$$

where thr represents the threshold value set based on $B^c(x,y)$ and is set according to the shooting environment.

The PWM pattern was proposed by Li et al. [46], which uses a triangular carrier $Tri(x,y)$ to modulate the ideal sinusoidal signal $I_k^p(x,y)$. Subsequently, Zuo et al. [47] reduced the amplitude of the triangular carrier by 3% to reduce the impact of harmonics. The triangular wave can be expressed mathematically as [48]:

$$\mathrm{Tri}(x,y) = \begin{cases} 1.94xf_c - 1.94Z + 0.015 & x \in (Z/f_c, (2Z+1)/2f_c) \\ -1.94xf_c + 1.94(Z+1) + 0.015 & x \in ((2Z+1)/2f_c, (Z+1)/f_c) \end{cases} \tag{8}$$

where $Tri(x,y)$ represents the amplitude of the triangular wave at (x,y), f_c is the frequency of the triangular wave, and Z is an integer. Setting $f_s = 36$, $f_c = 10f_s$ can obtain high-quality reconstruction results [34]. Then PWM can be expressed as:

$$PWM_k^p(x,y) = \begin{cases} 1 & I_k^p(x,y) > \mathrm{Tr}\,i(x,y) \\ 0 & I_k^p(x,y) \leq \mathrm{Tr}\,i(x,y) \end{cases} \tag{9}$$

The ideal PWM pattern obtained by Equation (9) is shown in Figure 2b. This paper uses the method proposed by Zheng et al. [41] to embed speckle into PWM to obtain SPPWM. The embedding steps are as follows: (1) The speckle image $S(x,y)$ is generated using the software proposed by Su et al. [49], and the speckle parameters are set the same as in [41]. $S(x,y)$ is binarized to obtain $S_{\mathrm{bin}}(x,y)$, as shown in Figure 2c. (2) According to Equation (10), embedding $S_{\mathrm{bin}}(x,y)$ into PWM can obtain SPPWM. The SPPWM pattern is shown in Figure 2d. This method embeds 1-bit speckle into 1-bit pattern PWM to obtain

1-bit SPPWM, so that SPPWM retains the dynamic range of the original pattern while increasing the information capacity.

$$\mathrm{SPPWM}_k^p(x,y) = \begin{cases} S_{bin}(x,y) & \mathrm{PWM}_k^p = 0 \\ \mathrm{PWM}_k^p(x,y) - S_{bin}(x,y) & \mathrm{PWM}_k^p = 1 \end{cases} \tag{10}$$

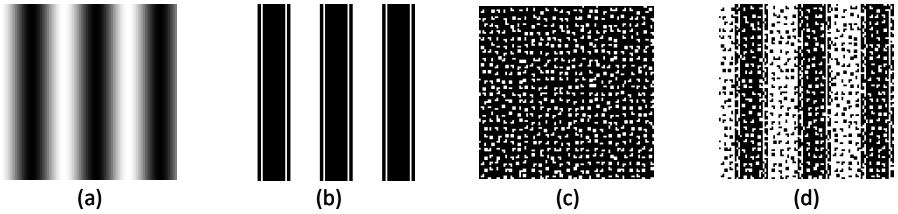


Figure 2. SPPWM generation process. (**a**) Sinusoidal fringe pattern; (**b**) PWM pattern; (**c**) Speckle pattern; (**d**) SPPWM pattern.

2.2. Calculating the Coarse Disparity Map

In our proposed framework, the framework consists of four main modules: wrapped phase recovery, mask computation, coarse disparity map computation, and coarse disparity map refinement. The overall flow chart is shown in Figure 3.

In order to efficiently obtain integer-level disparity estimation, SGM is used to calculate the coarse disparity map. Specifically: (1) Matching cost calculation: To enhance the robustness to illumination changes, generate a bit string descriptor of the local area of the image using Census transform based on a 5×5 window. The initial matching cost is obtained by calculating the Hamming distance between the Census descriptors of corresponding points in the left and right images. (2) Cost aggregation: Aggregate the initial matching costs along 8 directions to enhance the robustness of matching. (3) Disparity calculation: For each pixel, a Winner-Takes-All (WTA) strategy is used. The disparity value with the smallest aggregation cost is selected as the final coarse disparity estimate for that pixel.

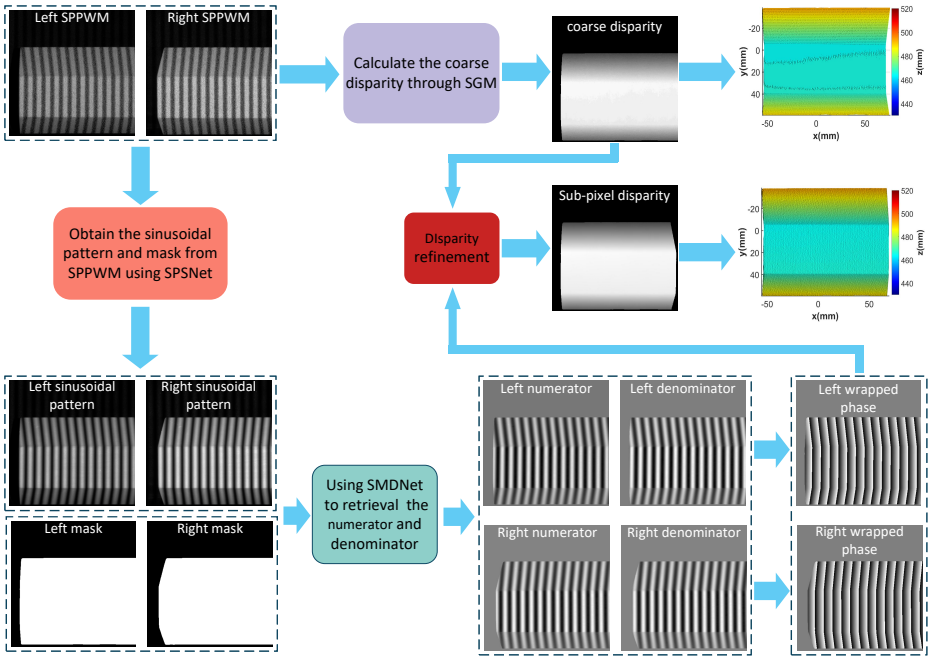


Figure 3. Schematic diagram of the composite deep learning framework for absolute 3D shape measurement using single-frame 1-bit patterns.

2.3. Neural Network for Recovering Sinusoidal Pattern from SPPWM Pattern and Calculating Mask

Since it is extremely difficult to directly recover the wrapped phase from SPPWM, our process of recovering the wrapped phase from SPPWM is not accomplished by a single network. First, we propose a Y-Net-structure network SPSNet to recover the sinusoidal pattern and mask from SPPWM, and then use a U-Net-structure network SMDNet to recover the wrapped phase from the sinusoidal pattern.

The input data of the SPSNet network is the image of SPPWM projected onto the surface of the object taken by the camera, expressed as $SPPWM_k^c$. The output of the network is a high-quality sinusoidal fringe pattern and mask, denoted by $\widehat{I_k^c}$ and $\widehat{\text{Mask}}^c$, respectively. The width and height of the input image are denoted by W and H respectively.

Our proposed SPSNet is shown in Figure 4. The encoding path on the left consists of four feature extraction modules with the same structure, the bottleneck layer in the middle, and two independent decoding paths on the right. In the lower right corner is the Recovery Sinusoidal Pattern Path (RSPP), whose task is to recover the sinusoidal pattern from the SPPWM. The upper left corner is the Semantic Segmentation Path (SSP), whose task is to obtain the mask from the semantic segmentation of SPPWM. Two independent decoding paths each consist of four repeated stages.

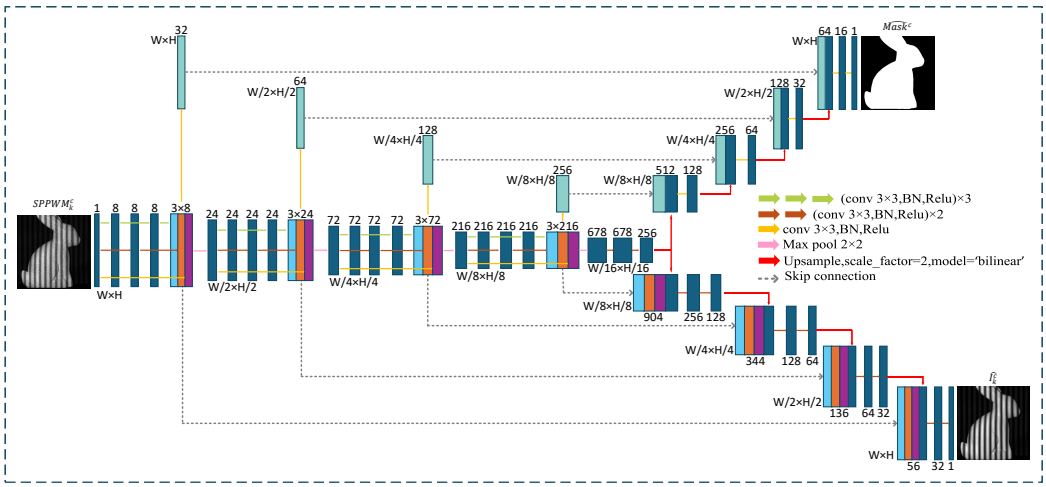


Figure 4. SPSNet structure diagram, where the blocks represent feature maps. The numbers above or below the blocks represent the number of feature maps. The labels with W and H below or to the left of the block represent the size of the feature map. The arrow on the right is the operating instructions.

In the encoding path, we use a single 3 × 3 convolution (yellow), two 3 × 3 convolutions (brown) and triple 3 × 3 convolutions (green) to extract three sets of feature maps; Then batch normalization (BN) and rectified linear unit (ReLU) are performed. These three sets of feature maps are concatenated, downsampled using a 2 × 2 max pooling operation with a stride of 2, and then passed to the next stage. After four passes, it is downsampled again using a 2 × 2 max pooling operation with a stride of 2 and then passed to the bottleneck layer. The bottleneck layer uses two double 3 × 3 convolutions (brown) followed by BN and ReLU to better extract high-level semantic features.

In RSPP, bilinear interpolation is used to upsample the feature maps. Skip connections concatenate the upsampled feature maps with those from the encoding path to integrate multi-scale information. Double 3 × 3 convolutions (brown) are then used, followed by BN and ReLU.

In SSP, since the semantic segmentation task is simpler than recovering the sinusoidal pattern, the concatenated feature maps at each stage of the encoding path undergo a single

3 × 3 convolution (yellow) followed by BN and ReLU to compress the number of channels. Bilinear interpolation is used to upsample the feature maps. The upsampled feature maps and the compressed feature maps are concatenated using skip connections to integrate multi-scale information. A single 3 × 3 convolution (yellow) is then used, followed by BN. Except for the last convolution layer which is activated by a sigmoid function, the rest are activated by ReLU.

When training the network, since the task of RSPP is more difficult than SSP, we first train the encoder, bottleneck layer and RSPP. After RSPP training is completed, the parameters of the encoder, bottleneck layer and RSPP are fixed, and then the SSP with relatively simple task difficulty is trained. In order to better train RSPP, we calculate the loss only over valid regions. The mask is calculated using Equation (7). In this paper, $thr = 0.015$ can meet the requirements of all scenarios.

The task of RSPP is to minimize the following loss function:

$$\mathrm{Loss}(I_\theta) = \frac{\left\|\mathrm{Mask}^c \odot \left(I_k^c - \widehat{I_k^c}\right)\right\|_2}{\|\,\mathrm{Mask}\,\|_1} \tag{11}$$

where I_θ represents the encoder, bottleneck layer, and RSPP parameters in the model, and $\widehat{I_k^c}$ represents the output of RSPP. $\|\cdot\|_2$ represents the L_2 norm, and $\|\cdot\|_1$ represents the L_1 norm. $\odot$ represents element-wise product, and the task of SSP is to minimize the following loss function:

$$\mathrm{Loss}(\mathrm{Mask}_\theta) = -\left[\mathrm{Mask}^c \cdot \log\left(\widehat{\mathrm{Mask}^c}\right) + \left(1 - \mathrm{Mask}^c\right) \cdot \log\left(1 - \widehat{\mathrm{Mask}^c}\right)\right] \tag{12}$$

where Mask_θ represents the parameters of SSP, $\widehat{\mathrm{Mask}^c}$ represents the output of SSP. Because SSP is finally mapped to $[0, 1]$ by a sigmoid function, the mask is finally obtained through Equation (13). When $\widehat{\mathrm{Mask}^c} \geq 0.5$, the pixel is considered to belong to the foreground; otherwise it is considered to belong to the background.

$$\mathrm{Mask} = \begin{cases} 1 & \widehat{\mathrm{Mask}^c} \geq 0.5 \\ 0 & \widehat{\mathrm{Mask}^c} < 0.5 \end{cases} \tag{13}$$

2.4. Neural Network for Recovering the Wrapped Phase from Sinusoidal Patterns

This section will introduce SMDNet, an improved U-Net. This network aims to recover the corresponding wrapped phase from the sinusoidal fringe pattern output by SPSNet in Section 2.3. However, since the inverse tangent function in Equation (5) is difficult to learn, we use the inverse tangent function to solve the wrapped phase by learning the numerator and denominator of the wrapped phase.

It is worth noting that the sinusoidal fringe pattern containing background noise output by SPSNet is not directly input into the network, by the mask before being input to the network and is denoted as $\mathrm{Mask} \odot \widehat{I_k^c}$. The numerator and denominator of the output are expressed as $\widehat{M^c}$ and $\widehat{D^c}$, respectively. The width and height of the input image are denoted by W and H respectively.

The SMDNet structure diagram is shown in Figure 5. It adopts the strategy of encoding first and decoding later. In the encoding path, we use 3 × 3 convolution (green) to extract the feature map, followed by BN and ReLU. Downsampling is performed using a 2 × 2 max pooling operation with a stride of 2 and then passed to the next stage. After three passes, it is downsampled again using a 2 × 2 max pooling operation with a stride of 2 and then passed to the bottleneck layer. To prevent overfitting, the dropout rate of the bottleneck layer is set to 0.2. Upsampling is done using bilinear interpolation in the decoding path. After each upsampling, a 3 × 3 convolution is performed (green), and then BN and ReLU are used to compress the feature map to half of its original size before concatenating it

with the shallow feature map. Then a 3 × 3 convolution (green) is used, followed by BN. Except for the last convolution layer, which does not use an activation function, other layers still use the ReLU function for activation.

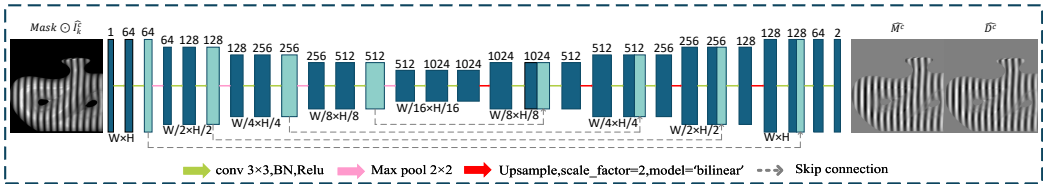


Figure 5. SMDNet structure diagram. The blocks represent feature maps; the numbers above the blocks represent the number of feature maps; the labels with W and H below the blocks represent the size of the feature maps; and the arrows on the bottom are operation instructions.

Likewise, the SMDNet network only calculates the loss of the valid area to output more accurate results. The task of the SMDNet network is to minimize the following loss function:

$$\text{Loss}(\theta) = \frac{\left\|\text{Mask}^c \odot \left(M^c - \widehat{M^c}\right)\right\|_2}{\left\|\text{Mask}^c\right\|_1} + \frac{\left\|\text{Mask}^c \odot \left(D^c - \widehat{D^c}\right)\right\|_2}{\left\|\text{Mask}^c\right\|_1} \tag{14}$$

where θ represents the parameters in the model, $\|\cdot\|_2$ represents the L_2 norm, $\|\cdot\|_1$ represents the L_1 norm, and $\odot$ represents element-wise product.

2.5. Refining the Disparity Map

In Sections 2.2–2.4, we obtain the coarse disparity map and the wrapped phase map and mask of the left and right images, where the coarse disparity map provides a rough matching position. Now all we need to do is use the integer disparity provided by the coarse disparity map, and according to the constraints of the wrapped phase image, the sub-pixel disparity map is calculated within a certain range. Due to the periodicity of the wrapped phase, the jump region will appear periodically. During the calculation process, the calculation of the jump region is more complicated. Therefore, when the reference pixel and the candidate pixel are in the phase jump region, it is very important to eliminate the impact of the phase jump. The disparity refinement method we propose is shown in Figure 6.

Gai et al. [30] discussed three scenarios in the non-jump region and four scenarios in the jump region. In the jump region, the phase value is made continuous by adding 2π, thereby eliminating the effect of the jump. This approach produced correct matches in most locations. However, Gai et al. [30] ignored the case where only one of the jump detection boxes in the reference and target images contains the jump region while the other does not. This results in incorrect matching of some areas. Figure 6a,b show the cases where the reference pixel $\phi_L(X)$ and the candidate pixel $\phi_R(X_m)$ are not located in the jump region at the same time. Here, $(X) = (x, y), (X_m) = ((x - d(x, y) + \sigma), y)$, where σ is the size of the search window around the coarse matching position provided by the coarse disparity map. In this paper, we take $\sigma = 0, \pm1, \pm2, \pm3$, a total of seven adjacent candidate pixels, and the jump detection box is also the same size.

Regardless of whether $\phi_L(X)$ and $\phi_R(X_m)$ are in the jump region or not, the same calculation method is used. The specific calculation process is as follows:

$$\phi_{\max} = \max\{\phi_L(X), \phi_R(X_m)\} \tag{15}$$

$$\phi_L(X) = \begin{cases} X + 2\pi & |\phi_{\max} - \phi_L(X)| > 1 \\ X & |\phi_{\max} - \phi_L(X)| < 1 \end{cases} \tag{16}$$

$$\phi_R(X_m) = \begin{cases} X_m + 2\pi & |\phi_{\max} - \phi_R(X_m)| > 1 \\ X_m & |\phi_{\max} - \phi_R(X_m)| < 1 \end{cases} \tag{17}$$

$$\phi_d(X, \sigma) = |\phi_L(X) - \phi_R(X_m)| \tag{18}$$

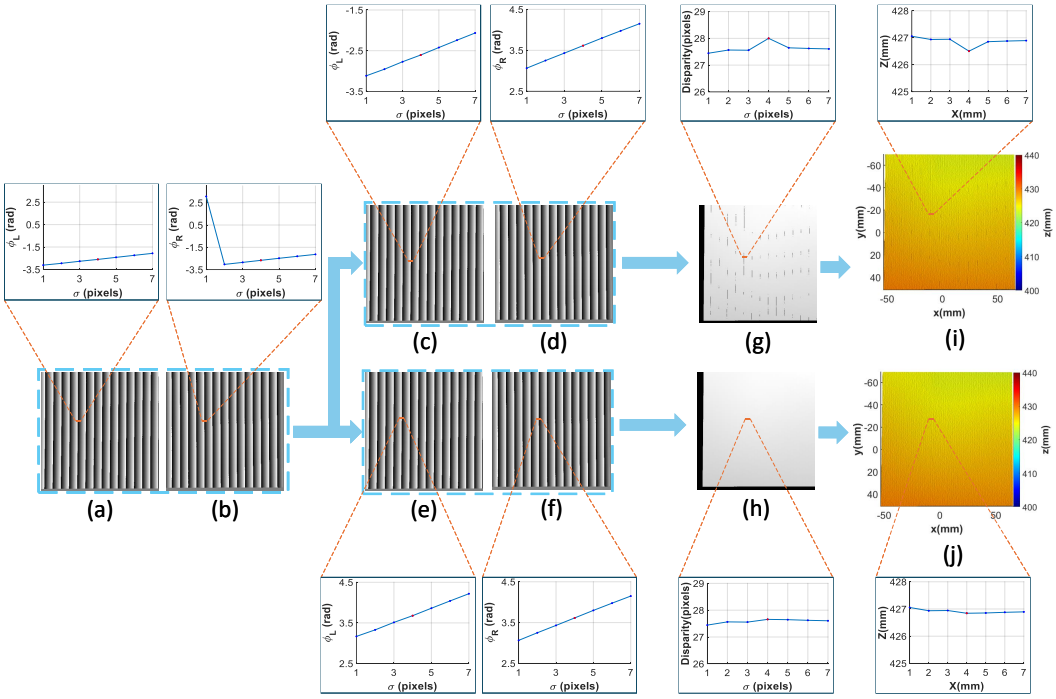


Figure 6. Disparity refinement process. (**a**) and (**b**) are the wrapped phase images ϕ_L and ϕ_R after epipolar correction, respectively. (**c**) and (**d**) are schematic diagrams of $\phi_L(X)$ and $\phi_R(X_m)$ when refining the disparity using the method in [30], respectively. (**e**) and (**f**) are schematic diagrams of $\phi_L(X)$ and $\phi_R(X_m)$ when refining the disparity using our method, respectively. (**g**) and (**h**) are sub-pixel disparity maps obtained using the method in [30] and our method, respectively. (**i**) and (**j**) are the reconstruction results of the sub-pixel disparity maps in (**g**) and (**h**), respectively.

First, use Equation (15) to calculate the maximum value $\phi_{\max}$ of $\phi_L(X)$ and $\phi_R(X_m)$. Then, use Equations (16) and (17) to calculate the absolute value of the difference between $\phi_L(X)$ and $\phi_{\max}$ and between $\phi_R(X_m)$ and $\phi_{\max}$, respectively. If it is greater than 1, it is considered that a phase jump occurs during this matching process, and the influence of the phase jump is eliminated by adding 2π. Then, use Equation (18) to calculate the absolute value of the difference between $\phi_L(X)$ and $\phi_R(X_m)$.

$$\sigma_{\min}(X) = \arg\min_{\sigma} \phi_d(X, \sigma) \tag{19}$$

$$d_{\phi}^{int}(X) = d(X) - \sigma_{\min}(X) \tag{20}$$

Using the idea of WTA in Equation (19), the best integer disparity matching point in the target image is determined, where d_{ϕ}^{int} is the pixel relationship between the reference pixel and the corresponding point. Finally, the sub-pixel disparity map is calculated by Equation (21):

$$d_{\phi}^{sub}(X) = \begin{cases} \frac{\phi_d(X,\sigma_{\min})}{\phi_d(X,\sigma_{\min}+1)-\phi_d(X,\sigma_{\min})} & \phi_d(X,\sigma_{\min}) > 0 \& \phi_d(X,\sigma_{\min}+1) < 0 \\ \frac{\phi_d(X,\sigma_{\min})}{\phi_d(X,\sigma_{\min})-\phi_d(X,\sigma_{\min}-1)} & \phi_d(X,\sigma_{\min}) < 0 \& \phi_d(X,\sigma_{\min}-1) > 0 \end{cases} \tag{21}$$

Figure 6h shows the sub-pixel disparity map $d_{\phi}^{\text{sub}}(X)$ obtained using our method. It can be seen that our proposed method can obtain accurate sub-pixel disparity maps from integer disparity maps. Figure 6g shows the disparity map computed using the method

in [30], which produces incorrect matches in the jump region. As shown in Figure 6i, the reconstruction result obtained using the method in [30] has many continuous abnormal points in the jump region. As shown in Figure 6j, the surface reconstructed using our proposed method has no abnormal points in the jump region, and high-precision 3D measurement can be achieved.

In Section 2.4, the numerator and denominator of the wrapped phase are recovered from SMDNet, and then the wrapped phase is obtained by the inverse tangent function. Compared with the real data, in the phase jump region, part of the phase inevitably changes from $-\pi$ to π or from π to $-\pi$. Our method cleverly eliminates this effect by directly finding the maximum value.

3. Experiments

3.1. Dataset Creation and Network Training

In order to verify the actual performance of our proposed composite deep learning framework, we built a binocular FPP system. The system consists of a DLP projector (DLP 4500, Texas Instruments, Dallas, TX, USA) with a resolution of 912 × 1140 and two monochrome industrial cameras (Hikvision CA004-10UM, Hangzhou Hikvision Digital Technology Co., Ltd., Hangzhou, China) with a resolution of 720 × 540. The distance between the two cameras is about 150 mm, and the distance between the entire system and the object to be measured is about 450 mm. The experimental equipment is shown in Figure 7.

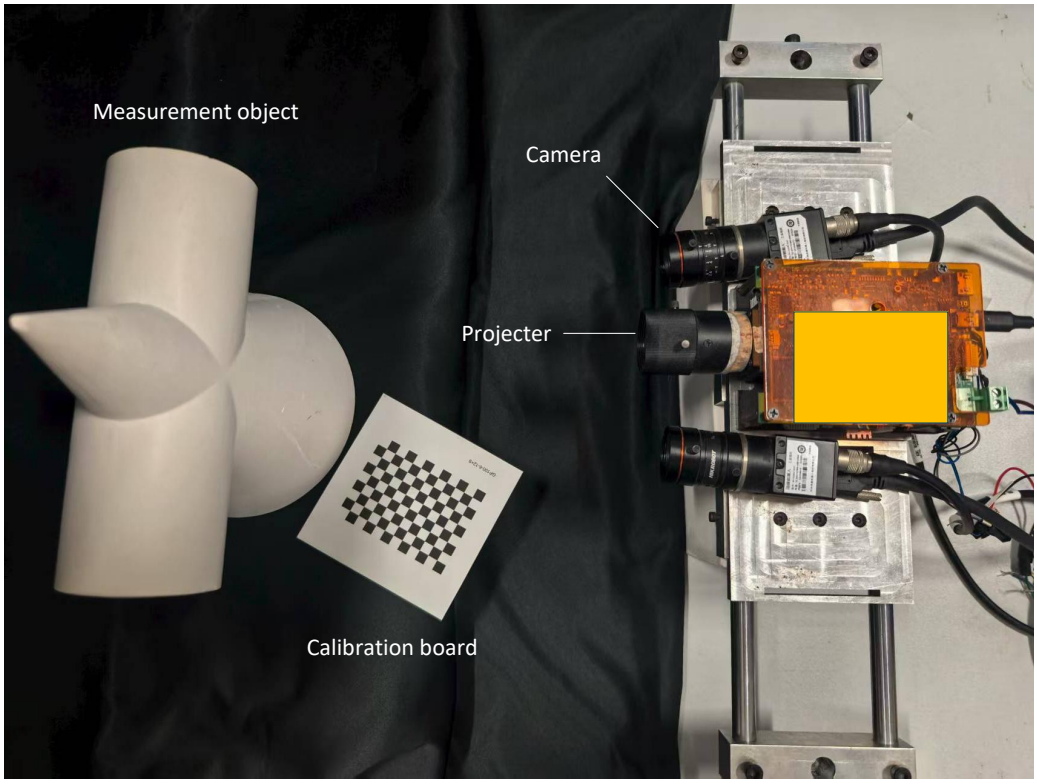


Figure 7. Experimental equipment.

Both networks in our proposed composite deep learning framework are data-driven and require a large amount of data to learn parameters and thus acquire powerful representation capabilities. When using deep learning for 3D reconstruction, the construction of the dataset is very important. An accurate dataset is directly related to the results of network prediction. We projected the twelve-step sinusoidal pattern with $N = 12$, $f_s = 1/36$ described

in Equation (1), and the SPPWM pattern corresponding to the sinusoidal fringes with $\delta_k = 0$, These patterns are modulated by the object's surface and captured synchronously by two cameras. After epipolar correction, we use Equations (3)–(7) to calculate the wrapped phase numerator, denominator and mask corresponding to the left and right cameras of the measured object as the data set. By randomly combining more than a dozen objects, 1300 sets of data were obtained, of which 1000 sets of data were used for training the network, 150 sets of data were used for verification, and 150 sets of data were used for testing. Figure 8 shows a part of our dataset. The first column is the SPPWM pattern, the second column is the sinusoidal pattern corresponding to SPPWM, the third and fourth columns are the numerator and denominator calculated by the twelve-step phase-shift method, and the fifth column is the mask.

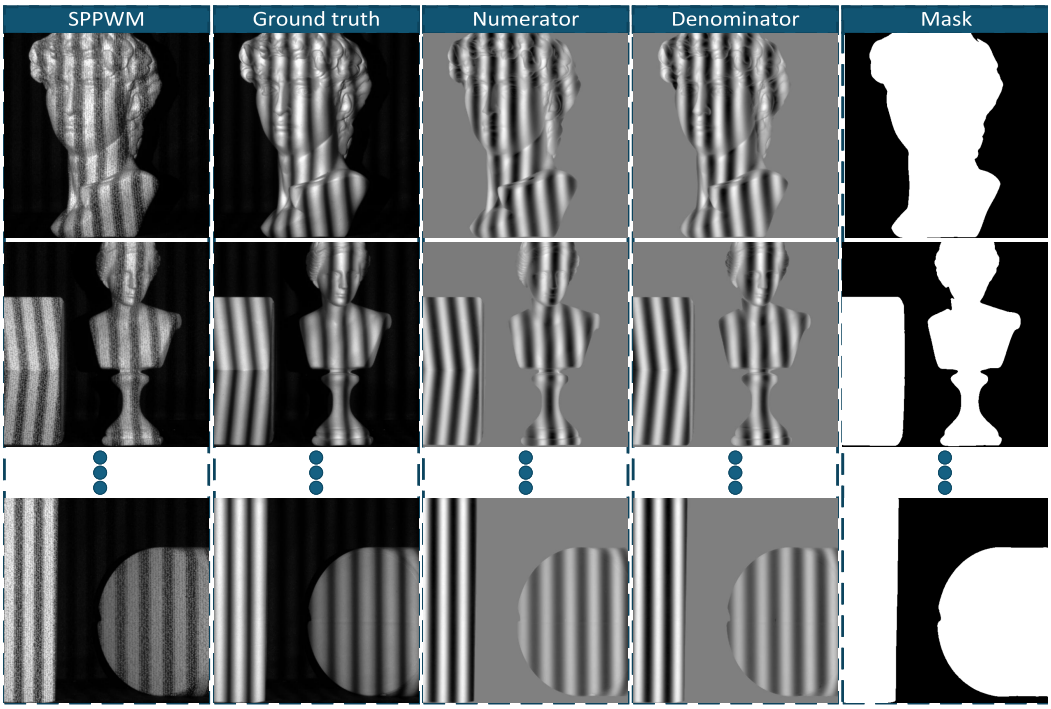


Figure 8. Part of the dataset used in the experiment.

When training the SPSNet network, we divide the input image by 255 for normalization to accelerate the convergence of the network. The initial learning rate of SPSNet is 2×10^{-5}, which is multiplied by 0.9 every 30 epochs. The batch size is set to 8. Due to the different difficulties of recovering sinusoidal patterns and semantic segmentation tasks, the number of training epochs for RSPP and SSP is set to 400 and 200 epochs, respectively. The total training time for the SPSNet network is 14.8 h, and the inference time is 14.2 milliseconds. The training and validation loss curves for RSPP are shown in Figure 9a, while the corresponding curves for SSP are presented in Figure 9b. The initial learning rate of SMDNet is 1×10^{-3}, which is multiplied by 0.9 every 30 epochs. The batch size is set to 8. The number of training epochs is set to 500. The total training time for the SMDNet network is 23.6 h, and the inference time is 22.8 milliseconds. The training and validation loss curves for the numerator component of the wrapped phase are shown in Figure 9c, while those for the denominator component of the wrapped phase are presented in Figure 9d. Both of these neural networks are implemented using the PyTorch framework (version 1.10.0) and were trained on an NVIDIA GeForce RTX 3090 (NVIDIA Corporation, Santa Clara, CA, USA).

To evaluate the performance of SPSNet, we randomly selected three SPPWM images from the dataset and used the SPSNet network to obtain the predicted sinusoidal fringe

patterns and the corresponding masks. As shown in Figure 10, the first column is the SPPWM pattern, the second column is the corresponding sinusoidal fringe pattern, the third column is the sinusoidal pattern predicted by SPSNet, and the fourth and fifth columns are the masks calculated using the phase-shift method and the masks predicted by SPSNet, respectively. Figure 11 shows the intensity of 512 pixels in row 200 and columns 0–512 of Figure 10k–m. It can be seen that our SPSNet eliminates speckles and harmonics in SPPWM and recovers high-quality sinusoidal patterns from SPPWM.

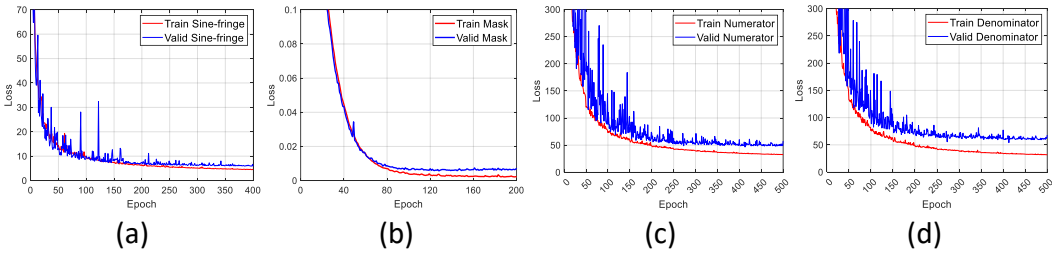


Figure 9. Network loss curve. (**a**) RSPP loss curve, (**b**) SSP loss curve, (**c**) Loss curves for numerator tasks as predicted by SMDNet, (**d**) Loss curves for denominator tasks as predicted by SMDNet.

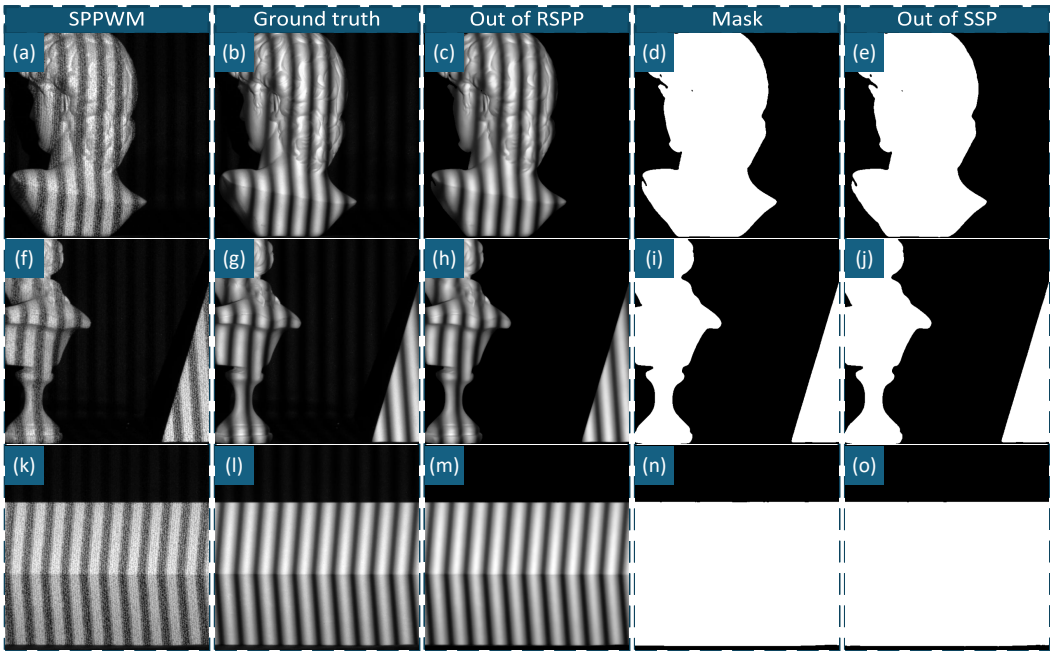


Figure 10. Prediction results of the SPSNet network. (**a**,**f**,**k**) are SPPWM images captured by the camera; (**b**,**g**,**l**) are the sinusoidal fringe patterns corresponding to SPPWM; (**c**,**h**,**m**) are the sinusoidal fringe patterns predicted by SPSNet; (**d**,**i**,**n**) are the masks calculated by the twelve-step phase-shift method; (**e**,**j**,**o**) are the masks predicted by SPSNet.

To test the performance of our proposed SMDNet, we selected three sets of data for testing. The images fed into the network are shown in the first column of Figure 12. These are all sinusoidal fringe patterns recovered from SPPWM by SPSNet. The numerator and denominator predicted by SMDNet are shown in the third and fifth columns of Figure 12, respectively. The ground-truth numerator and denominator are shown in the second and fourth columns of Figure 12, respectively. Their wrapped phase can be calculated using Equation (5). The first row of Figure 13 shows the true wrapped phases for the three scenarios mentioned above, and the second row shows the wrapped phases calculated based on the numerator and denominator predicted by SMDNet. To more intuitively show

the comparison between the true value and the predicted value, we plotted a comparison between the ground truth and the predicted values for the "David with Cube" composition at row 100, across 512 pixels. As shown in Figure 14, Figure 14a is a comparison of Figure 12l,m, Figure 14b is a comparison of Figure 12n,o, and Figure 14c is a comparison of their calculated wrapped phases. It can be seen that SMDNet recovers high-quality wrapped phase from SPPWM.

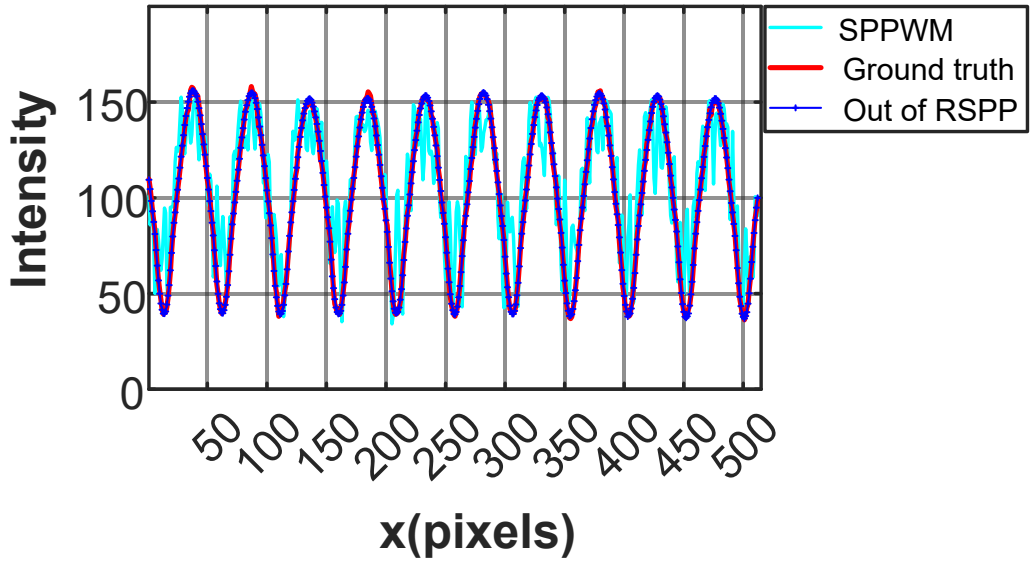


Figure 11. Comparison of pixel values of SPSNet prediction results. Comparison of pixel values in row 200 of Figure 10k–m.

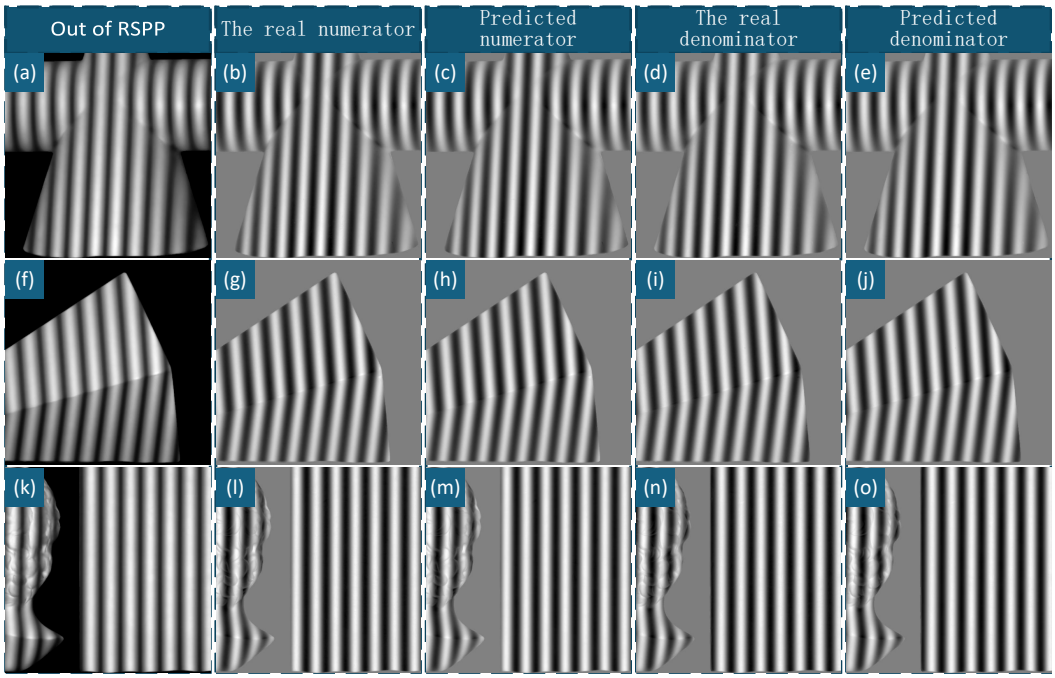


Figure 12. SMDNet network prediction results. (**a**,**f**,**k**) are the sinusoidal fringe patterns predicted by SPSNet; (**b**,**g**,**l**) are the true numerators; (**c**,**h**,**m**) are the numerators predicted by SMDNet; (**d**,**i**,**n**) are the true denominators; (**e**,**j**,**o**) are the denominators predicted by SMDNet.

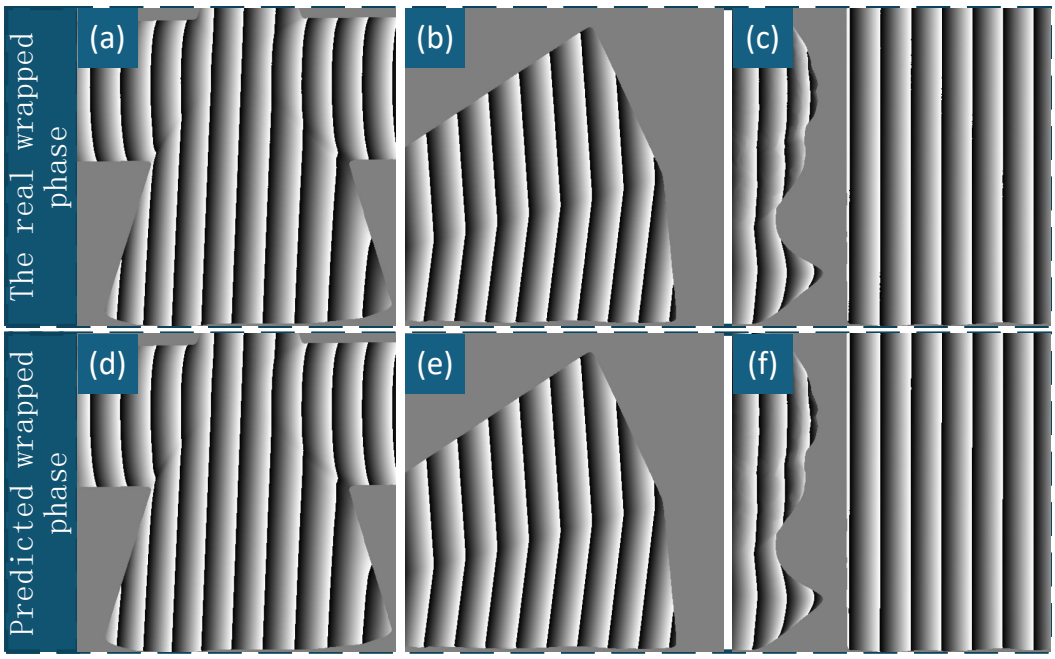


Figure 13. Comparison of the true parcel phase and the parcel phase predicted by SMDNet. (**a**–**c**) are the true wrapped phases of Figure 12a,f,k; (**d**–**f**) are the wrapped phases of Figure 12a,f,k predicted by SMDNet.

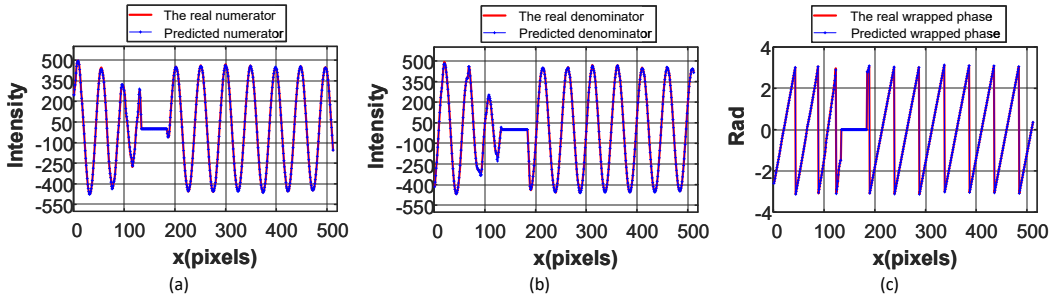


Figure 14. Comparison of pixel intensities between SMDNet prediction results and ground truth values. (**a**) Numerator comparison; (**b**) Denominator comparison; (**c**) Wrapped phase comparison.

3.2. Qualitative Analysis of 3D Reconstruction Accuracy

In this section, we use comparative experiments to evaluate the performance of our proposed method and other methods. We reproduced several other methods: (1) Recovering depth from a single-frame 8-bit sinusoidal fringe pattern via an end-to-end network (SFSPDR); (2) TFSPPWM proposed by Zheng et al. [41]; (3) 8-bit sinusoidal twelve-step phase-shifted algorithm (STSPSA). It should be noted that the coarse disparity map used by TFSPPWM is obtained through the SGM algorithm in Section 2.2. We selected three objects from the test dataset to conduct comparative experiments. Figure 15 presents the 3D reconstruction results of the measured object obtained using SFSPDR, STSPSA, TFSPPWM, and our proposed method, respectively. Figure 15a,e,i show the results for SFSPDR. SFSPDR uses an 8-bit sinusoidal pattern as input to directly generate a disparity map, and this algorithm does not incorporate any physical constraints; instead, it leverages deep learning to achieve a direct from fringe patterns to depth maps. This simple implementation process loses many details, resulting in poor reconstruction quality. Figure 15c,g,k show the results for TFSPPWM. This method requires the projection of only three 1-bit patterns to perform 3D measurement. Since the computational process is similar to that of the "three-step sinusoidal pattern method," it is susceptible to the influence of harmonics, which results in

a relatively rough reconstructed surface. The method we propose (Figure 15d,h,l) achieves reconstruction quality comparable to that of STSPSA (Figure 15b,f,j) while using only a single 1-bit image experimental results show that our method can achieve high-precision 3D measurement using a single-frame 1-bit pattern.

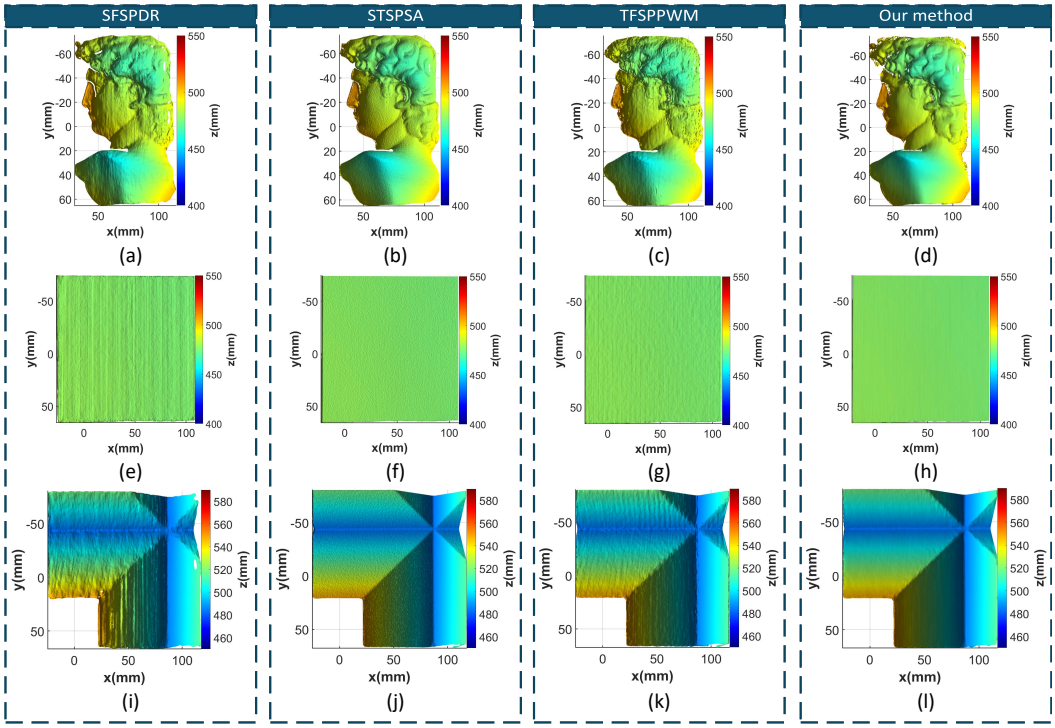


Figure 15. Performance comparison of 3D reconstruction results of vases using different methods. (**a**,**e**,**i**) Reconstruction result of SFSPDR; (**b**,**f**,**j**) Reconstruction result of STSPSA; (**c**,**g**,**k**) Reconstruction result of TFSPPWM; (**d**,**h**,**l**) Reconstruction result of our method.

3.3. Quantitative Analysis of 3D Reconstruction Accuracy

Finally, for quantitative analysis, we measured a standard sphere with a diameter of 40 mm and compared it with the other three methods mentioned in Section 3.2. Figure 16a–d show the results of 3D reconstruction of a standard sphere using SFSPDR, STSPSA, TFSPPWM, and our method, respectively. As shown in Figure 16a, due to the lack of physical constraints in SFSPDR, the reconstructed sphere contains many anomalous points. As shown in Figure 16b, the sphere reconstructed using STSPSA is essentially free of anomalous points. As shown in Figure 16c, there are still some abnormal points on the spherical surface reconstructed by TFSPPWM, but the reconstructed spherical surface is much smoother than that reconstructed by SFSPDR. As shown in Figure 16d, the spherical surface reconstructed using our method is smoother than those reconstructed by the other three methods. Figure 16e–h show the errors between the point clouds obtained by SFSPDR, STSPSA, TFSPPWM and our method and the standard point cloud, respectively. The mean absolute error (MAE) of SFSPDR is 0.1018 mm and the standard deviation (STD) is 0.1298 mm. The mean absolute error (MAE) of STSPSA is 0.0552 mm and the standard deviation (STD) is 0.0696 mm. The MAE of TFSPPWM is 0.0692 mm and the STD is 0.0849 mm. The MAE of our method is 0.0404 mm and the STD is 0.0561 mm. Figure 16i–l shows the statistical distribution of 3D reconstruction errors of SFSPDR, STSPSA, TFSPPWM and our method, respectively. Quantitative analysis experiments once again prove that our proposed method can achieve high-precision and smooth 3D measurement using only one frame of binary image.

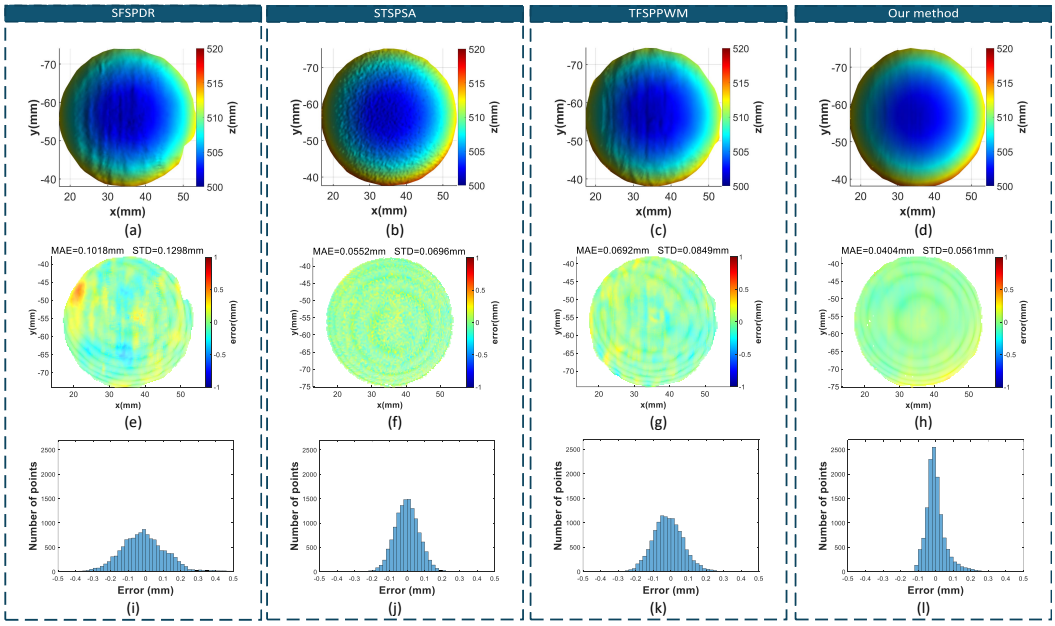


Figure 16. Quantitative analysis of a 40 mm standard sphere using different methods. (**a**) 3D reconstruction result of SFSPDR; (**b**) 3D reconstruction result of STSPSA; (**c**) 3D reconstruction result of TFSPPWM; (**d**) 3D reconstruction result of our method; (**e**) is the error distribution of (**a**); (**f**) is the error distribution of (**b**); (**g**) is the error distribution of (**c**); (**h**) is the error distribution of (**d**); (**i**) is the error statistical distribution of (**a**); (**j**) is the error statistical distribution of (**b**); (**k**) is the error statistical distribution of (**c**); (**l**) is the error.

4. Discussion and Conclusions

In this work, we propose a single-frame binary-fringe 3D measurement framework that integrates the FPP physical model with deep learning. In this method, binary speckle is embedded into PWM fringes to form SPPWM patterns. After epipolar rectification, SGM is used to obtain pixel-level coarse disparity, and two CNNs are then employed to regress the wrapped phase and foreground mask, respectively. Finally, geometric constraints between the wrapped phase and the coarse disparity are exploited for disparity refinement, thereby yielding sub-pixel disparity and eliminating phase ambiguity without phase unwrapping.

Experimental results show that 3D reconstruction can be achieved by projecting only a single 1-bit pattern: high-precision absolute 3D measurement with an accuracy of 41 μm is realized; compared with the end-to-end method SFSPDR, which lacks physical constraints, the proposed method reduces the number of outliers on the reconstructed surface (MAE decreased by approximately 60.3%); and compared with TFSPPWM, which also employs binary fringes, the proposed method achieves higher accuracy with fewer projected frames.

This study still has two limitations. First, the parameters of SGM need to be adjusted according to the scene, and parameter mismatch increases the coarse disparity error, thereby limiting the upper bound of refinement. Future work will replace SGM with learned coarse disparity estimation to achieve end-to-end coarse-to-fine two-stage prediction. Second, the method relies on the projector being in focus; defocus weakens the high-frequency features of the speckle and significantly degrades stereo matching quality. How to maintain accuracy under defocus or large depth-of-field conditions will be a key direction for the next step. In addition, further expanding scene coverage and material diversity is also an effective way to improve generalization ability.

Author Contributions: Conceptualization, Z.G. and Z.R.; methodology, Z.R.; software, Z.R.; validation, Z.L., Z.G. and Z.L.; formal analysis, Z.L.; investigation, Z.R.; resources, Z.L.; data curation, Z.R.; writing—original draft preparation, Z.R.; writing—review and editing, Z.R.; visualization, Z.R.; supervision, Z.L.; project administration, Z.G.; funding acquisition, Z.G. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially supported by the National Natural Science Foundation of China (Grant No: 12572273).

Data Availability Statement: Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Di Paola, F.; Morena, S.; Raimondi, G.; Antinozzi, S. The survey of the 'precious ones of Constance'. In Proceedings of the IMEKO TC-4, Rome, Italy, 19–21 October 2023; pp. 19–21.
2. Ding, D.; Ding, W.; Huang, R.; Fu, Y.; Xu, F. Research progress of laser triangulation on-machine measurement technology for complex surface: A review. *Measurement* **2023**, *216*, 113001. [CrossRef]
3. Catalucci, S.; Thompson, A.; Eastwood, J.; Zhang, Z.M.; Branson, D.T., III; Leach, R.; Piano, S. Smart optical coordinate and surface metrology. *Meas. Sci. Technol.* **2023**, *34*, 012001. [CrossRef]
4. Zuo, C.; Qian, J.; Feng, S.; Yin, W.; Li, Y.; Fan, P.; Han, J.; Qian, K.; Chen, Q. Deep learning in optical metrology: A review. *Light. Sci. Appl.* **2022**, *11*, 39. Correction in *Light. Sci. Appl.* **2022**, *11*, 74. [CrossRef]
5. Chen, S. Intraoral 3-D measurement by means of group coding combined with consistent enhancement for fringe projection pattern. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 5018512. [CrossRef]
6. Streli, P.; Jiang, J.; Rossie, J.; Holz, C. Structured light speckle: Joint ego-centric depth estimation and low-latency contact detection via remote vibrometry. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, San Francisco, CA, USA, 29 October–1 November 2023; pp. 1–12.
7. Staffenova, K.; Rakyta, M.; Binasova, V. The possibility of using 3D laser scanning as support for reverse engineering. *Int. Sci. J. About Simul.* **2022**, *8*, 41–45. [CrossRef]
8. Zhang, Z.; Wang, H.; Li, Y.; Li, Z.; Gui, W.; Wang, X.; Zhang, C.; Liang, X.; Li, X. Fringe-based structured-light 3D reconstruction: Principles, projection technologies, and deep learning integration. *Sensors* **2025**, *25*, 6296. [CrossRef]
9. Li, Y.; Qian, J.; Feng, S.; Chen, Q.; Zuo, C. Deep-learning-enabled dual-frequency composite fringe projection profilometry for single-shot absolute 3D shape measurement. *Opto-Electron. Adv.* **2022**, *5*, 210021. [CrossRef]
10. Lu, L.; Suresh, V.; Zheng, Y.; Wang, Y.; Xi, J.; Li, B. Motion induced error reduction methods for phase shifting profilometry: A review. *Opt. Lasers Eng.* **2021**, *141*, 106573. [CrossRef]
11. Bai, Y.; Zhang, Z.; Fu, S.; Zhao, H.; Ni, Y.; Gao, N.; Meng, Z.; Yang, Z.; Zhang, G.; Yin, W. Recent progress of full-field three-dimensional shape measurement based on phase information. *Nanomanuf. Metrol.* **2024**, *7*, 9. [CrossRef]
12. Zhang, S. Absolute phase retrieval methods for digital fringe projection profilometry: A review. *Opt. Lasers Eng.* **2018**, *107*, 28–37. [CrossRef]
13. Liu, X.; Yan, P.; Wang, Y. A robust phase extraction method for overcoming spectrum overlapping in shearography. *Light. Adv. Manuf.* **2023**, *4*, 54–65. [CrossRef]
14. Lai, X.; Xiao, Y.; Tong, S.; Zheng, S.; Chen, Z. Adaptive double sinusoidal-assisted EMD 3D shape profilometry measurement method. *Appl. Opt.* **2025**, *64*, 1901–1911. [CrossRef]
15. Wang, Z. Single-shot 3D reconstruction by fusion of Fourier transform profilometry and line clustering. *IEEE J. Sel. Top. Signal Process.* **2024**, *18*, 325–335. [CrossRef]
16. Yu, J.; Mai, S. Quasi-pixelwise motion compensation for 4-step phase-shifting profilometry based on a phase error estimation. *Opt. Express* **2022**, *30*, 19055–19068. [CrossRef]
17. Yin, W.; Chen, Q.; Feng, S.; Tao, T.; Huang, L.; Trusiak, M.; Asundi, A.; Zuo, C. Temporal phase unwrapping using deep learning. *Sci. Rep.* **2019**, *9*, 20175. [CrossRef]
18. Zheng, Y.; Duan, M.; Sun, Z.; Fan, X.; Jin, Y.; Zheng, J.; Zhu, C.; Chen, E. A universal self-correcting approach for abnormal jump errors in absolute phase retrieval. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 7006913. [CrossRef]
19. Huang, W.; Mei, X.; Wang, Y.; Fan, Z.; Chen, C.; Jiang, G. Two-dimensional phase unwrapping by a high-resolution deep learning network. *Measurement* **2022**, *200*, 111566. [CrossRef]
20. Wu, Z.; Zuo, C.; Guo, W.; Tao, T.; Zhang, Q. High-speed three-dimensional shape measurement based on cyclic complementary Gray-code light. *Opt. Express* **2019**, *27*, 1283–1297. [CrossRef]

21. Wang, L.; Cao, Y.; An, H. Gray-code fringe order jump error self-correction based on shifted phase encoding for phase measuring profilometry. *Opt. Commun.* **2022**, *524*, 128763. [CrossRef]
22. Zhang, H.; Cao, Y.; Li, H.; Xu, C.; An, H.; Wu, H.; Yang, N. Real-time computer-generated frequency-carrier Moiré profilometry with three-frequency heterodyne temporal phase unwrapping. *Opt. Laser Technol.* **2023**, *161*, 109201. [CrossRef]
23. Wei, Z.; Cao, Y.; Wu, H.; Xu, C.; Ruan, G.; Wu, F.; Li, C. Dynamic phase-differencing profilometry with number-theoretical phase unwrapping and interleaved projection. *Opt. Express* **2024**, *32*, 19578–19593. [CrossRef]
24. Li, Y.; Qian, J.; Feng, S.; Tao, T.; Chen, Q.; Zuo, C. Stereo phase unwrapping method based on feedback projection. In *Proceedings of the Optical Metrology and Inspection for Industrial Applications VII*; SPIE: Bellingham, WA, USA, 2020; Volume 11552, pp. 52–58.
25. Yin, W.; Feng, S.; Tao, T.; Huang, L.; Trusiak, M.; Chen, Q.; Zuo, C. High-speed 3D shape measurement using the optimized composite fringe patterns and stereo-assisted structured light system. *Opt. Express* **2019**, *27*, 2411–2431. [CrossRef]
26. Lohry, W.; Chen, V.; Zhang, S. Absolute three-dimensional shape measurement using coded fringe patterns without phase unwrapping or projector calibration. *Opt. Express* **2014**, *22*, 1287–1301. [CrossRef]
27. Lohry, W.; Zhang, S. High-speed absolute three-dimensional shape measurement using three binary dithered patterns. *Opt. Express* **2014**, *22*, 26752–26762. [CrossRef]
28. Li, T.; Zhou, G.; Zhang, Y.; Fan, L.; Zhang, S.; Hao, Q. Single-shot absolute 3D measurement based on speckle-embedded fringe projection. *Opt. Lasers Eng.* **2024**, *172*, 107884. [CrossRef]
29. Song, K.; Hu, S.; Wen, X.; Yan, Y. Fast 3D shape measurement using Fourier transform profilometry without phase unwrapping. *Opt. Lasers Eng.* **2016**, *84*, 74–81. [CrossRef]
30. Gai, S.; Da, F.; Dai, X. Novel 3D measurement system based on speckle and fringe pattern projection. *Opt. Express* **2016**, *24*, 17686–17697. [CrossRef]
31. Li, B.; Wang, Y.; Dai, J.; Lohry, W.; Zhang, S. Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques. *Opt. Lasers Eng.* **2014**, *54*, 236–246. [CrossRef]
32. Yuan, H.; Zeng, H.; Wang, J.; Xing, H.; Peng, R.; Yue, H.; Liu, Y. Superlarge depth range 3D measurement based on dual-focal optimization strategy. *Opt. Express* **2025**, *33*, 16041–16051. [CrossRef]
33. Ayubi, G.A.; Ayubi, J.A.; Di Martino, J.M.; Ferrari, J.A. Pulse-width modulation in defocused three-dimensional fringe projection. *Opt. Lett.* **2010**, *35*, 3682–3684. [CrossRef]
34. Zheng, Y.; Gao, Z.; Zuo, C. Complementary double pulse-width-modulation for 3D shape measurement of complex surfaces. *Opt. Laser Technol.* **2023**, *167*, 109765. [CrossRef]
35. Wang, J.; Li, Y.; Ji, Y.; Qian, J.; Che, Y.; Zuo, C.; Chen, Q.; Feng, S. Deep learning-based 3d measurements with near-infrared fringe projection. *Sensors* **2022**, *22*, 6469. [CrossRef]
36. Deng, L.; Chen, R.; Xu, Y.; Liu, W.; Guan, W.; Hu, Y.; Huang, X.; Xie, Z. Research on Single-Shot Wrapped Phase Extraction Using SEC-UNet3+. *Photonics* **2025**, *12*, 369. [CrossRef]
37. Chang, J.R.; Chen, Y.S. Pyramid stereo matching network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 5410–5418.
38. Zhang, F.; Prisacariu, V.; Yang, R.; Torr, P.H. Ga-net: Guided aggregation net for end-to-end stereo matching. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 185–194.
39. Feng, S.; Zuo, C.; Yin, W.; Gu, G.; Chen, Q. Micro deep learning profilometry for high-speed 3D surface imaging. *Opt. Lasers Eng.* **2019**, *121*, 416–427. [CrossRef]
40. Yin, W.; Zhong, J.; Feng, S.; Tao, T.; Han, J.; Huang, L.; Chen, Q.; Zuo, C. Composite deep learning framework for absolute 3D shape measurement based on single fringe phase retrieval and speckle correlation. *J. Phys. Photonics* **2020**, *2*, 045009. [CrossRef]
41. Zheng, Y.; Li, X.; Zuo, C.; Gao, Z. Three 1-bit speckle-embedded pulse-width modulation patterns for robust absolute 3D measurement. *Meas. Sci. Technol.* **2024**, *35*, 085009. [CrossRef]
42. Feng, S.; Chen, Q.; Gu, G.; Tao, T.; Zhang, L.; Hu, Y.; Yin, W.; Zuo, C. Fringe pattern analysis using deep learning. *Adv. Photonics* **2019**, *1*, 025001. [CrossRef]
43. Wan, M.; Kong, L.; Peng, X. Single-shot three-dimensional measurement by fringe analysis network. *Photonics* **2023**, *10*, 417. [CrossRef]
44. Nguyen, A.H.; Sun, B.; Li, C.Q.; Wang, Z. Different structured-light patterns in single-shot 2D-to-3D image conversion using deep learning. *Appl. Opt.* **2022**, *61*, 10105–10115. [CrossRef]
45. Yang, T.; Zhang, Z.; Li, H.; Li, X.; Zhou, X. Single-shot phase extraction for fringe projection profilometry using deep convolutional generative adversarial network. *Meas. Sci. Technol.* **2020**, *32*, 015007. [CrossRef]
46. Li, J.; Hassebrook, L.G.; Guan, C. Optimized two-frequency phase-measuring-profilometry light-sensor temporal-noise sensitivity. *J. Opt. Soc. Am. A* **2003**, *20*, 106–115. [CrossRef]
47. Zuo, C.; Chen, Q.; Feng, S.; Feng, F.; Gu, G.; Sui, X. Optimized pulse width modulation pattern strategy for three-dimensional profilometry with projector defocusing. *Appl. Opt.* **2012**, *51*, 4477–4490. [CrossRef]

48. Deng, W.; Chung, B.M. Optimal pulse width modulation technique combined with stair phase-coding method for absolute phase retrieval with projector defocusing. *Opt. Commun.* **2019**, *437*, 193–198. [CrossRef]
49. Su, Y.; Zhang, Q. Glare: A free and open-source software for generation and assessment of digital speckle pattern. *Opt. Lasers Eng.* **2022**, *148*, 106766. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.