Bandwidth Control in Computer-Generated Holography

Mian Wu, Lin Wu, Ying Qiu, Chao Yang, Zhixue He, Jin Tao, and Xi Xiao

***Abstract*—The Fourier transform between the time and frequency signals of light enables highly efficient band usage in optical communications. Analogously, the Fourier transform between the wavefield and wavevector distributions in the optical spatial diffraction is useful to understand the bandwidth considered in the diffraction simulations. In this work, we thoroughly investigate the numerical techniques for controlling the bandwidth in the design of computer-generated holography. The experiments reveal the effects of bandwidth control on both accuracy improvement of diffraction equations and suppression of higher hologram orders. A general formula for controlling the bandwidth directly using the spatial-frequency components is concluded.**

***Index Terms*—Bandwidth control, computer-generated holography, diffractive formula, fast fourier transform.**

I. Introduction

NUMERICAL calculation of the light's propagation in free space is of fundamental interest to many practical applications, such as digital holography [1], diffraction tomography [2], diffractive deep neural network [3], design of optical elements [4], and computer-generated holography (CGH) [5]. The Rayleigh-Sommerfeld (RS) diffraction formula paves the way for rigorous wavefield propagation simulations [6], [7], [8], [9]. Two numerical techniques are widely employed to solve the RS formula, as: angular spectrum (AS) and direct integration (DI) methods [10], [11], [12]. The calculations of both methods are based on a sampling of the analogue continuous field. The dense sampling of the input field and the diffraction equation imposes computational burden. To accelerate the calculations, the diffraction simulations are performed by using the fast Fourier transform (FFT) [13] in practical. The use of the FFT enables the manipulation of the wavevectors in the spatial-frequency domain.

The optical spatial diffraction is analogous to the temporal dispersion, namely: the conversions between the temporal and spectral profiles of a light pulse can be realized by the Fourier transform [14]. In optical communication networks, the signal processing in the time and frequency domain has led to several spectrally-efficient systems, such as orthogonal frequency division multiplexing (OFDM), Nyquist wavelength division multiplexing (Nyquist-WDM), and Nyquist optical time division multiplexing (Nyquist-OTDM) [15], [16]. Whereas, the applications of the signal processing for the optical spatial diffraction in the spatial and spatial-frequency domains have not been widely studied.

One application of controlling the signal bandwidth in the spatial-frequency domain is to improve the accuracy of the diffraction calculations. The FFT-based AS method has poor computing accuracy for long propagation distance due to the sampling issue [17], [18]. Several means have been proposed to address this problem, such as: the band-limited AS method by setting the spatial-frequencies outside the effective bandwidth to 0 [19], [20], the band-extended AS method by compressing the spatial-frequency sampling intervals to allow more sampling spatial-frequencies to be within the bandwidth [21], [22], [23], and the adaptive-sampling AS method to fully utilize the space-bandwidth product [24]. Another application of the diffraction signal processing lies in the field of CGH. The spatial-frequency bandwidth control can be realized optically by inserting a spatial filter in the Fourier plane of a 4-f configuration. This method can improve the quality of the CGH by eliminating the higher diffraction orders [25]. Recently, it has been demonstrated that the same effects can be realized digitally by embedding the diffraction bandwidth control into the iterative phase-retrieval-type algorithm [26], [27], [28], [29]. Thus, the noise in the CGH can be suppressed without the use of the specific optical filters. The above two applications study the diffraction bandwidth control in the diffraction equations and the CGH design algorithms separately. The effects of various bandwidth control methods on the CGH in terms of simultaneous accuracy improvement of diffraction equations and suppression of higher diffraction orders have not been clearly studied.

Manuscript received 11 June 2023; revised 24 October 2023; accepted 7 November 2023. Date of publication 13 November 2023; date of current version 24 November 2023. This work was supported in part by the National Natural Science Foundation of China under Grants 62005225 and 62275208, in part by the Natural Science Foundation of Hubei Province under Grant 2021CFA044. *(Mian Wu and Lin Wu contributed equally to this work.) (Corresponding authors: Lin Wu; Jin Tao.)*

Mian Wu, Lin Wu, Ying Qiu, and Chao Yang are with the State Key Laboratory of Optical Communication Technologies and Networks, China Information Communication Technologies Group Corporation (CICT), Wuhan 430074, China (e-mail: wumian_nanooptic@qq.com; wulin@wri.com.cn; yqiu@wri.com.cn; yangchao@wri.com.cn).

Zhixue He is with Peng Cheng Laboratory (PCL), Shenzhen 518055, China (e-mail: hezhx01@pcl.ac.cn).

Jin Tao is with the State Key Laboratory of Optical Communication Technologies and Networks, China Information Communication Technologies Group Corporation (CICT), Wuhan 430074, China, and also with Peng Cheng Laboratory (PCL), Shenzhen 518055, China (e-mail: taojin@wri.com.cn).

Xi Xiao is with the State Key Laboratory of Optical Communication Technologies and Networks, China Information Communication Technologies Group Corporation (CICT), Wuhan 430074, China, and with the National Information Optoelectronics Innovation Center, Wuhan 430074, China, and also with Peng Cheng Laboratory (PCL), Shenzhen 518055, China (e-mail: xxiao@wri.com.cn).

Digital Object Identifier 10.1109/JPHOT.2023.3332024

In this work, we summarize and derive the available numerical methods for manipulating the bandwidth in the spatial-frequency domain, and employ them for the CGH experiments. The bandwidth control is realized by changing the sampling intervals and numbers for the input and output signals in the spatial and spatial-frequency domain. The sampling changes are enabled by the scaled or nonuniform FFT (SFFT) [30], [31], [32], which allows the sampling properties in the spatial and spatial-frequency domains to be independent to each another. It is found that by controlling the bandwidth considered in the calculations, the quality of reconstructed hologram, the number of presented diffraction orders and the separation between hologram replicas can be tailored. The schemes of controlling the diffraction bandwidth via the modifications of the spatial and spatial-frequency components can improve the CGH performance for various optical applications [33], [34]. The CGH also provides a platform to visualize the effects of the parameter changes in the diffraction calculation on the computing bandwidth and accuracy.

Section II presents the basic formula for the AS and DI methods. For both methods, the identical and different sampling intervals in the input and output planes are calculated using the FFT and SFFT, respectively. Section III describes the indirect bandwidth control by changing the sampling properties in the spatial domain. The experimental results are explained via the band properties in the spatial-frequency domain. Section IV describes the methods of manipulating the spatial-frequency components for the direct bandwidth control. Finally, the equation relating the diffraction/signal bandwidth to all the controllable numerical parameters is concluded and a conclusive comparison of all methods is presented.

II. Basic Concept

The scalar diffraction of the monochromatic wave in an isotropic and homogeneous medium can be described by the scalar Helmholtz equation, as:

$$\left(\nabla^2 + k^2\right) U \;=\; 0 \tag{1}$$

where ∇^2 is the Laplacian, k is the wave vector, and U is the light field distribution. The RS diffraction integral is an exact solution to the scalar wave equation in (1) under Dirichlet boundary conditions [35]. It can be numerically calculated by either of two methods: the AS and DI [10].

The AS method models the light propagation in the spatial-frequency domain, where the light field is interpreted as the combination of a series of planewaves with various propagation directions. The method first transforms the light field in the source plane $(\alpha, \beta, 0)$ to the spatial-frequency domain, as:

$$A_0\left(u, v\right) = \iint_{-\infty}^{\infty} U_0\left(\alpha, \beta, 0\right) \exp\left[-i2\pi\left(u\alpha + v\beta\right)\right] d\alpha d\beta \tag{2}$$

in which, u, v are the spatial frequencies. After the wave propagates a distance of z, the field at the destination plane (x, y, z) parallel to the source plane can be given as:

$$U\left(x, y, z\right) = \iint_{-\infty}^{\infty} A_0\left(u, v\right) \exp\left[i2\pi\left(ux + vy\right) + ikz\sqrt{1 - \left(\lambda u\right)^2 - \left(\lambda v\right)^2}\right] dudv \tag{3}$$

In (3), the transfer function of free space is:

$$H_B\left(u, v, z\right) = \exp\left[ikz\sqrt{1 - \left(\lambda u\right)^2 - \left(\lambda v\right)^2}\right] \tag{4}$$

The DI method considers the wave propagation in the spatial domain, where light is treated as the coherent superposition of spherical sub-waves. The corresponding propagation formula is:

$$U\left(x, y, z\right) = U_0\ \left(\alpha, \beta, 0\right) * h\left(x, y, z\right) \tag{5}$$

In (5), the symbol $*$ denotes the two-dimensional (2D) convolution, and $h(x, y, z)$ is the impulse response (the propagation kernel), which is given by:

$$h\left(x, y, z\right) = \frac{1}{2\pi}\ \frac{z}{r}\left(\frac{1}{r} - ik\right)\frac{exp\left(ikr\right)}{r}, \tag{6}$$

in which, $r\ = \sqrt{x^2 + y^2 + z^2}$. (6) is obtained by the transfer function in (4) by using:

$$h\ \left(x, y, z\right) = \mathcal{F}^{-1}\ \left\{H_B\left(u, v, z\right)\right\} \tag{7}$$

where $\mathcal{F}^{-1}$ represents the inverse Fourier transform.

In the numerical calculations, the input and output wavefields are conventionally sampled using an equidistant grid, and the resulted discrete equations can be effectively calculated by means of FFT and inverse FFT (IFFT). However, the FFT-based methods only allow for identical sampling rates in the input and output planes. For more general cases, where the sampling intervals in the input and output planes are independent, the SFFT is required. In the following, the FFT-based and SFFT-based AS (FFT-AS and SFFT-AS) and DI (FFT-DI and SFFT-DI) methods are described.

A. AS Methods

The FFT-AS method samples the input light field $U_0(\alpha, \beta, 0)$ to a $N \times N$ equidistant grid, so the discrete version of (2) can be given as:

$$A_0\ \left(u_g, v_h\right) = \sum_{s=-N/2}^{N/2-1}\ \sum_{t=-N/2}^{N/2-1} U_0\left(\alpha_s, \beta_t, 0\right) \times \exp\left[-i2\pi\left(u_g\alpha_s + v_h\beta_t\right)\right]\Delta\alpha\Delta\beta \tag{8}$$

where *g, h, s, t* are the integers ranging from $-N/2$ to $N/2 - 1$, $\Delta\alpha, \Delta\beta$ are the sampling intervals on the source plane $U_0(\alpha_s, \beta_t, 0)$. The sampling intervals of the spatial frequencies u_g, v_h are:

$$\Delta u\ = \frac{1}{N\Delta\alpha}\ ,\ \Delta v\ = \frac{1}{N\Delta\beta} \tag{9}$$

The calculation of (8) can be accelerated by a two-dimensional (2D) FFT, namely,

$$A_0\left(u_g, v_h\right) = \ \mathrm{FFT}\left\{U_0\left(\alpha_s, \beta_t, 0\right)\right\} \tag{10}$$

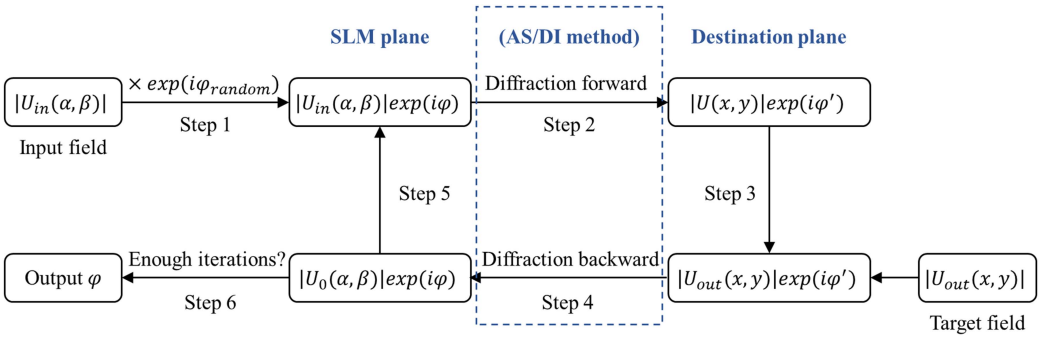

Fig. 1. Flow chart of GS algorithm.

Likewise, (3) can be written in a FFT form as:

$$U(x_m, y_n, z) = \mathrm{IFFT}\left\{\mathrm{FFT}\left\{U_0(\alpha_s, \beta_t, 0)\right\} \cdot \times H_B(u_g, v_h, z)\right\} \quad (11)$$

where m, n are the integers ranging from $-N/2$ to $N/2-1$, the symbol $\cdot\times$ denotes the element-by-element multiplication. The discrete transfer function in (11) is given by:

$$H_B(u_g, v_h, z) = \exp\left[ikz\sqrt{1-\left(\lambda\frac{g}{N\Delta\alpha}\right)^2-\left(\lambda\frac{h}{N\Delta\beta}\right)^2}\right] \quad (12)$$

To independently set the input and output sampling intervals, the SFFT is adopted, where a scale parameter q is introduced as: $(\Delta x, \Delta y) = q(\Delta\alpha, \Delta\beta)$. The scaled counterpart of the discrete Fourier transform in (8) writes:

$$A_0^{(q)}(u_g, v_h) = \sum_{s=-N/2}^{N/2-1}\sum_{t=-N/2}^{N/2-1} U_0(\alpha_s, \beta_t, 0) \times \exp\left[-i2\pi\frac{(u_g\alpha_s + v_h\beta_t)}{q}\right]\Delta\alpha\Delta\beta \quad (13)$$

Importantly, the scale parameter q needs to satisfy $|1/q| \le 1$ in (13) [30]. Therefore, the SFFT-AS method has two forms:

$$U(x_m, y_n, z) = \begin{cases} IFFT\left\{\mathrm{SFFT}\left\{U_0(\alpha_s, \beta_t, 0)\right\} \cdot \times H_B(u_g, v_h, z)\right\}, & q > 1 \\ ISFFT\left\{\mathrm{FFT}\left\{U_0(\alpha_s, \beta_t, 0)\right\} \cdot \times H_B(u_g, v_h, z)\right\}, & q \le 1 \end{cases} \quad (14)$$

where, SFFT and ISFFT denote the scaled fast Fourier transform and inverse scaled fast Fourier transform, respectively. The sampling intervals of the spatial frequencies u_g, v_h are:

$$\Delta u = \frac{1}{N\max(\Delta\alpha, \Delta x)}, \quad \Delta v = \frac{1}{N\max(\Delta\beta, \Delta y)} \quad (15)$$

So, the transfer function H_B is:

$$H_B(u_g, v_h, z) = \exp\left\{ikz\left[1-\left(\lambda\frac{g}{N\max(\Delta\alpha, \Delta x)}\right)^2 - \left(\lambda\frac{h}{N\max(\Delta\beta, \Delta y)}\right)^2\right]^{1/2}\right\} \quad (16)$$

In (14), the SFFT has been given by (13), and the ISFFT is expressed as:

$$U^{(q)}(x_m, y_n, z) = \sum_{s=-\frac{N}{2}}^{\frac{N}{2}-1}\sum_{t=-\frac{N}{2}}^{\frac{N}{2}-1} A_0 \cdot \times H_B \exp\left[i2\pi q\left(u_g x_m + v_h y_n\right)\right]\Delta u\Delta v\,. \quad (17)$$

(17) has sense when $|q| \le 1$.

Previously, the FFT-AS and SFFT-AS methods have been verified by calculating the diffraction patterns from known apertures [10], [30], [31]. Here, we examine these formulas using the CGH, which is designed by the Gerchberg-Saxton (GS) algorithm [36]. By changing the sampling intervals on the destination plane in the GS algorithm, we verify that the size of the hologram can be varied.

GS algorithm was first proposed in 1971 to solve the phase recovery problem. The flow chart of GS algorithm is shown in Fig. 1. Firstly, a random phase φ_{random} is superimposed onto the input amplitude distribution $|U_{in}(\alpha, \beta)|$. The light field on the SLM plane is then propagated to the destination plane using the AS/DI methods. Thirdly, the calculated amplitude distribution $|U(x, y)|$ is substituted with the target output field amplitude $|U_{out}(x, y)|$. The modified light field on the destination plane then propagates to the SLM plane by the AS/DI methods. If the number of iterations is enough for convergence, the phase distribution φ is returned. Otherwise, the calculated amplitude distribution $|U_0(\alpha, \beta)|$ is then replaced by the input amplitude distribution $|U_{in}(\alpha, \beta)|$. This completes the first iteration. Steps 2–5 are repeated until iterates enough times. The numerical diffraction methods are embedded in Step 2 and Step 4.

The experimental setup is depicted in Fig. 2. The CGH is enabled by a phase-only spatial light modulator (SLM) based on the liquid crystal on silicon (LCoS) technology. The wavelength of the incident laser is 1550 nm, and its polarization is controlled by a 3-paddle fiber polarization controller. The collimated plane wave covers 420 × 420 pixels on the SLM, and the size of each pixel is 8 × 8 μm^2. The propagation distance between the SLM and the camera is 97 mm, which is in a medium distance range for better experimental holography of all methods in this paper.

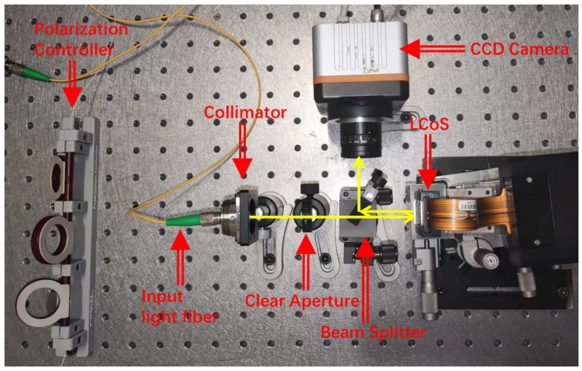

Fig. 2. Experimental setup for the computer-generated holography. The light path is indicated by the yellow line.

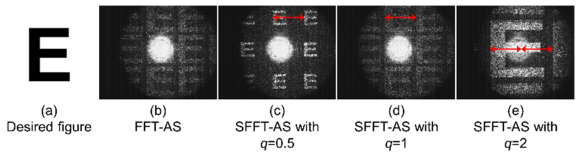

Fig. 3. Comparison of the experimental CGH results using the FFT-AS and SFFT-AS methods with various q values and the constant input sampling interval.

Here, we use the letter "E" as the target output light field, as shown in Fig. 3(a). The FFT-AS and SFFT-AS methods are employed in a GS algorithm to retrieval the phase diagrams. In the GS algorithm, the backward propagation formula is obtained by taking the conjugate of the transfer function in (12) or (16). Fig. 3(c)–(e) show the captured holograms generated using the SFFT-AS method with different output sampling intervals, i.e., the scale parameters q. The holograms generated using the FFT-AS and the SFFT-AS with the scale factor $q = 1$ have the same size, as shown in Fig. 3(b) and (d). The resulted images can be enlarged by increasing q.

From Fig. 3, the sampling interval changes also influence the spacing between the hologram replicas. This has been theoretically studied for the case of the Fresnel diffraction [37]. The spectral-based and direct calculations of the diffraction formula cause the separations between the hologram replicas of $1/\Delta u$ and $\lambda z/\Delta\alpha$, respectively. Here, λ is the wavelength. For the AS method, according to (15), the spacing between the replicas is $1/\Delta u = N\max(\Delta\alpha, \Delta x)$. In Fig. 3(c) and (d), $\Delta x \leq \Delta\alpha$, so increasing Δx will not change the sampling rate in the spatial-frequency domain Δu, the spacing between the replicas is retained. In Fig. 3(d) and (e), $\Delta x \geq \Delta\alpha$, so doubling Δx leads to halving of Δu and a two-fold increase in the spacing between the replicas. Although the result in Ref. [37] is obtained from the Fresnel diffraction, the relation between the spacing between the hologram copies and the sampling interval in the input plane is still valid for the RS formula, as shown in Fig. 3.

B. DI Methods

For the FFT-DI method, the discrete convolution for (5) can be written as:

$$U\ (x_m, y_n, z) = \sum_{s=-N/2}^{N/2-1} \sum_{t=-N/2}^{N/2-1} U_0\left(\alpha_s, \beta_t, 0\right) \times h\left(x_m - \alpha_s, y_n - \beta_t, z\right) \Delta\alpha\Delta\beta \quad (18)$$

Unlike (11), (18) cannot be directly calculated by a FFT and an IFFT, as the discrete Fourier transform involves periodicity in both Fourier and real spaces, which makes the convolution a circular one. To linearize the convolution, the matrices in (18) need to be properly zero padded. Ref. [10] provides an implementation of the FFT-DI method to calculate (18), as:

$$U' = \text{IFFT}\left\{\text{FFT}\left\{U_0^{'}\right\} \cdot \times \text{FFT}\left\{h'\right\}\right\} \quad (19)$$

In (19), the source field is zero padded to a matrix with the size of $(2N-1)\times(2N-1)$, as:

$$U_0^{'} = \begin{bmatrix} U_0 & 0 \\ 0 & 0 \end{bmatrix}_{(2N-1)\times(2N-1)} \quad (20)$$

The corresponding impulse response becomes:

$$h' = \begin{bmatrix} h\left(X_1, Y_1, z\right) & \cdots & h\left(X_1, Y_{2N-1}, z\right) \\ \vdots & \ddots & \vdots \\ h\left(X_{2N-1}, Y_1, z\right) & \cdots & h\left(X_{2N-1}, Y_{2N-1}, z\right) \end{bmatrix}_{(2N-1)\times(2N-1)} \quad (21)$$

$$X_j = \begin{cases} x_1 - \alpha_{N+1-j},\ j = 1,\ \ldots,\ N-1 \\ x_{j-N+1} - \alpha_1,\ j = N,\ \ldots,\ 2N-1 \end{cases} \quad (22)$$

$$Y_j = \begin{cases} y_1 - \beta_{N+1-j},\ j = 1,\ \ldots,\ N-1 \\ y_{j-N+1} - \beta_1,\ j = N,\ \ldots,\ 2N-1 \end{cases} \quad (23)$$

The output window U' should be clipped and reduced to the original size, this clipping depends on the position of the input window [20]. As the origin input field is at the left upper corner, the desired field on the output plane $U(x, y, z)$ is the right lower quadrant of U', as:

$$U' = \begin{bmatrix} \ddots & \vdots \\ \cdots & U_{N\times N} \end{bmatrix}_{(2N-1)\times(2N-1)} \quad (24)$$

The SFFT-DI method has been proposed in Ref. [32]. Similar to the FFT-DI method, the zero-padding is also required. Apart from the calculations for the case of $\Delta x \neq \Delta\alpha$ and $\Delta y \neq \Delta\beta$, in this work, we show that the method in Ref. [32] can also be used for the more general case of the different sampling points on the source and destination planes. We define the sampling points along both vertical and horizontal axis on the input and output planes are N_1 and N_2, respectively. The source field U_0 is zero padded to a new matrix $U_0^{'}$ with the size of $M \times M$, and M is defined as:

$$M = \max\left(N_1, N_2\right) \cdot \left(1 + \frac{\min\left(\Delta\alpha, \Delta x\right)}{\max\left(\Delta\alpha, \Delta x\right)}\right) \quad (25)$$

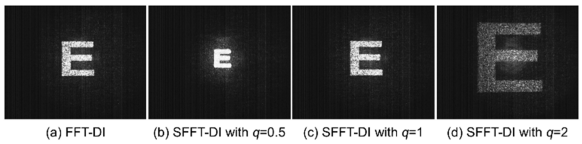

Fig. 4. Comparison of the experimental CGH results using the FFT-DI and SFFT-DI methods with various q values and the constant input sampling interval.

The SFFT-DI method is expressed as:

$$U' = \begin{cases} IFFT\left\{\text{SFFT}\left\{U_0'\right\} \cdot \times \text{FFT}\left\{h'\right\}\right\}, & q > 1 \\ ISFFT\left\{\text{FFT}\left\{U_0'\right\} \cdot \times \text{FFT}\left\{h'\right\}\right\}, & q \leq 1 \end{cases} \tag{26}$$

where h' is the transfer function derived from (6) and (21), with the difference of:

$$r = \sqrt{\left[g \cdot \max\left(\Delta\alpha, \Delta x\right)\right]^2 + \left[h \cdot \max\left(\Delta\beta, \Delta y\right)\right]^2 + z^2} \tag{27}$$

Here, the integers g and h are from $-M/2$ to $M/2-1$. The desired output field is clipped from U' following the same rule described in Ref. [20]. In the GS algorithm using the FFT-DI, the backward propagation formula is obtained by taking the conjugate of the FFT of the impulse response. Fig. 4 illustrates the CGH results using the FFT-DI and SFFT-DI methods. The captured holograms from the two methods have the same size when $q = 1$, as shown in Fig. 4(a) and (c). The size of the diffraction images can be tailored by changing the output sampling intervals, i.e., the scale parameter q.

Comparing with the results obtained using the AS methods, the zero-order beam is largely weakened, and the holograms in Fig. 4 generally have better quality than those in Fig. 3. This is due to that the DI method is more accurate than the AS method for long propagation distance. The AS method can be modified for expanding its applicable propagation range [20], [21], [22], [23], which will be discussed in Section IV.

In this work, we employ the same algorithm for solving the SFFT and ISFFT as that in Ref. [32], the detail is described in Appendix A in Ref. [32]. Other methods for calculating the scaled or nonuniform Fourier transform can be found in Ref. [38].

III. Indirect Bandwidth Control Using Spatial Components

Section II shows that the SFFT can change the size of hologram compared to the FFT by changing the sampling intervals in the destination plane. Besides, by changing the sampling intervals in the source plane, the resulted holograms can be tailored due to the bandwidth changes [26], [27], [28], [29]. In this section, we show the indirect bandwidth control by modifying the spatial components in the input plane using the SFFT-AS and SFFT-DI numerical calculations and CGH experiments.

A. Comparison of the SFFT-AS Method With Various Input Sampling Intervals

Fig. 5(a), (b) show the holograms generated by using the SFFT-AS method with various input sampling intervals

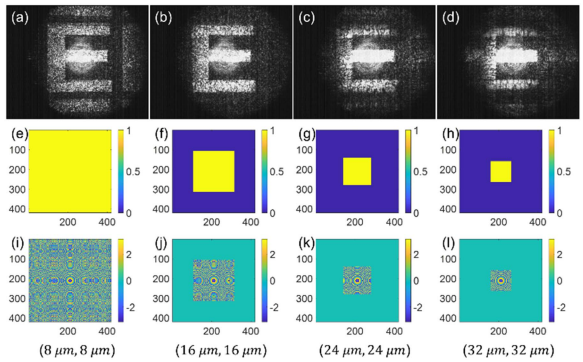

Fig. 5. Comparison of the experimental CGH results using the SFFT-AS method with various input sampling intervals. The output hologram size is (420 × 16) × (420 × 16) μm^2. (a)–(d) The resulted holograms. (e)–(h) The sampling numbers. (i)–(l) The transfer functions.

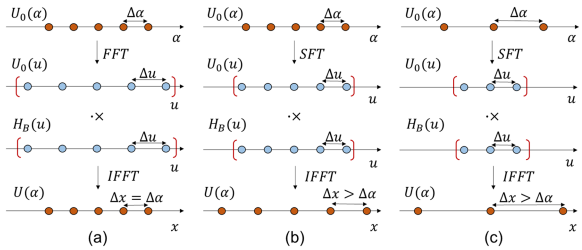

Fig. 6. Pipeline of (a) the FFT-AS method, (b) the SFFT-AS method with the increased output interval and (c) the SFFT-AS method with the increased input interval.

$(\Delta\alpha, \Delta\beta)$ and the identical output size. In the experiments, the input sampling rate can only be multiples of the pixel size of the SLM - 8 μm. The phase diagram on the SLM consists of supercells with multiple pixels according to the designed interval. By increasing $\Delta\alpha$ and $\Delta\beta$ from 8 to 16 μm, the high-order copies are eliminated. By further increasing the input sampling intervals, the area of the output images is narrowed till the hologram is incomplete. These can be interpreted by the bandwidth limitations for the CGH [20], [25], [26], [27], [28], [29]. As the size of the incident beam on the SLM is unchanged, increasing the input sampling intervals will reduce the sampling numbers in both the spatial and spatial-frequency domain, which is equivalent to applying a spatial filter as shown in Fig. 5(e)–(h). As a result, the transfer functions in (16) in the spatial-frequency domain are shown in Fig. 5(i)–(l).

It worth noting that the positions of the images remain unchanged when changing the input or output sampling rate using the SFFT-AS method. This is different from the results in Ref. [30], where the image is relocated to the center of output field by a correct term in the transfer function. This is due to the employment of a different SFFT method from Ref. [32].

Fig. 6 illustrates the calculating pipelines of the two AS methods. The spatial-frequency interval is given by (15). Accordingly, the effective bandwidth B is given as:

$$B = N\Delta u = \frac{1}{\max\left(\Delta\alpha, \Delta x\right)} \tag{28}$$

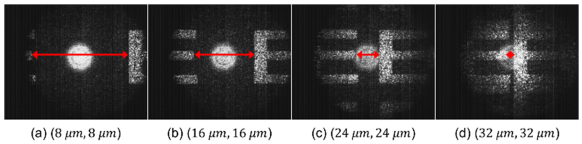

Fig. 7. Comparison of the experimental CGH results using the SFFT-DI method with various input sampling intervals. The output sampling interval is 12 μm, the output sampling matrix has the size of 430 $\times$ 430.

According to (28), increasing either Δx or $\Delta\alpha$, the effective bandwidth considered in the calculation will decrease. As depicted in Fig. 6(a) and (b), increasing Δx will decrease the bandwidth of the transfer function. Next, we increase $\Delta\alpha$ and retain the size of output images to show the change of B and the corresponding region size that can be calculated. In Fig. 6(c) Δx will change along with $\Delta\alpha$ in order to retain the size of output hologram, as the input and output sampling numbers are always the same in the AS method. By comparing Fig. 6(b) and (c), increasing $\Delta\alpha$ will decrease B, which lead to the reduction of the area that can be calculated, as shown in Fig. 5.

The analyses reveal that increasing the sampling rate in the spatial domain indeed limits the calculating bandwidth. Here, the input sampling intervals can only be an integral multiple of the SLM pixel. The more general and direct means to realize the band-limited AS method have been proposed [20], and will be discussed in Section IV-A.

B. Comparison of the SFFT-DI Method With Various Input Sampling Intervals

By using the SFFT-DI method, we can change both the sampling intervals and numbers. Here, we experimentally generate the holograms with various input sampling intervals. In the experiment, the input sampling number is 420, a different output sampling number of 430 is adopted, and the output sampling interval is 12 μm. Fig. 7 shows the experimental results, and the position of the hologram is shifted by adding a term to the axis in (27) to illustrate the change of the separation between hologram replicas.

From Fig. 7, the distance between the neighboring orders becomes smaller with the increase of the input sampling interval. From Ref. [37], when the diffraction is calculated by the DI method, the spacing between the replicas is $\lambda z/\Delta\alpha$. This inverse proportion relation between the spacing between the copies and the sampling interval in the input plane is verified by Fig. 7. Although the expressions for the spacing between replicas in Ref. [37] are derived from the Fresnel diffraction, they are still valid for the RS diffraction formula.

Fig. 8 illustrates the pipelines of the FFT-DI and SFFT-DI methods. In Fig. 8, the transfer function H_R is the discrete Fourier transform of the impluse response h'. Here, the bandwidth of H_R is not clipped, and the redundant sampling points are removed from the output matrix. Comparing with the band-limited AS methods in Fig. 6, more spatial-frequency components of the transfer function in the DI methods contribute to the diffraction calculation, leading to a higher computational

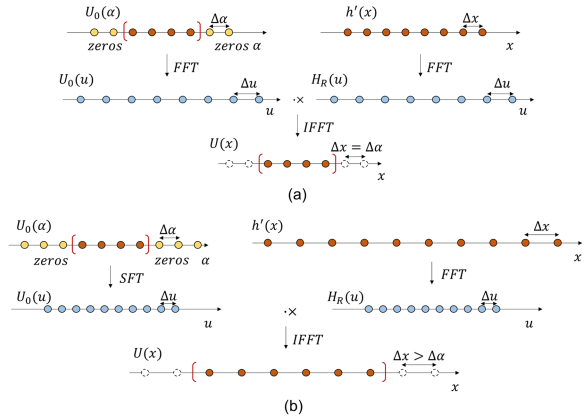

Fig. 8. Pipelines of (a) the FFT-DI and (b) the SFFT-DI methods with different sampling intervals and sampling numbers.

accuracy as shown in Fig. 4, more mathematical details referring to Table I. This calculating pipeline is similar to that of the band-extended AS methods [21], [22], [23]. More results on the band-extended AS methods are presented in Section IV-B. Fig. 8(b) also shows that both sampling intervals and sampling numbers can be modified in the SFFT-DI method.

IV. Direct Bandwidth Control Using Spatial-Frequency Components

In the above analyses, the SFFT-AS method can be regarded as a band-limited light propagation calculation, and the SFFT-DI is similar to a band-extended diffraction calculation. The band-limited and the band-extended AS methods are two typical means to increase the computational accuracy of the AS method for a wide propagation range. Whereas, the previous works only apply these methods to the FFT-based calculations. In this work, we embed the SFFT in both the band-limited and band-extended AS methods, and compare their performances in the application of CGH. The results show that the bandwidth is directly controlled by the spatial-frequency parameters in the band-limited and band-extended AS methods. By combining these two methods, the bandwidth can be directly controlled using three elements in the spatial-frequency domain, including the total sampling points M, the effective sampling points N' within the bandwidth and the sampling intervals Δu. In this section, we will describe the modifications of these three parameters.

A. Change of the Sampling Numbers in the Spatial-Frequency Domain

Fig. 9(a) depicts the pipeline of a more general method for performing the band-limited AS method than that shown in Fig. 6. Here, the input sampling number is extended to a user-defined parameter M by zero padding, then the transfer function H_B is also zero padded to meet the dimensions of the SFFT of the input field. The spatial-frequency interval and the effective bandwidth of the transfer function become:

$$\Delta u = \frac{1}{M\max\left(\Delta\alpha, \Delta x\right)} \tag{29}$$

TABLE I
SUMMARY OF ALL DIFFRACTION METHODS

Methods	$\Delta\alpha/\mu m$	q	$\Delta x'/\mu m$	M	$B/\mu m^{-1}$	$RMSE$	T/s	Diagrams
FFT-AS	Only for near-field diffraction, $q = 1$, $M = 420$, fast calculation							
	8	1	8	420	0.1250	0.5515	0.91	Fig. 3b
SFFT-AS	Only for near-field diffraction, variable q, $M = 420$, more calculation time than FFT-AS method							
	8	1	8	420	0.1250	0.5562	3.15	Fig. 3d
	8	2	16	420	0.0625	0.4349	3.14	Fig. 3e
FFT-DI	Only for far-field diffraction, $q = 1$, $M = 420$, more calculation time than FFT- and SFFT-AS methods							
	8	1	8	420	0.1250	0.2536	4.47	Fig. 4a
SFFT-DI	Only for far-field diffraction, variable q, variable M, more calculation time than FFT-DI method							
	8	1	8	430	0.1221	0.2337	12.38	Fig. 4c
	8	2	16	430	0.0610	0.5003	7.46	Fig. 4d
Band-limited SFFT-AS	For longer distance (than FFT-AS), variable q, variable M, calculate slower than SFFT-AS method with same M							
	8	2	16	500	0.0525	0.3606	7.03	Fig. 9b
Band-extended SFFT-AS	For both near- and far- field diffraction, variable q, variable M, variable $\Delta x'$, calculate slower than SFFT-AS method with same M							
	8	2	19	420	0.0526	0.3458	4.72	Fig. 10b
	8	2	16	1680	0.0625	0.3384	64.75	Fig. 11a

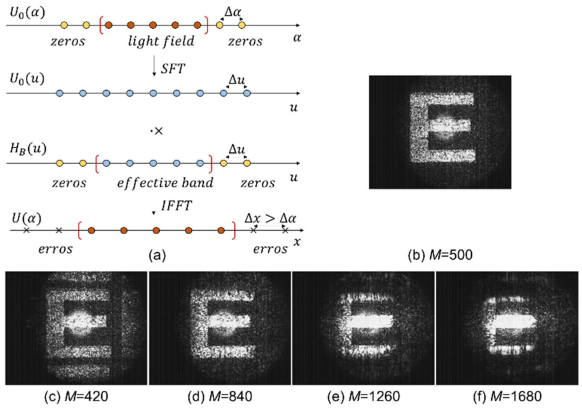

Fig. 9. (a) Pipeline of the band-limited SFFT-AS method based on zero-padding. (b)–(f) The experimental holograms with $N = 420$, $\Delta\alpha = 8\ \mu$m, $\Delta x = 16\ \mu$m, and various M.

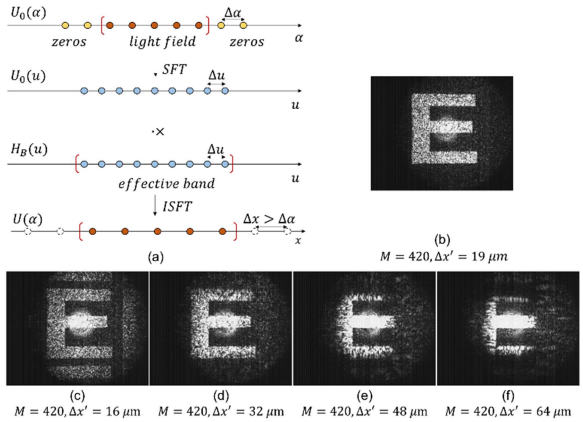

Fig. 10. (a) Pipeline of the band-extended SFFT-AS method. (b)–(f) The experimental holograms generated with $\Delta\alpha = 8\ \mu$m, $\Delta x = 16\ \mu$m, $N = M = 420$, and various $\Delta u(\Delta x')$.

$$B = N\Delta u = \frac{N}{M\max(\Delta\alpha, \Delta x)} \tag{30}$$

Different from (28), the bandwidth B can be tailored not only by the sampling rate in the input or output plane, but also by the user-defined dimension of the calculated matrices [39]. Fig. 9(b)–(f) show the holograms with various M. With the increase of M, the bandwidth considered in the diffraction calculation is narrowed, and an optimized hologram can be obtained. Different from the means of changing the input sampling rates in Section III-A, the value of M can be continuously varied. The bandwidth used in Fig. 9(c)–(f) is consistent with these in Fig. 5(a)–(d). The optimal hologram is realized when $M = 500$, as shown in Fig. 9(b), the corresponding bandwidth is 0.0525 μm^{-1}, which is not achievable by changing the input sampling intervals.

In Fig. 9(a), the sampling number of the transfer function is the same to the size of the output field. Next, we will discuss how to include more spatial-frequency components within the efficient bandwidth.

B. Change of the Sampling Intervals in the Spatial-Frequency Domain

Fig. 10(a) depicts the pipeline of the band-extended SFFT-AS method. Instead of forcing the transfer function components beyond a certain band to be zero, as show in Fig. 9(a), the band-extended method rearranges all the sampling points in the spatial-frequency domain into the required bandwidth. This is done by defining a new sampling interval for the transfer function H_B as:

$$\Delta u = \frac{1}{M\Delta x'} \tag{31}$$

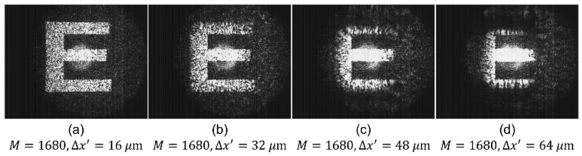

Fig. 11. Experimental holograms generated with the band-extended SFFT-AS method with $\Delta\alpha = 8\ \mu$m, $\Delta x = 16\ \mu$m, $N = 420$, $M = 1680$, and various $\Delta x'$.

As $\Delta x'$ can be different to both Δx and $\Delta\alpha$, both the SFFT and ISFFT are needed. Here, we vary $\Delta x'$ without zero padding, i.e., $M = N$. The experimental holograms are shown in Fig. 10(b)–(f). The bandwidths in Fig. 10(c)–(f) are designed to correspond to these in Fig. 5(a)–(d) and (c)–(f), and the similar results are obtained. As $\Delta x'$ can be continuously chosen, the optimal hologram is obtained at $\Delta x' = 19\ \mu$m. The corresponding bandwidth $M\Delta u = 1/\Delta x'$ is 0.0526 μm^{-1}, which is approximately equal to the optimal bandwidth in the band-limited SFFT-AS method shown in Fig. 9(b).

C. General Principle for Change of the Sampling Properties in the Spatial-Frequency Domain

By far, the modifications of the total sampling numbers and the sampling intervals of the wavevectors in the spatial-frequency domain have been discussed separately. Here, we simultaneously control them to show a general case as shown in Fig. 10(a). Different from Fig. 10(b)–(f), we pad zeros to the source field and increase M. The experimental results of the band-extended SFFT-AS method with $M = 1680$ and various $\Delta x'$ are shown in Fig. 11. The results show that increasing $\Delta x'$ will reduce the bandwidth used in the calculations. The hologram in Fig. 11(a) is slightly better than these generated by the band-limited AS method in Fig. 9, which confirms that the band-extended AS method retains more spatial-frequency components in the diffraction calculation, and thus, achieves a higher computational accuracy for the long propagation length than the band-limited AS method [21], [22], [23], referring to Table I. The band-extended SFFT-AS method can modify both $\Delta x'$ and M, which is responsible for setting the appropriate bandwidth and retaining more frequency components, respectively.

The bandwidth of the band-extended AS method in Fig. 10(a) is $1/\Delta x'$, which is irrelevant to M. This can be confirmed by the similarity between Fig. 10(d)–(f) and (b)–(c). However, Fig. 11(a) does not show high diffraction orders as Fig. 10(c) does. This is due to that increasing M will enlarge the spacing between the hologram copies, as the spacing between the hologram replicas produced by the AS method is $1/\Delta u = M\Delta x'$ [37].

Overall, both the total sampling numbers and the sampling intervals of the wavevectors in the spatial-frequency domain ($\Delta x'$ and M) can be continuously changed to control the bandwidth used in the calculations, in contrast, the input sampling intervals can only be chosen from the integral multiples of 8 μm constrained by LCoS, as shown in Section III. More generally, the number of spatial-frequency components within the effective bandwidth N' can also be freely chosen. As a result, the effective bandwidth can be set according to:

$$B = \frac{N'}{M\Delta x'} \tag{32}$$

where, M, N' and $\Delta x'$ can all be freely chosen as $N_1 \neq M \neq N' \neq N_2$ and $\Delta\alpha \neq \Delta x' \neq \Delta x$. This method of modifying the above three spatial-frequency components provides more flexible in optimizing the diffraction calculations for different application scenario.

D. Conclusive Comparison of All Methods

For better comparison, a summary table of all methods and some of the experimental results is presented in Table I. The calculation accuracy is defined as root mean squared error (*RMSE*) of the experimental CGH image and the desired figure in Fig. 3(a):

$$RMSE = \sqrt{\frac{1}{M^2}\sum_i\sum_j\left(U_e^{(i,j)} - U_d^{(i,j)}\right)^2} \tag{33}$$

where *M* represents the sampling number, U_e denotes the amplitude distribution matrix of the experimental CGH images corresponding to figures in the column "Diagrams", and U_d is the amplitude distribution matrix of the desired image "E". The calculation time *T* is investigated using a Lenovo computer with 1.7 GHz CUP and 32 GB of RAM and MATLAB R2017a.

As is shown in Table I, the DI methods produce nice results in the experimental distance of 97 mm and the sampling interval of 8 μm, but it is only accurate in long propagation distance, more complex for numerical calculation than AS methods, and with the same input and output sampling intervals and numbers. In practical experiments, the sampling intervals are also constrained by the pixel size of SLM.

The FFT-AS method is more widespread for its easy steps and fast calculation. The SFFT methods can have different input and output sampling intervals, and even different input and output sampling numbers. The band modulated AS methods compensate the distance limit of AS method without sacrificing much calculation complexity, and can direct optimize the bandwidth with several freedoms.

The numerical values of *RMSE* are relatively higher than expected. This is because hardware limitations, such as non-square light shaping by the collimator and 30 μm pixel pitch of CCD camera.

V. Conclusion

In this work, we systematically investigate the bandwidth control in the digital calculations for the diffraction pattern tailoring in the optical space. We first summarize the numerical methods for fast calculating the RS diffraction formula. By replacing the FFT in the conventional methods by the SFFT, the sampling intervals in the source and destination planes can be different, and the sampling rate in the spatial-frequency domain can also be independent to that in the spatial domain, i.e., $1/M\Delta u \neq \Delta\alpha \neq \Delta x$. Then, we change the sampling intervals in the spatial domain, which indirectly change the bandwidth

in the spatial-frequency domain, and subsequently change the performance of the CGH. Finally, the spatial-frequency components are modified to directly change the bandwidth. The effective bandwidth considered in the calculations can be increased by either increasing the sampling intervals in the spatial-frequency domain or decreasing the sampling number in the spatial-frequency domain. The sampling number changes can be realized by zero padding the sampling matrix, thus the sampling numbers used in the calculations can be larger than the actual size of the field needed to be calculated, i.e., $N_1 \neq M \neq N_2$. And the sampling intervals can be directly varied in the MATLAB codes. As a result, the computing accuracy of the diffraction equations and the suppression of higher diffraction orders can be both realized.

From the CGH experiments, the calculated bandwidth should be properly set to avoid both the losing of information due to small bandwidth and the presence of unwanted diffraction orders due to large bandwidth. Our experiments also confirm the previous theoretical results on the relation between the spacing among the imaging replicas and the sampling intervals in the spatial and spatial-frequency domains. The results also indicate that the number of the spatial-frequency components within the effective bandwidth has important impact on the hologram, as: the calculations with more spatial-frequency components not only produce a higher computational accuracy, but also cause a larger spacing between hologram replicas.

Our work provides a comprehensive study on the schemes of the bandwidth control by modifying the spatial and spatial-frequency components in the diffraction calculations. These methods can enhance the diffraction simulation accuracy and provides more freedoms of modulation for the CGH and various other optical image display and optical processing, all of which can be well realized by metasurface, a compact platform for arbitrarily complex operators without more geometries or optimization techniques.

REFERENCES

[1] F. Zhang, I. Yamaguchi, and L. P. Yaroslavsky, "Algorithm for reconstruction of digital holograms with adjustable magnification," *Opt. Lett.*, vol. 29, no. 14, pp. 1668–1670, Jul. 2004.

[2] S. S. Kou and C. J. R. Sheppard, "Image formation in holographic tomography," *Opt. Lett.*, vol. 33, no. 20, pp. 2362–2364, Oct. 2008.

[3] X. Lin et al., "All-optical machine learning using diffractive deep neural networks," *Science*, vol. 361, no. 6406, pp. 1004–1008, Jul. 2018.

[4] N. K. Fontaine, R. Ryf, H. Chen, D. T. Neilson, K. Kim, and J. Carpenter, "Laguerre-Gaussian mode sorter," *Nature Commun.*, vol. 10, no. 1, Apr. 2019, Art. no. 1865.

[5] A. E. G. Madsen, R. L. Eriksen, and J. Glückstad, "Comparison of state-of-the-art computer generated holography algorithms and a machine learning approach," *Opt. Commun.*, vol. 505, Feb. 2022, Art. no. 127590.

[6] J. A. C. Veerman, J. J. Rusch, and H. P. Urbach, "Calculation of the Rayleigh–Sommerfeld diffraction integral by exact integration of the fast oscillating factor," *J. Opt. Soc. Amer.*, vol. 22, no. 4, pp. 636–646, Apr. 2005.

[7] A. Wuttig, M. Kanka, H. J. Kreuzer, and R. Riesenberg, "Packed domain Rayleigh-Sommerfeld wavefield propagation for large targets," *Opt. Exp.*, vol. 18, no. 26, pp. 27036–27047, Dec. 2010.

[8] C. J. R. Sheppard, J. Lin, and S. S. Kou, "Rayleigh–Sommerfeld diffraction formula in k space," *J. Opt. Soc. Amer.*, vol. 30, no. 6, pp. 1180–1183, Jun. 2013.

[9] N. A. Ochoa, "Alternative approach to evaluate the Rayleigh-Sommerfeld diffraction integrals using tilted spherical waves," *Opt. Exp.*, vol. 25, no. 10, pp. 12008–12019, May 2017.

[10] F. Shen and A. Wang, "Fast-Fourier-transform based numerical integration method for the Rayleigh-Sommerfeld diffraction formula," *Appl. Opt.*, vol. 45, no. 6, pp. 1102–1110, Feb. 2006.

[11] D. Mas, J. Garcia, C. Ferreira, L. M. Bernardo, and F. Marinho, "Fast algorithms for free-space diffraction patterns calculation," *Opt. Commun.*, vol. 164, no. 4, pp. 233–245, Jun. 1999.

[12] J. Li, Z. Peng, and Y. Fu, "Diffraction transfer function and its calculation of classic diffraction formula," *Opt. Commun.*, vol. 280, no. 2, pp. 243–248, Dec. 2007.

[13] N. Delen and B. Hooker, "Free-space beam propagation between arbitrarily oriented planes based on full diffraction theory: A fast Fourier transform approach," *J. Opt. Soc. Amer.*, vol. 15, no. 4, pp. 857–867, Apr. 1998.

[14] A. W. Lohmann and D. Mendlovic, "Temporal filtering with time lenses," *Appl. Opt.*, vol. 31, no. 29, pp. 6212–6219, Oct. 1992.

[15] P. Guan et al., "All-optical ultra-high-speed OFDM to Nyquist-WDM conversion based on complete optical Fourier transformation," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 626–632, Jan. 2016.

[16] P. Guan et al., "Time lens based optical Fourier transformation for all-optical signal processing of spectrally-efficient data," *J. Lightw. Technol.*, vol. 34, no. 4, pp. 799–806, Feb. 2017.

[17] W. Zhao, C. Wei, C. Yuan, C. Chang, and J. Ma, and R. Zhu, "A flexible numerical calculation method of angular spectrum based on matrix product," *Opt. Lett.*, vol. 45, no. 21, pp. 5937–5940, Nov. 2020.

[18] T. Kozacki and K. Falaggis, "Angular spectrum-based wave-propagation method with compact space bandwidth for large propagation distances," *Opt. Lett.*, vol. 40, no. 14, pp. 3420–3423, Jul. 2015.

[19] S. Tian, L. Chen, and H. Zhang, "Optimized Fresnel phase hologram for ringing artifacts removal in lensless holographic projection," *Appl. Opt.*, vol. 61, no. 5, pp. B17–B24, Feb. 2022.

[20] K. Matsushima and T. Shimobaba, "Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields," *Opt. Exp.*, vol. 17, no. 22, pp. 19662–19673, Oct. 2009.

[21] X. Yu, T. Xiahui, Q. Y. Xiong, P. Hao, and W. Wei, "Wide-window angular spectrum method for diffraction propagation in far and near field," *Opt. Lett.*, vol. 37, no. 23, pp. 4943–4945, Dec. 2012.

[22] Y. - H. Kim et al., "Non-uniform sampling and wide range angular spectrum method," *J. Opt.*, vol. 16, no. 12, Nov. 2014, Art. no. 125710.

[23] W. Zhang, H. Zhang, and G. Jin, "Band-extended angular spectrum method for accurate diffraction calculation in a wide propagation range," *Opt. Lett.*, vol. 45, no. 6, pp. 1543–1546, Mar. 2020.

[24] W. Zhang, H. Zhang, and G. Jin, "Adaptive-sampling angular spectrum method with full utilization of space-bandwidth product," *Opt. Lett.*, vol. 45, no. 16, pp. 4416–4419, Aug. 2020.

[25] M. Agour, E. Kolenovic, C. Falldorf, and C. von Kopylow, "Suppression of higher diffraction orders and intensity improvement of optically reconstructed holograms from a spatial light modulator," *J. Opt. Soc. Amer.*, vol. 11, no. 10, Aug. 2009, Art. no. 105405.

[26] M. Gopakumar, J. Kim, S. Choi, Y. Peng, and G. Wetzstein, "Unfiltered holography: Optimizing high diffraction orders without optical filtering for compact holographic displays," *Opt. Lett.*, vol. 46, no. 23, pp. 5822–5825, Dec. 2021.

[27] X. Sui, Z. He, G. Jin, D. Chu, and L. Cao, "Band-limited double-phase method for enhancing image sharpness in complex modulated computer-generated holograms," *Opt. Exp.*, vol. 29, no. 2, pp. 2597–2612, Jan. 2021.

[28] L. Chen, S. Tian, H. Zhang, L. Cao, and G. Jin, "Phase hologram optimization with bandwidth constraint strategy for speckle-free optical reconstruction," *Opt. Exp.*, vol. 29, no. 8, pp. 11645–11663, Apr. 2021.

[29] L. Wu and Z. Zhang, "Domain multiplexed computer-generated holography by embedded wavevector filtering algorithm," *PhotoniX*, vol. 2, no. 1, Jan. 2021, Art. no. 250125.

[30] T. Shimobaba, K. Matsushima, T. Kakue, N. Masuda, and T. Ito, "Scaled angular spectrum method," *Opt. Lett.*, vol. 37, no. 19, pp. 4128–4130, Oct. 2012.

[31] T. Shimobaba et al., "Nonuniform sampled scalar diffraction calculation using nonuniform fast Fourier transform," *Opt. Lett.*, vol. 38, no. 23, pp. 5130–5133, Dec. 2013.

[32] V. Nascov and P. C. Logofătu, "Fast computation algorithm for the Rayleigh-Sommerfeld diffraction formula using a type of scaled convolution," *Appl. Opt.*, vol. 48, no. 22, pp. 4310–4319, Aug. 2009.

[33] Z. Huang and L. Cao, "High bandwidth-utilization digital holographic multiplexing: An approach using Kramers–Kronig relations," *Adv. Photon. Res.*, vol. 3, no. 2, Jan. 2022, Art. no. 2100273.

[34] Y. Li et al., "Spectrum sampling optimization for quantitative phase imaging based on Kramers–Kronig relations," *Opt. Lett.*, vol. 47, no. 11, pp. 2786–2789, Jun. 2022.
[35] R. K. Tyson, "Theory of optical propagation and diffraction," in *Principles and Applications of Fourier Optics*, 1st ed. Charlotte, USA: IOP Publishing, 2014, pp. 1–1–1-6.
[36] H. Wang, W. Yue, Q. Song, J. Liu, and G. Situ, "A hybrid gerchberg–Saxton-like algorithm for DOE and CGH calculation," *Opt. Lasers Eng.*, vol. 89, pp. 109–115, Feb. 2017.
[37] D. P. Kelly, "Numerical calculation of the Fresnel transform," *J. Opt. Soc. Amer.*, vol. 31, no. 4, pp. 755–764, Apr. 2014.
[38] L. Greengard and J. - Y. Lee, "Accelerating the nonuniform fast Fourier transform," *SIAM Rev.*, vol. 46, no. 3, pp. 443–454, Jan. 2004.
[39] W. Zhang, H. Zhang, and G. Jin, "Frequency sampling strategy for numerical diffraction calculation," *Opt. Exp.*, vol. 28, no. 26, pp. 39916–39932, Dec. 2020.