Noname manuscript No.
(will be inserted by the editor)

Progressive hologram transmission using a view-dependent scalable compression scheme

Anas El Rhammad[1,3] · **Patrick Gioia**[1,2] · **Antonin Gilles**[1] · **Marco Cagnazzo**[1,3]

Received: date / Accepted: date

Abstract Over the last few years, holography has been emerging as an alternative to stereoscopic imaging since it provides users with the most realistic and comfortable three-dimensional (3D) experience. However, high quality holograms enabling a free-viewpoint visualization contain tremendous amount of data. Therefore, a user willing to access to a remote hologram repository would face high downloading time, even with high speed networks. To reduce transmission time, a joint viewpoint-quality scalable compression scheme is proposed. At the encoder side, the hologram is first decomposed into a sparse set of diffracted light rays using Matching Pursuit over a Gabor atoms dictionary. Then, the atoms corresponding to a given user's viewpoint are selected to form a sub-hologram. Finally, the pruned atoms are sorted and encoded according to their importance for the reconstructed view. The proposed approach allows a progressive decoding of the sub-hologram from the first received atom. Streaming simulations for a moving user reveal that our approach outperforms conventional scalable codecs such as scalable H.265 and enables a practical streaming with a better quality of experience.

Keywords Digital Holography · Diffraction · Compression · Gabor Wavelets · Matching Pursuit · Streaming · Scalability

The authors can be reached at: anas.elrhammad@b-com.com & antonin.gilles@b-com.com & patrick.gioia@orange.com & marco.cagnazzo@b-com.com

[1] IRT b−com, Cesson Sévigné, France
[2] Orange Labs, Cesson-Sévigné, France
[3] LTCI, Télécom Paris, Institut polytechnique de Paris, France

1 Introduction

Contrary to conventional photography, holography has the ability to record and reconstruct both amplitude and phase information of the light wave scattered by a 3D object. In 1960, the invention of laser allows the optical recording of interference patterns on photographic plates [1]. With the rapid development of electronic devices such as Spatial Light Modulators (SLM), image sensors and high computational computers, the digital capture and display of dynamic color holograms became possible [2]. Digital holography has potential application in microscopy [3], particle measurement [4], security [5], storage [6], and 3D display systems [7,8]. The latter use case is particularly considered in this paper.

The most of conventional 3D visualization systems are based on stereoscopy to create a depth illusion from two planar views by means of human binocular vision. Despite its simple implementation, this technology necessitates special wearable glasses. Alternatively, modern 3D displays allow a depth perception to the naked eye based on autostereoscopic techniques. The main limitation of these (auto)-stereoscopic solutions is their inability to provide the focus cue at the real depth of the scene. Therefore, observer's eye accommodation is often in conflict with the vergence of eye-pair, which entails headaches and eye-strain [9]. By contrast, digital holography can account for all monocular and binocular visual cues including stereopsis, continuous parallax, occultation and accommodation. As a result, it is considered as the ultimate technology for a natural and comfortable 3D visualization.

Despite its attractive properties for 3D interactive displays, digital holography presents some hardware and software challenges that prevent for the time being, the design of consumer holographic applications. In this

study, we mainly address the signal processing issues related to compression and transmission. Indeed, a hologram pattern is an accumulation of the light wavefronts emitted by all the points of the scene. Consequently, holograms with high visual quality and large Field of View (FoV) will contain a massive volume of data (up to gigapixels) corresponding to all possible viewpoints and depths. A second bottleneck induced by holograms non-local character is the fact that their signal properties differ substantially from photographic images. Indeed, a digital hologram is a non-stationary signal with high frequency components, which makes the compression difficult when using classical image or video encoders. When considering these two constraints, the transmission of a $128K \times 128K$ monochromatic hologram using a Wi-Fi connection of 50Mbps bandwidth would require 1 hour and 45 minutes ! Thus, it is obvious that today's wired/wireless communications channels will not be able to transmit high resolution holograms in time compatible with practical applications.

To reduce the size of digital holograms and then enable a rapid display of the 3D contents, efficient compression algorithms must be applied to the hologram before its transmission to the user. Since digital holograms are encoded as 2D complex signals, a trivial solution was to apply image and video coding standards either on the real/imaginary or amplitude/phase representation format. Accordingly, the JPEG2000 standard was deployed for hologram compression in microscopy applications with an extension using directional adaptive wavelet transforms and full-packet decomposition [10]. The authors of [11] explored the use of video codecs H.264/AVC (Advanced Video Coding) [12] and H.265/HEVC (High Efficiency Video Coding) [13] to compress phase-shifting holographic sequences. In [14], an extension of HEVC was also proposed by designing optimized directional transforms from the residuals of a holographic training set encoded by HEVC. Despite the rate-distortion optimization of these coding standards, the obtained compression ratios are still very low for a good quality of reconstruction.

To improve the compression performance, the use of popular wavelets transforms has been investigated in the literature [15,16]. More specified wavelets schemes that match the statistical properties of holographic signals were designed using Bandelets transform [17] and a vector lifting scheme [18,19]. Another category of works considered the use of B-splines/Fresnelet [20,21] and Gabor/Morlet [22,23] wavelets as an efficient tool to deal with non-local nature of holograms due to their optimal space-frequency localization property. Recently, a new compression paradigm was introduced in [24], using Linear Canonical Transforms to construct a piece-wise operator that models the diffraction of non-planar surfaces. Despite the remarkable performance gain brought by these approaches compared to common compression standards, they did not provide any insights into an efficient scalable representation and coding.

Scalability has long been known as an effective tool allowing access to very large datasets [25] since it allows the user to quickly have a glimpse of the image while downloading the enhancement information. Scalable coding has mainly be considered for images and videos [26–28], but some preliminary work for scalable hologram compression exist. For example, in [29,30] the hologram is represented as a light-field by a spatial cutting into small apertures. The obtained sub-holograms are then encoded using a scalable video codec and transmitted with respect to the viewer's location. Unfortunately, a strong speckle noise is caused by these approaches due to the small aperture size. To overcome this limitation, a view-dependent compression scheme using Gabor wavelets to extract the light rays emitted in the viewer's pupil is introduced in [31]. Despite the good performance achieved by this technique in terms of compression and viewpoint scalability, the study did not take into account a progressive increase of quality for each sub-hologram.

In this paper, we propose a coding framework that combines viewpoint and quality scalability. This means that a user first receives the data allowing him to visualize the requested hologram from its current position, with quality that increases along the time. When the current viewpoint is received at its full quality level, neighboring viewpoints can be transmitted. This fine-grain level of scalability is possible thanks to the Gabor representation of the hologram, and may be the enabling technology for a practical transmission of high resolutions holograms. The remainder of the article is outlined as follows: in Section 2, we give a background on digital hologram representation, signal properties and challenges related to scalable coding. In Section 3, we present our proposed scalable compression scheme. Experimental results are given in Section 4 and conclusion in Section 5.

2 Digital holograms: representation, signal properties and scalable coding challenges

Digital holography is the process of recording interference patterns between an object wavefield and a reference beam. Digital holograms can be either optically acquired or generated: in the first case, a charge-coupled device (CCD) records the complex wavefield [32]. In the second case, the hologram is numerically generated from the synthetic 3D models or Multiview-Plus-Depth

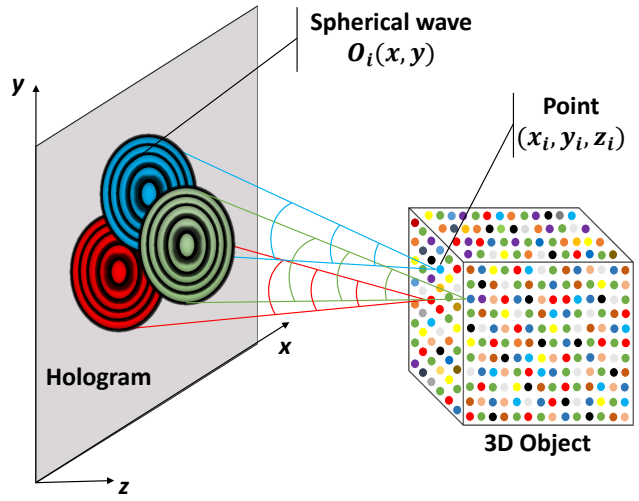


Fig. 1: Hologram formation from the scattered object wave.

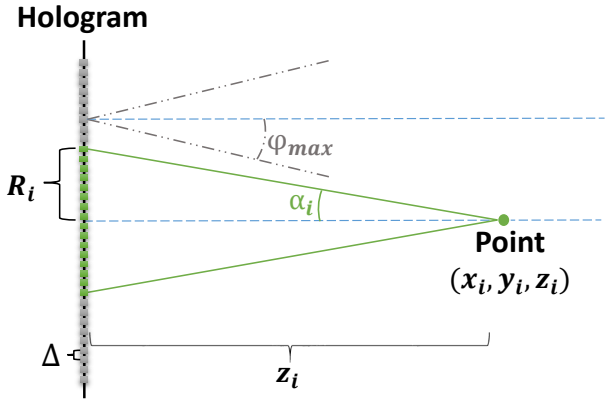


Fig. 2: Zone of light contribution by a given point.

data (MVD) of the scene [33–35]. The obtained hologram is called a *Computer Generated Hologram* (CGH). In this section we give an overview on different encoding formats, signal characteristics and scalable compression issues of digital holograms.

2.1 Digital holograms representation

A digital hologram corresponds to the encoding of the object light wave across a 2D surface. Two encoding data formats of this wavefield are commonly used: real-valued and complex-valued representations.

2.1.1 Real-valued representation

The object wave is encoded into a hologram with real positive values. The incident reference wave can be modulated using either amplitude or phase information.

An amplitude hologram $\mathbf{H}$ is recorded as the intensity of the interference between a reference wave $R : \mathbb{R}^2 \mapsto \mathbb{C}$ and an object wave $O : \mathbb{R}^2 \mapsto \mathbb{C}$, defined such that

$$H = (O + R)(O + R)^* = |O|^2 + |R|^2 + 2\Re\{OR^*\}, \quad (1)$$

where C^*, $|C|$ and $\Re\{C\}$ are respectively the conjugate, amplitude and real parts of the complex number C. The numerical computation of Equation (1) simulates the physical phenomenon of wave interference occurring in conventional optical holography [36]. Generally, the first two terms are not considered since they generate undesirable zero diffraction order artifacts. Thus, all the necessary holographic information is contained in the third term $2\Re\{OR^*\}$.

Phase holograms are obtained by modulating the reference wave phase, and then have a better diffraction efficiency than amplitude holograms. However, available optical sensors are not able to record phase information. Alternatively, iterative phase retrieval algorithms [37] has been proposed to record the phase of the object wave. Although the encoded data is reduced by half, real-valued representation cannot reproduce reliably the object wavefield since only amplitude or phase information is recorded.

2.1.2 Complex-valued representation

This is the most complete representation since it allows to control jointly the amplitude and phase of the object light wave, yielding an accurate replication of the captured scene. Complex-valued holograms are represented as complex numbers using either Real/Imaginary (R/I) or Amplitude/Phase (A/P) representation. Another possible complex representation is obtained by computing the difference between multiple amplitude hologram recordings with predefined shifted phases [32]. This technique is known as *Phase-shifting holography*. In the rest of this article, the digital holograms complex-value representation is considered.

2.2 Signal properties of digital holograms

For a better understanding of the holographic signal characteristics, one has to apprehend how the complex wavefield is formed by the light propagation.

Let us consider a 3D object represented with L discretized points as shown in Figure 1. The coordinate system (x, y, z) is defined so that the hologram lies on the $(x, y, 0)$ plane. Every point i with coordinates (x_i, y_i, z_i) is considered to emit a spherical light wave.

The object wave scattered by the scene into the hologram plane can thus be computed as the sum of the L spherical waves $\{O_i(x,y)\}_{1\leq i\leq L}$ emitted by the scene points, such that

$$O(x,y) = \sum_{i=1}^{L} O_i(x,y) = \sum_{i=1}^{L} \frac{a_i}{r_i} e^{[j(kr_i+\phi_i)]}, \quad (2)$$

where a_i is the amplitude of the point i, ϕ_i its phase defined randomly and r_i is the oblique distance, given by

$$r_i = \sqrt{(x-x_i)^2 + (y-y_i)^2 + z_i^2}. \quad (3)$$

As depicted from Equation 2, a hologram gathers intrinsically the superimposed beams of light diffracted by the object points. Thus, the 3D information may spread out over the whole hologram. Inversely, every hologram pattern could contain information about the entire captured scene.

Figure 2 presents the light beam emitted by a given point with an angle α_i. To avoid aliasing, α_i should not exceed the maximum angle of incidence on the hologram φ_{max} which is given by the grating equation [38] such that

$$\varphi_{max} = \arcsin(\lambda\, f_{max}), \quad (4)$$

where λ denotes the wavelength of light and f_{max} is the maximum spatial frequency of the hologram.

According to the Nyquist sampling principle, the maximum spatial frequency that can be represented with a sampling pitch Δ (distance between the centers of two adjacent pixels) is given by $f_{max} = (2\Delta)^{-1}$. Thus, according to Figure 2 and Equation (4), the maximal region that can be covered by the diffracted beam in the hologram plane is given by its maximum radius

$$R_{i,max} = z_i \tan(\varphi_{max}) = z_i \tan\left[\arcsin\left(\frac{\lambda}{2\Delta}\right)\right]. \quad (5)$$

As illustrated by Equation (5), the spreading out nature of the hologram depends both on the distance between the scene and the hologram plane as well as the pixel pitch. For a given wavelength and sampling pitch, the object points with deep depths will contribute to a large number of hologram pixels and then the 3D information will be more pervasive. On the other hand, for a given depth, holograms with very small pixel pitch (µm up to nm) presents a larger incidence angle which means that the object light wave will spread over all the pixels of the hologram. Consequently, the encoded holographic signal will present non-localized features with very high frequency components. This property is then the main difference compared to common images.

2.3 Scalable holographic coding challenges

The digital holograms scalable compression is a very challenging topic that has not been extensively studied yet. As stated before, two main hurdles prevent the use of scalable image or video codecs on holograms: the first one is the tremendous amount of information stored in its patterns and the second one is related to its signal properties which differ completely from semantically meaningful images.

2.3.1 Large volume of data

A digital hologram contains all the 3D information about the recorded scene at the cost of a massive data amount. For instance, let us consider a monochromatic laser ($\lambda = 640$ nm) and a target viewing angle of $\varphi_{max} = 55°$. This results in a pixel pitch of $\Delta = 0.4$ µm. Thus, generating a hologram of dimensions 10 cm×10 cm requires an ultra-high resolution of $250K \times 250K$ [39], that is more than 60 Gpixels.

Given such volume of data, the deployment of adaptive streaming solutions such as MPEG DASH [40] where the content is encoded at several quality levels, resolution and viewpoint, would be limited by the server computation and storage capabilities especially for holographic video. This calls for an effective and fine-grain scalable representation of digital holograms.

2.3.2 Specific hologram features

Since digital holograms have different signal characteristics compared to conventional images, scalable image/video coding schemes would be sub-optimal when applied to holographic signals. Indeed, each pixel contains information about the whole 3D scene and adjacent pixels are highly complementary to form the right wavefront in the reconstruction domain. Thus, directional spatial predictors and down/up sampling algorithms used for spatial scalability are ineffective on hologram patterns. Moreover, the relationship between the distortion in the hologram and reconstruction domains is not straightforward, which makes the quality scalability more complicated. Finally, traditional image and video scalable encoders would not explicitly take into account the specific features of holograms, and in particular their directionality: therefore, viewpoint scalability would not be possible with such schemes.

3 Proposed scalable compression scheme

In order to overcome the aforementioned issues, a possible solution consists in expanding the hologram into

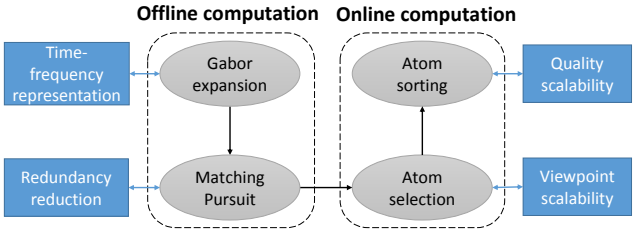


Fig. 3: Overall block of the proposed method.

a set of analyzing wavelets whose intrinsic properties are able to locally extract the directions of diffracted light rays. Accordingly, we propose a scalable compression scheme based on a Gabor wavelets decomposition which allows a scalability in terms of viewpoint and quality. Figure 3 shows the overall block diagram of our approach, which is composed of four steps. During the offline computation, the hologram is first decomposed over a family of 2D Gabor wavelets. Then, the overcomplete representation is sparsified using the Matching Pursuit algorithm. The online computation corresponds to the scalable part of our method. The viewpoint scalability consists in selecting only the wavelets which emit light rays to the observer window. Then, the pruned wavelets are sorted with respect to their importance for the reconstructed view, which enables a scalability in terms of quality.

3.1 Gabor expansion

As stated before, the propagation of light tends to spread out features that are initially well localized in the object domain. Thus, an adequate transform is required to minimize the energy spreading. Among various windowed basis functions, Gabor atoms are known to provide the optimal trade-off between space and frequency localization[41] with a compact support in both domains. The mother wavelet is the product of a complex exponential and a Gaussian function. It has been shown that Gabor wavelets are well suited for processing and reconstructing holographic signals [22,42].

3.1.1 2D discrete Gabor wavelets

A Gabor dictionary is obtained by shifting, dilating and rotating the mother wavelet. Thus, a family of discrete Gabor wavelets, in their 2D form, is given by

$$g_{l,p}[x,y] = \beta\, exp\left(-\frac{f_l^2}{\sigma^2}\left(a^2+b^2\right)\right) exp\left(2\pi i a f_l\right), \qquad (6)$$

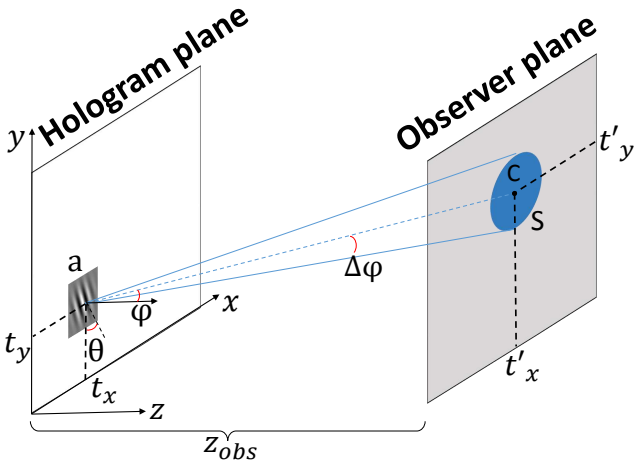


Fig. 4: Diffraction spectrum S emitted by a Gabor atom a.

where

$$\begin{cases} a = (x\cos(\theta_p) + y\sin(\theta_p))\Delta \\ b = (y\cos(\theta_p) - x\sin(\theta_p))\Delta, \end{cases} \qquad (7)$$

β is a constant chosen such that $g_{l,p}$ has a unit L_2 norm, and σ controls the standard deviation of the mother wavelet.

In Eq (6), $(x,y) \in \mathbb{Z}^2$ is the 2D shift discretized parameter such that:

$$\begin{cases} -\frac{M_x}{2} \leq x < \frac{M_x}{2} \\ -\frac{M_y}{2} \leq y < \frac{M_y}{2}, \end{cases} \qquad (8)$$

where (M_x, M_y) is the wavelet resolution.

The dilated frequency f_l and the rotation angle θ_p are discretized with L and P levels, respectively, such that

$$\begin{cases} f_l = \frac{f_{max}}{s_l} & 0 \leq l < L \\ \theta_p = \frac{2p\pi}{P} & 0 \leq p < P, \end{cases} \qquad (9)$$

where the integers l and p are the indexes of discretization corresponding to the dilation and rotation parameters, respectively.

3.1.2 Duality between Gabor wavelets and light rays

Figure 4 presents the example of a light ray emitted by a Gabor atom a from $\mathbf{t} = (t_x, t_y)$ in the hologram plan (HP). The direction of diffraction is given by $(\theta,\ \phi)$, where ϕ is related to the frequency of the Gabor wavelet according to the grating equation (4). In the observer plan (OP), the emitted light ray forms a diffraction

spectrum S with center C and angular dispersion $\Delta\phi$. The latter depends on the frequency dilation parameter.

Accordingly, a discrete Gabor expansion of a hologram $\mathbf{H}$ is generated by computing the inner-products $\{< \mathbf{H}, g_{l,p} >\}_{\substack{1\leq l\leq L \\ 1\leq p\leq P}}$ between $\mathbf{H}$ and the Gabor dictionary described in Equation (6). Thus, an overcomplete set of $N = N_x N_y LP$ atoms is obtained, where (N_x, N_y) denotes the spatial resolution of $\mathbf{H}$. Each atom corresponds to a diffraction spectrum emitted from a given position in the *HP*.

3.2 Matching pursuit

For compression purposes, when considering the entire hologram, a sparse approximation algorithm called Matching Pursuit (MP) is applied to reduce the size of the redundant Gabor expansion. MP consists on selecting greedily, from an overcomplete decomposition, the atoms that are the most correlated with the the current residual. At each iteration, the residual is updated by subtracting from the signal the contribution of the extracted atom. This process is repeated until the norm of the residual converge to a small threshold [43]. For a redundant Gabor representation of color digital holograms, the MP algorithm can be applied in its vector (VMP) or scalar (SMP) version as described in [23].

Thus, the MP K-expansion is a sparse representation of K atoms ($K \ll N$) such as

$$A_K = \{a_k = (\mathbf{t}_k, l_k, p_k, c_k), \qquad 1 \leq k \leq K\}. \tag{10}$$

Inversely, MP allows a hologram recovery from the K extracted atoms through the approximating linear combination

$$\hat{\mathbf{H}} = \sum_{k=0}^{K} c_k g_{l_k, p_k}. \tag{11}$$

To summarize the offline computation part, the digital hologram is first decomposed using a family of Gabor wavelets, then a sparse MP expansion is generated and stored in the server.

3.3 Viewpoint scalability

The hologram representation given in Equation (10) provides a convenient way to generate sub-holograms based on the observer's locations. This functionality is very important since it is useless to encode and send the full hologram when the user is only watching at a specific direction that could be represented by a much smaller sub-hologram [44].

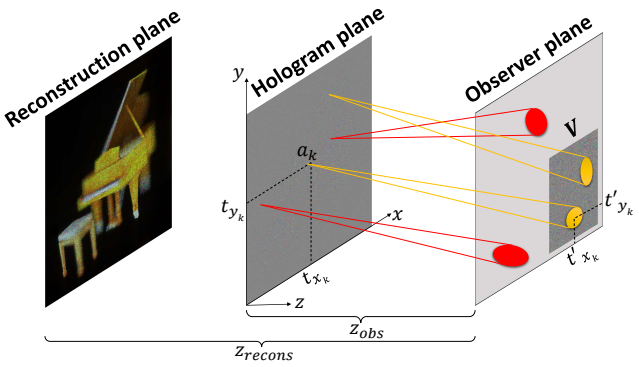


Fig. 5: Adaptive selection process. *The sub-hologram $\boldsymbol{H_V}$ is generated from the Gabor atoms having diffraction spectrums' centers inscribed in V*

Therefore, a viewpoint scalability may be achieved by encoding the hologram into different sub-holograms following the observer's trajectory [31]. This would drastically reduce the amount of data to transmit for a given perspective and thus the time needed for a first visualization of the scene.

A sub-hologram is generated from the set A_k by adaptively selecting the subset of Gabor atoms emitting light towards the viewer's retina. Thus, one has to exploit the duality between Gabor wavelets and light rays. As depicted in Figure 4, the coordinates of the diffraction spectrum's center $C(t'_x, t'_y)$ emitted by a Gabor atom a from a position (t_x, t_y) is given by the relation

$$\begin{cases} t'_x = t_x + z_{obs} \tan(\phi)\cos(\theta) \\ t'_y = t_y + z_{obs} \tan(\phi)\sin(\theta), \end{cases} \tag{12}$$

where z_{obs} is the distance between *HP* and *OP*.

Accordingly, Figure 5 illustrates the process of selecting the relevant wavelets to generate a sub-hologram with respect to the viewer's position. It consists on selecting all the Gabor atoms having diffraction spectrums' centers included in the viewing window $\mathbf{V}$ (yellow ones) and discard all the remaining atoms (red ones). Thus, the subset of M atoms defining the sub-hologram $\mathbf{H_V}$ corresponding to the field of view $\mathbf{V}$ is given by

$$A_V = \{a_k \in A_k \quad | \quad C_k \in \mathbf{V}, \quad 1 \leq k \leq K\}. \tag{13}$$

Finally, the numerical reconstruction $\mathbf{R_V}$ of the sub-hologram $\mathbf{H_V}$ is obtained by propagating the cropped wavefield in the OP to the reconstruction plane (RP). $\mathbf{R_V}$ is defined by

$$\mathbf{R_V} = \mathcal{P}_{-z_{recons}}(\mathcal{P}_{z_{obs}}(\mathbf{H_V})), \tag{14}$$

where $\mathcal{P}_z(\cdot)$ denotes the propagation operator for a distance z.

3.4 Quality scalability

In a streaming experience where the user may have several positions, i.e. viewpoints, it is highly preferable to encode each sub-hologram at different levels of quality in a scalable way. This would enable a progressive transmission of the holographic data based on the available network's bandwidth. Indeed, the purpose of quality scalability is to encode the sub-hologram of a given viewpoint into a single bitstream having several layers: a base layer to enable a rapid visualization of the sub-hologram, and refinement layers to progressively enhance the visual quality.

To achieve such goal, we propose to sort the atoms of the set A_V according to their importance for the reconstructed wavefield. Since the quality of a given sub-hologram depends on the amount of light emitted by the Gabor atoms into the viewing window, the sorting order depends on three criteria: (1) increasing distance of the diffraction spectrum's center C_m from the observation's center O, (2) decreasing light intensity given by the atom's amplitude $|c_m|$, (3) decreasing emission cone's size determined by the angular dispersion $\Delta\phi_m$.

Consequently, we define three functions $f_j : m \in [1, M] \mapsto [1, M]$ that give the sorted position of the m-th atom according to criterion (j). Then, the sorted set A_V^j allowing the quality scalability for the sub-hologram $\mathbf{H_V}$ is given by

$$A_V^j = \{a_{f_j(m)}, \qquad 1 \leq m \leq M\}. \tag{15}$$

Moreover, an adaptive order can be defined by combining the three criteria. For example, A_V^j may be divided into different subsets following a decreasing threshold on amplitude and/or angular dispersion and then each subset is sorted according to criterion (1).

For a static user, the set A_V^j is splitted into different layers which means that the hologram may be displayed from the first received atoms, i.e. from the base layer. Then, the visual quality is progressively enhanced until the highest quality is achieved.

When the user moves from a position O to O', the set A_V is updated to a set $A_{V'}$ by removing all the atoms that do not contribute for the new sub-hologram $\mathbf{H_{V'}}$, and adding those that emit light rays into the viewing window $\mathbf{V'}$ (viewpoint scalability). If $\mathbf{V}$ and $\mathbf{V'}$ overlap, only the atoms whose diffraction spectrum's centers are inscribed into the window $\mathbf{V'}\backslash\mathbf{V}$ are transmitted, following the sorting order given by Equation (15). To anticipate any viewpoint switching lag caused by a rapid observer movement, the atoms contributing for the peripheral views can be transmitted once a good reconstruction quality is attained for the current view.

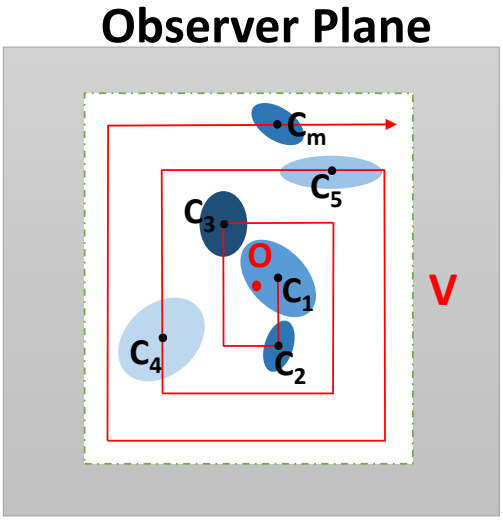


Fig. 6: Spiral order in the OP. *The diffraction spectrums' centers are encoded differentially in a spiral scan from the observation's center O*

3.5 Encoding consideration

For transmission purposes, a bitstream should be generated from the sorted set of atoms A_V^j. Thanks to the straightforward relation of Equation (12), the diffraction spectrums' centers in the observer window are encoded in the observer window instead of the atoms positions in the *HP*. Thus, the cost of encoding atoms positions is reduced from $log_2(N_x)+log_2(N_y)$ to $log_2(V_x)+log_2(V_y)$, where (V_x, V_y) is the spatial resolution of $\mathbf{V}$. Moreover, when considering the distance criterion, the centers $\{C_m\}_{1\leq m\leq M}$ are differentially encoded using a spiral scan order as shown in Figure 6. Then, the complex coefficients $\{c_m\}_{1\leq m\leq M}$, the dilation and rotations indexes are entropy encoded using context-adaptive arithmetic coding as detailed in [23]. Thus, the generated scalable bitstream for a sub-hologram $\mathbf{H_V}$ is given by

$$B_V = \{(\delta_m, l_m, p_m, [c_m]), \qquad 1 \leq m \leq M\}, \tag{16}$$

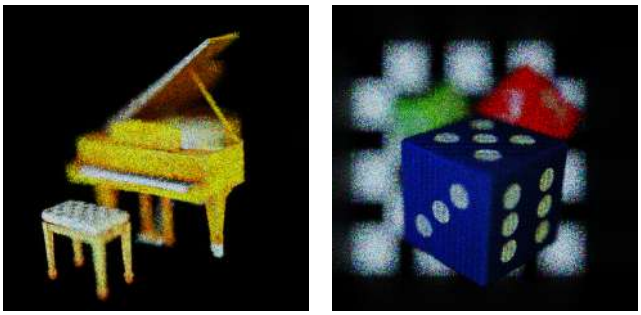
where δ_m is the difference between the coordinates of two consecutive diffraction spectrums' centers. $[c_m]$ is the quantified value of the m^{th} atom's coefficient.

4 Experimental results

In all our experiments we use as quality metric the PSNR (Peak Signal-to-Noise Ratio) between the images reconstructed from the original and compressed holograms. We consider the total rate needed to encode the real and imaginary part of the holographic signal

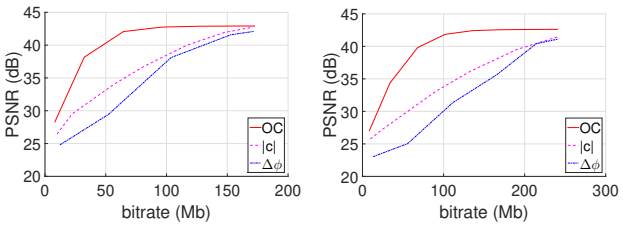
Parameter	Hologram	
	Piano8k : $\mathbf{H_1}$	**Dices8k : $\mathbf{H_2}$**
Resolution	8192×8192	8192×8192
Pixel pitch (µm)	0.4	0.4
$\mathbf{z_{obs}}$ (cm)	0.208	0.218
$\mathbf{z_{recons}}$ (cm)	0.505	0.603
Wavelength (nm)	(640, 532, 470)	(640, 532, 470)

Table 1: Hologram parameters.



(a) $\mathbf{H_1}$ (b) $\mathbf{H_2}$

Fig. 7: Numerical reconstructions of the original holograms for a central view.



(a) $\mathbf{H_1}$ (b) $\mathbf{H_2}$

Fig. 8: R-D graphs comparison of the three sorting criteria for a central view.

and we use the Bjontegaard Delta Rate [45] to compare different compression algorithms.

For the simulation tests, a K-expansion A_K is generated from color holograms using VMP, where $K = \frac{N_x N_y}{4}$. The obtained set of atoms is stored offline in the server's RAM memory. In the rest of the article, our viewpoint-quality scalable compression scheme is denoted by SVMP (Scalable VMP).

To evaluate the efficiency of the SVMP, we compare its compression performance to an existing scalable codec such as the Scalable extension of HEVC (SHVC) [46] which is an easy, off-the-shelf solution, and a natural benchmark for our proposed coding scheme.

4.1 Hologram database description

We notice that our proposed approach is relevant for both: optically acquired holograms and CGHs. Since they are easier to acquire, CGHs are considered in our experiments. Table 1 summarizes the parameters of two color CGHs. These holograms are generated from synthetic scenes using the methods described in [34,47]. The chrominance format of CGH is set to 4:4:4 with 8-bit depth for each color's channel (R, G and B). Since the R/I format allows a better bit-rate allocation than the A/P format [48], it is considered for all the encoding process. Figure 7 shows the numerical reconstructions of the non-compressed holograms for a central viewpoint, i.e. an observer located at O_{center} = (0.16 cm, 0.16 cm) according to the coordinate system given in Figure 1.

4.2 Comparison of the three sorting criteria for quality scalability

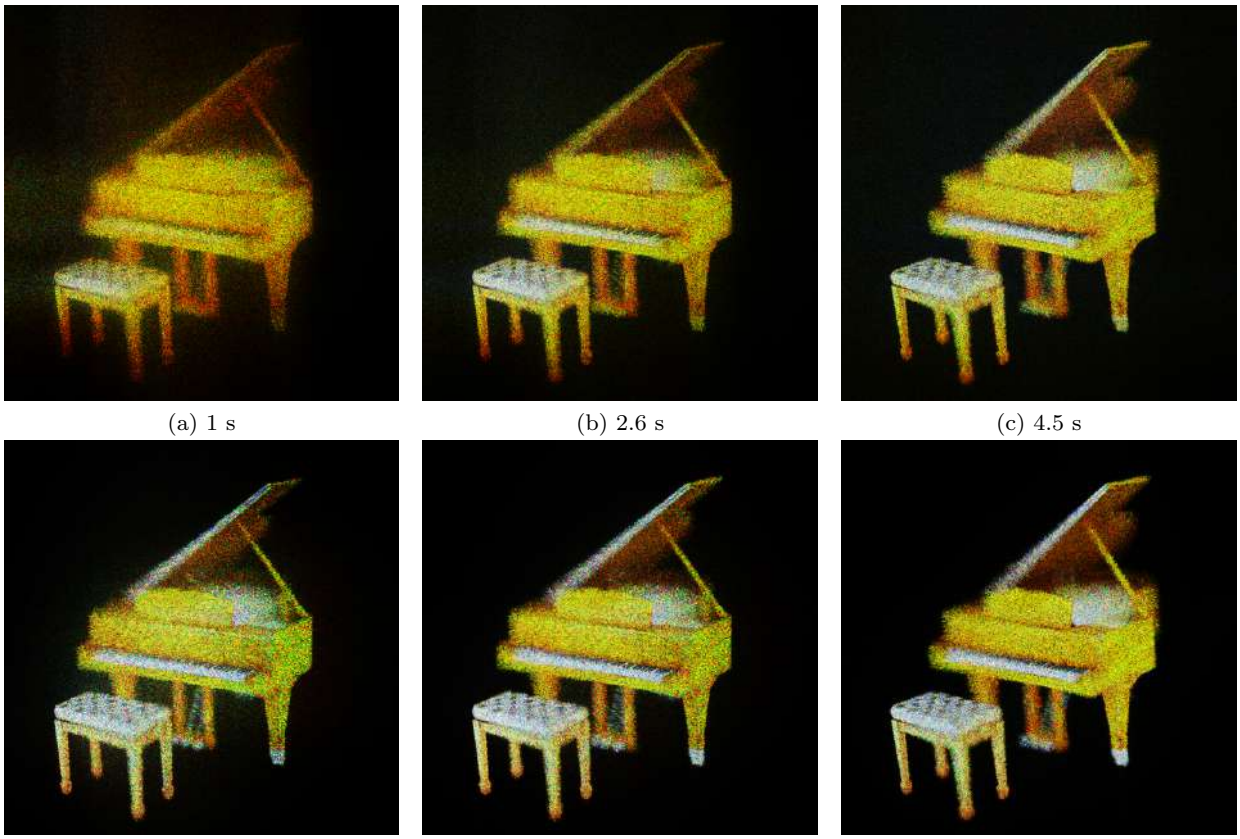
In order to find the most effective sorting criterion, we compare the compression results obtained by the three criteria defined in Section 3.4. In this part of experiments, a static user at position O_{center} is considered.

The R-D graphs corresponding to the numerical reconstructions of the decoded sub-hologram of both holograms $\mathbf{H_1}$ and $\mathbf{H_2}$ are presented in Figure 8. According to the obtained curves, the sorting function f_1 based on the distance OC (criterion 1) allows the best scalable coding performance. Indeed, it achieves compression gains of 49.5% and 65.7% (BD-RATE) over the amplitude and angular dispersion scalability, respectively, for the reconstructed view of $\mathbf{H_1}$. Moreover, the amplitude scalability outperforms the angular dispersion one, especially at low and medium bitrates.

Thus, we can conclude that the atoms which emit the nearest light rays to the viewer's center are the most important for the reconstruction and should be transmitted first. In the rest of the simulation tests, the quality scalability based on the proximity of the light rays to the observer's center, i.e. distance OC, will be considered.

4.3 Streaming simulation

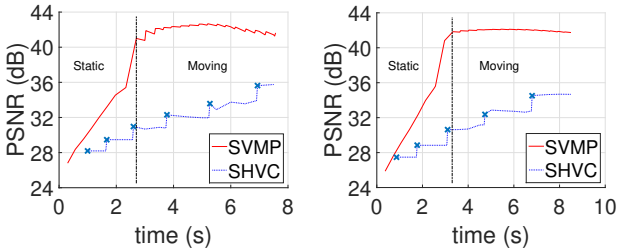
In the first part of streaming experiments, we consider a moving user and we simulate the streaming of the sub-holograms corresponding to its trajectory. We assume that the observer has an horizontal trajectory with two



(a) 1 s (b) 2.6 s (c) 4.5 s

(d) 1 s (e) 2.6 s (f) 4.5 s

Fig. 9: Numerical reconstructions of the decoded sub-holograms of Piano8k taken at different instants (see Online Resource 1). The first row corresponds to SHVC and the second one to SVMP (our approach).



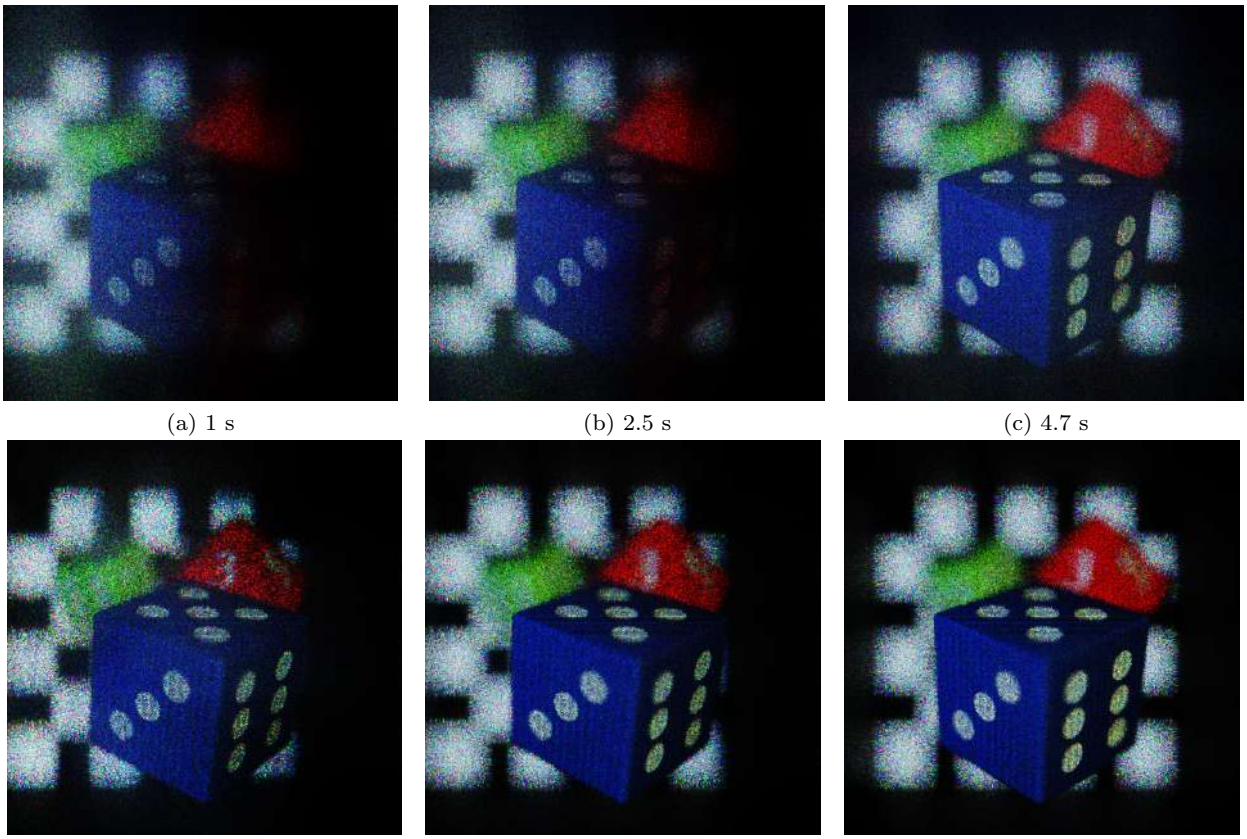
(a) $\mathbf{H_1}$ (b) $\mathbf{H_2}$

Fig. 10: T-D curves for a simulated streaming of a moving user using SHVC and SVMP (our approach) - The bandwidth is 30Mbps.

phases: the first phase in which the viewer is immobile at a position O_{left} = (0 cm, 0.16 cm), and the second phase where the observer moves uniformly from O_{left} to a position O_{right} = (0.33 cm, 0.16 cm). We consider a transmission bandwidth of 30Mbps, which corresponds to a relatively high-speed wireless connection [49]. We denote by Δ_1^{static} = 2.7s and Δ_1^{moving} = 4.8s the duration of the static and moving phases for the hologram $\mathbf{H_1}$. Δ_2^{static} = 3.3s and Δ_2^{moving} = 5.2s are the durations of the two phases for the hologram $\mathbf{H_2}$.

For the benchmark solution (SHVC), the reference software SHM12.4 is used to encode the real and imaginary parts of the original holograms. The base layer is encoded with a QP of 50, whereas for the enhancement layers the QP values vary from 48 to 40 with a step of 2. This choice is a convenient trade-off between the bitrate and distortion: for QP = 40, we achieve an acceptable image quality of 36dB with a bitrate of 200Mb. A smaller QP would lead to higher bitrate: 400Mb for a QP of 34. Since the software only support *YUV 4:2:0* format for multi-layers, each hologram channel (R, G and B) is encoded separately after converting it to the *YUV 4:2:0* format (using null chrominance channels). It is judicious to note that the whole information about the holographic signal is encoded offline by SHVC. The decoding process is performed online and is independent of the user's movements.

The time-distortion (T-D) graphs of the streaming simulation for both holograms using SHVC and SVMP are presented in Figure 10. At each time instant we



(a) 1 s (b) 2.5 s (c) 4.7 s

(d) 1 s (e) 2.5 s (f) 4.7 s

Fig. 11: Numerical reconstructions of the decoded sub-holograms of Dices8k taken at different instants (see Online Resource 2). The first row corresponds to SHVC and the second one to SVMP (our approach).

consider the viewpoint that can be reconstructed with each of the two methods, and we compute the PSNR with respect to the same viewpoint but reconstructed from uncompressed data. The blue cross markers corresponds to the moments where a new layer is decoded for SHVC. In that moments the PSNR increases sharply, then it slowly fluctuates since the reconstructed viewpoint changes as the user moves. In any case, the user has to wait for the full reception of the base layer in order to start the visualization.

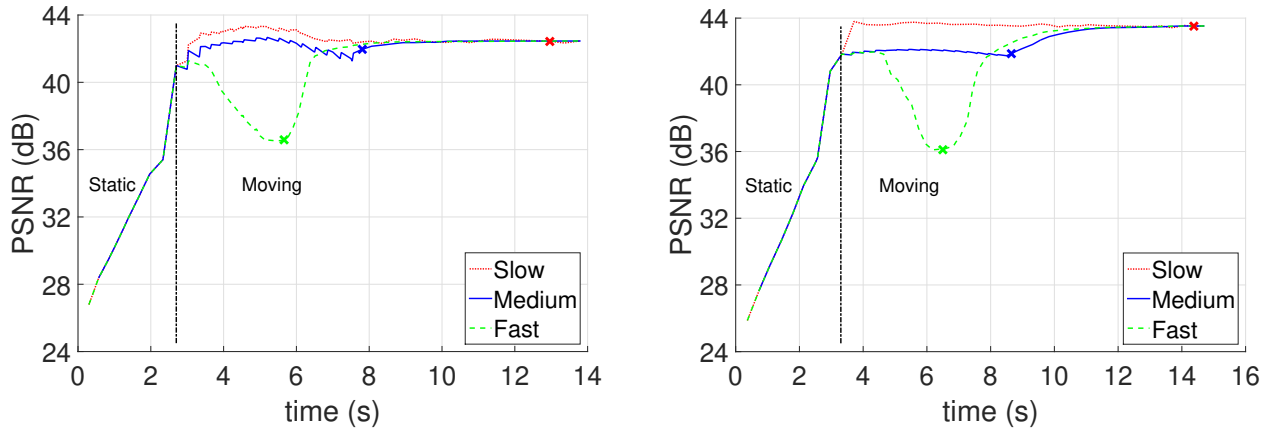
As illustrated by the curves, our approach allows a better performance in terms of scalability. Indeed, SVMP achieves a PSNR of 30.6dB for the hologram $\mathbf{H_1}$ after 1s, whereas SHVC attains only 28.2dB after the same duration. Also, it takes approximately 4s for SVMP to converge to a maximal reconstruction quality for the hologram $\mathbf{H_2}$, whereas SHVC would need smaller QPs to achieve the same quality. We note that the small fluctuations in the PSNR values during the moving phase are due to the view changes.

These objective results are visually confirmed by the numerical reconstructions given in Figure 9, Figure 11 and Online Resources 1 and 2. As shown in the reconstructions of Piano8k and Dices8k holograms using SHVC, the shape and contrast of the scenes are highly distorted after one second of transmission, which corresponds to the base layer. Then, the enhancement layers slowly improve the visual quality but still present some visible noise that affects the color information. On the other hand, these distortions are significantly reduced by SVMP, which enables a good visual quality after only one second of transmission and converge to a visually lossless quality after 4 seconds.

	V_s (mm/s)	V_m (mm/s)	V_f (mm/s)
$\mathbf{H_1}$	0.321	0.642	1.110
$\mathbf{H_2}$	0.296	0.593	1.026

Table 2: Speeds corresponding to slow (V_s), medium (V_m) and fast (V_f) viewpoint changes.

In the second test of streaming simulations, the influence of user's velocity on its quality of experience is studied. To this end, we consider three viewers' speeds:



(a) $\mathbf{H_1}$ (b) $\mathbf{H_2}$

Fig. 12: Streaming simulation for three viewing speeds: V_s (slow), V_m (medium) and V_f (fast) - The bandwidth is 30Mbps. (*The crossmarks correspond to moments when the user reachs its last position. After this moments, the user remains static.*)

V_s, V_m and V_f, corresponding to slow, medium and fast displacements, respectively. Table 2 summarizes the 3 speed values choosen for each hologram. These values are computed by considering a user traversing the whole horizontal trajectory (3.276mm) in a given duration. We assume that the streaming will finish when all the atoms of the K-expansion are downloaded. Moreover, once the user reachs its final position, he remains static until the end of streaming. The obtained T-D graphs for both holograms are presented in Figure 12.

The T-D curves reveal two important results:

(i) During the moving phase, the reconstruction quality is degraded as the viewer's speed increases.

(ii) At the end of the streaming, the reconstruction qualities converge to the same value, regradless the considered viewer's speed.

In the following, we provide some explanations of these phenomena.

In the moving phase, an average PSNR gain of 0.9dB and 1.5dB is achieved for slow displacements compared to medium ones, for $\mathbf{H_1}$ and $\mathbf{H_2}$, respectively. This result was expected, since for a slow viewpoint's change the user has more time to download the atoms of the next view. For fast displacements, the quality of reconstructed views decreases sharply after a short movement duration (0.8s for $\mathbf{H_1}$ and 1.3s for $\mathbf{H_2}$).

Once the last viewpoint is attained, the remaining atoms are sent progressively. This justify the fact that the quality increases a second time in case of medium and fast displacements, until reaching the same quality achieved by slow displacements.

5 Conclusion

A scalable compression scheme of digital holograms has been introduced in this paper. The proposed framework is based on a sparse Gabor wavelets decomposition of the holographic signal using Matching Pursuit. First, the viewpoint scalability is achieved by selecting the Gabor atoms that emit light rays to the viewer's field of view. Then, the quality scalability is obtained by sorting these atoms according to their importance for the reconstructed view. Finally, the generated bitstream allows a progressive decoding based on the observer's position and the quality of the reconstructed scene. This is possible thanks to the fine-grain scalability provided by the proposed encoding scheme.

The experimental results reveal that our approach outperforms the conventional scalable coding scheme based on HEVC and enables a rapid progressive decoding with a better quality of experience. Thus, the proposed method is a first step for an efficient streaming of digital holograms with high resolutions.

As a future work, we plan to extend our approach to holographic videos using an appropriate motion compensation model to exploit the temporal correlations in the 3D scene.

Acknowledgements This work has been achieved within the Institute of Research and Technology b-com, dedicated to digital technologies. It has been funded by the French government through the National Research Agency (ANR) Investment referenced ANR-A0-AIRT-07.

References

1. Emmett N Leith and Juris Upatnieks, "Reconstructed wavefronts and communication theory," *JOSA*, vol. 52, no. 10, pp. 1123–1130, 1962.
2. Joseph W Goodman and RW Lawrence, "Digital image formation from electronically detected holograms," *Applied physics letters*, vol. 11, no. 3, pp. 77–79, 1967.
3. Myung K Kim, "Principles and techniques of digital holographic microscopy," *SPIE reviews*, vol. 1, no. 1, pp. 018005, 2010.
4. Shigeru Murata and Norifumi Yasuda, "Potential of digital holography in particle measurement," *Optics & Laser Technology*, vol. 32, no. 7-8, pp. 567–574, 2000.
5. Bahram Javidi and Takanori Nomura, "Securing information by use of digital holography," *Optics letters*, vol. 25, no. 1, pp. 28–30, 2000.
6. Lambertus Hesselink, Sergei S Orlov, and Matthew C Bashaw, "Holographic data storage systems," *Proceedings of the IEEE*, vol. 92, no. 8, pp. 1231–1280, 2004.
7. Fahri Yaraş, Hoonjong Kang, and Levent Onural, "State of the art in holographic displays: a survey," *Journal of display technology*, vol. 6, no. 10, pp. 443–454, 2010.
8. David Blinder, Ayyoub Ahar, Stijn Bettens, Tobias Birnbaum, Athanasia Symeonidou, Heidi Ottevaere, Colas Schretter, and Peter Schelkens, "Signal processing challenges for digital holographic video display systems," *Signal Processing: Image Communication*, vol. 70, pp. 114–130, 2019.
9. David M Hoffman, Ahna R Girshick, Kurt Akeley, and Martin S Banks, "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of vision*, vol. 8, no. 3, pp. 33–33, 2008.
10. David Blinder, Tim Bruylants, Heidi Ottevaere, Adrian Munteanu, and Peter Schelkens, "Jpeg 2000-based compression of fringe patterns for digital holographic microscopy," *Optical Engineering*, vol. 53, no. 12, pp. 123102, 2014.
11. Yafei Xing, Béatrice Pesquet-Popescu, and Frederic Dufaux, "Compression of computer generated phase-shifting hologram sequence using avc and hevc," in *Applications of Digital Image Processing XXXVI*. International Society for Optics and Photonics, 2013, vol. 8856, p. 88561M.
12. Thomas Wiegand, Gary J Sullivan, Gisle Bjontegaard, and Ajay Luthra, "Overview of the h. 264/avc video coding standard," *IEEE Transactions on circuits and systems for video technology*, vol. 13, no. 7, pp. 560–576, 2003.
13. Gary J Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand, "Overview of the high efficiency video coding (hevc) standard," *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 12, pp. 1649–1668, 2012.
14. José Pedro Peixeiro, Catarina Brites, Joao Ascenso, and Fernando Pereira, "Holographic data coding: Benchmarking and extending hevc with adapted transforms," *IEEE Transactions on Multimedia*, vol. 20, no. 2, pp. 282–297, 2018.
15. Alison E Shortt, Thomas J Naughton, and Bahram Javidi, "Compression of digital holograms of three-dimensional objects using wavelets," *Optics Express*, vol. 14, no. 7, pp. 2625–2630, 2006.
16. Pavel A Cheremkhin and Ekaterina A Kurbatova, "Compression of digital holograms using 1-level wavelet transforms, thresholding and quantization of wavelet coefficients," in *Digital Holography and Three-Dimensional Imaging*. Optical Society of America, 2017, pp. W2A–38.
17. Le Thanh Bang, Zulfiqar Ali, Pham Duc Quang, Jae-Hyeung Park, and Nam Kim, "Compression of digital hologram for three-dimensional object using wavelet-bandelets transform," *Optics express*, vol. 19, no. 9, pp. 8019–8031, 2011.
18. Yafei Xing, Mounir Kaaniche, Béatrice Pesquet-Popescu, and Frédéric Dufaux, "Vector lifting scheme for phase-shifting holographic data compression," *Optical Engineering*, vol. 53, no. 11, pp. 112312, 2014.
19. Yafei Xing, Mounir Kaaniche, Béatrice Pesquet-Popescu, and Frédéric Dufaux, "Adaptive nonseparable vector lifting scheme for digital holographic data compression," *Applied optics*, vol. 54, no. 1, pp. A98–A109, 2015.
20. Michael Liebling, Thierry Blu, and Michael Unser, "Fresnelets: new multiresolution wavelet bases for digital holography," *IEEE Transactions on image processing*, vol. 12, no. 1, pp. 29–43, 2003.
21. Emmanouil Darakis and John J Soraghan, "Use of fresnelets for phase-shifting digital hologram compression," *IEEE transactions on image processing*, vol. 15, no. 12, pp. 3804–3811, 2006.
22. Kartik Viswanathan, Patrick Gioia, and Luce Morin, "Wavelet compression of digital holograms: towards a view-dependent framework," in *Applications of Digital Image Processing XXXVI*. International Society for Optics and Photonics, 2013, vol. 8856, p. 88561N.
23. Anas El Rhammad, Patrick Gioia, Antonin Gilles, Marco Cagnazzo, and Béatrice Pesquet-Popescu, "Color digital hologram compression based on matching pursuit," *Applied Optics*, vol. 57, no. 17, pp. 4930–4942, 2018.
24. David Blinder, Colas Schretter, Heidi Ottevaere, Adrian Munteanu, and Peter Schelkens, "Unitary transforms using time-frequency warping for digital holograms of deep scenes," *IEEE Transactions on Computational Imaging*, vol. 4, no. 2, pp. 206–218, 2018.
25. David Taubman and Michael Marcellin, *JPEG2000 image compression fundamentals, standards and practice: image compression fundamentals, standards and practice*, vol. 642, Springer Science & Business Media, 2012.
26. Steven McCanne, Van Jacobson, and Martin Vetterli, "Receiver-driven layered multicast," *ACM SIGCOMM Computer Communication Review*, vol. 26, no. 4, pp. 117–130, 1996.
27. J-R Ohm, "Advances in scalable video coding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 42–56, 2005.
28. Thomas André, Marco Cagnazzo, Marc Antonini, and Michel Barlaud, "Jpeg2000-compatible scalable scheme for wavelet-based video coding," *Journal on Image and Video Processing*, vol. 2007, no. 1, pp. 9–9, 2007.
29. Dae-Hyun Lee, Jae-Young Sim, Chang-Su Kim, and Sang-Uk Lee, "Viewing angle dependent coding of digital holograms," in *Signal Processing Conference, 2011 19th European*. IEEE, 2011, pp. 1367–1371.
30. Young-Ho Seo, Yoon-Hyuk Lee, Ji-Sang Yoo, and Dong-Wook Kim, "Scalable hologram video coding for adaptive transmitting service," *Applied optics*, vol. 52, no. 1, pp. A254–A268, 2013.
31. Anas El Rhammad, Patrick Gioia, Antonin Gilles, Marco Cagnazzo, and Béatrice Pesquet-Popescu, "View-dependent compression of digital hologram based on matching pursuit," in *Optics, Photonics, and Digital Technologies for Imaging Applications V*. International Society for Optics and Photonics, 2018, vol. 10679, p. 106790L.
32. Ichirou Yamaguchi and Tong Zhang, "Phase-shifting digital holography," *Optics letters*, vol. 22, no. 16, pp. 1268–1270, 1997.

33. Chris Slinger, Colin Cameron, and Maurice Stanley, "Computer-generated holography as a generic display technology," *Computer*, vol. 38, no. 8, pp. 46–53, 2005.
34. Antonin Gilles, Patrick Gioia, Rémi Cozot, and Luce Morin, "Computer generated hologram from multiview-plus-depth data considering specular reflections," *2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pp. 1–6, 2016.
35. Athanasia Symeonidou, David Blinder, Adrian Munteanu, and Peter Schelkens, "Computer-generated holograms by multiple wavefront recording plane method with occlusion culling," *Optics express*, vol. 23, no. 17, pp. 22149–22161, 2015.
36. Parameswaran Hariharan and P Hariharan, *Optical Holography: Principles, techniques and applications*, Cambridge University Press, 1996.
37. R.W. Gerchberg and A Saxton W. O., "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–250, 11 1971.
38. Joseph W Goodman, *Introduction to Fourier optics*, Roberts and Company Publishers, 2005.
39. Hiroshi Yoshikawa and Takeshi Yamaguchi, "Computer-generated holograms for 3d display," *Chinese optics letters*, vol. 7, no. 12, pp. 1079–1082, 2009.
40. Stefan Lederer, Christopher Müller, and Christian Timmerer, "Dynamic adaptive streaming over http dataset," in *Proceedings of the 3rd Multimedia Systems Conference*. ACM, 2012, pp. 89–94.
41. Tai Sing Lee, "Image representation using 2d gabor wavelets," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 18, no. 10, pp. 959–971, 1996.
42. Jingang Zhong, Jiawen Weng, and Cuiying Hu, "Reconstruction of digital hologram by use of the wavelet transform," in *Digital Holography and Three-Dimensional Imaging*. Optical Society of America, 2009, p. DWB16.
43. Stéphane G Mallat and Zhifeng Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on signal processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
44. A Schwerdtner, R Häussler, and N Leister, "Large holographic displays for real-time applications," in *Practical Holography XXII: Materials and Applications*. International Society for Optics and Photonics, 2008, vol. 6912, p. 69120T.
45. G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," in *VCEG Meeting*, Austin, USA, Apr. 2001.
46. Jill M Boyce, Yan Ye, Jianle Chen, and Adarsh K Ramasubramonian, "Overview of shvc: Scalable extensions of the high efficiency video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 20–34, 2016.
47. Antonin Gilles, Patrick Gioia, Rémi Cozot, and Luce Morin, "Hybrid approach for fast occlusion processing in computer-generated hologram calculation," *Applied optics*, vol. 55, no. 20, pp. 5459–5470, 2016.
48. Peter Schelkens, Touradj Ebrahimi, Antonin Gilles, Patrick Gioia, Kwan-Jung Oh, Fernando Pereira, Cristian Perra, and Antonio MG Pinheiro, "Jpeg pleno: Providing representation interoperability for holographic applications and devices," *ETRI Journal*, vol. 41, no. 1, pp. 93–108, 2019.
49. Vijay Garg, *Wireless communications & networking*, Elsevier, 2010.