ISO/IEC JTC1/SC29/WG1
(ITU-T SG16)

Coding of Still Pictures

JBIG	JPEG
Joint Bi-level Image Experts Group	Joint Photographic Experts Group

Title:	JPEG Pleno Holography Common Test Conditions 5.0
Source:	Raees Kizhakkumkara, Ayyoub Ahar, Tobias Birnbaum, Antonin Gilles, Saeed Mahmoudpour and Peter Schelkens
Project:	JPEG Pleno Holography
Version:	July 19, 2021
Status:	For distribution
Requested action:	None
Distribution:	Public

Contact:
ISO/IEC JTC1/SC29/WG1 Convener – Prof. Touradj Ebrahimi
EPFL/STI/IEL/GR-EB, Station 11, CH-1015 Lausanne, Switzerland
Tel: +41 21 693 2606, Fax: +41 21 693 7600, E-mail: Touradj.Ebrahimi@epfl.ch

Editorial Comments

This is a living document that goes through iterations. Proposals for revisions of the text can be delivered to the editor Peter Schelkens by sending it to `Peter.Schelkens@vub.be`.

If you have interest in JPEG Pleno Holography, please subscribe to the email reflector, via the following link: `http://jpeg-holo-list.jpeg.org`.

Contents

JPEG Pleno Holography Common Test Conditions 5.0

1 Scope

This document describes the JPEG Pleno Holography Common Test Conditions 5.0 for performance assessment of proposals submitted to the Call for Proposals on JPEG Pleno Holography and additionally defined Exploration Studies and Core Experiments. This document can also be considered as a guideline for testing various types of compression algorithms for holographic content. In Section 2, an overview of the test material is provided, summarizing the main properties of the content and download information. Section 3 defines the rate and quality metrics and subsequently discusses the measurement configurations, coding conditions and anchor specifications. Section 4, details image (hologram) and measurement data output configuration. The subjective test procedure is described in Section 5.

2 Test materials

This section describes the currently selected test material selected for JPEG Pleno Holography Call for Proposals (CfP), Core Experiments (CE) and Exploration Experiments (EE). The selection is justified by the diversity of the holograms in terms of intrinsic properties such a complexity and depth of the represented scene. The holograms are chosen to reflect diverse use cases and generation methods (see macroscopic complex-valued holograms in Table 1, macroscopic binary holograms in Table 2, and metrology holograms in Table 3). Note that a larger set of reference holograms is retrievable from `plenodb.jpeg.org`. These holograms can be classified by their use case into:

- **Holograms for visualization -** These holograms are intended for visualization and printing purposes and feature objects of sizes that are visible by human eye.
- **Microscopy and interferometry holograms -** These holograms are either (1) microscopic measurements of small objects like biological cells and microspheres or (2) metrology holograms that are usually characterized by large resolutions. Apart from static captures, microscopic holograms can also be used for time-lapse recordings and holographic tomography[9].

Holograms can be also classified by their generation method into

- **Computer generated holograms -** These are typically macroscopic holograms that are generated computationally using the principles of light wave propagation. The methods used to generate holograms can be broadly grouped under 4 categories - point cloud based synthesis, triangular mesh based synthesis, layer based synthesis and ray based synthesis [11, 2].
- **Optically captured holograms -** Optically recorded holograms are captured as actual physical measurements obtained typically by modulations of amplitude and phase.

Tab. 1: Floating-point holograms in JPEG Pleno Holography test set and associated parameters. CGH=computer-generated hologram; OCH=optically captured hologram; `plenodb.jpeg.org` DHs for the subjective test are ranked by priority. The final number will be decided upon in function of the number of proposals. The first reconstruction distance mentioned outside of brackets corresponds to the distance used for object plane compression of the anchor pipeline. The remaining reconstruction distances are subject to revision.

Hologram	Testset	Obj. test	Resolution	Aperture size	Pixel pitch (μm)	Wavelength (nm)	OCH/ CGH	Scene depth	Reconstruction distance (mm)	Ref. wave radius R (mm)	Background
Ball	Interfere-IV with WUT	x	2048×16384	2048×2048	3,45	532	CGH	Medium	701; (751)	694,5	No
CornellBox3 16k	Intervere-V	x	16384×16384	4096×4096	2	633	CGH	Medium	0.25307; (220; 228; 250; 269; 286.15)	Inf	Yes
DeepCornellBox 16k	Intervere-V	x	16384×16384	4096×4096	2	633	CGH	Medium	250; (209.0; 212.2; 227.0; 252.5; 255.0; 301.3; 342.5; 363.9; 405.1; 408.0; 418.3; 444.6; 489.1; 530.4; 532.0)	Inf	Yes
Chess	Interfere-IV with WUT	x	2048×16384	2048×2048	3,45	532	CGH	Deep	496.4; (648.6; 806.3)	694,5	Yes
DeepChess	Interfere-IV with WUT	x	2048×16384	2048×2048	3,45	532	CGH	Deep	998.6; (396.4; 1.6063)	998,6	Yes
Sphere	Interfere-IV with WUT	x	2048×16384	2048×2048	3,45	532,8	OCH	Medium	960	960	No
Squirrel	Interfere-IV with WUT	x	1792×27904	1792×3488	3,45	632,8	OCH	Deep	500; (465; 535)	500	No
Biplane16k-1	Interfere-III	x	16384×16384	2048×2048	1	633; 532; 460	CGH	Deep	43.55; (37.4; 45.5; 49.7)	Inf	Yes
Biplane16k-2	b-com	x	16384×16384	2048×2048	1	633; 532; 460	CGH	Deep	43.55; (37.4; 45.5; 49.7)	Inf	No
DeepDices2k	b-com	x	2048×2048	2048×2048	4,8	640; 532; 473	CGH	Deep	86.7; (5.07; 7.41; 166; 246)	Inf	Yes
DeepDices8k4k	b-com	x	7680×4320	7680×4320	4,8	640; 532; 473	CGH	Deep	14.8; (10.1; 17.3; 33.1; 49.2)	Inf	Yes
DeepDices16k	b-com	x	16384×16384	16384×16384	0,4	640; 532; 473	CGH	Deep	18.5; (3.38; 4.94; 18.5; 31.8; 45.9)	Inf	Yes
Dices16k	b-com	x	16384×16384	2048×2048	0,4	640; 532; 473	CGH	Deep	9.83; (6.58; 10; 13.1)	Inf	Yes
SpecularCar16k	b-com	x	16384×16384	2048×2048	0,4	640; 532; 473	CGH	Deep	5; (4.4; 10)	Inf	No
Piano16k	b-com	x	16384×16384	2048×2048	0,4	640; 532; 473	CGH	Deep	9.65; (6.8; 12.5)	Inf	No
Ring16k	b-com	x	16384×16384	2048×2048	0,4	640; 532; 473	CGH	Deep	8; (6; 8; 10;)	Inf	Yes
Ballet8k4k (Frame 22)	b-com	x	7680×4320	7680×4320	4,8	640; 532; 473	CGH	Deep	25; (5.84; 18.94; 21.45; 26.21; 51.87)	Inf	Yes
Breakdancers8k4k (Frame 22)	b-com	x	7680×4320	7680×4320	4,8	640; 532; 473	CGH	Deep	18.855; (4.18; 8.69; 20.60; 23.62; 25; 25.51; 33.53)	Inf	Yes
Astronaut	EmergImg-HoloGrail	x	2588×2588	1940×2588	2,2	632,8	OCH	Deep	-167.5; (-160; -165; -170; -172; -175)	Inf	No
Horse	EmergImg-HoloGrail	x	972×972	972×972	4,4	632,8	OCH	Deep	140; (135; 145; 150)	Inf	Yes
Lowiczanka Doll (OnAxis)	WUT	x	2016×59394	2016×2016	3,45	637; 532; 457	OCH	Medium	1052.5; (1030; 1060; 1075)	1060	No
Warsaw Mermaid (OnAxis)	WUT	x	2010×25730	2010×2010	3,45	632,8	OCH	Medium	350; (340; 355)	350	No

Tab. 2: Binary holograms in JPEG Pleno Holography test set and associated parameters. CGH=computer-generated hologram; OCH=optically captured hologram; `plenodb.jpeg.org` *: DHs were created from the complex-valued pendants and can be reconstructed with the NRSH Software 7.0. The reconstruction distances are subject to revision.

Hologram	Testset	Obj. test	Resolution	Aperture size	Pixel pitch (μm)	Wavelength (nm)	OCH/ CGH	Scene depth	Reconstruction distance (mm)	Ref. wave radius R (mm)	Background	
Ball*	Interfere-IV with WUT	x	2048×32768	2048×2048	3,45	532	CGH	Medium	701; 751	700	No	
CornellBox3 16k*	Intervere-V	x	32768×16384	4096×4096	2	633	CGH	Medium	250; 220; 228; 269; 286.15	Inf	Yes	
DeepCornellBox 16k*	Intervere-V	x	32768×16384	4096×4096	2	633	CGH	Medium	250; 209.0; 212.2; 227.0; 252.5; 255.0; 301.3; 342.5; 363.9; 405.1; 408.0; 418.3; 444.6; 489.1; 530.4; 532.0	Inf	Yes	
Chess*	Interfere-IV with WUT	x	2048×32768	2048×2048	3,45	532	CGH	Deep	496.4; 648.6; 806.3	700	Yes	
DeepChess*	Interfere-IV with WUT	x	2048×32768	2048×2048	3,45	532	CGH	Deep	396.4; 998.6; 1.6063	998,6	Yes	
Sphere*	Interfere-IV with WUT	x	2048×32768	2048×2048	3,45	532,8	OCH	Medium	960	960	No	
Squirrel*	Interfere-IV with WUT	x	1792×55808	1792×3488	3,45	632,8	OCH	Deep	500; 465; 535	500	No	
Biplane16k-1*	Interfere-III	x	32768×16384	2048×2048	1	633; 532; 460	CGH	Deep	45.5; 37.4; 49.7	Inf	Yes	
Biplane16k-2*	b-com	x	32768×16384	2048×2048	1	633; 532; 460	CGH	Deep	45.5; 37.4; 49.7	Inf	No	
DeepDices2k*	b-com	x	2	4096×2048	2048×2048	4,8	640; 532; 473	CGH	Deep	86.7 5.07; 7.41; 166; 246	Inf	Yes
DeepDices8k4k*	b-com	x	15360×4320	7680×4320	4,8	640; 532; 473	CGH	Deep	14.8; 10.1; 17.3; 33.1; 49.2	Inf	Yes	
DeepDices16k*	b-com	x	32768×16384	16384×16384	0,4	640; 532; 473	CGH	Deep	18.5 3.38; 4.94; 18.5; 31.8; 45.9	Inf	Yes	
Dices16k*	b-com	x	32768×16384	2048×2048	0,4	640; 532; 473	CGH	Deep	10; 6.58; 13.1	Inf	Yes	
SpecularCar16k*	b-com	x	32768×16384	2048×2048	0,4	640; 532; 473	CGH	Deep	5; 4.4; 10	Inf	No	
Piano16k*	b-com	x	32768×16384	2048×2048	0,4	640; 532; 473	CGH	Deep	10; 6.8; 12.5	Inf	No	
Ring16k*	b-com	x	32768×16384	2048×2048	0,4	640; 532; 473	CGH	Deep	10; 6; 14.6	Inf	Yes	
Astronaut*	EmergImg-HoloGrail	x	5176×2588	1940×2588	2,2	632,8	OCH	Deep	-172; -160; -165; -170; -175	Inf	No	
Horse*	EmergImg-HoloGrail	x	1944×972	972×972	4,4	632,8	OCH	Deep	140; 135; 145; 150	Inf	Yes	
Lowiczanka Doll* (OnAxis)	WUT	x	2016×118788	2016×2016	3,45	637; 532; 457	OCH	Medium	1060; 1030; 1075	1060	No	
Warsaw Mermaid* (OnAxis)	WUT	x	2010×51460	2010×2010	3,45	632,8	OCH	Medium	350; 340; 345; 355	350	No	
Bridge 100k	ETRI	x	100000 × 100000	-	0,5	660	CGH	Deep	41.29; 17; 50	Inf	Yes	
Dices 200k	b-com	x	204800 × 108000	-	0,24	660; 532; 460	CGH	Deep	10; 32.8	Inf	Yes	

Tab. 3: Holograms in JPEG Pleno Holography test set and associated parameters used for metrological testing. IP-OAH=Image Plane - Off axis hologram `plenodb.jpeg.org`

Hologram	Testset	Type	No. of frames	Resolution	Bitdepth	+1 Order bandwidth	Pixel pitch (μm)	Wavelength (nm)	Reconstruction distance (mm)
PMMA spheres	WUT-DHM	IP-OAH	1	2456x2058	8 bit unsigned integer	374x448	3,45	532	0
USAF	WUT-DHM	IP-OAH	1	2456x2058	8 bit unsigned integer	374x448	3,45	532	0
HaCaT single cell	WUT-DHM	IP-OAH	1	2456x2058	8 bit unsigned integer	374x448	3,45	532	0
HaCaT cell culture	WUT-DHM	IP-OAH	1	2456x2058	8 bit unsigned integer	374x448	3,45	532	0
Brain Tissue	WUT-DHM	IP-OAH	1	2456x2058	8 bit unsigned integer	374x448	3,45	532	0
Neuroblastoma cell line SHSY5Y	WUT-DHM	IP-OAH	121	2456x2058	8 bit unsigned integer	374x448	3,45	532	0
Keratinocyte cell line SHSY5Y	WUT-DHM	IP-OAH	138	2456x2058	8 bit unsigned integer	374x448	3,45	532	0
Nasal epithelium cells	WUT-DHM	IP-OAH	350	2456x2058	8 bit unsigned integer	374x448	3,45	532	0

3 Definition of performance metrics

3.1 Configuration quality metrics

The quality metrics to be computed for each type of hologram are shown in Tab. 4. Note that for macroscopic holograms, the PSNR, SSIM and VIFp scores are calculated for different reconstructions of hologram obtained using the reconstruction software (NRSH) mentioned in Section 4.1.2. The depths, viewing positions, aperture sizes and propagation method required for the NRSH software are defined in Table 1 and Table 2 for each test hologram.

Tab. 4: Deployment of quality metrics.

Hologram type	**Higher precision**		**Binary**		**Metrological**
Metric	**Hologram plane**	**Object plane**	**Hologram plane**	**Object plane**	–
SNR	Yes	–	Yes	–	Yes
PSNR	–	Yes	–	Yes	–
SSIM	Yes	Yes	–	Yes	–
VIFp	–	Yes	–	Yes	–
Hamming distance	–	–	Yes	–	–
SNR of first-order wavefield	–	–	–	–	Yes
RMSE of retrieved phase	–	–	–	–	Yes

3.2 Rate metrics

The bitrate, specified in the test conditions and reported for the experiments with the various codecs, accounts for the total number of bits necessary for generating the encoded file (or files) out of which the decoder can reconstruct a lossy or lossless version of the entire input hologram.

The main rate metric is defined as the number of bits per sample (pixel):

$$\text{Bitrate} = \frac{\text{Total number of bits}}{\text{Number of samples}} \tag{1}$$

where the numerator is the total file size of the encoded file and other files containing side information required for decoding in bits and the denominator is the number of samples(pixels) of the input hologram.

Please note that a sample can be complex valued, in this case the number of bits per sample is the sum of the number of bits for the real and imaginary components.

3.3 Quality metrics

The metrics used for evaluating the quality of the compressed holograms is given in Section 4. The measuring configuration to be used is given in Section 3.3 and depends on the type of hologram being compressed.

SNR and PSNR

The Signal to Noise Ratio (SNR) is defined as the ratio of the power of the signal to the power of the noise affecting the quality of the signal, while the Peak Signal to Noise Ratio (PSNR) is defined as the ratio between the maximum possible power of a signal and the power of noise. The SNR (in dB) is calculated on the complex valued wavefield in the hologram plane and is given by

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\sum\limits_{i=1}^{A}\sum\limits_{j=1}^{B}|X[i,j]|^2}{\sum\limits_{i=1}^{A}\sum\limits_{j=1}^{B}\left|X[i,j]-\hat{X}[i,j]\right|^2}\right) \tag{2}$$

where $X[*,*]$ and the lossy signal $\hat{X}[*,*]$ are the reference hologram and compressed hologram respectively.

The PSNR is used for evaluating the quality of reconstructions at the object plane. These real-valued reconstructions with integer bit-depth are obtained from the NRSH software given in Section 4.1.2 and the PSNR (in dB) is given by Eq. (3)

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{AB\left(2^n-1\right)^2}{\sum\limits_{i=1}^{A}\sum\limits_{j=1}^{B}\left|X[i,j]-\hat{X}[i,j]\right|^2}\right) \tag{3}$$

where n is the bit-depth and $X[*,*]$ and the lossy signal $\hat{X}[*,*]$ are the reconstructions of the reference hologram and compressed hologram obtained from NRSH respectively.

Bjøntegaard metric

The Bjøntegaard metric compares the rate-distortion performance of two coding solution across some rate/distortion region by computing the surface area that lies between the rate-SNR/SNR-rate curves of the two codecs, where the rate axis is logarithmically scaled [1].

SSIM

The Structural SIMilarity (SSIM) index is a full-reference perceptual metric to quantify the visual quality degradation measured by perceived change in structural information [17]. For complex valued data, the SSIM is obtained as the mean of the SSIM of the real and imaginary parts. The SSIM index is bounded between -1 to 1 where, values closer to 1 indicate high correlation and better perceptual quality while values closer to -1 indicates negative correlation. For compression, the range of values will lie closely in the range 0 to 1.

VIFp

The Visual Information Fidelity in pixel domain (VIFp) [13] is a faster implementation of the Visual Information Fidelity (VIF) which performs multi-scale analysis in spatial domain instead of originally utilized wavelet domain in VIF. In it's core, VIF approaches the overall visual process through the human visual system (HVS) as a baseline distortion channel which is added to every input data and models it using a stationary, zero mean, additive white Gaussian noise. Next, the mutual information is calculated between the source model (represented by the natural scene statistics) and the test image after adding the HVS baseline distortion. The value then is normalized by the value of another mutual information similarly calculated for the reference image. VIF is bounded below by 0, which indicates that all information about the reference image has been lost in the distortion channel. In case of no distortion (reference compared to itself), VIF is exactly unity. However, its upper bound is not limited to 1. For example, in case of a linear contrast enhancement of the reference image that does not add noise to it, will result in a VIF value larger than one.

SNR of first-order wavefield

For off-axis holograms the relevant information is encoded in the first-order wavefield. The fidelity of the compressed first-order wavefield is measured by the signal to noise ratio (SNR) metric given in Eq. (4).

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\sum\limits_{u=-B_u}^{B_u}\sum\limits_{v=-B_v}^{B_v}|U_f[u,v]|^2}{\sum\limits_{u=-B_u}^{B_u}\sum\limits_{v=-B_v}^{B_v}|U_f[u,v]-\hat{U}_f[u,v]|^2}\right) \tag{4}$$

where the demodulated first order wavefield in the frequency domain is denoted by $U_f[*,*]$ and its compressed version by $\hat{U}_f[*,*]$ while $[-B_u, B_u]$ and $[-B_v, B_v]$ is the bandwidth of the first-order term.

RMSE of retrieved phase

For quantitative phase imaging, the retrieved phase can provide additional insights on the effect of compression on meterological accuracy in practice. Phase-retrieval is a non-linear process due to the phase unwrapping being performed, which can sometimes introduce strong unwrapping errors even for small errors in the compression. The root mean squared error (RMSE) of the retrieved phase is calculated as shown in Eq. (5)

$$\mathrm{RMSE} = \sqrt{\sum_{i=L_a}^{L_b}\sum_{j=B_a}^{B_b}\frac{\left(\Phi[i,j]-\hat{\Phi}[i,j]\right)^2}{\left(L_b-L_a\right)\left(B_a-B_b\right)}} \tag{5}$$

where $[L_a, L_b]$ and $[B_a, B_b]$ describes the spatial boundary of the phase functions $\Phi[*,*]$ and $\hat{\Phi}[*,*]$ retrieved from the original hologram and the compressed hologram respectively. **Please note that the phase functions refer to the unwrapped phase in radians.**

The phase unwrapping functions to be used is based on efficient multiscale phase unwrapping methodology with modulo wavelet transform [4] applied on the the phase unwrapping via graph cuts (PUMA) algorithm [7].

Hamming distance

For binary holograms $X[*,*]$, the average Hamming distance between the compressed hologram $\hat{X}[*,*]$ is given as

$$\mathrm{H} = \frac{1}{AB}\sum_{i=1}^{A}\sum_{j=1}^{B}\left(X[i,j]\oplus\hat{X}[i,j]\right) \tag{6}$$

where $\oplus$ is the XOR operator.

3.4 Handling of colour information

Currently no validated procedures exist to de-correlate **colour information** in holography. For compression using anchor codecs, the three color channels are compressed independently. For quality evaluation, colour holograms are not converted to another colour space. The quality metrics are computed for each colour channel independently and the arithmetic mean is calculated as well as

$$\mathrm{M} = \frac{\mathrm{M_R} + \mathrm{M_G} + \mathrm{M_B}}{3} \tag{7}$$

where $\mathrm{M_R}$, $\mathrm{M_G}$, $\mathrm{M_B}$ refers to the quality metric for red, green and blue components respectively.

4 Testing pipeline

4.1 Pipeline for anchor codecs

4.1.1 Introduction

Unfortunately, so far no standards have been specified to address coding of holographic content. Hence, only codecs that have originally designed for natural image or binary image content can be deployed as anchor codecs. An additional problem is the fact that these anchor codecs typically do not depict a marvellous rate-distortion performance when directly applied to the hologram itself. Because of this reason two anchor codec pipelines have been devised. In a first pipeline, called the hologram plane coding pipeline, the anchor codec is directly applied to the hologram itself, requiring only a mapping of the typically deployed floating-point in the hologram domain to an integer representation that can be processed by the anchor codec. The second pipeline, the object plane coding pipeline, first the hologram is propagated to the object plane, subsequently converted to integer precision and finally encoded by the anchor codec. Inverting these steps delivers in both cases the decoded hologram, which can then be compared through quality assessment procedures with the original, reference hologram. The different quality assessment procedures deployed are discussed in Section 4.3.

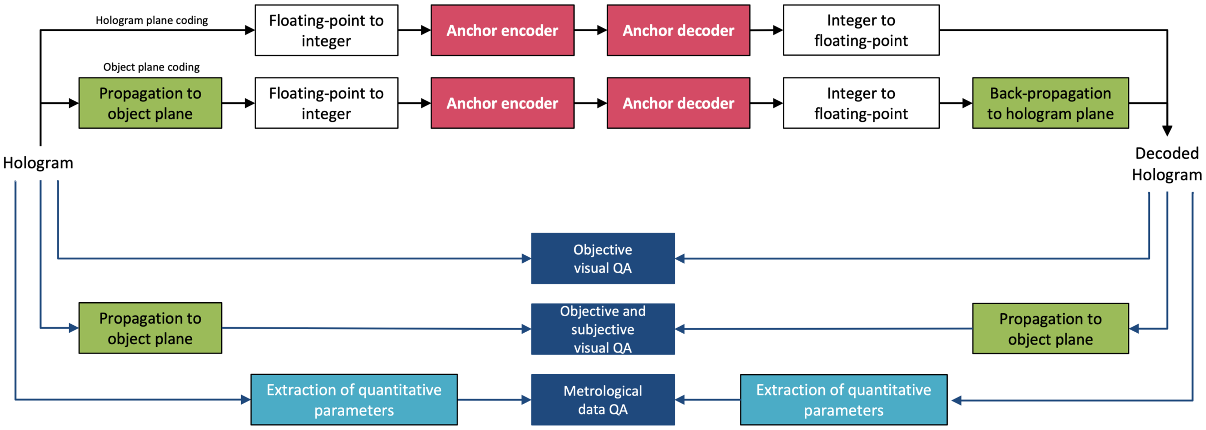

Fig. 1: The anchor codecs are tested in two pipelines, one performing the encoding in the hologram in the hologram plane, the other in the object plane. Visual quality assessment is performed in both planes, except for the subjective visual quality assessment, which is solely performed in the object plane. Metrological data quality is measured directly on the metrological data extracted from the uncompressed (original) and compressed holograms.

Note, compression in object or hologram plane is implemented in the same reference test pipeline software available at `https://gitlab.com/wg1/jpeg-pleno-holo-ctc` (Note: access can be obtained by contacting `Peter.Schelkens@vub.be`).

4.1.2 Propagation to object plane and back-propagation to hologram plane

To assess the objective and subjective visual quality of a hologram in the object plane, the hologram is reconstructed using the numerical reconstruction software (NRSH 7.0) specified in document no.WG1N92032 [15]. This tool comprises two main functions: nrsh and getSettings.

nrsh.m

The nrsh function generates the reconstructions at specified reconstruction points – viewing angles and focus planes – as listed in Table 1 and Table 2. It runs in Matlab 2017b (or higher) with the command:

```
>> [hol_rendered, clip_min, clip_max] = nrsh(hol, rec_dists, info, varargin)
```

where the input parameters are:

- hol: hologram to be reconstructed. It can be a matrix that has been previously loaded in the Workspace, or it can be a path to a folder (provided as character array) (e.g. './holograms/Dices8K/') that contains the file(s) representing the hologram; If left empty, the hologram is loaded manually through a GUI interface.
- rec_dists: reconstruction distance(s) in meters. It can be a single value, or a vector of values;
- info: Reconstruction parameters, initialized with getSettings.

The user can also overwrite the parameters defined in structure info by passing them in this specific order as additional input parameters through varargin: usagemode, ap_sizes, h_pos, v_pos, clip_min, clip_max.

The output parameters are:

- hol_rendered: reconstruction of the input hologram, returned as unsigned integer image (8 or 16 bpp, according to the value set in the configuration file). Note that in case of multiple reconstructions, hol_image is the last reconstruction performed.
- clip_min_out: minimal intensity of the numerical reconstructions. In case of multiple reconstructions, one value per reconstruction is returned.
- clip_max_out: maximal intensity of the numerical reconstructions. In case of multiple reconstructions, one value per reconstruction is returned.

The behavior of nrsh depends on the usagemode member of structure info:

- If usagemode = 'exhaustive', the software calculates all possible combinations of rec_dists, ap_sizes, h_pos, v_pos and performs a reconstruction for each combination of values;

- If usagemode = ‘individual’, it performs a reconstruction for each individual viewpoint listed as input;
- If usagemode = ‘dynamic’, it performs a reconstruction for each individual viewpoint listed as input and saves them as a pseudo-video sequence for subjective testing;
- If usagemode = ‘complex’, it skips all non-invertible transforms (apertures, filters, clipping and resizing) to obtain the complex-valued wavefield in the object plane (used for object-plane coding).

The software automatically resizes, crops or uses any "resize" function supplied as resize_fun argument to avoid writing large size images to the disk. Resizing is performed on the real data, before any conversion to integers. Clipping is done by default with a percentile value for the first reconstruction, and then using the same absolute threshold for the next ones. It is also possible to provide absolute clipping values clip_min and clip_max per reconstruction.

The reconstructions can be saved as MAT files and/or as PNG images (8 or 16 bpp) and are stored in the ./outfolderpath/ConfigurationFileName/ path. The file names are structured as follows:

```
<name_prefix><ConfigurationFileName>_<Hpos>_<Vpos>_<Ap_size>_<Rec_dist>.{mat|png}
```

Finally, if usagemode = ‘dynamic’, the software calculates per specified viewpoint a reconstruction and generates subsequently a video using `ffmpeg -c:v libx264/AVC -qp 0`. The video file using is written to the same folder as the figures of nrsh, i.e.

<outFolderPath>/<ConfigurationFileName>/.

A log file is written to the same folder. Temporary frames in this folder are removed after video generation. The file names of intermediate frames are structured as follows:

```
[<name_prefix> num2str(viewpointId, ’fID%04.0f’) ’.mat’]
```

The video filename is formatted as:

```
<name_prefix><ConfigurationFileName>_nFrames<#Frames>_at<fps>FPS<suffix>.mp4
```

where suffix=’_LR’ in case of diffraction limited reconstructions and suffix remains empty otherwise.

getSettings.m

In NRSH, most input reconstruction setting are read through a structure called info. This structure is initialized by function getSettings, whose declaration is given by:

```
>> info = getSettings(varargin)
```

This function takes as input a list of parameter name and value pairs, such that

```
>> info = getSettings('paramName1', paramValue1, 'paramName2', paramValue2, ...)
```

It returns structure info with fields paramName1 = paramValue1, paramName2 = paramValue2, etc. Allowed parameters are the following:

- usagemode: (string or char array) – optional, default is 'exhaustive'. Usage mode of nrsh. It can take four different values:
  - 'exhaustive': use combination of all possible viewpoints
  - individual': use individual viewpoints as listed
  - dynamic': use individual viewpoints as listed and save them as a video
  - complex': reconstruct the complex light field in the object plane, and disable non-invertible transforms (apertures, clipping, filters, zero-padding)
- ap_sizes: (numeric cell or array) – optional, default is empty. Synthetic aperture size. If the synthetic aperture declaration is based on angles, it must be a single value (or a vector of values) expressed in degrees . If the synthetic aperture declaration is based on pixel, it must be a $1 \times n$ cell array, in which every element is a 1×2 vector that expresses the aperture size in pixels (height x width); more information can be found in the NRSH user guide;
- h_pos: (numeric array) – optional, default is 0. If the synthetic aperture declaration is expressed as angles, it represents the horizontal angles, in degrees, at which the synthetic aperture will be placed. If the synthetic aperture declaration is expressed in pixels, it represents the horizontal position at which the synthetic aperture will be placed, expressed in the range [-1, 1] where -1 is the leftmost position, while 1 is the rightmost position. In both cases (angle or pixel based) it can be a single value or a row or column vector of values.
- v_pos: (numeric array) – optional, default is 0. If the synthetic aperture declaration is expressed as angles, it represents the vertical angles, in degrees, at which the synthetic aperture will be placed. If the synthetic aperture declaration is expressed in pixels, it represents the vertical position at which the synthetic aperture will be placed, expressed in the range [-1, 1] where -1 is the lowermost position, while 1 is the uppermost position. In both cases (angle or pixel based) it can be a single value or a row or column vector of values.
- apertureinpxmode: (boolean) – optional, default is true. True to use pixel-based apertures, false for angle-based.
- clip_min: (numeric array) – optional, default is empty. Minimal intensity values for clipping. It can be a single value or a row or column vector of values (one per reconstruction). If left empty, a percentile clipping may be performed.
- clip_max :(numeric array) – optional, default is empty. Maximal intensity values for clipping. It can be a single value or a row or column vector of values (one per reconstruction). If left empty, a percentile clipping may be performed.

- use_first_frame_reference: (boolean) – optional, default is true. True to use the computed absolute clipping values of the first reconstruction for the next ones, false otherwise.
- dataset: (string or char array) – optional, default is empty. Dataset to which the hologram (hol) belongs. It should be one of the following character arrays:
  - bcom8
  - bcom32
  - bcom32_bin
  - interfere
  - interfere_bin
  - interfere4
  - interfere4_bin
  - emergimg
  - emergimg_bin
  - wut_disp
  - wut_disp_on_axis
  - wut_disp_on_axis_bin
- cfg_file: (string or char array) – optional, default is empty. Path to configuration file. It should be a character vector.
- name_prefix: (string or char array) – optional, default is empty. Name prefix for the files written.
- outfolderpath: (string or char array) – optional, default is './figures'. Path to the output folder for figures.
- direction (string or char array) – optional, default is 'forward'. Propagation direction. It has effect only if usagemode = 'complex'. To reverse the propagation, the mathematical inverse of the propagation kernel is used. Note, that for some propagation kernels, this is not the same as applying the same propagation kernel with the negative object plane distance. It should take one of the following values:
  - 'forward': forward diffraction transform (propagation towards the object plane)
  - 'inverse': inverse diffraction transform (propagation towards the hologram plane)
- resize_fun: (string, char array or function handle) – optional, default is empty. Resize, clipping or down-sampling function handle to use on reconstructions. If 'DR' is provided, diffraction-limited reconstruction is performed using phase-space bandwidth limitation to reduce the resolution of reconstructed image (cf. document wg1m89038 [3]). If left empty, no resizing is performed.
- targetres: (numeric array) – optional, default is empty. Target resolution of the final video, when using resize_fun = 'DR'. No frame will have higher resolution. A single aperture size will be calculated for all frames. If left empty and resize_fun = 'DR', the diffraction-limited reconstruction will be based on the input aperture size ap_sizes.
- fps : (scalar number) – optional, default is 10. Frame rate of final video. It has effect only if usagemode = 'dynamic'.

4.1.3 Floating-point to integer conversion

Since not all anchor codecs operate at floating-point precision, the holographic content is mapped from floating-point representation to a 16-bit integer representation, before encoding. This process is inversed immediately after decoding.

The mapping is based on a uniform mid-rise quantizer to convert the floating point inputs to integer bit-depths. For any given distribution, a Lloyd max quantizer will asymptotically iterate towards the mapping that minimizes the mean-squared error (MSE). However, for sufficiently large bit-depths, the Lloyd-max quantizer will approach the uniform quantizer [6]. The dequantized output for the uniform mid-rise quantizer is given by:

$$\mathcal{Q}(x, L, X_{\max}) = \left\{ \begin{array}{ll} \left(\frac{-L}{2} + 0.5\right) \frac{2X_{\max}}{L} & \text{else if } x < -X_{\max} \\ \left(\left\lfloor \frac{xL}{2X_{\max}} \right\rfloor + 0.5\right) \frac{2X_{\max}}{L} & \text{else if } -X_{\max} \leq x \leq X_{\max} \\ \left(\frac{L}{2} - 0.5\right) \frac{2X_{\max}}{L} & \text{otherwise} \end{array} \right\}. \tag{8}$$

where $L = 2^{16}$, while $X_{\max}$ refers to the value that minimizes the MSE for the uniform quantizer. The choice of $X_{\max}$ represents a trade-off between the granular error (increases as $X_{\max}$ increases) and the overflow error (decreases as $X_{\max}$ increases till the largest value to be quantized) [16].

Since we use high bit-depths, the overflow error dominates the granular error and the value of $X_{\max}$ is almost always the largest floating point value to be quantized. However, for some cases we notice that $X_{\max}$ has a unimodal relationship with respect to its possible values, where $X_{\max}$ is lesser than the largest floating point value. Hence we also use the ”golden section search” numerical optimization technique to obtain another candidate of $X_{\max}$ [10], from which the final candidate is chosen.

Please note, that solely anchor codecs are subjected to this process. Proponent codecs shall be able to handle floating-point data at the input and output. The supported internal precision of the codecs under test is at the discretion of the proponents.

4.1.4 Anchor codecs

Three anchor codecs are selected for reference purposes: H.265/HEVC, JPEG 2000 and JBIG-2 As indicated in Table 5, they are not deployed in every setting.

Tab. 5: Anchor codecs and their employment during testing

Hologram type	**Higher precision**		**Binary**	
Anchor codec	**Hologram plane**	**Object plane**	**Hologram plane**	**Object plane**
H.265/HEVC	Yes	Yes	–	–
JPEG 2000	Yes*	Yes	–	–
JBIG-2	–	–	Yes	–

Note: currently VVC and JPEG XL are also being validated for serving as anchor codecs. At the 92^{nd} WG 1 meeting (July 2021) a decision will be taken with respect to this matter.

H.265/HEVC is configured in intra-frame mode. HM version 16.22 is being deployed in the experiment. The software can be downloaded from: `https://vcgit.hhi.fraunhofer.de/jct-vc/HM/-/releases/HM-16.22` and should be compiled for as 64bit binary after enabling internal 16bit representation by changing `#define RExt__HIGH_BIT_DEPTH_SUPPORT 0` to `1` in the file `sources/Lib/TLibCommon/TypeDef.h`.

The exact configuration files of the codec can be found below. The codec was called as

`TAppEncoder.exe -c HEVC_genConf.cfg -c HEVC_specConf.cfg` .

Listing 1: HEVC_genConf.cfg

```
#======== Profile definition ===============
Profile                         : monochrome16
Tier                            : main

#======== Unit definition ================
MaxCUWidth                      : 64
MaxCUHeight                     : 64
MaxPartitionDepth               : 4
QuadtreeTULog2MaxSize           : 5

QuadtreeTULog2MinSize           : 2

QuadtreeTUMaxDepthInter         : 5
QuadtreeTUMaxDepthIntra         : 5

#======== Coding Structure =============
IntraPeriod                     : 1
DecodingRefreshType             : 0
GOPSize                         : 1

#============ Misc. ============
InputColourSpaceConvert         : UNCHANGED
InputChromaFormat               : 400
InternalBitDepth                : 16
WaveFrontSynchro                : 1
SummaryVerboseness              : 1
```

Listing 2: HEVC_specConf.cfg

```
InputFile          : *YUV input filename*
InputBitDepth      : 16
BitstreamFile      : *Temporary bitstream filename*
ReconFile          : *YUV compressed + decoded output filename*
Level              : 8.5
QP                 : *qpn value*
SourceWidth        : *nr. of columns of input*
SourceHeight       : *nr. of rows of input*
FrameRate          : 1
```

```
FrameSkip          : 0
FramesToBeEncoded  : 1
```

JPEG 2000 can handle floating-point input, but is constraint to operate at 16-bit such as H.265/HEVC. The software can be downloaded from: `https://kakadusoftware.com/`. The exact configuration of the codec can be found below:

```
kdu_compress –i <image>.bmp –o tmp.jp2 –precise -no_weights Qstep=2e-16 .
```

JBIG-1 is included as anchor since it can handle binary image input. Hence, it will only be deployed during encoding of binary test data. The software can be downloaded from: `https://www.cl.cam.ac.uk/~mgk25/jbigkit/`. The exact configuration of the codec can be found below:

```
pbmtojbg <input-file.pbm> <output-file.jbg>
```

4.1.5 Coding conditions

For lossy coding the target bit rates for the experiments are provided in Table 6. Encoded bit streams should be at a bit rate which is within an error tolerance bound of maximum 5 % of the required bit rates in Table 6.

The bit rates for the colour holograms are three times as large as for the monochrome holograms since it is assumed the colour planes are encoded separately without exploiting potential correlations. Hence, every colour plane is attributed one third of the bit budget during encoding with the anchor encoders. Note also that both lossy and lossless – if supported – compression behaviour is tested for the floating point holograms. For the binary holograms, solely lossless compression is tested.

In addition, and if supported by the proponent's codec, also lossless coding results need to be provided for the holograms at 32 bit integer precision, 32-bit floating-point ...

Tab. 6: Target bit rates for the JPEG Pleno Holography objective test set.

Dataset	**Target Bitrate (bpp)**					
Monochrome holograms	0.1	0.25	0.5	1	2	4
Colour holograms	0.3	0.75	1.5	3	6	12

In addition to the bit rates mentioned in Table 6, holograms from the subjective test set should be evaluated **also** at the bit rates listed in Table 7 on a per hologram basis.

4.2 Pipeline for codecs under test

The proponents codecs will be evaluated in a similar fashion as the anchor codecs. The main difference is though that neither the floating-point to integer mapping and propagation modules are

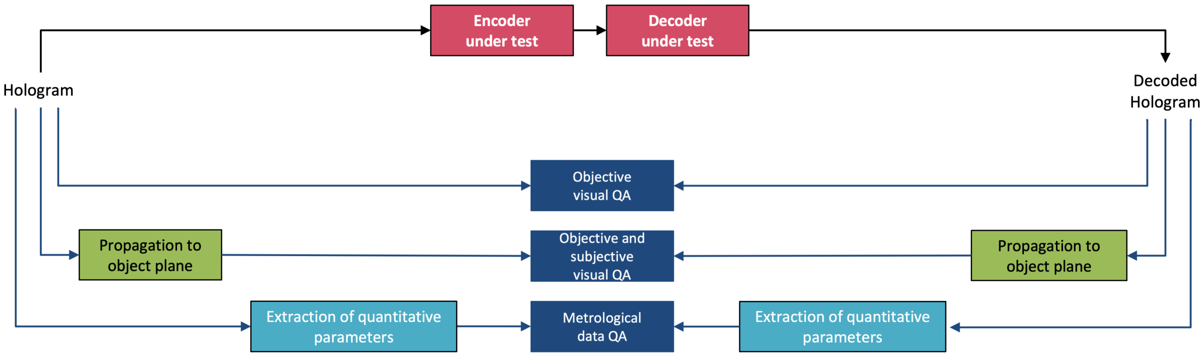

Fig. 2: Pipeline for the codec under test.

explicitly present in the testing pipeline. Note though that proponents might opt for including them in the codec's modules depending on the advocated type of coding architecture/solution.

Encoders will be evaluated at the target bit rates reported in Table 6 as the case for the anchor codecs.

4.3 Quality assessment

4.3.1 Objective visual quality assessment

Objective visual quality assessment is performed in both the hologram plane as in the object plane. For the latter numerical reconstruction software (NRSH Software 7.0) is applied to generate the reconstructions at specified reconstruction points – viewing angles and focus planes – as listed in Table 1.

Since not all introduced metrics are suitable for deployment in both planes or all types of holograms, Table 4 lists which metrics are utilized for testing of visual quality of holograms.

Because of computational complexity reasons measurements are currently only taken at various reconstruction positions and to the extend allowed by the characteristics of the holograms, namely center, left, top-center and top-left to account for different viewing angles and this in up to three depth planes. Also the average performance of these measurements will be calculated.

4.3.2 Subjective visual quality assessment

Due to the lack of availability of high-end holographic display and the costly nature of holographic printing for subjective testing purposes, subjective visual quality assessment will be performed on numerical reconstructions displayed on high-end 2D monitors.

Two test procedures are defined: (1) a **static subjective quality assessment procedure** where holograms are rendered for a particular reconstruction plane and viewing angle, allowing for stress-testing the codecs under test and (2) a **dynamic subjective quality assessment procedure** where the holograms are rendered as a pseudo-video to allow for better evaluation of the 3D features in the reconstructed holograms.

Note that in later phases of the standardization process, other types of subjective tests might be carried out providing additional evidence in support of the decision process.

To facilitate a sound subjective evaluation, the holograms will be reconstructed and displayed according to the procedure outlined below. Because of time constraints to run the experiments only a selection of the holograms in 3 bitrates, listed in Table 7 will be involved. The encoded material is the same as used during object evaluation in section 4.3.1.

Tab. 7: Selected holograms and target bit rates for the subjective test. For holograms 3 4, only 1 reconstruction plane is tested. In case of too many proponent codecs to be tested, holograms 6, 7 and 8, will be tested at less reconstruction planes. Bit rates for colour holograms indicates the sum of the bit rates for all colour channels.

	Hologram	**Selected bitrates(bpp)**
1	DeepCornellBox_16K	0.1, 0.25, 1
2	DeepDices2k	0.75, 1.5, 3
3	Lowiczanka Doll	0.3, 0.75, 3
4	Astronaut	0.1, 0.25, 0.5
5	DeepChess	0.1, 0.5, 2
6	Biplane16k (Interfere)	0.3, 3, 6
7	Dices16k	0.75, 1.5, 3
8	Piano16k	0.75, 1.5, 3

View Reconstruction Holograms will be reconstructed using the reference software (NRSH) V7.0 [15] and should be reconstructed for, and viewed on, a professional Eizo CG318-4K 2D monitor with 4K UHD resolution (4096×2160 pixels) and 10-bit colour depth, which is recommended for use in visual test laboratories [5]. The colour representation mode is set to ITU-R BT.709-6. The monitor is calibrated using the build-in sensor on the monitor, operated by the ColorNavigator-7 Color Management Software. The calibration is set for the sRGB Gamut, D65 white point, 120 cd/m^2 brightness, and minimum black level of 0.2 cd/m^2.

Moreover, holograms should be reconstructed such that the intrinsic resolution of the rendered scene matches the display resolution. To do so, a sliding synthetic aperture of the size 2048×2048 pixels is used to extract the views for reconstruction. For holograms reconstructed with larger apertures, cropping will be applied to fit the 2048×2048 pixel patch size. The exact cropping zone will be at the discretion of the test lab.

Image and Video Production For the *static test procedure*, reconstructed views for each reference and test hologram will be generated. The reconstruction positions are not communicated to the proponents and determined by the test lab(s).

In case of the *dynamic test procedure*, a pseudo video sequence will be generated from the reconstructed views for each reference and test hologram. The exact path used for sliding the aperture on holograms and extracting the video sequences will vary depending on the content type and will only be known to a limited set of people responsible for subjective testing and not involved as proponent in the process. The view paths are not revealed to proponents to avoid proposals being explicitly tailored according to content view paths.

Please note that for the dynamic test a similar procedure is followed as for the Call for Proposals on Light Field Coding [8] and the Call for Evidence on Point Cloud Coding [12], while of course accounting for the particular properties of the holographic modality.

Viewing Conditions Viewing conditions should follow ITU-R Recommendation BT.500.13 [BT50013]. In case of the dynamic test procedure, the MPV video player will be used for displaying the videos.

Displays used in the subjective testing should have anti-aliasing disabled.

Training Before Subjective Evaluation The test subjects are required to pass the Snellen visual acuity test and the Ishihara colourblindness test. Prior to subjective evaluation there will be a training period to acclimatize participants to the purpose of the experiment and their task.

This training period involves showing participants images or video sequences similar to the ones used in the test, guidance regarding presence of the speckle noise and the scoring protocol. **Test subjects will be instructed to ignore the speckle noise in their evaluation.** Participants are requested to score the perceived quality of the rendered hologram in relation to the uncompressed reference.

Subjective Evaluation Protocol The DSIS simultaneous test method will be used with a 5-level impairment scale, including a hidden reference for sanity checking. Both the reference and the degraded stimuli will be simultaneously shown to the observer, side-by-side, and every subject asked to rate the visual quality of the processed with respect to the reference stimulus. The reference will always shown on the same location.

Analysis of Results Outlier detection algorithm based on ITU-R Recommendation BT.500-13 [BT50013] should be applied to the collected scores, and the ratings of the identified outliers will be discarded. The scores are then averaged to compute mean opinion scores (MOS) and 95% Confidence Intervals (CIs) computed assuming a Student's t-distribution.

4.3.3 Metrological quality assessment

The exact specification of the metrological quality assessment is currently subject of an exploration study. For QPI (Quantitative Phase Imaging), the acceptable values of RMSE error on the retrieved wrapped phase is around 0.05 rads [14] and serves as the region of interest shown in Table 8. The Matlab code (including the objective evaluation pipeline with anchor codecs) is provided in wg1n91009. The code is also available at `https://gitlab.com/wg1/jpeg-pleno-holo-ctc-metrology` (Note: access can be obtained by contacting `Peter.Schelkens@vub.be`).

Tab. 8: Target bit rates for the JPEG Pleno Holography test set

Dataset	**Target Bitrate (bpp)**					
Microscopy holograms	0.1	0.25	0.5	1	2	4

References

[1] G. Bjontegaard. Calculation of average PSNR differences between RD-curves. *ITU-VCEG*, 2001.

[2] 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*, 70:114–130, 2019.

[3] David Blinder and Tobias Birnbaum. Minimizing the required viewing resolution of reconstructed holograms with phase space models, September 2020. WG1N89038, 89th JPEG Meeting, Online (Geneva, Switzerland).

[4] David Blinder, Heidi Ottevaere, Adrian Munteanu, and Peter Schelkens. Efficient multiscale phase unwrapping methodology with modulo wavelet transform. *Optics Express*, 24:23094–23108, 10 2016.

[5] EIZO. ColorEdge CG318-4K. `https://www.eizoglobal.com/products/coloredge/cg318-4k/index.html`, 2019 (accessed Feb 15, 2019).

[6] H. Gish and J. Pierce. Asymptotically efficient quantizing. *IEEE Transactions on Information Theory*, 14(5):676–683, 1968.

[7] Miguel Arevallilo Herráez, David R. Burton, Michael J. Lalor, and Munther A. Gdeisat. Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path. *Appl. Opt.*, 41(35):7437–7444, Dec 2002.

[8] ISO/IEC JTC1/SC29/WG1. JPEG Pleno Call for Proposals on Light Field Coding, January 2017. WG1N74014, 74th JPEG Meeting, Geneva, Switzerland.

[9] Małgorzata Kujawińska, Michał Ziemczonok, Arkadiusz Kuś, and Wojciech Krauze. Metrological studies of limited angle holographic tomography systems based on a phase phantom mimicking biological cell. In *Digital Holography and Three-Dimensional Imaging 2019*, page Th2B.2. Optical Society of America, 2019.

[10] Raees Kizhakkumkara Muhamad, David Blinder, Athanasia Symeonidou, Tobias Birnbaum, Osamu Watanabe, Colas Schretter, and Peter Schelkens. Exact global motion compensation for holographic video compression. *Appl. Opt.*, 58(34):G204–G217, Dec 2019.

[11] Jae-Hyeung Park. Recent progress in computer-generated holography for three-dimensional scenes. *Journal of Information Display*, 18(1):1–12, January 2017.

[12] Perry, Stuart. JPEG Pleno Point Cloud Coding Common Test Conditions v3.2, April 2020. WG1N87037, 87th JPEG Meeting, Online (Erlangen, Germany).

[13] H. R. Sheikh and A. C. Bovik. Image information and visual quality. *IEEE Transactions on Image Processing*, 15(2):430–444, 2006.

[14] P. Stępień, R. K. Muhamad, D. Blinder, P. Schelkens, and M. Kujawińska. Spatial bandwidth-optimized compression of image plane off-axis holograms with image and video codecs. *Opt. Express*, 28(19):27873–27892, Sep 2020.

[15] Antonin Gilles Raees Kizhakkumkara Muhamad Cristian Perra Tobias Birnbaum, Roberto Corda and Peter Schelkens. Numerical Reconstruction Software for Holography (NRSH) 7.0, July 2021. WG1N92032, 92th JPEG Meeting, Online.

[16] Yuli You. *Audio Coding:Theory and Applications*. Springer US, 2010.

[17] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.