An exploratory study towards objective quality evaluation of digital hologram coding tools

Roberto Corda, Antonin Gilles, Kwan-Jung Oh, Antonio Pinheiro, Peter Schelkens, et al.


Event: SPIE Optical Engineering + Applications, 2019, San Diego, California, United States
An exploratory study towards objective quality evaluation of digital hologram coding tools

Roberto Corda\textsuperscript{a}, Antonin Gilles\textsuperscript{b}, Kwan-Jung Oh\textsuperscript{c}, Antonio Pinheiro\textsuperscript{d}, Peter Schelkens\textsuperscript{e,f}, and Cristian Perra\textsuperscript{a}

\textsuperscript{a}DIEE, UdR CNIT, University of Cagliari, Italy
\textsuperscript{b}IRT b<>com, Cesson-Sévigné, France
\textsuperscript{c}Electronics and Telecommunications Research Institute (ETRI), Korea
\textsuperscript{d}Universidade da Beira Interior and Instituto de Telecomunicações, Portugal
\textsuperscript{e}Vrije Universiteit Brussels, Belgium
\textsuperscript{f}imec, Leuven, Belgium

ABSTRACT

Holography is an acquisition and reproduction technique of visual content which allows, theoretically, for the reconstruction of the acquired scene without any difference with its real-world counterpart. The objective quality assessment of digital holograms coding tools is a very challenging problem because the signal properties of holograms are significantly different from those of regular images. Several approaches can be devised for holography compression and objective quality evaluation. The exploratory study presented in this paper aims at assessing a procedure for objective quality evaluation of data compression tools when applied to the hologram plane.

Keywords: Digital holography, JPEG Pleno, compression, quality

1. INTRODUCTION

Holography is an acquisition and reproduction technique of visual contents which allows, theoretically, for the reconstruction of the acquired scene without any difference with its real-world counterpart.\textsuperscript{1} The most popular 3D visualization technologies, which often are based on the stereoscopic principle,\textsuperscript{2–4} are affected by the vergence-accommodation conflict (VAC) as major problem and difference with respect to the real world vision. On the contrary, the VAC is absent in holography.

Recently, digital holography received a lot of attention because it simplifies acquisition, reproduction and transmission processes when compared to analog holography. Microscopy, tomography and interferometry are examples of the main use cases.\textsuperscript{3} In microscopy applications, digital holography can bring several benefits, as it aids to overcome the limited field of depth at high magnifications, without the need for laborious mechanical refocusing.\textsuperscript{6,7} In screening applications it can be used as a fast labelling-free technique, allowing for cell analysis without inducing alterations.\textsuperscript{8} In tomography, morphology and chemical information about the cell can be obtained studying its refractive index (RI) and high resolution 3D RI tomograms can be numerically reconstructed to obtain useful information regarding the pathophysiology of diseases.\textsuperscript{9,10} Interferometric acquisition allows for non-invasive measurements of displacements and/or deformations of materials. This acquisitions allow for nanometer precision depth measurements.\textsuperscript{11} This investigation technique can be applied in a wide range of fields,
ranging from cortical bone compression and load tests, to wood deformation caused by humidity but also to perform real-time measurements of oscillation models of large structures as buildings.

Display applications is another context in which holography could introduce thrilling innovations. Different types of holographic displays have been proposed so far, including head-mounted prototypes, which are less penalized from a technical point of view, due to the limited display size and the short distance between the display and the observers eyes. Head-mounted devices could therefore be the first commercially available holographic displays in the near future, while bigger high quality multi-user displays are not supposed to be available in the near future.

Different research challenges still need to be addressed, and data compression is one of those: digital holograms are characterized by a vast amount of data, having several different features compared to classical image and video content. Data compression is thus an aspect of strategic importance in the development and deployment of holography as imaging technique. Standard lossless compression algorithms have been employed and different quantization methods have been compared on optically acquired holograms. Transforms that are successfully employed on standard image and video contents, such as the Discrete Wavelet Transform and the Discrete Cosine Transform, have been applied also on digital holograms, as well as other types of transforms such as Fresnelets and Gabor wavelets, better suited for holographic contents, have been proposed and tested on digital holographic data. Standard still image and video codecs have been also evaluated as well as modifications to the standard codec pipelines.

To improve rate-distortion performance, the hologram coding after the back-propagation has been investigated and relying on this concept, recently a new lossless integer Fresnel transform method has been proposed and tested on computer generated holograms. Moreover, machine learning based methods have been proposed: a convolutional neural network has been developed to enhance holograms after the compression with the JPEG codec, but also an entire deep neural network compression and decompression scheme has been proposed.

A procedure for objective quality evaluation of digital hologram coding tools is presented and discussed. The procedure considers a basic setting where the hologram plane is the input to the coding process. Signal to noise ration (SNR) on the hologram plane and average peak SNR (PSNR) on the reconstructed images are chosen as objective quality metrics. Some digital holographic samples are selected from the JPEG Pleno database as reference data for testing the overall procedure. The experimental evaluation takes into consideration several coding tools such as JPEG 2000, H.265/HEVC, Google’s VP9 and AOMedia Video 1 (AV1).

The paper is organized as follows. The JPEG Pleno dataset is discussed in Section 2. The proposed framework for objective quality evaluation of digital holography compression is presented in Section 3. Section 4 reports the experimental analysis. Finally, Section 5 draws the conclusions.

## 2. JPEG PLENO DATASET

To evaluate the performance of coding technologies it is of utmost importance to define a test data set that is sufficiently diverse and challenging such that the targeted application domains are sufficiently covered in terms of content behavior and consequently the investigated coding technologies effectively stress-tested. Hence, in the context of the JPEG Pleno standardization effort of which it is the intention to produce among others a
specification for holographic content coding, the JPEG committee created the JPEG Pleno Database, which is currently composed of the following eight datasets: B-Com Repository, ERC Interfere I, II and III, Tomocube, EmergImg-HoloGrail v1 and v2, WUT Display and WUT Microscopy. The holograms have been generated with different methods and they target different application fields. In particular, B-Com, EmergImg-HoloGrail, ERC Interfere, and WUT Display are dedicated to display applications, while Tomocube and WUT Microscopy target microscopy and tomography applications. The database comprises 63 still samples and 7 video sequences (the latter belonging to B-Com Repository and Emerg-Img Holograil v1). The B-Com Repository and ERC Interfere samples are Computer Generated Holograms (CGH), for a total of 45 CGH. The EmergImg-HoloGrail, Tomocube and WUT Display/Microscopy have been optically acquired, for a total of 18 optically acquired samples.

The holographic database created by the Institute of Research and Technology B-Com is composed of different CGH, generated with two different algorithms. The dataset comprises 21 holograms acquired from synthetic scenes and have resolutions from 1920 × 1080 to 16384 × 16384 with pixel pitch from 6.4µm to 0.4µm. Two video sequences at different resolutions of real scenes are also proposed.

Interfere I comprises both monochromatic holograms generated from 2D and 3D objects and produced with the multiple wavefront recording plane CGH method with occlusion culling. All the holograms have a pixel pitch equal to 8µm and a resolution equal to 1920 × 1080, except for the sample 3D Cat, whose pixel pitch is equal to 2µm and it has resolution of 8192 × 8192. Improving the algorithm of Interfere I, the authors have proposed Interfere II, adding the simulation of diffuse light reflection on surfaces. The data set comprises a total of 12 monochromatic holograms. The resolution is 8192 × 8192 with a pixel pitch of 1µm for all holograms. The most recent Interfere III data base, consists of 6 color CGH. The generation algorithm based on includes an extension of the Phong illumination model that allows ambient, diffuse and specular reflection support. The holograms have resolutions between 1920 × 1080 and 16384 × 16384, with pixel pitches ranging from 8µm to 1µm.

Other than computer-generated holograms, the JPEG Pleno Database includes also some examples of optically acquired datasets of macroscopic objects. The EmergImg-HoloGrail is one of of them, and each hologram is derived from the acquisition and combination of four interferograms in order to retrieve phase information. It is organized in two sets of monochromatic holograms. The first set (v1) includes three holograms with a resolution of 972 × 972 and 4.4µm of pixel pitch, and a video sequence consisting of 54 holograms at 600 × 600 resolution. The second set (v2) comprises six 2588 × 2588 samples with pixel pitch equal to 2.2µm.

The WUT Display dataset includes one monochromatic and one color hologram, acquired with a synthetic aperture technique. The former has resolution of 19794 × 2020 while the latter’s resolution is 5394 × 2016, and both have a pixel pitch of 3.45µm.

As it concerns microscopy applications, the JPEG Pleno Database comprises the WUT Microscopy dataset, recorded with an off-axis configuration and composed of three holograms with resolution 2464 × 2056 and pixel pitch of 3.45µm. The last dataset currently included in the JPEG Pleno Database is released by Tomocube, and comprises four monochromatic samples with resolution from 512 × 512 to 24478 × 16641 and pixel pitch from 0.08µm to 0.366µm.

In the B-Com Repository, ERC Interfere, EmergImg-HoloGrail and Tomocube datasets, the samples are provided as complex matrices. The WUT holograms are represented as interferograms, and also EmergImg-HoloGrail v1 provides the interferograms from which the complex representation is derived. The datatype of WUT samples is 8 bit unsigned, while the Interfere samples are expressed in 32 bit unsigned float. B-com Repository samples are provided both in 8 bit unsigned and in 32 bit float. Finally, the complex representation of EmergImg-HoloGrail samples is expressed in 64 bit float type. The data set features described above are summarized in Table 1, while in Fig. 1 some examples of reconstructed holograms of these databases are provided.
3. DIGITAL HOLOGRAPHY COMPRESSION

In this Section, the framework employed for the objective quality evaluation and comparison of different coding tools on digital holograms is described. The codecs under test are AOMedia Video 1 (AV1) 0.1.0, H.265/HEVC (reference software 16.18), VP9 (FFmpeg 4.0) and JPEG 2000 (Kakadu Software 7.10.2). The first three are video codecs, while the last is a still image codec. The scheme of the framework is reported in Fig. 2 and will be described in the hereafter.

The holograms that have been employed in the compression experiments are solely CGHs that are represented by a matrix, in which each element is a complex number. Standard codecs typically do not support this type of representation as input data, thus pre and post processing operations are performed on the data before the compression stage and after the decompression stage. The Hologram plane data block is the two-dimensional array of complex values, from which the real (Re) and imaginary (Im) components are calculated within the Complex to components block. Although the complex data could be converted in other forms (such as the amplitude-phase format), the real-imaginary format has been chosen since it is known to provide better compression performances. The real and imaginary components, are both mapped in the range \([0, 1]\) in the Range mapping & Quantization block, and then each component is quantized to 8 bit per sample (bps), resulting in the range \([0, 255]\). This step was included to provide a compatible input to all codecs under test. In the Coding block, the two components are independently encoded with each codec at seven different quality levels. Subsequently, the Decoding block performs the decoding of the coded components. At this stage, the
effects of the compression are evaluated through the SNR metric, defined as:

$$SNR = 10 \log_{10} \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |r(x,y)|^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |r(x,y) - t(x,y)|^2}$$  

where \( r(x,y) \) is the reference data, while \( t(x,y) \) denotes the data under test, in order to compare the two components before and after the coding process.

Subsequently, the complex matrix that represents the hologram is restored through the Components to complex block, and the holographic image is reconstructed in the Rendering block in a predetermined reconstruction plane, aka object plane. This operation is performed on the hologram that has been compressed, represented by Complex data (decoded), but also on the original hologram represented by Complex data (raw). The two reconstructed images are finally compared through the PSNR metric:

$$PSNR = 10 \log_{10} \frac{(2^n - 1)^2}{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |r(x,y) - t(x,y)|^2}$$

For each quality level of the codecs, the PSNR evaluation is performed (when allowed by the hologram under test) on several reconstructions, carried out at different depths and different viewpoints.

4. EXPERIMENTAL ANALYSIS

In this Section, the experimental results obtained applying the compression scheme described in Section 3 are shown. In this paper, we limit the reporting to a subset of holograms selected from the aforementioned datasets. The holograms under test are 3D Multi from Interfere I, Ball 8KD from Interfere II, Venus from Interfere III and Specular Car 8K from B-Com Repository.

For each hologram, the compression results of the real and imaginary components (in terms of SNR) and the objective quality assessment in the reconstructed planes (in terms of PSNR) are reported. The four holograms under test allow for the image reconstruction at different distances, which fall in a specific interval defined by the hologram itself. Different portions of the scene can be focused by varying the reconstruction distance. Taking this feature into account, for Ball 8KD, Venus and Specular Car 8K, the PSNR is thus evaluated, for every quality level of the codecs, at 30 different reconstruction distances. These reconstruction distances cover uniformly the specific reconstruction distance range allowed by each hologram. The graphs below show the mean PSNR obtained with these 30 different reconstructions, along with the standard deviation. For the 3D Multi hologram instead, the authors recommend three different reconstruction distances, each of which brings in focus one of the three dices depicted in the holographic image, therefore for this hologram only three different reconstruction distances are evaluated and averaged.

The Ball 8KD and Specular Car 8K holograms allow also for the reconstruction at different viewpoints, therefore for these two holograms, the PSNR is evaluated at 36 different reconstruction viewpoints, for every quality level of the codecs. The 36 different viewpoints have been chosen in order to uniformly cover the range of different viewpoints allowed by the specific hologram under test. Also in this case, the mean PSNR along with its standard deviation is reported in the graphs.

In Figs. 3–4 the compression results obtained respectively with 3D Multi and Ball 8KD are shown. Both samples are gray scale; however, the former does not support reconstruction from different viewpoints, as opposed to the second. In terms of components, it can be noted as the real (Fig. 3a and Fig. 4a), and the imaginary part (Fig. 3b. Fig. 4b) shows almost identical results on both holograms, so they react in a very similar manner to compression. JPEG 2000 is performing the poorest on both samples, for bitrates up to 2.5bps on 3D Multi and up to 4bps on Ball 8KD. For bitrate values close to zero and up to 1.9bps, in both holograms the performance difference among the video codecs is negligible, while above 1.9bps it can be noted the superior performances of
HEVC and AV1, on both samples, while the distortion of VP9 saturates as the bitrate increases. On the contrary, JPEG 2000 increases gradually its performance as the bitrate increases. With regard to the average quality of the reconstructed image with the variation of reconstruction distance, the results on 3D Multi are reported in Fig. 3c. Each point in the graph represents the mean PSNR value, while the vertical interval represents the standard deviation. As could be expected from the component analysis, the video codecs are the best up to 4.5bps, while the JPEG 2000 is the worst. Beyond this value, the difference between AV1, HEVC and JPEG 2000 is negligible, while VP9 is the worst. The average quality of the reconstructed image with the variation of reconstruction distance for Ball 8KD is reported in Fig. 4c. It can be noted that also in this case the results are consistent with the results obtained in terms of components, with the video codecs that provide the best performances at low bitrates, while JPEG 2000 improves at higher bitrates.

Similar observations can be made analyzing the image quality with the change in viewpoint (Fig. 4d). The PSNR standard deviation is slightly higher at medium-low and at high bitrates. However the highest standard deviation values never exceed ±2dB. It can also be noted that Ball 8KD, at the same bitrates, shows lower SNR and PSNR values than 3D Multi. Ball 8KD is a “more advanced” sample, that supports diffuse reflections and full parallax effects, and these features affect also the compression performances.

Fig. 5 and Fig. 6 show the results obtained respectively from the Venus and Specular Car 8K samples. Also in this case, only the second hologram supports reconstructions from different viewpoints; however, in this case the two samples are RGB. It can be noted that even on these two samples, the real and imaginary parts react in the same way on the compression process. On Venus (Fig. 5a–5b) the performances of the four codecs are similar up to 11bps, while above this threshold HEVC and JPEG 2000 have the best performances. The worst codec is VP9, while the most recent AV1 has lower performances at high bitrates, if compared to HEVC and JPEG 2000. The results of Specular Car 8K in terms of components (Fig. 6a–6b) show instead that up to 4bps the best codecs are the most recent (and complex) AV1 and HEVC. However, beyond this threshold, AV1 does not maintain performances similar to HEVC, but on the contrary its performances become similar to the worst codec, VP9, as the bitrate increases. JPEG 2000 also in this case improves its performances as the bitrate increases, but surpasses HEVC only at high bitrates, close to 14bps. In both samples the overall PSNR trends (Fig. 5c and Fig. 6c–6d) are consistent with those of the SNR in terms of components, thus the same consideration apply. For what concerns the results obtained with Venus in terms of mean PSNR with the change in reconstruction distance, it can be noted that the standard deviation is near zero for all codecs: there are not appreciable variations in image quality with varying distances. However, the Specular Car 8K shows a different behavior. Both with the variation in the reconstruction distance (Fig. 6c) and in the viewpoint (Fig. 6d) the standard deviation is not equal to zero and it shows a gradual reduction with the increase in the bitrate. Also in the worst cases, which occur especially at medium-low bitrates, it never exceeds ±2dB, similar to 3D Multi and Ball 8KD. Variations of that entity should be rarely perceptible by the observer.

In some cases it can be noted an unpredictable trend of the Rate-Distortion curves, as in Ball 8KD, with the variation in the reconstruction distance with the JPEG 2000 codec (Fig. 4c). This fact can be caused by the

![Figure 3: 3D Multi results: SNR real part (a), SNR imaginary part (b), mean PSNR at different reconstruction distances (c).](image-url)
lack of data content-awareness of the codecs: the modifications of the standard lossy coding (that cannot takes the hologram features into account) can have unpredictable effects of the reconstructed image quality.

Finally, it should be noted that the process of hologram rendering into the object plane originates the so called speckle noise which drastically affects the PSNR measures. The influence of speckle noise on the reconstructed image quality is an open research problem.

Figure 4: Ball 8KD results: SNR real part (a), SNR imaginary part (b), mean PSNR at different reconstruction distances (c), mean PSNR from different viewpoints (d).

Figure 5: Venus results: SNR real part (a), SNR imaginary part (b), mean PSNR at different reconstruction distances (c).

Figure 6: Specular Car 8K results: SNR real part (a), SNR imaginary part (b), mean PSNR at different reconstruction distances (c), mean PSNR from different viewpoints (d).
5. CONCLUSIONS

A digital hologram is a representation of the plenoptic function, which carries more information of a conventional digital photo. It allows for the reconstruction of a scene at different viewing angles and reconstruction planes. A digital hologram can correspond to a considerable amount of samples which creates a demand for compression when being used in practical applications. This paper presents an objective quality evaluation framework. It should be noted that the compression is performed naively in the hologram plane and that the codecs under test have not been tuned to operate in this domain. The sole purpose of this experiment is to evaluate the overall procedure. More advanced settings can also include backprojection operation to allow for compression in the reconstruction plane and/or deployment of more specialized codecs that are better suited to handle holographic content.

ACKNOWLEDGMENTS

The research leading to these results received funding from the Cagliari2020 project (MIUR, PON04a2 00381), the DigitArch Cluster Top-Down project (POR FESR, 2014-2020), the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement n. 617779 (INTERFERE), and the Giga-KOREA project (GK19D0100, Development of Telecommunications Terminal with Digital Holographic Table-top Display).

REFERENCES


[34] “JPEG Pleno Database.” https://jpeg.org/plenodb/. (Online; accessed on 1 July 2019).


