Optics & Laser Technology 164 (2023) 109542

Contents lists available at ScienceDirect

Optics and Laser Technology

journal homepage: www.elsevier.com/locate/optlastec

Full length article

Accurate defocusing fringe projection profilometry in a large depth-of-field

Haotian Yu[1], Xiaoyu Chen[1], Yucheng Zheng, Lianfa Bai, Dongliang Zheng *, Jing Han **

School of Electronic and Optical Engineering, Nanjing University of Science and Technology, No. 200 Xiaolingwei Street, Nanjing, Jiangsu Province 210094, China
Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China

ARTICLE INFO

Keywords:
Defocusing fringe projection profilometry
Large depth-of-field measurement
Deep learning

ABSTRACT

Defocusing fringe projection profilometry (DFPP) has been one of the most popular 3-D measurement techniques. The measurement error caused by the low-contrast patterns becomes non-ignorable in large depth-of-field (DoF) DFPP. Traditional methods sacrifice the measurement speed, 3-D details or ignore the influence of the projector defocusing, which limit the performance of extending the system DoF. In this paper, a deep learning-based fringe-enhancing method (DFEM) is proposed, which transforms three patterns with different phase shifts captured at a fixed focal length into the desired phase. DFEM divides multiple sub-DoFs for reducing the difficulty of pattern transformation in the training. In the testing, the geometric constraint is introduced for determining the sub-DoF in which the object is located. DFEM can achieve accurate 3-D reconstruction in a DoF large to 1420 mm, which improve the DoF of DFPP up to 4.7 times of the traditional one. The provided experiments demonstrate the accurateness of the resulted large-DoF DFPP.

1. Introduction

Fringe projection profilometry (FPP) has been widely used for 3-D shape measurement due to its attributes of high-speed, high-precision and non-contact [1,2]. FPP requires to calculate the phase by using the transform-based [3,4] or phase-shifting algorithms [5]. The former can reconstruct the 3-D shape from a single sinusoidal fringe pattern but preserves unsatisfactory edges for objects with complex surface [6]. The latter can be used for objects with complex surface but requires at least three phase-shifted sinusoidal fringe patterns [7]. Defocusing fringe projection profilometry (DFPP) substitutes sinusoidal patterns with binary ones, which improves the 3-D sensing speed to thousands of frames per second [8].

DFPP has a relatively small depth-of-field (DoF) of 3-D reconstruction compared with traditional FPP because it generates sinusoidal patterns with projector defocusing [9]. This projector defocusing will reduce the fringe contrast, and generate phase error in the calculated phase, thus reducing the measurement DoF [10]. The light attenuation of the projector also reduces the fringe contrast when the object is located far from the ideal DoF [11]. In addition, the camera blur also reduces the fringe contrast for objects located outside its ideal DoF [12].

The measurement DoF can be enlarged by improving the fringe contrast [13] or compensating the corresponding phase error [14]. The fringe contrast can be improved by alleviating the defocusing level of the projector [15] or changing the focal length of the camera with the electrically tunable lens (ETL) [16]. The defocusing alleviation usually selects the pattern optimization method (e.g., dithering [17] or pulse width modulation [18]). But it still performs not well when exceeding the desired projector defocusing range. By using the driver to adjust the drive current, the focal length of the ETL can be changed and the focused patterns can be captured [19]. But the ETL-based method does not consider the projector defocusing [19,20]. The phase error can be compensated simply by increasing patterns [13] or spectral filtering [3], which sacrifice the measurement speed or 3-D details, respectively [21].

For some difficult 3-D reconstruction tasks, deep learning-based methods can obtain higher accuracy with fewer patterns compared with traditional methods, e.g., dynamic, motion error correction, high dynamic range, super-reconstruction and noise measurement conditions [22–29]. In this paper, a deep learning-based fringe-enhancing method (DFEM) is proposed for accurate DFPP with a large DoF. DFEM directly transforms the captured patterns into the desired phase by using a fringe-enhancing convolutional neural network (FENet). The large DoF is divided into multiple sub-DoFs to reduce the computation complexity of FENet. Each sub-DoF has an individual deep learning

* Corresponding author at: School of Electronic and Optical Engineering, Nanjing University of Science and Technology, No. 200 Xiaolingwei Street, Nanjing, Jiangsu Province 210094, China.
** Corresponding author at: School of Electronic and Optical Engineering, Nanjing University of Science and Technology, No. 200 Xiaolingwei Street, Nanjing, Jiangsu Province 210094, China.
E-mail addresses: dlzheng@njust.edu.cn (D. Zheng), eohj@njust.edu.cn (J. Han).
[1] Co-first author with equal contribution

https://doi.org/10.1016/j.optlastec.2023.109542
Received 9 December 2022; Received in revised form 25 February 2023; Accepted 26 April 2023
Available online 8 May 2023
0030-3992/© 2023 Elsevier Ltd. All rights reserved.

model. The actual sub-DoFs of the measured object can be flexibly selected by introducing geometric constraint. DFEM can achieve accurate 3-D reconstruction in a large DoF. The provided experimental results demonstrate that the proposed DEFM can improve the measurement DoF of DFPP up to 4.7 times of the traditional one.

The rest of the paper is organized as follows. Section 2 introduces the proposed DFEM. Section 3 gives experiments. Section 4 concludes this paper.

2. Principle of the proposed DFEM

2.1. Basics of DFPP

By properly defocusing the projector lens, DFPP generates the desired sinusoidal patterns by projecting 1-bit binary patterns [30]. The squared binary patterns can be expressed as [31]

$$f^p(x^p, y^p) = \begin{cases} 0, & 0 \le MOD(x^p, T) < T/2 \\ 1, & T/2 \le MOD(x^p, T) < T, \end{cases} \tag{1}$$

where (x^p, y^p) denotes the projector coordinate, T is the period of the pattern and MOD is the modulo operation. The projector defocusing can be approximated by a low-pass Gaussian filter, and by filtering the squared binary patterns, the sinusoidal fringe patterns can be generated as [22,32]

$$I_n^p(x^p, y^p) = a^p + b^p cos\left(2\pi x^p/T + \delta_n\right), n = 1, 2, \dots, N, \tag{2}$$

where a^p is the mean value, b^p is the amplitude, $\delta_n = 2\pi(n-1)/N$ is the phase shift amount and N is the number of phase steps.

The captured patterns can be expressed as [33]

$$I_n(x, y) = a(x, y) + b(x, y)cos\left[\varphi(x, y) + \delta_n\right], n = 1, 2, \dots, N, \tag{3}$$

where (x, y) denotes the camera coordinate, a denotes the background, b denotes the modulation and φ denotes the desired phase. By using a least-squares algorithm [34], φ can be calculated by

$$\varphi(x, y) = \tan^{-1}\frac{M(x, y)}{D(x, y)} = \tan^{-1}\frac{\sum_{n=1}^{N} I_n(x, y)\sin\delta_n}{\sum_{n=1}^{N} I_n(x, y)\cos\delta_n}, \tag{4}$$

where M and D represent the numerator and denominator of arctangent function respectively. Because φ is wrapped in the range of $(-\pi, \pi]$, the phase unwrapping process is required. The phase unwrapping method aims to calculate the fringe order K and obtain the absolute phase by [35]

$$\Phi(x, y) = \varphi(x, y) + 2\pi \times K(x, y). \tag{5}$$

For simplicity, we omit the notation (x, y) hereafter. The projector coordinate can be calculated from the absolute phase according to the fringe period [5]. Combining the calibrated system parameters and two corresponding 2-D coordinates of the camera and projector, the 3-D reconstruction can be achieved based on the triangulation-based method [36].

2.2. Procedure of DFEM

The proposed DFEM contains two steps of training and testing. DFEM divides the large DoF into multiple sub-DoFs. In each sub-DoF, the constructed FENet is trained to transform three phase-shifted patterns into numerator and denominator for the phase calculation. Using the trained model, DFEM can achieve accurate DFPP in the large DoF.

In large-DoF DFPP, the low-contrast patterns are caused by the compound effect of the projector and camera. For reducing the difficulty of pattern transformation, the sub-DoF is divided. For each sub-DoF, a set of proper focus lengths of the projector and camera can capture the desired sinusoidal patterns in the depth range. For the projector, the phase difference can be used to evaluate whether the focal length is proper [28]. Usually, the phase calculated by using the high-step phase-shifted method (i.e., fifteen-step) can be considered as the ideal phase [37]. The actual phase is calculated by using the three-step phase-shifted method. When the difference between the ideal phase and the actual phase is minimal, the proper focus length of the projector can be selected. When the camera is focused, the captured scene has the maximum grayscale gradient [16]. The proper focus length of the camera can be evaluated according to the grayscale gradient.

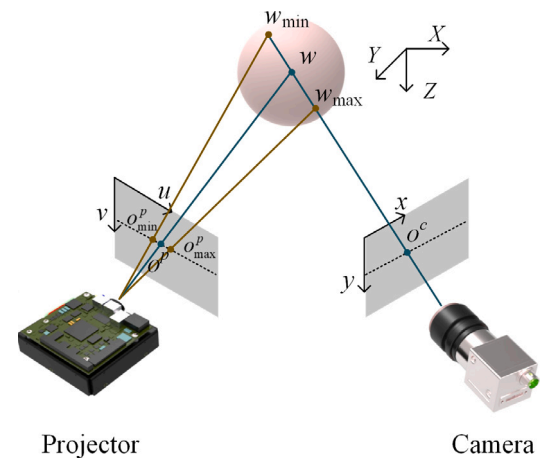

Fig. 1. The schematic diagram of the geometric constraint.

In the training, three phase-shifted patterns are captured under a fixed exposure time for each sub-DoF. Considering that the measured objects may be located under multiple positions, the fixed exposure time is required and selected manually for avoiding overexposure in the large DoF. The ground-truth can be obtained by eliminating the phase error introduced by the low-contrast patterns. For the projector, we can increase the number of phase steps (i.e., fifteen steps) firstly. Then, selecting a proper exposure time (i.e., the captured patterns have the highest grayscale without intensity saturation) can be used to improve the signal-to-noise ratio [13]. For the camera, using the ETL to adjust to a proper focal length is required [16].

In the testing, the captured patterns are inputted in the trained model for obtaining the desired phase. Because the location of the measured object is unknown, DFEM needs to locate the corresponding sub-DoF. The phase unwrapping method can provide the geometric constraint [38]. The schematic diagram of the geometric constraint is illustrated in Fig. 1. In the world coordinate system, the point w of the object reflects the projected light of the coordinate o^p onto the image point o^c. The depth range of w can be pre-determined when the correct fringe order is obtained. Based on Eq. (5), the absolute phase range $(\Phi_{min}, \Phi_{max}]$ can be determined as $(2\pi K - \pi, 2\pi K + \pi]$. From Φ_{min} and Φ_{max}, the corresponding projector coordinates o^p_{min} and o^p_{max} and 3-D point w_{min} and w_{max} can be calculated. Based on w_{min} and w_{max}, the sub-DoF can be determined.

For clarity, two white toys are selected for illustrating the flowchart of DFEM as shown in Fig. 2. The large DoF is divided into S sub-DoFs. In each sub-DoF, the network is trained, and the training step is illustrated in the first row of Fig. 2. In the training step, the object is placed in the sth sub-DoF. The ETL is installed in the camera for obtaining the ground-truth. The inputted patterns $I_n, n = 1, 2, 3$ are illustrated in first column, which are captured under the fixed drive current $i_{in} = 0$ mA of the ETL and the fixed exposure time e_{in}. In this case, the ETL is invalid and the captured pattern is the same as the one taken with the original lens. FENet outputs the patterns M^s_{out} and D^s_{out} as shown in the third column. In the fifth column, the ground-truths M^s_{gt} and D^s_{gt} are illustrated, which are obtained in advance. The corresponding fifteen-step phase-shifted patterns are captured under the current i^s_{gt} and a proper exposure time e^s_{gt} for eliminating the

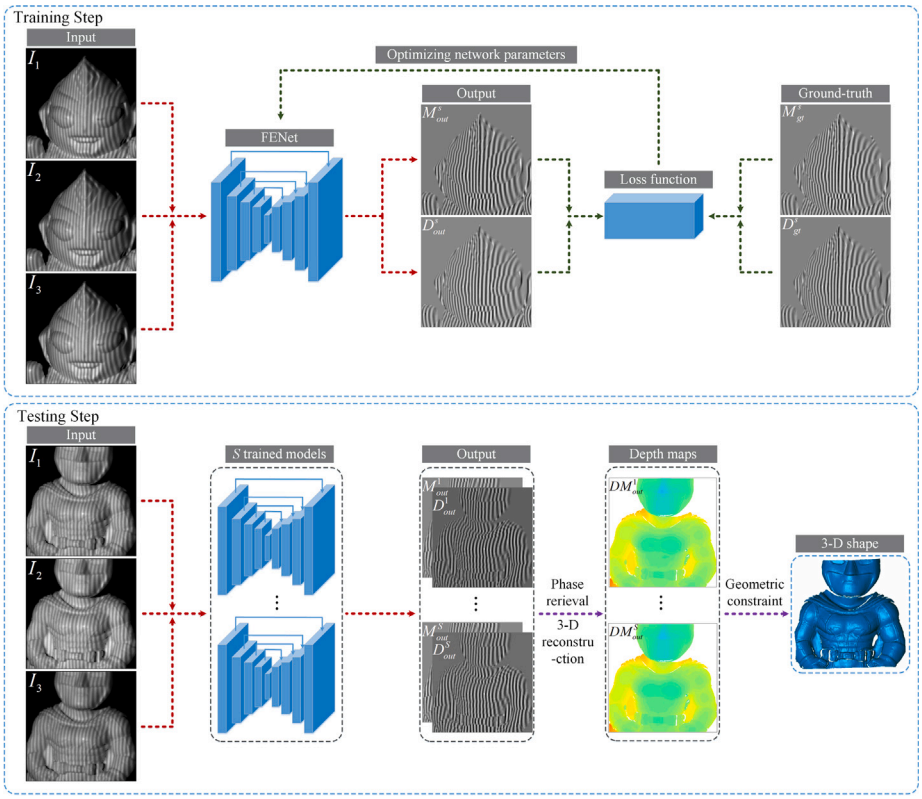

Fig. 2. Flowchart of the proposed DFEM.

phase error caused by the projector and camera. The loss function is calculated based on the difference between the output and the ground-truth. By minimizing the loss function, the network parameters can be optimized iteratively. By training the network in each sub-DoF, S models can be obtained. The loss function of FENet is expressed as

$$Loss(\theta) = \frac{1}{2p}\left(\left\|M^s_{out} - M^s_{gt}\right\|^2 + \left\|D^s_{out} - D^s_{gt}\right\|^2\right), \tag{6}$$

where θ denotes the parameter space of the network, p denotes the number of pixels and $\|\cdot\|^2$ is ℓ_2-norm.

The testing step is illustrated in the second row of Fig. 2. The ETL is removed and the measured object is placed randomly. The first column is the inputted patterns, which are captured under e_{in}. Using S models, FENet outputs S sets of patterns as shown in the third column. Based on Eq. (4), S phases can be calculated from the outputted patterns. The absolute phases can be retrieved by using the gray code-based method [39]. Combining with the calibrated parameters [40], the 3-D reconstruction can be achieved and the depth maps $DM^s_{out}, s = 1, 2, \ldots, S$ can be obtained. In the large DoF, due to the adverse effect of the system, the unwrapping errors may be generated along 2π phase jumps, which can be effectively removed by introducing a median filter [41]. The gray code-based phase retrieval method provides the geometric constraint, which allows $DM^s_{out}, s = 1, 2, \ldots, S$ to be located to the same sub-DoF. The desired model and corresponding 3-D shape can be selected.

Taking the trained patterns with the resolution of $W \times H$ as examples, the structure of FENet is illustrated in Fig. 3. FENet consists of an U-shaped structure [42,43], which improves the computing efficiency of the network and achieves better restoration of details. The U-shaped structure contains the operations: convolution (Conv) [42] and rectified linear unit (Relu) [44]; max pooling operation [45], Conv and Relu by four times; upsampling by a convolution (Up-conv) [43]; Conv, Relu and Up-conv by three times; Conv; skip connection [46]. The inputted patterns are downsampled and upsampled spatially, and the feature maps of different resolutions can be extracted, which help the network achieve accurate transformation from the input to output. Meanwhile, the structure reduces the risk of falling into local optimum. The skip connections between the encoder and the decoder are selected in the structure, which can recover the details lost during the downsampling and improve the accuracy of the outputted patterns. FENet is trained by the adaptive moment estimation optimizer [47]. We set a batch size of 4 and start with a learning rate of 0.01, which is divided by a factor of 2 each time (usually every 600 epochs) when the training error becomes stagnant.

3. Experiment

We first verify the performance of the proposed DFEM. Then we compare it with the previous method given in Ref. [20]. Based on the focal sweep technique, Ref. [20] only projects a set of patterns to achieve large-DoF DFPP using the ETL. The FPP system includes a projector (DLP6500, Texas Instruments) with a resolution of 1920 × 1080, a CMOS camera (Basler acA800-510um) with a resolution of 800 × 600, and a lens of 12 mm focal length. To fit the scale of the object, the size of the captured image is reduced to 496 × 496. The ETL (Optotune EL-16-40-TC) with a high-precision electrical lens driver (Optotune Lens driver 4i) is used to obtain the ground-truth. The camera aperture is set to be the maximum ($f/1.4$). The binary patterns with a period of 15 pixels are projected. Without ETL, the camera captures focused patterns and the projector projects desired sinusoidal patterns in the DoF of (1200 mm, 1500 mm). FENet is implemented by using Python and the framework of Pytorch on a PC with Intel Core i9-7900X CPU (3.30 GHz), 32 GB of RAM, and the GeForce GTX Titan RTX (NVIDIA).

To verify the proposed method, the large DoF of (680 mm, 2100 mm) is selected and divided into seven sub-DoF (i.e., (680 mm, 800 mm), (750 mm, 900 mm), (850 mm, 1050 mm), (1000 mm, 1250 mm), (1200 mm, 1500 mm), (1400 mm, 1750 mm), (1650 mm, 2100 mm)). By requiring the image gradient to be maximum and stable, the depth range of a sub-DoF can be determined [16]. It should be noted

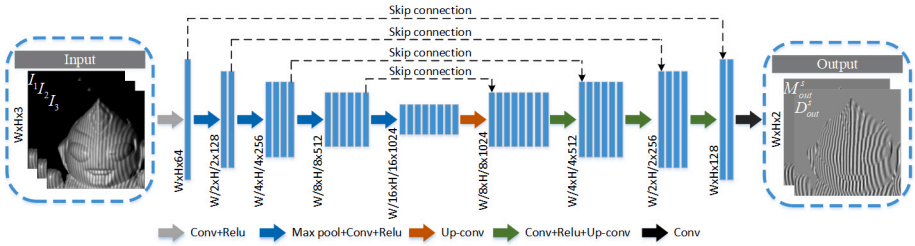

Fig. 3. Network structure of the designed FENet.

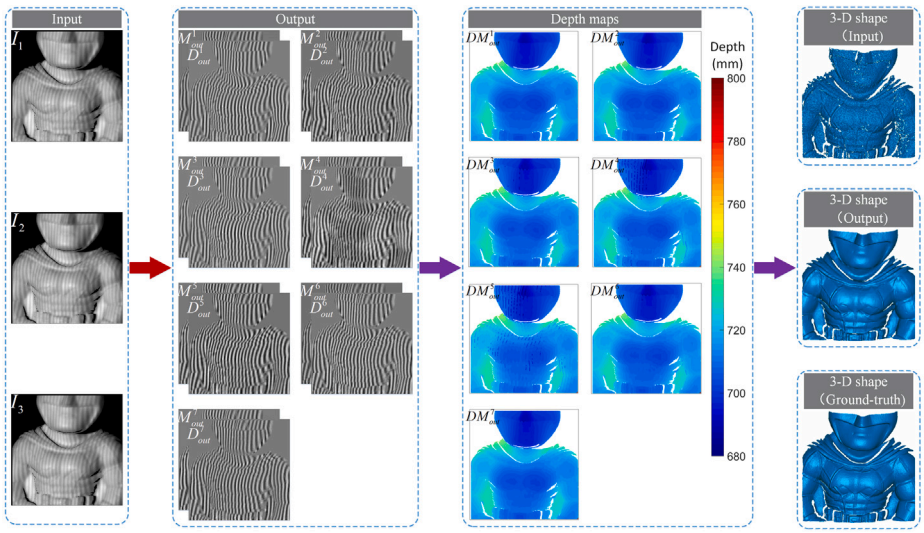

Fig. 4. Experimental results of DFEM for the toy located in the first sub-DoF.

that, two adjacent sub-DoFs should have some overlap for avoiding fusion errors caused by strict edges.

In each sub-DoF, the training set and the validation set contain 50 samples and 20 samples, respectively, which are captured by using different objects. For each scene, the inputted three patterns with respective phase shift amounts 0, $2/3\pi$ and $4/3\pi$ are captured under $i_{in} = 0$ mA and $e_{in} = 3.9$ ms. The ground-truth are calculated by using fifteen phase-shifted patterns, which are captured under i_{gt} and e_{gt}. In each sub-DoF, i_{gt} are set as 33.40 mA, 22.96 mA, 15.45 mA, 7.15 mA, 0 mA, −15.45 mA and −22.60 mA, and e_{gt} are set as 3.9 ms, 5.5 ms, 6.5 ms, 8.5 ms, 10 ms, 14 ms and 17 ms. It should be noted that, the captured patterns are required not to be overexposed when the object is located in the first sub-DoF. Meanwhile, the phase optimization ability of DFEM is limited, and the fringe modulation should not be too low for the distant object [14].

The testing set contains 180 scenes, where 30 different objects are placed in the large DoF randomly. The low-contrast patterns can be caused by the projector and the camera. Thus, the influence of different equipment on the measurement results is further demonstrated. The methods of "EP" and "EC" are constructed. The first one requires to capture three phase-shifted patterns under i_{gt} and e_{in}, which contains the error introduced by the projector. The second one requires to capture fifteen phase-shifted patterns under e_{gt} without ETL, which contains the error introduced by the camera. Correspondingly, the method only uses the inputted patterns to be described as "EP + EC". We emphasize that all the experimental results are obtained from the testing set.

In our datasets, the captured samples are independent and identically distributed, which can obtain similar results in all datasets. Thus, it can prove that our datasets contain sufficient samples and have extensively covered the sample space in our experiment environment. It should be noted that, the proposed method is suitable for the patterns with different resolutions and periods. When the system is changed, the trained models may not obtain the desired results because of the data dependent. At this time, additional samples are necessary for re-training.

3.1. Verification of DFEM

We select a white toy to verify the proposed DFEM. The experimental results are illustrated in Fig. 4, where the white toy is located in the first sub-DoF. The same white toy is also placed in other sub-DoFs and the results are provided in Visualization 1.

In Fig. 4, the first column shows the inputted patterns $I_n, n = 1, 2, 3$. The second column shows the outputted patterns M^s_{out} and $D^s_{out}, s = 1, 2, \ldots, 7$. The third column shows the calculated depth maps $DM^s_{out}, s = 1, 2, \ldots, 7$. The fourth column shows the 3-D shapes calculated by using the inputted patterns, outputted patterns and ground-truth, respectively. FENet transforms three low-contrast patterns into seven sets of outputted patterns using seven trained models, which can be used to calculate seven depth maps. The gray code-based method provides the geometric constraint, and $DM^s_{out}, s = 1, 2, \ldots, 7$ are located in the same depth range. Thus, the white toy can be determined in the first sub-DOF, and we can select M^1_{out} and D^1_{out} to achieve accurate 3-D reconstruction. Compared with the 3-D shape calculated by using the inputted patterns, DFEM can obtain the more smooth result, which is more similar with the ground-truth.

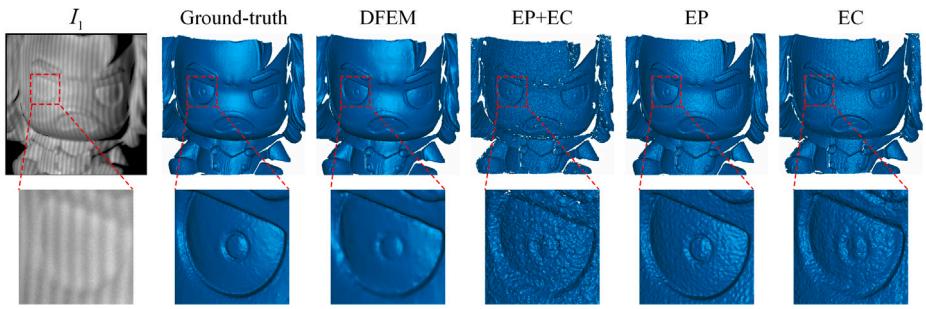

Fig. 5. 3-D shapes of the toy located in the first sub-DoF.

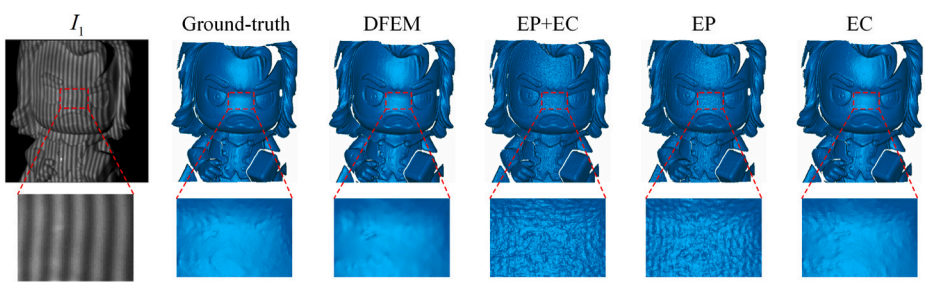

Fig. 6. 3-D shapes of the toy located in the third sub-DoF.

3.2. Performance of DFEM

Another white toy is selected and the influence of low-contrast patterns is illustrated. The white toy is placed in each sub-DoF. The reconstructed 3-D shapes are shown in Figs. 5–7, where the toy is located in the first, third and sixth sub-DoFs, respectively. Each column represents the inputted pattern I_1, the results of ground-truth, DFEM, EP + EC, EP and EC, respectively. For clarity, the results of the white toy placed in other sub-DoFs are provided in Visualization 2.

In Fig. 5, DFEM performs better than the other methods in complex and smooth textures, which preserves the details and obtains smooth and similar 3-D shape with the ground-truth. The measurement error is reduced obviously compared with EP + EC. Due to the improper defocusing level, the modulation is reduced and the phase error introduced by the random noise is increased [14]. It causes the speckles in smooth textures for the 3-D shape calculated by EP. For the 3-D shape calculated by EC, the serious camera defocusing causes the blurred texture around the eyeball and eye socket as shown in the enlarged area.

In Fig. 6, DFEM also can obtain similar result with the ground-truth, which performs better than EP + EC and EP. The random noise still prevents EP + EC and EP from obtaining the smooth 3-D shape, as shown in the enlarged area. The camera defocusing is slight and does not influence the measurement result obviously. Thus, the 3-D shape calculated by EC is similar with DFEM.

In Fig. 7, DFEM also performs better than the other methods, which can recover the destroyed details and smooth the surface compared with EP + EC. The 3-D shape calculated by EP contains the wrinkles and speckles as shown in the enlarged area. The wrinkles are caused by the improper defocusing level and the speckles are caused by the light attenuation of the projector. The 3-D shape calculated by EC contains the blurred texture caused by the camera defocusing, as shown in the enlarged area. The proposed DFEM can be used for the object located in different sub-DoF, and obtain the desired results.

For illustrating the performance further, a sphere with the diameter of 50.8 mm is measured, and the mean measurement errors are provided in Table 1. Compared with EP + EC, the proposed DFEM can reduce the measurement error obviously, and the accuracy is similar with that of the ground-truth. The measurement results of EP are influenced by the composite of the improper defocusing level and the light attenuation of the projector. For EC, in the first, sixth and seventh sub-DOF, due to serious camera defocusing, the measurement accuracy is significantly reduced.

The proposed DFEM also can be used for the scene with multiple objects. These objects are placed in different sub-DoFs. The results are shown in Fig. 8. The objects of Fig. 8(a)–(c) are placed in the first and seventh sub-DoFs, the second and sixth sub-DoFs, and the third and fifth sub-DoFs, respectively. The first, second, third and fourth columns represent I_1, the 3-D shape obtained by the ground-truth, DFEM and EP + EC, respectively. DFEM separates multiple isolated objects and fuses the corresponding 3-D shapes based on the geometric constraint. DFEM can reduce the measurement error of EP + EC and obtain desired 3-D shape, which is similar to the scene containing a single object.

For a single object placed in two adjacent sub-DoFs, we can set the weights according to the optimized depth. The weight is lower when the depth is closer to the edge of the sub-DoF. Based on the normalized weights, two sets of 3-D data are summed and fused. Because two adjacent sub-DoFs contain the common depth range, the fused 3-D data by weights can be transited smoothly and avoid the measurement error introduced by the edges of sub-DoFs. Taking a white toy placed in first and second sub-DoFs as an example, the results are shown in Fig. 9. The common range of two adjacent sub-DoFs is enlarged in the red box. It can be found that the edges of two adjacent sub-DoFs do not introduce measurement errors.

3.3. Comparison between DFEM and the previous method

We compare the proposed DFEM with the previous method [20] under different scenes. First, the measurement accuracy is compared

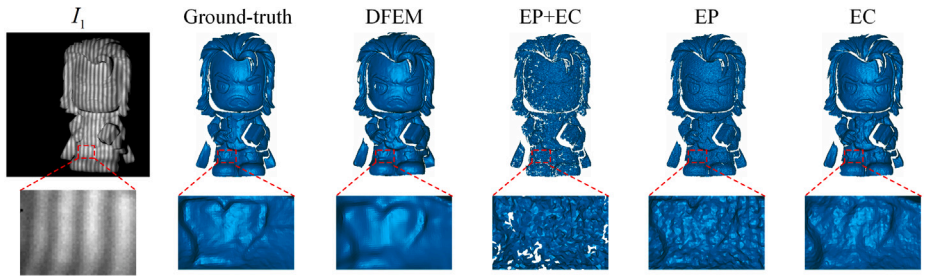

Fig. 7. 3-D shapes of the toy located in the sixth sub-DoF.

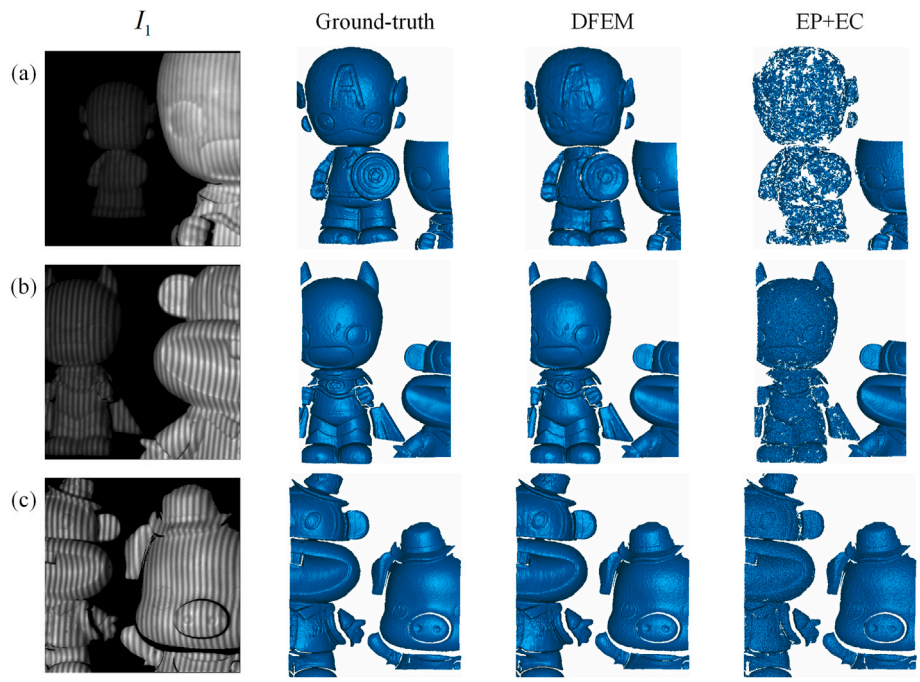

Fig. 8. 3-D reconstruction for the scene with multiple objects. The objects are placed in (a) the first and seventh sub-DoFs, (b) the second and sixth sub-DoFs, (c) the third and fifth sub-DoFs.

Table 1
Measurement error in each sub-DoF.

Serial number of sub-DoF	First	Second	Third	Fourth	Fifth	Sixth	Seventh
Error of ground-truth (mm)	0.037	0.031	0.040	0.025	0.049	0.078	0.131
Error of DFEM (mm)	0.040	0.027	0.045	0.033	0.055	0.085	0.143
Error of EP + EC (mm)	0.109	0.077	0.080	0.075	0.130	0.389	0.724
Error of EP (mm)	0.048	0.054	0.068	0.076	0.130	0.188	0.286
Error of EC (mm)	0.058	0.041	0.041	0.045	0.049	0.120	0.238

by using a sphere with the diameter of 50.8 mm. The sphere is placed in all sub-DOFs. In the first to fifth sub-DoFs, the previous method can reduce the measurement error caused by the camera defocusing, and the deconvolution algorithm can smooth the 3-D shape. Both the two methods have similar accuracy. In the sixth and seventh sub-DoFs, the light attenuation of the projector causes serious error, which is not considered in the previous method. Thus, the measurement accuracy of DFEM is 0.05 mm and 0.1 mm higher than it, respectively.

Then, the reconstructed 3-D shapes of the white toys are compared. In Figs. 10 and 11, two white toys are selected and placed in the first and sixth sub-DoFs, respectively. Each column represents the 3-D shape obtained by the ground-truth, DFEM, previous method and EP+EC, respectively. In Fig. 10, the previous method can recover the details by reducing the influence of the camera defocusing. But the speckles caused by the projector cannot be effectively removed as shown in the enlarged area. In Fig. 11, the influence of the projector causes serious measurement error. The previous method cannot recover the details as shown in the enlarged area. DFEM can obtain desired 3-D shapes, which are similar to the ground-truth. For the object placed in the seventh sub-DoF, the results are similar to Fig. 11. In other sub-DoFs, the performance of the two methods is similar.

4. Conclusions

In this paper, a deep learning-based fringe-enhancing method (DFEM) is proposed by designing a fringe-enhancing convolutional neural network (FENet). Three patterns with different phase shifts are captured at a fixed focal length, which contain the phase error

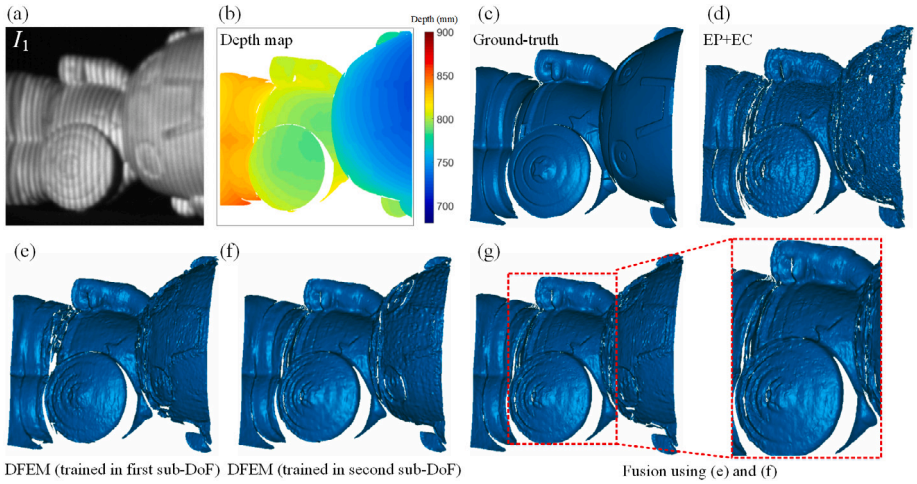

Fig. 9. The performance for a single object placed in first and second sub-DoFs. (a) The inputted pattern I_1, (b) the depth map, (c)–(f) the 3-D shapes obtained by the ground-truth, EP + EC, DFEM (trained in first sub-DoF) and DFEM (trained in second sub-DoF), respectively, (g) the fused 3-D shape using (e) and (f).

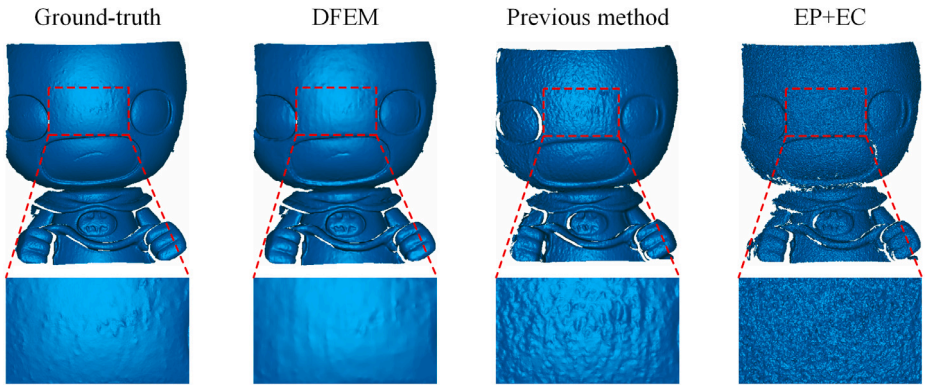

Fig. 10. 3-D shapes of the toy located in the first sub-DoF for comparison.

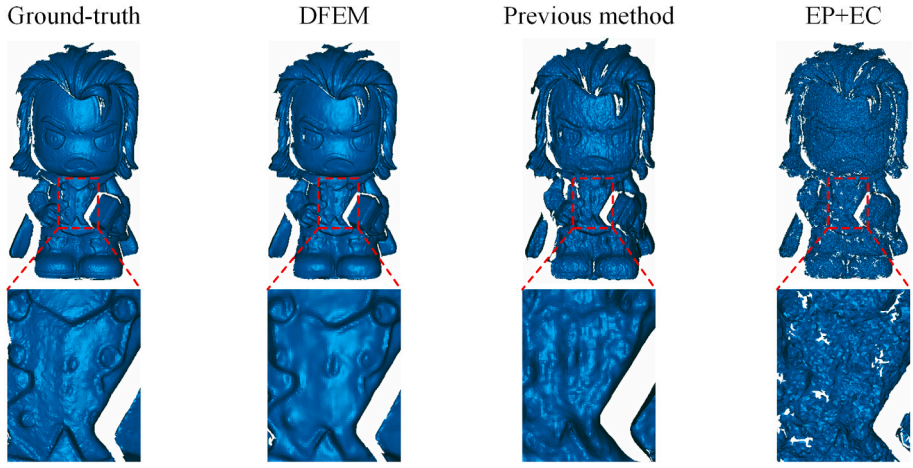

Fig. 11. 3-D shapes of the toy located in the sixth sub-DoF for comparison.

introduced by the low fringe contrast. These patterns are transformed into the desired phase by using FENet. DFEM takes full account of the influence of the projector and the camera on the measurement results. DFEM can obviously reduce the measurement error and recover the 3-D details. It is verified on different scenes located in the large DoF. DFEM does not require additional patterns without sacrificing the measurement speed. It has the potential for online processing by selecting a higher performance GPU and optimizing the network structure. Compared with the previous method, DFEM can further improve the measurement DoF of DFPP.

CRediT authorship contribution statement

Haotian Yu: Conceptualization, Methodology, Writing – original draft. **Xiaoyu Chen:** Software, Formal analysis. **Yucheng Zheng:** Visualization, Investigation. **Lianfa Bai:** Supervision. **Dongliang Zheng:** Writing – review & editing. **Jing Han:** Validation, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request

Acknowledgments

This work was sponsored by the National Natural Science Foundation of China (62031018, 62201261), Jiangsu Provincial Key Research and Development Program, China (BE2022391), China Postdoctoral Science Foundation, China (2022M721620) and Jiangsu Funding Program for Excellent Postdoctoral Talent (2022ZB256).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.optlastec.2023.109542.

References

[1] J.-S. Hyun, S. Zhang, High-speed three-dimensional absolute shape measurement with three projected binary patterns, Opt. Eng. 59 (2) (2020) 024104.
[2] L. Lu, Y. Ding, Y. Luan, Y. Yin, Q. Liu, J. Xi, Automated approach for the surface profile measurement of moving objects based on PSP, Opt. Express 25 (25) (2017) 32120–32131.
[3] Q. Kemao, Two-dimensional windowed Fourier transform for fringe pattern analysis: principles, applications and implementations, Opt. Lasers Eng. 45 (2) (2007) 304–317.
[4] X. Su, W. Chen, Fourier transform profilometry:: a review, Opt. Lasers Eng. 35 (5) (2001) 263–284.
[5] C. Lin, D. Zheng, Q. Kemao, J. Han, L. Bai, Spatial pattern-shifting method for complete two-wavelength fringe projection profilometry, Opt. Lett. 45 (11) (2020) 3115–3118.
[6] S. Zhang, Absolute phase retrieval methods for digital fringe projection profilometry: A review, Opt. Lasers Eng. 107 (2018) 28–37.
[7] D. Zheng, Q. Kemao, J. Han, J. Wang, H. Yu, L. Bai, High-speed phase-shifting profilometry under fluorescent light, Opt. Lasers Eng. 128 (2020) 106033.
[8] W. Yin, C. Zuo, S. Feng, T. Tao, Y. Hu, L. Huang, J. Ma, Q. Chen, High-speed three-dimensional shape measurement using geometry-constraint-based number-theoretical phase unwrapping, Opt. Lasers Eng. 115 (2019) 21–31.
[9] B. Li, Y. Wang, J. Dai, W. Lohry, S. Zhang, Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques, Opt. Lasers Eng. 54 (2014) 236–246.
[10] J. Zhang, B. Luo, X. Su, Y. Wang, X. Chen, Y. Wang, Depth range enhancement of binary defocusing technique based on multi-frequency phase merging, Opt. Express 27 (25) (2019) 36717–36730.
[11] H. Jiang, H. Zhao, X. Li, High dynamic range fringe acquisition: a novel 3-D scanning technique for high-reflective surfaces, Opt. Lasers Eng. 50 (10) (2012) 1484–1493.
[12] M. Zhong, X. Hu, F. Chen, C. Xiao, D. Peng, S. Zhang, Autofocusing method for a digital fringe projection system with dual projectors, Opt. Express 28 (9) (2020) 12609–12620.
[13] C. Zuo, S. Feng, L. Huang, T. Tao, W. Yin, Q. Chen, Phase shifting algorithms for fringe projection profilometry: A review, Opt. Lasers Eng. 109 (2018) 23–59.
[14] H. Yu, D. Zheng, J. Fu, Y. Zhang, C. Zuo, J. Han, Deep learning-based fringe modulation-enhancing method for accurate fringe projection profilometry, Opt. Express 28 (15) (2020) 21692–21703.
[15] J.-S. Hyun, B. Li, S. Zhang, High-speed high-accuracy three-dimensional shape measurement using digital binary defocusing method versus sinusoidal method, Opt. Eng. 56 (7) (2017) 074102.
[16] X. Hu, G. Wang, J.-S. Hyun, Y. Zhang, H. Yang, S. Zhang, Autofocusing method for high-resolution three-dimensional profilometry, Opt. Lett. 45 (2) (2020) 375–378.
[17] Y. Wang, S. Zhang, Three-dimensional shape measurement with binary dithered patterns, Appl. Opt. 51 (27) (2012) 6631–6636.
[18] C. Zuo, Q. Chen, S. Feng, F. Feng, G. Gu, X. Sui, Optimized pulse width modulation pattern strategy for three-dimensional profilometry with projector defocusing, Appl. Opt. 51 (19) (2012) 4477–4490.
[19] X. Hu, G. Wang, Y. Zhang, H. Yang, S. Zhang, Large depth-of-field 3D shape measurement using an electrically tunable lens, Opt. Express 27 (21) (2019) 29697–29709.
[20] X. Hu, S. Zhang, Y. Zhang, Y. Liu, G. Wang, Large depth-of-field three-dimensional shape measurement with the focal sweep technique, Opt. Express 28 (21) (2020) 31197–31208.
[21] B. Lin, S. Fu, C. Zhang, F. Wang, Y. Li, Optical fringe patterns filtering based on multi-stage convolution neural network, Opt. Lasers Eng. 126 (2020) 105853.
[22] S. Feng, C. Zuo, W. Yin, G. Gu, Q. Chen, Micro deep learning profilometry for high-speed 3D surface imaging, Opt. Lasers Eng. 121 (2019) 416–427.
[23] H. Nguyen, N. Dunne, H. Li, Y. Wang, Z. Wang, Real-time 3D shape measurement using 3LCD projection and deep machine learning, Appl. Opt. 58 (26) (2019) 7100–7109.
[24] Y. Zheng, S. Wang, Q. Li, B. Li, Fringe projection profilometry by conducting deep learning from its digital twin, Opt. Express 28 (24) (2020) 36568–36583.
[25] J. Tan, W. Su, Z. He, N. Huang, J. Di, L. Zhong, Y. Bai, B. Dong, S. Xie, Deep learning-based method for non-uniform motion-induced error reduction in dynamic microscopic 3D shape measurement, Opt. Express 30 (14) (2022) 24245–24260.
[26] Y. Li, W. Guo, J. Shen, Z. Wu, Q. Zhang, Motion-induced phase error compensation using three-stream neural networks, Appl. Sci. 12 (16) (2022) 8114.
[27] L. Zhang, Q. Chen, C. Zuo, S. Feng, High-speed high dynamic range 3D shape measurement based on deep learning, Opt. Lasers Eng. 134 (2020) 106245.
[28] Y. Zhao, R. Zhu, K. Zhang, H. Yu, L. Bai, D. Zheng, J. Han, Accurate dynamic 3-D shape measurement based on the fringe pattern super-reconstruction technique, Measurement 200 (2022) 111575.
[29] Y. Li, J. Shen, Z. Wu, Q. Zhang, Passive binary defocusing for large depth 3D measurement based on deep learning, Appl. Opt. 60 (24) (2021) 7243–7253.
[30] S. Zhang, Flexible 3D shape measurement using projector defocusing: extended measurement range, Opt. Lett. 35 (7) (2010) 934–936.
[31] X. Kang, Z. Yin, S. Dong, C. Liu, X. He, Evaluating binary defocusing quantitatively in real-time for fringe projection profilometry, Opt. Eng. 60 (6) (2021) 064110.
[32] D. Zheng, F. Da, H. Huang, Phase unwrapping for fringe projection three-dimensional measurement with projector defocusing, Opt. Eng. 55 (3) (2016) 034107.
[33] D. Zheng, Q. Kemao, F. Da, H.S. Seah, Ternary Gray code-based phase unwrapping for 3D measurement using binary patterns with projector defocusing, Appl. Opt. 56 (13) (2017) 3660–3665.
[34] W. Yin, S. Feng, T. Tao, L. Huang, M. Trusiak, Q. Chen, C. Zuo, High-speed 3D shape measurement using the optimized composite fringe patterns and stereo-assisted structured light system, Opt. Express 27 (3) (2019) 2411–2431.
[35] H. Yu, B. Han, L. Bai, D. Zheng, J. Han, Untrained deep learning-based fringe projection profilometry, APL Photonics 7 (1) (2022) 016102.
[36] T. Tao, Q. Chen, S. Feng, J. Qian, Y. Hu, L. Huang, C. Zuo, High-speed real-time 3D shape measurement based on adaptive depth constraint, Opt. Express 26 (17) (2018) 22440–22456.
[37] H. Yu, X. Chen, Z. Zhang, C. Zuo, Y. Zhang, D. Zheng, J. Han, Dynamic 3-D measurement based on fringe-to-fringe transformation using deep learning, Opt. Express 28 (7) (2020) 9405–9418.
[38] D. Zheng, F. Da, Self-correction phase unwrapping method based on Gray-code light, Opt. Lasers Eng. 50 (8) (2012) 1130–1139.
[39] Z. Wu, W. Guo, Y. Li, Y. Liu, Q. Zhang, High-speed and high-efficiency three-dimensional shape measurement based on Gray-coded light, Photonics Res. 8 (6) (2020) 819–829.
[40] K. Liu, Y. Wang, D.L. Lau, Q. Hao, L.G. Hassebrook, Dual-frequency pattern scheme for high-speed 3-D shape measurement, Opt. Express 18 (5) (2010) 5229–5244.
[41] D. Zheng, F. Da, Q. Kemao, H.S. Seah, Phase-shifting profilometry combined with gray-code patterns projection: unwrapping error removal by an adaptive median filter, Opt. Express 25 (5) (2017) 4700–4713.
[42] E. Romera, J.M. Alvarez, L.M. Bergasa, R. Arroyo, Erfnet: Efficient residual factorized convnet for real-time semantic segmentation, IEEE Trans. Intell. Transp. Syst. 19 (1) (2017) 263–272.
[43] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.

[44] V. Nair, G.E. Hinton, Rectified linear units improve restricted boltzmann machines, in: Icml, 2010.
[45] G. Spoorthi, S. Gorthi, R.K.S.S. Gorthi, PhaseNet: A deep convolutional neural network for two-dimensional phase unwrapping, IEEE Signal Process. Lett. 26 (1) (2018) 54–58.
[46] K. Wang, Y. Li, Q. Kemao, J. Di, J. Zhao, One-step robust deep learning phase unwrapping, Opt. Express 27 (10) (2019) 15100–15115.
[47] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, 2014, arXiv preprint arXiv:1412.6980.