

Received March 21, 2020, accepted April 2, 2020, date of publication April 7, 2020, date of current version April 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2986225

Phase Difference-3D Coordinate Mapping Model of Structural Light Imaging System Based on Extreme Learning Machine Network

SHANSHAN LV^{1,2}, MINGSHUN JIANG¹, CHENHUI SU¹, LEI ZHANG¹, FAYE ZHANG¹, QINGMEI SUI¹, AND LEI JIA^{1,2}

¹School of Control Science and Engineering, Shandong University, Jinan 250061, China

²Institute of Marine Science and Technology, Shandong University, Qingdao 266237, China

Corresponding author: Mingshun Jiang (jiangmingshun@sdu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61873333, in part by the Fundamental Research Fund of Shandong University under Grant 2018JCG06, in part by the Natural Science Foundation of Shandong Province under Grant ZR2017PEE023 and Grant ZR2017BF007, and in part by the National Key Research and Development Project under Grant 2018YFE02013.

ABSTRACT To meet the requirements of high accuracy and high efficiency in three-dimensional (3D) measurement, a phase difference-3D coordinate mapping model is proposed based on extreme learning machine (ELM) network. First, the reconstruction model of the ideal measurement system is set following the geometric structure of the system. Subsequently, by generalizing camera and world coordinates, a generalized measurement model is built. Lastly, ELM network is employed to solve the mapping coefficients. During measurement, only one phase difference map is required to complete the 3D reconstruction of the object, which simplifies the data processing process and saves time. The result indicates that the mean square errors (MSEs) of the X, Y and Z of the testing sample are 3.5955×10^{-4} mm, 9.5113×10^{-4} mm and 4.4×10^{-3} mm, respectively. Moreover, the reconstruction experiments of objects with different geometric structures are performed to demonstrate the general application of the proposed method.

INDEX TERMS Phase difference, structural light, 3D reconstruction, ELM network.

I. INTRODUCTION

3D imaging technology can be split into contact and non-contact measurement. In terms of the contact measurement, the coordinate measuring machine and articulated arm measuring machine are primarily included. The non-contact measurement covers optics [1]–[3], X-ray [4], [5], acoustics [6], [7] and magnetics [8], etc. The former measurement technology is not limited by the light and color of the object; it exhibits high accuracy. Nevertheless, the contact measurement technology has great limitations in the industry application since its probe cannot traverse the surface of the object, and it can easily damage the object. In contrast, the non-contact measurement primarily based on computer image processing has broader applications. To be specific, the 3D measurement based on digital fringe projection [9]–[11], as a type of surface measurement technology, exhibits several

advantages (e.g., fast measurement speed, high resolution, as well as simple system structure); it has been extensively applied in manufacturing and other fields.

Digital fringe projection technology is explained as follows. First, a group of digital coded fringes is projected onto the object surface with the use of a projector. Subsequently, the camera is adopted to acquire the deformed 2D fringe images. Lastly, given the pixel coordinates and phase information covered in the camera image, the 3D measurement of the object is achieved. According to whether a reference plane exists in the measurement system, the digital reconstruction method can fall into phase difference-coordinates [12]–[14] and phase-coordinates mapping models [15]–[17].

In fact, the phase difference mapping model has been built by setting the geometric relationships between the object, reference plane, camera and the projector. The model built in the early stage raises high requirements on the geometric constraints of the system structure [18], [19] (for example, the line between the optical center of the projector and the optical

The associate editor coordinating the review of this manuscript and approving it for publication was Andrea F. F. Abate¹.

center of the camera is required to be parallel to the reference plane), which limits its practical application. To generalize the model, An *et al.* [12], [13] substituted additional geometric parameters (e.g., the angle between the camera optical axis and the world coordinate axis) into the measurement system. The unknown coefficients in the mapping model can be solved by the least square algorithm based on a sufficient number of samples. However, considering the lens distortion [20], high order items of pixel coordinates are required in the model, thereby to some extent limiting the accurate calibration of the parameters. Cai *et al.* [14] proposed an optical field imaging method by installing a microlens array in front of the imaging sensor, in which each ray has its own phase-depth mapping coefficient. The method can achieve 3D reconstruction with high dynamic range, whereas the low spatial resolution causes the reconstruction result to lose some details.

In the phase mapping model, the projector is considered the reverse camera employed for building a binocular stereo vision system. Li *et al.* [15], [16] adopted camera calibration method to calibrate the projector. On the whole, the calibration accuracy of the projector is lower than that of the camera due to the gamma effect [21]. The binary defocusing technique proposed by Li *et al.* [22] can eliminate the gamma effect on calibration and accurately calibrate the inner parameters of the projector. However, the time-consuming corresponding points search causes low efficiency of 3D reconstruction. Based on the epipolar geometry constraint [23], Cai *et al.* [17] established a coefficient look-up table (LUT) to map the absolute phase to 3D coordinate. The result shows that the methods can be as accurate as the stereo vision method, while exhibiting higher reconstruction efficiency. However, the mapping coefficient of each imaging point should be solved, which makes the calibration process complex.

As revealed from the mentioned analysis, the direct calibration of the system parameters makes the calibration process flexible, whereas the measurement accuracy requires enhancement. If the mapping model is built for each imaging point, the complexity of the calibration process will increase. Moreover, the selection of calibration algorithm significantly impacts the measurement accuracy. Given the high requirements of accuracy and efficiency for 3D measurement, a concise and efficient measurement method is proposed in this paper. The sample database is directly established according to the collected image information and the measured 3D coordinate. Moreover, the ELM network with good generalization performance is adopted to calibrate system parameters. First, based on the principle of camera imaging and projection, the mapping model in general measurement system is deduced in detail. Then, the principle of ELM network used in nonlinear fitting and the calibration process of the system are introduced. Lastly, the effectiveness of the proposed method is verified by building the measurement system and analyzing the experimental results.

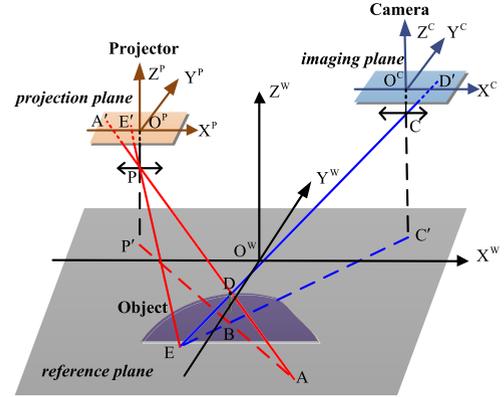


FIGURE 1. Ideal measurement system: camera coordinate system (CCS) and projection coordinate system (PCS) are built based on imaging plane and projection plane, respectively; the axes of world coordinate system (WCS), CCS and PCS are parallel to each other.

II. 3D RECONSTRUCTION BASED ON DIGITAL FRINGE PROJECTION

A. IDEAL MEASUREMENT SYSTEM

In the measurement system (Fig. 1), O^C and O^P refer to the projection points of the lens optical center C and P , respectively. The plane $X^W O^W Y^W$ is taken as the reference plane (RP), D denotes a point on the measured object, and B refers to its projection point on RP. C' and P' indicate the projection points of C and P on RP, respectively. A , E represent the intersections of RP with lines PD and CD , respectively. Based on the similar triangles $EBD \sim EC'C$ and $ABD \sim AP'P$, the coordinates of D in WCS can be expressed as:

$$\begin{cases} X_D^W = \frac{X_A^W Z_P^W (-X_E^W + X_C^W) + X_E^W Z_C^W (X_A^W - X_P^W)}{Z_P^W (-X_E^W + X_C^W) + Z_C^W (X_A^W - X_P^W)} \\ Y_D^W = \frac{(X_E^W - X_A^W)(Y_E^W - Y_C^W) Z_P^W}{Z_P^W (-X_E^W + X_C^W) + Z_C^W (X_A^W - X_P^W)} + Y_E^W \\ Z_D^W = \frac{(X_A^W - X_E^W) Z_P^W Z_C^W}{Z_P^W (-X_E^W + X_C^W) + Z_C^W (X_A^W - X_P^W)} \end{cases} \quad (1)$$

In Eq. (1), the superscript W denotes the coordinates in WCS. X_C^W , Y_C^W , Z_C^W , X_P^W and Z_P^W represent the structural parameters of the system, which are constants. According to the PCS-WCS and WCS-CCS mapping model deduced in Sect. 3 of Ref. [24], the following results are obtained:

$$\begin{cases} X_A^W = -\frac{T_0 Z_P^W}{2\pi l_p} \Delta\varphi_{AE} + X_E^W \\ X_E^W = \frac{-Z_C^W}{l_c} X_{D'}^C + X_C^W \\ Y_E^W = \frac{-Z_C^W}{l_c} Y_{D'}^C + Y_C^W \end{cases} \quad (2)$$

where l_p and l_c denote the lens focal length of the projector and camera, respectively. φ_{AE} is the phase difference between A and E . T_0 is the period of projected sinusoidal fringe pattern. $(X_{D'}^C, Y_{D'}^C)$ is the coordinate of D' in CCS.

In Fig. 1, the phase of D' is identical to that of A . Nevertheless, when there is no object to be measured, the phase

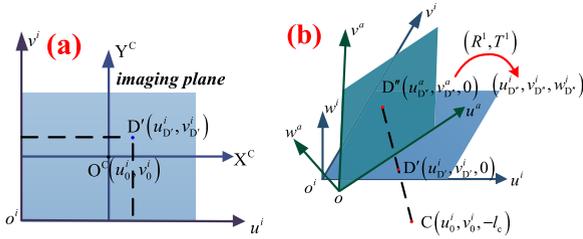


FIGURE 2. Imaging plane: (a) ideal imaging plane $u^i v^i$; (b) ideal imaging coordinate system $u^i v^i w^i$ and actual imaging coordinate system $u^a v^a w^a$.

of D' is equated with the phase of E. By substituting $\Delta\varphi_{AE}$ for $\Delta\varphi_{D'}$ ($\Delta\varphi_{D'}$ denotes the phase difference of D' between the actual measurement and the RP measurement), and combining Eqs. (1) and (2), the world coordinates of D can be expressed as Eq. (3).

$$\begin{cases} X_D^W = f_X(X_{D'}^C, Y_{D'}^C, \Delta\varphi_{D'}) \\ Y_D^W = f_Y(X_{D'}^C, Y_{D'}^C, \Delta\varphi_{D'}) \\ Z_D^W = f_Z(X_{D'}^C, Y_{D'}^C, \Delta\varphi_{D'}) \end{cases} \quad (3)$$

where f_X , f_Y and f_Z represent the mapping functions of $(X_{D'}^C, Y_{D'}^C, \Delta\varphi_{D'})$ to X_D^W, Y_D^W and Z_D^W , respectively.

B. GENERALIZATION OF MEASUREMENT SYSTEM

1) GENERALIZATION OF CAMERA COORDINATE

It is assumed that the physical size of each camera pixel along the u^i -axis and v^i -axis are denoted as dx and dy , the camera coordinate of D' is then expressed as Eq. (4). Where the superscript i denotes the pixel coordinates in the ideal imaging coordinate system (Fig. 2(a)).

$$\begin{cases} X_{D'}^C = (u_{D'}^i - u_0^i) dx \\ Y_{D'}^C = (v_{D'}^i - v_0^i) dy \end{cases} \quad (4)$$

As revealed in Fig. 2(b), since C, D' and D'' are collinear, Eq. (5) can be yielded.

$$\begin{cases} \frac{u_{D'}^i - u_0^i}{u_{D''}^i - u_0^i} = \frac{v_{D'}^i - v_0^i}{v_{D''}^i - v_0^i} = \frac{l_c}{w_{D''}^i + l_c} \\ \Delta\varphi_{D'} = \Delta\varphi_{D''} \end{cases} \quad (5)$$

\mathbf{R}^1 and \mathbf{T}^1 are assumed as the rotation and translation matrices of the coordinate system $u^a v^a w^a$ to $u^i v^i w^i$. Subsequently, the mapping relation between the coordinates of D'' in the two imaging coordinate systems is defined in Eq. (6).

$$\begin{aligned} \begin{bmatrix} u_{D''}^i \\ v_{D''}^i \\ w_{D''}^i \end{bmatrix} &= \mathbf{R}^1 \begin{bmatrix} u_{D''}^a \\ v_{D''}^a \\ 0 \end{bmatrix} + \mathbf{T}^1 \\ &= \begin{bmatrix} r_{11}^1 & r_{12}^1 & r_{13}^1 \\ r_{21}^1 & r_{22}^1 & r_{23}^1 \\ r_{31}^1 & r_{32}^1 & r_{33}^1 \end{bmatrix} \begin{bmatrix} u_{D''}^a \\ v_{D''}^a \\ 0 \end{bmatrix} + \begin{bmatrix} t_1^1 \\ t_2^1 \\ t_3^1 \end{bmatrix} \end{aligned} \quad (6)$$

Given the distortion of camera lens [17], there exhibits polynomial relation between the undistorted imaging coordinate $(u_{D''}^a, v_{D''}^a)$ of point D'' and the distorted imaging

coordinate $(u_{D''}, v_{D''})$ (i.e., the pixel coordinate on the image captured by the camera). Subsequently, combine Eqs. (4)-(6) to get Eq. (7), where f_u^p and f_v^p indicate non-linear functions.

$$\begin{cases} X_{D'}^C = f_u^p(u_{D''}, v_{D''}) \\ Y_{D'}^C = f_v^p(u_{D''}, v_{D''}) \\ \Delta\varphi_{D'} = \Delta\varphi_{D''} \end{cases} \quad (7)$$

2) GENERALIZATION OF WORLD COORDINATE

(X_D, Y_D, Z_D) is assumed as the coordinate of D in the general WCS, \mathbf{R} and \mathbf{T} represent the rotation and translation matrix from the ideal WCS to the general WCS, respectively. Subsequently, the general coordinates of D can be obtained by linear transformation of its ideal coordinates, as expressed in Eq. (8).

$$\begin{bmatrix} X_D \\ Y_D \\ Z_D \end{bmatrix} = \mathbf{R} \begin{bmatrix} X_D^W \\ Y_D^W \\ Z_D^W \end{bmatrix} + \mathbf{T} \quad (8)$$

3) GENERAL EXPRESSION OF MEASUREMENT MODEL

By substituting Eqs. (3) and (7) into Eq. (8), and by extending the measurement model of D to the whole imaging plane, Eq. (9) can be yielded.

$$\begin{cases} X = F_X(u, v, \Delta\varphi) \\ Y = F_Y(u, v, \Delta\varphi) \\ Z = F_Z(u, v, \Delta\varphi) \end{cases} \quad (9)$$

F_X, F_Y and F_Z represent the nonlinear mapping functions from (u, v, φ) to X, Y and Z, respectively. Accordingly, once the mapping function is determined, the 3D information of the object can be determined in accordance with the image information captured by the camera.

III. ELM NETWORK FOR IMAGING SYSTEM

A. ELM PRINCIPLE FOR NONLINEAR FITTING

ELM refers to a special type of artificial neural network, as employed to solve classification and regression problems [25], [26]. It exhibits the performance comparable to that of the error back propagation algorithm-based classical multi-layer perceptron training, whereas the training time can be narrowed by 6 orders of magnitude. During the ELM network training, the connection weight and threshold values between input and hidden layers are randomly generated. Only by setting the number of hidden layer neurons can the unique optimal solution be obtained. Besides, the solution exhibits fast learning speed and prominent generalization performance.

The structure of an ELM network with m input vectors (g_1, g_2, \dots, g_m) and n output vectors (s_1, s_2, \dots, s_n) is illustrated in Fig. 3, where l denotes the number of neurons in the hidden layer. h_{ij} indicates the connection weight between the i -th input neuron and j -th hidden layer neuron, and b_j represents the threshold value of j -th hidden layer neuron, all of which are randomly generated. Node “1” is the unit input node, and the value multiplied by threshold b_j is the bias of the hidden

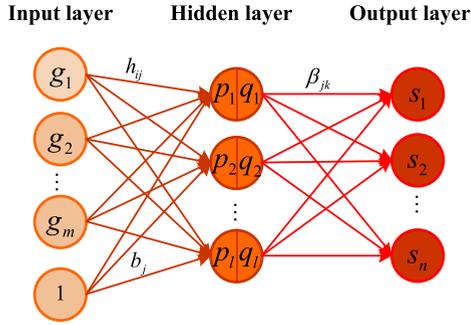


FIGURE 3. The structure of ELM network.

layer neuron. β_{jk} refers to the connection weight between the j -th hidden layer neuron and the k -th output neuron, acting as the parameter to be solved.

$f(\bullet)$ is assumed as the activation function, the inputs and outputs of the hidden layer can be expressed as:

$$\mathbf{P} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_l \end{bmatrix} = \begin{bmatrix} h_{11} & h_{21} & \cdots & h_{m1} \\ h_{12} & h_{22} & \cdots & h_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{1l} & h_{2l} & \cdots & h_{ml} \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \\ 1 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_l \end{bmatrix} = \mathbf{HG} + \mathbf{B} \quad (10)$$

$$\mathbf{Q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_l \end{bmatrix} = \begin{bmatrix} f(p_1) \\ f(p_2) \\ \vdots \\ f(p_l) \end{bmatrix} = f(\mathbf{P}) \quad (11)$$

Next, the output of the network is written in Eq. (12).

$$\mathbf{S} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{21} & \cdots & \beta_{l1} \\ \beta_{12} & \beta_{22} & \cdots & \beta_{l2} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{1n} & \beta_{2n} & \cdots & \beta_{ln} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_l \end{bmatrix} = \boldsymbol{\beta}\mathbf{Q} \quad (12)$$

If the actual output of the training sample is \mathbf{T} , the loss function is defined as $\mathbf{J} = \boldsymbol{\beta}(\mathbf{Q}-\mathbf{T})^T(\boldsymbol{\beta}\mathbf{Q}-\mathbf{T})$. For an activation function $f(\bullet)$, which is infinitely differentiable in any interval, when the l is equated with the number of training samples, ELM can approximate the result of samples with zero error for any \mathbf{H} and \mathbf{B} . Under the large number of samples, to reduce the calculation amount, l is generally smaller than the number of samples, and the training error of ELM can approximate any small error $\varepsilon > 0$. The connection weight $\boldsymbol{\beta}$ between the hidden layer and the output layer that minimizes the loss function can be calculated by Eq. (13), \mathbf{Q}^+ denotes the Moore-Penrose generalized inverse matrix of \mathbf{Q} .

$$\hat{\boldsymbol{\beta}} = \mathbf{S}\mathbf{Q}^+ \quad (13)$$

B. SYSTEM CALIBRATION

Figure 4 shows that the calibration of the imaging system is split into ELM network training and ELM network testing. In the part of network training, a group of network parameters exhibiting better performance are determined based on

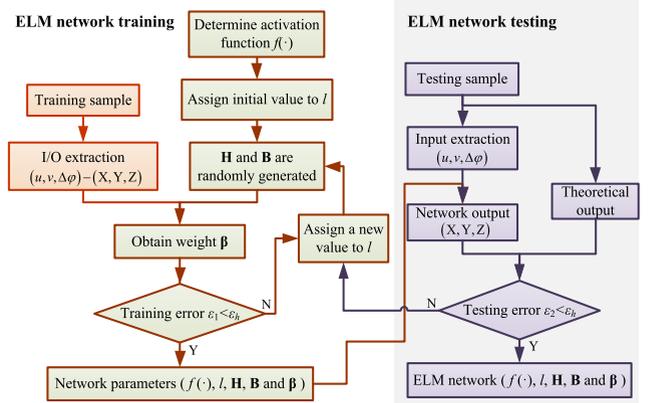


FIGURE 4. System calibration: the input vector of ELM network is the pixel coordinate (u, v) and phase difference $\Delta\varphi$ of the imaging point, and the output vector is the 3D world coordinate (X, Y, Z) of the corresponding object point.

the training samples. Network testing aims to employ the testing samples to verify the performance of the obtained ELM network. If the test error satisfies the requirements, the network is suggested to apply to 3D measurement. Otherwise, the number of hidden layer neurons is reset, and the solving process continues.

The specific system calibration is illustrated as follows.

- (1) Network initialization. The activation function $f(\bullet)$ and the number of neurons in the hidden layer l are determined.
- (2) Given the number of neurons in the hidden layer, the connection weight \mathbf{H} and threshold \mathbf{B} are randomly generated.
- (3) Based on training sample, the weight $\boldsymbol{\beta}$ is calculated by Eq. (13).
- (4) If the training error $\varepsilon_1 < \varepsilon_h$, the network parameters $(f(\bullet), l, \mathbf{H}, \mathbf{B}, \boldsymbol{\beta})$ will be recorded, and then step (5) continues. Otherwise, l will be reset, and then it turns to step (2).
- (5) The output of the testing sample is calculated based on the obtained network parameters $(f(\bullet), l, \mathbf{H}, \mathbf{B}, \boldsymbol{\beta})$ and the input of the testing sample.
- (6) The output of the testing sample calculated by the network is compared with the theoretical output. If the testing error $\varepsilon_2 < \varepsilon_h$, the ELM network is saved for subsequent application; otherwise, l is reset, and turn to step (2).

According to the imaging model presented in Sect. 2, the pixel coordinate (u, v) and phase difference $\Delta\varphi$ of each point are inputted by ELM network, while the output is the 3D world coordinate (X, Y, Z) . With the use of training error and testing error for the double check of the performance of the ELM network, the measurement accuracy of the model can be effectively ensured in the practical application.

IV. EXPERIMENT AND ANALYSIS

To verify the performance of the proposed method, the fringe projection 3D measurement system is built using

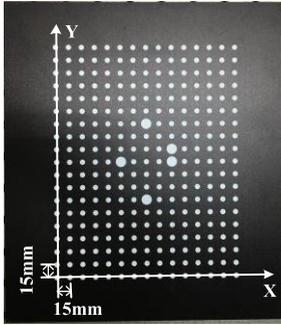


FIGURE 5. Calibration plate.

a CMOS camera (MER-500-14GM) and a DLP projector (IN2128HDx).

A. SYSTEM CALIBRATION AND TESTING

Figure 5 depicts a calibration plate printed with 15×19 white standard circles (the center distance is 15 mm). The calibration plate exhibits the size of $275\text{mm} \times 335\text{mm}$, which is fixed vertically on the slide rail (SGX57H56, accuracy of 0.05mm). This study takes the calibration plate as the reference plane, takes the center of the lower left circle as the origin, and takes the direction of the roller movement as the Z-axis to build WCS.

In this study, the standard 4-step phase-shifting algorithm is adopted to calculate the wrapped phase. The four sinusoidal patterns captured by camera are expressed as:

$$\begin{cases} I_1 = I' + I'' \cos(\varphi - \pi) \\ I_2 = I' + I'' \cos(\varphi - \pi/2) \\ I_3 = I' + I'' \cos(\varphi) \\ I_4 = I' + I'' \cos(\varphi + \pi/2) \end{cases} \quad (14)$$

where $I_1 \sim I_4$ denote the reflected intensities, I' is the background intensity, I'' is the modulated intensity. φ is the absolute phase associated with the height of the object. Based on the trigonometric function algorithm, the wrapped phase φ_w of the detecting area is calculated by Eq. (15).

$$\varphi_w = \arctan \frac{I_4 - I_2}{I_1 - I_3} \quad (15)$$

Given the significant difference of the light intensity on the surface of the calibration plate, to ensure the accuracy of the absolute phase, two phase-coding patterns are employed to solve the phase order N [27]. Next, the absolute phase is calculated by Eq. (16).

$$\varphi = \varphi_w + 2\pi \cdot N \quad (16)$$

Figures 6(a)-6(f) shows the sinusoidal and phase-coding patterns acquired by the camera when the calibration plate is placed at the position of $Z=0\text{mm}$. Figures 6(g) and 6(h) are the calculated phase order and unwrapped phase maps, respectively.

Figure 7(a) presents the captured image of the calibration plate at $Z=82.425\text{mm}$, and the enlarged image gives the

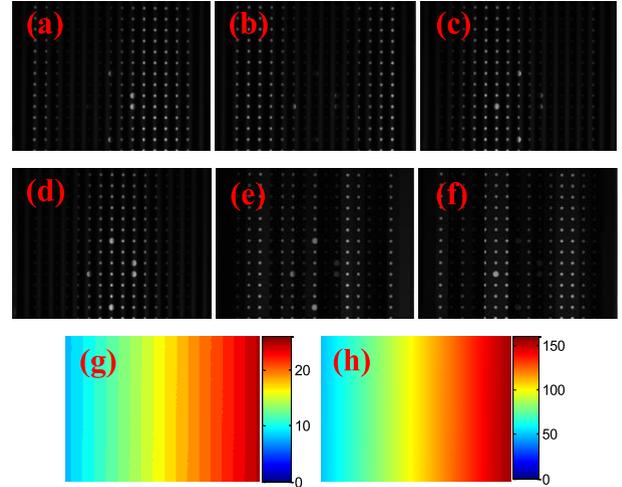


FIGURE 6. Calibration plate with $Z=0\text{mm}$: (a-d) 4-step phase-shifting patterns; (e-f) phase-coding patterns; (g) phase order map; (h) unwrapped phase map.

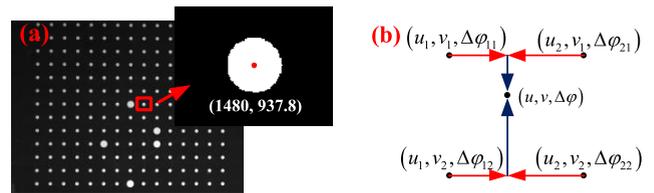


FIGURE 7. Calibration plate with $Z=82.425\text{mm}$: (a) circle center identification; (b) bilinear interpolation method for phase difference calculation (u_1, u_2, v_1 and v_2 are integer pixel coordinates adjacent to the circle center).

partial result of circle center recognition (marked with red dot) with the gray centroid method. Since the solved center coordinates are not necessarily integer, the bilinear interpolation method is adopted to calculate the phase difference of non-integer coordinates based on the phase difference of adjacent pixels. Linear interpolation is first performed in the direction of u and then in the direction of v . The calculation process is illustrated in Fig. 7(b), and the calculation formula is written in Eq. (17).

$$\Delta\varphi = \frac{v - v_1}{v_2 - v_1} \cdot \left(\frac{u - u_1}{u_2 - u_1} \Delta\varphi_{22} + \frac{u_2 - u}{u_2 - u_1} \Delta\varphi_{12} \right) + \frac{v_2 - v}{v_2 - v_1} \cdot \left(\frac{u - u_1}{u_2 - u_1} \Delta\varphi_{21} + \frac{u_2 - u}{u_2 - u_1} \Delta\varphi_{11} \right) \quad (17)$$

Within the measuring range of 0-350 mm, 9 groups of circle center data of calibration plate at different positions are taken as the training samples. Another 3 groups of calibration plate data act as the testing samples. The network performance when the number of neurons in the hidden layer meets the interval [20, 200] (with an interval of 10) is presented in Fig. 8. This figure reveals that in the initial stage, the MSE of the prediction results decreases rapidly as the number of neurons rises. At the number of neurons more than 100, the network performance tends to be stable, or the error becomes larger. Since too many neurons in the hidden layer are not-so-obvious to enhance the performance of the network while

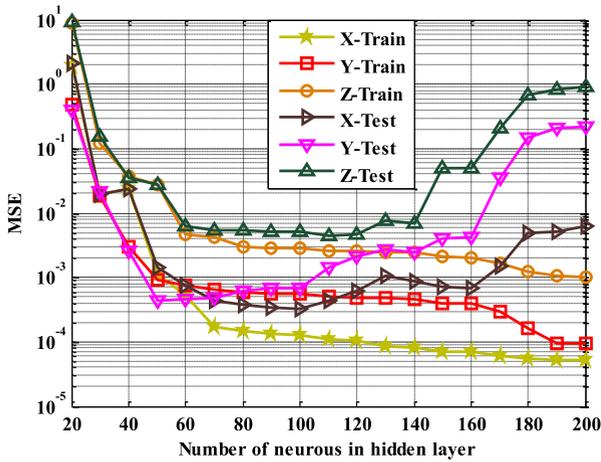


FIGURE 8. ELM network performance.

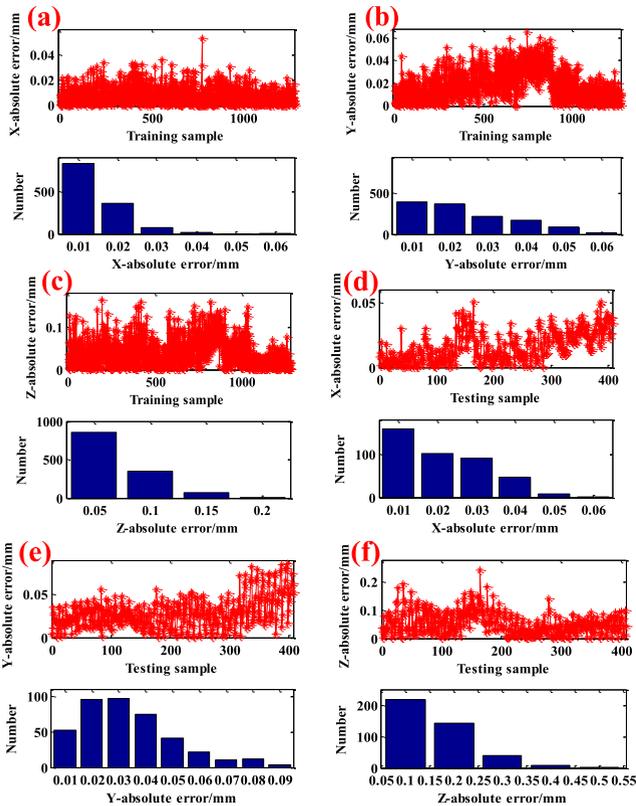


FIGURE 9. Absolute error of network output: (a)-(c) error of training sample; (d)-(f) error of testing sample.

extending the calculation time, the number of neurons in the hidden layer is set to 100.

With 100 neurons in the hidden layer, the MSEs of the predicted X, Y and Z of the training sample are 1.1118×10^{-4} mm, 5.3353×10^{-4} mm and 2.7×10^{-3} mm, respectively. The MSEs of the X, Y and Z of the testing sample are 3.5955×10^{-4} mm, 9.5113×10^{-4} mm and 4.4×10^{-3} mm, respectively. As revealed from the absolute error of network output presented in Fig. 9: (1) the Z coordinate exhibit lower prediction accuracy than the X and Y coordinates, probably associated with the accuracy of the slide rail and the fact

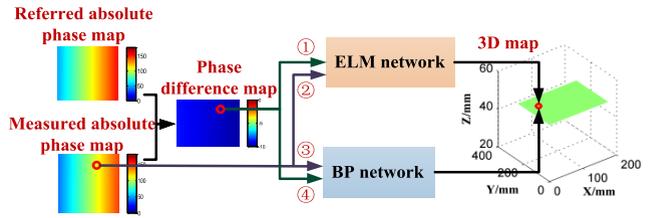


FIGURE 10. Four mapping models based on phase difference-3D coordinate and phase- 3D coordinate.

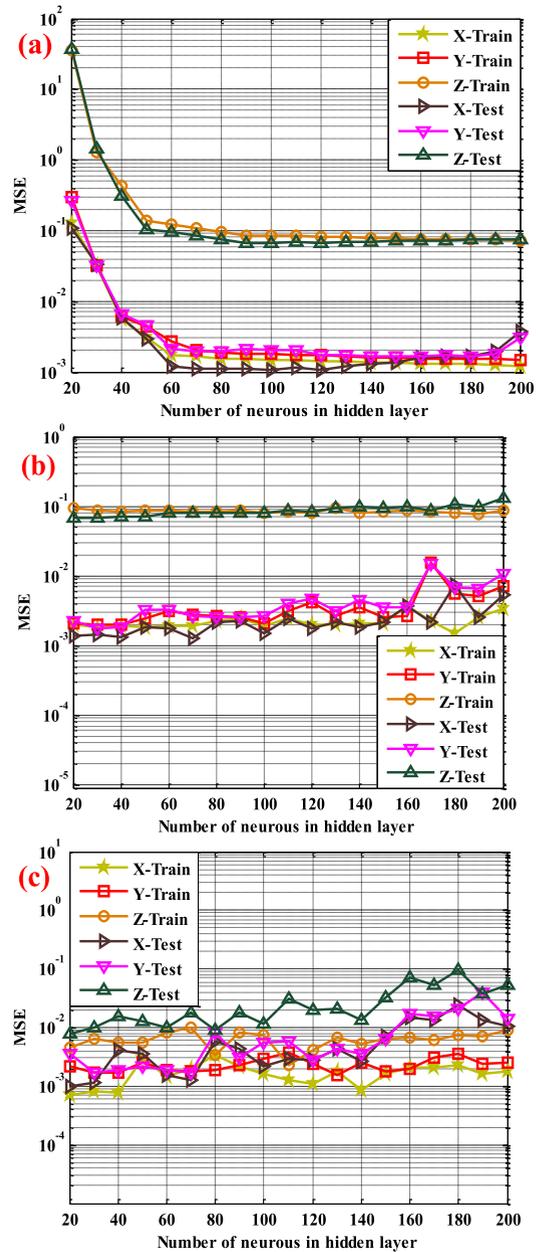


FIGURE 11. The performance curves of the three comparative mapping models: (a) method 2; (b) method 3; (c) method 4.

that its moving direction is not rigorously perpendicular to the calibration plate; (2) the testing sample exhibits slightly higher error than the training sample, whereas the difference is slight on the whole.

TABLE 1. Result comparison of the four measurement methods when the number of neurons in the hidden layer is 100.

Method	Training samples				Testing samples			
	MSE-X (10^{-3} mm)	MSE-Y (10^{-3} mm)	MSE-Z (10^{-2} mm)	Time(ms)	MSE-X (10^{-3} mm)	MSE-Y (10^{-3} mm)	MSE-Z (10^{-2} mm)	Time(ms)
①	0.11118	0.53353	0.27	4.836	0.35955	0.95113	0.44	2.476
②	1.5	1.8	8.56	5.505	1.1	2.1	6.75	1.416
③	1.9	2.8	8.14	123.196	1.4	2.5	9.32	43.484
④	1.5	2.8	0.4	121.568	2.0	4.3	2.25	47.755

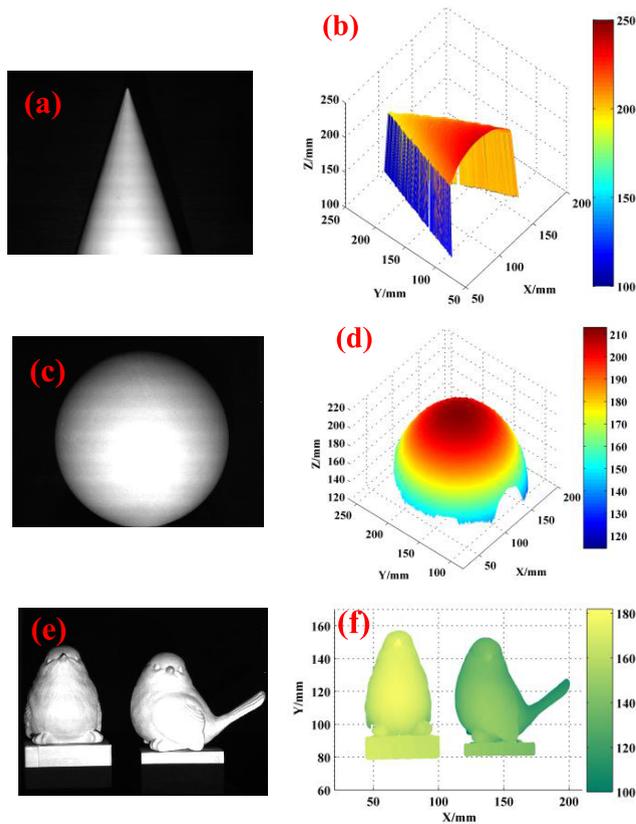


FIGURE 12. 3D reconstruction result: (a)-(b) physical map and reconstructed map of the cone; (c)-(d) physical map and reconstructed map of the sphere; (e)-(f) physical map of the two bird sculptures and top view of reconstructed map.

B. SYSTEM EVALUATION

Li *et al.* [16] proposed to use back propagation (BP) network to develop the mapping relationship between the imaging coordinates (x_c, y_c, x_p) and 3D coordinate (X, Y, Z) of the measured object. The parameter x_p is calculated based on the absolute phase, i.e., the mapping model belongs to a phase-3D coordinate mapping model based on BP network. As suggested from the experimental result, the measurement accuracy based on BP network is higher than that achieved with the general parameter calibration method. Based on the existing research, to verify the advantages of the proposed method, three groups of comparative models are set up (Fig. 10). Method ① is the method employed in this study;

method ③ is the phase-3D coordinate mapping model based on BP network proposed; methods ② and ④ refer to the phase-3D coordinate mapping model based on ELM network and the phase difference-3D coordinate mapping model based on BP network, respectively.

Several parameters of BP network are set as follows: the expected error is 10^{-7} , the maximum number of iterations is 200, the training method is Levenberg-Marquardt algorithm, and the activation function of hidden layer is “logsig” function. The performance curves of the three comparative mapping models for the training sample and testing sample are plotted in Fig. 11. As revealed from the results shown in Fig. 8 and Fig. 11, BP neural network is less affected by the number of neurons in the hidden layer. At the number of hidden layer neurons of 100, the prediction accuracy of each network has almost reached the optimal performance.

The MSEs of predicted X, Y, Z and the prediction time are used as evaluation indexes. Table 1 shows the prediction results of the four methods when the number of neurons in the hidden layer is 100. It can be seen that: (1) it is obvious that the prediction time of ELM network is far less than that of BP network; (2) the latter three methods have little difference in the prediction accuracy of X and Y, it may be because the influence of phase (or phase difference) on X and Y is far less than the influence of pixel coordinates; (3) the accuracy of the two methods based on phase difference is higher than that of the other two methods based on phase; (4) the prediction accuracy of ELM network is higher than that of BP network, whether the prediction model is based on phase difference or phase. Overall, the proposed measurement method shows the best performance in terms of accuracy and speed. Mean-time, the experiment result also reveals that the selection of calibration algorithm and model features significantly impact the measurement accuracy.

C. 3D RECONSTRUCTION

Three groups of reconstruction experiments are performed with the calibrated network parameters, as shown in Fig. 12. Since the 3D reconstruction method proposed in this paper is point-to-point mapping, when the accurate phase difference map is generated, the geometry and color of the object to be measured will not affect the stability of the system. As

suggested from the imaging results, the 3D surface obtained is complete and clear, revealing that the proposed method is feasible.

V. CONCLUSION

In this paper, a novel mapping model of structural light imaging system is proposed. The principle of mapping model and ELM network are elucidated. First, the sample library is established based on the circle center data of the calibration plate. By analyzing the network output error of samples, the number of neurons in hidden layer is determined. Subsequently, ELM network is used to predict the output of training samples and testing samples respectively. The result shows that the predicted absolute error of the X and Y coordinates is concentrated within 0.05mm, while the absolute error of the Z coordinates is concentrated within 0.3mm. Finally, three groups of 3D reconstruction experiments are performed with the calibrated network parameters. The experimental results verify the feasibility of the proposed method, so this method provides a new idea for the research of 3D measurement model.

Nevertheless, the prediction accuracy of Z coordinate is slightly less than that of X and Y coordinates. This finding may be associated with the fact that the Z-axis and the reference plane are not strictly vertical in WCS, thus leading to some errors in the built sample library. Second, due to the gamma effect of the projector, there is a certain error between the measured value and the real value of the phase. Hopefully, in our follow-up study, the measurement accuracy can be enhanced from the two aspects of sample data establishment and network performance enhancement.

REFERENCES

- [1] K. Khoshelham and S. O. Elberink, "Accuracy and resolution of kinect depth data for indoor mapping applications," *Sensors*, vol. 12, no. 2, pp. 1437–1454, Feb. 2012, doi: [10.3390/s120201437](https://doi.org/10.3390/s120201437).
- [2] B. Milosevic, F. Bertini, E. Farella, and S. Morigi, "A SmartPen for 3D interaction and sketch-based surface modeling," *Int. J. Adv. Manuf. Technol.*, vol. 84, pp. 1625–1645, Sep. 2016, doi: [10.1007/s00170-015-7828-1](https://doi.org/10.1007/s00170-015-7828-1).
- [3] B. R. De Araújo, G. Casiez, J. A. Jorge, and M. Hachet, "Mockup builder: 3D modeling on and above the surface," *Comput. Graph.*, vol. 37, no. 3, pp. 165–178, May 2013, doi: [10.1016/j.cag.2012.12.005](https://doi.org/10.1016/j.cag.2012.12.005).
- [4] L. Salvo, M. Suéry, A. Marmottant, N. Limodin, and D. Bernard, "3D imaging in material science: Application of X-ray tomography," *Comp. Rendus Phys.*, vol. 11, nos. 9–10, pp. 641–649, Nov. 2010, doi: [10.1016/j.crhy.2010.12.003](https://doi.org/10.1016/j.crhy.2010.12.003).
- [5] T. Sohmura, K. Wakabayashi, R. Lowmunkong, H. Hojo, N. Kusumoto, H. Okuda, T. Kojima, T. Nakamura, H. Yatani, and J. Takahashi, "3D shape measurement of dental casts using medical X-ray CT," *Dental Mater. J.*, vol. 23, no. 2, pp. 121–128, Mar. 2004, doi: [10.4012/dmj.23.121](https://doi.org/10.4012/dmj.23.121).
- [6] Y. Ji, S. Kwak, A. Yamashita, and H. Asama, "Acoustic camera-based 3D measurement of underwater objects through automated extraction and association of feature points," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst. (MFI)*, Sep. 2016, pp. 224–230.
- [7] S. Negahdaripour, H. Sekkati, and H. Pirsiavash, "Opti-acoustic stereo imaging, system calibration and 3-D reconstruction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
- [8] G. D. Bell, R. S. Rowland, M. Rutter, M. Abu-Sada, S. Dogramadzi, and C. Allen, "Colonoscopy aided by magnetic 3D imaging: Assessing the routine use of a stiffening sigmoid overtube to speed up the procedure," *Med. Biol. Eng. Comput.*, vol. 37, no. 5, pp. 605–611, Sep. 1999, doi: [10.1007/BF02513355](https://doi.org/10.1007/BF02513355).
- [9] S. Zhang, "High-speed 3D shape measurement with structured light methods: A review," *Opt. Lasers Eng.*, vol. 106, pp. 119–131, Jul. 2018, doi: [10.1016/j.optlaseng.2018.02.017](https://doi.org/10.1016/j.optlaseng.2018.02.017).
- [10] Y. He and Y. Cao, "A composite-structured-light 3D measurement method based on fringe parameter calibration," *Opt. Lasers Eng.*, vol. 49, no. 7, pp. 773–779, Jul. 2011, doi: [10.1016/j.optlaseng.2011.03.013](https://doi.org/10.1016/j.optlaseng.2011.03.013).
- [11] P. S. Huang, "High-resolution, real-time three-dimensional shape measurement," *Opt. Eng.*, vol. 45, no. 12, Dec. 2006, Art. no. 123601, doi: [10.1117/1.2402128](https://doi.org/10.1117/1.2402128).
- [12] D. An, F. P. Da, S. Y. Gai, and K. Lu, "New system calibration method on fringe projection profilometry," *J. Appl. Opt.*, vol. 35, no. 1, pp. 81–84, Jan. 2014.
- [13] D. An, S. Y. Gai, and F. P. Da, "A new model of three-dimensional shape measurement system based on fringe projection," *Acta Optica Sinica*, vol. 34, no. 5, May 2014, Art. no. 0512004, doi: [10.3788/AOS201434.0512004](https://doi.org/10.3788/AOS201434.0512004).
- [14] Z. Cai, X. Liu, X. Peng, Y. Yin, A. Li, J. Wu, and B. Z. Gao, "Structured light field 3D imaging," *Opt. Express*, vol. 24, no. 18, p. 20324, Sep. 2016, doi: [10.1364/oe.24.020324](https://doi.org/10.1364/oe.24.020324).
- [15] Z. Li, "Accurate calibration method for a structured light system," *Opt. Eng.*, vol. 47, no. 5, May 2008, Art. no. 053604, doi: [10.1117/1.2931517](https://doi.org/10.1117/1.2931517).
- [16] Z.-W. Li, Y.-S. Shi, C.-J. Wang, D.-H. Qin, and K. Huang, "Complex object 3D measurement based on phase-shifting and a neural network," *Opt. Commun.*, vol. 282, no. 14, pp. 2699–2706, Jul. 2009, doi: [10.1016/j.optcom.2009.04.055](https://doi.org/10.1016/j.optcom.2009.04.055).
- [17] Z. Cai, X. Liu, A. Li, Q. Tang, X. Peng, and B. Z. Gao, "Phase-3D mapping method developed from back-projection stereovision model for fringe projection profilometry," *Opt. Express*, vol. 25, no. 2, pp. 1262–1277, Jan. 2017, doi: [10.1364/oe.25.001262](https://doi.org/10.1364/oe.25.001262).
- [18] M. Takeda and K. Mutoh, "Fourier transform profilometry for the automatic measurement of 3-D object shapes," *Appl. Opt.*, vol. 22, no. 24, p. 3977, Dec. 1983, doi: [10.1364/ao.22.003977](https://doi.org/10.1364/ao.22.003977).
- [19] V. Srinivasan, H. C. Liu, and M. Halioua, "Automated phase-measuring profilometry of 3-D diffuse objects," *Appl. Opt.*, vol. 23, no. 18, pp. 3105–3108, Sep. 1984, doi: [10.1364/ao.23.003105](https://doi.org/10.1364/ao.23.003105).
- [20] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000, doi: [10.1109/34.888718](https://doi.org/10.1109/34.888718).
- [21] Z. Wang, D. A. Nguyen, and J. C. Barnes, "Some practical considerations in fringe projection profilometry," *Opt. Lasers Eng.*, vol. 48, no. 2, pp. 218–225, Feb. 2010, doi: [10.1016/j.optlaseng.2009.06.005](https://doi.org/10.1016/j.optlaseng.2009.06.005).
- [22] B. Li, N. Karpinsky, and S. Zhang, "Novel calibration method for structured-light system with an out-of-focus projector," *Appl. Opt.*, vol. 53, no. 16, pp. 12–3415, Jun. 2014, doi: [10.1364/AO.53.003415](https://doi.org/10.1364/AO.53.003415).
- [23] S. Negahdaripour, "Epipolar geometry of opti-acoustic stereo imaging," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1776–1788, Oct. 2007, doi: [10.1109/TPAMI.2007.1092](https://doi.org/10.1109/TPAMI.2007.1092).
- [24] S. Lv, M. Jiang, C. Su, L. Zhang, F. Zhang, Q. Sui, and L. Jia, "Flexible calibration method of an FPP system based on a geometrical model and NLSM with fewer parameters: Publisher's note," *Appl. Opt.*, vol. 58, no. 5, p. 1298, Feb. 2019, doi: [10.1364/ao.58.001298](https://doi.org/10.1364/ao.58.001298).
- [25] Z. Liouane, T. Lemlouma, P. Roose, F. Weis, and H. Messaoud, "An improved extreme learning machine model for the prediction of human scenarios in smart homes," *Int. J. Speech Technol.*, vol. 48, no. 8, pp. 2017–2030, Aug. 2018, doi: [10.1007/s10489-017-1062-5](https://doi.org/10.1007/s10489-017-1062-5).
- [26] G. Li, M. Liu, and M. Dong, "A new online learning algorithm for structure-adjustable extreme learning machine," *Comput. Math. with Appl.*, vol. 60, no. 3, pp. 377–389, Aug. 2010, doi: [10.1016/j.camwa.2010.03.023](https://doi.org/10.1016/j.camwa.2010.03.023).
- [27] S. Lv, M. Jiang, C. Su, L. Zhang, F. Zhang, Q. Sui, and L. Jia, "Improved unwrapped phase retrieval method of a fringe projection profilometry system based on fewer phase-coding patterns," *Appl. Opt.*, vol. 58, no. 32, pp. 8993–9001, Nov. 2019, doi: [10.1364/ao.58.008993](https://doi.org/10.1364/ao.58.008993).

...