Optics and Lasers in Engineering 191 (2025) 108994

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

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

High-quality binary amplitude hologram generation for digital micromirror device based holographic display

Alejandro Velez-Zea *, César Antonio Hoyos-Peláez, John Fredy Barrera-Ramírez

Grupo de Óptica y Fotónica, Instituto de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín, Colombia

ARTICLE INFO

Keywords:
Computer generated holography
Amplitude hologram
Gradient descent
Digital micromirror device
Fourier transform

ABSTRACT

In this work, we demonstrate a high-performance binary amplitude hologram generation method and verify its performance with an experimental holographic projection scheme based on binary amplitude modulation by means of a digital micromirror device. The proposed method consists of a subsampling procedure applied to a target scene prior to hologram generation with binarized stochastic gradient descent (B-SGD). In combination with temporal multiplexing, this approach leads to holographic reconstructions with significantly improved quality when compared to direct application of the B-SGD and other common binary amplitude hologram generation methods. Furthermore, our proposal achieves this improved quality with nearly the same computational speed as the conventional B-SGD. Numerical results confirm the validity of our proposal, and an experimental holographic projection scheme is implemented to demonstrate its feasibility.

1. Introduction

Since the first demonstrations of computer-generated holography (CGH) by Brown & Lohmann in 1966 [1], the potential of this technique for high-performance, high-fidelity display applications, and precise control of optical fields has been evident. Thanks to significant research and the rapid advancements in high-performance computing and spatial light modulation, CGH has found applications in many fields such as optogenetics [2–4], optical trapping [5,6], and visualization [7,8], to name a few.

CGH generation consists in calculating backward propagation of light from a virtual scene to a defined plane. The field in this plane is the computed hologram. Depending on the complexity of the scene and the target application, this computation can be very time-consuming, and the resulting hologram is usually a complex-valued function. From this hologram, a 3D scene can be numerically reconstructed by propagating the hologram. However, for many of the mentioned applications of CGH, experimental reconstruction of computed holograms is required. This imposes a critical limitation: there is no direct method to achieve complex light modulation for introducing a complex CGH into an experimental setup. In practice, programmable spatial light modulator (SLM) devices are used to experimentally reconstruct CGHs. However, these devices are restricted to either phase-only or amplitude-only modulation. Furthermore, they present limited resolution and relatively large pixel sizes. This significantly limits the size and quality of the CGHs that can be experimentally reconstructed. Furthermore, encoding a scene into a phase-only or amplitude-only hologram requires specialized algorithms with increased computational cost compared to complex CGH generation, and typically results in lower quality. This is because a phase-only or amplitude-only hologram generally can only produce an approximation of the target scene. Thus, in most cases, CGH techniques face a trade-off between reconstruction quality and computation speed.

Most research efforts in CGH have been directed towards generation of phase-only holograms. This is because phase modulation offers superior diffraction efficiency compared to amplitude modulation. The most well-known approach for phase-only hologram generation is the Gerchberg-Saxton (GS) algorithm [9], an iterative projection method where the phase connecting an input plane and its Fourier transform plane can be calculated by imposing amplitude constraints in each plane. This algorithm was later extended to calculation of Fresnel phase-only holograms of 2D [10], 3D [11], and multiplane objects [12–15]. Despite its broad usefulness, the GS algorithm and its variations are prone to stagnation, leading to holograms of limited quality. On the other hand, improved phase CGH methods are based on optimization algorithms like stochastic gradient descent (SGD) [16]. The application of these optimization approaches to phase holography was challenging due to the need for the calculation of gradients of

* Corresponding author.
E-mail address: alejandro.velezz@udea.edu.co (A. Velez-Zea).

https://doi.org/10.1016/j.optlaseng.2025.108994
Received 25 November 2024; Received in revised form 2 April 2025; Accepted 3 April 2025

complex-valued functions. However, the proposal of Wirtinger holography [17], where gradient calculus is performed by means of Wirtinger derivatives, has helped overcome this issue, leading to vastly improved phase holograms.

Beyond these iterative approaches, there is also a broad range of non-iterative phase-only hologram generation methods, each with its own limitations. A broadly used approach is double-phase holography [18], where a complex hologram is encoded into two interleaved phase-only holograms. This approach offers good quality, with the drawback of halving the effective resolution of the hologram. More recently, a similar approach named binary amplitude encoding has been proposed, offering improved performance and diffraction efficiency, and replacing the resolution trade-off with localized noise dependent on the aspect ratio of the hologram [19]. Another approach is based on the optimization of carrier waves, such as the use of optimized random phases [20,21] or quadratic phases [22,23]. In these approaches, the target scene is illuminated with a specially tailored wave to ensure that the result after backpropagation is as close as possible to a phase-only hologram. The application of these methods depends on the characteristics of the scene, which limits their usefulness. Another method to quickly and accurately obtain phase holograms relies on the use of deep learning approaches [24]. These methods usually require many holograms obtained through a conventional approach to train convolutional neural networks. However, once training is completed, they can produce high quality holograms with significantly reduced computation time compared to conventional algorithms. Furthermore, the characteristics of the experimental setup can be included during training by means of a camera-in-the-loop setup [25], leading to some of the best holographic reconstructions experimentally demonstrated to date.

Nevertheless, the use of phase-only SLMs has significant drawbacks. High-performance phase-only SLMs are typically liquid crystal on silicon (LCOS) devices [26], which have slow response times, with frame rates limited to 60 Hz, and at most 180 Hz in some of the most advanced devices. Additionally, they are relatively costly, and as the pixel size becomes smaller, they suffer from pixel crosstalk, which represents a barrier to many applications. To overcome these issues, there has been growing interest in the usage of digital micromirror devices (DMDs) instead of LCOS-SLMs. DMDs are a type of micro-electromechanic system (MEMS) consisting of arrays of millions of micrometer sized mirrors that can pivot at high speeds between two positions. This movement enables binary amplitude modulation of the incoming light. DMDs offer significantly lower cost compared to LCOS-SLMs, often by an order of magnitude for similar resolution and pixel size, have framerates in the range of tens of kilohertz, and do not suffer from pixel crosstalk. These advantages have led to DMDs being used in a wide range of light processing applications, particularly in video projectors [27], fringe profilometry 3D scanners [28–31], beam shaping [32,33], and in holography [34–37].

However, binary amplitude modulation imposes a challenging constraint on CGH algorithms: it not only suffers from the comparatively lower efficiency of amplitude modulation, but it also introduces the effects of maximum quantization due to only two effective amplitude levels. Due to these issues, binary amplitude hologram (BAH) generation methods have received relatively less attention compared to phase CGH methods. An initial approach to BAH generation consists of the direct quantization of either phase or amplitude holograms [38,39]. This approach usually leads to significantly degraded reconstructions due to the introduction of strong quantization noise. To mitigate this issue, several different binarization methods have been studied to determine the best approach. Furthermore, although the direct binarization introduces severe noise, it can also be applied to optimized phase holograms and partially preserve their increased quality over non-optimized holograms [40]. Nevertheless, this approach has limited effectiveness due to the presence of noise but has the advantage of very fast calculation. As a method to produce higher quality BAH, we find direct binary search (DBS) [41]. In DBS, an initial random binary amplitude is propagated, emulating the reconstruction calculation of a CGH. Then, a quality metric between this reconstruction and the target scene is calculated. Afterwards, a pixel in the random binary amplitude is switched in value, and the reconstruction is recalculated. If the quality metric improves (i.e., the reconstruction is closer to the target scene), the value of the changed pixel is retained, and a new pixel is tested. Otherwise, the change in value is reverted before moving to the next pixel. This process is repeated until all pixels in the initial hologram are tested, and it is iteratively repeated several times until the desired quality is achieved. This approach can find the BAH that best approximates a target scene after reconstruction, but it is extremely slow. Some works have found methods to improve the convergence of DBS by means of block-wise processing [42] and altering the order in which pixels are processed [42,43]. However, the slow computation remains a significant limitation of these approaches.

Another useful method for generating BAH is error diffusion [44,45]. In this approach, a complex or amplitude hologram of a scene is binarized, and the error in the first pixel between the hologram before and after binarization is calculated. This error is distributed to the neighboring pixels by means of set of weights. The calculation is then repeated for the next pixel until the entire hologram is processed. Error diffusion requires significantly less computation time than DBS. However, it introduces high-frequency noise in the reconstruction that limits the effective size of the target scene. Several proposals have been made to improve the performance of error diffusion for BAH generation, by calculating optimal weight factors [46], or by using a specifically tailored carrier wave [47]. More recently, a binarized stochastic gradient descent (B-SGD) BAH generation method was proposed by Lee *et al.* [48]. The main roadblock to applying an SGD approach to BAH generation was the fact that a binarization operation is, by definition, a non-derivable function. This meant that calculating the gradients needed by the SGD optimization was not possible directly. Lee *et al.* solve this problem by replacing the undefined derivative of the sign operation, used for binarization, with a hard hyperbolic tangent function. This enables B-SGD, where an amplitude hologram is optimized to minimize the loss in the reconstruction before and after binarization. This approach shows very high-quality and faster computation compared to DBS and enables the use of a broad range of techniques previously developed for SGD generation of phase-only holograms.

However, despite the effectiveness of this approach, the maximum quality that can be achieved is still well below that of state-of-the-art phase-only hologram generation methods such as Wirtinger holography. For this reason, the authors of the B-SGD proposal still recommend the application of temporal multiplexing for speckle reduction. In holographic display applications, temporal multiplexing is a technique where several different holograms of the same target scene are projected sequentially, and the resulting reconstructions are averaged [49]. If the initial conditions used for hologram generation are different in each case, each reconstruction will present statistically independent speckle noise that will be averaged with this process, resulting in a speckle reduction proportional to the number of multiplexed holograms. Temporal multiplexing is effective; however, it demands additional computation time due to the need for the calculation of several holograms for each scene to be projected. This method also demands a fast SLM to ensure that the projection of each hologram is performed at a sufficient speed to be averaged by the user or recording medium. The B-SGD was demonstrated experimentally using a ferroelectric phase-only SLM configured for binary amplitude modulation by using adequate polarization. Ferroelectric SLMs are capable of frame rates of up several hundred Hz, faster than LCOS, but significantly slower than DMDs. This means that the full performance of the B-SGD on a true binary amplitude modulator such as a DMD, has yet to be tested.

Taking into account the previously discussed points, we propose an improved approach to synthetizing and displaying scenes using BAH. This method, called subsampling binarized stochastic gradient descent (SSB-SGD), involves dividing the target scene into multiple subsamples,

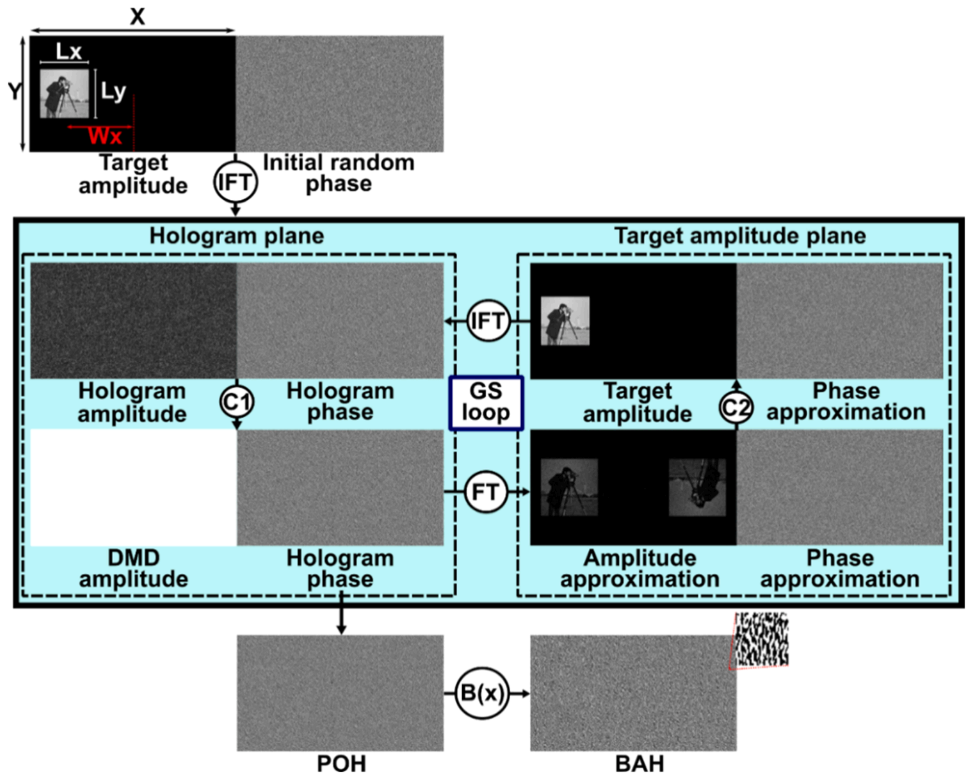

Fig. 1. Flowchart of the GS method for BAH generation (X and Y: target amplitude size; Lx and Ly: object size; Wx: distance from the center of the target amplitude to the center of the object; FT: Fourier transform; IFT: inverse Fourier transform; C1 and C2: applied constraints in during steps 2 and 4 of the B-GS algorithm; B(x): thresholding function).

each containing information from different pixels of the same scene. The BAH of each subsample is generated independently using the conventional B-SGD. After reconstructing the BAHs of all subsamples and applying temporal multiplexing, a significant increase in quality over the direct application of B-SGD with temporal multiplexing is achieved. This improvement arises because the sampling process in SSB-SGD enhances the convergence of the B-SGD algorithm and introduces a separation between pixels in the reconstructed object, thereby reducing speckle noise. Additionally, we implement a holographic projection setup based on a DMD to further validate and experimentally confirm the effectiveness of our proposed method.

2. Binary amplitude hologram generation methods

To demonstrate the advantages of our proposal compared to several well-known CGH methods, we will first introduce two common algorithms used for BAH generation. Then, we will explain the B-SGD, and finally, we will introduce our proposal. The first, and most straightforward method, is a threshold binarization of a random phase hologram. We will refer to this method as binarized one-step phase retrieval (B-OPR). This approach, while presenting the lowest overall accuracy, is notable for its computational simplicity. The algorithm is executed as follows

1. A target amplitude with dimensions X $\times$ Y is defined. This amplitude is a zero-valued array where the target object with dimensions Lx $\times$ Ly is placed at a distance Wx from the center.
2. The target amplitude is multiplied by an initial random phase.
3. The inverse Fourier transform (IFT) of the previous result is calculated to obtain the corresponding optical field in the hologram plane.
4. The amplitude information of the field in step 3 is discarded, preserving the phase.
5. The phase obtained in step 4 is converted into a BAH using direct global thresholding binarization with the following equation

$$B_f(x) = \left\{ \begin{matrix} 0 & if & 0 < x \leq \pi \\ 1 & if & \pi < x \leq 2\pi \end{matrix} \right\} \tag{1}$$

where the phase is in the range $[0, 2\pi]$. Several thresholding methods can be employed to generate a hologram at this step. For example, Pavel *et al.* [50] performed a comparative study of the different approaches to hologram binarization. The expression of Eq. 1 was selected since it represents the most straightforward approach to phase binarization within range $[0, 2\pi]$, effectively avoiding the need for iterative analysis of the phase histogram or the computationally intensive calculations involved in other binarization methods.

Although this algorithm can generate BAHs, its effectiveness is limited due to significant amounts of speckle noise affecting the reconstructed object, mainly caused by the binarization process. The primary value of this approach is to offer a baseline for comparison with more effective methods, and in cases where the visual quality of the reconstruction is not as critical as fast computation.

The next common approach for BAH generation is based on the threshold binarization of a Gerchberg-Saxton optimized phase-only holograms [9,51]. We will refer to this approach as binarized Gerchberg-Saxton method (B-GS). Its steps are described below

1. An initial field in the hologram plane is obtained by following steps 1-3 of the B-OPR algorithm described previously.
2. The first amplitude constraint (C1) is applied to the field in the hologram plane, by replacing the amplitude obtained in step 1 with a constant value while preserving the phase.
3. A Fourier transform (FT) is applied to the field resulting from the previous step, obtaining the field in the reconstruction plane.

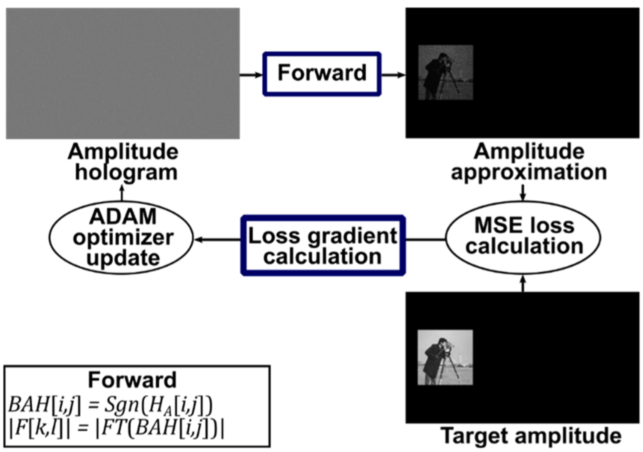

Fig. 2. Flowchart of the B-SGD for BAH generation. $H_A[i,j]$: amplitude hologram, $F[k,l]$: reconstructed field from binarized amplitude hologram, $[i,j]$: hologram plane pixel coordinates, $[k,l]$: reconstruction plane pixel coordinates.

4. A second constraint (C2) is applied, by replacing the amplitude in the reconstruction plane with the target amplitude while preserving the phase.
5. The IFT of the result of the previous step is calculated, obtaining once again the field in the hologram plane. Steps 2-5 are repeated iteratively, leading to a better approximation of the target amplitude in each iteration.

The GS cycle continues until a quality metric measured between the object region in the amplitude-approximated reconstruction and the target object reaches a level that satisfies a predetermined standard. After finalizing these iterations, the resulting phase in the hologram plane is extracted and converted into a BAH using the thresholding operation shown in Eq. 1. The scheme of this approach can be seen in Fig. 1.

The B-GS produces BAHs that can be reconstructed with significant lower speckle noise than the reconstructions from BAHs obtained with B-OPR. The reason for this higher effectiveness is that the B-OPR is essentially the same algorithm as the B-GS with a single partial iteration, meaning that no significant phase optimization is achieved.

However, there is still noticeable degradation. For this reason, to further improve the performance of BAH generation, Lee *et al.* introduced a high-performance binary hologram optimization framework that minimizes the quantization noise introduced when an amplitude hologram is binarized [48]. This approach is called binarized stochastic gradient descent (B-SGD).

In B-SGD, an initial amplitude hologram corresponding to a given target object is binarized using the signum (or sign) operation. This amplitude hologram can be obtained by means of sideband encoding. This can be achieved by following the first three steps of the B-OPR algorithm, as previously described. The real part of the resulting field in the hologram plane can be used as an amplitude hologram of the desired target amplitude.

After binarizing this amplitude hologram, the mean square error (MSE) loss function is calculated between the reconstructed object from the hologram before and after binarization. SGD is then used to minimize this loss, ensuring that the hologram reconstruction after binarization closely matches the target amplitude.

To implement this optimization, an initial amplitude hologram $H_A(x,y)$ is defined, where x and y are the spatial coordinates in the hologram plane. This amplitude hologram is converted into a BAH by applying the signum function, defined as:

$$Sgn(x) = \begin{cases} 1 & if\ x > 1 \\ -1 & if\ x < -1 \\ 0 & if\ x = 0 \end{cases} \tag{2}$$

The resulting BAH is then given by:

$$BAH(x,y) = Sgn(H_A(x,y)) \tag{3}$$

From this BAH, the reconstruction field is obtained via a Fourier transform:

$$F(v,w) = FT(BAH(x,y)) \tag{4}$$

Here, v and w are coordinates in the reconstruction plane.

Subsequently, the B-SGD algorithm is applied to the generation of the BAH in order to optimize $H_A(x,y)$, such that the corresponding amplitude of $F(v,w)$ minimizes the loss function, given by

$$L = \sum_{k=0}^{n} \sum_{l=0}^{m} \frac{1}{nm} (|F[k,l]| - A[k,l])^2 \tag{5}$$

Here, k and l are pixel coordinates in the reconstruction plane, $A[k,l]$ represents the target amplitude to be reconstructed from the hologram, and $m \times n$ is the hologram resolution. To minimize this loss, SGD calculates the gradient of Eq. 5 for the initial amplitude hologram obtained through sideband encoding. Then, during each iteration of the SGD optimizer, the amplitude hologram is updated according to the following equation

$$H_A{}^k = H_A{}^{k-1} - \alpha \frac{\partial L}{\partial H_A{}^{k-1}} \tag{6}$$

where k is an iteration counter, and α is a parameter that controls the magnitude of the modifications performed upon H_A during each iteration, commonly known as the learning rate. A very high learning rate can cause overshooting the H_A that minimizes the loss, while a low learning rate may require many iterations to achieve the desired optimization.

The primary limitation of the conventional SGD method applied for BAH generation is that the signum operation has zero gradient at all points except at zero. Consequently, the partial derivative component of Eq. 6 becomes zero. This vanishing gradient prevents the adequate updating of the amplitude hologram when applying Eq. 6. To avoid this issue, Lee *et al.* [48] proposed the B-SGD method, which replaces the gradient of the sign operation with the hard hyperbolic tangent function, defined as

$$\text{Hardtanh}(x) = \begin{cases} 1 & if\ x > 1 \\ -1 & if\ x < -1 \\ x & if\ -1 < x < 1 \end{cases} \tag{7}$$

By using this definition when calculating the gradient of the loss, we avoid the vanishing gradient problem and ensure an adequate operation for BAH generation. For the numerical implementation of B-SGD, modern programming languages have specialized libraries with a broad array of optimizers based on the SGD. These optimizers adjust the learning rate in Eq. 6 using additional hyperparameters, which are updated at each iteration. Among them, adaptative moment estimation (ADAM) is the most widely used [52]. In this work, although we continue to refer to the BAH generation method with gradient descent as B-SGD, we use an ADAM implementation for numerical calculations.

The B-SGD optimization significantly reduces the presence of speckle noise in the reconstructed binary hologram. Fig. 2 illustrates the process of using B-SGD to generate a BAH, which consists of the following steps:

1. During the forward propagation, the signum function is applied to the amplitude hologram, assigning each pixel's amplitude a value of -1 (negative), 0 (zero), or $+1$ (positive). A FT is applied to the

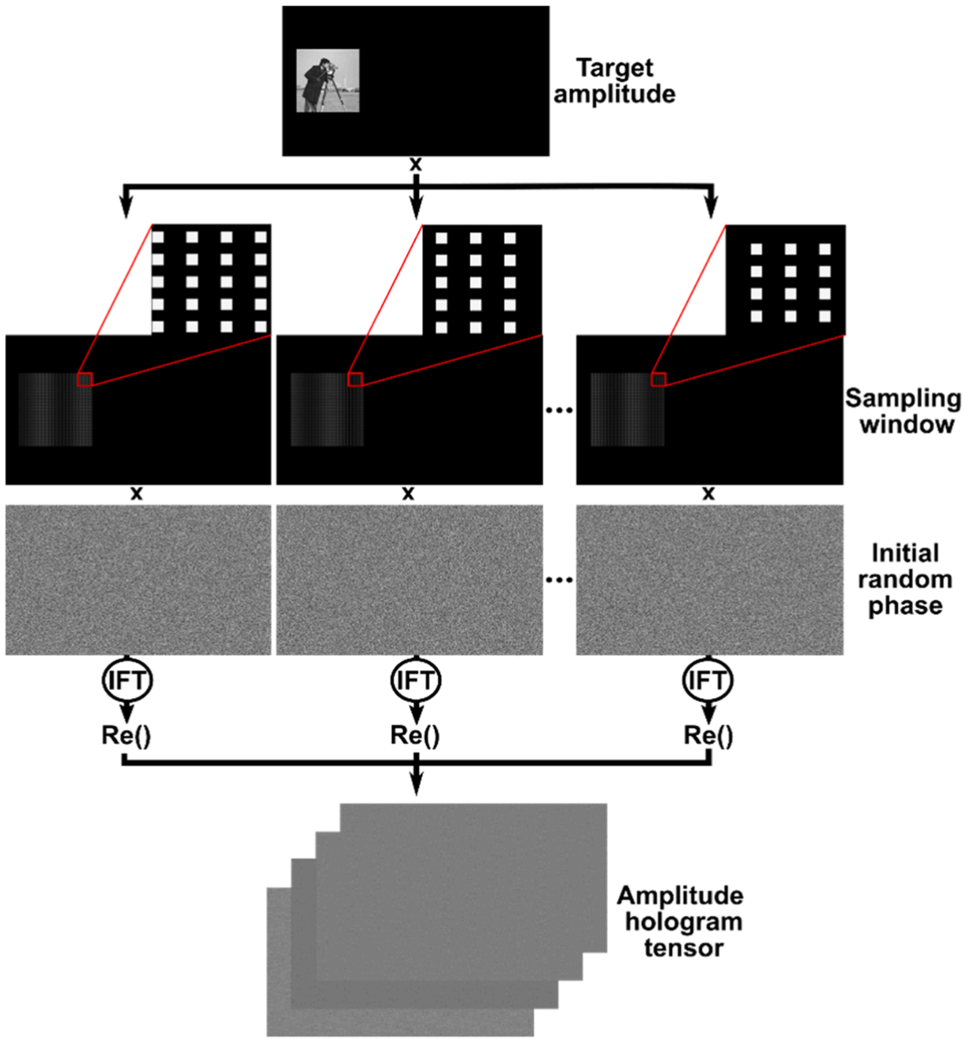

Fig. 3. Subsampling process to obtain the amplitude hologram tensor (IFT: inverse Fourier transform; Re(): real part of the complex field).

outcome and then squared to compute the intensity pattern, which serves as the predicted hologram reconstruction.

2. The MSE is used as a loss function, as defined in Eq. 5, to quantify the error between the forward propagation output and the target object.
3. The backward propagation employs the hard hyperbolic tangent (Eq. 7) function to compute gradients based on the loss function, facilitating a consistent gradient transition that is critical for subsequent optimization.
4. The optimization process, using ADAM [52], updates the hologram based on the insights gained from backward propagation, setting the stage for the ensuing epoch.

This cycle is repeated across multiple iterations, progressively enhancing the fidelity of the reconstructed object found in the amplitude approximation to closely match the target object. Once this is achieved, the sign function is applied to the final amplitude hologram to obtain the BAH.

While B-SDG produces significantly improved results compared to B-OPR and B-GS, the reconstructed object still maintains a degree of speckle noise. To deal with the effects of this remaining speckle noise, Lee *et al.* proposed the use of temporal multiplexing, averaging the reconstruction of several BAHs obtained with B-SDG, each initialized with different random phases, to achieve high-quality reconstructions. Although this approach is effective, the need for several BAHs to perform multiplexing increases the computational requirements of any application using this methodology.

To further enhance the effectiveness of the B-SGD approach in scenarios where temporal multiplexing is required, we introduce a novel alternative method which we call subsampling binarized stochastic gradient descent (SSB-SGD). This approach divides the target object into several subsamples to minimize the unwanted interference between nearby points of the reconstructed object, which is one of the main causes of speckle noise in holographic displays [53]. The SSB-SGD algorithm proceeds as follows:

1. A comb function is generated with samples every N pixels in the X direction and every M in the Y direction. A higher value of $N \times M$ requires more subsamples, increasing computation time but also leading to higher quality reconstructions after temporal multiplexing.
2. The target object is sampled with this comb function, obtaining a subsample. The comb is shifted one pixel in the X direction, and the target object is sampled again. This process is repeated shifting the comb function N times, and then the comb is shifted one pixel in the y direction, repeating the sampling process until $N \times M$ subsamples are obtained.
3. An amplitude hologram of each subsample is generated using side-band encoding.

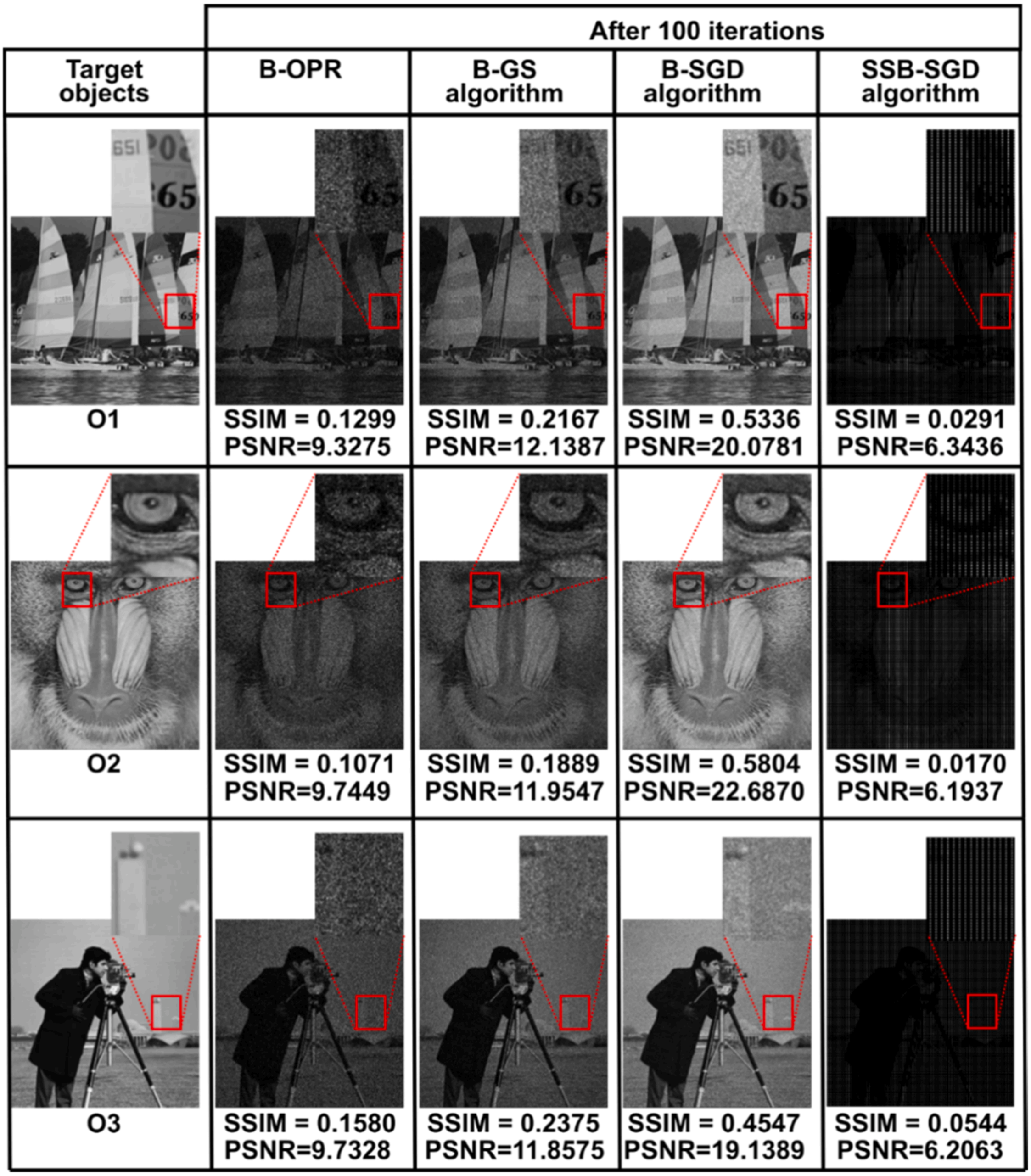

Fig. 4. Computational reconstruction of binary amplitude holograms (SSIM: structural similarity index measure, PSNR: peak signal-to-noise ratio).

4. We construct a $(N \times M) \times W \times H$ tensor that contains the amplitude hologram for each sub-sampled amplitude, where $W \times H$ is the desired hologram resolution.

In Fig. 3, we show a flowchart illustrating how the initial amplitude hologram tensor is obtained through these steps. We then employ GPU-based parallel processing to optimize the entire tensor simultaneously, using the same procedures described in steps 1-4 of the B-SGD algorithm. Here, the loss function is computed as the sum of the MSE between the reconstruction of each $M \times N$ channel of the binarized amplitude hologram tensor and the corresponding subsample amplitude target for that channel.

It is worth noting that with SSB-SGD, the result of this process is a tensor containing $N \times M$ BAHs, each corresponding to a subsample of the target, unlike B-SDG which reconstructs the entire amplitude with a single hologram. Thus, to obtain the entire object, the BAHs of all individual subsamples must be reconstructed, which, in an experimental implementation, requires temporal multiplexing.

This leads us to the key requirement of the proposed approach: as with all temporal multiplexing holographic display methods, an adequate reconstruction of the complete object can only be achieved if the projection device operates at a sufficiently high speed to display the BAHs of all subsamples within a short time interval. For example, with the parameters used in this work, achieving a display rate of at least 30 frames per second for complete objects while using 8 subsamples requires a BAH projection device with a refresh rate of at least 240 Hz. This exceeds the capabilities of most currently available phase spatial light modulators (SLMs) and even LCD screens. However, it is feasible with ferroelectric liquid crystal SLMs, digital micromirror devices (DMDs) as used in this work, or with extremely high-end and costly phase SLMs. Nevertheless, as we will show, when combining the reconstruction of each of the holograms in the SSB-SGD tensor, the results significantly surpass the quality of B-SDG reconstructions after applying temporal multiplexing with the same number of BAHs.

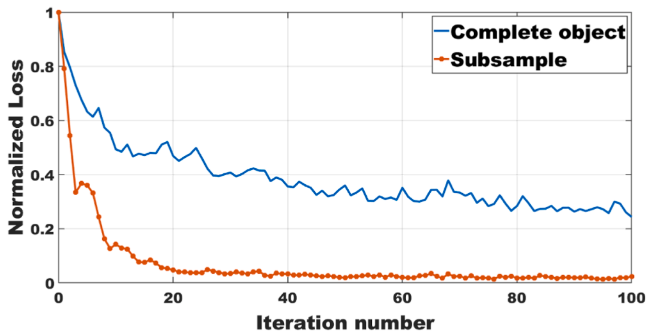

Fig. 5. Normalized loss as a function of the number of iterations of the B-SGD used to generate a BAH of the complete object O3 and a subsample of the same object.

3. Numerical results

We now proceed to test the four distinct methods for generating BAHs explained in the previous section: B-OPR, B-GS, B-SGD, and SSB-SGD. For our simulations, the target amplitude and the corresponding holograms have a resolution of 2048×1200 pixels. Likewise, we establish a 560×560 resolution for the target object, positioned at a distance of 70 pixels from the amplitude target's edge. For the SSB-SGD, the sampling is performed with $N = 4$ and $M = 2$, leading to 8 subsamples. For the B-GS, B-SGD, and SSB-SGD methods, 100 iterations were applied to obtain each result. For each result, we evaluate quality by computing the structural similarity index measure (SSIM) and peak signal-to-noise ratio (PSNR) between the reconstructed object and the target object.

The SSIM is a perceptually inspired image quality metric that evaluates the structural, luminance, and contrast similarity between two images [54]. The SSIM between two images A and B is given by

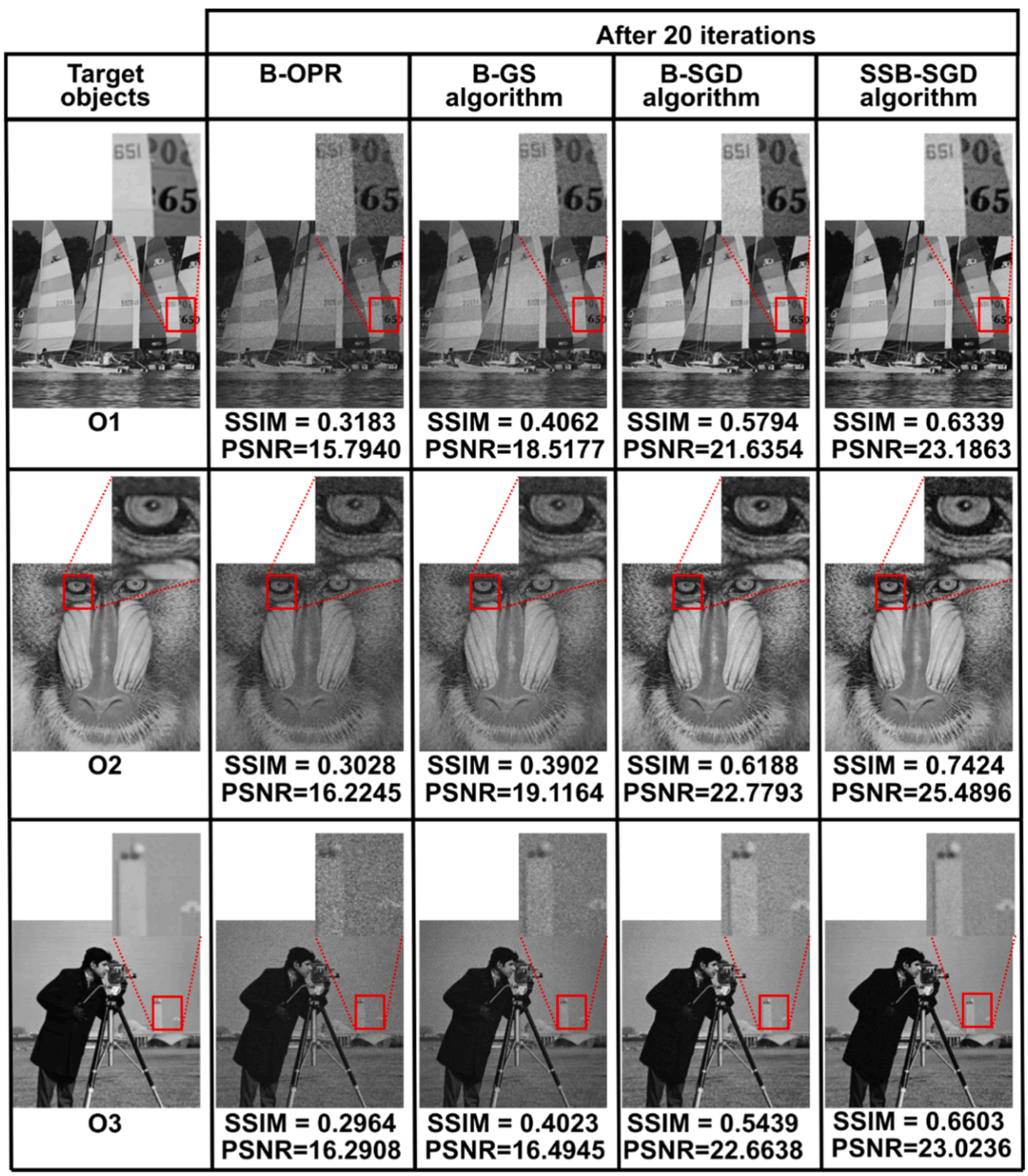

Fig. 6. Computational reconstruction of binary amplitude holograms with multiplexing (SSIM: structural similarity index measure, PSNR: peak signal-to-noise ratio).

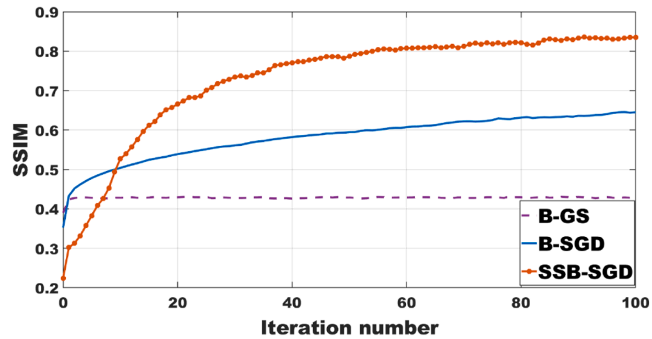

Fig. 7. Structural similarity index measure between O3 and the reconstructed object from eight temporally multiplexed binary amplitude holograms generated with B-GS, B-SGD, and SSB-SGD.

Table 1
Computation time for generating one and eight BAHs of O3.

		20 Iterations		
	B-OPR	B-GS	B-SGD	SSB-SGD
1 BAH (s)	0.0528	0.5187	0.2198	0.2851
8 BAHs (s)	0.4062	3.8167	0.6199	0.6647

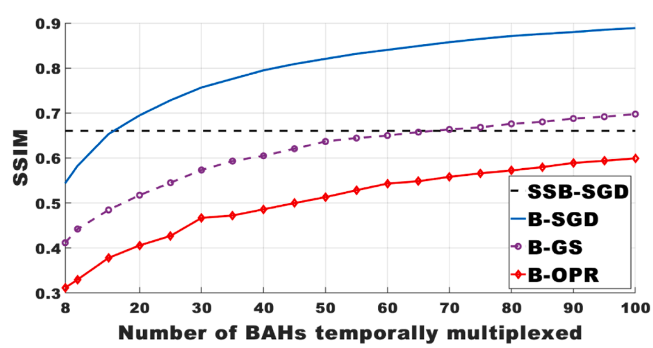

Fig. 8. Structural similarity index measure between O3 and the reconstructed objects from different number of multiplexed BAHs generated with B-SGD, B-GS and B-OPR. The dashed line without marker represents the SSIM achieved when multiplexing 8 BAH generated with SSB-SGD.

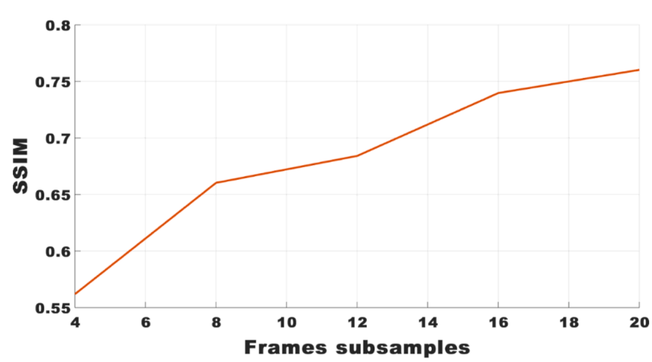

Fig. 9. SSIM of the reconstructed object obtained after temporal multiplexing of an increasing number of subsamples BAHs generated with SSB-SGD.

$$SSIM(A,B) = \frac{(2\mu_A\mu_B + C_1)(2\sigma_{AB} + C_2)}{(\mu_A{}^2 + \mu_B{}^2 + C_1)(\sigma_A{}^2 + \sigma_B{}^2 + C_2)} \tag{8}$$

where μ_A and μ_B are the mean values of the image A and B, respectively; σ_A and σ_B denote their standard deviations; σ_{AB} is their cross-correlation; and C_1 and C_2 are small constants introduced to prevent unstable results in the SSIM when $(\mu_A{}^2 + \mu_B{}^2)$ or $(\sigma_A{}^2 + \sigma_B{}^2)$ are close to zero. An SSIM of 1 means signifies perfect similarity between the two images.

The PSNR, on the other hand, is a quality metric derived from the mean square error, that quantifies the absolute per-pixel differences between the two images. It is commonly used to assess image degradation caused by lossy compression, or other lossy image processing techniques, making it well-suited for evaluating holographic reconstruction quality. PSNR is given by

$$PSNR(A,B) = 20\log_{10}\left(\frac{I_{\max}}{\sqrt{\frac{1}{nm}\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\left(A_{ij} - B_{ij}\right)^2}}\right) \tag{9}$$

where $I_{\max}$ denotes the maximum value of a pixel in the images being compared, $n \times m$ represents the image resolution, and i and j are pixel coordinates. For standard 8-bit images, like those used in this work, $I_{\max}$ is 255. A higher PSNR indicates greater similarity between the compared images.

In Fig. 4, the reconstructions of three objects (O1, O2 and O3) from the BAHs generated with each method are shown. The reconstructions derived from the B-SGD algorithm achieved the highest SSIM and PSNR values, indicating superior structural fidelity and visual quality. Conversely, objects reconstructed using the SSB-SGD algorithm show lower quality, since they only correspond to a single subsample, not the entire target object.

Nevertheless, it is important to highlight one of the main advantages of using the B-SGD to generate the hologram of only a subsample of the object: the overall convergence of the algorithm improves significantly, resulting in a hologram of the subsample with higher accuracy.

Fig. 5 illustrates the behavior of normalized loss across different iterations of B-SGD when generating a BAH for a complete object—specifically, O3 from Fig. 4—compared to generating a BAH for a subsample of that object. The results indicate a significantly greater loss reduction when generating the BAH for the subsample, as opposed to the complete object. This finding shows that generating BAHs for subsamples with B-SGD not only reduces noise due to increased pixel separation but also enhances the overall efficiency of the stochastic gradient descent optimization process. This highlights the effectiveness of the SSB-SGD approach. Based on these results, we find that 20 iterations of B-SGD are generally sufficient to produce satisfactory results when generating the hologram of a subsample. Therefore, unless otherwise stated, for all other results in this paper we will apply 20 optimization iterations when generating holograms for all methods where this is a necessary parameter.

Since SSB-SGD with the applied sampling produces 8 holograms, a correct measurement of its performance requires the averaging of the reconstruction of all 8 BAHs [55]. This can be achieved by using temporal multiplexing. To compare with the other three methods, specifically B-OPR, B-GS, and B-SGD we use 8 different initial random phases to produce 8 distinct BAHs of the same target amplitudes, and also perform the averaging of the reconstruction from all holograms generated with each method. The results of this multiplexing test are shown in Fig. 6. Here, it is evident that for all objects, our proposed SSB-SGD method yields the highest SSIM and PSNR values, superior visual quality, and minimal noise in the reconstructed object, confirming the effectiveness of our proposal.

Additionally, for each iteration, the SSIM value between the target

Fig. 10. Additional examples of computational reconstruction of binary amplitude holograms using temporal multiplexing (SSIM: structural similarity index measure, PSNR: peak signal-to-noise ratio).

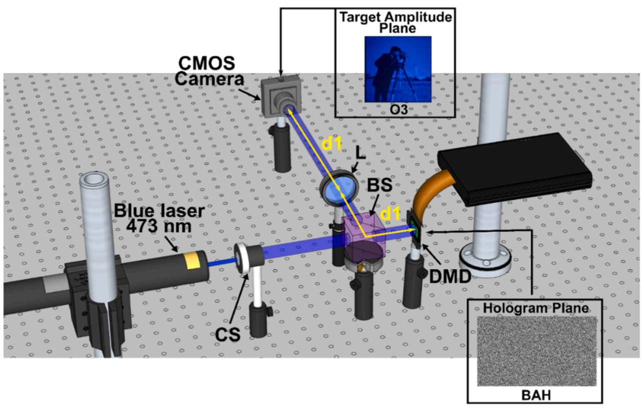

Fig. 11. Scheme for the experimental reconstruction of binary holograms (CS: collimation system; L: lens; BS: beam splitter; DMD: digital micromirror device, d1: lens focal length, equal to 12.5 cm).

object and the reconstructed object from the multiplexing of 8 BAHs is calculated, as shown in Fig. 7. We observe that the reconstruction quality improves as the number of iterations increases across all algorithms. In the case of the B-GS, this increase is minimal after 4 iterations, demonstrating its limited ability to produce high-quality BAHs. Both B-SGD and the SSB-SGD significantly outperform B-GS as the number of iterations increase beyond 3 and 8 respectively. Additionally, the SSB-SGD method not only demonstrates higher SSIM values but also exhibits a faster increase in quality, indicating a more effective optimization process. The results shown in Fig. 7 were obtained only for O3, as shown in Figs. 4 and 6. For all numerical results reported in this paper, we used GPU computing with a NVIDIA RTX4090.

Moreover, Table 1 presents the computation time required for generating a single hologram and for generating eight holograms for the application of temporal multiplexing. This table shows that the fastest algorithm for generating BAHs is B-OPR, but due to its deficiencies in reconstruction, it cannot compete with the other analyzed methods. The B-GS shows better quality than the B-OPR but is also inferior to the B-SGD and SSB-SGD in both computation speed and quality. It is noteworthy that the computation time for the B-SGD and SSB-SGD methods is quite similar. This highlights the fact that our algorithm not only delivers superior quality when temporal multiplexing is utilized but also offers computation time comparable to that of the B-SGD algorithm.

To further validate the effectiveness of our approach, we assessed the reconstruction quality by measuring the SSIM for temporal multiplexing using more than eight BAHs generated with conventional methods (B-SGD, B-GS, and B-OPR) and compared it to the quality achieved by multiplexing eight BAHs generated with our proposed approach. In the case of the B-GS, B-SGD, and SSB-SGD methods, 20 iterations were applied to generate each hologram to be multiplexed.

The results of this test are presented in Fig. 8. As shown, even when multiplexing up to 100 BAHs generated using B-OPR, the SSIM values remain lower than those achieved by multiplexing only 8 BAHs generated with SSB-SGD. On the other hand, the conventional B-GS and B-SGD can match or even exceed the SSIM values achieved with 8 BAHs from SSB-SGD, but they require multiplexing 66 and 16 BAHs, respectively, to achieve this. Generating these 66 BAHs with B-GS takes approximately 27.4197 seconds, and generating 16 BAH with B-SGD requires approximately 1.1814 seconds, compared to the 0.6647 seconds needed to produce 8 BAHs with SSB-SGD (see Table 1). This means that B-SGD requires nearly double the computation time, while B-GS requires 41 times more computation time to achieve a quality comparable to our approach. This underscores the significant advantage of the SSB-SGD.

Now, we show how the performance of the SSB-SGD depends on the number of subsamples of the object. To achieve this, we divided object O3 from Fig. 4 and 6 into an increasing number of subsamples, ranging from 4 to 20, and generated the corresponding BAHs using SSB-SGD. Then, we calculated the SSIM between the multiplexing of all the reconstructed BAHs and the target object.

The result of this test can be seen in Fig. 9. As expected, SSIM increases as more subsamples are used. However, this advantage must be balanced against the need to compute and multiplex an increasing number of BAHs. In this work, eight subsamples corresponding to N=4 and M=2 are selected, as this choice provides a significant increase in quality over the conventional SGD, while maintaining a manageable number of subsamples to be reconstructed and temporally multiplexed. In the next section, we will expand upon the experimental considerations involved in selecting this number of subsamples.

As a final test, we numerically evaluate the general performance of the SSB-SGD in comparison with the conventional SGD, the B-GS, and the B-OPR across a broader variety of objects. This evaluation is conducted using the same parameters as those employed in the results presented in Fig. 6.

The results of this test are presented in Fig. 10, demonstrating that, for all objects, our proposed SSB-SGD method achieves the highest SSIM and PSNR values, along with superior visual quality and minimal noise in the reconstructed objects. Furthermore, these results highlight the generalizability of our method, which performs effectively for both binary patterns and grayscale images, thereby confirming the robustness of our approach.

4. Experimental results

We now proceed to experimentally test our proposed hologram generation method. To achieve this, we implemented a holographic projection system based on a digital micromirror device, as shown in Fig. 11. The setup involves the following components

1. A coherent laser source operating at a wavelength of 473 nm with an output power of 150 mW. The beam undergoes expansion, collimation, and filtering through a collimation system (CS) composed of a spatial filter with a 60x microscope objective, a 25 um pinhole, and a positive lens with a focal length of 12.5 cm.
2. For the projection of BAHs, a DLPLCR50XEVM DMD featuring a resolution of 2048×1200 pixels and a pixel pitch of $5.4\mu m \times 5.4\mu m$.
3. An additional positive lens (L) with a focal length of 12.5 cm is used to perform the Fourier transform of the field reflected by the DMD in order to reconstruct the object.
4. The objects reconstructed from BAHs, projected via the DMD, are recorded using an IDS imaging U3-3800CP camera, which boasts a resolution of 5536×3692 pixels and a pixel pitch of $2.4\mu m$. An integration time of 0.133 s is used for recording, averaging the reconstruction of the projected holograms. The recorded objects occupied an area equivalent to $338 \times$ 338camera pixels in the recording plane.

In this study, we aimed to minimize speckle in the experimental reconstruction of objects by adopting a temporal multiplexing technique, which has been demonstrated to enhance reconstruction quality both numerically and experimentally. This method involves the creation of eight BAHs, which are then sequentially projected in the DMD at a framerate of 16.1 kHz.

Here, we highlight additional motivations for selecting eight subsamples for SSB-SGD and eight BAHs for temporal multiplexing in the other methods used in this study. The DMD has two modes of operation: "pattern-on-the-fly" mode and "video pattern" mode. In "pattern-on-the-fly" mode, the BAHs must be uploaded directly to the DMD's internal memory. This mode allows the full 16.1 kHz refresh rate but is limited to a maximum of 400 preloaded BAHs. In contrast, in "video pattern" mode, the BAHs are streamed to the DMD via an HDMI connection. In this streaming mode, the DMD can receive a 60 Hz video signal, where each

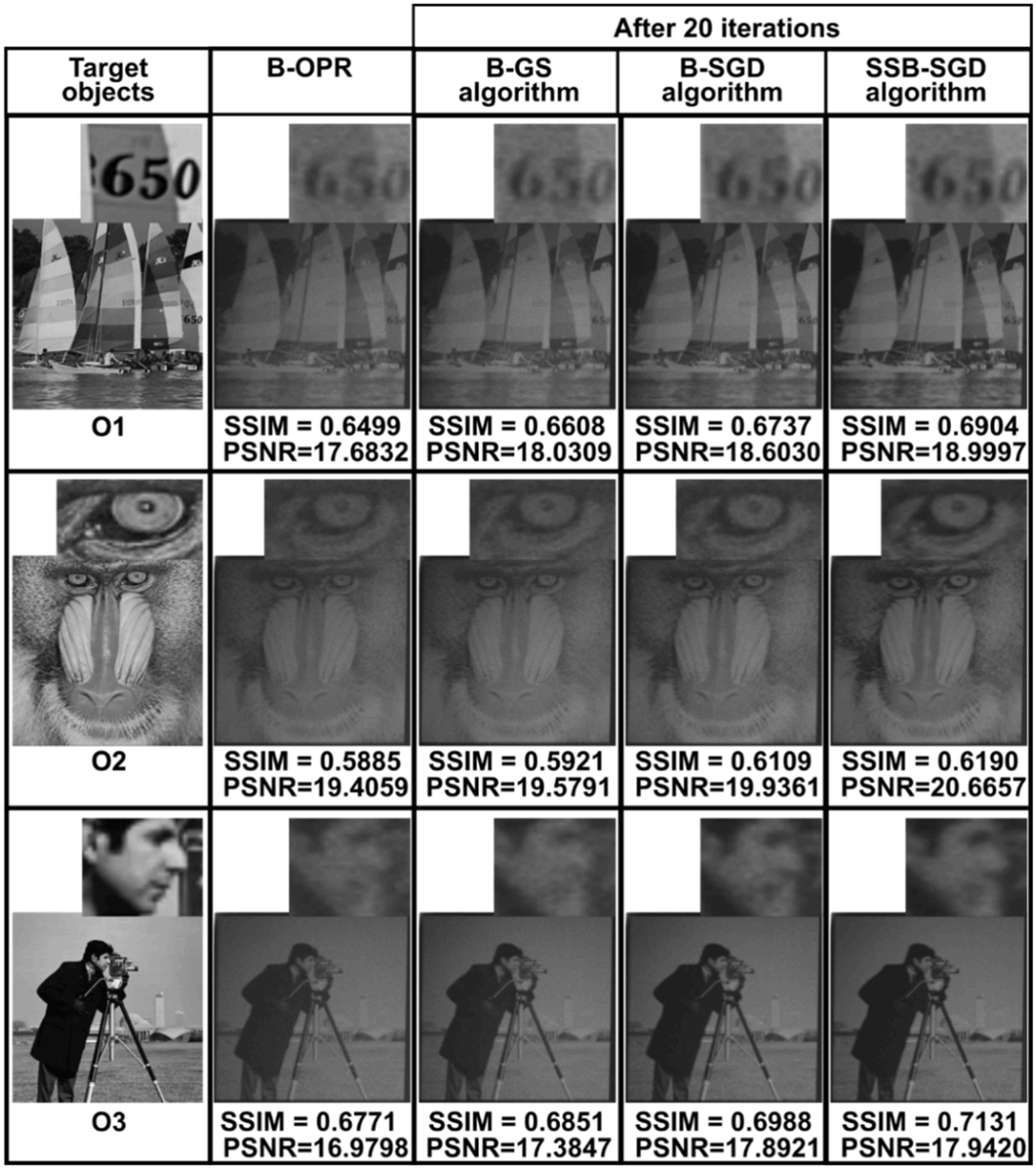

Fig. 12. Experimental reconstruction of binary amplitude holograms (SSIM: structural similarity index measure, PSNR: peak signal-to-noise ratio).

frame is composed of 24 binary patterns. Typically, eight binary patterns are assigned to each color channel, resulting in a projection rate of 1440 Hz per pattern. This mode is particularly advantageous, as it removes restrictions on the number of BAHs that can be projected, enabling applications such as color holographic video with no duration constraints. Consequently, projecting multiples of eight BAHs is optimal, as it ensures maximum flexibility and compatibility with our specific DMD configuration.

The experimental reconstructions from BAHs for various objects, depicted in Fig. 12, illustrate that reconstruction using the SSB-SGD algorithm consistently exhibit superior quality across all objects. Thus, we experimentally confirm the numerical results, demonstrating that the SSB-SGD method significantly enhances reconstruction quality when combined with temporal multiplexing. This finding is supported by the higher SSIM values reported in Figs. 6, 7, 8, 10, and 12, as well as in the PNSR values of Figs. 6, 10, and 12 for both numerical and experimentally evaluations, visually affirming the superiority of our approach over the B-SGD method. It is worth noting, that, while the experimental results show higher SSIM values compared to the numerical results, these two cases should not be directly compared, due to the fact that the experimental reconstructions were recorded with a lower resolution than the numerical results.

5. Conclusions

In this work, we implement an approach based on subsampling the target scenes, named SSB-SGD, which reduces speckle noise and significantly improves quality in applications where temporal multiplexing of holograms is applied. Our proposal can achieve the same reconstruction quality as B-SGD with nearly half the computation time and half the number of temporally multiplexed holograms, or significantly better quality with nearly the same computation time. Furthermore, we demonstrate the application of B-SGD to generate BAHs for projection in an experimental holographic setup scheme based on binary amplitude modulation by means of a digital micromirror device.

Comparisons with other common BAH generation methods, such as those based on the GS algorithm and on one-step phase retrieval, are also presented. The effectiveness of our proposal is tested both numerically and experimentally. All results confirm that the SSB-SGD offers the best quality without a trade-off in computation speed, unlike most common CGH methods. While further research is needed to extend the SSB-SGD to generating holograms of 3D or multiplane scenes, it is worth noting that the SSB-SGD has no limitations on the target object characteristics, potentially simplifying this application. Additionally, a similar approach to the SSB-SGD could also be effectively applied to the generation of phase holograms with SGD.

Funding

Comité para el Desarrollo de la Investigación—CODI (Universidad de Antioquia—UdeA, Colombia); Optica Foundation; Sistema General de Regalías de Colombia (BPIN 2020000100464).

Author contributions

A.V conceived the initial proposal. All authors carried out the experimental design and simulation. All authors contributed to the discussion of the results. The manuscript was written with contributions from all co-authors.

Data availability

The data that supports the results within this paper is available from the corresponding authors upon reasonable request.

Code availability

The custom code and mathematical algorithms used to obtain the results within this paper are available from the corresponding authors upon reasonable request.

CRediT authorship contribution statement

Alejandro Velez-Zea: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **César Antonio Hoyos-Peláez:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **John Fredy Barrera-Ramírez:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

[1] Brown BR, Lohmann AW. Complex spatial filtering with binary masks. Appl. Opt. 1966;5:967–9.
[2] Junge S, Schmieder F, Sasse P, Czarske J, Torres-Mapa ML, Heisterkamp A. Holographic optogenetic stimulation with calcium imaging as an all optical tool for cardiac electrophysiology. J. Biophotonics 2022;15:e202100352.
[3] Faini G, Tanese D, Molinier C, Telliez C, Hamdani M, Blot F, Tourain C, de Sars V, Del Bene F, Forget BC, Ronzitti E, Emiliani V. Ultrafast light targeting for high-throughput precise control of neuronal networks. Nat. Commun. 2023;14:1888.
[4] Carmi I, De Battista M, Maddalena L, Carroll EC, Kienzler MA, Berlin S. Holographic two-photon activation for synthetic optogenetics. Nat. Protoc. 2019;14:864–900.
[5] Abacousnac J, Grier DG. Dexterous holographic trapping of dark-seeking particles with Zernike holograms. Opt. Express 2022;30:23568–78.
[6] Bhebhe N, Williams PAC, Rosales-Guzmán C, Rodriguez-Fajardo V, Forbes A. A vector holographic optical trap. Sci. Rep. 2018;8:1–9.
[7] Jang C, Bang K, Chae M, Lee B, Lanman D. Waveguide holography for 3D augmented reality glasses. Nat. Commun. 2024;15:1–12.
[8] Zhong C, Sang X, Yan B, Li H, Xie X, Qin X, Chen S. Real-time 4K computer-generated hologram based on encoding conventional neural network with learned layered phase. Sci. Rep. 2023;13:1–10.
[9] Gerchberg RW, Saxton WO. A practical algorithm for the determination of the phase from image and diffraction plane pictures. Optik 1972;35:237–46.
[10] Dorsch RG, Lohmann a W, Sinzinger S. Fresnel ping-pong algorithm for two-plane computer-generated hologram display. Appl. Opt. 1994;33:869–75.
[11] Su P, Cao W, Ma J, Cheng B, Liang X, Cao L, Jin G. Fast computer-generated hologram generation method for three-dimensional point cloud model. J. Disp. Technol. 2016;12:1688–94.
[12] Velez-Zea A, Torroba R. Mixed constraint in global and sequential hologram generation. Appl. Opt. 2021;60:1888–95.
[13] Piestun R, Spektor B, Shamir J. Wave fields in three dimensions: analysis and synthesis. J. Opt. Soc. Am. A 1996;13:1837–47.
[14] Makowski M, Sypek M, Kolodziejczyk A, Mikula G, Suszek J. Iterative design of multiplane holograms: experiments and applications. Opt. Eng. 2007;46:045802.
[15] Velez-Zea A, Barrera-Ramírez JF, Torroba R. A modified constraint for iterative phase-only multiplane hologram generation. In: Opt. InfoBase Conf. Pap. 2021; 2021. DTh7C.4.
[16] Chen C, Lee B, Li N-N, Chae M, Wang D, Wang Q-H, Lee B. Multi-depth hologram generation using stochastic gradient descent algorithm with complex loss function. Opt. Express 2021;29:15089–103.
[17] Chakravarthula P, Peng Y, Kollin J, Fuchs H, Heide F. Wirtinger holography for near-eye displays. ACM Trans. Graph. 2019;38:213.
[18] Arrizón V. Improved double-phase computer-generated holograms implemented with phase-modulation devices. Opt. Lett. 2002;27:595–7.
[19] Shimobaba T, Wang F, Starobrat J, Kowalczyk A, Suszek J, Ito T. Comparison of double-phase hologram and binary amplitude encoding: Holographic projection and vortex beam generation. Appl. Opt. 2023;62:7471–9.
[20] Velez-Zea A, Barrera-Ramírez JF, Torroba R. Optimized random phase only holograms. Opt. Lett. 2018;43:731–4.
[21] Velez-Zea A, Torroba R. Optimized random phase tiles for non-iterative hologram generation. Appl. Opt. 2019;58:9013–9.
[22] Pang H, Wang J, Cao A, Deng Q. High-accuracy method for holographic image projection with suppressed speckle noise. Opt. Express 2016;24:22766–76.
[23] Chen L, Zhang H, Cao L, Jin G. Non-iterative phase hologram generation with optimized phase modulation. Opt. Express 2020;28:11380–92.
[24] Zhang Y, Zhang M, Liu K, He Z, Cao L. Progress of the computer-generated holography based on deep learning. Appl. Sci. 2022;12:8568.
[25] Peng Y, Choi S, Padmanaban N, Wetzstein G. Neural holography with camera-in-the-loop training. ACM Trans. Graph. 2020;39:185.
[26] Lazarev G, Chen P-J, Strauss J, Fontaine N, Forbes A. Beyond the display: phase-only liquid crystal on Silicon devices and their applications in photonics [Invited]. Opt. Express 2019;27:16206–49.
[27] Packer O, Diller LC, Verweij J, Lee BB, Pokorny J, Williams DR, Dacey DM, Brainard DH. Characterization and use of a digital light projector for vision research. Vision Res 2001;41:427–39.
[28] Zhang C, Huang PS, Chiang F-P. Microscopic phase-shifting profilometry based on digital micromirror device technology. Appl. Opt. 2002;41:5896–904.
[29] Miao Y, Yang Y, Hou Q, Wang Z, Liu X, Tang Q, Peng X, Gao BZ. High-efficiency 3D reconstruction with a uniaxial MEMS-based fringe projection profilometry. Opt. Express 2021;29:34243–57.
[30] Kim J, Jeon P, Lee H, Kim DY. High-frequency sinusoidal structured light generation with the anti-blaze condition of a digital micromirror device. Opt. Lasers Eng. 2024;176:108053.
[31] Hu Y, Chen Q, Feng S, Zuo C. Microscopic fringe projection profilometry: a review. Opt. Lasers Eng. 2020;135:106192.
[32] Ren YX, De Lu R, Gong L. Tailoring light with a digital micromirror device. Ann. Phys. 2015;527:447–70.
[33] Ayoub AB, Psaltis D. High speed, complex wavefront shaping using the digital micro-mirror device. Sci. Rep. 2021;11:1–10.
[34] Goorden SA, Bertolotti J, Mosk AP. Superpixel-based spatial amplitude and phase modulation using a digital micromirror device. Opt. Express 2014;22:17999–8009.
[35] Jiao S, Zhang D, Zhang C, Gao Y, Lei T, Yuan X. Complex-amplitude holographic projection with a digital micromirror device (DMD) and error diffusion algorithm. IEEE J. Sel. Top. Quantum Electron. 2020;26:1–8.
[36] Jaramillo-Osorio A, Bustamante S, Munoz B, Velez-Zea A, Barrera-Ramírez JF, Torroba R. Experimental Fresnel and Fourier digital holography using a digital micro-mirror device. J. Opt. 2021;23:035701.
[37] Yan H, Sun Y, Lin Y, Chu F, Wan W. Multi-color complex spatial light modulation with a single digital micromirror device. Opt. Express 2023;31:22649–59.
[38] Cheremkhin PA, Kurbatova EA. Binarization of digital holograms by thresholding and error diffusion techniques. In: Opt. InfoBase Conf. Pap.; 2019. p. 1–2.

[39] Cheremkhin PA, Kurbatova EA, Evtikhiev NN, Krasnov VV, Rodin VG, Starikov RS. Adaptive digital hologram binarization method based on local thresholding, block division and error diffusion. J. Imaging 2022;8:15.
[40] Hoyos-Peláez CA, Velez-Zea A, Barrera-Ramírez JF. Non-iterative generation of binary amplitude holograms applied to holographic projection with digital micromirror devices. J. Opt. 2024;26:035602.
[41] Seldowitz MA, Allebach JP, Sweeney DW. Synthesis of digital holograms by direct binary search. Appl. Opt. 1987;26:2788–98.
[42] Kang JH, Leportier T, Kim M, Park MC. Non-iterative direct binary search algorithm for fast generation of binary holograms. Opt. Lasers Eng. 2019;122: 312–8.
[43] Jennison BK, Sweeney DW, Allebach JP. Efficient design of direct-binary-search computer-generated holograms. J. Opt. Soc. Am. A 1991;8:652–60.
[44] Tsang P, Poon TC, Cheung WK, Liu JP. Computer generation of binary Fresnel holography. Appl. Opt. 2011;50:B88–95.
[45] Barnard E. Optimal error diffusion for computer-generated holograms. J. Opt. Soc. Am. A 1988;5:1803–17.
[46] Yang G, Jiao S, Liu J-P, Lei T, Yuan X. Error diffusion method with optimized weighting coefficients for binary hologram generation. Appl. Opt. 2019;58: 5547–55.
[47] Min K, Park J-H. Quality enhancement of binary-encoded amplitude holograms by using error diffusion. Opt. Express 2020;28:38140–54.
[48] Lee B, Kim D, Lee S, Chen C, Lee B. High-contrast, speckle-free, true 3D holography via binary CGH optimization. Sci. Rep. 2022;12:2811.
[49] Takaki Y, Yokouchi M. Speckle-free and grayscale hologram reconstruction using time-multiplexing technique. Opt. Express 2011;19:7567–79.
[50] Cheremkhin PA, Kurbatova EA. Comparative appraisal of global and local thresholding methods for binarisation of off-axis digital holograms. Opt. Lasers Eng. 2019;115:119–30.
[51] Velez-Zea A, Quinchia SB, Barrera-Ramírez JF, Torroba R. Fast computation of binary amplitude holograms with optimized random phases. Opt. InfoBase Conf. Pap. 2021;2021. DTu7B.5.
[52] Kingma, D.P., Ba, J.L., Adam: a method for stochastic optimization, 3rd International Conference for Learning Representations. ICLR 2015 - Conf. Track Proc. 1–15 (2015).
[53] Makowski M. Minimized speckle noise in lens-less holographic projection by pixel separation. Opt. Express 2013;21:29205–16.
[54] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 2004;13:600–12.
[55] Martienssen W, Spiller S. Holographic reconstruction without granulation. Phys. Lett. A 1967;24:126–8.