applied optics

INTERFERE, short-time Fourier-transform-based compression of complex-valued holograms with bit depth and range-adaptive quantization

Raees Kizhakkumkara Muhamad,[1,2,†,*] Tobias Birnbaum,[1,2,†] David Blinder,[1,2,3,†] and Peter Schelkens[1,2]

[1]Department of Electronics and Informatics (ETRO), Vrije Universiteit Brussel (VUB), Pleinlaan 2, B-1050 Brussels, Belgium
[2]Department of Digital and User Centric Solutions (DUCS), imec, Kapeldreef 75, B-3001 Leuven, Belgium
[3]Graduate School of Engineering, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba, Japan
[†]These authors contributed equally to this work.
[*]raees.kizhakkumkara.muhamad@vub.be

Received 2 June 2024; revised 13 August 2024; accepted 14 August 2024; posted 14 August 2024; published 28 August 2024

With digital holographic display and recording setups steadily improving and the advent of realistic super-high-resolution holograms (>100 megapixels), the efficient compression of digital holograms (DHs) becomes an urgent matter. Therefore, JPEG Pleno holography is undergoing a standardization effort to address this challenge. The accepted, current baseline coding solution for lossy compression of complex-valued DHs, entitled INTERFERE, is presented in this paper. Its features include a simple and modular overall architecture, high scalability, view-selective coding, low decoder complexity, and the highest rate-distortion performance among state-of-the-art solutions. We also introduce, to our knowledge, a novel meta-quantization strategy that can be used for signals exhibiting large variations in dynamic range in the domain being quantized. We were able to demonstrate on the versatile JPEG Pleno hologram database BD-rate reductions between 16% and 272% (average of 119%) over HEVC for achieving an SNR in the range 5–25 dB. With this first compression standard on DHs, we hope to provide an essential building block for their future commercialization in large-scale consumer markets.

https://doi.org/10.1364/AO.531422

1. INTRODUCTION

A. Problem Statement and Objectives

Holographic [1] techniques record the interference pattern produced by light waves emanating from three-dimensional scene objects on a flat two-dimensional plane, such that the object wavefront may be reconstructed later from the recording. Holography has many practical applications, ranging from interferometry for metrological applications to the visual consumption of 3D multimedia. Interferometric modalities like digital holographic microscopy [2] and tomography [3] can be used in wide-ranging biological or material science applications to perform quantitative volumetric measurements of 3D shapes/refractive indices, study turbulence, heat distributions, etc.

Since holography reproduces physically accurate wavefields, it is also considered the holy grail for 3D display applications like tabletop displays and VR/AR headsets [4–6]. Used as an imaging or display technique, holography provides all depth cues without suffering from many of the issues seen in conventional 3D displays, such as the accommodation-vergence conflict [7], incorrect or even missing focal cues, discretized motion parallax, etc.

Emerging applications in this domain face large data storage and processing requirements [8], which hinders widespread adoption. For example, extremely small pixel sizes in the range of a few hundred nanometers are required for visualization. Here, the pixel pitch (p) used and wavelength (λ) of the emitted light are related to the angular field of view θ^{FOV} of the display by the grating equation [9] as

$$\theta^{\mathrm{FOV}} = 2\sin^{-1}\left(\frac{\lambda}{2p}\right). \tag{1}$$

Displays with a high space-bandwidth product, i.e., simultaneously supporting a large field of view and viewing aperture size, have enormous pixel counts and require massive amounts of data to be driven. Signaling only the scene information, which is usually much smaller than the native hologram representation, and generating the hologram on the client side are one potential solution to the data transmission problem. However, computationally generating holograms is a challenging task as holography is non-local, i.e., each scene point can potentially

affect all pixels in the holograms. This limits that approach to only simple scenes and/or requires extremely powerful hardware at the client side [10].

If this bandwidth problem were not to be addressed, it could significantly delay the onset of holographic display technology. Holographic signals have properties and statistics that differ substantially from natural photographic imagery because they consist of complicated interference patterns. That is why conventional image and video codecs are sub-optimal for this purpose. Therefore, novel adapted hologram coding solutions must be designed to solve the problem of compressing and transmitting digital holograms.

B. State-of-the-Art in Hologram Compression

Data compression techniques can be applied to mitigate transmission costs. Naturally, the decoding procedure should require much less computation than generation to make compression worthwhile. Quantization techniques that aim to reduce the representation accuracy of the hologram in their native format are the most straightforward of such compression approaches [11]. Scalar and vector quantization techniques were studied for the compression of phase-shifted holograms in [12] and off-axis Fresnel holograms in [13]. Compression approaches eschewing explicit quantization also exist, like the optical compression protocol [14] proposed for off-axis Fresnel holographic videos, which utilizes a virtual lens to scale down the input signal without affecting the fringe structure. To compress phase-only holograms, [15] obtains a lossy alternative phase representation with fewer discontinuities and lower entropy. In the case of phase-only holographic video, [16] proposes an alternative hologram generation method that introduces an inter-frame correlation that can be exploited by a computationally simple delta coding step while improving compression performance over conventional video codecs.

Most existing image and video compression techniques also satisfy the computational requirement but show sub-optimal rate-distortion performance for DHs. DHs can have significant content all over the frequency spectrum; see, e.g., [17]. This is in contrast to natural imagery, which is predominantly low pass and where the amplitude of the spectra follows a reciprocal relation with increasing frequency values [18]. This difference can be perceived when inspecting raw hologram data. The microstructure of DHs appears as noise due to the higher frequency content. Compression strategies in conventional image coding tools strive to preserve predominantly low-frequency content. For example, JPEG 2000 Part I uses the Mallat decomposition-based, multi-resolution wavelet transform, which provides a sparse representation for smoothly varying content [19]. JPEG/HEVC profiles often employ quantization tables, so more bits are spent for low-pass content [20]. The motivation is that the human visual system's CSF (contrast sensitivity function) is less sensitive to errors in higher spatial frequencies. However, these approaches do not apply to holography, which encodes interference patterns. Here, the higher spatial frequencies of the hologram also relate to larger viewing angles. Quantizing or pruning those higher spatial frequencies will reduce the reconstruction quality of views at larger angles.

Another approach compresses DHs by propagating them to an in-focus/object plane, so some objects are in focus [21]. Object plane compression works best if there is a single, small, and shallow object in the hologram field of view, as minimal scene information will be out of focus. Then, classical image compression solutions would perform well [22]. A unitary transform that can apply piecewise depth propagation in the time-frequency domain given a depth map of the scene was proposed in [23], which allows this approach to be generalized to complex, deep scenes at the expense of computational complexity. Nonetheless, the propagation step in these approaches incurs a vast computational burden and requires additional information about the scene composition. While free-space light propagation is mathematically invertible [23], it is potentially vulnerable to aliasing issues when employing lossy compression. Conversely, a key advantage of the above solutions is that integrating well-established image codecs will represent lesser upfront costs to design specialized hardware/software. Thus, significant efforts have also been expended to modify conventional codecs to handle holographic data better. For example, Fresnel hologram sequences were compressed using AVC [24] after a 2D-discrete cosine transform was applied so that the fringe patterns could be better characterized [25]. In a similar vein, the compression efficiency of HEVC on phase-only holograms can be markedly improved by techniques like phase unwrapping [26] or by modifying clipping operations performed within the codec itself [27].

Several transform-based compression techniques have been proposed over the years for DH. We shall also follow such an approach. In one of the first works [28], the application of the 2D-discrete Fourier transform (DFT) with a user-specified low-pass filter was proposed for the compression of electronic speckle pattern interferometry (ESPI) images. Similarly, 2D-DFTs of size 8×8 were used in [29] for lossy compression, where compression is achieved by setting some n lowest coefficient magnitude values out of the 64 to zero. Other transforms that have been explored targeted the distinct nature of oriented fringes and/or phase space (i.e., the space spanned by spatial coordinates and spatial frequencies [17]) locality [30–32]. In [33], the hologram is compressed using the Fresnelets basis obtained by Fresnel transforming B-spline wavelets. This Fresnelet transform at some wavelet scale relates to the hologram reconstruction at a corresponding distance. In [34], Gabor wavelet bases were proposed for compression and view-dependent transmission, as they are known to be excellent tools for signal analysis. However, this technique is overcomplete and thus contains redundant information.

Recently, there has also been a trend in applying deep learning [35] in the context of hologram compression, often in conjunction with classical image coding tools [36–38]. Deep learning has shown promise in uncovering complex relationships and finding novel applications in many related domains. However, the gains demonstrated by deep learning for hologram compression remain somewhat marginal and/or were only demonstrated in restrictive parameter settings, for example, shallow scenes, large pixel pitches, and small field of view. Ultimately, none of the aforementioned techniques has advanced into more than an academic proposal so far, and henceforth, no solution for the effective compression of DHs for visualization exists. For more

insight into the state-of-the-art of hologram coding and the associated challenges, we defer to [39–41].

C. Proposed Solution and Contributions

To address the compression requirements in holography and other plenoptic modalities, the JPEG committee (ISO/IEC JTC 1/SC 29/WG 1) has been working on a suite of standard specifications called JPEG Pleno [42]. JPEG Pleno Holography (ISO/IEC 21794-5) aims to attain high compression efficiency while enabling efficient hardware/software implementation, ensuring device interoperability and supporting various holographic modalities and representations [43]. A call for coding solutions for holography [44] was launched in 2021 by the JPEG committee, where the INTERFERE solution — presented here — was chosen as the baseline compression method for standardization. INTERFERE comprises two parts [45]: one for the lossless compression of binary DHs and another for the lossy compression of (non-binary) complex-valued holograms. The lossless compression of binary DHs is described in [46]. This work describes the core compression framework utilized for the lossy compression of complex-valued DHs.

We demonstrate that the short-time Fourier transform (STFT) paves the way for computationally efficient lossy hologram compression. The STFT representation is not only conducive to sparsification but also allows for the selection of viewports, i.e., given the position of the user(s), only a fraction of the hologram needs to be signaled depending on the viewing angles. Both factors can help in significantly reducing bandwidth requirements. We also introduce a novel two-stage quantization mechanism for effectively storing signal representations exhibiting large dynamic range variations, like typical holographic data in the STFT domain.

Our compression framework can operate in both lossy and near visually lossless regimes and supports the following features: simple and modular overall architecture, high scalability (easy to parallelize and transfer to specialized hardware), view-selective coding, asymmetric en-/decoder complexity, as well as the highest rate-distortion performance of any existing solution to our knowledge.

In summary, our specific contributions are as follows:

- we give a detailed account and rationale for the codec standardized for JPEG Pleno Holography;
- we present a scalable and random-access supporting architecture based on the STFT, utilizing a four-dimensional (4D) spatio-angular coefficient representation compatible with many holographic display and recording types;
- we present novel two-step quantization, rate-distortion optimization, and entropy coding procedures tailored for hologram coding;
- we report BD-rate reductions ranging from 16% and 272% over H.265/HEVC (average of 119%) and ranging from 36% and 470% over JPEG 2000 (average of 209%) for an SNR in the range of 5–25 dB on the holograms from the JPEG Pleno hologram database.

An initial version of the compression framework without the two-step quantization procedure and entropy coding was developed in [47], while preliminary compression results achieved for a single hologram with the new framework were communicated as part of [45]. This work describes and motivates the processing steps used in the framework and includes comprehensive experimental evaluations. The following section provides an overview of this paper's organization and a high-level description of the en-/decoder, data organization, and processing flow.

2. OVERVIEW OF THE PAPER

The compression procedure starts with a monochrome DH spatially partitioned into tiles. This allows for limiting the memory requirements of the codec in the case of very large hologram resolutions. The encoder independently processes each tile as shown in Fig. 1(a).

Each tile undergoes an STFT transformation as described in Section 3. The transform coefficients are organized into quantization blocks (QBs) and subject to a novel meta-quantization strategy, which will be discussed in Section 4. The parameters governing the quantization procedure are obtained by application of a rate-distortion optimization (RDO) step on the QB data as described in Section 5.

The quantized data and quantizer side information are subject to fixed-point arithmetic encoding using adaptive context modelling and will be elaborated upon in Section 6. The QB information is organized into codeblocks (CBs) for entropy

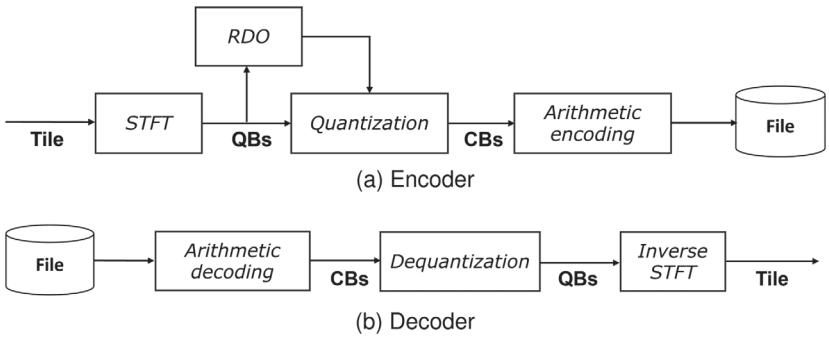

Fig. 1. Overview of compression proposal defined for a spatially partitioned (not shown) tile. Data structures are given in bold font, while processes are in italics. STFT: short-time Fourier transform, QB: quantization block, RDO: rate-distortion optimization, CB: codeblock.

coding. CBs are the random access unit in our compression system—that is, each CB can be decoded independently. Each QB collects STFT coefficients from a neighborhood in 4D phase space. Likewise, each CB is obtained as a collection of QBs from a 4D neighborhood partitioning the 4D phase space volume of an STFT-transformed tile.

The decoder, shown in Fig. 1(b), performs a straightforward reversal of the steps performed by the encoder, starting from one or more CBs. The CBs are subject to fixed-point arithmetic decoding to retrieve the quantized data and quantizer side information. After obtaining this information, dequantization is applied to the quantized STFT coefficients, followed by the inverse STFT procedure to obtain the decoded tile.

Exhaustive benchmarks for the compression performance regarding state-of-the-art image compression codecs, with both objective and subjective analysis, are provided in Section 7. Finally, the work is concluded in Section 8.

The notation used in this work is discussed below.

Variables are denoted in italics, constants are in upright font, vectors are in boldface, and scalars are in normal font. The index for addressing an element of a vector is given as a subscript. For example, x_i is the ith element of vector $\boldsymbol{x}$. The expression $i{:}d$ refers to d contiguous indices $\{i, i+1, \ldots, i+d-1\}$. Vectors can be defined using the following notation: $\boldsymbol{x} = \{x_i \in \boldsymbol{f} | i \in \boldsymbol{I}\}$. This implies the vector is constructed by iterating over all indices $i \in \boldsymbol{I}$. $\boldsymbol{f}$ denotes the domain of x_i. Vectors can be multi-dimensional, where commas separate the subscript indices for each dimension.

$\mathbb{R}^+$ are the non-negative real numbers. $\emptyset$ is the empty set. j is the imaginary unit.

3. SHORT-TIME FOURIER TRANSFORMATION

Holograms are complex-valued amplitude measurements corresponding to a superposition of coherent electromagnetic waves propagating through a medium, typically measured across a 2D plane. For a monochromatic wavelength λ, the hologram is denoted by $h(x, y)$, where x and y represent the two spatial dimensions of the 2D plane. Here the Fourier transform value $H(u, v)$ calculated for some frequency coordinate u, v represents the complex amplitude contribution of the plane wave $e^{2\pi j(ux+vy)}$ travelling with directional cosines $(\lambda u, \lambda v, \sqrt{1-(\lambda u)^2-(\lambda v)^2})$ with respect the hologram plane [48]:

$$H(u, v) = \iint_{-\infty}^{+\infty} h(x, y) e^{-2\pi j(ux+vy)} \mathrm{d}x \mathrm{d}y. \quad (2)$$

The objects in the scene emitting light will excite the various spatial regions in the hologram with plane waves sent in different directions, as guided by the scene geometry [17]. In order to sparsify DHs, we need to leverage the sparsity existing in the plenoptic representation, which should allow for simultaneous representation of space and (spatial) frequency. Many types of such representations are possible, denoted as phase-space representations [49]. The Heisenberg uncertainty relation bounds any given representation's spatial and frequency resolution.

The discrete short-time Fourier transform (STFT) is one such representation. It is special because it is reversible, can form a critically sampled ortho-normal basis, and is easy to implement in hardware. We propose to use a normalized rectangular window with zero overlap (i.e., no redundancy) for our compression scheme. We will hereafter simply refer to this specific variant as STFT. For a 2D hologram tile of size $\mathrm{N_1} \times \mathrm{N_2}$, and using a STFT transform block of dimensions $\mathrm{F_1} \times \mathrm{F_2}$, the 4D STFT is given by Eq. (3) where $u \in \{0{:}\mathrm{F_1}\}$, $v \in \{0{:}\mathrm{F_2}\}$, $x \in \left\{0{:}\frac{\mathrm{N_1}}{\mathrm{F_1}}\right\}$, and $y \in \left\{0{:}\frac{\mathrm{N_2}}{\mathrm{F_2}}\right\}$. The procedure is visualized in Fig. 2:

$$H_{u,v,x,y} = \sum_{x'=0}^{\mathrm{F_1}-1} \sum_{y'=0}^{\mathrm{F_2}-1} h_{x'+\mathrm{F_1}x, y'+\mathrm{F_2}y} \cdot e^{-j2\pi \cdot \left(\frac{ux'}{\mathrm{F_1}} + \frac{vy'}{\mathrm{F_2}}\right)}. \quad (3)$$

The discrete short-time Fourier transform can decompose the hologram into the most dominant plane waves received across the different spatial regions, thus aiding *sparsification*. The STFT was also deployed in our previous work [47] targeting holographic video compression. Here, it was used to compress residual data obtained after subtracting the motion compensation reference hologram from the original hologram.

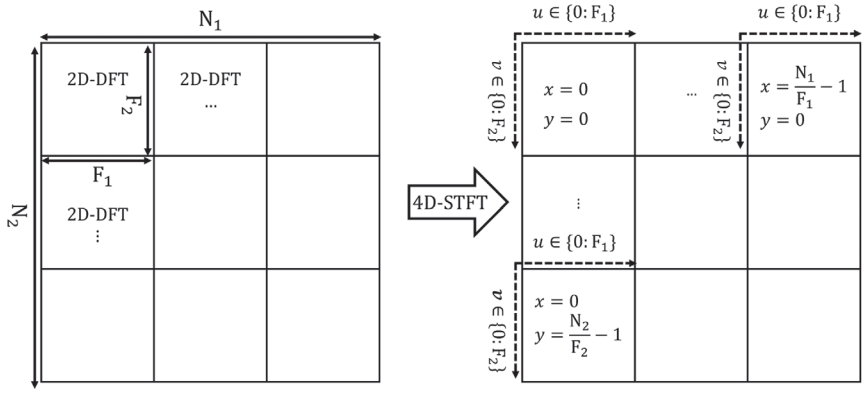

Fig. 2. Illustration of the non-redundant, short-time Fourier transform (STFT) with rectangular spatial windows used in this work. The left-hand side shows that the 4D-STFT is obtained by applying 2D-DFTs on blocks of $\mathrm{F_1} \times \mathrm{F_2}$ pels each for a hologram tile with $\mathrm{N_1} \times \mathrm{N_2}$ pels. The right-hand side shows that the frequency coordinates $\{u, v\}$ of the 2D-DFT obtained from the $\{x, y\}$ block are mapped to the 4D STFT coordinates $\{u, v, x, y\}$. As the 2D-DFT block size $\mathrm{F_1} \times \mathrm{F_2}$ increases, the frequency localization of the STFT representation increases while the spatial localization decreases.

This residual data consists of missing information that can be localized spatially and motion compensation errors that tend to be localized in frequency. The STFT permits trading off spatial and frequency localization with respect to each other, such that a sparse signal representation is obtained. Similarly, in off-axis digital holography microscopy and tomography, only the information belonging to a well-localized region in the frequency spectrum corresponding to the +1-order term is used [50,51].

Apart from sparsity, this decomposition using the STFT can also enable view-dependent hologram coding, where only the information that is visible to a user is transmitted—depending on the position and viewing orientation with respect to the hologram plane [34,52]. The transform coefficient in Eq. (3) represents the component of the complex amplitude of light received along the direction represented by plane wave $e^{j2\pi\cdot(\frac{ux'}{F_1}+\frac{vy'}{F_2})}$ for the spatial segment of size $F_1 \times F_2$ whose indices are given by $\{F_1x{:}F_1, F_2y{:}F_2\}$. Grouping STFT coefficients from neighboring spatial and frequency precincts into independently decodable units provides *spatio-angular random access* that facilitates the selective signaling of views or data only visible from specific viewing positions.

Lastly, the discrete Fourier transform is already *ubiquitously* used in the processing of holographic content [48] for a wide variety of purposes like numerical depth propagation [53,54] in accelerated computer-generated holography (CGH), [55–57], motion compensation [47,58], and tomographic reconstruction [59].

The inverse/forward transform in Eq. (3) can be efficiently computed using a cascade of row-wise and column-wise 1D-(I)FFTs of length F_1 and F_2. The (I)FFT has log-linear computational complexity and is one of the most widely used tools in signal processing, with efficient hardware implementations available for many platforms [60,61].

4. QUANTIZATION

Quantization increases the compressibility of a signal by reducing the total number of possible signal states. As a consequence, it introduces a distortion in the signal representation.

We define our quantizers on 4D hyperrectangles of dimension $QB_1 \times QB_2 \times QB_3 \times QB_4$ in the STFT domain and call them quantization blocks (QBs), as given in Eq. (4):

$$\boldsymbol{QB}_{u,v,x,y} = \boldsymbol{H}_{QB_1u:QB_1,QB_2v:QB_2,QB_3x:QB_3,QB_4y:QB_4}. \quad (4)$$

Thereby, $u \in \{0{:}\frac{F_1}{QB_1}\}$, $v \in \{0{:}\frac{F_2}{QB_2}\}$, $x \in \{0{:}\frac{N_1}{QB_3F_1}\}$, and $y \in \{0{:}\frac{N_2}{QB_4F_2}\}$ in Eq. (4). We shall denote the number of all QBs in a tile as N^{QB} and the set of 4D indices containing the QB positions as $\mathbf{N}^{QB}$. The number of coefficients inside a QB shall be denoted by S^{QB}.

Before we elaborate on the quantization strategy employed to aid sparsification, we will motivate it based on the statistical distribution of the STFT coefficients obtained from typical holograms.

A. Motivation

STFT coefficients of holograms can have a large dynamic range. The key idea here is to choose a small S^{QB} such that quantization can be adapted to the statistical distribution of coefficients in each hyperrectangle. For demonstration, we use the hologram CGH *Venus*, computed from a point cloud by superposing spherical waves originating from each point. Since the CGH is generated using a rigorous and analytical approach, it is free from approximation artifacts. Therefore, it will also serve as a representative hologram for analysis throughout this work.

We take an exemplary QB of size $32 \times 32 \times 1 \times 1$ and plot the histogram (zero-order statistical distribution) of 2048 real/imaginary STFT coefficients belonging to it in Fig. 3(a). Here, we are interested in the change in statistical distribution when these coefficients are further subgrouped into 4 QBs of size $16 \times 16 \times 1 \times 1$ as shown in Fig. 3(b). The variances (alternatively viewed as the energy) and the kurtosis obtained for the distribution from each region are also showcased. In general, as the size of the QB gets smaller, the hologram is split into clusters with higher and lower variances. In other words, the hologram energy polarizes into higher- and lower-energy QBs. Since the representation accuracy of each QB quantizer can be adapted, using smaller-sized clusters helps better localize the hologram energy. Additionally, it can be seen that the distribution of the smaller-sized clusters, shown in Fig. 3(b), tends to be more uniform than the larger-sized QB in Fig. 3(a) as also indicated by their kurtosis. When discretely representing a uniformly distributed signal, the maximum increase in SNR for each additional bit can be calculated from rate-distortion theory as around 6 dB, which is always greater than the maximum increase possible for peaked distributions like Gaussian/Laplacian [62].

However, the tradeoff in opting for a small-sized QB is that the quantizer information of each QB must also be transmitted.

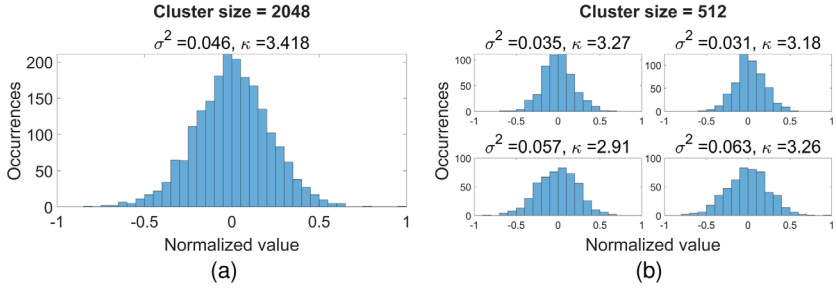

Fig. 3. Histogram (zero-order statistical distribution) of 2048 neighboring real/imaginary STFT coefficients chosen from hologram *Venus* when viewed as (a) single QB of size $32 \times 32 \times 1 \times 1$. (b) Four smaller QBs of size $16 \times 16 \times 1 \times 1$. The real/imaginary values have been normalized by the largest absolute valued real/imaginary coefficient (σ^2—variance, κ—kurtosis).

Our quantization procedure discussed below strives to minimize the quantizer side information while still maintaining quantization adaptivity.

B. Quantization Procedure

For scalar quantization, the Lloyd-Max quantizer design in [63] has been shown to achieve the lower bound for ℓ_2 distortion under certain weak assumptions. Despite its optimality, the Lloyd-Max quantizer is not used in many practical compression scenarios. The determination procedure is iterative and computationally tedious. Additionally, if the signal statistics are not available at the dequantizer, then the mapping that describes the dequantization for the quantization output states needs to be also transmitted,

Often, simpler quantization approaches suffice, like truncating the least significant bits of the signal. It was shown in [64] that at asymptotically high rates, the relation between the entropy of the output signal — after rounding/truncation — and the distortion approaches the Shannon rate-distortion bound for typically encountered signals and distortion functions. This work simplifies quantization using such a uniform scalar mid-rise quantizer (MRQ).

The MRQ quantizer can be uniquely defined by its bit depth and quantization range. For an input $x \in \mathbb{R}$, an integer quantization bit depth $b \geq 1$, and quantization range $X \in \mathbb{R}^+$, the output of the MRQ centered at zero is given by

$$\mathcal{Q}(x, b, X) = \begin{Bmatrix} -2^{b-1}, & \text{if } x - X \\ \lfloor \frac{2^{b-1}x}{X} \rfloor, & \text{if } -X \leq x \leq X \\ 2^{b-1} - 1, & \text{otherwise} \end{Bmatrix}. \quad (5)$$

Applying dequantization on a quantized output x using the uniform mid-rise dequantizer (MRDQ) yields

$$\mathcal{Q}^{-1}(x, b, X) = (x + 0.5)\frac{X}{2^{b-1}}. \quad (6)$$

The transformation undergone by some input due to the quantization procedure is denoted as $\mathcal{Q}^{-1}\mathcal{Q}$:

$$\mathcal{Q}^{-1}\mathcal{Q}(x, b, X) := \mathcal{Q}^{-1}(\mathcal{Q}(x, b, X), b, X). \quad (7)$$

For $b = 0$, $x \in \mathbb{R} \cup \emptyset$, and $X \in \mathbb{R}^+ \cup \emptyset$, the operations of the MRQ and MRDQ are defined as $\mathcal{Q}(x, 0, X) = \emptyset$ and $\mathcal{Q}^{-1}(\emptyset, 0, X) = 0$.

Using the MRQ, we quantize the STFT coefficients within a QB as follows: we split the complex values into real/imaginary parts [65,66] and quantize each part using the same MRQ within a QB. Per QB, the coefficient bit depth $b \in \mathbf{B} = \{0 : \mathrm{b}^{\max} + 1\}$ and QB quantization range $X \in \mathbb{R}^+$. If $b = 0$, then the coefficients in the QB are essentially skipped, and the other bit depths shall be denoted by $\mathbf{B}^+$.

As motivated in Section 4, we pursue the quantization strategy of keeping the QB dimensions small, such that the quantization can better address the statistical variation that occurs in the STFT domain for holographic content when optimizing for the minimum overall error. However, the drawback for permitting such a fine level of adaptivity is that the coefficient quantization bit depth — used for all QBs and the quantization range for all QBs with non-zero coefficient quantization bit depth — needs to be transmitted as side information to the dequantizer.

The mean-squared error $d(R)$ of the Lloyd-Max quantizer for a signal is approximated as a function of its rate R (in bit) by $d(R) \approx \epsilon^2 \sigma^2 \cdot 2^{-2R}$ (Eq. 3.20 in [62]), where ϵ is a parameter derived from the signal's probability distribution function. This relation hints that the target bit allocation for a QB will be correlated to the (signal) energy in a logarithmic manner. We can expect the target bit depth allocation for the QBs to vary smoothly across phase space and, therefore, be amenable to entropy coding.

In the case of the target QB coefficient quantization range, we also propose quantizing the range. That is, the representation of the quantizer is quantized itself, leading to the meta-quantization philosophy. We can expect a correlation between the target QB bit depth allocation and the target QB range, which is exploited in this design. For all QBs whose coefficients are quantized with some $b \in \mathbf{B}^+$, its target quantization range is further quantized with a uniform MRQ having a bit depth $q \in \mathbf{Q} = \{0 : \mathrm{q}^{\max} + 1\}$. For positive q, the MRQ is centered around an offset $Q^{\text{off}} \in \mathbb{R}^+$, and has a range $Q \in \mathbb{R}^+$. Here, the offset Q^{off} is used because the target QB quantization ranges are not zero-mean and consequently help to reduce the effective dynamic range encountered by the QB range quantizer. In case $q = 0$, the QB range is assigned a constant value of Q^{off}.

Summarized, the following target parameters need to be determined for the quantization procedure at the encoder, also shown in Fig. 4(a).

- The bit depth allocation vector $\boldsymbol{b} = \{b_{\boldsymbol{i}} \in \mathbf{B} | \boldsymbol{i} \in \mathbf{N}^{\mathrm{QB}}\}$ containing the bit depths used for quantizing the STFT coefficients of each QB.

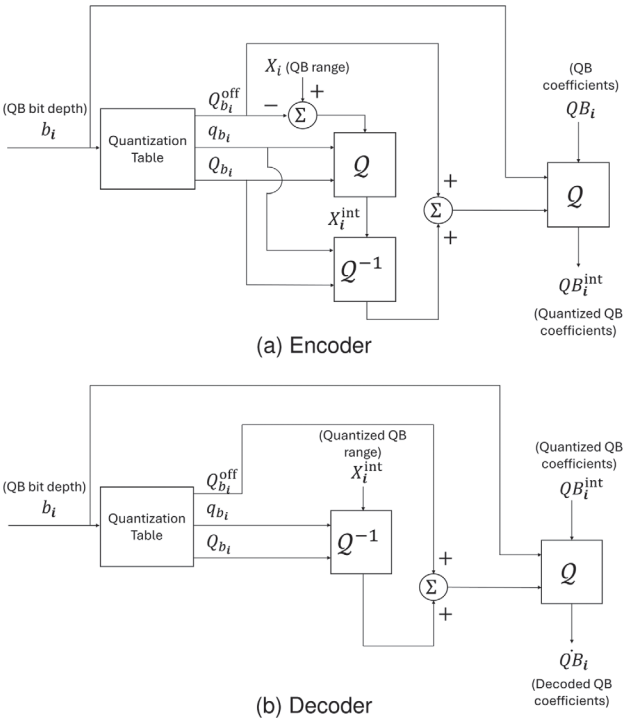

Fig. 4. Quantization and dequantization procedures performed at the encoder and decoder for the $\boldsymbol{i}$th QB are shown.

• The quantization range vector $\boldsymbol{X} = \{X_{\boldsymbol{i}} \in \mathbb{R}^+ \cup \emptyset | \boldsymbol{i} \in \mathbf{N}^{\mathrm{QB}}\}$ for each QB. The value $\emptyset$ is assigned for those QBs whose coefficients were quantized with a bit depth $b = 0$.

• The quantization table represents the one-to-one mapping between the QB coefficient bit depth and the parameters used for the QB range quantization.

- The meta-quantizer bit depth vector $\boldsymbol{q} = \{q_b \in \mathbf{Q} | b \in \mathbf{B}^+\}$.

- The range offset vector $\boldsymbol{Q}^{\mathrm{off}} = \{Q_b^{\mathrm{off}} \in \mathbb{R}^+ \cup \emptyset | b \in \mathbf{B}^+\}$. The value $\emptyset$ is assigned to a bit depth b if it does not occur in $\boldsymbol{b}$, i.e., when no QBs were assigned a coefficient quantization bit depth b.

- The meta-quantizer range vector $\boldsymbol{Q} = \{Q_b \in \mathbb{R}^+ \cup \emptyset | b \in \mathbf{B}^+\}$. The value $\emptyset$ is assigned to a bit depth b if it does not occur in $\boldsymbol{b}$ or if the meta-quantizer bit depth q_b mapped to b is zero.

The quantized QB range and quantized QB coefficients belonging to the $\boldsymbol{i}$th QB are obtained at the encoder as shown in Fig. 4(a). These quantized values and QB bit depth will be stored after entropy coding with the quantization table.

The decoder will utilize these components to perform the dequantization procedure to obtain the dequantized QB coefficients as shown in Fig. 4(b).

A natural question follows: "How are these target parameters determined at the encoder?" This forms the crux of the discussion given in the next section. Here, we utilize a rigorous approach for estimating the influence of the various parameters on compression performance by computing the true distortion from quantization. Additional heuristics, like parametrically estimating distortion [62], can be used to speed up the estimation procedure, but this is beyond the purview of this work. In any case, the meta-quantization does not appreciably increase the complexity of the dequantization done at the decoder, as the additional steps are trivial and performed once per QB.

5. RATE-DISTORTION OPTIMIZATION

Lossy compression codecs typically require a bitrate or distortion target such that the codec can adjust its internal operating parameters to reach the same. This adjustment process is called rate-distortion optimization (RDO). We use a global ℓ_2 distortion target D^{tar} calculated per hologram tile.

In the case of a non-orthonormal transform, the problem stated above is intractable due to the extremely large number of possible parameter sets. Furthermore, the global distortion for some parameter sets can only be calculated after applying the inverse sparsifying transform over the entire hologram tile — quantized with the corresponding parametrization. For ortho-normal transforms, like the non-redundant STFT variant used here, the theorem of Plancherel holds, and the ℓ_2 error of the coefficients is the ℓ_2 error in the original hologram domain. Therefore, the ℓ_2 error in the hologram domain can be optimized independently on small hyperrectangles in the transform domain, rendering computation feasible.

The RDO procedure utilizes the Lagrangian multiplier method in which the distortion of the quantizers is calculated explicitly [67]. Other Lagrangian-based optimization strategies estimate rate-distortion behavior rather than calculating exact values. In contrast, the employed approach does not require convexity assumptions on quantizer rate-distortion behavior. The following additional numerical optimization procedures are used for determining the most suitable RDO parameter set.

• `UniMini` $(f(\cdot), \{y_1, y_2\})$ returns y, where $f(y)$ is minimal in $[y_1, y_2]$ for a unimodal function. We solve this problem with the golden section search [68].

• `MonoSearch` $(f(\cdot), t, \{y_1, y_2\})$ returns y, where $f(y) = t$ given that t lies between $f(y_1)$ and $f(y_2)$ for a monotonic function. We use the binary search algorithm [69].

A. RDO Overview

The RDO proceeds as follows. In the first step, an oracle vector $\mathbf{X}^{\mathrm{ora}}$ containing the optimal values of quantization ranges $X_{\boldsymbol{i},b}^{\mathrm{ora}}$ to be used for each QB $\boldsymbol{i} \in \mathbf{N}^{\mathrm{QB}}$, at each potential bit depth $b \in \mathbf{B}^+$ its coefficients may be quantized at, is determined:

$$\mathbf{X}^{\mathrm{ora}} = \left\{X_{\boldsymbol{i},b}^{\mathrm{ora}} \in \mathbb{R}^+ | \boldsymbol{i} \in \mathbf{N}^{\mathrm{QB}}, b \in \mathbf{B}^+\right\}. \tag{8}$$

The optimal oracle quantization range to be used for the uniform MRQ quantizer depends on the tradeoff between two diametrical errors: the granular and clipping errors. The granular error represents the quantization error that occurs when the input is within the quantization range. When increasing the quantization range while keeping the bit depth constant, the granular error will increase due to the increase in quantization step (bin) size.

The clipping error refers to the error encountered for inputs above or below the quantization range. It decreases as the quantization range increases. Figure 5 shows for an exemplary QB how the distortion produced by the MRQ changes with respect to the quantization range for different bit depths. It can be seen that the distortion, which is the summation of these diametrical errors, is a unimodal function of the quantization range. As the bit depth of the MRQ quantizer increases, the granular error reduces, and the value of the optimal quantizer range increases [47]. To determine the oracle quantization ranges, we use the `UniMini` numerical optimization procedure as shown in Eq. (9):

$$X_{\boldsymbol{i},b}^{\mathrm{ora}} = \texttt{UniMini}\left(D^{\mathrm{QB}}(\boldsymbol{i}, b, \cdot), \{y_1, y_2\}\right). \tag{9}$$

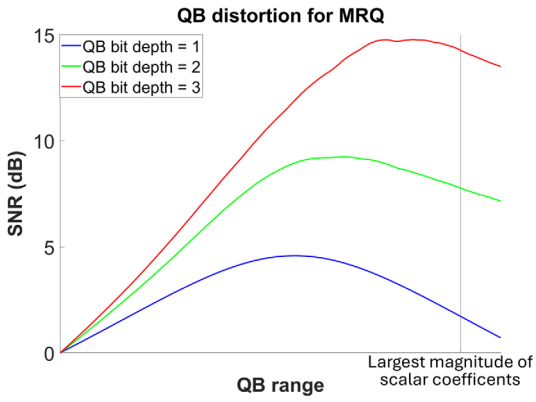

Fig. 5. ℓ_2 distortion dependence of an exemplary QB on MRQ bit depth and range. Result shown for *Venus* with $\mathbf{F} = \{512, 512\}$, $\mathbf{QB} = \{8, 8, 1, 1\}$.

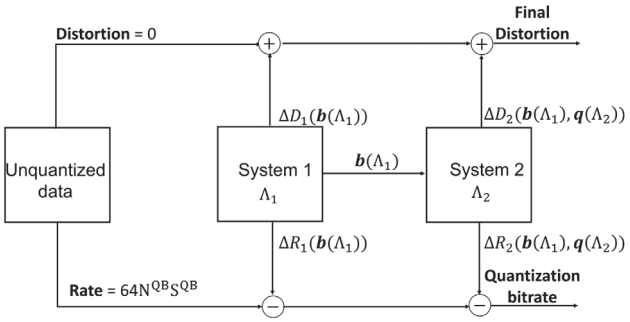

Fig. 6. The quantization procedure is modelled via two coupled sub-systems, Systems 1 and 2. The final bitrate reduction and distortion functions induced by quantization are expressed additively from the corresponding functions of each system. Lagrangian optimization performed individually at some multiplier value Λ_1 for System 1 and Λ_2 for System 2 finds the solution that reduces the bitrate most among all solutions having the same distortion within each system. The RDO strives to pick the tuple $\{\Lambda_1, \Lambda_2\}$ among all such solutions that meet the final distortion target with the lowest bitrate.

The unimodal objective is the ℓ_2 distortion for $\boldsymbol{QB_i}$ when using a bit depth b and quantization range y; see Eq. (10):

$$D^{\mathrm{QB}}(i, b, y) = \left\| \mathcal{Q}^{-1}\mathcal{Q}(\boldsymbol{QB_i}, b, y) - \boldsymbol{QB_i} \right\|_2. \quad \textbf{(10)}$$

The search starts with the endpoints $y_1 = 0$ and y_2, where the latter is the QB's largest absolute valued scalar coefficient. The oracle vector is determined independently of the global ℓ_2 distortion target D^{tar}.

After $\boldsymbol{X}^{\mathrm{ora}}$ is calculated, the RDO strives to find the optimal allocation of bits to the QB coefficients $\boldsymbol{b}^{\mathrm{tar}}$, and the bits for the QB quantization range $\boldsymbol{q}^{\mathrm{tar}}$ along with the other meta-quantizer parameters.

We use two coupled quantization-optimization systems for the RDO: Systems 1 and 2, which we will elaborate upon next.

System 1 performs quantization of the QB coefficients using some (coefficient) bit depth allocation $\boldsymbol{b}$, where we do not consider the influence of the subsequent quantization of the QB range. Here, from the previously determined oracle ranges, $X^{\mathrm{ora}}_{\boldsymbol{i},b_{\boldsymbol{i}}}$ will be used as the quantization range for the QB indexed by $\boldsymbol{i}$. The quantization range is represented as a 32-bit floating point number. The associated reduction of the total rate and the induced distortion of the quantization procedure shall be given by $\Delta R_1(\boldsymbol{b})$ and $\Delta D_1(\boldsymbol{b})$. Out of all possible solutions, the bit depth allocation is restricted to $\boldsymbol{b}(\Lambda_1)$ obtained as the solution minimizing the Lagrangian rate-distortion cost-function obtained by System 1, for some Lagrangian multiplier $\Lambda_1 \in \mathbb{R}^+$. The chosen bit depth allocation shall satisfy the following with respect to any bit depth allocation $\boldsymbol{b}$: $\Delta D_1(\boldsymbol{b}(\Lambda_1)) \leq \Delta D_1(\boldsymbol{b})$ if $\Delta R_1(\boldsymbol{b}(\Lambda_1)) \leq \Delta R_1(\boldsymbol{b})$ [67].

System 2 represents the second stage of the quantization procedure and determines the meta-quantizer bit depth allocation $\boldsymbol{q}$ for the QB quantization ranges, along with $\boldsymbol{Q}^{\mathrm{off}}$ and $\boldsymbol{Q}$. It requires the QB coefficient bit depth allocation $\boldsymbol{b}$ from System 1 as input. Let $\Delta R_2(\boldsymbol{b}, \boldsymbol{q})$ and $\Delta R_1(\boldsymbol{b}, \boldsymbol{q})$ be the induced rate-reduction and distortion due to the quantization by System 2. In an analogous manner to the solution of System 1, the bit depth allocation of System 2 is determined as $q(\Lambda_2)$ for some $\Lambda_2 \in \mathbb{R}^+$, where $\Delta D_2(\boldsymbol{b}, \boldsymbol{q}(\Lambda_2)) \leq \Delta D_2(\boldsymbol{b}, \boldsymbol{q})$ if $\Delta R_2(\boldsymbol{b}, \boldsymbol{q}(\Lambda_2)) \leq \Delta R_2(\boldsymbol{b}, \boldsymbol{q})$.

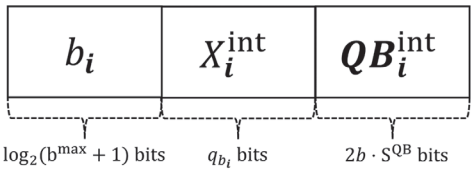

Fig. 7. Payload belonging to the $\boldsymbol{i}$th QB before application of entropy encoding.

Thus, the bit depth allocations are determined using a greedy approach, where the chosen solution is optimal with respect to both systems individually. Note that the ideal compression performance can only be achieved when solving both systems simultaneously, which is computationally intractable.

The final rate and distortion spanned by all operating points under this quantization approach as shown by Fig. 6 are given in Eq. (11), where we have assumed non-quantized STFT coefficients to be 32-bit floating point numbers:

$$D(\Lambda_1, \Lambda_2) = \Delta D_1(\boldsymbol{b}(\Lambda_1)) + \Delta D_2(\boldsymbol{b}(\Lambda_1), \boldsymbol{q}(\Lambda_2)),$$

$$R(\Lambda_1, \Lambda_2) = 64N^{\mathrm{QB}}S^{\mathrm{QB}} - \Delta R_1(\boldsymbol{b}(\Lambda_1)) - \Delta R_2(\boldsymbol{b}(\Lambda_1), \boldsymbol{q}(\Lambda_2)). \quad \textbf{(11)}$$

Out of all possible tuples $\{\Lambda_1, \Lambda_2\}$, the RDO with distortion target strives to pick the best values, such that $D(\Lambda_1, \Lambda_2) = D^{\mathrm{tar}}$ while minimizing $R(\Lambda_1, \Lambda_2)$. For a given D^{tar}, there is only a range of values for Λ_1 for which the distortion constraint can be satisfied. Within this range, for some Λ_1 the required value of Λ_2 can be obtained by solving for the distortion constraint D^{tar}. The `MonoSearch` procedure can be used, as the system distortion (and rate) functions are monotonic with respect to the Lagrangian multiplier.

The search procedure is detailed in Supplement 1. The procedure identifies a search range for Λ_1 and then finds the value Λ_1^{tar} within this search range that minimizes the overall rate. While here we have used a distortion target, the RDO can also be performed using a rate target. However, in this case, the target will serve as an upper bound as the rate will become smaller after entropy coding. On the other hand, distortion is unchanged as entropy coding is lossless.

We demonstrate the utility of the meta-quantization (i.e., the coupled System 1 and System 2) by comparing its compression performance with that of just utilizing System 1 for the hologram *Venus* (as given in Fig. 9). The obtained compression gain reduces as the QB size S^{QB} increases because the cost of the side information is now spread over $2S^{\mathrm{QB}}$ real/imaginary coefficients. Our experiments that evaluate INTERFERE's compression performance as a function of S^{QB} in Supplement 1 show that the best performance is obtained when using small QB sizes, typically in the range 16–64. See Fig. S3 for a few examples.

6. ENTROPY CODING

The final bitstream is split into two parts: the header information and payload information. The header information

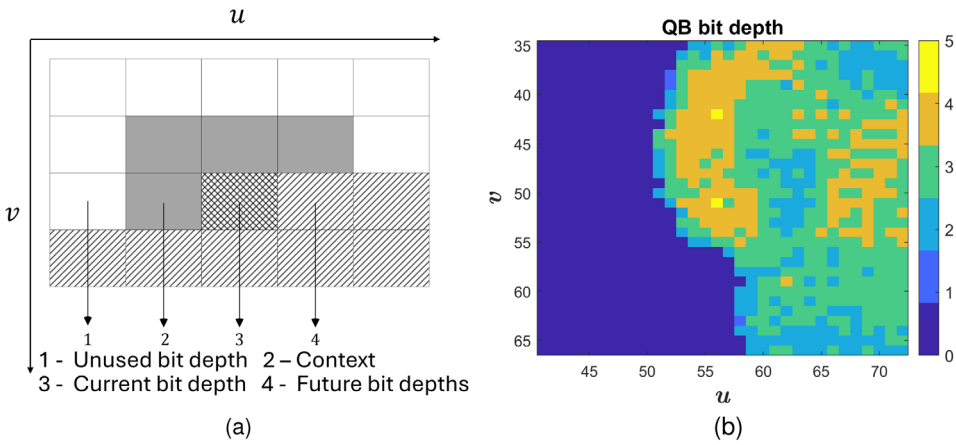

Fig. 8. Entropy coding performed for the QB bit depth is illustrated. (a) Context determination used in the encoding of the QB bit depth. (b) QB bit depths allocated for an exemplary transform block from the *Venus* hologram—$\mathbf{N} = \{4096, 4096\}$, $\mathbf{F} = \{512, 512\}$, $\mathbf{QB} = \{4, 4, 1, 1\}$, $\mathbf{CB} = \{128, 128, 1, 1\}$, and a distortion target of 15 dB.

represents a tiny fraction of the total information and is stored as is, once per hologram tile. The header comprises the following:

- data organization: describes the dimensions used in the construction of the data organization units, namely, the hologram tile, STFT block, quantization block, and codeblock ($\mathbf{N}$, $\mathbf{F}$, $\mathbf{QB}$, $\mathbf{CB}$);
- quantizer side information: contains the maximum bit depths, b^{max} and q^{max}, that the QB coefficient and QB range quantizer can use, respectively. It also consists of the quantization table ($\boldsymbol{q}$, $\boldsymbol{Q}^{off}$, $\boldsymbol{Q}$).

The payload information corresponds to the remaining information on which entropy coding is applied. Here, the payload consists of the QB coefficient bit depths $\boldsymbol{b}$, and the quantized QB range and QB coefficients $\mathbf{X}^{int}$ and $\mathbf{QB}^{int}$.

As explained earlier, we will split the payload information into units called codeblocks (CBs). Each codeblock is a grouping of the payload from $CB_1 \times CB_2 \times CB_3 \times CB_4$ QBs belonging to the 4D space-frequency neighborhood. Each CB, like the QB, is a contiguous 4D partition of the phase-space information of the hologram tile and permits hologram transmission with simultaneous spatial and angular viewports. Apart from enabling random access, partitioning into codeblocks also permits parallel processing for the entropy en-/decoding procedures, which will certainly be required to support the throughput requirements imposed by holographic content. We will now elaborate on how the information is organized within a codeblock.

A. Entropy Coding Procedure

For entropy coding, the payload information belonging to all QBs per CB is grouped together as shown in Fig. 7. As indicated, the QB coefficient bit depth needs to be en-/decoded first, followed by the quantized QB range and the quantized QB STFT coefficients.

Real and imaginary QB STFT coefficients are interleaved. For obtaining the codeblock payload, the QB payloads within a codeblock are strung together. Thus, each CB payload comprises a stream of integer-valued symbols representing the information belonging to the constituent QBs.

For the entropy coding of the CB payload, we use adaptive entropy coding with the fixed point arithmetic coder given in [70], where we assume *no a priori* information about the hologram. Simply put, the entropy coder "learns" from the information that was previously en-/decoded, such that for a stationary signal, the compression performance improves over time.

Three symbol types are used depending on whether the symbol is the QB bit depth, the QB range, or the QB coefficients. The QB bit depth is encoded using a fixed template Markov-based histogram model comprising bit depths of the four immediate previously coded QB neighbors in the same STFT transform block, as shown in Fig. 8(a). The primary motivation behind this statistical model is the observation that bit depths from the neighboring regions are highly correlated. An example is given in Fig. 8(b).

The QB range and QB coefficients are coded using zero-order histogram models. A separate model is mapped to each QB bit depth for either.

The compression gain obtained by the entropy coding module is given in Fig. 9 for the hologram *Venus*. It can be seen

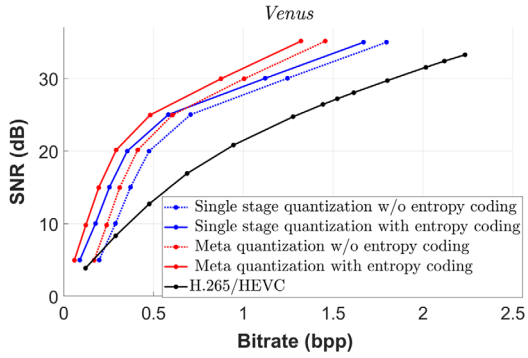

Fig. 9. Rate-distortion curves obtained by the different compression modes of INTERFERE—$\mathbf{N} = \{4096, 4096\}$, $\mathbf{F} = \{512, 512\}$, $\mathbf{QB} = \{4, 4, 1, 1\}$, $\mathbf{CB} = \{64, 64, 1, 1\}$.

that enabling entropy coding brings noticeable benefits. Also noteworthy is the influence of adaptive quantization on the compression gain seen by the entropy coder. When the quantizer is ill-tuned to the data, most quantizer outputs are mapped to only a small subset of its output alphabets. Entropy coding can be highly beneficial here, helping to regain some of the "lost" performance from poor adaptivity in quantization. In other words, when less adaptive quantization models are used, the burden of compression shifts to the entropy coding stage.

7. RESULTS

In this section, we discuss the compression performance achieved by INTERFERE. The experiments reported in this paper have been carried out according to the Common Test Conditions [71] defined for the JPEG Pleno Holography standardization effort. The tests were conducted on a Windows 10 PC equipped with an i7 11800H processor and dual-channel DDR4 3200 MHZ RAM. The results have also been independently verified in the context of evaluation conducted by the JPEG committee [72–77].

As test data, a diverse selection of complex-valued DHs is used. The test data contains both computer-generated holograms (CGHs) and optically captured holograms (OCHs) with varying scene complexities. All holograms used [78–84], except for *Venus*, are available at the JPEG Pleno database [85]. The recording and generation parameters of these holograms are given in Table S2 in Supplement 1. The dimensional information describing the various data organization units used by the INTERFERE codec for these holograms is included. The influence of the choice of these parameters on compression performance is also discussed. The encoder and decoder software of INTERFERE was written in ISO C ++17 using Microsoft Visual Studio 2022.

For comparison, JPEG 2000 (Kakadu 8.0.5 [86]) and intra-mode H.265/HEVC (HM 16.20) are used. The coding conditions used for both are specified in Supplement 1. Since JPEG 2000 and HEVC require unsigned integer input representations, the floating-point complex-valued holograms are split into real and imaginary components and quantized to integer representation before compression [71]. 16-bit input modes are used for both codecs to introduce only minimum loss from this procedure. In the case of color holograms, each wavelength/color channel is en-/decoded independently for all compression solutions.

The compression performance of lossy compression methods can be compared objectively by rate-distortion curves that plot a selected distortion metric calculated on the lossy decoded hologram $\dot{\boldsymbol{b}}$ as a function of the achieved bitrate. The bitrate is measured as

$$\text{bitrate (bpp)} = \frac{\text{Number of bits in encoded file}}{\text{Total pixels in hologram}}. \tag{12}$$

The distortion is measured by the SNR calculated over the global hologram wavefield as given in Eq. (13). Other distortion metrics, such as SSIM (variants) and even subjective scores, were studied in [75]:

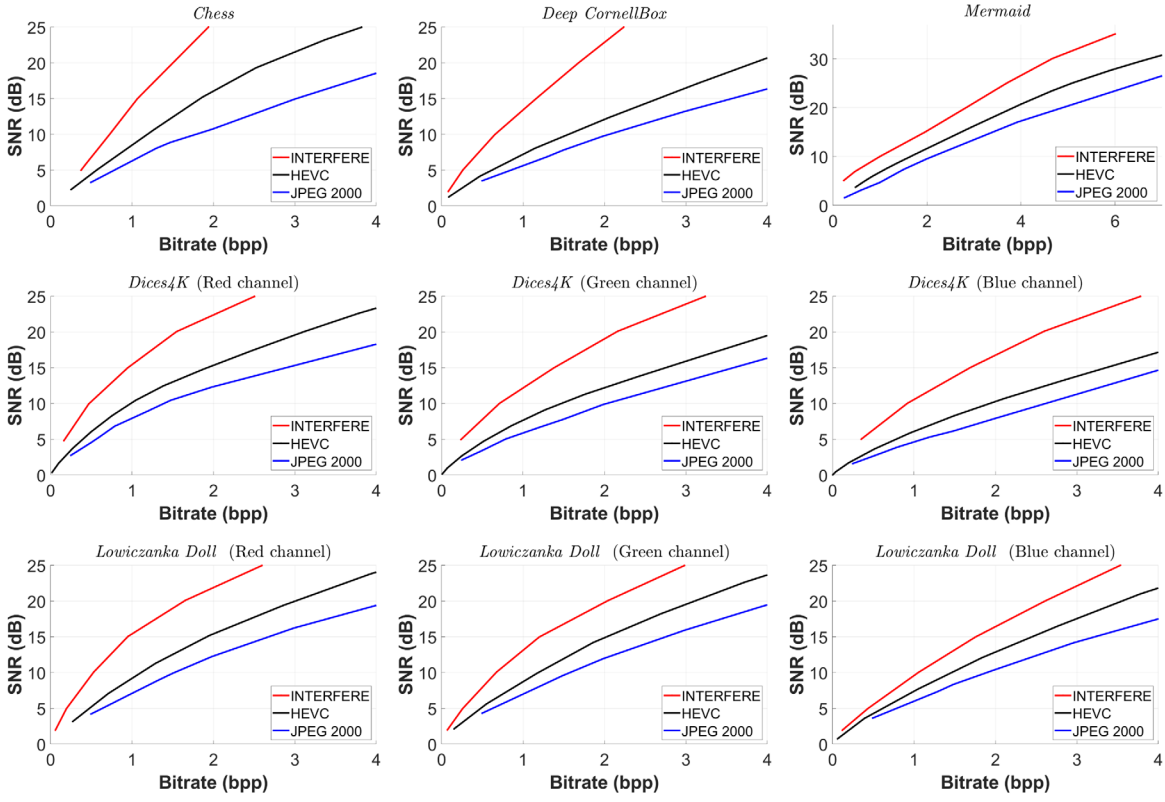

Fig. 10. Rate-distortion plots comparing INTERFERE with JPEG 2000 and HEVC are shown for a selection of test holograms from the JPEG Pleno Database. See Fig. 11 to compare numerical reconstructions from different compression solutions.

$$\mathrm{SNR\ (dB)} = 10\log_{10}\left(\frac{\|\boldsymbol{b}-\dot{\boldsymbol{b}}\|_2}{\|\boldsymbol{b}\|_2}\right). \tag{13}$$

The rate-distortion plots comparing INTERFERE with other compression solutions are shown in Fig. 10 for a selection of test holograms. It can be seen that INTERFERE obtains a convincing improvement over conventional image codecs.

The BD-rate metric can summarize the relative rate-distortion compression performance between two codecs across a quality region of interest $[D_1, D_2]$, as shown in Eq. (14):

$$\mathrm{BD\text{-}Rate}(\%) = \frac{100\cdot\int_{D_1}^{D_2}(R_\mathrm{B}(D)-R_\mathrm{A}(D))\mathrm{d}D}{\int_{D_1}^{D_2}R_\mathrm{A}(D)\mathrm{d}D}, \tag{14}$$

where $R_\mathrm{A}(D)$ and $R_\mathrm{B}(D)$ are the bitrates for the codecs for achieving distortion D. A positive value corresponds to a better relative performance of the reference codec A. We use the hologram wavefield's SNR as the distortion metric, with the INTERFERE codec as a reference. The distortion is measured across two quality regions of interest. We select a perceptually lossy region (5–15 dB) as well as a perceptually lossless region (15–25 dB).

The BD rates were calculated for H.265/HEVC Intra and JPEG 2000 across all tested holograms.

For HEVC, the geometric mean of the BD rates was +98% for the perceptually lossy region and +81% in the perceptually lossless region. In the case of JPEG 2000, the geomean was 182% and 145%, respectively. Table S2 in Supplement 1 lists the BD rates obtained for each test hologram.

The objective results demonstrate bitrate gains of 40%–400% across the entire dataset. Only holograms with resolutions smaller than 2048 × 2048 show less impressive gains of 10%–50%. This showcases the robustness of the proposed method to optical recording noise, CGH algorithms, hologram parameters, and scene content. Albeit this, slightly larger improvements can be seen in the case of CGH compared to OCH. The presence of measurement noise in the case of OCH reduces the gain provided by the quantization procedure. Furthermore, in the case of CGH, higher utilization of the space-bandwidth is possible. This can allow for more spatio-angular information to be present, like in the *Deep CornellBox* hologram, leading to an increased "stress test" for the codecs.

The hologram is propagated to the plane containing the object to evaluate visual quality using NRSH (Numerical Reconstruction Software for Holography) V8.0 [87, 88]. In the case of color holograms, the same bitrates are used for all wavelengths. This results in one or more reconstructions generated per hologram. The number of reconstructions depends on the number of objects in the scene and the viewing angle range supported by the DH. NRSH utilizes different reconstruction algorithms depending on the diffraction regime of the hologram type under test. The generated reconstruction is of 8-bit integer precision and can be viewed on a conventional 2D-display [83]. In fact, elaborate subjective quality experiments [83] indicate that test subjects could perceive compression artifacts more easily on 2D displays when compared to holographic displays.

We show reconstructions obtained by operating the codes at approximately the same bitrates for a selection of lossy compressed holograms in Fig. 11, where we have chosen low bitrates to highlight each codec's compression artifacts better. Since holograms encode 3D information, is it important to assess whether the visual information is well-preserved from all viewpoints. Conventional codecs give higher weight to lower frequencies of the signal, leading to non-uniform distortion of the angular information, as opposed to our proposed STFT-based codec. The PSNR, which objectively measures the visual quality of a given reconstruction, is also indicated alongside the reconstructions:

$$\mathrm{PSNR\ recon.(dB)} = 10\log_{10}\left(\frac{\|\boldsymbol{r}-\dot{\boldsymbol{r}}\|_2}{255^2}\right). \tag{15}$$

The improvement of INTERFERE over the two codecs measured objectively by the SNR is also perceptible from the numerical reconstructions. We see a pervasive presence of unstructured noise in the reconstructions of conventional codecs. Both these codecs have aliasing issues with some holograms (see, e.g., *Mermaid*), where unwanted copies of scene objects are visible. This issue is more pronounced in the case of JPEG 2000 [89]. Furthermore, a vignetting effect is seen in the reconstructions of these codecs, especially for corner views. As mentioned earlier, this arises from the design of these codecs prioritizing low spatial frequencies inspired by the contrast sensitivity function of the human visual system. As a result, more bits are spent on on-axis content, while the off-axis content is skipped, especially at lower bitrates. For example, center views of *DeepCornellBox* show a comparable level of detail with respect to INTERFERE for both the codecs, but most of the information in the side view is gone. In the case of OCH *Mermaid*, the reconstruction from INTERFERE looks even "better" than the uncompressed data. This can be attributed to the frequency filtering induced by the adaptive quantization procedure that removes most of the measurement noise at low bitrates. For a more detailed quality evaluation of the INTERFERE compression proposal with respect to these codecs, the reader is directed to [75].

Encoding the *Dices4K* hologram having 4096 × 4096 pixels using INTERFERE with parameters $\mathbf{F} = \{256,256\}$, $\mathbf{QB} = \{4, 4, 1, 1\}$, and $\mathbf{CB} = \{64,64, 1, 1\}$ while targeting an SNR of 10 dB took around 100 s and is dominated by the RDO step. Decoding takes around 2.22 s. For achieving the same quality level with Kakadu 8.0.5 and HM 16.20, encoding takes around 0.42 s and 380 s, while decoding takes 0.66 s and 5.19 s, respectively. Note that Kakadu 8.0.5 is commercial software, while HM 16.20 and INTERFERE are verification models written to serve as a reference implementation. Therefore, the processing times of HM 16.20 and INTERFERE are not fully indicative of achievable performance.

8. CONCLUSION

In this work, we have described an end-to-end pipeline for the compression of static holograms. It utilizes the STFT using non-redundant rectangular windows as a sparsifying transform, along with novel quantization, rate-distortion optimization, and entropy coding procedures. With the employed scheme, variations in dynamic range across STFT regions seen for typical holograms can be efficiently addressed.

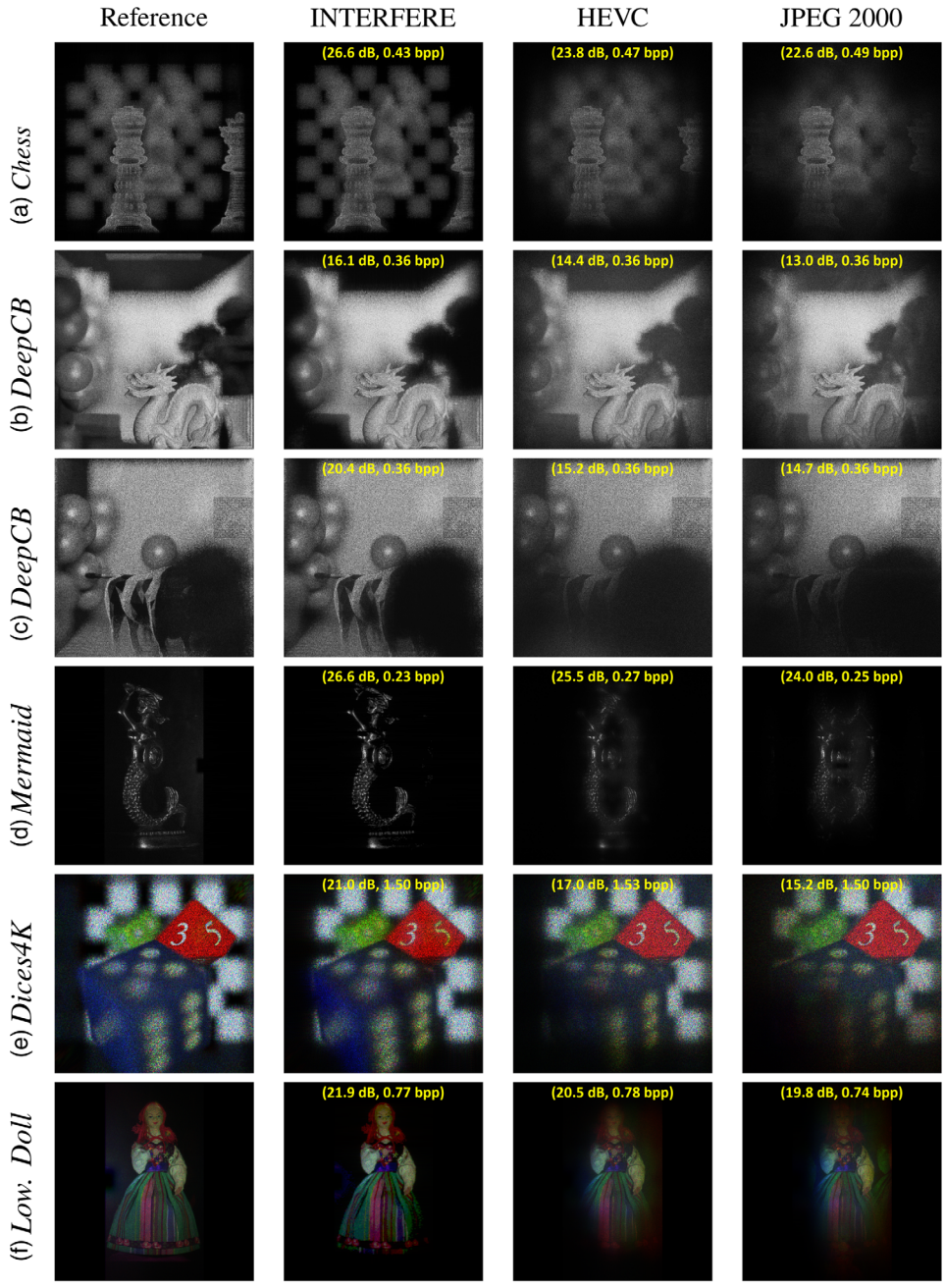

Fig. 11. Numerical reconstructions produced by NRSH V8.0 [88] from holograms compressed by different codecs (columns). PSNR and bitrate shown in brackets for viewports. (a) Center view, 496 mm. (b) Center view, 250 mm. (c) Top right view, 418 mm. (d) Center view, 450 mm. (e) Top right view, 2.5 mm. (f) Center view,1060 mm. *DeepCB—DeepCornellBox*, *Low. Doll—LowiczankaDoll*. See Fig. 10 for rate-distortion plots.

The shortcomings of conventional image coding solutions, widely used for hologram compression in practice, were identified. The proposed solution demonstrably obtains higher compression levels compared to conventional codecs when measured objectively by the rate-distortion performance and when evaluated subjectively from numerical reconstructions for a wide selection of optically acquired and computer-generated holograms. Moreover, the compression was achieved with spatio-angular random access and supports a low-complexity decoder design. The data to be compressed is also organized into convenient data structures for scalability.

The proposed coding technology has been accepted as the basis for the first international standard addressing the compression of holographic content, JPEG Pleno Holography (ISO/IEC 21794). This standard is expected to be published in late 2024. We hope this contribution is helpful in the development of practical solutions targeting high-end holographic displays and digital holographic microscopy and tomography.

Funding. FP7 Ideas: European Research Council (617779); Fonds Wetenschappelijk Onderzoek (12ZQ223N, G0B3521N); Japan Society for the Promotion of Science (P22752).

Acknowledgment. The authors thank Colas Schretter and the ISO/IEC JTC1/SC29/WG1 members for their constructive criticism and feedback.

Disclosures. The authors declare no conflicts of interest.

Data availability. Most of the holograms used for evaluating compression performance in this work are available in Ref. [85].

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. D. Gabor, "A new microscopic principle," Nature **161**, 777–778 (1948).
2. P. Marquet, B. Rappaz, P. J. Magistretti, *et al.*, "Digital holographic microscopy: a noninvasive contrast imaging technique allowing quantitative visualization of living cells with subwavelength axial accuracy," Opt. Lett. **30**, 468–470 (2005).
3. V. Balasubramani, A. Kus, H.-Y. Tu, *et al.*, "Holographic tomography: techniques and biomedical applications," Appl. Opt. **60**, B65–B80 (2021).
4. Y. Takaki and N. Okada, "Hologram generation by horizontal scanning of a high-speed spatial light modulator," Appl. Opt. **48**, 3255–3260 (2009).
5. S. Reichelt, R. Häussler, G. Fütterer, *et al.*, "Full-range, complex spatial light modulator for real-time holography," Opt. Lett. **37**, 1955–1957 (2012).
6. A. Maimone, A. Georgiou, and J. S. Kollin, "Holographic near-eye displays for virtual and augmented reality," ACM Trans. Graph. **36**, 1–16 (2017).
7. D. M. Hoffman, A. R. Girshick, K. Akeley, *et al.*, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," J. Vis. **8**(3):33, 33 (2008).
8. D. Blinder, A. Ahar, S. Bettens, *et al.*, "Signal processing challenges for digital holographic video display systems," Signal Process. Image Commun. **70**, 114–130 (2019).
9. A. W. Lohmann, R. G. Dorsch, D. Mendlovic, *et al.*, "Space–bandwidth product of optical signals and systems," J. Opt. Soc. Am. A **13**, 470–473 (1996).
10. D. Blinder, T. Birnbaum, T. Ito, *et al.*, "The state-of-the-art in computer generated holography for 3D display," Light Adv. Manuf. **3**, 572–600 (2022).
11. G. A. Mills and I. Yamaguchi, "Effects of quantization in phase-shifting digital holography," Appl. Opt. **44**, 1216–1225 (2005).
12. Y. Xing, B. Pesquet-Popescu, and F. Dufaux, "Comparative study of scalar and vector quantization on different phase-shifting digital holographic data representations," in *3DTV-Conference: The True Vision—Capture, Transmission and Display of 3D Video (3DTV-CON)* (2014), pp. 1–4.
13. P. A. Cheremkhin and E. A. Kurbatova, "Numerical comparison of scalar and vector methods of digital hologram compression," Proc. SPIE **10022**, 1002227 (2016).
14. S. Trejos, M. Gómez, A. Velez-Zea, *et al.*, "Compression of 3D dynamic holographic scenes in the Fresnel domain," Appl. Opt. **59**, D230–D238 (2020).
15. A. V. Zea, A. L. V. Amado, M. Tebaldi, *et al.*, "Alternative representation for optimized phase compression in holographic data," OSA Contin. **2**, 572–581 (2019).
16. A. V. Zea and R. Torroba, "Compression protocol for optimized random phase color holographic videos by delta encoding," J. Opt. **21**, 085706 (2019).
17. T. Birnbaum, T. Kozacki, and P. Schelkens, "Providing a visual understanding of holography through phase space representations," Appl. Sci. **10**, 4766 (2020).
18. G. J. Burton and I. R. Moorhead, "Color and spatial structure in natural scenes," Appl. Opt. **26**, 157–170 (1987).
19. D. Blinder, T. Bruylants, H. Ottevaere, *et al.*, "JPEG 2000-based compression of fringe patterns for digital holographic microscopy," Opt. Eng. **53**, 123102 (2014).
20. L.-W. Chang, C.-Y. Wang, and S.-M. Lee, "Designing JPEG quantization tables based on human visual system," in *International Conference on Image Processing (ICIP)* (1999), Vol. **2**, pp. 376–380.
21. E. Darakis and J. J. Soraghan, "Reconstruction domain compression of phase-shifting digital holograms," Appl. Opt. **46**, 351–356 (2007).
22. M. V. Bernardo, E. Fonseca, P. Fiadeiro, *et al.*, "A digital hologram compression scheme for representation on the object plane," Proc. SPIE **10752**, 107520J (2018).
23. D. Blinder, C. Schretter, H. Ottevaere, *et al.*, "Unitary transforms using time-frequency warping for digital holograms of deep scenes," IEEE Trans. Comput. Imaging **4**, 206–218 (2018).
24. Y.-H. Seo, H.-J. Choi, and D.-W. Kim, "3D scanning-based compression technique for digital hologram video," Signal Process. Image Commun. **22**, 144–156 (2007).
25. H. Yoshikawa and J. Tamai, "Holographic image compression by motion picture coding," Proc. SPIE **2652**, 2–9 (1996).
26. J.-K. Kim, K.-J. Oh, J.-W. Kim, *et al.*, "Intra prediction-based hologram phase component coding using modified phase unwrapping," Appl. Sci. **11**, 2194 (2021).
27. K.-J. Oh, H. Ban, S. Choi, *et al.*, "HEVC extension for phase hologram compression," Opt. Express **31**, 9146–9164 (2023).
28. J. L. Vago, H. C. Vermeulen, and A. Verga, "Fast Fourier transform based image compression algorithm optimized for speckle interferometer measurements," Opt. Eng. **36**, 3052–3063 (1997).
29. T. J. Naughton, Y. Frauel, B. Javidi, *et al.*, "Compression of digital holograms for three-dimensional object reconstruction and recognition," Appl. Opt. **41**, 4124–4132 (2002).
30. L. T. Bang, Z. Ali, P. D. Quang, *et al.*, "Compression of digital hologram for three-dimensional object using Wavelet-Bandelets transform," Opt. Express **19**, 8019–8031 (2011).
31. D. Blinder, T. Bruylants, E. Stijns, *et al.*, "Wavelet coding of off-axis holographic images," Proc. SPIE **8856**, 88561L (2013).
32. T. Birnbaum, A. Ahar, D. Blinder, *et al.*, "Wave atoms for digital hologram compression," Appl. Opt. **58**, 6193–6203 (2019).
33. M. Liebling, T. Blu, and M. Unser, "Fresnelets: new multiresolution wavelet bases for digital holography," IEEE Trans. Image Process. **12**, 29–43 (2003).
34. K. Viswanathan, P. Gioia, and L. Morin, "Wavelet compression of digital holograms: towards a view-dependent framework," Proc. SPIE **8856**, 88561N (2013).
35. T. Shimobaba, D. Blinder, T. Birnbaum, *et al.*, "Deep-learning computational holography: a review (invited)," Front. Photon. **3**, 854391 (2022).
36. S. Jiao, Z. Jin, C. Chang, *et al.*, "Compression of phase-only holograms with JPEG standard and deep learning," Appl. Sci. **8**, 1258 (2018).
37. T. Shimobaba, D. Blinder, P. Schelkens, *et al.*, "Deep-learning-based dynamic range compression for 3D scene 699 hologram," in *International Conference on Optics & Electro-optics (ICOL)*, K. Singh, A. K. Gupta, S. Khare, *et al.*, eds. (Springer, 2021), pp. 41–44.
38. W. Kim, J.-K. Kim, B.-S. Park, *et al.*, "Phase-only hologram video compression using a deep neural network for up-scaling and restoration," Appl. Opt. **61**, 10644–10657 (2022).
39. P. Schelkens, A. Ahar, A. Gilles, *et al.*, "Compression strategies for digital holograms in biomedical and multimedia applications," Light Adv. Manuf. **3**, 40 (2022).
40. F. Dufaux, Y. Xing, B. Pesquet, *et al.*, "Compression of digital holographic data: an overview," Proc. SPIE **9599**, 95990I (2015).
41. R. Guerra, E. Fonseca, A. Pinheiro, *et al.*, "Holography challenges on coding and quality assessment," in *10th European Workshop on Visual Information Processing (EUVIP)* (2022), pp. 1–6.
42. T. Ebrahimi, S. Foessel, F. Pereira, *et al.*, "JPEG pleno: toward an efficient representation of visual reality," IEEE Multimedia **23**, 14–20 (2016).
43. P. Schelkens, T. Ebrahimi, A. Gilles, *et al.*, "JPEG Pleno: providing representation interoperability for holographic applications and devices," ETRI J. **41**, 93–108 (2019).
44. "Final call for proposals on JPEG pleno holography, WG1N91022, 91st JPEG Meeting," ISO/IEC JTC1/SC29/WG1 (2021).
45. R. K. Muhamad, T. Birnbaum, D. Blinder, *et al.*, "INTERFERE: a unified compression framework for digital holography," in *Digital Holography and 3-D Imaging* (Optica, 2022), paper Th4A.2.
46. R. K. Muhamad, T. Birnbaum, D. Blinder, *et al.*, "Binary hologram compression using context based Bayesian tree models with

adaptive spatial segmentation," Opt. Express **30**, 25597–25611 (2022).
47. R. K. Muhamad, D. Blinder, A. Symeonidou, *et al.*, "Exact global motion compensation for holographic video compression," Appl. Opt. **58**, G204–G217 (2019).
48. J. Goodman, *Introduction to Fourier Optics*, McGraw-Hill Physical and Quantum Electronics Series (W. H. Freeman, 2005).
49. T. Birnbaum, D. Blinder, R. K. Muhamad, *et al.*, "Object-based digital hologram segmentation and motion compensation," Opt. Express **28**, 11861–11882 (2020).
50. E. Sánchez-Ortiga, A. Doblas, G. Saavedra, *et al.*, "Off-axis digital holographic microscopy: practical design parameters for operating at diffraction limit," Appl. Opt. **53**, 2058–2066 (2014).
51. P. Stepien, R. K. Muhamad, D. Blinder, *et al.*, "Spatial bandwidth-optimized compression of image plane off-axis holograms with image and video codecs," Opt. Express **28**, 27873–27892 (2020).
52. Z. Dong, Y. Ling, C. Xu, *et al.*, "Gaze-contingent efficient hologram compression for foveated near-eye holographic displays," Displays **79**, 102464 (2023).
53. U. Schnars and W. Jüptner, "Digital recording and numerical reconstruction of holograms," Meas. Sci. Technol. **13**, R85–R101 (2002).
54. T. Birnbaum, D. Blinder, R. K. K. Muhamad, *et al.*, "A standard way for computing numerical reconstructions of digital holograms," Proc. SPIE **12138**, 121380O (2022).
55. D. Blinder and P. Schelkens, "Accelerated computer generated holography using sparse bases in the STFT domain," Opt. Express **26**, 1461–1473 (2018).
56. A. Gilles and P. Gioia, "Real-time layer-based computer-generated hologram calculation for the Fourier transform optical system," Appl. Opt. **57**, 8508–8517 (2018).
57. D. Yasuki, D. Blinder, T. Shimobaba, *et al.*, "Hologram calculation in sparse Fourier bases using fixed-point operations, and its circuit architecture," Proc. SPIE **11353**, 1135305 (2020).
58. D. Blinder, C. Schretter, and P. Schelkens, "Global motion compensation for compressing holographic videos," Opt. Express **26**, 25524–25533 (2018).
59. P. Müller, M. Schürmann, and J. Guck, "The theory of diffraction tomography," arXiv, (2016).
60. X. Liu, F. Chen, R. K. Muhamad, *et al.*, "Bitwidth-optimized energy-efficient FFT design via scaling information propagation," in *58th ACM/IEEE Design Automation Conference (DAC)* (2021), pp. 613–618.
61. M. Garrido, K. Möller, and M. Kumm, "World's fastest FFT architectures: breaking the barrier of 100 GS/s," IEEE Trans. Circuits Syst. I **66**, 1507–1516 (2019).
62. D. Taubman and M. Marcellin, *JPEG2000 Image Compression Fundamentals, Standards and Practice* (Springer, 2013).
63. S. Lloyd, "Least squares quantization in PCM," IEEE Trans. Inf. Theory **28**, 129–137 (1982).
64. H. Gish and J. Pierce, "Asymptotically efficient quantizing," IEEE Trans. Inf. Theory **14**, 676–683 (1968).
65. P. Cheremkhin and E. Kurbatova, "Wavelet compression of off-axis digital holograms using real/imaginary and amplitude/phase parts," Sci. Rep. **9**, 7561 (2019).
66. A. El Rhammad, P. Gioia, A. Gilles, *et al.*, "Progressive hologram transmission using a view-dependent scalable compression scheme," Ann. Telecommun. **75**, 201–214 (2020).
67. Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers (speech coding)," IEEE Trans. Acoust. Speech Signal Process. **36**, 1445–1453 (1988).
68. J. Kiefer, "Sequential minimax search for a maximum," Proc. Am. Math. Soc. **4**, 502–506 (1953).
69. L. F. Williams, "A modification to the half-interval search (Binary Search) method," in *14th Annual Southeast Regional Conference*, Association for Computing Machinery, New York, New York, USA, 1976, Vol. **ACM-SE 14**, pp. 95–101.
70. I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," Commun. ACM **30**, 520–540 (1987).
71. "Common test conditions 9.0 for JPEG pleno holography, WG1N100341, 97th JPEG meeting," ISO/IEC JTC1/SC29/WG1 (2022).
72. A. Ahar, D. Blinder, T. Bruylants, *et al.*, "Subjective quality assessment of numerically reconstructed compressed holograms," Proc. SPIE **9599**, 95990K (2015).
73. A. M. G. Pinheiro, J. Prazeres, A. Gilles, *et al.*, "Definition of common test conditions for the new JPEG pleno holography standard," Proc. SPIE **12138**, 121380N (2022).
74. A. Ahar, M. Pereira, T. Birnbaum, *et al.*, "Validation of dynamic subjective quality assessment methodology for holographic coding solutions," in *13th International Conference on Quality of Multimedia Experience (QoMEX)* (2021), pp. 7–12.
75. J. Prazeres, A. Gilles, R. Kizhakkumkara Muhammad, *et al.*, "Quality evaluation of the JPEG Pleno Holography Call for Proposals response," in *14th International Conference on Quality of Multimedia Experience (QoMEX)* (2022), pp. 1–6.
76. R. K. Muhamad, T. Birnbaum, A. Gilles, *et al.*, "JPEG Pleno holography: scope and technology validation procedures," Appl. Opt. **60**, 641–651 (2021).
77. T. Birnbaum, D. Blinder, and P. Schelkens, "Diffraction limited perspective, numerical reconstruction of macroscopic DH," in *OSA Imaging and Applied Optics Congress 2021 (3D, COSI, DH, ISA, pcAOP)* (Optica, 2021), paper DW6C.2.
78. D. Blinder, M. Chlipala, T. Kozacki, *et al.*, "Photorealistic computer generated holography with global illumination and path tracing," Opt. Lett. **46**, 2188–2191 (2021).
79. A. Symeonidou, D. Blinder, A. Munteanu, *et al.*, "Computer-generated holograms by multiple wavefront recording plane method with occlusion culling," Opt. Express **23**, 22149–22161 (2015).
80. M. V. Bernardo, P. Fernandes, A. Arrifano, *et al.*, "Holographic representation: hologram plane vs. object plane," Signal Process. Image Commun. **68**, 193–206 (2018).
81. A. Gilles, P. Gioia, R. Cozot, *et al.*, "Computer generated hologram from multiview-plus-depth data considering specular reflections," in *IEEE International Conference on Multimedia & Expo Workshops (ICMEW)* (2016), pp. 1–6.
82. A. Gilles, P. Gioia, R. Cozot, *et al.*, "Hybrid approach for fast occlusion processing in computer-generated hologram calculation," Appl. Opt. **55**, 5459–5470 (2016).
83. A. Ahar, M. Chlipala, T. Birnbaum, *et al.*, "Suitability analysis of holographic vs light field and 2D displays for subjective quality assessment of Fourier holograms," Opt. Express **28**, 37069–37091 (2020).
84. A. Gołoś, W. Zaperty, G. Finke, *et al.*, "Fourier RGB synthetic aperture color holographic capture for wide angle holographic display," Proc. SPIE **9970**, 99701E (2016).
85. "JPEG Pleno database," https://plenodb.jpeg.org/.
86. D. Taubman, A. Naman, and R. Mathew, "High throughput JPEG 2000 (HTJ2K): algorithm, performance and potential," (2019).
87. "NHRT 1.0—numerical holography reconstruction tool 1.0, WG1N83060, 83rd JPEG meeting," ISO/IEC JTC1/SC29/WG1 (2019).
88. "Numerical reconstruction software for holography (NRSH) V8.0, WG1N100197, 95th JPEG meeting," ISO/IEC JTC1/SC29/WG1 (2022).
89. "Wavelet compression in the hologram plane and associated aliasing artifacts after reconstruction in object plane, WG1M88009, 88th JPEG meeting," ISO/IEC JTC1/SC29/WG1 (2020).