Research Article Vol. 5, No. 1 / 15 Jan 2026 / Optics Continuum 170
OPTICS CONTINUUM

Deep learning-based complex hologram compression enhanced by ringing reduction

MAKOTO SEKIGUCHI,[1,6] CHUNG-HSUAN HUANG,[2] DAVID BLINDER,[3,4] SHOSEI YABUKI,[1] FAN WANG,[1] PETER SCHELKENS,[3,4] HAN-YEN TU,[2] CHAU-JERN CHENG,[5] TOMOYOSHI ITO,[1] AND TOMOYOSHI SHIMOBABA[1,7]

[1]*Graduate School of Engineering, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan*
[2]*Department of Electrical Engineering, Chinese Culture University, Taipei 11114, Taiwan*
[3]*Department of Electronics and Informatics (ETRO), Vrije Universiteit Brussel (VUB), Pleinlaan 2, B-1050 Brussel, Belgium*
[4]*IMEC, Kapeldreef 75, B-3001 Leuven, Belgium*
[5]*Institute of Electro-Optical Engineering, National Taiwan Normal University, Taipei 11677, Taiwan*
[6]*24wm4307@student.gs.chiba-u.jp*
[7]*shimobaba@faculty.chiba-u.jp*

Abstract: Recent advances in computer-generated holography (CGH) have significantly improved visual quality through high-resolution rendering; however, the accompanying increase in data size has become a critical obstacle to practical deployment. Conventional image compression techniques such as JPEG and high efficiency video coding (HEVC) do not adequately account for the unique statistical and spectral characteristics of holograms, thereby limiting both compression efficiency and reconstruction quality. This study proposes an efficient compression framework specialized for digital holograms by integrating a ringing reduction technique for diffraction calculations with CompressAI, a deep learning-based image compression framework. We demonstrate that ringing artifacts are a major factor hindering efficient hologram compression, and show that neural networks can achieve superior performance by learning the intrinsic statistical and spectral features of holograms. Our method achieves a superior balance between compression efficiency and reconstruction quality compared to conventional approaches, particularly at low bit rates. Furthermore, by introducing a multi-channel input representation, our method achieves higher compression ratios.

1. Introduction

With the rapid advancement of computer technology, the resolution of computer-generated holograms (CGHs) has increased dramatically, leading to substantial improvements in the quality of reconstructed images. Recent studies have reported the realization of large-scale CGH generation techniques. Time-sequential holographic displays use multiple CGHs to mitigate speckle noise and expand the field of view [1,2]. They are driving progress toward practical applications of large field-of-view holographic displays in next-generation VR/AR systems.

Alongside these developments, the demand for storage and transmission of CGHs and digital holograms has grown significantly, rendering efficient delivery of holographic data over networks indispensable [3,4]. However, holograms inherently contain fine interference fringes and rich high-frequency components, resulting in data sizes that are orders of magnitude larger than those of natural images [5,6]. As resolution continues to increase, storage and transmission requirements have become major bottlenecks. To address this issue, establishing highly efficient compression methods for holographic data is urgently required [7,8]. A variety of compression methods tailored explicitly to holograms have been explored. Early approaches include vector

#581618 https://doi.org/10.1364/OPTCON.581618
Journal © 2026 Received 13 Oct 2025; accepted 24 Dec 2025; published 12 Jan 2026

OPTICS CONTINUUM

quantization-based coding [9,10], as well as methods exploiting sparse representations and domain partitioning [11–13].

Previous studies have attempted to apply conventional codecs such as JPEG [14], JPEG 2000 [15], and high efficiency video coding (HEVC) [16,17] to holograms. Since these codecs are designed based on the assumption that most spectral energy of natural images lies in low-frequency regions, they are not always suitable for holograms, which exhibit broad spectral distributions.

JPEG Pleno Holography [18–21] is a standardization effort by the JPEG committee focused on the efficient representation and compression of holographic data. JPEG Pleno aims to provide interoperability, scalability, and optimized storage for hologram data, with applications ranging from medical imaging and cultural heritage to virtual reality and scientific research.

The emergence of deep learning has spurred active research on learning-based hologram compression [22]. Various learning-based approaches have been introduced in recent years. These include end-to-end trainable methods that optimize the entire compression pipeline, covering both phase hologram generation and coding [23]; integrative frameworks such as neural holographic video compression (NHVC) [24], which jointly address generation and video compression; and joint optimization techniques that explicitly account for reconstruction quality degradation during the coding process [25]. Additionally, deep learning-based compression specifically targeting complex-valued holograms has been explored [26].

In addition to fully trainable models, hybrid schemes have been investigated wherein holograms undergo conventional compression (e.g., JPEG) followed by deep learning-based enhancement to recover lost high-frequency details and suppress artifacts [27]. These methods are particularly effective for improving the fidelity of phase-only hologram reconstructions. Their significance lies in the seamless integration with existing compression standards, combined with the specialized capacity of neural networks to counteract typical holographic degradations introduced during bit-depth reduction or transform-based coding.

Perceptual optimization frameworks that incorporate human visual characteristics have also been developed, and more recently, hologram compression leveraging foveated rendering—representing the foveal region in high resolution while coarsely encoding the periphery—has achieved high compression ratios [28].

In this work, we extend the deep learning-based image compression library CompressAI [29] to propose a novel compression method tailored for complex holograms. Specifically, we employ the scale hyperprior model [30], an extension of the hyperprior architecture that captures spatial dependencies to minimize bitrate. We train it using the real and imaginary components of holograms to achieve efficient coding beyond the capabilities of natural-image-oriented codecs. We further investigate the effect of ringing artifacts, an issue largely overlooked in prior studies, on hologram compression and demonstrate that their reduction enables higher compression ratios in both conventional codecs and learning-based methods. Two ringing reduction techniques are considered: the fast Fresnel integrals method [31] and the blank aperture method [32]. In addition, we introduce a multi-channel input representation that allows multiple holograms to be jointly encoded, reducing redundancy and enabling even higher compression. To validate the effectiveness of the proposed approach, we conducted comparative experiments against standard codecs such as JPEG 2000 and HEVC, as well as state-of-the-art learning-based methods. The results show that our method achieves superior compression efficiency and reconstruction quality for both holograms and their reconstructed images.

In the remainder of this paper, Section 2 presents the proposed method, Sections 3 and 4 describe the experimental results, and Section 5 concludes the paper.

2. Method

This section describes the proposed framework for hologram compression and its implementation details.

OPTICS CONTINUUM

2.1. Outline

We aim to achieve efficient compression of hologram image data by leveraging CompressAI [29], an open-source deep learning–based image compression framework implemented in PyTorch. CompressAI provides end-to-end training and evaluation pipelines for image compression, upon which we build a framework tailored explicitly to complex holograms.

An overview of the processing pipeline is shown in Fig. 1. The input consists of complex holograms, represented as paired real and imaginary components for both training and inference. Furthermore, by flexibly increasing the number of input channels, the framework can jointly process multiple holograms in a single pass, thereby allowing the network to exploit inter-hologram correlations and reduce redundancy. The C-channel input holograms are first transformed by an encoder into a latent representation, which is subsequently decoded to reconstruct the original holograms.

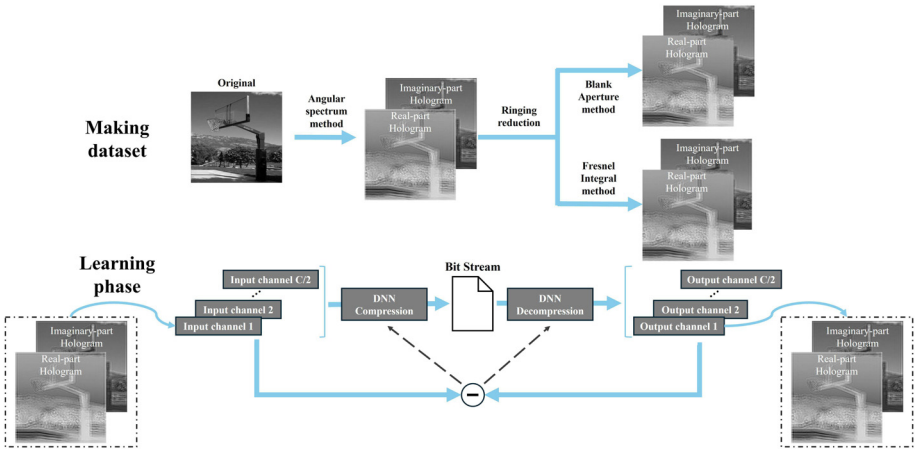

Fig. 1. Overview of the proposed method.

As a preprocessing step, we apply two ringing reduction techniques — the Fresnel integrals method and the blank aperture method — to mitigate diffraction-induced ringing artifacts. Both ringing-processed and unprocessed holograms are independently used as input datasets, allowing for a comparative evaluation of the impact of ringing reduction on compression performance. In this way, the proposed framework effectively captures and compresses the complex spatial structures and rich high-frequency content characteristic of holograms.

2.2. Scale hyperprior model

This study adopts the scale hyperprior model [30] as the backbone for learning hologram compression. The factorized prior model [33], which serves as a baseline approach, assumes that each element in the latent representation is statistically independent and models their distributions with a fully factorized density. The distribution parameters are typically shared globally across all spatial locations in the latent space. Compared to the factorized prior, the hyperprior model captures spatial dependencies in the latent representation by introducing an additional set of latent variables that model the distribution parameters. This leads to a more accurate entropy model and significantly improves compression efficiency.

By contrast, the scale hyperprior model introduces additional mechanisms to capture spatially-varying statistical properties of the latent representation, leading to improved compression performance.

OPTICS CONTINUUM

The network architecture is illustrated in Fig. 2. In addition to the conventional encoder–decoder structure, the scale hyperprior introduces a hyperencoder h_a and a hyperdecoder h_s to estimate the statistical properties of the latent representation.

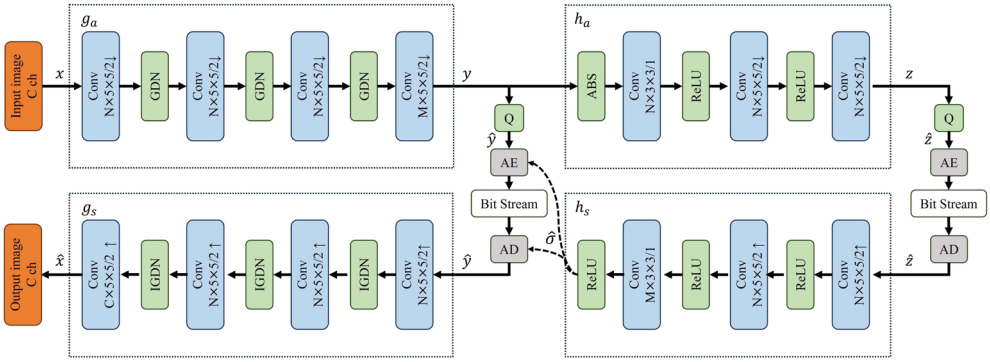

Fig. 2. Network architecture of the scale hyperprior model used in the proposed method. Top-left: encoder (analysis transform) g_a. Top-right: hyperencoder h_a. Bottom-left: decoder (synthesis transform) g_s. Bottom-right: hyperdecoder h_s. The input hologram is transformed by g_a into the latent y, while the hyperpath extracts an auxiliary latent z. Both y and z are quantized (Q) and entropy-coded (AE). During decoding, $\hat{z}$ provides a prior for optimal entropy decoding of y, and g_s reconstructs the output image. (Conv: convolution; GDN/IGDN: Generalized Divisive Normalization/Inverse GDN; ReLU: activation function; Q: quantization; AE/AD: arithmetic encoder/decoder. ABS denotes the absolute value operation.

Given a C-channel input hologram x, the encoder (analysis transform) g_a converts it into a latent representation y. The encoder consists of four convolutional layers with generalized divisive normalization (GDN), progressively reducing spatial resolution while compressing features. The latent y is quantized (Q) to $\hat{y}$ and entropy-coded into a bitstream using an arithmetic encoder (AE).

To improve entropy modeling, the scale hyperprior extracts auxiliary features z from y via the hyperencoder h_a, which is implemented using convolutional layers with ReLU activations. The auxiliary latent z is quantized to $\hat{z}$ and entropy-coded as well. On the decoder side, the arithmetic decoder (AD) first recovers $\hat{z}$, which is then fed to the hyperdecoder h_s to predict spatially varying scale parameters $\hat{\sigma}$ for y. Conditioning the probability model $p(y \mid \hat{z})$ on these scales enables a more precise entropy model during arithmetic decoding, thereby improving coding efficiency compared to simpler priors.

Finally, the recovered latent $\hat{y}$ is passed to the decoder (synthesis transform) g_s, which mirrors the encoder and employs inverse GDN (IGDN) and transposed convolutions to progressively restore spatial resolution, yielding the C-channel reconstructed hologram $\hat{x}$.

The scale hyperprior model achieves efficient compression and faithful reconstruction of holograms by accurately capturing statistical dependencies in the latent space through side information. This is particularly effective for holographic data, where redundancy and regularity are intricately intertwined.

2.3. *Training strategy and implementation details*

We optimize the network using a rate–distortion loss [34] that balances distortion between the reconstructed and original holograms and the resulting bitrate. The loss L is defined as:

$$L = \alpha D + R, \tag{1}$$

OPTICS CONTINUUM

where D denotes the distortion (we use mean squared error), R is the bitrate after entropy coding, and α is a weight parameter. Higher α prioritizes reconstruction quality, while lower α aims for higher compression. During training, quantization is approximated by additive uniform noise, while rounding is used at inference time. The rate R is computed as the expected code length under the scale-Gaussian entropy model provided by the hyperprior. The weighting parameter α was determined through empirical evaluation: we tested values of 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, and 1.0, and selected $\alpha = 0.01$ as it yielded the optimal balance between reconstruction quality and compression efficiency across the evaluated configurations.

In the multi-channel setting, we jointly input the real and imaginary components of the holograms, so that cross-channel correlations are learned. This configuration enables the model to potentially exploit inter-component redundancy, leading to lower bitrate R and, in some cases, improved reconstruction accuracy D. Extensions to more than two channels (e.g., concatenating multiple holograms) are also possible, but in this work, we focus on real and imaginary pairs.

It is known that hologram generation via diffraction can introduce ringing artifacts. Prior work on hologram compression has largely ignored this issue, even though such artifacts degrade reconstructed image quality and add high-frequency noise that is challenging to compress. To investigate the impact of ringing on compression performance, we construct three datasets from identical source images: (i) original holograms without ringing reduction [31], (ii) holograms processed with the Fresnel integrals method, and (iii) those processed with the blank aperture method [32]. By using the same compression architecture and hyperparameters for all experiments, any performance difference can be attributed solely to the effect of ringing reduction.

The blank aperture method [32] suppresses ringing by dividing the diffraction result on the hologram by the diffraction result of a unit-amplitude aperture:

$$u_2(x_2, y_2) = \frac{P\{u_1(x_1, y_1)\}}{P\{1\}}, \tag{2}$$

where $u_1(x_1, y_1)$ is the hologram, $u_2(x_2, y_2)$ is the ringing-reduced hologram, and $P\{\cdot\}$ denotes a diffraction operator (e.g., angular spectrum or Fresnel propagation). $P\{1\}$ is the diffraction result of an aperture whose pixel values are all ones.

The fast ringing reduction method [31] based on the Fresnel integrals is computed as follows. Unlike Eq. (2), this approach requires no additional diffraction computations or memory: the diffraction result of the unit-amplitude aperture, $F(x_2, y_2)$, is obtained analytically via the Fresnel integrals in advance, and ringing is suppressed by

$$u_2(x_2, y_2) = \frac{P\{u_1(x_1, y_1)\}}{F(x_2, y_2)}, \tag{3}$$

where

$$\begin{aligned} F(x, y) = \frac{\exp(ikz)}{2i} &\Big([C(\alpha_2) - C(\alpha_1)] + i[S(\alpha_2) - S(\alpha_1)]\Big) \\ &\times \Big([C(\beta_2) - C(\beta_1)] + i[S(\beta_2) - S(\beta_1)]\Big), \end{aligned} \tag{4}$$

where $C(\cdot)$ and $S(\cdot)$ denote the Fresnel cosine and sine integrals, respectively. The variables are defined as $\alpha_1 = -\sqrt{2/(\lambda z)}\,(\omega + x)$, $\alpha_2 = -\sqrt{2/(\lambda z)}\,(\omega - x)$, $\beta_1 = -\sqrt{2/(\lambda z)}\,(\omega + y)$, and $\beta_2 = -\sqrt{2/(\lambda z)}\,(\omega - y)$, with $k = 2\pi/\lambda$ being the wavenumber. λ and z denote the wavelength and propagation distance, respectively. Here, ω indicates the calculation window size, so the computational domain is $2\omega \times 2\omega$.

OPTICS CONTINUUM

3. Implementation details

3.1. Dataset

We used the Caltech256 dataset [35] as the source images. Each image was resized to 2048×2048 pixels. Under the optical settings (wavelength 520 nm, pixel pitch 6.4 μm), we synthesized complex holograms at propagation distances from 0.20 m to 1.00 m in 0.20 m increments.

From 300 source images, we generated a total of 1,500 complex holograms (300 images $\times$ 5 distances), each represented by its real and imaginary components. Each hologram was then divided into non-overlapping 256×256 patches, which yielded a data set of 96,000 samples. This patching strategy increases the amount of training data and helps mitigate overfitting. We used 80% of the data (76,800 samples) for training and 20% (19,200 samples) for validation.

3.2. Environment

Table 1 summarizes the hyperparameters and the environment. We used Adam for optimization and the rate–distortion loss in Eq. (1). The batch size was fixed to 32. The learning rate was initialized at 1×10^{-4} and decayed during training.

Table 1. Experimental settings

Parameter	Value
Number of pixels [px]	2048×2048
Wavelength [nm]	520
Pixel pitch [μm]	6.4
Propagation distance [m]	0.20–1.00 (step = 0.20)
Batch size	32
Loss function	Rate–distortion loss
α	0.01
Optimization algorithm	Adam
Learning rate	1×10^{-4} (with decay)

The computing environment consisted of an AMD EPYC Genoa CPU, 256 GB RAM, and an NVIDIA RTX 6000 Ada Generation GPU. The software stack included Python 3.12.3, PyTorch 2.6.0, and CompressAI 1.2.8.

4. Result

4.1. Compression rate

The bits per pixel (bpp) of a compressed hologram are computed as

$$bpp = \frac{8 \times B}{C \times H \times W}, \tag{5}$$

where B is the compressed bitstream size in bytes, H and W are the image height and width in pixels, and C is the total number of input channels (e.g., the real and imaginary parts of one or more holograms). Here, we assume the uncompressed data are represented by 8 bits per pixel per channel. Accordingly, the compression ratio (CS) is given by

$$\mathrm{CS} = \frac{8}{bpp} = \frac{\mathrm{C} \times \mathrm{H} \times \mathrm{W}}{\mathrm{B}}. \tag{6}$$

A higher CS denotes greater data reduction relative to the original uncompressed holograms.

OPTICS CONTINUUM

4.2. Relationship between input channel and bitrate

In the proposed framework, we can input multiple pairs of real and imaginary holograms as well as concatenate holograms corresponding to different propagation distances to expand the input channel count. This design improves robustness across propagation distances and aggregates cross-channel correlations and redundancy into the latent space, enabling more efficient coding. We therefore evaluated the relationship between the number of input channels and the compressed bitrate.

Figure 3 illustrates the variations in bpp, PSNR, and MS-SSIM as the number of input channels increases from 2 to 256 in powers of two. The results show that *bpp* decreases approximately exponentially as the number of channels increases. In particular, for eight channels and above, the scale hyperprior model consistently achieves lower *bpp* than the factorized prior, indicating that it captures inter-channel correlations more effectively. Furthermore, in terms of PSNR, the scale hyperprior consistently outperforms the factorized prior across nearly all channel configurations. A similar trend is observed for MS-SSIM, where the scale hyperprior yields superior scores in most cases; however, at the 32-channel mark, the factorized prior exhibits a marginal advantage. This minor discrepancy is likely attributable to the inherent training variability of deep learning models. These results suggest that the synergy between the scale hyperprior architecture and channel expansion enhances compression efficiency beyond the bitrate reduction afforded by simply increasing the channel count. Consequently, the individual contributions of model architecture and channel capacity to overall performance can be clearly decoupled.

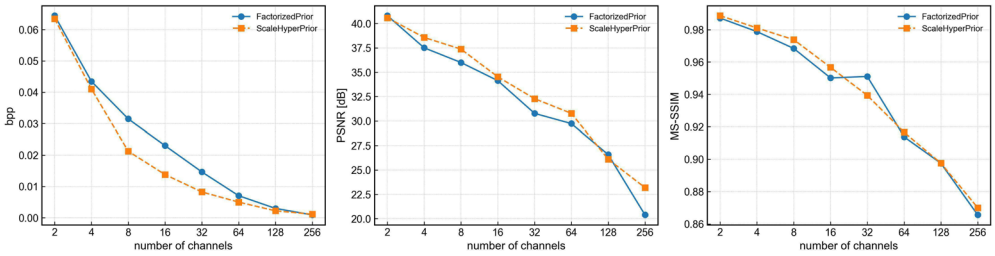

Fig. 3. Relationship between the number of input channels and compression performance. The figure illustrates the changes in bpp, PSNR, and MS-SSIM as the number of input channels increases from 2 to 256 in powers of 2. As the number of channels increases, BPP gradually decreases. For most channel configurations, the scale hyperprior achieves a lower bitrate and higher reconstruction quality than the factorized prior, demonstrating the synergy between the architectural differences and channel expansion.

These findings indicate that multi-channel input offers intrinsic advantages for compression beyond mere data augmentation. The model can represent the data with shorter code lengths by integrating duplicated structures shared between the real and imaginary components and across multiple propagation distances into the latent representation. We therefore conclude that the multi-channel design is an effective strategy for reducing data volume in hologram compression.

4.3. Evaluation based on reconstructed images

In this section, we evaluate the image quality based on numerically reconstructed images obtained after hologram compression. Figure 4 presents reconstructions for holograms generated from Caltech256 test images that were not included in the training set. Each hologram was compressed and decoded with the compared methods, followed by numerical reconstruction. The reconstruction wavelength was 520 nm, and to assess robustness with respect to the propagation

OPTICS CONTINUUM

distance, we used z = 0.20, 0.50, and 1.00 m. As quality metrics, we report the peak signal-to-noise ratio (PSNR) and multi-scale structural similarity (MS-SSIM) [36], together with the bitrate (*bpp*).

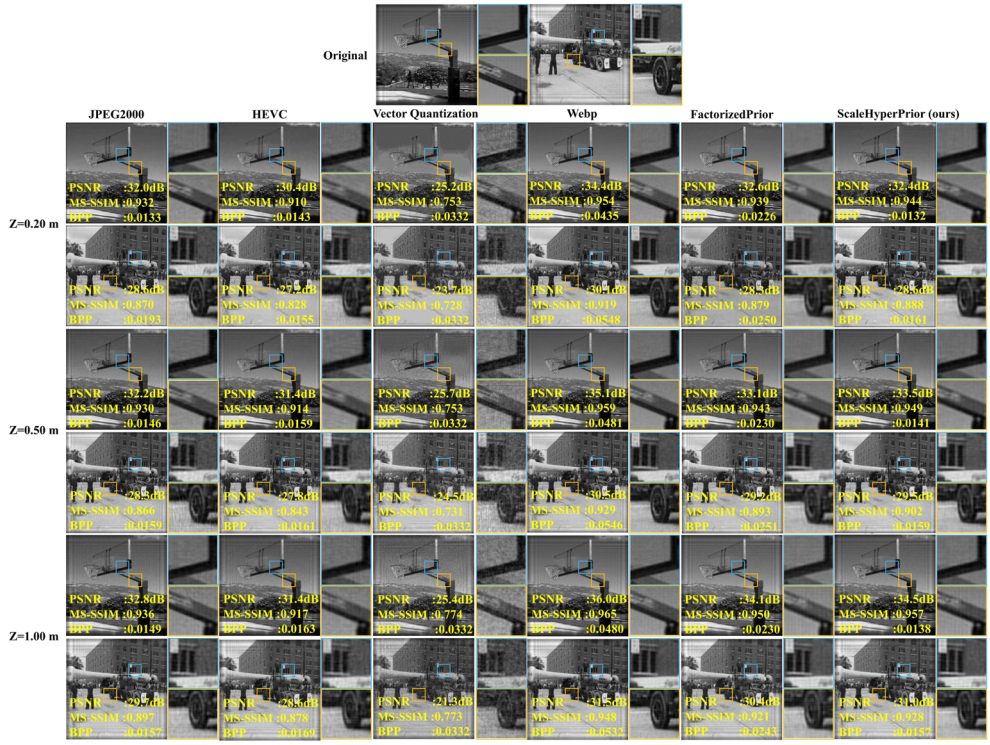

Fig. 4. Reconstructed images evaluated by PSNR, MS-SSIM, and bitrate (*bpp*). These results are based on holograms generated from test images not included in the training set. Each hologram was compressed, decoded, and then numerically reconstructed. Here, z denotes the propagation distance. Evaluation at multiple propagation distances enables a robust, distance-independent performance assessment.

For comparative analysis, we included results obtained from conventional general-purpose codecs, specifically JPEG2000, WebP, and high efficiency video coding (HEVC), along with a vector quantization-based approach. For the latter, the Linde–Buzo–Gray (LBG) algorithm [37] was employed. To achieve high compression ratios, the codebook size and block size were set to 4 and 8, respectively. WebP compression was implemented via the FFmpeg library, and its performance was evaluated across various compression levels by adjusting the quality parameter. To ensure a fair comparison, both the factorized prior and scale hyperprior versions of the proposed method were trained and evaluated using a fixed input channel count of 16.

Although prior studies commonly computed PSNR and MS-SSIM directly on the compressed holograms, such pixel-domain errors do not necessarily correlate with the quality of the reconstructed images, which are the final observable. Therefore, we evaluated the quality by comparing the reconstructed images before and after compression.

The upper row of Fig. 4 shows results for a basketball-hoop image. JPEG 2000 and HEVC exhibit loss of high-frequency content and blocking artifacts, leading to blurred edges in the reconstructions. Vector quantization–based coding suffers from severe quality degradation due to aggressive compression, while WebP compression likewise struggles to preserve fine structural details. The factorized prior model attains PSNR comparable to conventional codecs but produces

OPTICS CONTINUUM

higher MS-SSIM, indicating better preservation of structural features. Our scale hyperprior model achieves the best performance in both PSNR and MS-SSIM; in particular, PSNR is higher by approximately 1.7 dB compared to JPEG 2000. This suggests that by modeling the latent distribution more precisely, the scale hyperprior preserves both global and local structures as well as high-frequency components without degradation. Consequently, these results demonstrate that the proposed method achieves superior reconstruction quality relative to conventional codecs, particularly under low-bitrate conditions.

4.4. Comparison of rate–distortion (R–D) characteristics

Sections 4.2 and 4.3 demonstrated that increasing the number of input channels markedly improves coding efficiency, confirming the effectiveness of the multi-channel design for reducing data volume. This section further assesses the proposed method from an image-quality perspective by evaluating its rate–distortion (R–D) characteristics.

For comparison, we considered factorized prior, JPEG 2000, WebP, HEVC, and vector quantization-based coding. Conventional codecs (JPEG 2000, WebP, and HEVC) and vector quantization were applied directly to complex holograms, while factorized prior and scale hyperprior models were trained on the same hologram dataset under identical settings.

Figure 5 shows the R–D curves: the horizontal axis is the bitrate (*bpp*), and the vertical axis reports PSNR (left) and MS-SSIM (right). The results indicate that the conventional codecs (JPEG 2000 and HEVC) generally underperform the learned methods, with pronounced degradation in the low-bitrate regime. In contrast, the learned factorized prior and scale hyperprior models substantially outperform the conventional codecs. Moreover, the scale hyperprior model surpasses the factorized prior across the entire bitrate range, with a particularly notable gap around 0.01 bpp. This improvement is attributable to the hyperprior's more precise modeling of the latent distribution.

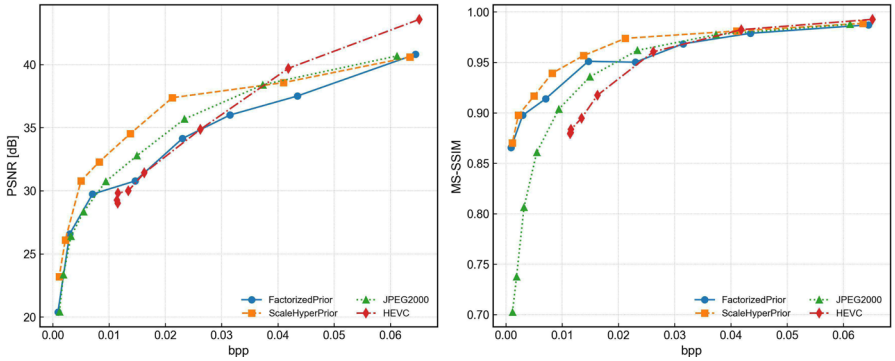

Fig. 5. Rate–distortion characteristics for reconstructed images. PSNR versus bitrate (*bpp*) (left) and MS-SSIM versus bitrate (*bpp*) (right).

These findings confirm that the proposed multi-channel scale hyperprior model is effective for hologram compression and can significantly mitigate quality degradation at low bitrates.

4.5. Computational cost evaluation

Beyond compression performance and reconstruction fidelity, we evaluated the computational complexity of the proposed method to assess its practical viability. Specifically, we report the number of model parameters alongside the execution times for both encoding and decoding.

Regarding model complexity, the factorized prior model comprises approximately 3.1×10^6 parameters, whereas the scale hyperprior model contains roughly 5.2×10^6 parameters. This

OPTICS CONTINUUM

represents an approximately 1.7-fold increase, primarily attributed to the introduction of the hyper-encoder and hyper-decoder networks. Nevertheless, given the gains in compression performance and reconstruction quality established in Sections 4-1 and 4.2, the overall model size remains within a practical and relatively lightweight range.

Next, we evaluated the execution time of the proposed method, using the widely adopted JPEG2000 standard as a baseline for comparison. For the proposed approach, we employed the scale hyperprior model with 16 input channels and measured the average execution time required to compress and decompress 320 holograms. Table 2 summarizes the encoding, decoding, and total execution times per hologram.

Table 2. Comparison of execution time per hologram

Method	Encode [s]	Decode [s]	Total [s]
JPEG2000	0.39807	0.00014	0.39821
Scale hyperprior	0.01833	0.02250	0.04083

As shown in Table 2, the proposed method achieves an approximately 9.8× reduction in total processing time per hologram compared to JPEG2000. Notably, a substantial acceleration is observed in the encoding stage; this can be attributed to the inherent parallel processing capabilities of the deep learning architecture and the efficiency of batch-based compression for multiple holograms.

On the other hand, regarding decoding time, JPEG 2000 is exceptionally efficient as it involves only entropy decoding and the inverse discrete wavelet transform. In contrast, the proposed method relies on deep neural network inference, leading to a longer decoding latency. We acknowledge this as a current limitation and identify it as a key area for future optimization. Nevertheless, the substantial reduction in encoding time demonstrates that our approach offers distinct advantages for applications requiring the rapid compression and transmission of large-scale holographic data.

Overall, while the learning-based nature of the proposed method results in a higher computational cost than non-learning-based codecs like JPEG 2000, both the encoding and decoding times remain within a practical range, and the model size is relatively lightweight. Given the substantial improvements in compression efficiency and reconstruction quality established in Sections 4.3 and 4.4, we conclude that this additional computational overhead is well justified by the resulting performance gains.

4.6. Evaluation of compression performance with ringing suppression

In this section, we evaluate model performance when ringing reduction is applied to the dataset. To isolate the impact of ringing artifacts on compression, we processed existing complex holograms with two methods—Fresnel integrals and blank aperture—and constructed two additional datasets. Together with the original dataset (without ringing reduction), we trained and evaluated the same scale hyperprior model under identical settings.

Figure 6 shows the relationship between input channel count and *bpp* for each dataset. We observe that in the range of 16–64 channels, the datasets with ringing reduction yield notably lower *bpp* than the original dataset. This suggests that ringing reduction removes unnecessary high-frequency components, improving the efficiency of the latent representation and thereby reducing the bitrate. In other words, ringing reduction is effective for achieving higher compression with learned codecs.

Figure 7 presents reconstructed images obtained from holograms generated using Caltech256 test images not included in the training set. Following compression and decoding with each method under ringing-reduction datasets, numerical reconstruction was performed with a wavelength of $\lambda = 520$ nm and propagation distances of $z = 0.20$, 0.50, and 1.00 m. For the proposed method, the input comprised 32 channels. The results show that the proposed

OPTICS CONTINUUM

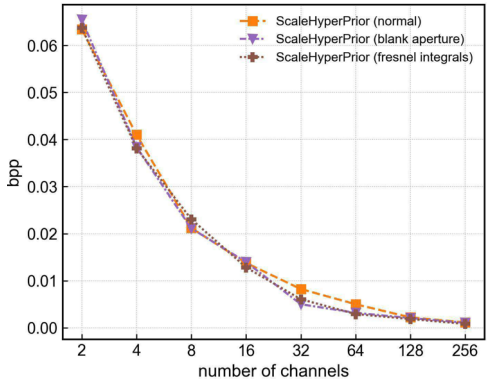

Fig. 6. Relationship between input channel count and bitrate (*bpp*) for each dataset. Datasets with ringing suppression show reduced *bpp*, particularly for 16–64 channels.

method maintains a high compression ratio while minimizing degradation across all propagation distances. In particular, the scale hyperprior model combined with blank-aperture preprocessing achieves an approximately $1,600\times$ compression ratio with a PSNR of 31.2 dB, outperforming conventional codecs such as JPEG2000 and HEVC. Furthermore, even at short propagation distances ($z = 0.20$ m), the method accurately reproduces high-frequency components without loss, effectively preserving edges and text regions.

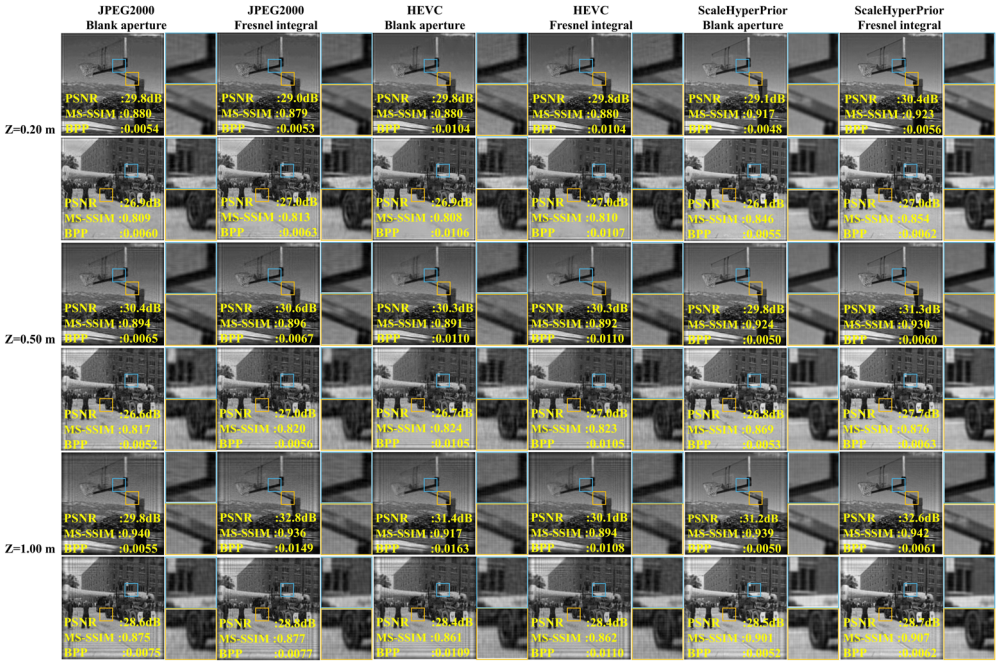

Fig. 7. Numerical reconstructions from Caltech256 test images after compression and decoding with each method under ringing-suppressed datasets ($\lambda = 520$ nm, $z = 0.20$, 0.50, and 1.00 m; 32 input channels for the proposed method). The proposed method maintains a high compression ratio with minimal degradation across all propagation distances. In particular, the scale hyperprior with blank-aperture preprocessing achieves approximately $1,600\times$ compression and 31.2 dB PSNR at $z = 1.00$ m.

OPTICS CONTINUUM

These findings demonstrate that the proposed framework exhibits low dependency on propagation distance and successfully learns both high-frequency and spatial characteristics inherent in holograms. The combination of ringing reduction and a learned codec enables more efficient coding in the latent space.

Figure 8 shows the rate–distortion characteristics across the three datasets. As in Fig. 4, we evaluate the image quality of numerically reconstructed results for holograms generated from the basketball-hoop image. For all datasets, PSNR and MS-SSIM increase with bitrate. However, in the low-bitrate regime (0.003–0.015 *bpp*), the datasets with ringing reduction exhibit superior quality metrics compared with the original dataset, for both the proposed method and conventional codecs (JPEG 2000 and HEVC). Between the two suppression methods, the Fresnel-integrals dataset yields relatively higher PSNR and MS-SSIM than blank aperture. A plausible explanation is that Fresnel-integral-based suppression preserves components closer to the original hologram structure, thereby maintaining compressibility.

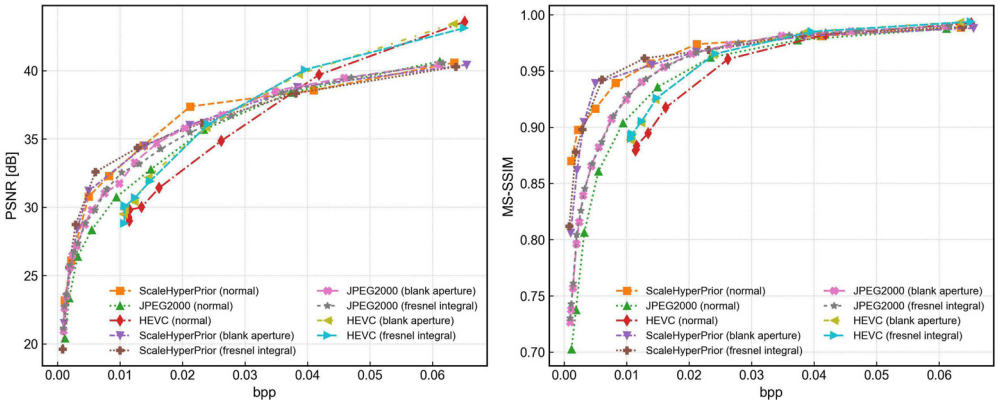

Fig. 8. Rate–distortion characteristics across datasets. In the low-bitrate regime, ringing reduction yields better quality metrics for both neural network and conventional codecs.

Overall, these results suggest that suppressing ringing reduces high-frequency noise components, stabilizes latent learning in the scale hyperprior model, and improves compression efficiency. Ringing artifacts are thus a non-negligible factor limiting compression performance at low bitrates; removing them as a preprocessing step benefits both neural codecs and conventional image codecs in terms of compression efficiency and reconstruction quality.

4.7. Evaluation on high-resolution dataset

Finally, to evaluate the generalization capability of the proposed method, we applied the model trained on the Caltech256 dataset to a high-resolution dataset, DIV4K-50 [38], and assessed its compression performance and reconstruction quality. DIV4K-50 consists of 4K-resolution images (4096×4096 pixels). In this experiment, two images were selected, and the central 2048×2048 regions were used as evaluation data.

Figure 9 shows numerically reconstructed images from holograms generated using DIV4K-50 after compression and decoding with each method. Numerical reconstruction was performed at a wavelength of $\lambda = 520$ nm for propagation distances of $z = 0.20$ m, 0.50 m, and 1.00 m. In general, high-resolution images contain abundant high-frequency components, which are prone to loss during compression, often leading to severe degradation in reconstruction quality. However, the experimental results confirm that the proposed method successfully maintains high compression efficiency (low bpp) while minimizing reconstruction degradation across all propagation distances. In particular, at $z = 1.00$ m, the proposed approach achieved substantially

OPTICS CONTINUUM

higher PSNR and MS-SSIM values than conventional methods, demonstrating its ability to reconstruct high-quality images even under low-bitrate conditions. These results indicate that the proposed model, although trained solely on Caltech256, generalizes well to high-resolution datasets such as DIV4K-50, achieving both high compression performance and faithful image reproduction. This confirms the strong generalization capability of the proposed framework.

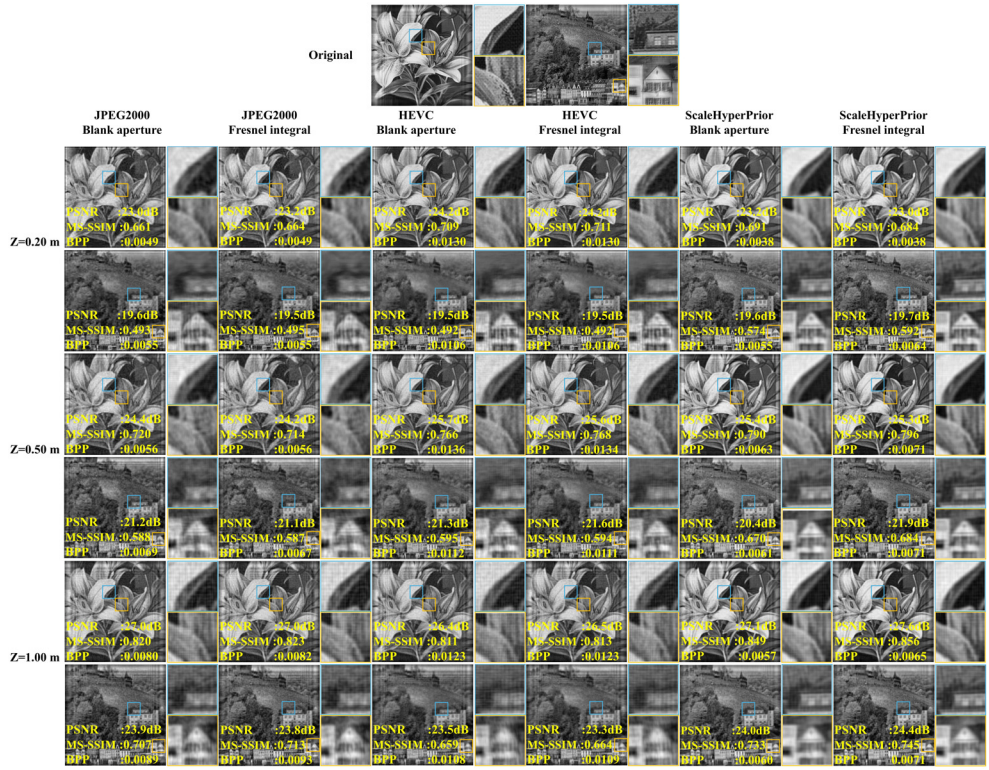

Fig. 9. Numerical reconstructions from the DIV4K-50 dataset after compression and decoding with each method under ringing-suppressed conditions (λ = 520 nm, z = 0.20 m, 0.50 m, and 1.00 m). Despite being trained on Caltech256, the proposed method achieves high compression ratios and superior reconstruction quality across all propagation distances, demonstrating strong generalization to high-resolution data.

4.8. Applicability to 3D holograms

While this study primarily focuses on holograms generated from 2D images, efficient compression of holograms derived from 3D object models is crucial for practical holographic applications. Accordingly, to investigate the generalization capability of the proposed method, we evaluated its compression performance on holograms generated from a 3D object model.

The 3D object model of a "papillon" butterfly [39] was employed as test data. Holograms were generated under optical reconstruction conditions featuring a propagation distance of 0.50 m and an object thickness of 1 cm. The wavelength and pixel pitch were configured according to the parameters specified in Table 1. Following compression and decoding, numerical reconstruction was conducted at multiple propagation distances (z = 0.50 m and 0.51 m), where the proposed scale hyperprior method was benchmarked against JPEG2000 and HEVC. To mitigate artifacts, a ringing reduction technique based on the Fresnel integrals method was applied. The resulting comparisons are presented in Fig. 10.

OPTICS CONTINUUM

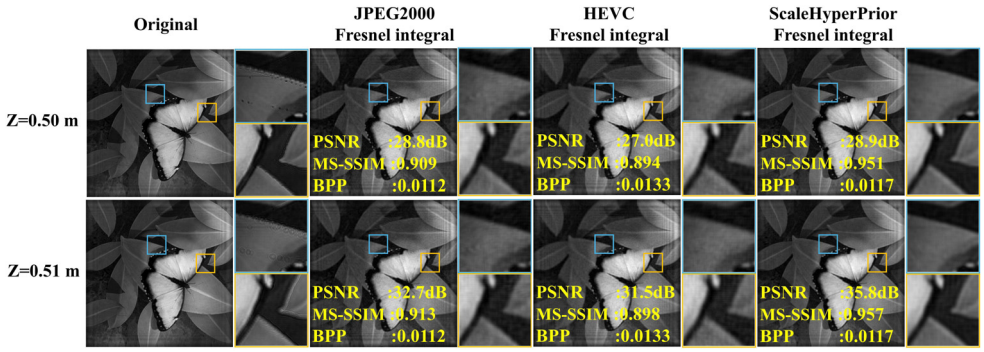

Fig. 10. Numerical reconstructions of holograms generated from a 3D papillon object model [39] after compression and decoding with each method. The proposed method maintains a high compression ratio with minimal degradation and achieves a PSNR of 35.8 dB at a propagation distance of z = 0.51 m.

As illustrated in Fig. 10, the proposed method can be effectively applied to holograms generated from 3D objects. Compared with conventional methods, the proposed approach achieves high objective quality metrics. In particular, for numerical reconstruction at a propagation distance of z = 0.51 m, the proposed method achieves a PSNR of 35.8 dB while maintaining a high compression ratio. The proposed method also yields higher MS-SSIM values than the conventional methods, indicating superior perceptual quality.

Furthermore, it is confirmed that the depth information intrinsic to 3D objects is effectively preserved throughout the compression process. As illustrated in Fig. 10, the reconstructed images at distinct propagation distances (0.50 m and 0.51 m) exhibit selective focus and blurring depending on the focal plane, thereby demonstrating that three-dimensional spatial information is faithfully reproduced. This result indicates that the proposed method is not merely limited to compressing two-dimensional representations but is capable of encoding holograms without compromising the depth-related cues inherent to 3D scenes.

It is worth emphasizing that the model utilized in this evaluation was trained exclusively on holograms derived from 2D images in the Caltech256 dataset, as detailed in Section 4.1, and did not include any 3D object-based holograms during the training phase. Nevertheless, the proposed method achieves favorable compression performance on 3D object holograms, suggesting that the deep learning-based framework effectively captures generalized feature representations even from a limited data domain. Furthermore, by fine-tuning the training datasets and hyperparameters, it is anticipated that models specifically optimized for the unique characteristics of 3D objects can be further developed.

These results demonstrate that the proposed method is effective not only for holograms generated from 2D images but also for those derived from 3D object models, thereby exhibiting a high degree of generalization. This confirms that the proposed framework is versatile and can be broadly applied to a diverse range of holographic applications.

4.9. Optical experiments

For the optical experiments, we utilized holograms of the 3D object "Papillon" [39], as previously evaluated in Section 4.8. First, the complex holograms compressed and decoded by each method were obtained as numerical data. Subsequently, these complex holograms were converted into phase-only holograms using the binary amplitude encoding method [40,41]. Finally, the corresponding optical reconstruction images were captured using a dedicated optical setup.

OPTICS CONTINUUM

Figure. 11 presents the optical reconstruction results. The proposed method preserves edges and intricate structural details more effectively than conventional methods, demonstrating high consistency with the numerical reconstruction results shown in Fig. 10. While some degradation is present relative to the original hologram, the overall geometry and depth information of the 3D object are sufficiently reproduced. These results confirm that the essential information required for optical reconstruction is accurately retained even under high compression ratios.

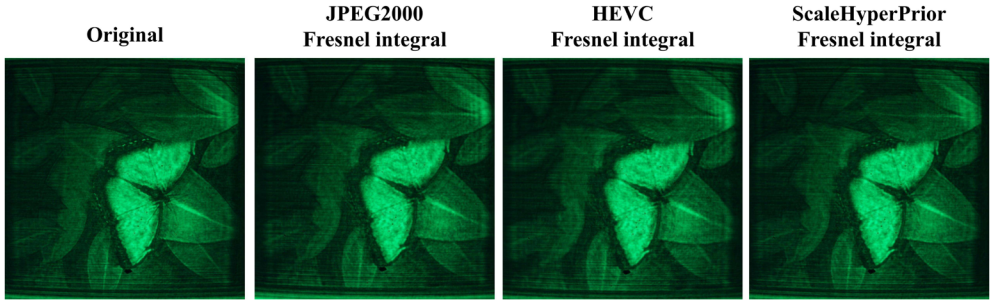

Fig. 11. Optical reconstruction results of the Papillon hologram. Although slight degradation is observed compared with the original hologram, the proposed method (scale hyperprior) enables successful optical reconstruction. Compared with conventional methods, edges and fine structural details are better preserved.

5. Conclusion

We extended CompressAI to develop a hologram-aware learned compression method tailored to complex holograms. By training a scale hyperprior model on paired real and imaginary components, the proposed approach achieved more efficient encoding of hologram-specific high-frequency content compared with conventional codecs designed for natural images. Furthermore, introducing a multi-channel input representation that jointly encodes multiple holograms (e.g., at different propagation distances) effectively exploited inter-hologram redundancy, yielding higher compression ratios.

We further investigated the impact of diffraction-induced ringing on compression by constructing datasets in which ringing was suppressed using the Fresnel integrals method and the blank aperture method, and by training and evaluating the same model under identical settings. The ringing-suppressed datasets improved rate–distortion performance, especially at low bitrates, confirming that ringing components hinder coding efficiency.

In addition, experiments applying the model trained on the Caltech256 dataset to the high-resolution DIV4K-50 dataset demonstrated that the proposed method maintains high compression performance and reconstruction quality even for unseen, higher-resolution data. This confirms that the proposed framework exhibits strong generalization and adaptability across different holographic data distributions.

These results validate the effectiveness of learned compression specialized for holographic data. Future work includes designing loss functions that directly optimize the quality of numerically or optically reconstructed images, and extending end-to-end training to full-resolution holograms.

Funding. Japan Society for the Promotion of Science (22H03607, 19H01097, 24H00687); JST Japan-Taiwan Collaborative Research Program (JPMJKB2310); National Science and Technology Council (113-2923-E-003-001-MY3); IAAR Research Support Program, Chiba University, Japan; Fonds Wetenschappelijk Onderzoek (12ZQ223N, G089424N, G0A3O24N).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

OPTICS CONTINUUM

References

1. Y. Takaki, Y. Matsumoto, and T. Nakajima, "Color image generation for screen-scanning holographic display," Opt. Express **23**(21), 26986–26998 (2015).
2. B. Lee, D. Yoo, J. Jeong, *et al.*, "Wide-angle speckleless dmd holographic display using structured illumination with temporal multiplexing," Opt. Lett. **45**(8), 2148–2151 (2020).
3. M. Makowski, A. Kowalczyk, M. Bieda, *et al.*, "Miniature holographic projector with cloud computing capability," Appl. Opt. **58**(5), A156–A160 (2019).
4. K. Yamaguchi, M. Onishi, and Y. Sakamoto, "Evaluation of transmission performance for streaming cgh video based on computer holography and electro-holography," in *2024 International Conference on Information Networking (ICOIN)*, (IEEE, 2024), pp.689–694.
5. D. Blinder, A. Ahar, S. Bettens, *et al.*, "Signal processing challenges for digital holographic video display systems," Signal Processing: Image Communication **70**, 114–130 (2019).
6. Y. Pan, X. Xu, X. Liang, *et al.*, "Holographic display system using spatial light modulators," in *Proceedings of SPIE*, vol. 8644 (2013), pp.86440F.
7. Y. Xing, M. Kaaniche, B. Pesquet-Popescu, *et al.*, *Digital holographic data representation and compression* (Academic Press, 2015).
8. P. Schelkens, A. Ahar, A. Gilles, *et al.*, "Compression strategies for digital holograms in biomedical and multimedia applications," Light: advanced manufacturing **3**(3), 1–621 (2022).
9. P. Tsang, K. W. K. Cheung, and T. C. Poon, "Low-bit-rate computer-generated color fresnel holography with compression ratio of over 1600 times using vector quantization [invited]," Appl. Opt. **50**(34), H42–H49 (2011).
10. A. E. Shortt, T. J. Naughton, and B. Javidi, "Nonuniform quantization compression techniques for digital holograms of three-dimensional objects," in *Proceedings of SPIE, Optical Information Systems II*, vol. 5557 (2004), pp.0-0.
11. H. Yoshikawa and J. Tamai, "Holographic image compression by motion picture coding," in *Practical Holography X*, vol. 2652 (SPIE, 1996), pp.2–9.
12. P. Memmolo, M. Paturzo, A. Pelagotti, *et al.*, "New high compression method for digital hologram recorded in microscope configuration," in *Proceedings of SPIE, Modeling Aspects in Optical Metrology III*, vol. 8083 (2011), pp.80830W.
13. A. Chen, J. Gui, X. Ma, *et al.*, "Compression of color digital hologram using wavelet thresholds and two-times quantization," Opt. Commun. **537**, 129439 (2023).
14. T. J. Naughton, Y. Frauel, B. Javidi, *et al.*, "Compression of digital holograms for three-dimensional object reconstruction and recognition," Appl. Opt. **41**(20), 4124–4132 (2002).
15. D. Blinder, T. Bruylants, H. Ottevaere, *et al.*, "Jpeg 2000-based compression of fringe patterns for digital holographic microscopy," Opt. Eng. **53**(12), 123102 (2014).
16. J. P. Peixeiro, C. Brites, J. Ascenso, *et al.*, "Holographic data coding: Benchmarking and extending hevc with adapted transforms," IEEE Trans. Multimedia **20**(2), 282–297 (2018).
17. K.-J. Oh, H. Ban, S. Choi, *et al.*, "Hevc extension for phase hologram compression," Opt. Express **31**(6), 9146–9164 (2023).
18. P. Schelkens, T. Ebrahimi, A. Gilles, *et al.*, "Jpeg pleno: Providing representation interoperability for holographic applications and devices," ETRI journal **41**(1), 93–108 (2019).
19. R. Kizhakkumkara Muhamad, T. Birnbaum, A. Gilles, *et al.*, "Jpeg pleno holography: scope and technology validation procedures," Appl. Opt. **60**(3), 641–651 (2021).
20. T. Birnbaum, R. K. Muhammad, C. Perra, *et al.*, "Jpeg pleno holography presents the numerical reconstruction software for holograms: an excursion in holographic views," Appl. Opt. **62**(10), 2462–2469 (2023).
21. R. K. Muhamad, D. Blinder, and P. Schelkens, "Interfere: a high-throughput, view-selective codec for holography," in *Practical Holography XXXIX: Displays, Materials, and Applications*, vol. 13390 (SPIE, 2025), pp.40–46.
22. Y. Rivenson, Y. Zhang, H. Günaydin, *et al.*, "Phase recovery and holographic image reconstruction using deep learning in neural networks," Light:Sci. Appl. **7**(2), 17141 (2017).
23. H. Ko and H. Y. Kim, "Deep learning-based compression for phase-only hologram," IEEE Access **9**, 79735–79751 (2021).
24. H. Ban, S. Choi, J. Y. Cha, *et al.*, "NHVC: Neural holographic video compression with scalable architecture," in *2024 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, vol. 9446 (IEEE, Orlando, FL, USA, 2024), pp.969–978.
25. J. Seo, J. Lee, J. Lee, *et al.*, "Deep compression network for enhancing numerical reconstruction quality of full-complex holograms," Opt. Express **31**(15), 24573 (2023).
26. L. Shi, R. Webb, L. Xiao, *et al.*, "Neural compression for hologram images and videos," Opt. Lett. **47**(22), 6013–6016 (2022).
27. S. Jiao, Z. Jin, C. Chang, *et al.*, "Compression of phase-only holograms with jpeg standard and deep learning," Appl. Sci. **8**(8), 1258 (2018).
28. Z. Dong, Y. Ling, C. Xu, *et al.*, "Gaze-contingent efficient hologram compression for foveated near-eye holographic displays," Displays **79**, 102464 (2023).
29. J. Bégaint, F. Racapé, S. Feltman, *et al.*, "Compressai: a pytorch library and evaluation platform for end-to-end compression research," arXiv (2020).
30. J. Ballé, D. Minnen, S. Singh, *et al.*, "Variational image compression with a scale hyperprior," arXiv (2018).

OPTICS CONTINUUM

31. T. Shimobaba, C.-H. Huang, T. Tahara, *et al.*, "Fast ringing artifact reduction in diffraction calculations using fresnel integrals," J. Opt. **27**(5), 055705 (2025).
32. Y. Nagahama, "Reducing ringing artifacts for hologram reconstruction by extracting patterns of ringing artifacts," Opt. Continuum **2**(2), 361–369 (2023).
33. J. Ballé, V. Laparra, and E. P. Simoncelli, "End-to-end optimized image compression," in *Proceedings of the 5th International Conference on Learning Representations (ICLR)*, (Toulon, France, 2017).
34. Y. Blau and T. Michaeli, "Rethinking lossy compression: The rate–distortion–perception tradeoff," in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, (2019), pp.675–685.
35. G. Griffin, A. Holub, and P. Perona, "Caltech-256 object category dataset," Tech. rep., California Institute of Technology (2007).
36. Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Proceedings of the 37th Asilomar Conference on Signals, Systems and Computers*, (2003), pp.1398–1402.
37. Y.-C. Lin and S.-C. Tai, "A fast linde-buzo-gray algorithm in image vector quantization," IEEE Trans. Circuits Syst. II **45**(3), 432–435 (1998).
38. Y. Zuo, Q. Zheng, M. Wu, *et al.*, "4kagent: agentic any image to 4k super-resolutionquantization," arXiv (2025).
39. S. Wanner, S. Meister, and B. Goldluecke, "Datasets and benchmarks for densely sampled 4d light fields," *VMV*, vol. 13, (2013), pp.225–226.
40. T. Shimobaba, T. Takahashi, Y. Yamamoto, *et al.*, "Simple complex amplitude encoding of a phase-only hologram using binarized amplitude," J. Opt. **22**(4), 045703 (2020).
41. T. Shimobaba, F. Wang, J. Starobrat, *et al.*, "Comparison of double-phase hologram and binary amplitude encoding: holographic projection and vortex beam generation," Appl. Opt. **62**(28), 7471–7479 (2023).