Results in Physics 41 (2022) 105904

Contents lists available at ScienceDirect

Results in Physics

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

Invalid point removal method based on error energy function in fringe projection profilometry

Kaifeng Zhu[a,b], Xin He[a,*], Yi Gao[c], Ruidong Hao[a,b], Zhonghui Wei[a], Bing Long[d], Zhiya Mu[a], Jun Wang[a,*]

[a] *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China*
[b] *University of Chinese Academy of Sciences, Beijing 100049, China*
[c] *College of Criminal Science and Technology, Criminal Investigation Police University of China, Shenyang 110035,China*
[d] *Key Laboratory of Criminal Inspection of Sichuan Province, Sichuan Police College, Luzhou 646000,China*

ARTICLE INFO

Keywords:
3D reconstruction
Invalid points removal
Fringe analysis
Error energy
Image segmentation

ABSTRACT

Invalid points, such as shadow and background, in the captured fringe patterns of fringe projection profilometry (FPP) are often inevitable due to the limited field of view measurements of three-dimensional (3D) imaging equipment. To ensure the quality of 3D reconstruction data, these invalid points must be identified and removed. FPP captures co-frequency-based fringe pattern sequences and approximately distributes this data along an ideal cosine curve. We propose an invalid points removal method based on an error energy function. By analyzing the relationship between the pixel values of a series of captured co-frequency patterns and an ideal cosine curve, we quantize the error energy by using a Gaussian weighted Euclidean distance. An improved Gaussian filtering method based on modulation intensity is used to significantly increase the error energy difference between the valid points of the target and the invalid points of the shadow/background area. Finally, the points whose error energy is greater than a certain error threshold are removed as invalid points. Experimental results indicate that this method can efficiently remove invalid points in fringe patterns and performs better than existing traditional methods.

Introduction

Fringe projection profilometry (FPP) [1–7] has become the most widely used 3D vision measurement technology because of its simplicity of operation, fast measurement speed, high accuracy and applicability to weakly textured targets. However, due to several limitations, like optical imaging mechanisms of projectors and cameras, and the influence of imaging noise, the captured fringe patterns inevitably have invalid points, such as shadow and background, which seriously affects the post-processing of reconstructed point clouds. In recent years, researchers have proposed some methods to identify and remove invalid points in fringe projection profilometry. The existing invalid point removal methods are mainly divided into two categories: unwrapping phase characteristics [8–12] and threshold design based on modulation [13–18].

Song et al and Zhang et al. [8,9] removed noisy points by unwrapping phase monotonicity, but this method may eliminate some valid points and was not suitable for some discontinuous objects. Chen et al. [10] identified invalid points by comparing the least-squares fitting of unwrapping phase with different fringe frequencies with a set threshold, but some of the invalid points also have small least-squares fitting errors. Similarly, Huang et al. [11] identified invalid points by a proper reliability graph R and the root mean square error (RMSE) in the temporal phase unwrapping method. After comparing two modulation maps, Feng et al. [12] used the phase relationship of adjacent pixels and Gaussian filter to identify and remove invalid points.

Su et al. [13] removed invalid points via the histogram of modulation intensity, but a proper threshold is difficult to determine directly. Lau et al. [14] applied the discrete Fourier transform to the captured fringe sequence in the time axis direction, and took the pixels with sufficiently large modulations and sufficiently small first harmonic values as valid points. Zhang et al. [15] and Du et al. [16] respectively employed a new weighting factor to improve the Otsu method, and used the modulation histogram to calculate the threshold. The modulation value, which is less than the threshold was regarded as background and shadow pixels, but the threshold is biased to the category with larger variance due to

---

* Corresponding authors.
E-mail addresses: hexin6627@sohu.com (X. He), zif1005@163.com (J. Wang).

https://doi.org/10.1016/j.rinp.2022.105904
Received 31 May 2022; Received in revised form 2 August 2022; Accepted 13 August 2022

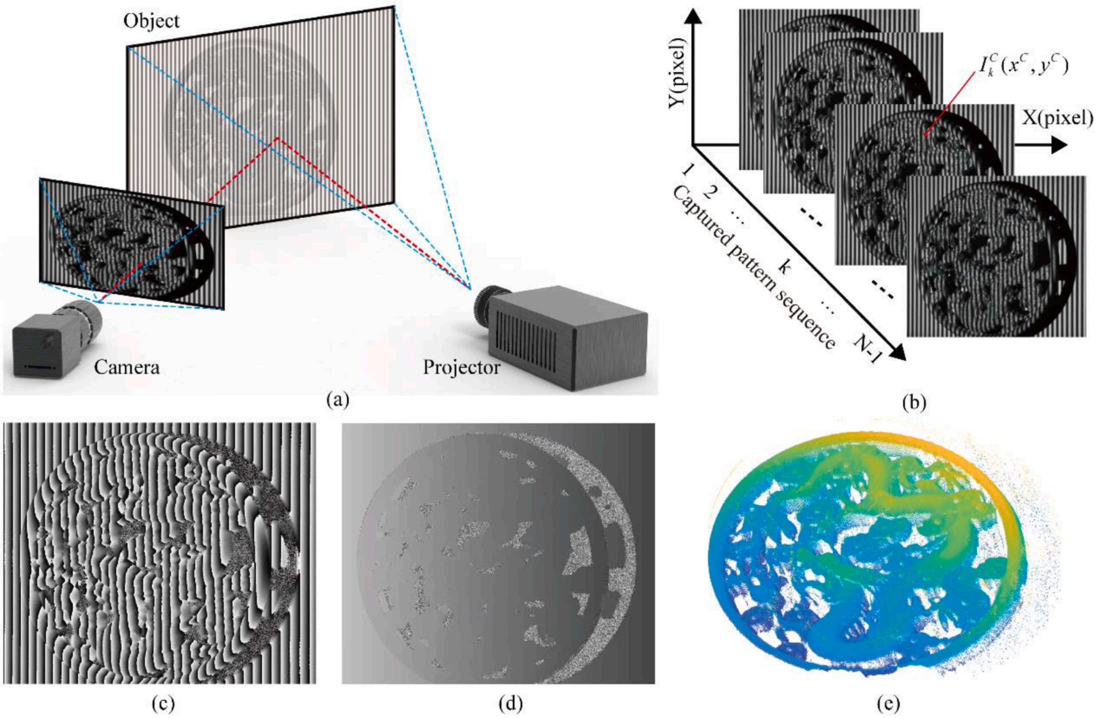


Fig. 1. Monocular structured light measurement system: (a) monocular profilometry principle; (b) initial images; (c) wrapped phase; (d) unwrapped phase; (e) retrieved 3D result.

the influence of intra-class variance of the measured object. Lu et al. [17] established the mapping relationship between the reconstructed 3D data and the digital micro-mirror device (DMD) of the projector and identified the invalid points by judging the coordinates of the mapping points on the DMD plane and the corresponding data modulation. Qi et al. [18] established an epipolar constraint between the camera image plane and the projector image plane and took the distance between the projector image coordinate and the epipolar line of a pixel as the criterion to remove invalid points. However, these two methods [17,18] need to calibrate the relationship between the projector image plane and the camera image plane in advance; additionally, the operation is more complex.

Researchers are also actively looking for non-traditional alternatives to prevent shadows/backgrounds caused by fringe patterns. For example, Skydan et al. [19] used multiple projectors to find shadowless surfaces from different angles. Du et al. [20] used robotic arms to obtain three-dimensional data of objects from different views. However, the additional hardware/equipment increases the cost, and for objects with complex surfaces, it is difficult to eliminate invalid points.

Qi et al. [18] believe that thresholding is an accurate and simple method to remove invalid points without needing additional fringe frequency. The method uses a proper threshold to treat pixels with low modulation as invalid points. Le et al. [21] pointed out that in optical measurements, there are complex optical phenomena such as diffraction or interference at the edge of an object, and the intensity modulations at the edge often changes greatly. In many of the existing methods, when the intensity modulations of the object surface are small or generated by low-quality fringe patterns, some valid modulations and invalid modulations will be intermixed with each other, so it is difficult to achieve the ideal segmentation effect via a single modulation threshold alone.

In view of the shortcomings of traditional methods for identifying and removing invalid points in fringe patterns, we propose an invalid points removal method using the error energy function in fringe projection profilometry. First, we use a Gaussian weighted Euclidean distance formula to obtain the error energy between the pixel values of the captured co-frequency pattern sequence and the ideal cosine curve. The error energy difference between the invalid points of the shadow/the background area and the target's valid points is not obvious. Thus, an improved Gaussian filtering method based on intensity modulation is used to increase the error energy difference. The points whose error energy is greater than a certain error threshold are removed as invalid points. Because this method uses the error energy of a set of captured co-frequency fringe patterns as the invalidation criterion, the method is not sensitive to object reflectivity or modulation; therefore, it can be used to remove invalid points of complex objects in multiple scenes. The experimental results indicate that this method can effectively remove the invalid points in the fringe patterns, and the performance is better than the existing methods.

The rest of this article is arranged as follows. Section 2 describes the principle of the proposed method. The experimental results are discussed in section 3. Section 4 is a summary of the work.

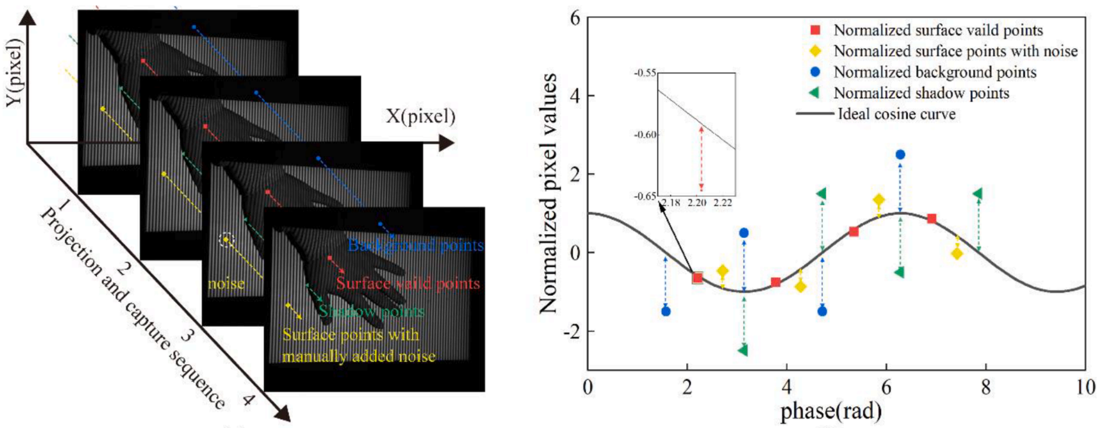
Principle

Principle of fringe projection profilometry

Fig. 1 depicts the workflow of the FPP measurement system. First, a series of standard cosine fringe images are projected to the measured object. The wrapped phases are then obtained using fringe analysis techniques [22] and the unwrapped phase is obtained by the phase unwrapping methods [23]. Finally, the unwrapped phase is converted to three-dimensional coordinates [24]. For the N-step phase shifting method, the standard cosine fringe patterns projected by the projector can be expressed as:

$$I_k^P(x^P, y^P) = A^P(x^P, y^P) + B^P(x^P, y^P)cos[\varphi^P(x^P, y^P) + 2k\pi/N]. \tag{1}$$

where (x^P, y^P) is the coordinate of arbitrary point in the projector plane, $\varphi^P(x^P, y^P)$ is the phase, $I_k^P(x^P, y^P)$ is the intensity, $A^P(x^P, y^P)$ is the background intensity, $B^P(x^P, y^P)$ is the amplitude, and N is the total number of phase shifts,$k = 0, 1, \ldots, N-1$. In Fig. 1(a), the digital light processing (DLP) projector projects the fringe patterns of standard cosine



(a) (b)

Fig. 2. Common types of a set of sequence points on the captured fringe images: (a) a series of deformed fringe images obtained by 4-step phase shift; (b) the normalized distribution of four common sequence point types from (a) relative to the ideal cosine curve.

distribution onto the object surface, and the height of the object surface modulates the cosine fringes. The deformation of the light fringes captured by the camera can be expressed as:

$$I_k^C(x^C,y^C) = A^C(x^C,y^C) + B^C(x^C,y^C)cos[\varphi^C(x^C,y^C) + 2k\pi/N] \tag{2}$$

where $I_k^C(x^C,y^C)$ refers to the intensity value of arbitrary coordinate point (x^C,y^C) in the k-th fringe pattern as shown in Fig. 1(b), $\varphi^C(x^C,y^C)$ denotes the desired wrapped phase, and $A^C(x^C,y^C)$ is the background intensity and $B^C(x^C,y^C)$ is the intensity modulation. The quantities can be expressed as

$$\varphi^C(x^C,y^C) = \arctan\frac{\sum_{k=0}^{N-1} I_k^C(x^C,y^C)\sin(2k\pi/N)}{\sum_{k=0}^{N-1} I_k^C(x^C,y^C)\cos(2k\pi/N)} \tag{3}$$

$$A^C(x^C,y^C) = \frac{\sum_{k=0}^{N-1} I_k^C(x^C,y^C)}{N} \tag{4}$$

$$B^C(x^C,y^C) = \frac{2}{N}\sqrt{\left[\sum_{k=0}^{N-1} I_k^C(x^C,y^C)\sin(2k\pi/N)\right]^2 + \left[\sum_{k=0}^{N-1} I_k^C(x^C,y^C)\cos(2k\pi/N)\right]^2} \tag{5}$$

where the wrapped phase $\varphi^C(x^C,y^C)$ is wrapped by the inverse tangent function in the range $[-\pi,\pi)$ with 2π discontinuities, as shown in Fig. 1(c). To obtain the continuous phase distribution in the whole field, the wrapped phase needs to be unwrapped by the phase unwrapping formula, which is expressed as follows:

$$\phi(x^C,y^C) = \varphi^C(x^C,y^C) + 2\pi K(x^C,y^C) \tag{6}$$

where $K(x^C,y^C)$ represents the fringe order map, $\phi(x^C,y^C)$ is the unwrapping phase function determined by $K(x^C,y^C)$. Currently, the two main phase unwrapping methods are temporal phase unwrapping and spatial phase unwrapping. Compared with the spatial phase unwrapping method, the temporal phase unwrapping method has the advantages of independent of neighborhood pixels and is insensitive to noise. In this study, by using the multi-wavelength temporal phase unwrapping method, several wrapped phases with different frequencies are extended to a synthetic wavelength with lower frequency, and the synthetic wavelength is used to determine the integer order of the high frequency phase. The equation is written as follows:

$$K(x^C,y^C) = Round\left[\frac{(f_h/f_{eq})\varphi_{eq}(x^C,y^C) - \varphi_h^C(x^C,y^C)}{2\pi}\right] \tag{7}$$

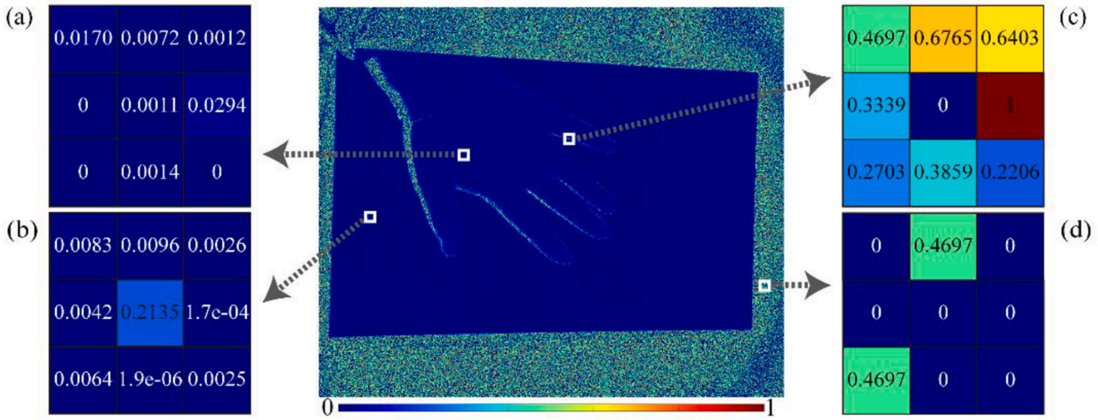
where function $Round(\cdot)$ represents rounding $\cdot$ to its closest integer, f_{eq} is the synthetic wavelength frequency, f_h is the high frequency wavelength frequency, $\varphi_{eq}(x^C,y^C)$ is the wrapped phase of the synthetic wavelength and $\varphi_h^C(x^C,y^C)$ is the high frequency wrapped phase. Note that the superscript "C" have been removed later for easier reading. The unwrapping phase map in Fig. 1(d) is obtained by Eq. (6) (7). As shown in Fig. 1(e), the unwrapping phase is converted to three-dimensional coordinates by pre-calibrated system parameters [24]. However, we can see from Fig. 1(e) that there are several discrete points with invalid height information in the resulting 3D point cloud model because of the influence of shadow, background, or clutter. Therefore, the removal of these invalid points is of great significance to generate high-quality 3D models.

Invalid point removal based on error energy function

In FPP, the projector is usually the main light source of the measurement environment, so the intensity variation of pixels in the camera image plane is mainly affected by the structured light reflected from various areas within the camera field of view. Taking the four-step phase shift as an example, Fig. 2(a) shows four representative pixel sequence points in the captured fringe patterns. The red points are a set of surface valid points which indicate that the surface of an object is correctly encoded by structured light. The yellow points are a set of surface points that become invalid after the manual addition of strong noise, which simulates the fringe patterns contaminated by a strong system noise. The blue and the green points are a set of background points and a set of shadow points, respectively, which are not directly illuminated by the structured-light due to the limitations of the measurement system. As calculated by Eq. (5), surface valid points have high intensity modulation, background/shadow points have low intensity modulation, and surface points with strong noise tend to have intensity modulation that is significantly different from background/shadow but of no practical significance. Therefore, it is sometimes difficult to filter out such invalid points when only the modulation threshold is used as the segmentation criterion. The background light intensity and intensity modulation of different areas in the camera field of view are different due to the influence of reflectivity and ambient light on the measured object, so the sequence points of different regions need to be converted to the same scale for analysis. By changing the form of Eq. (2) we get:

$$cos[\varphi(x,y) + 2k\pi/N] = \frac{I_k(x,y) - A(x,y)}{B(x,y)} \tag{8}$$

where on the right side of Eq. (8), the pixel sequence points $I_k(x,y)$ are normalized by background intensity $A(x,y)$ and modulation intensity $B(x,y)$, and the relationship between the normalized sequence values



(a)

0.0170	0.0072	0.0012
0	0.0011	0.0294
0	0.0014	0

(b)

0.0083	0.0096	0.0026
0.0042	0.2135	1.7e-04
0.0064	1.9e-06	0.0025

(c)

0.4697	0.6765	0.6403
0.3339	0	1
0.2703	0.3859	0.2206

(d)

0	0.4697	0
0	0	0
0.4697	0	0

Fig. 3. Error energy map calculated by Gaussian weighted Euclidean distance and displayed in pseudo color. (a) is the region of surface valid points; (b) is a surface invalid point surrounded by surface valid points; (c) and (d) are the regions of shadows and background with different amounts of low error energy, respectively.

and the ideal cosine curve $\cos(\cdot)$ in a single period is obtained. Fig. 2(b) shows the distribution of the four types of sequence pixel points in Fig. 2 (a) after normalization with respect to the ideal cosine curve, whose horizontal coordinates are determined by Eq. (3) and $\varphi(x,y) + 2k\pi/N$. The valid sequence points (red squares) on the object surface fit well on the ideal cosine curve, the variation of the surface sequence points with strong noise (yellow prisms) is similar to that of the ideal cosine curve, but it is obvious that the error caused by strong noise is propagated to the whole sequence points, while the normalized background points (blue circles) and shadow points (green triangles) obviously deviate from the ideal cosine curve and show disordered distribution. Chen [10] pointed out that the wrapped phase $\varphi(x,y)$ calculated by Eq. (3) is not the real phase of the object surface point, and the phase calculation at some discontinuities of the object becomes meaningless due to the camera sampling mechanism. Similarly, the phase of invalid points in areas such as background and shadows calculated by Eq. (3) also becomes meaningless. Thus Eq. (8) holds only when the sequence points on the captured fringe patterns are correctly encoded by the structured light. To obtain the cosine nature of the co-frequency pixel sequence points, we use Gaussian weighted Euclidean distance to quantize the dispersion between the sequence points and the ideal cosine curve as the error energy, as shown in Eq. (9):

$$Error(x,y) = \sqrt{\sum_{i=k}^{N-1} \frac{w_i}{W}(error_i(x,y))^2} \tag{9}$$

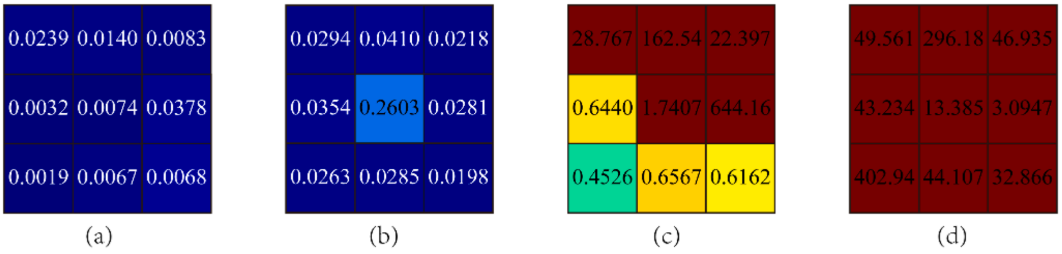
where $Error(x,y)$ is the error energy, w_i is the Gaussian weight, W is the normalization factor, and $error_k(x,y)$ is the deviation of the pixel relative to the ideal cosine curve. Then w_i, W and $error_k(x,y)$ can be defined respectively as:

$$w_i = e^{\left(-|error_i(x,y)|^{-2}/2\sigma_w^2\right)},\ W = \sum_{i}^{N-1} w_i \tag{10}$$

$$error_k(x,y) = cos[\varphi(x,y) + 2k\pi/N] - \frac{I_k(x,y) - A(x,y)}{B(x,y)} \tag{11}$$

Where the standard deviation σ_w determines the degree of influence of deviation $error_k(x,y)$ on Gaussian weight w_i, and k is the serial number of the captured fringe patterns,$k = 0, 1, ..., N-1$, N is total number of phase-shifting steps. As opposed to [12,25] where only the cumulative errors of the pixel sequence are considered, Eq. (9) also considers the effect between individual elements, that is, the larger the deviation $error_k(x,y)$, the more weight it has in the error energy $Error(x,y)$. Ideally, the error $Error(x,y)$ is approximately zero when the surface area of the object corresponding to pixel $I(x,y)$ in the camera image plane is correctly encoded by the structured-light, while the $Error(x,y)$ of the background/shadow area and the surface area contaminated by strong noise has a larger magnitude.

The phase shift method is essentially a least-squares fitting method that uses the cosine function to fit the measured phase shift values [11,26]. Therefore, in practical measurements, when some of the invalid sequence points change also approximating some cosine variation, the error energy obtained by Eq. (9) for these points as shown in Fig. 3 is similar to the error energy obtained by the valid sequence points. Fig. 3 (a) shows that the valid areas of the surface have low error energy. Fig. 3 (b) shows that the noise-contaminated invalid point in the center has a significant magnitude of error energy relative to the surrounding valid points. Fig. 3(c) and (d) show the error energy of the shadow and background regions, respectively, which theoretically should be large. However, Fig. 3(c) and (d) demonstrate that there are still low error energy points in this region. Therefore, it is difficult to obtain high-quality effective regions by using error thresholds to segment $Error(x,y)$. Considering the prior knowledge that invalid points are usually distributed concentratively and tend to have low intensity modulations, we propose an improved Gaussian filtering method based on modulation intensity to establish a new error energy function. The method incorporates the modulation information of pixel values and



(a)

0.0239	0.0140	0.0083
0.0032	0.0074	0.0378
0.0019	0.0067	0.0068

(b)

0.0294	0.0410	0.0218
0.0354	0.2603	0.0281
0.0263	0.0285	0.0198

(c)

28.767	162.54	22.397
0.6440	1.7407	644.16
0.4526	0.6567	0.6162

(d)

49.561	296.18	46.935
43.234	13.385	3.0947
402.94	44.107	32.866

Fig. 4. Error energy results after using the proposed algorithm; (a)-(d) corresponds to (a)-(d) in Fig. 3.

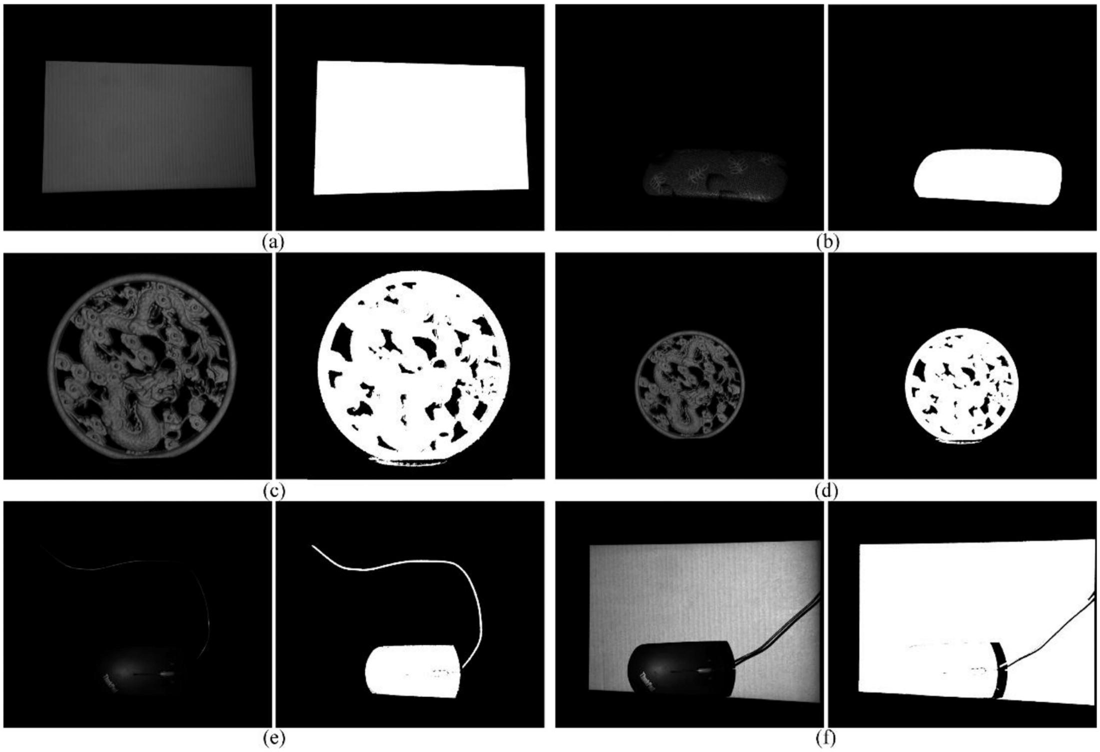


Fig. 5. Modulation images and ground truth for the fringe images: (a) white paper, (b) curved box with different colors, (c) a big dragon model, (d) a small dragon model, (e) a black mouse, (f) a black mouse with white background.

neighborhood error information according to the error energy characteristics, which can significantly improve the error energy difference between valid and invalid points. The proposed new error energy function model is as follows:

$$E(x,y) = EG(x,y) \times M(x,y) \tag{12}$$

Among:

$$EG(x,y) = Error(x,y) + (Gau \otimes Error)(x,y) \tag{13}$$

$$M(x,y) = \begin{cases} e^{\lambda(\alpha - B(x,y))}, & B(x,y) \leqslant \alpha \\ 1, & B(x,y) > \alpha \end{cases}, \quad 0.7 \leqslant \alpha \leqslant 5 \tag{14}$$

where $E(x,y)$ is the new error energy, Gau is a two-dimensional Gaussian window, α is the segmented point that affects the action range of weight factor $M(x,y)$, and λ determines the strength of the weighting factor $M(x,y)$. $EG(x,y)$ is an error energy function based on an improved Gaussian filtering method, which uses neighborhood error energy to compensate its own error energy. The error energy of the invalid points located in the background / shadow regions can be greatly affected by the neighborhood error energy, this function does not change error energy distinction between the invalid points and the valid points in the valid regions. $M(x,y)$ is a weighting factor based on modulation intensity, which corrects the calculated error energy function $EG(x,y)$ based on the prior knowledge that invalid pixels often have low modulation information on low-quality fringes in FFP. Fig. 4(a)-(d) shows the error energy results for Fig. 3(a)-(d) after using the proposed algorithm. Fig. 4(a) and (b) shows that the proposed method does not change the difference between the error energy of the valid points and the invalid points while retaining the low error energy at the valid points, and the invalid points with low error energy in Fig. 4 (c) and (d) are corrected. Finally, we compare the new error energy function with the error threshold T_{error} to get a mask template $MASK(x,y)$:

$$MASK(x,y) = \begin{cases} 1, & E(x,y) \leqslant T_{error} \\ 0, & E(x,y) > T_{error} \end{cases} \tag{15}$$

The detailed steps of our method are described as follows:

Step 1: Calculate the wrapped phase and intensity modulation of the initial images by Eq. (3) and (5), respectively.

Step 2: Calculate the error energy $Error(x,y)$ using the weighted Gaussian Euclidean distance Eq. (9).

Step 3: The new error energy $E(x,y)$ is obtained by correcting the error energy $Error(x,y)$ point by point by the proposed method Eq. (12).

Step 4: Select a proper error threshold T_{error}, and obtain a mask template $MASK(x,y)$ by marking all points less than T_{error} as 1 and the remaining points as 0.

Step 5: The invalid points in the background and shadow are removed by $MASK(x,y)$, in which the pixel of 1 in $MASK(x,y)$ is regarded as a valid point.

The adaptive selection of error threshold T_{error}

In the proposed method, the error threshold T_{error} is used to distinguish whether the point is valid at the final stage. Considering that the error energy of the valid point is concentrating in the low values region, error values ranging from 0 to L in error energy $E(x,y)$ can be utilized to denote the error-level histogram, where the error values L of 1 ~ 6 is determined to be appropriate for a reliable result. After establishing the error-level histogram, the proposed method automatically determines the error threshold T via the cumulative probability distribution of error energy. The formula is as follows:

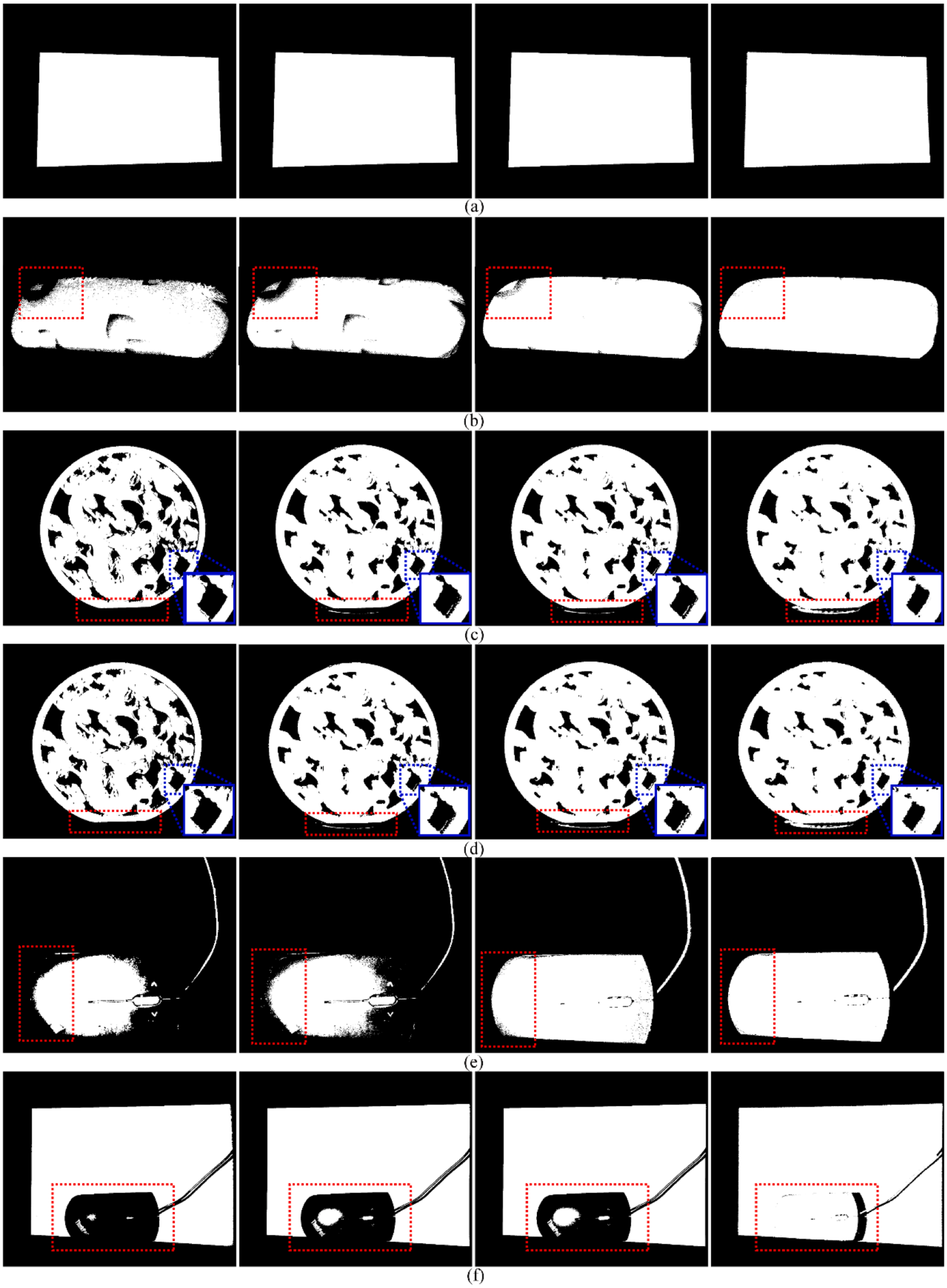


Fig. 6. Segmentation results of different objects in Fig. 5, where valid points are labeled in white. From left to right column: Wang's method, Du's method, Zhang's method, and our method.

$$T = \underset{0 \leqslant i \leqslant L}{Arg\ Min} \left| c - CDF(T) \right|, \quad CDF(T) = \sum_{i=0}^{T} n_i \Big/ \sum_{i=0}^{L} n_i \tag{16}$$

Where n_i is the number of pixels with error values i, c is a constant in the range of $[0, 1]$, $CDF(T)$ is cumulative probability. Eq. (16) indicates that the error threshold is the T when the cumulative probability $CDF(T)$ is close to c. After many experiments, the setting of value c less than 0.995 can effectively reduce the fluctuation of the threshold T caused by the different selection of the value L. Therefore, we set c to 0.995 and L to 3. Considering that the error values at the edge of the measured object are often larger than that inside the object, the error threshold T_{error} is set as

$$T_{error} = \beta T \tag{17}$$

Where β is the amplification ratio. We set β to 1.5 to identify more edge points with relatively large errors.

Experiments

The experimental system mainly consists of a DLP projector (LightCrafter4500, TI) with a resolution of 912 × 1140, a CMOS black-and-white camera (MV-CH050-10UM, HIKVISION) with a resolution of 2048 × 2448, and a computer with an Intel(R) Core(TM) i7-11800H @ 2.30 GHz. The projection scheme used in this study uses a four-step phase shift to obtain the wrapped phase and a three-frequency multi-wavelength phase unwrapping method to unwrap the phases. The fringe patterns with three frequencies (f = 70, 64, 59) are projected sequentially onto the measured object surface. The four-step phase shift method has the advantage of suppressing high-order harmonics [27].

Fig. 5 gives the modulation images of several different objects with different states and their manual segmentation of the ground truth. Fig. 5(a) is a white paper, Fig. 5(b) is a curved box with different textural complexity on the surface, Fig. 5(c) is a large wooden dragon model, Fig. 5(d) is a small wooden dragon model, Fig. 5(e) is a black mouse with overall intensity modulation close to the background, and Fig. 5(f) is a black mouse with a white background. Wang's method [28], Du's method [16], Zhang's method [15] and our method are compared on Fig. 5 and the segmentation results are shown in Fig. 6, where valid points are marked in white and invalid points are marked in black.

The objects shown in Fig. 6(a) and (b) are used to test the performance of the methods under different texture complexities. For the white paper with low texture complexity, the segmentation results of the four methods are similar. However, for the regions circled by red dashed lines in Fig. 6(b), where the texture is close to the background, our method can better identify these valid points. The surface shape of the measured object in Fig. 6(c) is more complex than that in Fig. 6(a) and (b). Du's method, Zhang's method and our method are better for valid points identification of complex surfaces. However, based on the local magnification of the area circled by the blue dashed line located in the lower right corner of Fig. 6(c) and the area circled by the red dashed line, our method can identify more valid points at the edges of the target. Fig. 6(d) is a different size dragon model from Fig. 6(c), but has the same conclusions as in Fig. 6(c). The error potential of each point in our method is only affected by the neighborhood error values and its own intensity modulation weight, so the segmentation performance is not affected by the measured target size. Compared to Fig. 6(a)-(d) where the overall texture of the object is significantly distinguished from the background/shadow, the intensity modulation of the mouse surface in Fig. 6(e) is close to that of the background. It can be seen that when the inter-class variance between the background/shadow and the target is small, the first two methods fail to identify regions where the intensity modulation is close to the background; see the area circled by the red dashed line. Zhang's method has a relatively reasonable segmentation result in Fig. 6(e) due to the enhanced weighting effect of the modulated histogram valleys. But in fact, Zhang's method only enhances the valley bottom point, which is easily affected by missing modulation level or noise leading to the failure of invalid points removal. Because our method directly aims at the fringe quality as the segmentation criterion, it can identify and remove invalid points in Fig. 6(e), which is better than the first three methods. The background of the images in Fig. 6(a)-(e) is black, and Fig. 6(f) shows the object with a white background. The white background area as a valid area significantly affects the ratio of pixels between the foreground and background. For the target whose intensity modulation is close to the background, the first three methods end to judge them as background pixels, while our method is independent of the measurement scene because it relies on the pixel's own error energy as the segmentation criterion. In Fig. 6, we tested the performance of the four methods in removing invalid points for different objects with different states, and our methods performed better and showed good robustness.

Table 1
MIoU of four methods for different images.

	Wang's method	Du's method	Zhang's method	Our method
Fig. 6(a)	0.9857	0.9920	0.9921	0.9981
Fig. 6(b)	0.9228	0.9568	0.9901	0.9927
Fig. 6(c)	0.8316	0.9147	0.9202	0.9815
Fig. 6(d)	0.9087	0.9544	0.9671	0.9906
Fig. 6(e)	0.7519	0.7635	0.9834	0.9891
Fig. 6(f)	0.8218	0.8410	0.8444	0.9936

Table 2
ME of four methods for different images.

	Wang's method	Du's method	Zhang's method	Our method
Fig. 6(a)	0.0071	0.0040	0.0039	0.0009
Fig. 6(b)	0.0159	0.0088	0.0020	0.0015
Fig. 6(c)	0.0917	0.0442	0.0411	0.0092
Fig. 6(d)	0.0233	0.0116	0.0083	0.0024
Fig. 6(e)	0.0387	0.0369	0.0025	0.0017
Fig. 6(f)	0.0970	0.0855	0.0834	0.0031

Qualitative results of the effects of four different methods are shown in Fig. 6. To quantitatively evaluate these methods, this study introduces the mean intersection over union (MIoU) [29] and the misclassification error (ME) [30] to reflect the image segmentation performance, respectively. MIoU is a standard accuracy assessment metric in the field of image segmentation, which averages the intersection and union ratio of each class of ground truth and predicted values; the closer MIoU is to 1, the better the prediction accuracy of the foreground object is. ME reflects the percentage of foreground pixels incorrectly assigned to the background, and ME varies from 1 for a completely mis-segmented image to 0 for an ideal segmentation result. For the binary classification problem, MIoU and ME can be simply expressed as follows, respectively:

$$MIoU = \frac{1}{2}\left(\frac{|B_o \cap B_T|}{|B_o \cup B_T|} + \frac{||F_o \cap F_T||}{|F_o \cup F_T|}\right) \tag{16}$$

$$\mathrm{ME} = 1 - \frac{|B_o \cap B_T| + |F_o \cap F_T|}{|B_o| + |F_o|} \tag{17}$$

where B_o and F_o denote the ground truth of background and foreground pixels, B_T and F_T denote the predicted results of background and foreground pixels, and $|\cdot|$ is the total number of elements in the set. The manually segmented ground truth is shown in Fig. 5. The MIoU and ME results shown in Fig. 6 are listed in Table 1 and Table 2. The results are consistent with the qualitative analysis results in Fig. 6. The proposed method gives the maximum MIoU value and the minimum ME value in the given test images. For the images in Fig. 6(a), (c), (d) with relatively large differences in object and background intensity modulation, the corresponding MIoU and ME values in the table show that Du's method and Zhang's method have similar prediction accuracy and classification error, while the MIoU of our method is 0.61 %, 6.25 % and 2.43 % higher than the highest MIoU values of the first three methods, respectively. The classification error is 4–10 times higher than that of the highest ME values of the first three methods. For both scenarios of Fig. 6 (b), (e), the MIoU and ME values of Zhang's method are similar to our method, but usually our proposed method gives better metric results. From Table 1 and Table 2, we can see that our method has the best segmentation precision for different measured objects with different states.

High-quality fringe patterns help to obtain reliable 3D reconstruction data. Fringe pattern areas that are heavily corrupted by additive noise or

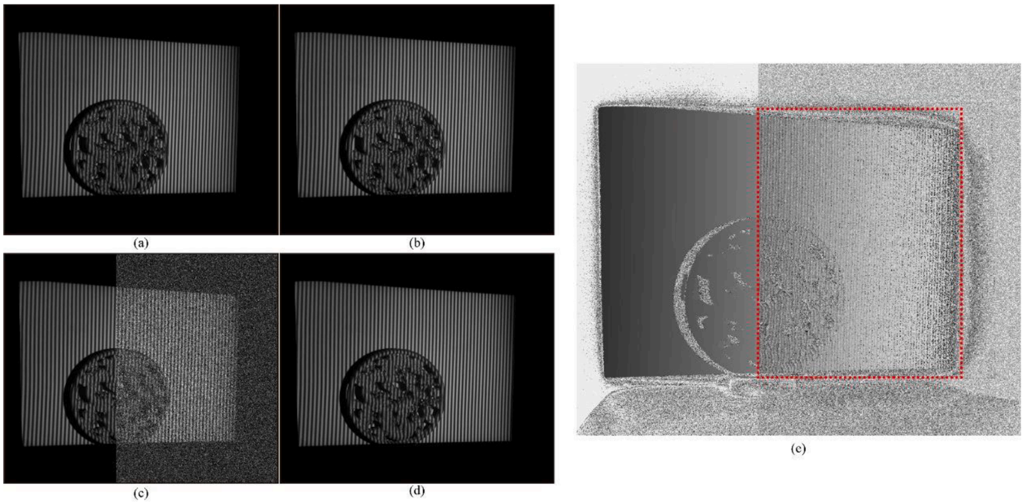


Fig. 7. Noise fringe images: (a) the captured first-step deformed patterns, (b) the captured second-step deformed patterns, (c) the captured third-step deformed patterns and mix strong noise on the right side, (d) the captured fourth-step deformed patterns, (e) unwrapped phase.

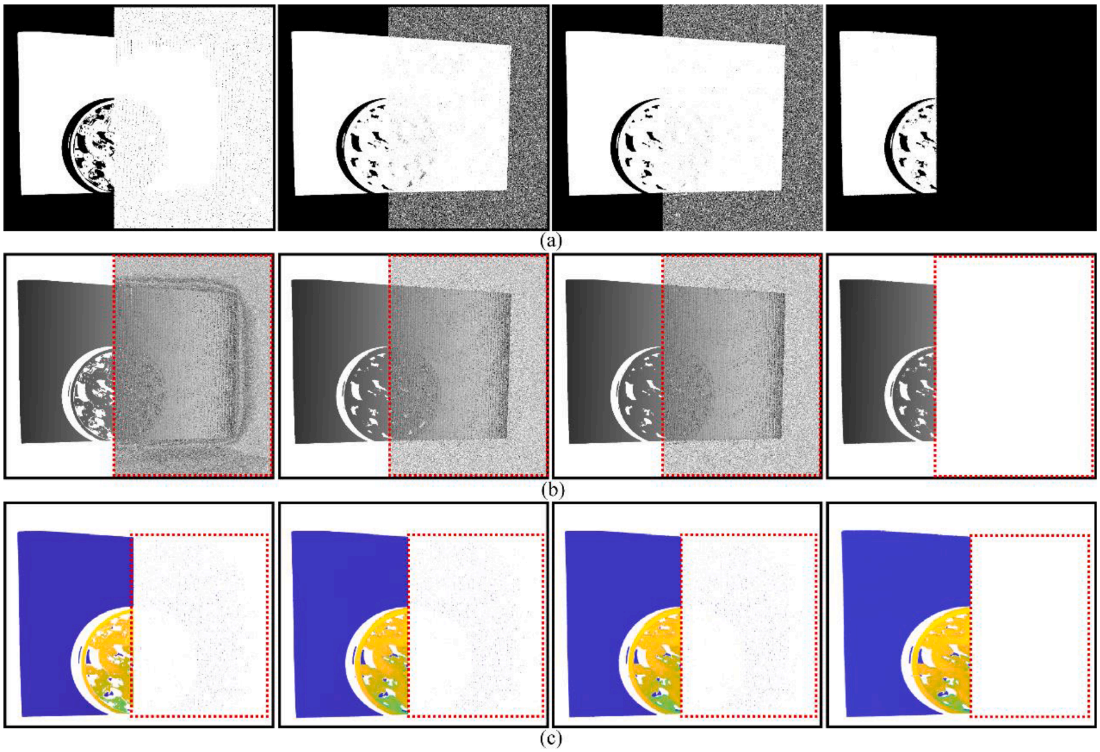


Fig. 8. Segmentation results for noise fringe image in Fig. 7. From left to right column: Wang's method, Du's method, Zhang's method, and our method: (a) the Segmentation results of four methods, (b) the Segmentation results of unwrapped phase using the four methods, (c) the range maps of segmentation results after using monotonicity on (b).

ambient glare do not carry any valid information about the measured objects, so these invalid points need to be identified and removed. To demonstrate the segmentation quality of our method, as shown in Fig. 7 (a)-(d), Gaussian noise is randomly added to the right side of one of the fringe images to corrupt the fringe information. The unwrapping phase map of this fringe sequence is obtained by Eqs. (3)-(6), also shown in Fig. 7(e). The resulting unwrapped phase is severely damaged in the red dashed boxed area. Fig. 8 shows the segmentation results after using Wang's method and Du's method, Zhang's method and our method are compared in Fig. 7. Fig. 8(a) shows the mask template obtained by the four methods, in which a pixel of 1 in mask template is regarded as a valid point. The unwrapping phase map is shown in Fig. 7(e) and it is extracted by applying the filter shown Fig. 8(a). The extracted results are shown in Fig. 8(b). The red dashed box region in Fig. 8(b) shows that the first three methods misclassify many invalid points as valid points, while our method identifies and removes these invalid points very well. The phase monotonicity check [7,8,10] is a general method to remove phase unwrapping errors. The results after using the phase monotonicity check

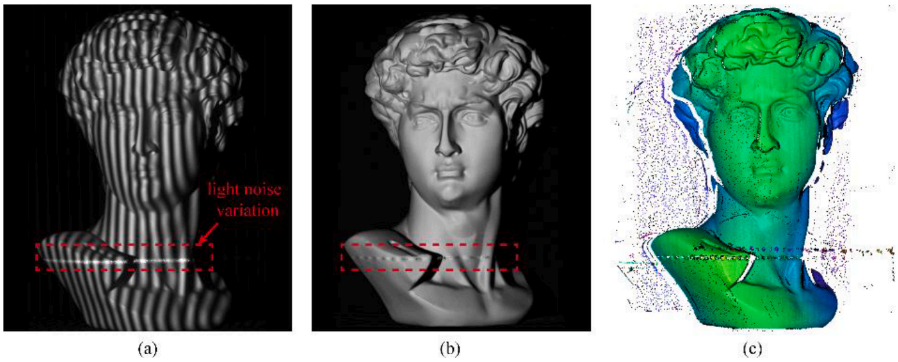


Fig. 9. The David statue model with significant linear light noise variation. (a) One of the captured fringe images, (b) modulation image, (c) original result of 3D reconstruction, (d) the result of manual removal on (c).

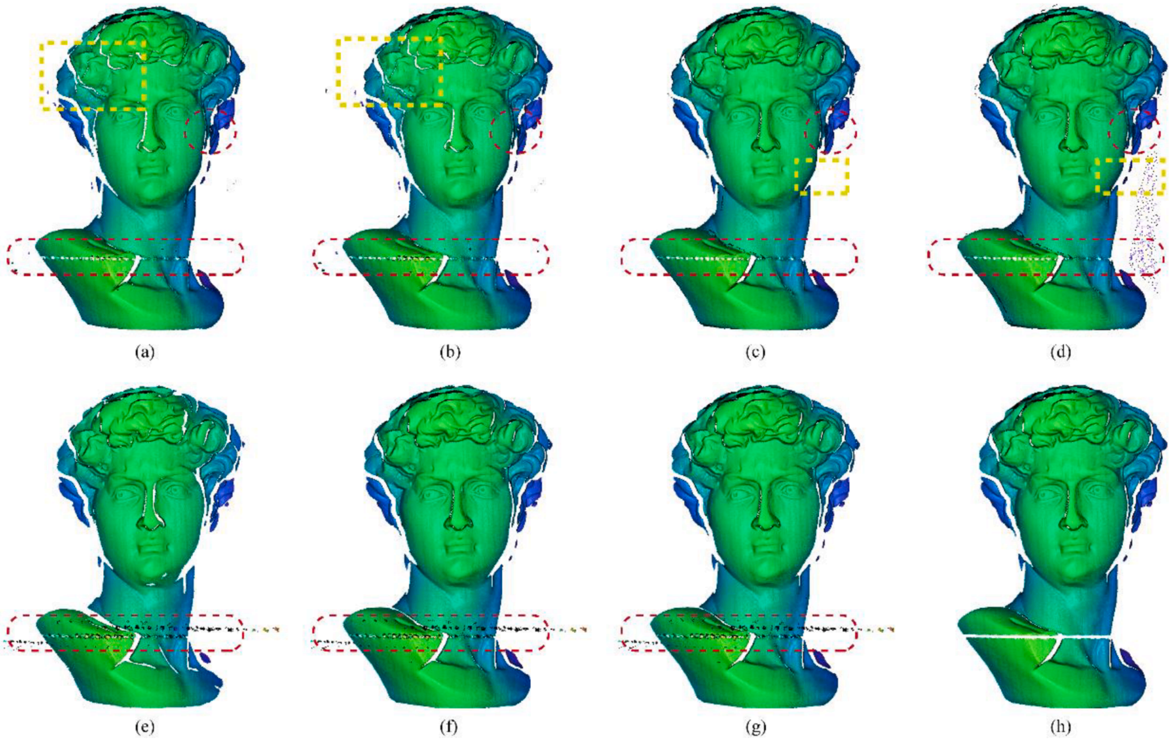


Fig. 10. Segmentation 3D results of the Fig. 9 using post-processing process, the comparison methods and the proposed method respectively. (a) and (b) are the filtering results using radius outlier removal and statistical outlier removal, respectively; (c) and (d) are the clustering results using Euclidean cluster extraction with different minimum distance parameters; (e)-(h) are the comparative results of Wang’s, Du’s, and Zhang’s method, respectively; (h) is the result of the proposed method.

for Fig. 8(b) are shown as range maps in Fig. 8(c). However, there are still many invalid points in the red dashed box area of the first three methods in Fig. 8(c) that cannot be removed, and these invalid points cause the reconstructed 3D model to be mixed with low quality data. The results of our method, shown in Fig. 8(c), remains consistent with Fig. 8(b). Compared with the first three methods, our method can obtain higher quality segmentation results.

To verify the effect of invalid point removal under light noise variation, as seen in Fig. 9(a), when DLP projector projects the fringe patterns onto the David statue model, we intermittently project the line laser as light noise variable onto the model, thus affecting the quality of the local fringe pattern. Fig. 9(b) is the obtained modulation image. Fig. 9(c) is the original point cloud without any processing, and it can been seen that the invalid points in this scene come not only from the background and shadow areas, but also from the areas polluted by light noise variation. We use the proposed method to remove the invalid points in this scene. For comparison, invalid points are also removed by the post-processing process and the threshold modulation, respectively. The result of manual removal in Fig. 9(d) is regarded as the ground truth and the invalid-point removal result is shown in Fig. 10.

Fig. 10(a)-(d) are the results of removing invalid points by using two post-processing methods including filtering and clustering. As shown in Fig. 10(a) and (b), radius filter and statistical filter can remove most of the invalid points, but some outliers still exist and some valid points at the edge of the contour are mistakenly filtered; see the area circled by the yellow dashed line in Fig. 10(a) and (b). Fig. 10(c) and (d) successfully removes most of the invalid points by using Euclidean cluster extraction with different minimum distance parameters of 35 and 40, but the clustering method is easy to identify the invalid points as object components when pursuing the integrity of the measured objects; see the area circled by the yellow dashed line in Fig. 10(c) and (d). What's more, the post-processing method shown in the red dashed line area of Fig. 10(a)-(d) has a poor effect on the removal of point clouds existing in the interior of the object. Fig. 10(e)-(g) are the results of removing invalid points by threshold modulation. From left to right column: Wang's method, Du's method, and Zhang's method. It can be seen that the invalid points in the background and shadow areas are well removed, but the deletion of the invalid points caused by the light noise variation fails; see the area circled by the red dashed line in Fig. 10(e)-(g). In fact, it can be seen from Fig. 9(b) that the modulation of the region polluted by light noise variation (see the area circled by the red dashed line) and the modulation of the valid region are intertwined, so it is difficult to achieve the ideal removal effect only through the threshold modulation. In contrast, the proposed method successfully removes all invalid points; see Fig. 10(h).

The post-processing process in Fig. 10(a)-(d) often requires the establishment of complex logical relationships between spatial points, while the proposed method directly starts from the reconstruction mechanism of 3D point clouds and removes invalid points according to the validity of point clouds rather than spatial relations. It is worth mentioning that effectively removing invalid points before generating large-scale point clouds helps to reduce the time consumption of 3D reconstruction. Compared with the threshold modulation method in Fig. 10(e)-(g), which reflects the light intensity fluctuation on the surface of the measured object without considering whether the fluctuation is cosine or not, the proposed method uses the error energy function to reflect the light intensity fluctuation of each imaging point under the joint influence of projected light and light noise variation.

Conclusion

We propose a method to identify and remove invalid points in fringe projection profilometry using the error energy function. The method quantifies the relationship between a set of captured co-frequency pattern pixel values and the ideal cosine curve as error energy using a Gaussian-weighted Euclidean metric. An improved Gaussian filtering method based on modulation intensity will significantly increase the error energy difference between the invalid pixels of the shadow/background regions and the valid pixels of the target. The valid region is determined by categorizing the size of the new error energy model with respect to a certain the error threshold. Since the method uses error energy as the invalidation criterion, the method is insensitive to object reflectivity or intensity modulation and can be applied in different test environments. The experimental results show that the proposed method is suitable for the task of removing invalid points in complex measurement scenario, and its segmentation accuracy, segmentation robustness and segmentation quality are better than the existing methods.

Deep learning techniques have been proved to effectively improve the performance of fringe analysis, fringe denoising, phase unwrapping and 3D reconstruction. Inspired by successes of deep learning in FPP, we further incorporate deep learning techniques into our work from two directions. First, the neural networks are used to replace some modules of the proposed method to achieve higher quality invalid point removal, for example, when only a single fringe is provided, multiple sets of phase-shifting sinusoidal fringe images [31] or high-accuracy phase [32] required by the proposed method can be obtained by the neural network. Second, using the proposed method or other prior information as the input source, the deep neural network (e.g., DCNN [33,34]) is used to automatically optimize the parameters to segment the valid region.

Funding

This work was supported by the Opening Project of Key Laboratory of Sichuan Universities of Criminal Examination (2018YB03).

CRediT authorship contribution statement

Kaifeng Zhu: Writing – original draft, Conceptualization, Methodology, Software. **Xin He:** Writing – review & editing, Supervision. **Yi Gao:** Funding acquisition, Investigation. **Ruidong Hao:** Formal analysis, Validation. **Zhonghui Wei:** Resources, Validation. **Bing Long:** Project administration. **Zhiya Mu:** Software, Resources. **Jun Wang:** Writing – review & editing, Conceptualization.

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.

References

[1] Wang YF, Zhao HJ, Jiang HZ, Li XD, Li YX, Xu Y. Paraxial 3d shape measurement using parallel single-pixel imaging. Opt Express. 2021;29:30543-57. https://doi.org/10.1364/oe.435470.

[2] He X, Kemao Q. A comparative study on temporal phase unwrapping methods in high-speed fringe projection profilometry. Opt Lasers Eng 2021;142:106613.

[3] Landmann M, Heist S, Dietrich P, Lutzke P, Gebhart I, Templin J, et al. High-speed 3d thermography. Opt Lasers Eng 2019;121:448–55.

[4] Zuo C, Qian J, Feng S, Yin W, Li Y, Fan P, et al. Deep learning in optical metrology: A review. Light-Sci Appl 2022;11(1).

[5] Heist S, Dietrich P, Landmann M, Kühmstedt P, Notni G, Tünnermann A. Gobo projection for 3d measurements at highest frame rates: A performance analysis. Light-Sci Appl 2018;7(1).

[6] Wang J, Zhou Y, Yang Y. A novel and fast three-dimensional measurement technology for the objects surface with non-uniform reflection. Results Phys 2020; 16:102878.

[7] Rivenson Y, Zhang Y, Günaydın H, Teng Da, Ozcan A. Phase recovery and holographic image reconstruction using deep learning in neural networks. Light-Sci Appl 2018;7(2).

[8] Song L, Chang Y, Li Z, Wang P, Xing G, Xi J. Application of global phase filtering method in multi frequency measurement. Opt Express 2014;22(11):13641.

[9] Zhang S. Phase unwrapping error reduction framework for a multiple-wavelength phase-shifting algorithm. Opt Eng 2009;48(10):105601.

[10] Chen F, Su X, Xiang L. Analysis and identification of phase error in phase measuring profilometry. Opt Express 2010;18(11):11300.

[11] Huang L, Asundi AK. Phase invalidity identification framework with the temporal phase unwrapping method. Meas Sci Technol 2011;22(3):035304.

[12] Feng S, Chen Q, Zuo C, Li R, Shen G, Feng F. Automatic identification and removal of outliers for high-speed fringe projection profilometry. Opt Eng 2013;52(1): 013605.

[13] Su X-Y, von Bally G, Vukicevic D. Phase-stepping grating profilometry: Utilization of intensity modulation analysis in complex objects evaluation. Opt Commun 1993; 98(1-3):141–50.

[14] Lau DL, Liu K, Hassebrook LG. Real-time three-dimensional shape measurement of moving objects without edge errors by time-synchronized structured illumination. Opt Lett 2010;35(14):2487.

[15] Zhang W, Li W, Yan J, Yu L, Pan C. Adaptive threshold selection for background removal in fringe projection profilometry. Opt Lasers Eng 2017;90:209–16.

[16] Du H, Chen X, Xi J. An improved background segmentation algorithm for fringe projection profilometry based on otsu method. Opt Commun 2019;453:124206.

[17] Lu L, Xi J, Yu Y, Guo Q, Yin Y, Song L. Shadow removal method for phase-shifting profilometry. Appl Opt 2015;54(19):6059.

[18] Qi Z, Wang Z, Huang J, Xing C, Gao J. Invalid-point removal based on epipolar constraint in the structured-light method. Opt Lasers Eng 2018;105:173–81.

[19] Skydan OA, Lalor MJ, Burton DR. Using coloured structured light in 3-d surface measurement. Opt Lasers Eng 2005;43(7):801–14.

[20] Du H, Chen XB, Xi JT, Yu CY, Zhao B. Development and verification of a novel robot-integrated fringe projection 3d scanning system for large-scale metrology. Sensors. 2017;17:13.https://doi.org/10.3390/s17122886.

[21] Le M-T, Chen L-C, Lin C-J. Reconstruction of accurate 3-d surfaces with sharp edges using digital structured light projection and multi-dimensional image fusion. Opt Lasers Eng 2017;96:17–34.

[22] Zuo C, Feng S, Huang L, Tao T, Yin W, Chen Q. Phase shifting algorithms for fringe projection profilometry: A review. Opt Lasers Eng 2018;109:23–59.

[23] Zuo C, Huang L, Zhang M, Chen Q, Asundi A. Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review. Opt Lasers Eng 2016;85:84–103.

[24] Feng S, Zuo C, Zhang L, Tao T, Hu Y, Yin W, et al. Calibration of fringe projection profilometry: A comparative review. Opt Lasers Eng 2021;143:106622.

[25] Su X, Chen W. Reliability-guided phase unwrapping algorithm: A review. Opt Lasers Eng 2004;42(3):245–61.

[26] Zhang C, Zhao H, Gao X, Zhang Z, Xi J. Phase unwrapping error correction based on phase edge detection and classification. Opt Lasers Eng 2021;137:106389.

[27] Dunin-Barkowski I, Kim J. Influence of the projected grid pattern distortions on measurement accuracy for phase shift based 3d inspection. 2005.https://doi.org/10.1117/12.648848.

[28] Wang H, Kemao Q, Soon SH. Valid point detection in fringe projection profilometry. Opt Express 2015;23(6):7535.

[29] Garcia-Garcia A, Orts-Escolano S, Oprea S, Villena-Martinez V, Martinez-Gonzalez P, Garcia-Rodriguez J. A survey on deep learning techniques for image and video semantic segmentation. Appl Soft Comput 2018;70:41–65.

[30] Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. J Electron Imaging 2004;13(1):146.

[31] Yu H, Chen X, Zhang Z, Zuo C, Zhang Yi, Zheng D, et al. Dynamic 3-d measurement based on fringe-to-fringe transformation using deep learning. Opt Express 2020;28(7):9405. https://doi.org/10.1364/oe.387215.

[32] Feng S, Chen Q, Gu G, Tao T, Zhang L, Hu Y, et al. Fringe pattern analysis using deep learning. J Adv Photon 2019;1:025001. https://doi.org/10.1117/1.AP.1.2.025001.

[33] Luo X, Bai S, Song W, Zhu S, Yu Q, Su J, et al. Learning-based invalid points detection for fringe projection profilometry. In: Eighth Symposium on Novel Photoelectronic Detection Technology and Applications. https://doi.org/10.1117/12.2627287.

[34] Yan K, Yu Y, Huang C, Sui L, Qian K, Asundi A. Fringe pattern denoising based on deep learning. Opt Commun 2019;437:148–52. https://doi.org/10.1016/j.optcom.2018.12.058.