ÓPTICA PURA Y APLICADA
www.sedoptica.es

Type: Review
Section: Optical Communications

Optical and digital methods for holographic data compression

Métodos de compresión óptica y digital para datos holográficos

Alejandro Velez-Zea[1,*], John Fredy Barrera-Ramírez[1], Roberto Torroba[2]

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

2. Centro de Investigaciones Ópticas (CONICET La Plata-CIC-UNLP) C.P 1897, La Plata, Argentina.

(*) E-mail: alejandro.velezz@udea.edu.co S: miembro de SEDOPTICA / SEDOPTICA member

Received: 01/11/2021 Accepted: 22/02/2022
DOI: 10.7149/OPA.55.1.51075

ABSTRACT:

The development of novel digital holographic techniques made the research on efficient manage of holographic data more relevant than ever. Specifically, holographic data compression is the subject of intense research, with several proposals based on digital, optical, and optodigital techniques. For this reason, we provide an overview of several data compression techniques and their application to holographic data. We include a discussion about the potential compression ratios achievable with each technique and their inherent limitations. Special attention is given to the problem of compressing the phase obtained from holograms of diffuse objects.

Keywords: Digital holography, compression, scaling, quantization, phase unwrapping.

REFERENCES AND LINKS / REFERENCIAS Y ENLACES

[1] B. Javidi, A. Carnicer, A. Anand, G. Barbastathis, W. Chen, P. Ferraro, J. W. Goodman, R. Horisaki, K. Khare, M. Kujawinska, R. A. Leitgeb, P. Marquet, T. Nomura, A. Ozcan, Y. Park, G. Pedrini, P. Picart, J. Rosen, G. Saavedra, N. T. Shaked, A. Stern, E. Tajahuerce, L. Tian, G. Wetzstein, and M. Yamaguchi, "Roadmap on digital holography," Opt. Express **29**, 35078–116 (2021)

[2] T. O'Connor, A. Anand, B. Andemariam, and B. Javidi, "Deep learning-based cell identification and disease diagnosis using spatio-temporal cellular dynamics in compact digital holographic microscopy," Biomed. Opt. Express **11**, 4491–508 (2020)

[3] G. Pedrini, I. Alekseenko, G. Jagannathan, M. Kempenaars, G. Vayakis, and W. Osten, "Feasibility study of digital holography for erosion measurements under extreme environmental conditions inside the International Thermonuclear Experimental Reactor tokamak [invited]," Appl. Opt. **58**, A147-55 (2019)

[4] J. Park, K. R. Lee, and Y. K. Park, "Ultrathin wide-angle large-area digital 3D holographic display using a non-periodic photon sieve," Nat. Commun. **10**, 1–8 (2019)

[5] D. Blinder, A. Ahar, S. Bettens, T. Birnbaum, A. Symeonidou, H. Ottevaere, C. Schretter, and P. Schelkens, "Signal processing challenges for digital holographic video display systems," Signal Process. Image Commun. **70**, 114–30 (2019)

[6] D. Salomon, *Data Compression: The Complete Reference*, Third edit (Springer-Verlag, 2004)

[7] D. A. Huffman, "A Method for the Construction of Minimum-Redundance Codes," Proc. I.R.E. **40**, 1098–101 (1952)

[8] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," IEEE Trans. Inf. Theory **23**, 337–43 (1977)

[9] T. J. Naughton, Y. Frauel, B. Javidi, and E. Tajahuerce, "Compression of digital holograms for three-dimensional object reconstruction and recognition.," Appl. Opt. **41**, 4124–32 (2002)

[10] E. Darakis and J. J. Soraghan, "Use of fresnelets for phase-shifting digital hologram compression," IEEE Trans. Image Process. **15**, 3804–11 (2006)

[11] B. E. Usevitch, "A tutorial on modern lossy wavelet image compression: Foundations of JPEG 2000," IEEE Signal Process. Mag. **18**, 22–35 (2001)

[12] P. A. Cheremkhin and E. A. Kurbatova, "Quality of reconstruction of compressed off-axis digital holograms by frequency filtering and wavelets," Appl. Opt. **57**, A55-64 (2018)

[13] S. Trejos, J. F. Barrera-Ramírez, A. Velez-Zea, M. Tebaldi, and R. Torroba, "Optical approach for the efficient data volume handling in experimentally encrypted data," J. Opt. **18**, 065702 (2016)

[14] A. Velez-Zea, J. F. Barrera-Ramírez, S. Trejos, M. Tebaldi, and R. Torroba, "Optical field data compression by opto-digital means," J. Opt. **18**, 125701 (2016)

[15] E. M. Gómez-Valencia, S. Trejos, A. Velez-Zea, J. F. Barrera-Ramírez, and R. Torroba, "Experimental holographic movie compression using optical scaling and sampling," J. Opt. **22**, 035703 (2020)

[16] S. Trejos, M. Gómez-Valencia, A. Velez-Zea, J. F. Barrera-Ramírez, and R. Torroba, "Compression of 3D dynamic holographic scenes in the Fresnel domain," Appl. Opt. **59**, D230-8 (2020)

[17] U. Schnars and W. Juepner, *Digital Holography,* 1st ed. (Springer, 2005)

[18] E. Cuche, P. Marquet, and C. Depeursinge, "Spatial filtering for zero-order and twin-image elimination in digital off-axis holography," Appl. Opt. **39**, 4070–5 (2000)

[19] D. Kermisch, "Image Reconstruction from Phase Information Only," J. Opt. Soc. Am. **60**, 15–7 (1970)

[20] D. Blinder, C. Schretter, and P. Schelkens, "Global motion compensation for compressing holographic videos," Opt. Express **26**, 25524–33 (2018)

[21] J. Ziv and A. Lempel, "A Universal Algorithm for Sequential Data Compression," IEEE Trans. Inf. Theory **23**, 337–43 (1977)

[22] A. H. H. Robinson and C. Cherry, "Results of a prototype television bandwidth compression scheme," Proc. IEEE **55**, 356–64 (1967)

[23] J. W. Goodman and A. M. Silvestri, "Some Effects of Fourier-domain Phase Quantization," IBM J. Res. Dev. **14**, 478–84 (1970)

[24] W. J. Dallas, "Phase Quantization—a Compact Derivation," Appl. Opt. **10**, 673–4 (1971)

[25] W. J. Dallas, "Phase Quantization in Holograms—a Few Illustrations," Appl. Opt. **10**, 674–6 (1971)

[26] W. J. Dallas and a W. Lohmann, "Phase quantization in holograms-depth effects.," Appl. Opt. **11**, 192–4 (1972)

[27] G. K. Wallace, "The JPEG Still Picture Compession Standard," Commun. ACM **34**, 31–44 (1991)

[28] R. Shahnaz, J. F. Walkup, and T. F. Krile, "Image compression in signal-dependent noise," Appl. Opt. **38**, 5560–7 (1999)

[29] D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals , Standards and Practice* (Springer Science+Bussiness Media, 2002)

[30] P. Schelkens, Z. Alpaslan, I. Tabus, T. Ebrahimi, K.-J. Oh, A. M. G. Pinheiro, Z. Chen, and F. M. Pereira, "JPEG Pleno: a standard framework for representing and signaling plenoptic modalities," in *Applications of Digital Image Processing XLI*, A. G. Tescher, ed. (SPIE, 2018).

[31] J. a Davis and D. M. Cottrell, "Random mask encoding of multiplexed phase-only and binary phase-only filters," Opt. Lett. **19**, 496–8 (1994)

[32] A. Velez-Zea, J. F. Barrera-Ramírez, and R. Torroba, "Cross-talk free selective reconstruction of individual objects from multiplexed optical field data," Opt. Lasers Eng. **100**, 90–7 (2018)

[33] A. Velez-Zea, J. F. Barrera-Ramírez, and R. Torroba, "One-step reconstruction of assembled 3D holographic scenes," Opt. Laser Technol. **75**, 146–50 (2015)

[34] J. D. Armitage and a. W. Lohmann, "Theta Modulation in Optics," Appl. Opt. **4**, 399–405 (1965)

[35] S. Trejos, J. F. Barrera-Ramírez, M. Tebaldi, and R. Torroba, "Experimental opto-digital processing of multiple data via modulation, packaging and encryption," J. Opt. **16**, 055402 (2014)

[36] Y. Duan, F. Zhang, M. Pu, Y. Guo, T. Xie, X. Ma, X. Li, and X. Luo, "Polarization-dependent spatial channel multiplexing dynamic hologram in the visible band," Opt. Express **29**, 18351–61 (2021)

[37] G. Situ and J. Zhang, "Multiple-image encryption by wavelength multiplexing," Opt. Lett. **30**, 1306–8 (2005)

[38] M. A. Schofield and Y. Zhu, "Fast phase unwrapping algorithm for interferometric applications," Opt. Lett. **28**, 1194–6 (2003)

[39] H. Kadono, H. Takei, and S. Toyooka, "A noise-immune method of phase unwrapping in speckle interferometry," Opt. Lasers Eng. **26**, 151–64 (1997)

[40] J. M. Huntley and H. O. Saldner, "Shape measurement by temporal phase unwrapping: Comparison of unwrapping algorithms," Meas. Sci. Technol. **8**, 986–92 (1997)

[41] B. Osmanoglu, T. H. Dixon, S. Wdowinski, and E. Cabral-Cano, "On the importance of path for phase unwrapping in synthetic aperture radar interferometry," Appl. Opt. **50**, 3205-20 (2011)

[42] S. Chavez, Qing-San Xiang, and L. An, "Understanding phase maps in MRI: a new cutline phase unwrapping method," IEEE Trans. Med. Imaging **21**, 966–77 (2002)

[43] A. Velez-Zea, A. L. V. Amado, M. Tebaldi, and R. Torroba, "Alternative representation for optimized phase compression in holographic data," OSA Contin. **2**, 572–81 (2019)

[44] E. M. Gómez-Valencia, S. Trejos, A. Velez-Zea, J. F. Barrera-Ramírez, and R. Torroba, "Quality guided alternative holographic data representation for high performance lossy compression," J. Opt. **23**, 075702 (2021)

1. Introduction

Digital holography (DH) has become a potent technique, enabling a broad range of interesting applications [1]. Digital holography consists of the full phase and amplitude recording of an optical field, and its subsequent digital storage, manipulation, and reconstruction. In this way, digital holography has become an essential bridge between optical and digital data processing techniques, driving recent advances in microscopy [2], metrology [3], and novel display technologies [4], to name a few.

Parallel to the development of the diverse applications of DH, there has been a growing interest in methods to efficiently manage holographic data [5]. This is because the need for high-resolution digital holograms leads to large amounts of information that must be digitally stored and processed. In this sense, the application and development of holographic data compression techniques have become fundamental for managing these large data volumes.

There are two general types of compression methods. These are lossy and lossless compression techniques. Lossy compression consists of the use of methods that eliminate less relevant data from the input, resulting in an overall reduction in volume at the cost of information loss. Lossy compression is only limited by the amount of loss that can be tolerated, and as such can enable extreme amounts of volume reduction [6]. For this reason, some of the most powerful compression algorithms and formats used today, like mp3, mp4, and JPEG rely on lossy compression.

On the other hand, lossless compression consists in using a more compact representation of the input, without discarding any data. This often relies upon identifying periodicities or data redundancies that can be encoded into a dictionary. The potential compression that can be achieved with lossless compression is strongly related to the entropy of the input data, and as such, these techniques often result in significantly less volume reduction than lossy approaches. Some examples of lossless compression are Huffman coding [7] and the LZ77 algorithm [8], used in the .zip compression file format.

A first approach to the compression of holographic data was to use common digital algorithms. However, it was rapidly found that these were ill-suited for this purpose [9]. This is because holographic data usually presents high entropy, limiting the usefulness of lossless compression.

This also means that discarding certain frequencies or the use of quantization, common in lossy compression, can lead to more severe degradation compared to their application in other kinds of data. For this reason, the development of digital compression methods specifically tailored to holographic data is an ongoing effort, with algorithms based on wavelet transform showing the most promise [10–12].

A more recent alternative to digital compression methods is the research in methods inspired by optical processes [13–16]. These have the potential to be far more effective when applied to holographic data since they operate under the same principles and limitations of an optical system. These methods can be fully optical, or implemented with virtual optical setups, and can be successfully combined with digital algorithms, further enhancing their performance and effectiveness.

Given the above-mentioned points, we believe that as a starting point for researchers interested in the field of holographic data compression, it is important to review some of the basic methods. In this paper, we will show some holographic data compression techniques, including basic digital methods, some optical approaches, and finally, recent advances in optodigital compression applied to digital holograms of diffuse 3D objects.

2. Digital hologram filtering

A digital hologram is an intensity recording of the interference between an object beam and a reference beam, registered using a digital recording medium, like a CMOS or CCD camera. These cameras contain a certain number of light-sensitive elements, or pixels, each with a finite size. The number of pixels determines the resolution of the recorded hologram. After reconstruction, the hologram resolution is related to the maximum spatial frequency that can be reproduced, and the pixel size with the maximum size of the recorded object or scene. From a digital standpoint, the recording process consists in reading each pixel and converting its signal to a binary value. This conversion has a precision that will depend on the camera and determines how many different intensity values it can detect. This precision is measured as the number of bytes necessary to read the value of a pixel and is called bit-depth. A conventional CMOS camera has a bit depth of 8 bits, which means that it can detect up to 256 different intensity levels. More advanced scientific cameras can have bit depths of 10, 12, and even 16 bits.

Thus, the digital hologram will have a data volume in bits given by

$$V_H = N \times M \times B \tag{1}$$

where N is the horizontal number of pixels in the digital camera, M is the vertical number of pixels, and B is the bit depth. This will be our initial data volume, which we seek to compress using the techniques shown in this paper.

To introduce compression approaches, first we must understand what data is present in the hologram. Digital holograms can be recorded with two types of schemes, depending on how the reference and object beam propagates through the holographic setup. These can be on-axis and off-axis holograms. In an on-axis hologram, the reference and object beam propagate to the camera along the same axis, while in an off-axis setup, there is a certain angle between both beams.

A discussion of the merits of these two approaches is outside the scope of this work. However, it can be found in ref [17]. In general, on-axis setups allow for the recording of larger and more complex objects but require phase shift methods to eliminate the DC term from the reconstructed hologram. Off-axis setups do not have this requirement; therefore, we will center our attention on these kinds of holograms.

Furthermore, holograms can be classified as Fourier and Fresnel holograms depending on how the light propagates from the object to the camera plane. If this propagation is through free space, we have a Fresnel hologram. If we place a lens between the object and the camera, so that the lens performs the FT of the object in the camera plane, we have a Fourier hologram.

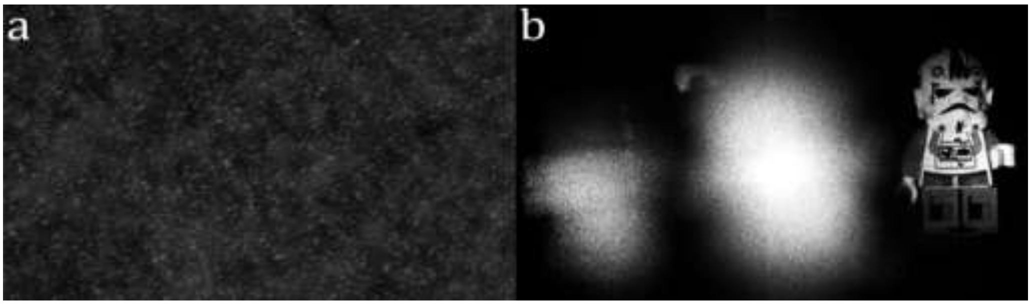

Figure 1: a) Digital Fresnel hologram, and b) reconstruction from a).

In figure 1, we show an example of an off-axis Fresnel hologram and its reconstruction. As can be seen in figure 1b, in the reconstruction plane we obtain three distinct terms. First, we have a central white noise cloud. This is the DC term and is essentially the self-convolution of the object and reference wave. Then, we have the object, and another noise cloud, which is the unfocused complex conjugate of the object, known as twin image. The spatial separation between the object and the DC term is given by the angle between the object and the reference beam used when registering the hologram.

From this figure, we can see that the object does not occupy the entire reconstruction plane, whose size is given by the hologram resolution. Thus, a basic holographic data compression method is filtering this reconstruction, discarding the area of the DC term and the unfocused complex conjugate of the object [18]. This can be done by performing the Fourier transform (FT) of the hologram and applying a bandpass filter to the result centered on the object information. We then take the filtered area, and after an inverse Fourier transform (IFT), we obtain the optical field data (OFD). The OFD is a complex-valued function that represents the sampled object beam at the camera plane and contains all information necessary to reconstruct the object.

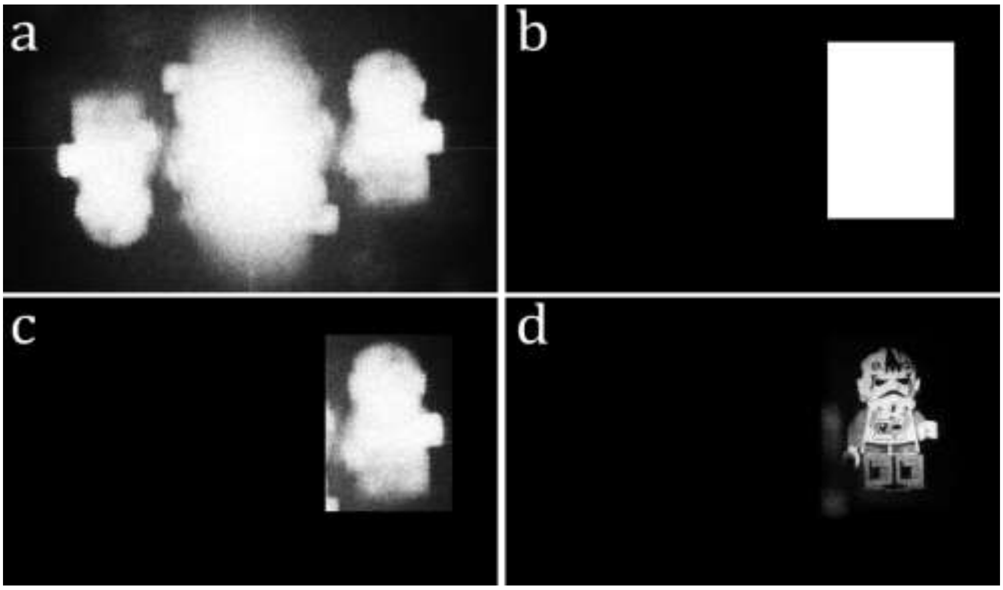

Figure 2: a) FT of the Fresnel hologram of figure 1, b) bandpass filter, c) bandpass filter applied to a), and d) filtered and reconstructed object.

In figure 2 we show an example of this filtering process. Filtering is an important starting point for the holographic compression procedure, as it allows the removal of redundant information from the hologram. The data volume of the OFD is given by

$$V_{OFD} = 2 \times N_F \times M_F \times B \tag{2}$$

here N_F and M_F are the number of vertical and horizontal pixels in the filtered area, and B is the bit-depth. Although the factor 2 caused by the need to store both phase and amplitude information (since the OFD is complex-valued) increases the data volume; it is worth noting that to adequately register an off-axis hologram, the object must occupy a number of pixels in the reconstruction plane equal to or less than 1/4 of the total hologram resolution along the spatial separation axis with the DC term. This restriction can be deduced by taking into account the size of the object, and the separation between the orders necessary to ensure that there is no crosstalk [17]. For this reason, filtering usually results in a net reduction of holographic data volume.

Another basic compression approach that can be very effective when dealing with diffuse objects is discarding the amplitude data of the OFD. In this way, we can reduce its data volume by half. This is a lossy procedure, which causes degradation of the reconstructed object. However, when dealing with diffuse objects this loss is minimal. This is because in this case the amplitude is a speckle pattern with low dynamic range, and most of the information of the object is therefore encoded into the phase of the optical field [19].

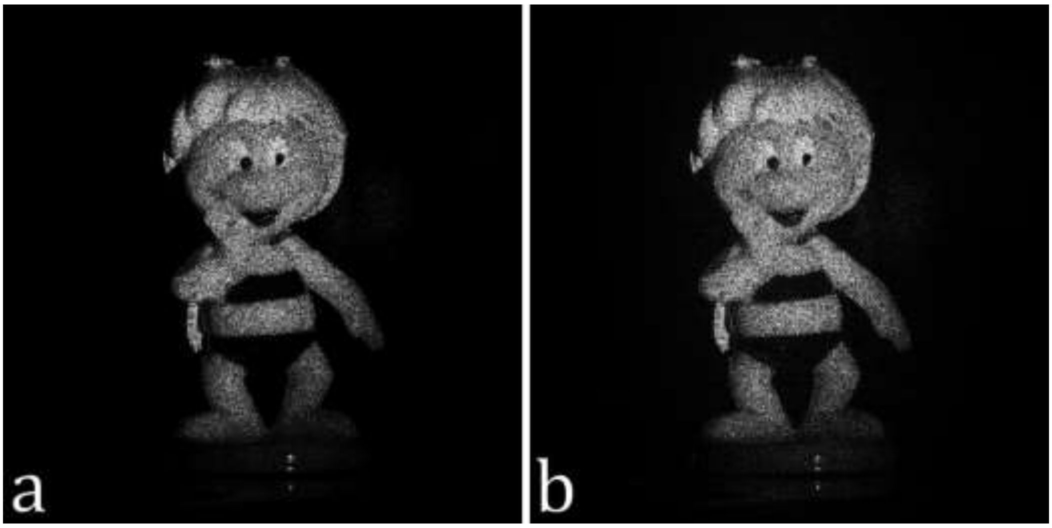

Figure 3: a) Reconstructed object from the full OFD, and b) the same object reconstructed only with the OFD's phase.

In figure 3a, we show an object reconstructed from the full OFD extracted from a Fourier hologram, and in figure 3b the same object was reconstructed only from the phase information of that OFD. As mentioned above, the difference in visual quality is minimal.

As an example of these procedures, the hologram of figure 1 was registered with a CMOS camera with a resolution of 3840×2748 pixels and an 8-bit depth. This leads to an initial hologram with a data volume of 10.06 megabytes (MB). After filtering, the OFD had a resolution of 960×1500 pixels, for a data volume of 2.74 MB. Finally, by discarding the amplitude, we obtain a final data volume of 1.37 MB. A common way to evaluate the level of compression achieved with a procedure is the compression ratio (CR), which given an initial data volume V_i and a data volume after compression V_f, is defined as

$$C_R = V_i/V_f \tag{3}$$

Taking this into account, the combination of filtering and discarding the amplitude of the OFD leads to a CR of 7.34. As a baseline comparison, modern lossy compression methods can lead to CR of between 10 and 20, reaching up to 50 in the case of state-of-the-art video compression algorithms with motion compensation [20]. On the other hand, lossless compression usually achieves CRs between 2 and 3, depending on the input data entropy. For this reason, filtering and discarding the amplitude information is not enough to satisfy the need for high compression ratios but offers a reasonable starting point with modest compression and low data loss.

3. Digital methods.

3.1 Lossless compression.

As mentioned in the introduction, lossless compression methods consist in analyzing the input data to find redundancies, which are then encoded with alternative symbols that can be stored more efficiently. As a simple example, if we have a text where the same word is repeated several times, we can replace that word with a codeword with fewer characters. Thus, the more times the word is repeated, the more we can reduce the data volume by using the codeword. This is a lossless procedure because we can easily replace again the codeword with the original data without loss. This is the working principle of Huffman coding [7], which in turn is the basis for more sophisticated approaches like the Lempel-Ziv (LZ77) algorithm [21]. In the other extreme, we have very simple methods, like PACKBITS, which only compact repeated contiguous values in a data file. In this case, if we have the same character repeated one after another, we can replace the entire repeated set with only one character and the number of times it is repeated. For example, AAAAA would become A5. For this reason, these kinds of algorithms are called run-length encoding [22].

All of these lossless methods rely on finding data redundancy. A way to measure the redundancy of a data set is entropy. Large entropy values mean that the data is highly random, and there is low redundancy, where low entropy means that there can be a large amount of redundant or repeated data. Holograms, due to the effect of speckle, interference, and the wrapped nature of phase information, usually present extremely high entropy. This means that algorithms like LZ77 when applied to holographic data result in compression ratios close to 1. This was demonstrated in detail by Naughton *et al.* [9]. For this reason, most research about holographic data compression has centered around lossy methods, which we discuss next.

3.2. Lossy compression

From the previous section, we found that for an off-axis hologram, in the worst case (with the maximum size object) after filtering and amplitude discarding, we have an OFD's phase with a data volume of

$$V_P = \frac{1}{4} N \times M \times B \tag{4}$$

We can see that the most direct way to further diminish the data volume is to reduce the resolution or the bit-depth of the OFD's phase. Let's first explore the effect of reducing the bit depth.

Reducing the bit depth is equivalent to performing a quantization procedure, reducing the number of discrete values that the phase can take. Due to the limitations of the devices used to print the first computer-generated holograms, the quantization of holographic data has been extensively explored. In particular, Goodman & Silvestri [23] first demonstrated the effects of quantization of the FT of an object, and later works demonstrated that the digital storage of phase data with less than 4 bits lead to a significant loss in quality and the appearance of twin images and other artifacts [24–26].

In figure 4, we show the effects of reducing the bit depth of from the OFD's phase on the reconstructed object. To quantitively measure the changes in quality, we calculated and included the correlation coefficient (CC) between the reconstructed object from the 8-bit depth OFD's phase and from the OFD's phase with lower bit depth. The CC is calculated as

$$CC = \frac{\sum_{p,q}^{N,M} (I[p,q] - \bar{I})(R[p,q] - \bar{R})}{\sqrt{\left(\sum_{p,q}^{N,M} (I[p,q] - \bar{I})^2\right)\left(\sum_{p,q}^{N,M} (R[p,q] - \bar{R})^2\right)}} \tag{5}$$

where I and R are the two images to compare, p and q are pixel coordinates, $\bar{I}$ and $\bar{R}$ are the mean values of I and R respectively. The CC ranges from 0, which means that there is no correlation between images, to 1, which represents an exact match.

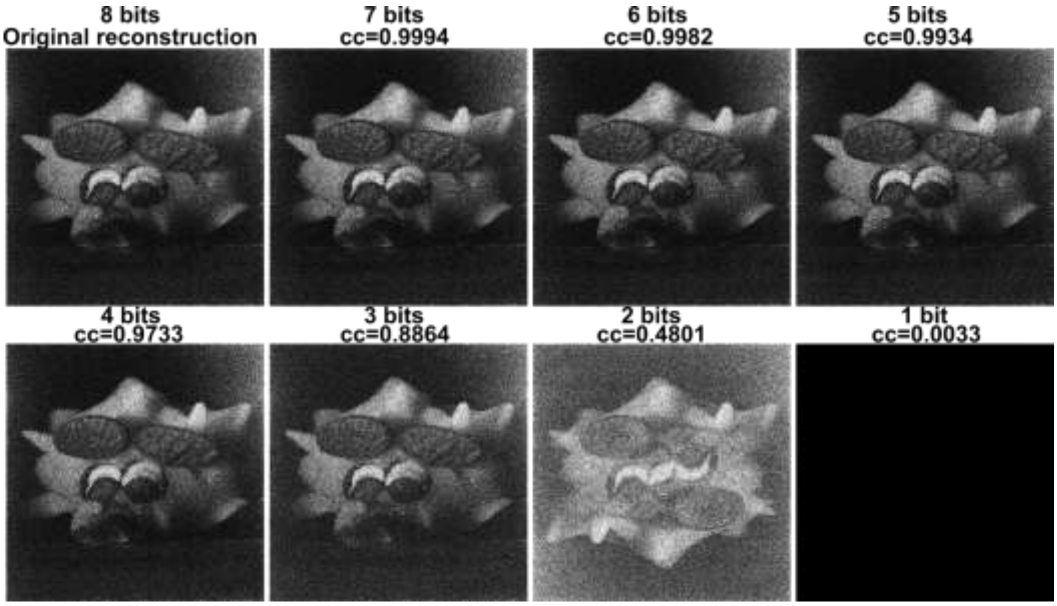

Figure 4: Reconstruction of an object from the OFD's phase quantized with different bit depths.

The result of figure 4 shows how below a bit depth of 4, the CC and visual quality fall sharply. This means that in practice, the maximum CR that can be achieved via direct quantization of the OFD's phase while maintaining reasonable quality is 2. For this reason, this is not a commonly used compression approach.

We now turn our attention to the other variable we can modify to reduce the volume, the resolution of the OFDs' phase. A direct way to reduce this resolution is to perform a digital scaling of the OFD's phase, which is equivalent to resampling the holographic data with a lower resolution and larger pixel size. This can be done for example, by generating a new compressed OFD's phase where the phase value of each pixel is the mean value of 4 adjacent pixels in the original input. Most types of digital algorithms for scaling are optimized for continuous-tone images, not for phase information of diffuse objects, and as result lead to enormous losses after reconstruction from the scaled OFD's phase, even when the scale factor is only slightly lower than 1 [9].

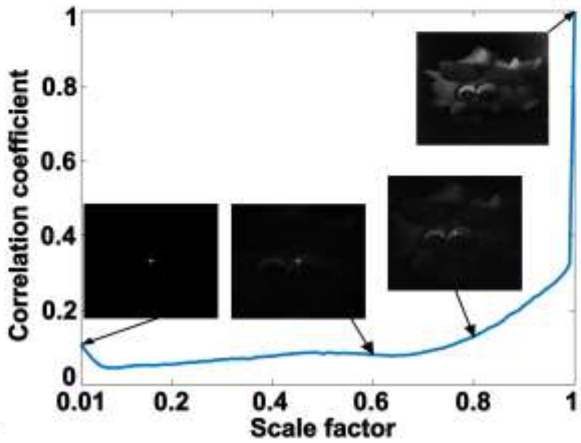

Figure 5: Correlation coefficient between the object reconstructed from the OFD's phase and the same object reconstructed from digitally scaled OFD's phase with different scale factors.

In figure 5, we show how the correlation coefficient between the object reconstructed from an uncompressed and a digitally scaled OFD's phase when different scale factors are used. The scale factor is the ratio between the initial resolution and the resolution after scaling. This result shows that digital scaling is not suitable for holographic compression applications.

More sophisticated digital compression methods, like the joint photographic expert group (JPEG) algorithm [27], make use of spectral quantization [28]. This method consists in performing the FT of the input and reducing the bit-depth in the regions with less spectral content. Spectral quantization is especially powerful when applied to natural image data, like photographs. This is because the spectral content of these images is concentrated in the low frequencies, allowing for the safe quantization of the high frequency data. A more general approach, considering that some images can have some regions with high spatial frequencies and others with low spatial frequencies, is dividing the image into blocks of pixels, and then quantizing the spectrum of each block individually. The JPEG image format uses this blockwise spectral quantization, dividing the image in blocks of 64x64 pixels and using the discrete cosine transform instead of the Fourier transform to avoid complex valued data.

If we apply spectral quantization to holographic data, for example to a Fourier hologram, we find that its spectral information is the reconstructed object. Since the object has a limited size, the pixels outside the area of the object (which correspond to high spectral frequencies of the hologram) can be quantized with a lower bit depth than the pixels with the object information. Nevertheless, this procedure is of limited usefulness for holographic data because filtering already eliminates the pixel information outside the object.
Regarding blockwise spectral quantization, the main difficulty is that if the input has random noise or very high entropy, in the transform of each block the weight of both high and low frequencies will be relatively similar. In this case, spectral quantization cannot result in significant compression ratios without excessive data loss. This effect was demonstrated by Shahnaz *et al.* [28] when dealing with noisy images. This is the case when dealing with OFD's phase from diffuse holograms. The entropy of these phases is very high, and as a result, block-wise spectral quantization cannot produce optimal results.

More recently, the concept of spectral quantization has been used with other types of mathematical transforms, like the wavelet transform. This is the basis of the JPEG2000 compressed image format [29]. Nowadays there is intense research regarding wavelet transform methods applied to the compression of holographic data. In particular, the same group responsible for JPEG and JPEG2000 is currently establishing the groundwork for the JPEG Pleno format [30]. This format is expected to be able to compress images, holograms, point clouds, and integral images effectively.

Despite the advances in this field, there is another approach worth exploring to achieve holographic data compression: the use of techniques inspired by optical systems and procedures.

4. Optical methods

4.1. Optical scaling

In the previous section, we found that the digital scaling of the hologram produces a large amount of degradation, even when the reduction in resolution is very small. This is because the

OFD's phase contains discontinuities due to the wrapped nature of the phase, and the interpolation applied in digital scaling "blurs" these discontinuities. This, in turn, leads to a massive loss of quality in the reconstruction.

To solve this issue, Sorayda *et al*. [13] proposed optical scaling for the compression of optically encrypted data. The basis of this method is using a virtual optical imaging system with a magnification of less than one. This results in an output with a smaller resolution. Additionally, since the procedure is performed optically simulating the propagation of light and the phase shift of the lens, no phase interpolation is performed, and as a result, there is a much smaller loss in quality.

As shown in figure 6, if we place an OFD at a distance d_1 of a positive lens with focal length f, we will obtain an image at a distance d_2 behind the lens with a magnification E given by

$$E = \frac{f}{f - d_1} \tag{6}$$

If the input has a resolution of $N_f \times M_f$, after scaling its resolution will be $E^2 \times N_f \times M_f$, with E the magnification. This procedure will result in a compression ratio equal to $1/E^2$.

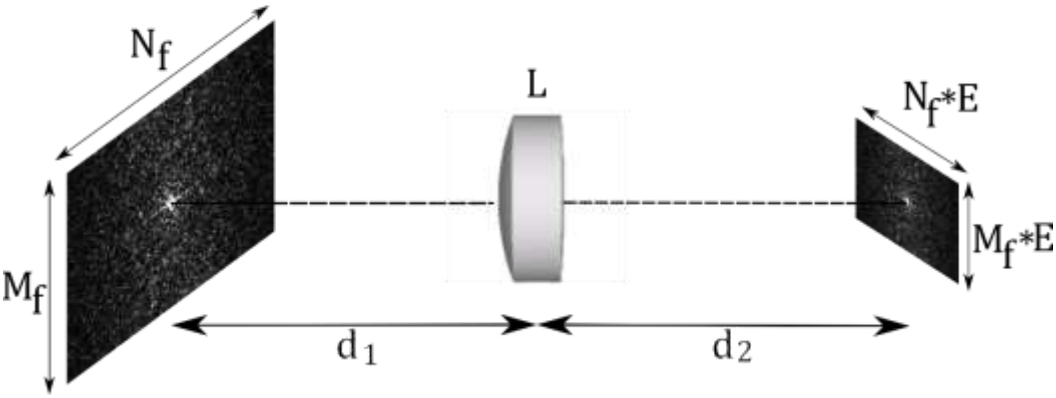

Figure 6: Scheme of the virtual optical setup used for optical scaling.

Although optical scaling was initially applied to optically encrypted data, Velez *et al.* [14] demonstrated its application to the compression of digital holograms. In particular, optical scaling was applied to the full complex-valued OFD extracted from the hologram of a diffuse object. As a result, it was shown that optical scaling produced more degradation than JPEG compression for low compression ratios. Nevertheless, when large compression ratios are necessary, optical scaling demonstrates better performance. Another interesting result is that optical scaling method performed equally well for both phase and amplitude information, while the JPEG approach demonstrated much lower performance for phase compression. This behavior is a manifestation of the difficulties of phase compression with spectral quantization discussed in the previous section.

4.2. Random sampling

Another method for holographic data compression inspired by optical methods is random sampling with binary masks. This approach was first introduced by Davis & Cottrell [31], to multiplex holographic filters, and then applied to the selective multiplexing of Fresnel holograms by Velez *et al.* [32]. Multiplexing consists in combining several data channels into a single one.

Multiplexing has been performed optically using different approaches, like spatial [33], angular [34,35], polarization [36], and wavelength multiplexing [37]. In general, multiplexing does not lead to data compression, since often the data volume of the multiplexed package must be greater than the individual holograms to avoid issues like crosstalk and significant degradation after reconstruction. However, random sampling with binary masks can effectively lead to data compression.

Random sampling with binary masks was initially used to multiplex the full OFD from several holograms with the same resolution. It consists in defining a binary mask of the same resolution of the holographic data to be multiplexed, where a percentage of the pixels are set to a value of 1 and the remaining pixels are set to a value of 0. Then, the OFD is multiplied by this binary mask. As a result, the information of the pixels where the binary mask takes the value of 0 are lost, causing degradation after reconstruction.

However, holographic data is highly resistant to data loss due to this random sampling. For example, in figure 7, we show the reconstruction from OFDs with different sampling percentages, and the original reconstruction from the full OFD.

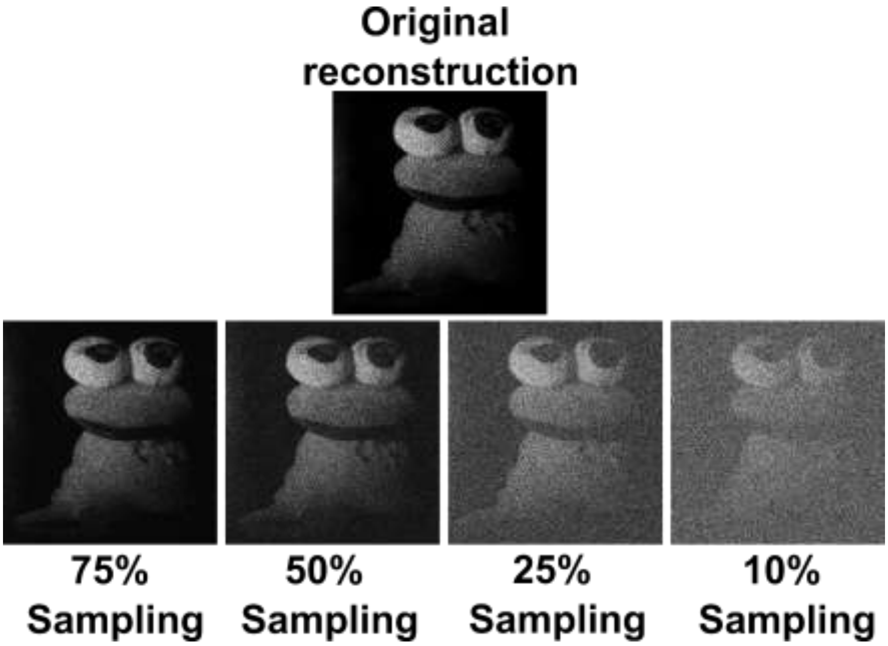

Figure 7: Reconstruction of the same object with different sampling percentages.

As can be seen, the reconstruction quality is proportional to the sampling percentage, with less than 50% sampling leading to significant degradation.

To achieve the maximum compression using this method, we can apply it to directly to the OFD's phase. In this case, we first we generate a set of orthogonal binary masks, so that the 0 valued pixels of one mask coincide with the pixels with value 1 of the other. Sampling the OFD's phases with different masks from this set, and then multiplexing the result, allows the combination of all sampled OFD's phases into a single multiplexed package. This package is a phase function with the same resolution and the same data volume as single initial OFD's phase. We can select the object to be reconstructed by multiplying the package by the binary mask with which the OFD's phase of the desired object was sampled. Then, after reconstruction, we obtain only the selected object without any crosstalk. A scheme of this method is shown in figure 8.

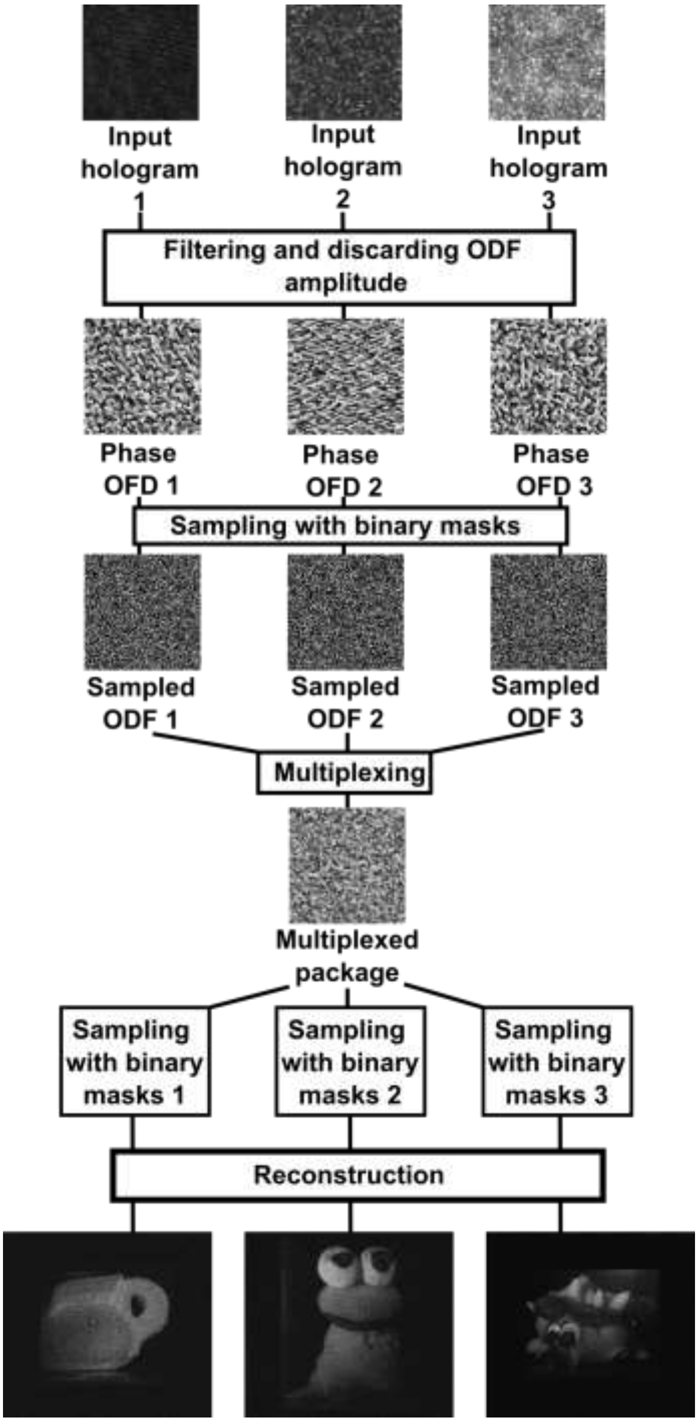

Figure 8: Scheme of the sampling with binary masks multiplexing and compression procedure**.**

Random sampling with binary masks enables a compression ratio of $1/N$, where N is the number of multiplexed holograms. In practice, obtaining compression ratios of more than 3 with this method implies the use of low sampling percentages, introducing severe degradation.

Both the optical scaling and random sampling procedures lead to limited compression ratios. Further compression with these techniques causes significant degradation. Despite of this, Trejos *et al.* and Gomez-Valencia *et al.* demonstrated that it is possible to successfully combine these procedures for compression of both Fourier [15] and Fresnel [16] holographic videos, achieving compression ratios of up to 15.91 compared to an initial full OFD with acceptable degradation.

5. Alternative phase representations

Another approach to data compression is the combination of optical and digital methods. In particular, there are methods common in optics that can be potentially used to prepare holographic data in such a way that they can be successfully processed with digital compression methods. We will now discuss a way to achieve this.

As mentioned previously, holographic data has very high entropy, which limits the usefulness of lossless algorithms. This entropy is especially high in the OFD's phase, which is caused mainly by the discontinuities due the fact that phase values are limited to a range between 0 and 2π. If the phase of a light field is beyond this range, it is increased or decreased by an integer multiple of 2π. This is called phase wrapping. A commonly used approach to avoid this issue is encoding the phase information into a real-imaginary representation. The real and imaginary parts of a phase function are continuous, and as such present lower entropy, which can make lossless compression algorithms more efficient.

An alternative to the real-imaginary representation to reduce the entropy of phase information is the use of phase unwrapping. Phase unwrapping covers a broad range of algorithms that analyze the wrapped phase to eliminate its discontinuities and produce a continuous function. These methods are particularly used in metrology [38–40], radar [41], and medical [42] applications.

In general, the unwrapped phase will have lower entropy than the wrapped phase due to the elimination of discontinuities, but it will also present a higher dynamic range. This can be an issue if we only have a limited bit-depth to store the unwrapped phase, leading to quantization errors. To address this issue, Velez *et al.* [43] introduced a partial phase unwrapping method, where the phase was only unwrapped up to maximum dynamic range. The partial phase unwrapping, although not completely devoid of discontinuities, has lower entropy than the initial phase, allowing improved lossless compression. A basic scheme of this alternative phase representation approach is shown in figure 9. Compression ratios of up to 4.5 were achieved by applying the DEFLATE lossless compression algorithm [6] to the partially unwrapped phase while maintaining a CC of 0.92.

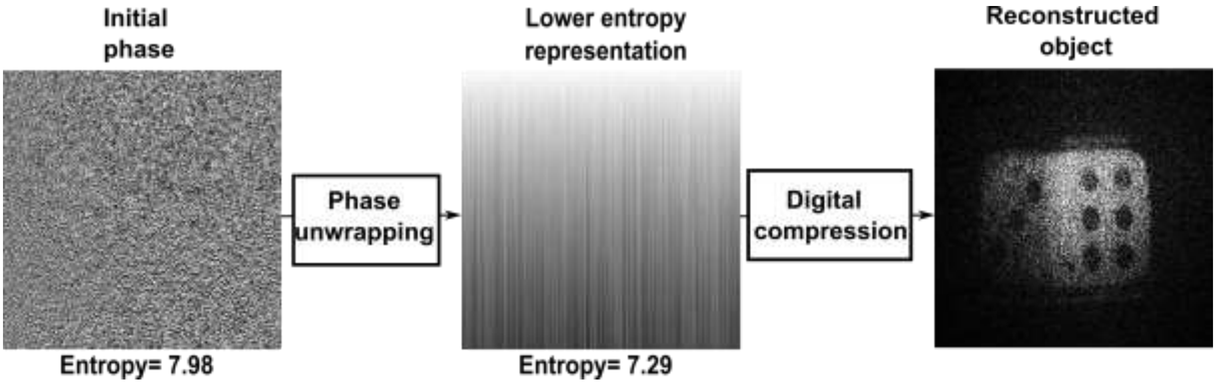

Figure 9: Scheme of the alternative phase representation with the phase unwrapping compression method.

A further refinement of this method was later introduced [44], making use of a quality guided phase unwrapping algorithm. The combination of this approach with the JPEG lossy compression format allowed to reach compression ratios between 17 and 26, with CCs of 0.80. Given these results, alternative phase representations combined with digital methods offer the potential for much higher compression than the individual use of digital or optical approaches.

6. Conclusions

The techniques shown in this work are only a subsection of a large and active research field, they give a first approximation about the different possible approaches to contribute to the holographic data compression field. The results show a large data reduction by taking advantage of filtering and then discarding the amplitude of the resulting OFD, especially when dealing with diffuse objects. Then, we introduce some common digital compression approaches, highlighting the challenges faced when applying digital techniques to holographic information. Afterwards, we show two techniques based on optical techniques, which show better or comparable performance to lossy digital methods. Further research in optical methods for holographic data compression, which preserve the features of phase information, maybe the key to future holographic compression codecs. Finally, we discuss the alternative representation method, an optodigital technique that has the potential to enable the large compression ratios common in video and image compression when applied to holographic data.

Nevertheless, the demand for large volumes of holographic data seems to be increasing faster than the capability for highly efficient compression. In this sense, we believe that the path is open for novel contributions that help bridge this gap.

Acknowledgments

Comité para el Desarrollo de la Investigación—CODI (Universidad de Antioquia—UdeA, Colombia); Fondo Nacional de Financiamiento para la Ciencia, la Tecnología y la Innovación 'Francisco José de Caldas' No. 2019-848 (MINCIENCIAS-Colombia); Agencia Nacional de Promoción Científica y Tecnológica No. 2018-04558 (Argentina); CONICET No. 0849/16 (Argentina).