Calibration Method of Liquid-Crystal Spatial Light Modulator Based on a 1-D Phase Retrieval Algorithm

Yifu Zhou, Hanyue Wei, Liyong Ren ¹⁰, Member, IEEE, and Jian Liang

Abstract — Liquid-crystal spatial light modulator (LC-SLM) has been widely applied as a programmable digital device. However, the LC-SLM can only manipulate on light fields accurately under designated wavelengths since, when being uploaded a specific grayscale image onto it, the phase retardance offered by the LC-SLM is relevant to the wavelength of the incident light. This means that the calibration of LC-SLM is indispensable once the working wavelength changes. In this paper, based on a phase retrieval algorithm, a novel phase calibration method with high efficiency and accuracy is proposed for scaling LC-SLM. In the method, a 1-D phase retrieval algorithm for recovering the phase of a 1-D light field distribution is used to measure the voltage-phase characteristic curve of LC-SLM, where the gradient descent algorithm with a Root Mean Square propagation is introduced to obtain the phase. Simulations and experiments show that this method is stable and has the ability of anti-noise on some conditions and can eliminate the influence caused by crosstalk between pixels on the calibration. Compared with the traditional diffraction-based method, our method improves the calibration error up to 30% under the same experimental conditions.

Index Terms—Spatial light modulator, far-field diffraction, phase retrieval algorithm.

I. INTRODUCTION

S PATIAL light modulators (SLMs) are kinds of optical instruments that can spatially manipulate amplitude or phase distributions of incident light fields. Liquid-crystal spatial light modulator (LC-SLM) and digital micromirror devices (DMD) are common SLMs. Particularly, LC-SLM, as a useful tool for manipulating phase distribution of light fields, has been widely used in various application fields such as optical trapping [1], [2], holographic displays [3], [4], lithography [5], [6] and adaptive optics [7].

LC-SLM can be regarded as a programmable pixelated phase mask, where a grayscale image can be uploaded and then modulates the phase distribution of the incident light according to the corresponding grayscale to each pixel. In the process,

The authors are with the Xi'an Key Laboratory of Optical Information Manipulation and Augmentation (OMA), School of Physics and Information Technology, Shaanxi Normal University, Xi'an 710119, China (email: 1070759771@qq.com; weihanyyue@163.com; renliy@snnu.edu.cn; jianliang@snnu.edu.cn).

Digital Object Identifier 10.1109/JPHOT.2023.3262919

the grayscale of each pixel, through a look-up-table (LUT), will be mapped to a specific voltage and thus being added to the control electrode for modulating the pixel unit of the LC-SLM, then the liquid-crystal upon the electrode will offer a voltage-related phase retardance for the incident light with a specific wavelength. Ideally, the phase retardance offered by each modulation unit on the working plane of LC-SLM is expected to be proportional to the grayscale of corresponding pixel on grayscale image, that means the grayscale between 0 to 255 should be mapped to a phase retardance between 0 to 2π linearly. However, the phase delay offered by each modulation unit of LC-SLM under a specific control voltage is relevant to the wavelength of incident light. Although the manufacturer of LC-SLM might have offered a LUT for a specific working wavelength to ensure the linear relationship between the grayscale and the resultant phase retardance, the LUT would be incompatible while the working wavelength changes. Therefore, it is essential to calibrate the LC-SLM for establishing a series of appropriate LUT corresponding to each different wavelength.

In fact, measuring the voltage-phase characteristics of LC-SLM is an important step to calibrate it. In this step, phase delays under different control voltages are measured when a beam of light with a specific wavelength illuminates the LC-SLM, as a result, a voltage-phase mapping relationship corresponding to this wavelength, i.e., a LUT, is constructed accordingly. Note that, ideally, such a LUT intends to ensure the phase delay varying with the grayscale linearly, i.e., producing a linear grayscalephase relationship. Unfortunately, measuring the phase of a light is usually more difficult than measuring its intensity, therefore, accurately measuring the voltage-phase characteristic curve of LC-SLM is an important step for the calibration.

Based on some different ways to measure the voltage-phase characteristic curve, some different calibration methods have been demonstrated so far, which include the self-referenced method [8], [9], the out-referenced method [10], [11], the diffraction pattern analyzed method [12], [13], holographic method [14], and polarization-based method [21], [22], [23], etc. The self-referenced and out-referenced methods belong to interferometry-based methods, which measure the phase modulation of a LC-SLM based on the fact that the interference fringes are sensitive to the phase difference between the reference and the modulated lights. The diffraction pattern analyzed method uses the diffraction pattern of the modulated light to measure the phase modulation of the LC-SLM. The holographic method reconstructs the phase distribution of the modulated light based on the digital holography. The polarization-based method measures

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

Manuscript received 2 March 2023; revised 21 March 2023; accepted 26 March 2023. Date of publication 29 March 2023; date of current version 10 April 2023. This work was supported in part by the Natural Science Foundation of Shaanxi Province under Grants 2021JM-204 and 2022JQ-612 and in part by Xi'an Scientific and Technological Projects under Grant 2020KJRC0013. (*Corresponding author: Liyong Ren.*)

the phase modulation of LC-SLM by analyzing the polarization state changes of the modulated light.

In general, both the out-referenced method and the polarization-based method can reach high calibration precisions relatively, but the former needs to move the LC-SLM and the latter requires a complicated data processing [20]. While the polarization-based method needs to introduce additional polarizers and retarders into the calibration system, leading to the system complicated and a less calibration accuracy due to more uncontrollable factors involved. On the other hand, the selfreferenced method and the holographic method use similar optical setup [13], [14], where a part of the reflective light of the LC-SLM is selected to be the reference light, leading to an inherited phase distortion and thus measurement errors. In fact, the diffraction pattern analyzed method shows a trade-off between the calibration accuracy and the system complexity, which is simple in the optical setup but relatively low in the accuracy as compared with both the out-referenced method and the polarization-based method in general. However, the diffraction pattern analyzed method is more practical for in-situ calibration [20].

Based on a phase retrieval algorithm, a calibration method for LC-SLM is also proposed and demonstrated in this paper. This method is a kind of diffraction pattern analyzed method that aims to acquire the phase distribution of light field modulated by LC-SLM from far-field diffraction pattern of the light. By changing the control voltage on each modulation unit, intensity distributions of far-field diffraction patterns of different modulated light fields can be collected, as a result, a voltage-phase characteristic curve of LC-SLM can be measured by analyzing the relationship between the intensity distribution and the phase distribution. In this method, a non-convex optimization problem is introduced to describe the relationship between the far-field diffraction pattern and the phase distribution yielded by LC-SLM, and a 1-D phase retrieval algorithm is proposed to solve this problem to acquire the phase distribution from the diffraction pattern. Numeric simulations show that this method has the ability of anti-noise and is stable on some conditions. Besides, experimental results show that this method can eliminate the influence of the crosstalk between pixels [15] on calibration. These results indicate that, as a kind of diffraction pattern analyzed methods, the proposed method is more accurate than traditional diffractive methods owing to the phase retrieval algorithm.

II. 1-D PHASE RETRIEVAL ALGORITHM

A. Problem Descriptions

Given a scalar complex amplitude distribution $E_{z=0}(x, y)$ of a light beam on the plane of z = 0 in \mathbb{R}^3 , and assuming that the light propagates along positive z direction and wavelength is λ , then the far-field diffracted pattern I(x, y) on a plane of $z = z_0$ satisfies the equation that

$$I(x,y) = \frac{1}{\lambda^2 z_0^2} |F(x,y)|^2,$$
(1)

where F(x, y) is the 2-D Fourier transform of the complex amplitude distribution $E_{z=0}(x, y)$, that is $F(f_x, f_y)$, with $f_x = x/\lambda z_0$ and $f_y = y/\lambda z_0$. Phase retrieval problem can be simply described as recovering the complex amplitude distribution $E_{z=0}(x, y)$ from the intensity distribution I(x, y) of its diffraction pattern. As usual, I(x, y) is captured by digital imaging devices such as a CCD or a CMOS camera, meaning that the distribution of intensity is discretized. Assuming the pixel numbers of the camera alone x-axis and yaxis are M and N, respectively, then the elements in the intensity signal matrix obtained from the camera can also be expressed by

$$I(m,n) = \frac{1}{\lambda^2 z_0^2} |F(m,n)|^2,$$
(2)

where F(m, n) denotes the discretized signal matrix of 2-D Fourier transform F(x, y). Besides, F(m, n) can be approximately regarded as the *M*-*N* point 2-D discrete Fourier transform (DFT) of a matrix $E \in \mathbb{C}^{M_0 \times N_0}$, where *E* is a finite sampling of $E_{z=0}(x, y), M_0$ and N_0 are the sampling numbers alone *x*-axis and *y*-axis, respectively. From this point of view, let the *M*-*N* point 2-D DFT of *E* be D(m, n), then the discretized phase retrieval problem can be described as

Find
$$\boldsymbol{E} \in \mathbb{C}^{M_0 \times N_0}$$
 to minimize $f(\boldsymbol{E})$
$$f(\boldsymbol{E}) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \left| |D(m,n)|^2 - |F(m,n)|^2 \right|^2 .$$
(3)

Obviously, f(E) = 0 if and only if $|D(m, n)|^2 = |F(m, n)|^2$, meaning that E is the solution of the discretized phase retrieval problem. In this case, the problem is a minimization problem. To solve the problem (3), some algorithms have already been proposed, which include the Wirtinger Flow [16] and the Phase Lift [17]. Note that the former solved the problem (3) directly by using Wirtinger derivatives, while the latter relaxed this problem to a convex optimization problem.

Similar to the above general problem (3), the calibration task of LC-SLM is a little bit easier since it can be simplified to a 1-D phase retrieval problem as follows:

Find
$$\varphi \in \mathbb{R}^{M_0}$$
 to minimize $f(\varphi)$

$$f(\varphi) = \sum_{m=0}^{M-1} \left| |D(m)|^2 - |F(m)|^2 \right|^2, \qquad (4)$$

where F(m) is the 1-D DFT of the 1-D complex amplitude distribution of a light beam with a fixed amplitude *A*, and D(m)is the 1-D DFT of *E* with the expression of $E = A\exp(j\varphi)$, where $\exp(x)$ represents the exponential function for each component of vector *x*. In this problem, the 1-D complex amplitude of light field with a fixed amplitude will be recovered from the intensity of its 1-D DFT, meaning that only the phase of it should be acquired. It should be pointed out that the problem (4) is a simplified problem because only the phase should be acquired, which makes the problem become a real optimization problem from a complex one, and less parameters will be recovered.

B. Algorithm

In fact, as for the 1-D phase retrieval problem (4), D(m) can be expressed by

$$D(m) = \boldsymbol{f}_m^H \boldsymbol{E},\tag{5}$$

where f_m^H is the *m*-th row of the Fourier matrix with a size of $M \times M_0$, the superscript H denotes the conjugate transpose. The gradient of the loss function $f(\varphi)$, *i.e.*, $\nabla f(\varphi)$, can be expressed by

$$\nabla f(\boldsymbol{\varphi}) = 2A^2 \sum_{m=0}^{M-1} \left[|D(m)|^2 - |F(m)|^2 \right] \boldsymbol{b}_m, \quad (6)$$

where \boldsymbol{b}_m is a vector in \mathbb{R}^{M_0} and is relevant to *m*, which can be expressed as

$$\boldsymbol{b}_{m} = \operatorname{Im}\left[\left(\boldsymbol{f}_{m}^{H}\boldsymbol{f}_{m}\boldsymbol{E}\right)\odot\boldsymbol{E}^{*}\right],\tag{7}$$

where Im(x) is a vector that is constructed by the imaginary part of each component of x. The superscript * denotes the complex conjugate, and the symbol \odot denotes the Hadamard product.

Then, the gradient descent algorithm is adopted to solve the problem (4). Given an initial guess φ_0 , the corresponding iterative formula is expressed by

$$\varphi_{k+1} = \varphi_k - \eta \nabla f\left(\varphi_k\right), \tag{8}$$

where η is the learning rate, i.e., the step length of the iteration. For an appropriate η , $f(\varphi_k)$ will decrease with the increase of the numbers of iterations *k*.

In general, the maximum times of iterations, the initial guess φ_0 and the learning rate η should be set manually in advance. But in this work, an adaptive algorithm, i.e., the Root Mean Square propagation (RMSprop) [18], is also introduced to decide the learning rate dynamically and locally, where a vector $\eta_k \in \mathbb{R}^{M_0}$ is introduced to be the learning rate instead of the scalar learning rate η :

$$\boldsymbol{\varphi}_{k+1} = \boldsymbol{\varphi}_k - \boldsymbol{\eta}_{k+1} \odot \nabla f\left(\boldsymbol{\varphi}_k\right). \tag{9}$$

Equation (9) means that each of the component of the decision variable φ_k has a different dynamical learning rate which changes with the iteration. The learning rate vector η_k can be calculated by

$$\eta_k = \eta \frac{1}{\sqrt{G_k + 1}},\tag{10}$$

where G_k is a vector in \mathbb{R}^{M_0} and $1/\sqrt{G_k + 1}$ represents a vector which is constructed by the scalar function $1/\sqrt{x+1}$ of each component x of vector G_k . Note that η is a constant to be set manually in general and it determines the initial learning rate. Besides, the calculation of η_k shown in (10) is different with the original RMSprop algorithm. We used $1/\sqrt{G_k + 1}$ instead of $1/\sqrt{G_k}$, which is used in the original algorithm, to avoid η_k being too large when G_k is too small. And G_k is calculated by the iteration through

$$\boldsymbol{G}_{k+1} = \rho \boldsymbol{G}_k + (1-\rho) \left[\nabla f\left(\boldsymbol{\varphi}_k\right) \odot \nabla f\left(\boldsymbol{\varphi}_k\right) \right].$$
(11)

As usual, the initial value G_0 is set to be **0**. And ρ is a constant in (0, 1) and is also the so-called forgetting rate. It is seen from (1, 1) that G_k is the weighted sum of squares of the present and the past descent directions and the forgetting rate ρ determines the ratio of the present and the past imformation in G_k . According to (10), if a component of G_k is large enough, then the corresponding learning rate in η_k for the component will be

Algorithm 1: 1-D phase retrieval. Input: Observed intensity signal $|F(m)|^2$ with length M, Initial learning rate η , Maximal number of iterations K, Length of output signal M_0 , Forgetting rate ρ . Initialize: Normalize $|F(m)|^2$ to $\sum_{M=0}^{M-1} |D(m)|^2 = M \times M_0$. Set initial guess $\varphi_0 \in \mathbb{R}^{M_0}$ and $\varphi_0 = \mathbf{0}$. Set initial signal $E_0 = A\exp(j\varphi_0)$, where A = 1. Set $G_0 \in \mathbb{R}^{M_0}$ and $G_0 = 0$. Loop: **For** k = 0 to K - 1 **do:** Calculate $\nabla f(\varphi_k)$ via (6) and (7), where A = 1. Calculate G_{k+1} via (11). Calculate η_{k+1} via (10). Calculate φ_{k+1} via (9). Calculate E_{k+1} via $E_k = A \exp(j\varphi_k)$, where A = 1. End Output: Recovered phase distribution φ_K .

quite small. So, the mechanism of RMSprop can be described as that the learning rate for a component of φ_k is determined by two factors, one is the learning rate of a component that is inversely proportional to the corresponding component of descent direction $\nabla f(\varphi_k)$, the other is the information of the past descent direction which influences the present learning rate. The former ensures the learning rate is not too large or too small; the latter introduces a momentum mechanism in gradient descent that can make convergence faster.

For simplicity, the amplitude of the signal to be recovered, *A*, is fixed to be 1 due to the Parseval's equality of DFT:

$$\|\boldsymbol{E}\|_{2}^{2} = \frac{1}{M} \sum_{M=0}^{M-1} |D(m)|^{2}, \qquad (12)$$

where $\|\cdot\|_2$ denotes the 2-norm for vectors. It's easy to obtain $\|\boldsymbol{E}\|_2^2 = \|A\exp(j\varphi)\|_2^2 = M_0$ and $\sum_{M=0}^{M-1} |D(m)|^2 = M \times M_0$ for A = 1. In the experiment, $|F(m)|^2$ is captured by a camera and is normalized to ensure $\sum_{M=0}^{M-1} |F(m)|^2 = M \times M_0$ for satisfying the energy conservation.

Based on (6), (7), (9), (10) and (11), the detailed phase retrieval algorithm for problem (4) is given in follows.

The output of the algorithm, i.e., the recovered 1-D phase distribution φ_K , could be regarded as the solution of problem (4) if $f(\varphi_K)$ is sufficiently small. This algorithm will be used to recover the phase retardant distribution of LC-SLM in Section III. In the experiment, $|F(m)|^2$, as the 1-D intensity distribution of the far-field diffractive pattern on a line, is captured by a camera and introduced into the algorithm, then the output, φ_K , gives the phase retardant distribution of LC-SLM on the line accordingly. By recovering phase retardant distributions under different phase mask, the voltage-phase characteristic curve could be measured.



Fig. 1. Original rectangle phase distribution and recovered phase distribution.

C. Numerical Simulations and Analyses

In this section, some numerical simulations on the 1-D phase retrieval algorithm proposed in Section II-A are carried out, where phase distributions to be recovered with a rectangle and a sinusoidal profile are supposed, respectively. As for the simulation parameters, the sampling length of the intensity pattern capered by a camera, M, is set to be 128; the length of the phase distribution uploaded onto the LC-SLM and thus to be recovered, M_0 , is also set to be 128; the maximal number of iterations, K, is 4000, the forgetting rate ρ is 0.99, and the initial learning rate η is 0.001.

Fig. 1 shows the original rectangular phase distribution with an amplitude of π and the corresponding recovered distribution. It is seen that the profile of the recovered phase distribution is accorded with the original one, but there exists an overall vertical phase shift between them. This result shows an important fact that there is a constant phase shift between the recovered and the original phase distributions, and, indeed, this algorithm has the ability to recover the phase distribution profile. It should be emphasized that diffractive characteristics of light is fully dependent on the phase difference distribution (instead of the absolute phase) of the LC-SLM, meaning that it is enough for measuring the voltage-phase characteristic curve of LC-SLM provided that the profile of the phase difference distribution is obtained accurately. In other words, the overall constant phase shift will not affect the calibration.

In principle, the overall phase shift in Fig. 1 is related to the linear property of DFT. Let X(m) be the *M*-point DFT of signal $x(m_0)$, this linear property of DFT can be expressed as

$$DFT_M \left[\exp(j\theta) x(m_0) \right] = \exp(j\theta) X(m), \qquad (13)$$

where the operator $DFT_M(\cdot)$ denotes the *M*-point 1-D DFT, θ is a real constant. Equation (13) means that recovering signal $x(m_0)$ from $|X(m)|^2$ will lead to an ambiguity, i.e., the recovered distribution may have a phase shift θ due to $|DFT_M[x(m_0)]|^2 = |DFT_M[\exp(j\theta)x(m_0)]|^2$. In fact, it can be observed that the constant phase shift is random when the initial guess φ_0 of the algorithm is random.

Besides, in the view of amplitude, the average amplitude of the recovered distribution is 0.996π , there also exists an amplitude error of 0.004π between the recovered distribution and the original one.



Fig. 2. Original sinusoidal phase distribution and recovered phase distribution.

Fig. 2 shows the original sinusoidal phase distribution with an amplitude of $\pi/2$ and the corresponding recovered distribution. Also, there exists a small overall vertical phase shift between them. In addition, it is shown in Fig. 2 that an obvious horizontal cyclic displacement is occurred in the recovered phase distribution. But the horizontal cyclic displacement does not change the profile of the phase distribution when the distribution is periodic. In fact, the calibration method used in this paper will upload and only upload a series of periodic phase mask onto the LC-SLM, meaning that such a horizontal cyclic displacement does not influence the calibration process.

Similarly, the horizontal cyclic displacement is attributed to the cyclic displacement property of DFT owing to the theorem:

$$DFT_{M}\left[\tilde{x}\left(m_{0}+k\right)\right] = \exp\left(-j\frac{2\pi mk}{M}\right)X\left(m\right), \quad (14)$$

where $\tilde{x}(m_0 + k)$ denotes the result of *k*-point cyclic displacement for $x(m_0)$. Equation (14) shows another ambiguity, being yielded when recovering signal $x(m_0)$ from $|X(m)|^2$, that the recovered distribution might be the result of cyclic displacement for the original distribution due to $|\text{DFT}_M[x(m_0)]|^2 = |\text{DFT}_M[\tilde{x}(m_0 + k)]|^2$.

The simulation results shown in Figs. 1 and 2 indicate that there are essential difficulties in this kind of phase retrieval problem, i.e., there always yields an inherent information difference between the distribution of the complex amplitude and that of the intensity, meaning that it is irreversible once the complex amplitude of light field has been converted to intensity.

Some extra numeric simulations are conducted to illustrate excellent stability and strong anti-noise ability of the algorithm when $M > M_0$. Fig. 3(a) shows the recovered phase distributions of the rectangular phase distribution used in Fig. 1, where M is set to be 128, while M_0 is set to be 12, 64 and 128, respectively. And Fig. 3(b) shows the normalized intensity distributions in the frequency domain of the original signal and the recovered signals under different M_0 . Besides, Fig. 3(c) shows the normalized mean square error (MSE) curves of gradient descent algorithm results under different M_0 . The MSE curves show the changing of normalized MSE of the used gradient descent algorithm with iterations, where the normalized MSE is the normalized value of the cost function we used, because we are dealing with a nonlinear least squares problem.



Fig. 3. (a) Original phase distribution and recovered phase distributions under different M_0 ; (b) normalized intensity distributions in frequency domain of original signal and recovered signals under different M_0 . (c) Normalized MSE curves of gradient descent algorithm results under different M_0 .

It is seen from Fig. 3(a) that the recovered phase distributions for those cases when $M > M_0$ just reflect parts of the original phase distribution. In this point of view, it is M_0 that determines how much the length of the recovered part of the original phase distribution could be. This is owing to the fact that the length of the phase distribution which could be recovered is depending on the length of the intensity distribution $|F(m)|^2$. Note that, if the length of $|F(m)|^2$ is M, then the maximal M_0 to be set at M is enough. Once $M < M_0$, the part of the recovered phase distribution, whose length is greater than M, will be a periodic extension of the preceding part, meaning that those extended parts are redundant for phase retrieval.

On the other hand, Fig. 3(b) shows that in the frequency domain the difference between the recovered complex amplitude distribution and the original one will increase with the decrease of M_0 . It is seen from Fig. 3(b) that, as compared with the original signal, there exists just a slight difference when $M_0 = 128$ (corresponding to the nearly zero MSE in Fig. 3(c)); but there exist obvious differences when $M_0 = 64$ and $M_0 = 12$



Fig. 4. Original phase distribution and recovered distributions with noise under different M_0 .

TABLE I AVERAGE AMPLITUDE AND ITS ABSOLUTE ERROR OF RECOVERED PHASE DISTRIBUTIONS

Phasedistribution	Average amplitude	Absolute
		error
Original	1.000π	0.000π
Recovered at $M_0 = 128$	1.011π	0.011π
Recovered at $M_0 = 64$	1.007π	0.007π
Recovered at $M_0 = 12$	0.999π	0.001π

(corresponding to the large MSEs in Fig. 3(c)). However, shown in Fig. 3(a) indicates that the profiles of phase distributions are all accurate in the spatial domain, even though in the frequency domain the differences between the recovered complex amplitude distributions and the original ones are obvious for $M_0 =$ 64 and $M_0 = 12$. Therefore, a conclusion could be made that the algorithm inherently has the ability to obtain the profile of phase distribution in the spatial domain even though there may exist larger errors in the frequency domain. And based on this ability, as a solution of phase retrieval problem, this algorithm is effective accordingly.

Fig. 4 shows the results of another simulation that is basically similar to the one shown in Fig. 3, but the original intensity distribution in the frequency domain, i.e., $|F(m)|^2$, is added Gaussian noise with SNR = 40 (SNR is the signal-to-noise ratio). Here *K* is set to be 6000 to ensure the sufficiency of iteration. It can be seen from Fig. 4 that the profiles of phase distributions under *M* $< M_0$ are closer to the original one obviously, meaning that this algorithm has an anti-noise ability when $M < M_0$.

Considering these recovered distributions features related to the sampling length, the random vertical phase shift and the random horizontal cyclic displacement, it is reasonable to evaluate the accuracy of them in aspect of amplitude. Table I shows the average amplitude and its absolute error of different phase distribution shown in Fig. 4 with respect to the original one. It indicates that the error decreases with the decrease of M_0 , meaning that, if the phase distribution to be recovered is periodic, it is effective and enough to recover a part of the signal (for example, only one period) instead of the entire.

1



Fig. 5. Optical path used to measure voltage-phase characteristic curve.

III. CALIBRATION METHOD

In this section, a calibration method and the principle will be introduced. According to previous discussions in introduction, the measurement of voltage-phase characteristic curve is an important step of the calibration of LC-SLM, so this section will stress on how to measure this curve.

First of all, it is necessary to initialize the LUT of LC-SLM to ensure that the grayscale can be linearly mapped to the control voltage because the grayscale distribution of the digital image uploaded to LC-SLM can be regarded as the control voltage distribution on working area in this condition. In this time, measuring the voltage-phase characteristic curve is equivalent to measuring the grayscale-phase characteristic curve.

The optical path used to measure the voltage-phase characteristic curve is schematically shown in Fig. 5. A linearly polarized planar light beam illuminates a reflective LC-SLM vertically, which provides the wavelength of light to be calibrated. According to Fig. 5, an image with a periodic stripe pattern is uploaded to LC-SLM, that means the phase distribution of light field on the working area of LC-SLM is also striped, like a Ronchi grating. The grayscale of black stripes is fixed to be 0 (that means the control voltage on these areas is also zero), the grayscale of white stripes is set in $0\sim 255$ manually.

Under the coordinate system shown in Fig. 5, assuming that the widths of the working area along the *x*-axis and the *y*-axis are a_x and a_y , respectively, then the modulated light field, noted by E(x, y), on the plane where the working area is located can be described by

$$E(x,y) = A \exp\left[j\varphi \operatorname{rect}\left(\frac{x}{2T}\right) \otimes \operatorname{comb}\left(\frac{x}{T}\right)\right] \operatorname{rect}\left(\frac{x}{a_x}\right)$$
$$\operatorname{rect}\left(\frac{x}{a_y}\right), \tag{15}$$

where *T* is the period of the striped phase modulation distribution, *A* is the constant amplitude of light source, \otimes denotes convolution operator, φ is the difference of phase delay between areas located on black stripes and white stripes. φ could be regarded as the phase retardance offered by the modulation unit located on white stripes if the phase delay offered by modulation unit located on black stripes is regarded as 0. Obviously, φ is relevant with grayscales of white stripes. One can measure the voltage-phase characteristic curve to establish the relationship between φ and the grayscale.



Fig. 6. (a) Extracting 1-D intensity distribution from far-field diffraction pattern; (b) recovered phase distribution of *E* along *x*-axis from observed 1-D intensity distribution and measuring φ from recovered phase distribution.

It is reasonable to ignore the effect of the limited width of LC-SLM since the period T can be set to a value that is much less than a_x and a_y . In this point of view, there is an approximation that can be expressed by

$$E(x,y) \approx A \exp\left[j\varphi \operatorname{rect}\left(\frac{x}{2T}\right) \otimes \operatorname{comb}\left(\frac{x}{T}\right)\right].$$
 (16)

According to (16), actually *E* is only relevant to *x*. Set 1-D Fourier transform of *E*, with respect to *x*, to be $\mathcal{F}_x(E)$, and the 2-D Fourier transform of *E* to be $\mathcal{F}(E)$, there is an equation that

$$\mathcal{F}(E) = \mathcal{F}_x(E)\,\delta\left(f_y\right),\tag{17}$$

where δ is the Dirac Delta function. (17) means that the energy of the diffracted light will concentrate on a line on the plane of the far-field diffraction pattern. That is the reason why light spot of the far-field diffraction pattern for light field modulated by a classical 1-D grating such as the binary grating will distribute on a line.

In Fig. 5, the far-field pattern of *E* will be collected by a camera. The intensity distribution on image plane, i.e., I(x, y), satisfies (1) with $z_0 = f$, where *f* is the focal length of the lens. According to (16) and (17), it is predictable that the pattern will be consist of some discreate light spots (because *E* is periodic), and the light spots will distribute on a line. In ideal conditions, light spots are ideal points without area, but in the experiment, as shown in Fig. 5 these spots have certain areas because of the diffraction limit.

In order to establish the relationship between φ and the grayscale of white stripes, it is feasible to measure φ from far-field diffraction patterns under different grayscale of white stripes. In this paper, the phase retrieval algorithm used in previous numeric experiments will be used to measure φ from far-field diffraction patterns. The measurement processes are shown in Fig. 6.

Firstly, as shown in Fig. 6(a), the left side shows the intensity distribution of the diffraction pattern. The intensity distribution

of a zone that centered on zero-order light spot with size of $5 \times M$ will be measured and a data matrix with the same size of $5 \times M$ will be obtained accordingly. Then, a 1-D intensity distribution with a length of M, which can be regarded as an approximation of sampling of $|\mathcal{F}_x(E)|^2$, can be constructed by summing all row vectors of the data matrix, which is shown in the right side of Fig. 6(a). The reason why the zone with a size of $5 \times M$ should be measured instead of $1 \times M$ is that the energy of the diffractive light focuses on the zones centered on each light spot instead of ideal points without area because of the diffraction limit. Summing all row vectors of the $5 \times M$ data matrix can concentrate energy on a line with a size of $5 \times M$, it can correct the 1-D intensity distribution to some extent.

Then, as shown in Fig. 6(b), the measured 1-D intensity distribution will be regarded as an approximation of sampling of $|\mathcal{F}_x(E)|^2$, and as the input of phase retrieval algorithm. Before inputting the 1-D intensity distribution into the algorithm, the lower-order light spot should be shifted to periphery of the distribution from the center, the result is shown in the left side of Fig. 6(b). The output of the algorithm can be regarded as an approximation of phase distribution of E along x-axis. φ can be obtained by measuring average amplitude of Fig. 6(b).

Recovered phase distribution obtained by this method can be regarded as a sampling of phase distribution of light field E on a certain finite zone along x-axis. Assuming Δx_0 is the spatial distance between a pair of adjacent sampling points in the recovered phase distribution, which is given by

$$\Delta x_0 = \frac{\lambda f}{M \Delta x},\tag{18}$$

where Δx is the spatial distance between a pair of adjacent sampling points in the observed intensity distribution (i.e., the pixel spacing along x-axis of camera). It can be deduced from (18) that decreasing f or increasing M are conducive to improve the resolution of the recovered phase distribution when λ and Δx are fixed. However, in practice, there are two factors to be considered, which limit the resolution. The first factor comes from the paraxial approximation, which requires a larger focal length f for obtaining smaller diffraction angles, since the paraxial approximation is relatively accurate when the diffraction angle is small. The second factor is the limited dynamic range of camera, which requires a smaller sampling length M, since the intensities of light spots decay rapidly with the increase of their orders. This factor will be discussed in detail in Section IV.

By using the grayscale-voltage relationship described in the initialized LUT, the voltage-phase characteristic curve can be measured from the grayscale-phase relationship obtained in advance. Then, a new LUT corresponding to the new working wavelength can be established.

IV. RESULT AND DISCUSSION

A. Result of Calibration

In this section, an experiment of calibration is carried out to illustrate the effeteness of the calibration method demonstrated in this paper. The reflective LC-SLM to be calibrated is produced



by Meadowlark Optics. It has 1920×1152 modulation units, the pixel pitch is $9.2 \ \mu m \times 9.2 \ \mu m$, the size of the working area is 17.6 mm × 10.7 mm. The camera used in the experiment is produced by HAMAMATSU, it has 2048×2048 pixels, with a pixel length of 6.5 μ m. The light source is a He-Ne laser with a wavelength of 633 nm, the focal length of lens, *f*, is 50 mm. Other parameters for phase retrieval algorithm used in experiments are: $K = 8000, \eta = 0.004, \rho = 0.99, M_0 = 190$ and M = 512. In this case, according to (18), the resolution of the recovered phase distribution is $\Delta x_0 = 9.51 \ \mu$ m, which is close to the pixel pitch of LC-SLM. Besides, the period of the binary mask (i.e., the periodic stripe mask) uploaded is 16 pixels.

Fig. 7(a) shows the relationship between φ and the grayscale of white stripes, which is obtained by using our calibration method. The curve is normalized to ensure that the maximum theoretical value of φ is π to compensate the measurement error. Besides, there is a problem of phase wrapping, i.e., φ and $\varphi + k\pi$ ($k \in \mathbb{Z}$) cannot be distinguished because they will produce the same diffraction pattern in this case. Fig. 7(b) shows the unwrapped result of the curve that is shown in the Fig. 7(a).

Based on data shown in Fig. 7(b), a new LUT is established to complete the calibration. Fig. 7(c) shows the grayscale-phase curve after the calibration and unwrapping. It can be seen that





Fig. 8. (a) 2-D diffraction intensity distribution of a binary grating; (b) 1-D intensity distribution of (a) on the x-axis; (c) the effective part of (b).

the curve is approximately a straight line. Root mean square error (RMSE) between the curve shown in Fig. 7(c) and the ideal straight line is 0.0247π .

B. Impact of Limited Dynamic Range on the Calibration

The proposed method use a camera to measure the diffraction intensity distribution of a binary grating. However, in this case, the energy of diffractive light is mostly concentrated on several light spots with low diffraction orders. That means the power difference between low diffraction orders and high diffraction orders may be so large that might exceed the dynamic range of the camera. In Section III, the limited dynamic range of the camera is regarded as a factor that limits the spatial resolution of the recovered phase distribution. In fact, as shown in the following, this point needs to be further clarified since it is not always the case if the specific conditions are selected in the experiment.

Fig. 8(a) shows a 2-D diffractive intensity distribution of a binary grating we used in the calibration. The intensity distribution is captured by the camera. Besides, the 1-D intensity distribution of Fig. 8(a) on the x-axis is shown in Fig. 8(b). To explain the problem more clearly, in Fig. 8(b), the coordinate of x-axis is in pixels, where the coordinate of y-axis is in grayscales. In order to obtain the intensity distribution accurately, the exposure time of the camera must be set to an appropriate value so that the most powerful diffraction order (zero-order in Fig. 8) will not be overexposed. We can see that the grayscale of the zero-order is about 27000, meaning that there is no overexposure in Fig. 8 (note that the bit depth of our camera is 16-bit, so the grayscale is up to 65535), even there is a bit of an underexposure.

If the intensity of a light spot is less than the noise intensity, the light spot will be invisible in the image. It can be seen in Fig. 8(b) that the light spots with high diffraction orders are invisible since the exposure time is too short for them. Therefore, the effective length of the 1-D intensity distribution is limited. This is the limitation caused by the limited dynamic range. We should use an effective part of the intensity distribution to calibrate the LC-SLM, where all the diffraction orders should be visible. The effective part of Fig. 8(b) is shown in the Fig. 8(c), of which the length is 512 pixels. We can see that all the diffraction orders in this part is visible, and intensities of most of them is much greater than the noise intensity, so we can use this part to recover the phase distribution accurately because of the high signal-to-noise ratio of this part. This is the reason why M is set to 512 pixels in the previous experiment.

According to the discussion above, the limited dynamic range will limit the effective length of the 1-D diffraction intensity distribution we can use. On the other hand, the spatial resolution of the recovered phase distribution, i.e., Δx_0 , is dependent on the length of the intensity distribution, as what shown in (18). Under the parameters of the previous experiment, it can be calculated from (18) that Δx_0 is about 9.51 μ m, which is close to the pixel pitch of LC-SLM (9.2 μ m), meaning that the recovered phase distribution can provide enough information to measure the phase modulation of LC-SLM.

In summary, the limited dynamic range limits the effective length of the diffraction intensity distribution, and the limited effective length of the diffraction intensity distribution limits the spatial resolution of the recovered phase distribution further. Considering the fact that the spatial resolution is acceptable, we think the influence of limited dynamic range on the calibration is within an acceptable range.

C. Effect of Crosstalk on Calibration

It is relatively convenient to measure the phase of light from far-field diffraction pattern because the pattern can be collected by setting up a simple optical path. Some calibration methods, such as [12] and [13], adopted similar idea used in this paper to measure the voltage-phase characteristic curve. But these methods introduced priori hypothesis about the distribution of phase modulation offered by LC-SLM, it might cause error when physical conditions do not accord with the hypothesis. For example, [13] demonstrates a calibration method that can also be described by Fig. 5, but they used a simple equation, as expressed by (19), to obtain φ from the far-field diffraction pattern instead of the phase retrieval algorithm. Let *I* be the intensity of zero order light spot in the far-field diffraction pattern of *E* described



Fig. 9. Effect of crosstalk between pixels.

by (16), then the relationship between *I* and φ can be expressed as

$$I = \frac{I_0}{2} \left(1 + \cos \varphi \right), \tag{19}$$

where I_0 is the intensity of the zero-order light spot when φ is 0. (19) means that φ can be obtained from *I* directly. But it should be emphasized that (19) is based on a fundamental hypothesis that the complex distribution of light field could be described by (16) precisely. Therefore, there exists an inevitable error if the complex distribution does not satisfy with such a hypothesis completely. In fact, some studies [15], [19] have pointed out that (16) is imprecise to describe the light field when a striped mask is uploaded to LC-SLM due to the crosstalk between pixels. The effect of the crosstalk is schematically shown in Fig. 9. When a white-black stripe mask is uploaded, it is reasonable that there is a continuous change in the real phase modulation on the boundary of white stripes and black stripes along x-axis, instead of an ideal step change manner which is described by (16). The crosstalk is caused by leakage electric field between different modulation unit of LC-SLM, i.e., the electric field upon a control electrode will influence the distribution of electric field upon other control electrodes.

The recovered phase distributions along *x*-axis for those masks with different grayscales of white stripes being uploaded onto the LC-SLM but without calibration are shown in Fig. 10. The result points out that phase distribution cannot change suddenly on the boundary between black stripes and white stripes, especially while the phase modulation is relatively low. It can be seen from Fig. 10 that the change of phase modulation will be more acute for large grayscale difference between black and white stripes, it is reasonable since large difference of control voltage will cause drastic change of phase modulation. In this point of view, the voltage-phase characteristic curve measured by (19) is inaccurate to some degree since the phase distribution shown in Fig. 10 cannot be described by (16) accurately.

To verify the influence of the crosstalk effect on the traditional calibration method, the actual relationships between I and φ



Fig. 10. Recovered phase distributions under different grayscale of white stripes without calibration.



Fig. 11. Grayscale-intensity curves of the zero-order light spot under different period T of the binary mask.

under different period of the binary mask before calibration is measured, which is shown in Fig. 11, where *T* is the period in pixel and the normalized intensity is I/I_0 . It can be seen in Fig. 11 that the height of the second peak of the grayscale-intensity curves, corresponding to the case of $\varphi = 2\pi$, is lower than the first peak, meaning that the intensity-phase relationship cannot satisfy the cosine law that described by (19), especially when the period *T* is low. Obviously, the effect of crosstalk increases with the decrease of the period, because the height of the second peak is decrease with the decrease of the period, which cannot be described by (19) precisely for the traditional method.

In the general, for the traditional method, people usually believe that the second peak of the grayscale-intensity curve is corresponding to the intensity under $\varphi = 2\pi$ and calibrate LC-SLM based on (19) and the data between the first peak and the second peak. Besides, in aspect of data processing, the second peak being lower than the first peak is inevitable, so one should normalize the curve to make the height of the second peak equal to the first [12]. It is obvious that there are two flaws in the traditional method. The first is that the second peak may not correspond to the case of $\varphi = 2\pi$, which can be seen in Fig. 11 (we can see the second peak is shift to right with the decrease of



Fig. 12. (a) Grayscale-phase curves before calibration measured by traditional method under different period T and proposed method under T = 16; (b) grayscale-phase curves after calibration.

T). The second flaw is the normalization of the curve. There is no scientific basis to make the height of the second peak equal to the first peak artificially. The two flaws will introduce error into the calibration. Essentially, the two flaws are caused by the crosstalk effect of LC-SLM.

For comparison, Fig. 12 shows the grayscale-phase curves before and after calibration with the traditional and the proposed method. Fig. 12(a) shows five curves that measured by traditional method under different period T, and a curve measured by the proposed method under T = 16. Based on the six curves before calibration, six LUTs is established. But only the LUT based on the proposed method is used to calibrate the LC-SLM in practice. After the calibration, the grayscale-phase curve is remeasured, which is shown in Fig. 12(b) (i.e., the curve with legend of "Proposed: T = 16" in Fig. 12(b)). Based on the grayscale-phase curve after calibration of proposed method and the difference between the six LUTs, we estimated the gravscalephase curves after calibration of the traditional method, which are also shown in Fig. 12(b). We calculated RMSEs between the curves shown in Fig. 12(b) and the ideal straight line, which are shown in Table II. It can be seen in Table II that the RMSE of traditional method increases with the decrease of period T basically. The increasing RMSE of traditional method reflects

 TABLE II

 RMSEs of Calibration Result of Traditional Method Under

 Different Period T and the Proposed Method Under T = 16

Method and period T	RMSE
Traditional: $T = 32$	0.0355π
Traditional: $T = 16$	0.0354π
Traditional: $T = 8$	0.0347π
Traditional: $T = 4$	0.0513π
Traditional: $T = 2$	0.0617π
Proposed: $T = 16$	0.0247π

the influence of crosstalk effect on calibration. For comparison, the RMSE of the proposed method is always less than the traditional method, because the proposed method has ability to avoid the potential error caused by the crosstalk effect inherently. According to the data shown in Table II, a conclusion can be made that compared with the traditional method, the RMSE of the proposed method is reduced by 30% under the same condition (the same condition refers to T = 16).

According to the discussion above, the phase retrieval algorithm can obtain more accurate phase modulation of the LC-SLM. On one hand, this method avoids the potential error caused by crosstalk since it introduces less priori hypothesis about the phase distribution and the data obtained in the experiment for instantiating the voltage-phase characteristic curve inherently involve the effect of crosstalk, meaning that the recovered phase is consistent with the phase distribution on the LC-SLM. On the other hand, this method obtains φ by analyzing information about several light spots on the diffraction pattern instead of only using one of them in the traditional calibration method. Besides, it should be noted that the spatial resolution of the calibration method can access the level of pixel size, which has been demonstrated in the experiment. In general, traditional diffraction-based calibration method cannot access this level of spatial resolution [20]. Meanwhile, the calibration method we proposed also has the advantages of the traditional diffraction pattern analyzed method, i.e., the convenience and the anti-noise. On the aspect of convenience, all the data used in calibration can be collected by a simple optical path shown in Fig. 5 without sophisticated instruments and optical setup, even the data analysis can be completed automatically by programming. On the aspect of anti-noise, the 1-D phase retrieval algorithm has good ability of anti-noise, which has discussed in Section II. In summary, our method not only keeps the advantages of traditional diffraction-based calibration method, but also has high spatial resolution, which allows us to avoid the potential error caused by the crosstalk effect.

V. CONCLUSION

In conclusion, a calibration method of LC-SLM is proposed in this paper based on a 1-D phase retrieval algorithm. The algorithm for measuring the voltage-phase characteristic curve uses the gradient descent algorithm with RMSprop to solve the 1-D phase retrieval problem. Numerical simulations indicate that the algorithm is stable and has the ability of anti-noise. Experiments demonstrate that the calibration method inherently avoid potential error caused by the crosstalk between pixels since recovered phases are consistent with phase distributions on LC-SLM.

REFERENCES

- Y. N. Cai et al., "Rapid tilted-plane Gerchberg-Saxton algorithm for holographic optical tweezers," *Opt. Exp.*, vol. 28, no. 9, pp. 12729–12739, 2020.
- [2] K. S. Malik and B. R. Boruah, "Optimal trap velocity in a dynamic holographic optical trap using a nematic liquid crystal spatial light modulator," *J. Opt.*, vol. 24, no. 3, 2022, Art. no. 034004.
- [3] Y. Shen, E. C. Martinez, and C. Rosales-Guzman, "Generation of optical skyrmions with tunable topological textures," *Amer. Chem. Soc. Photon.*, vol. 9, no. 1, pp. 296–303, 2022.
- [4] H. Li et al., "Optimized dual spatial light modulators holographic display based on wavefront frequency decomposition," *Opt. Laser Technol.*, vol. 149, 2022, Art. no. 107866.
- [5] R. Yoshizaki et al., "High-efficiency microdrilling of glass by parallel transient and selective laser processing with spatial light modulator," *Opt. Laser Technol.*, vol. 154, 2022, Art. no. 108306.
- [6] J. Strobelt et al., "Optical microstructure fabrication using structured polarized illumination," *Opt. Exp.*, vol. 30, no. 5, pp. 7308–7318, 2022.
- [7] P. Rajaeipour, K. Banerjee, A. Dorn, H. Zappe, and Ç. Ataman, "Cascading optofluidic phase modulators for performance enhancement in refractive adaptive optics," *Adv. Photon.*, vol. 2, no. 6, 2020, Art. no. 066005.
- [8] A. Bergeron, J. Gauvin, F. Gagnon, D. Gingras, H. H. Arsenault, and M. Doucet, "Phase calibration and applications of a liquid-crystal spatial light modulator," *Appl. Opt.*, vol. 34, no. 23, 1995, Art. no. 5133.
- [9] J. Fuentes, E. J. Fernandez, P. M. Prieto, and P. Artal, "Interferometric method for phase calibration in liquid crystal spatial light modulators using a self-generated diffraction-grating," *Opt. Exp.*, vol. 24, no. 13, pp. 14159–14171, 2016.
- [10] L. Yang, J. Xia, C. L. Chang, X. B. Zhang, Z. M. Yang, and J. H. Chen, "Nonlinear dynamic phase response calibration by digital holographic microscopy," *Appl. Opt.*, vol. 54, no. 25, pp. 7799–7806, 2015.

- [11] H. X. Zhang, J. Zhang, and L. Y. Wu, "Evaluation of phase-only liquid crystal spatial light modulator for phase modulation performance using a Twyman-Green interferometer," *Meas. Sci. Technol.*, vol. 18, no. 6, pp. 1724–1728, 2007.
- [12] X. Chen et al., "A calibration algorithm for the voltage-phase characteristic of a liquid crystal optical phased array," *Proc. SPIE*, vol. 9795, pp. 10–15, 2015.
- [13] Z. Zhang, G. Lu, and T. S. Francis, "Simple method for measuring phase modulation in liquid crystal televisions," *Opt. Eng.*, vol. 33, no. 9, pp. 3018–3022, 1994.
- [14] R. J. Li, Y. H. Gao, and L. C. Cao, "In situ calibration for a phase-only spatial light modulator based on digital holography," *Opt. Eng.*, vol. 59, no. 5, 2020, Art. no. 053101.
- [15] E. Ronzitti, M. Guillon, V. de Sars, and V. Emiliani, "LCoS nematic SLM characterization and modeling for diffraction efficiency optimization, zero and ghost orders suppression," *Opt. Exp.*, vol. 20, no. 16, pp. 17843–17855, 2012.
- [16] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *IEEE. Trans. Inf. Theory*, vol. 61, no. 4, pp. 1985–2007, Apr. 2015.
- [17] E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Commun. Pure Appl. Math.*, vol. 66, no. 8, pp. 1241–1274, 2013.
- [18] T. Tieleman and G. Hinton, "Lecture 6.5—RMSprop: Divide the gradient by a running average of its recent magnitude," *Coursera: Neural Netw. Mach. Learn.*, vol. 2, pp. 26–31, 2012.
- [19] J. E. Stockley, D. Subacius, and S. A. Serati, "Influence of the interpixel region in liquid crystal diffraction gratings," in *Proc. Int. Soc. Opt. Eng.*, vol. 3635, pp. 127–136, 1999.
- [20] R. J. Li and L. C. Cao, "Progress in phase calibration for liquid crystal spatial light modulators," *Appl. Sci.*, vol. 9, no. 10, 2019, Art. no. 2012.
- [21] J. E. Wolfe and R. A. Chipman, "Polarimetric characterization of liquidcrystal-on-silicon panels," *Appl. Opt.*, vol. 45, no. 8, pp. 1688–1703, 2006.
- [22] F. J. Martínez et al., "Averaged stokes polarimetry applied to evaluate retardance and flicker in PA-LCoS devices," *Opt. Exp.*, vol. 22, no. 12, pp. 15064–15074, 2014.
- [23] T. Vipin, Y. Pandey, and N. S. Bisht, "Spatially addressable polarimetric calibration of reflective-type spatial light modulator using Mueller-Stokes polarimetry," *Front. Phys.*, vol. 9, 2021, Art. no. 709192.