Suitability analysis of holographic vs light field and 2D displays for subjective quality assessment of Fourier holograms

Ahar, Ayyoub; Chlipala, Maksymilian; Birnbaum, Tobias; Zaperty, Weronika; Symeonidou, Athanasia; Kozacki, Tomasz; Kujawinska, Malgorzata; Schelkens, Peter

Published in:
Optics Express

DOI:
10.1364/OE.405984

Publication date:
2020

License:
Other

Document Version:
Final published version

Citation for published version (APA):
Ahar, A., Chlipala, M., Birnbaum, T., Zaperty, W., Symeonidou, A., Kozacki, T., ... Schelkens, P. (2020). Suitability analysis of holographic vs light field and 2D displays for subjective quality assessment of Fourier holograms. Optics Express, 28(24), 37069-37091. [24]. https://doi.org/10.1364/OE.405984
Suitability analysis of holographic vs light field and 2D displays for subjective quality assessment of Fourier holograms

AYYOUB AHAR,1,2,* MAKSYMILIAN CHLIPALA,3 TOBIAS BIRNBAUM,1,2 WERONIKA ZAPERTY,3 ATHANASIA SYMEONIDOU,1,2 TOMASZ KOZACKI,3 MALGORZATA KUJAWINSKA,3 AND PETER SHELKENS1,2

1Dept. of Electronics and Informatics (ETRO), Vrije Universiteit Brussel (VUB), Pleinlaan 2, B-1050, Brussels, Belgium
2Imec, Kapeldreef 75, B-3001, Leuven, Belgium
3Warsaw University of Technology, Institute of Micromechanics and Photonics, 8 Sw. A. Boboli St., 02-525 Warsaw, Poland
*ayyoub.ahar@vub.be

Abstract: Measuring the impact of compression on the reconstruction quality of holograms remains a challenge. A public subjectively-annotated holographic data set that allows for testing the performance of compression techniques and quality metrics is presented, in addition to a subjective visual quality assessment methodology. Moreover, the performance of the quality assessment procedures is compared for holographic, regular 2D and light field displays. For these experiments, a double-stimulus, multi-perspective, multi-depth testing methodology was designed and implemented. Analysis of the quality scores indicated that in the absence of a suitable holographic display and under the presented test conditions, non-holographic displays can be deployed to display numerically reconstructed holograms for visual quality assessment tasks.

© 2020 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Digital holography, in theory, is considered to be the holy grail of 3D imaging solutions [1]. While the concept and its initial realizations have been around for almost half a century, they have been re-investigated for the purpose of 3D visualization only recently. This is due to steady growth of available computational power and significant improvements in nano-electronics, optical hardware, and photonics technologies. However, there are hardware and signal processing challenges yet to be addressed in order to facilitate an immersive 3D experience via a complete pipeline for high-quality dynamic holography with full-parallax and wide field of view (FoV) [2].

In this regard, one of the core challenges is modelling the perceived visual quality of the rendered holograms, which has a vital impact on steering the other components of the holographic imaging pipeline. Since holograms have a totally different frequency content than photographic imagery, are potentially complex-numbered representing interferometric content and can be huge for realistic 3D scenes, holographic compression methods are totally different from natural image codecs [2,3]. Hence, also rate-distortion optimization is based on different criteria than classical codecs would require and appropriate distortion measurement and assessment has to be developed. Also while the design of highly efficient numerical methods in Computer-Generated Holography (CGH) [4–9] and efficient encoders for holographic content [2,3,10–12] is gaining momentum, visual quality assessment of holograms has a rather long way to reach its primary milestones due to various open problems along the way [2,3]. Indeed, conducting a systematic subjective test is the very first step. But holographic modalities and in extension, plenoptic
modalities, pose specific challenges as it concerns evaluating their visual quality [2,13,14]. The main challenges to conduct a reliable holographic subjective experiment are threefold:

First, creating a large-enough database is required. It should contain a diverse set of reference holograms as well as their distorted versions which have been contaminated by multiple types of distortions with varying levels of magnitude. Few earlier efforts took place, proposing open-access reference-only test data; such as: the B-Com Repository [15,16], ERC Interfere I [17], II [18] and III [19], all of them contain CGHs acquired from 3D point-clouds or 2D+depth data. Also EmergImg-HoloGrail v1 and v2 [12] provide a limited set of Optically Recorded Holograms (ORH).

Second, there is no widely-accepted standard methodology for visual quality assessment of plenoptic content and especially for holographic data. This is due to the fact that the plenoptic function "allows for the reconstruction of every possible view, at every moment, from every position, at every wavelength within the space-time wavelength region under consideration" [20]. As a consequence, to meet its time and resource constraints, a subjective experiment has to be limited to evaluate only a necessary subset of this 7D space. However, the selection of such a subset and the testing methodology in general is directly related to the chosen technology for displaying plenoptic content. For instance, several subjective quality assessment approaches have been reported for 4D light fields. At IEEE ICME 2016 a Grand Challenge on Light Field Compression was organized [21] deploying a double stimulus continuous quality scale (DSCQS) methodology [22]. In this solution uncompressed and decoded views are shown side-by-side on a high-end monitor while selecting a limited set of views and focus points per light field. In the context of the JPEG Pleno Light Field Coding standardization effort and an associated Grand Challenge on Light Field Coding organized at IEEE ICIP 2017, a double stimulus comparison scale (DSCS) methodology was employed with side-by-side rendering of the light field as a pseudo video sequence and using a discrete quality scale ranging from -3 to 3 [23,24]. Viola et al. [25] assessed the impact of an interactive approach to determine perceived quality thereby enabling the evaluation of a larger fraction of the light field compared to the passive approach [24]. The same authors applied the advocated solution also to evaluate a larger set of light field compression techniques [26]. In the case of point cloud data only very few subjective quality assessment experiments have been conducted. In [27], two subjective testing methodology namely Absolute Category Rating (ACR) and Double Stimulus impairment Scale (DSIS) [22] were compared for subjective quality assessment of point clouds degraded by gaussian noise and Octree-pruning. To evaluate point cloud compression techniques, Javaheri et al. [28] rendered a video sequence by having a virtual camera spiralling around the point cloud object. The DSIS methodology was adopted and the pseudo-video sequences of the impaired and original point cloud were shown sequentially. A similar method is advocated by the JPEG Pleno initiative for the evaluation of compression methods for point clouds [29]. In the case of digital holography, earlier in [30] and then recently in [31], methodologies to evaluate perceptual quality of compressed holograms with partial-reconstructions shown on a 2D display were proposed. Also here JPEG published a JPEG Pleno Common Test Conditions document for holographic content coding [32], which was partially inspired by the results presented in this paper.

Third, holographic displays with acceptable visual characteristics are still rare [33,34]. Moreover, configuring and operating such displays requires advanced technical skills. One alternative, is to record hard-copy holograms of the scenes and then illuminate them for the subjects to rate their visual quality. The advantage of this method is that very high definition holograms of deep scenes can be printed to illustrate the full potential of the holograms for recreation of a 3D scene or object [35–40]. However, using hard-copy holograms for subjective experiments is facing serious challenges. Those holograms can not be evaluated by objective quality metrics. From practical point of view, it requires flawless printing of a large number of holograms which normally will be several times bigger than the final hologram set utilized for
the subjective test. This is due to the fact that preliminary experiments and mock-up tests are required to shortlist and choose the right scenes and objects, depth of the scenes and distortion levels to span over the full visual quality range. Furthermore, it would severely prolong the testing procedure, as the holograms need to be permuted through and optically aligned multiple times per test subject. Finally, if quality evaluation of digital holographic displays in any form is to be investigated then naturally, experimenting with printed holograms is not helpful.

Researchers have been rendering numerically reconstructed holograms on non-holographic displays, including regular 2D displays or more recently multi-view light field displays, to alleviate this problem [19,30,41]. However, potential perceptual differences between visualization on holographic displays and numerical reconstructions rendered on non-holographic displays have not been investigated thoroughly before to the best of our knowledge. Some of the most evident issues include loss of visual cues related to the depth perception on 2D displays and light field displays, FoV, and appearance of different types of speckle noise. These issues are inter-connected with the display chosen for visualizing the holographic content. For example, for a 2D display, only a specific focus plane and perspective of the hologram can be rendered. For a light field display each view of the hologram needs to be reconstructed with a particular focus plane utilizing a suitable aperture. As such, only a section of the 3D scene volume, contained in the hologram, is rendered properly. Nevertheless, both display types support high spatial resolutions and large display sizes. In contrast, holographic displays can render the complete plenoptic scene, but their resolution and overall size are currently limited, which in practice results in a tiny viewing window (VW) to explore the visualized hologram. These fundamentally diverse properties require different strategies and procedures per display type to conduct a subjective test.

Since all of these three main problems, concerning the subjective quality assessment of holograms, are connected, in this research we put forward an ambitious effort to address all three problems together for the first time. We report here the creation of the first comprehensive visually annotated database of CGH and ORH. We utilize the created database to conduct three separate subjective experiments on a holographic display, a light field display, and a 2D display. For each subjective test we utilize and report a tailored methodology to meet the limitations and requirements for each tested display and then finally, we provide a suitability analysis of using non-holographic displays for visualization of digital holograms backed up by statistical analysis of the results from our subjective experiment. Figure 1 represents a high-level representation of the experimental pipeline used in this research. The full set of the test holograms along with their numerical reconstructions, acquired quality scores and other related data to these experiment are publicly available (Ref. [42]).

The main contributions of this manuscript include:

1. Creation of the first publicly available database of optically recorded and computer-generated holograms annotated with subjective test results.

2. Design and implementation of a test methodology for subjective quality evaluation of holograms.

3. Comparison of holographic versus non-holographic displays based on the visual appearance of same set of holograms.

4. Evaluation of computer-generated against optically recorded Fourier holograms.

The details about the numerical and optical methods to produce the holograms used in this experiment as well as the content preparation are provided in section 2. In section 3, the details of the tested holographic, light field and regular 2D display are described. Section 4 introduces the subjective test methodology including its details for each setup and training of the test subjects. In section 5, we explain the statistical post-processing of the experimental results and provide
Fig. 1. Overview of the experimental pipeline.

2. Capturing and generation of test data

For a successful subjective experiment the test data should provide sufficient diversity in terms of the features of the represented 3D scenes, production of the holograms, and the distortions introduced. Having a set of holograms obtained from a diverse set of objects is vital to avoid any bias in the results. This is particularly important in holography since each hologram is an interferogram. Therefore, characteristics of the recorded scene (e.g., object positions, their distances to the recording plane, surface properties, and occlusions) will affect the entire interference footprint on the recorded hologram.

On top of the scene characteristics, the method which produces the hologram must be taken into account. Holograms can either be optically recorded or numerically computed. The latter category, referred to as CGHs, can nowadays be calculated at very high spatial resolutions allowing for high SBPs, efficient occlusion handling, and bidirectional reflectance distribution functions (BRDFs). However, photo-realistic quality is difficult to achieve. Therefore, it is also important to include ORHs in the test data. Moreover, ORHs have different characteristics such as the presence of measurement noise.

The holograms of the real objects were recorded using a lensless Fourier holographic capture system [43], described in section 2.1. The holograms of the synthetic objects, represented as point clouds, are generated with a multiple Wavefront Recording Plane (WRP) method [4] shortly discussed in section 2.2.

All considered holograms are created with respect to a spherical reference wave with focal point in the scene center. In the Fourier holographic capture system a spherical reference point source is placed at the center of the object plane. In the CGH calculation framework a demodulation with a Fresnel approximated spherical phase factor is performed numerically after initial propagation to the hologram plane. This way, the advantage of the lensless Fourier holographic capture system in terms of best SBP usage [44] is beneficial in both scenarios and furthermore the hologram pixel count does not limit the maximal object dimensions but instead the maximal angular FoV. ORHs and CGHs are obtained for a high resolution of 16384×2048 pixel. An exception is the OR-Squirrel which was obtained with the resolution of 27904×1792 pixel. Table 1 summarizes the analysis as well as a discussion of the outcomes. Finally, section 6 presents the concluding remarks.
the important characteristics of the holograms produced for this experiment. Figure 2 shows the center view of numerical reconstructions for these holograms.

![Holograms](image)

**Fig. 2.** Center view numerical reconstructions for the CGHs (top) and ORHs (bottom).

**Table 1.** Characteristics of the objects geometry utilized to generate and reconstruct the holograms. (PC=Point cloud; WRP=Wavefront recording plane)

| Hologram   | Acquisition Method | PC density/Material | Number WRP | Recording Dist.(mm) | Obj. Size: W×H×D(mm) | Rec. Dist./Depth | Angular FoV (deg) |
|------------|--------------------|---------------------|-------------|---------------------|-----------------------|------------------|------------------|
| OR-Mermaid | ORH                | Polished Metal      | -           | 450                 | 27 × 53 × 5           | 90.00            | 8.8              |
| OR-Ball    | ORH                | 3D-Print            | -           | 960                 | 65 × 65 × 65          | 14.76            | 8.8              |
| OR-Squirrel| ORH                | Brushed Metal       | -           | 500                 | 43 × 85 × 70          | 7.14             | 10.5             |
| OR-Wolf    | ORH                | Plastic             | -           | 780                 | 50 × 60 × 80          | 9.12             | 8.8              |
| CG-Ball    | CGH                | 1.313.280           | 101         | 700                 | 50 × 50 × 50          | 14.00            | 8.8              |
| CG-Chess   | CGH                | 219.100             | 200         | 491                 | 38 × 43 × 310         | 1.58             | 8.8              |
| CG-Earth   | CGH                | 306.372             | 101         | 706                 | 46 × 46 × 46          | 15.34            | 8.8              |
| CG-Plane   | CGH                | 9.999.079           | 200         | 716                 | 53 × 46 × 71          | 10.08            | 8.8              |

2.1. **Optical acquisition**

Four ORHs from objects of various dimensions, surface characteristics, and capture distances, were selected, which are shown in Fig. 1. The first object is a ”Mermaid” figurine with small depth, while the second object is a ”Squirrel” figurine with larger depth. Both objects are characterized by a glossy, metallic surface. The third and fourth object, ”Wolf”, a rubber toy, and ”Sphere”, the 3D printed model based on the input content of the CGH ”Ball”, respectively have diffuse surfaces and rather large depths.

For optical acquisition a lensless, Fourier synthetic aperture holographic capture system [43,45] is employed (Fig. 3). In this system, the laser beam is divided into a reference and an object beam by a polarizing beam splitting cube PBS. The intensity ratio of both beams is adjusted with an achromatic half-wave plate λ/2, to obtain a high contrast of the interference
fringes. The reference beam is formed and directed by the following set of elements: a pinhole \( PH \), an achromatic collimating lens \( C(F_C = 300 \text{ mm}, NA_C = 0.13) \), and mirrors \( M_1 \) and \( M_2 \). The reference point source \( S \) is generated at the object plane by an achromatic objective \( L(F_L = 60 \text{ mm}, NA_L = 0.21) \). The lenses \( C \) and \( L \) are selected such that the generated spherical reference wave covers the entire area of the synthetic aperture hologram capture. For the object beam, the diffusers \( D_1 \) and \( D_2 \) create a double-sided illumination with the help of the mirrors \( M_3, M_4, M_5 \) and another beam splitting cube \( BS \). The analyzer \( A \), placed in front of the camera, improves the hologram contrast by filtering out non-interfering light. The hologram is recorded by a charge-coupled device (CCD) camera (Basler piA2400-12gm) with a pixel pitch of 3.45 \( \mu m \) and a resolution of 2448\times2050 pixel. To realize the synthetic aperture, the CCD is translated in the horizontal direction with the use of motorized linear stage with steps of 2.85 mm over a range of 60 mm. This results in an overlap of 60\% between adjacent captured sub-holograms and enables data stitching with sub-pixel precision using a correlation-based routine [46]. The obtained off-axis, synthetic aperture lensless Fourier hologram is composed of 20 sub-holograms that have a physical size of approximately 56.5 \times 7.1 \text{ mm}, which corresponds to 16384\times2048 pixel. The maximum scene size is reduced by half in the horizontal direction, due to the presence of a twin image in the hologram. Holograms were recorded at distances \( R_m \) adapted to the scene size and with either 532 nm or 632.8 nm laser beam having wavelengths \( \lambda_n \): “Mermaid”, \( R_1 = 450 \text{ mm}, \lambda_1 = 532 \text{ nm}; “Squirrel”, \( R_2 = 500 \text{ mm}, \lambda_2 = 632.8 \text{ nm}; “Wolf”, \( R_3 = 780 \text{ mm}, \lambda_3 = 532 \text{ nm}; and "Sphere", \( R_4 = 960 \text{ mm} \) with \( \lambda_4 = 532 \text{ nm}. For holograms captured with the green and red light sources, the angular FoV equals 8.8° and 10.5°, respectively.

![Fig. 3. Lensless Fourier synthetic aperture holographic capture system.](image)

### 2.2. Computer-generated holograms

The CGHs, used for the subjective tests, were generated from point clouds with an extension of the WRP method [4]. This method employs multiple parallel wavefront recording planes and pre-computed look-up tables. Moreover, it includes an occlusion handling technique. As shown in Fig. 4, the contribution of each point – starting from the point furthest away to the hologram plane to the closest point – is added to the respectively closest WRP. When all the points that belong to the current WRP are accounted for, the wavefield is propagated to the next WRP and so forth. To simulate diffuse reflection, we assign a random phase to the point spread function during the calculation of the LUTs, as presented in [18]. However, there is a very important difference, compared to the previously published methods. To exploit the SBP advantage of the Fourier holographic approach, the wavefield at the last WRP plane is converted to comply with the Fourier hologram configuration, contrary to the in-plane configuration that it supported before. This is done in two steps. First, hologram is propagated to its proper viewing distance using the angular spectrum method [47,48] and subsequently demodulated with a quadratic Fresnel phase kernel corresponding to the axial distance between the hologram and the last WRP.
The second step approximates a spherical wavefront with focus in the center-plane of the object by using the Fresnel approximation. The four CGHs are full parallax and have the same setup parameters: the pixel pitch is 3.45 µm, the wavelength of the reference beam is 532 nm and the scene center plane were located 700 mm from the hologram plane.

2.3. Content preparation

Finally, to facilitate the subjective test and to examine the suitability of each used display, the produced holograms have to be processed such that they are available at different visual quality levels. This enables testing of the sensitivity of each display for quality degradation of the holographic content. An added complexity here is the selection of suitable distortion types. Nonetheless, compression artifacts are a good starting point for a holographic dataset. From a practical point of view, the performance of the available compression methods can be compared. Also, compression artifacts normally stem from a combination of multiple distortion types as a result of different processes undergone inside the encoders. Though other distortions used in visual quality testing could be considered, it is important to realize that the end-user will observe the reconstructed hologram in the object plane and not in the hologram plane. During the reconstruction or back-propagation process the propagated data from each point on the fringe pattern updates each and every point of the reconstructed scene. Hence, the reconstructed scene is particularly resilient to local artifacts or even complete loss of information in some small regions of the hologram. As an example, salt and pepper noise, which in regular imaging significantly degrades the visual quality, almost completely vanishes after reconstruction of the hologram. Therefore, in this experiment we constrained the distortions to those that have a more global impact on the hologram, in particular compression distortions. We employed three coding engines each of which was used to encode real and imaginary parts of the holograms separately: JPEG 2000 [49,50], intra h.265/HEVC [51] and wave atom coding (WAC) [52].

All three encoders compressed the 8-bit quantized real and imaginary parts of each hologram separately. The holograms were compressed at bitrates 0.25 bpp, 0.5 bpp, 0.75 bpp and 1.5 bpp. These bitrates were determined via a series of mock-up tests in which the holograms were compressed at 9 different bit-depths ranging from 0.15 bpp to 4 bpp and their visual appearance was tested on 2D and light field setups. Also, on the holographic display, 3 sample holograms were tested for all 9 bit-depths under consideration. The remaining holograms were only verified at the chosen bit depths. The goal was to ensure that the distortion levels resulted in a broad range of visual quality levels ranging from very poor to imperceptible.
The IRIS-JP3D software package was deployed to implement the JPEG 2000 compression [53]. The default configuration for JPEG 2000 was utilized using a 4-level Mallat decomposition and CDF 9/7 wavelets with 64×64 pixel sized code blocks.

For the implementation of the h.265/HEVC compression standard, revision HM-16.18 [54] was used. Since all tested images were in grayscale format, we used 4 : 0 : 0 subsampling and inserted the images as an 8-bit luminance channel with empty chrominance channels. Hence, cross-component prediction and motion search settings were disabled. "Frame rate" and "frames to be encoded" were set to 1. The desired compression level was achieved by tuning the quantization parameter (QP). All other parameters were set to their default values.

The WAC leverages the orthonormal wave atom transform. This non-adaptive multi-resolution transform has good space-frequency localization and its orthonormal basis is suitable for sparsifying holographic signals. WAC is based on a JPEG 2000 coding architecture where the CDF 9/7 wavelet transform is replaced by the 2D wave atom transform for which the spatial footprint of each atom scales parabolically across resolutions, while the quantization and Embedded Block Coding by Optimized Truncation (EBCOT) [55] are further deployed. EBCOT code blocks of size 128×128 pixel are used.

3. Display systems

The test holograms presented in the previous section will be subjectively tested on three types of displays: an holographic, a regular 2D and a light field display, which are described hereafter. Please note, that the content described above has been produced such that it meets the specifications of the holographic display; the specifications of the other display types are met through the numerical reconstruction procedure. Also, all holograms were displayed using a reference beam with wavelength of 532 nm, due to the fact that human eye is more sensitive to the green zone of the light spectrum.

3.1. Holographic display

In this work, a Fourier holographic display with an incoherent LED source is employed. It is placed on an optical table. The system provides high-quality orthoscopic reconstructions of large and deep objects [56], which can be viewed with a naked eye. The display setup is presented in the Fig. 5. In this system, a phase-only spatial light modulator (SLM) (Holoeye 1080P, 1920×1080 pixel, pixel pitch 8 μm) is illuminated by a normal plane wave, which is formed by an LED source (Doric Lenses, center wavelength λG = 515 nm and fiber core of 960 μm) and a collimating lens L_C (f_C = 400 mm). The SLM displays the object wave with removed spherical phase factor. Next, the reflected beam passes through the imaging module, which introduces a magnification and facilitates the complex wave coding. The first imaging element is realized by a 4F afocal imaging system composed of the lenses L_1 (f_1 = 100 mm) and L_2 (f_2 = 600 mm) with magnification ratio M = −6. The 4F system and the field lens L_f conjugate the SLM plane with a 3D hologram reconstruction volume focused at the VW. The complex coding scheme is experimentally supported with the absorbing cut-off filter in the Fourier plane of the 4F system [57].

In this experiment holographic setup is covered using black barriers such that no environmental light would enter the black box, i.e. the display setup. A small slit was carved into the box and a metal chinrest and forehead holder were put in front of the slit such that all subjects could easily observe the displayed holograms as soon as they would position their eye at the VW. Lensless Fourier holography enabled the reconstruction of a 1:1 orthoscopic copy of the 3D object with no visible distortions for the holographic display described above. This is achieved with the help of Confocal Fresnel Propagation methods to adjust focus mismatch [56]. During the visual tests the center plane of all holograms was 100 mm away from the image of the SLM, which is about 700 mm away from the observer.
For this distance the full object is viewed by the naked eye with a maximum size of 107 mm. With the available Space Bandwidth Product (SBP) [58,59] of the SLM this results in an angular FoV as stated in Table 1 and an angular resolution of display which is comparable to the resolution of the human eye for dark observation conditions. Theoretically, our holographic display can reconstruct details with the angular resolution of 0.3 arc minutes, while the human’s eye can resolve 1 arc minute. The 2D and light field displays discussed below are based on 2D reconstructions of a single or multiple views, respectively. Imaging on both displays benefits from the convention of Fourier holography as well since it fully uses the SBP and thereby achieves the highest quality during the recording/generation process.

3.2. 2D display
The issued 2D display is a professional Eizo CG318-4K monitor with 4K UHD resolution (4096×2160 pixel) and 10-bit color depth, which is recommended for use in visual test laboratories [60]. The color representation mode was set to ITU-R BT.709-6. The monitor was calibrated using the build-in sensor on the monitor, operated by the ColorNavigator-7 Color Management Software. The calibration was done according to the following profile: sRGB Gamut, D65 white point, 120 cd/m² brightness, and minimum black level of 0.2 cd/m². On this display, numerical reconstructions of the object were rendered for a chosen reconstruction plane and perspective.

3.3. Light field display
The light field display system is a HoloVizio-722RC by Holografika [61]. This 72 inch display has a horizontal angular FoV of 70° with a total 3D resolution of 73 Mpixel. It provides a 2D equivalent resolution of 1280×768 pixel for each of the 72 views. It features a 24 bit RGB color system with a brightness of ≈1000 cd/m². For each fixed chosen reconstruction plane, the holograms were rendered on this display by calculating a set of numerical reconstructions at the particular distances for all views supported by the display.

4. Test methodology
4.1. Generic procedure for subjective quality assessment
As mentioned in section 1, the test method should be adapted to the specific limitations and different technical requirements of each display type. The holograms were shown largely following the procedure for DSIS. In our method the reference and distorted stimuli are sequentially shown to the subject and then the subject scores the second stimulus (impaired version) based on the first (reference). As it will be explained in section 4.4, only in the case of the 2D display, which required separate renders per focal distance and perspective, the reference and the impaired
version were shown side by side to reduce the test duration. The hologram sequences were shown in a fully randomized order. The presentation order was also randomized for each subject. Also, in each setup, every subject scored only a subset of the total set of hologram pairs. Nonetheless, this was arranged such that every test condition received scores from 20 test subjects.

The scoring procedure for all three setups followed the standard DSIS scoring protocol, providing 5 quality scales. Depending on the perceived mismatch, the subject chooses a quality number from 1 to 5 representing one of the impairment scales: Very Annoying, Annoying, Slightly Annoying, Perceptible but not Annoying, and Imperceptible. The test-lab conditions corresponded to ITU-R BT.500-13 recommendations [22] and recommendations described in Annex B of ISO/IEC 29170-2 (AIC Part-2). The subjective test on the holographic display was conducted in the laboratory of Warsaw University of Technology while the other two tests were conducted in the Visual Quality Test lab of Vrije Universiteit Brussel. It should be noted that the utilized methodologies for each display setup are designed to adhere to the main objectives of current study (i.e. to facilitate maximum comparability between the gathered MOS values for the holographic and non-holographic displays, providing detailed ground-truth data for objective quality assessment and the investigation of the compression artifacts). Although, if only one of these displays are utilized for subjective experimentation, the test method can be further tuned to better account for their particular capabilities which are beyond the scope of this study.

4.2. Subjective quality assessment on holographic display

Here, the holograms were shown following the DSIS procedure. From each hologram a synthetic aperture was used to extract sub-holograms of 2048×2048 pixel, for visualization and scoring the center and right-corner views. Captured images of the exemplary reconstructed holograms in the holographic display are shown below. Note that these images are for the demonstration purposes and may not present exact details as it would appear to the eyes of the human subjects. In this particular case, the room lighting condition did not meet the standard ITU-R BT.500-13 recommendations. However, the environmental luminance does not impact the visibility of the stimuli because the subject has to position his eye on the watching-slit, as explained in section 3.1. Since subjects watch the stimuli with one eye only (due to the limited size of the screen), we preferred to use a dark room to guarantee the repeatability of the experiment. This way subjects keep the other eye open without being affected by the environmental light. During the mock-up test session we realized that subjects felt fatigued significantly sooner compared to non-holographic displays due to the stronger effort required to focus the eye on the content. Hence, the test was divided into 4 sessions of at most 10 minutes each. The experiment was performed in two days such that each subject participated in only 2 sessions per day. A compulsory minimum of 5 minute rest was facilitated by the test operator before starting the next session. The maximum rest time was not limited and subjects had the freedom to take larger recuperation periods in case they felt it to be necessary.

4.3. Subjective quality assessment on light field display

For the light field display, again the DSIS method was implemented following the ITU-BT.500-13 recommendations [22]. Although the display provides a wide angle, simultaneously rendering multiple views. Thus each subject was required to observe and score only the center and right-corner view of each displayed hologram-pair. To facilitate a repeatable procedure, the positions, where subjects had to stand to see the required views, were marked on the floor. The distance from the screen was chosen 3.2 times the height of the screen. For each tested hologram, the subject starts standing in the center-position and the operator displayed the reference and impaired holograms sequentially. After recording the score, the subject moved to the right-corner position and again both reference and impaired holograms were displayed by the operator followed by the scoring. According to Table 2 the number of test-conditions per subject was twice the
number of test-conditions per subject in the holographic display test. This is due to the fact that for each hologram, test subjects scored the visual quality at two different reconstructions (focal) distances for the light field display. The test in this setup was conducted in 2 sessions with a target duration of 20 minutes. The subjects were required to stand and move multiple times to designated positions during the test. Therefore, 1 hour rest was recommended before starting the second session.

Table 2. Details of the test conditions and the gathered scores per display setup, (* full depth of the scene is provided at once, the subject refocuses his eye on different parts of the scene. † the average number of reconstructions per hologram. See section 5.1 for description).

| Display          | Tested Holograms | Distortions | Distortion Levels | Perspectives | Focal Distances | Total Conditions | Scores Per Condition | Tested Subjects | Lab     |
|------------------|------------------|-------------|-------------------|--------------|----------------|-------------------|----------------------|-----------------|---------|
| Holographic      | 8                | 3           | 4                 | 2            | 2*             | 192               | 20                   | 40              | WUT     |
| Light field      | 8                | 3           | 4                 | 2            | 2†             | 384               | 20                   | 40              | VUB     |
| Regular 2D       | 8                | 3           | 4                 | 2            | 2†             | 384               | 20                   | 40              | VUB     |

4.4. Subjective quality assessment on 2D display

For the 2D setup, see Fig. 6 (a), each reconstructed hologram was shown for the 2 perspective positions corresponding to the ones used for the light field display and the holographic display. Per perspective two reconstruction (focal) distances corresponding to the light field display test were used. In order to stay within the recommended test durations the reconstructed reference and the impaired holograms were displayed side by side, reducing the test time per subject by half. Figure 7 shows captured images of the exemplary reconstructed holograms in the holographic display.

4.5. Training of test subjects

For each setup, 40 subjects participated. 22 subjects participated in the tests on both non-holographic displays. The remaining subjects differed between display setups. As the test for the holographic display was conducted in a different country, its 40 participants could not be tested on the same non-holographic displays. Nonetheless, Mean Opinion Score (MOS) results are utilized precisely for the purpose of avoiding dependencies to the opinions of specific persons. From the total of 120 participants, the number of female and male participants were 54 and 66 respectively. Their age ranged between 18 to 30 years. Prior to the test, subjects were required to pass the Snellen visual acuity test. Though, all the content shown to the subjects were monochromatic,
Fig. 7. Captured images from the center view of CG-Ball (top) and OR-Sphere (bottom) as they appear in the holographic display. Three compressed versions from each hologram are also shown, all compressed at 0.25 bpp for completeness the Ishihara test for the detection of colourblindness was performed as well. Prior to the first test session in each setup, a 5 minute training session was conducted where the test and the scoring procedures were explained and rehearsed.

4.6. Reliability analysis of the obtained MOS

First, a reliability analysis is performed on the acquired opinion scores for the three setups. Before performing any post-processing on the scores and calculating the MOS for each test condition (see procedure described in section 4.7), it is important to check whether the average is a reliable representative of the underlying distribution per condition. To determine the MOS reliability, one should ideally identify the distribution model of the data. Though, considering our limited sample size (20 scores per condition, from each setup) conventional statistical modelling may not necessarily reach to a conclusive result. Instead, a kurtosis analysis has been recommended in standards like ITU.BT500.13 [22], where a score distribution with a kurtosis value of 2 to 4 is interpreted as a representative of the normal probability model. However, this is a vague and flawed assumption. It is indeed correct that the kurtosis of a normal distribution model is equal to 3, though mathematically this is a necessary but not sufficient condition. Moreover, by definition its only unambiguous interpretation is in terms of distribution tail extremity [62]. Nonetheless, no score set (per condition) in our dataset showed any irregular kurtosis value.

Next, we seek to answer two questions: (1) whether the subject scores for a specific test condition reach a consensus about the visual quality score for this test condition or not?; (2) in case the answer to the first question is positive, to what extend can that consensus be represented by the mean of these scores? To compactly address both, first we standardize the scores per
condition. The Z-score of each raw score is calculated by subtracting the mean of the obtained scores for the same test condition and then dividing it by the standard deviation of the same score set. The advantage of Z-scores is that their normalization enables direct comparison of individual scores across all conditions and even different setups. Nonetheless, the Z-score value does not provide any information about the actual visual quality level. It gives the distance of each individual score from the average opinion score (in units of standard deviation $\sigma$). This way a histogram of all scores for a particular setup (Fig. 8) gives an abstract view on the overall agreement among the test subjects. Notably, for all setups, a significantly good agreement is available around the mean opinion values, such that more than 69.5% and 96.5% of the individual scores fall within only 1 and 2 standard deviation(s) from their corresponding mean, respectively. Note, in a perfect normal distribution 68.27 and 95.45% of the samples are located within 1 and 2 standard deviation. Based on this and the fact that no specific skewness can be seen around the tails of shown distributions, we believe our MOS values can appropriately represent opinions of the majority of tested subjects.

![Fig. 8. Histogram of Z-scores per display setup calculated per condition from the raw scores before outlier removal and averaging. The percentage of Z-scores, which falls within 1 and 2 standard deviation(s) $\sigma$ from their mean (their MOS after outlier removal) are shown on the graphs.](image)

4.7. Statistical processing of results

The distributions of Fig. 8 shows that a very small portion of scores per setup is more than 4 standard deviations away from the average scores per condition. Therefore, an outlier detection and removal was performed on the test results. Following the procedure used in [63] and [30], the 25th ($Q_1$) and 75th ($Q_3$) percentiles were calculated. A score $u$ was considered as an outlier if $u>Q_3+w(Q_3−Q_1)$ or $u<Q_1−w(Q_3−Q_1)$, where $w$ was the maximum whisker length. In this experiment, $w$ was set to 1.5 which corresponds to a 99.3% coverage for normal distributed data. Our results also showed that no test subject had more than 15% of outlier scores. Consequently,
no test subject had to be removed from the dataset. After removal of the outlier scores, the average of the remaining scores for a particular test condition was combined into the final MOS.

5. Results and analysis

In this section, we provide the results of our subjective experiments and further investigate various aspects of the outcomes, potential similarities as well as correlations among the scores gathered from the three test setups.

5.1. MOS analysis based on 2D reconstructions with different focuses

First, the MOS values were evaluated at different reconstruction distances for light field and 2D displays. Figure 9 shows the overall comparison between MOS obtained from each hologram while the front and the back of the scene was in focus. Each MOS is averaged between the two perspectives (center and right-corner views). Exceptions were the holograms Mermaid (reconstructed at 1 focal distance) and Chess (reconstructed at 3 focal distances). It is obvious that the MOS from both depths are nearly following the same trend. Nonetheless, the non-averaged MOS are also visualized in the scatter plots of Fig. 9 (c, d). Therein points are colored different by perspective. At this point, the results do not show any meaningful differences. Therefore, to limit the degrees of freedom for our analysis, we use in the next subsections the MOS, which have been averaged for both focal distances. This means the number of MOS’ for light field and 2D display experiments equals to the number of MOS’ for the holographic display test (96 scores per perspective and a total of 192 scores per setup).

5.2. MOS analysis based on perspective

Next, the correlation between the scores for the two tested perspectives is evaluated. Figure 10 shows per setup the comparison of the MOS from the center view with the right-corner view. For each setup, first the center scores were sorted and the sorting indices were used to plot the corner view MOS. The 95% confidence intervals for each perspective are shown as well. To avoid clutter and to further clarify the trend, only 4th-degree polynomial fit lines for the mentioned data are shown in Fig. 10(a,b,c). To provide more detail, scatter plots of the center vs right corner MOS’ are shown in Fig. 10(d,e,f). The scatter plots clearly show a distinct trend across the setups where central views regularly obtain a higher MOS compared to the corner views of the same hologram. However, the score difference evolves across the quality range. More specifically, for all setups, the corner view MOS for high quality holograms (holograms with center view MOS higher than 3.5) remains within the confidence interval fits of the center view MOS. For holograms in the lower end of quality range, the differences increase. This is perfectly in line with the way some encoders compress the holograms [64,65]. When performing lossy compression the general objective is to weigh the transform components in the space-frequency domain, which carry more visually important information. However, if very high compression ratios are demanded, this will translate into complete elimination of the weakest coefficients. This leads, in the case of the chosen WAC variant, to an introduction of overlapping first diffraction orders by imperfect coefficient cancellation, which is more pronounced away from the center. In the case of the other selected methods, it leads to an elimination of high frequency components, which correspond to high diffraction angles (corner view information). The MOS variations experimentally reveal this shortcoming of the current holographic encoders. The scatter plots show furthermore that there are some cases that do not follow this trend of differences. In some extreme cases the center MOS is 1.5 points higher than the corner MOS.
5.3. Inter-setup comparison

In this section, the MOS results obtained with the different display systems are compared. First, the overall trend of the scores is evaluated. Thereafter, a more detailed analysis is performed related to the influence of the object characteristics and the bit-depths used to encode them in this experiment.

Figure 11 shows the inter-setup comparison for the center view. Similar to Fig. 10, the MOS results of the holographic setup in (a,b) and of the light field display in (c) were sorted and their order was used to plot MOS scores for the corresponding setup. The lines (solid and dashed) depict again the 3rd-degree polynomial fits for the corresponding data and their confidence intervals. Figure 11 (a, d) provide the comparison between the holographic and the light field setup. Interestingly, a specific gap exists between the two setups. For each hologram the MOS obtained for the holographic setup is typically higher than the MOS for the light field setup - except in the very low quality range. A similar trend can be observed in Fig. 11 (b,e) where the holographic setup was compared with a 2D display. However, the gap for the mid-range quality holograms (scores 2.5 - 4) is slightly larger than in Fig. 11 (a, d). The graphs shown in Fig. 11 (c, f) demonstrate a rather close agreement between the MOS of the light field setup and 2D display. A comparison for the right corner view perspective is provided in Fig. 12.
corner view scores follow closely the trend found for the center view. It should be noted that not all the penalized artifacts in non-holographic displays are visible on the holographic display and that to our opinion explains the gaps in gathered MOS both for the center and corner views. See Fig. 11 (a,b) and Fig. 12 (a,b).

Additionally, for the cases where quality scores of a real holographic setup are not available, while having access to the 2D or light field scores; one can use the fit-functions evaluated and shown as blue lines in the scatter plots of Fig. 11 (d,e) and Fig. 12 (d,e) to predict the scores for the same data reconstructed in an holographic setup.

Table 3 shows the coefficients for the 4th-degree polynomials (Fig. 11 and Fig. 12), which are best fit in a least-squares sense, for each of the three pairwise comparisons of the display setups. During our experiments, the 4-th degree polynomial showed the lowest regression error while not over-fitting the data when comparing the fitting behaviour of polynomials of degree 1 to 7. In the same table, the Pearson and Spearman correlation coefficients are shown before and after application of the fit functions to the data. A concern, which arises here, is the robustness of the provided fit functions. For example, consider the case when a test subject changes its opinion about a hologram. The last column of the Table 3 represents the maximum possible change of the fitted values, if a single subject changes the score for a test condition by ±1 unit. The reported errors are in unit scores as well. These results indicate that a high correlation exists between the MOS obtained for the three display systems and, in particular, after polynomial fitting. Both Pearson and Spearman correlation coefficients are very large in the latter case, thus underlining the predictive power of 2D and light field displays with respect to a holographic setup. Nonetheless, it is important to understand these fits cannot be transferred to other display setups and a calibration process will always be needed. When we consider the non-fitted MOS, it
Fig. 11. center view MOS comparison for holographic versus Light Field display (a,d), holographic versus 2D display (b,e) and Light Field versus 2D display (c,f).

Fig. 12. Right corner view MOS comparison for holographic versus Light Field display (a,d), holographic versus 2D display (b,e) and Light Field versus 2D display (c,f).
is interesting to observe that the 2D and light field displays are more sensitive to artifacts than the holographic display. This is partially related to the higher quality of the issued 2D and light field displays, but also due to the fact they are displaying numerically reconstructed holograms which contain more coherent speckle noise than the content rendered on the holographic display, for which a partially coherent LED illumination reduced speckle noise. Also non-optimal optics further reduce naturally the spatial beam coherency and thereby the speckle noise. The test subjects were not familiar with the phenomenon of speckle noise and were instructed to ignore it for the 2D and light field displays. Though, it might have influenced their scoring.

Table 3. Calculated coefficients of the fit functions, which enable to map the gathered MOS from different setups into another for the center and the corner views. Pearson and Spearman correlation coefficients are provided to evaluate the accuracy of the functions. As for the robustness, the last column shows the maximum absolute error in unit score for the predicted fit, if one of the test subjects changes a score for a condition with ±1 unit score.

|                      | Before Fit | After Fit | Error  |
|----------------------|------------|-----------|--------|
|                      | p1 | p2 | p3 | p4 | p5 | Pearson | Spearman | Pearson | Spearman | Pearson | Spearman | Error  |
| Center View          |    |    |    |    |    |         |          |         |         |         |          |        |
| LF → OPT            | 0.03923 | -0.56276 | 2.72965 | -4.27371 | 3.72549 | 0.9179 | 0.9210 | 0.9483 | 0.9992 | 0.0273 |
| 2D → OPT            | 0.05112 | -0.65713 | 2.8338 | -3.80237 | 3.14057 | 0.8824 | 0.8975 | 0.9946 | 0.9998 | 0.0279 |
| 2D → LF             | 0.03264 | -0.38953 | 1.52786 | -1.27431 | 1.17860 | 0.9587 | 0.9650 | 0.9938 | 0.9999 | 0.0299 |
| Corner View          |    |    |    |    |    |         |          |         |         |         |          |        |
| LF → OPT            | 0.00585 | -0.13114 | 0.74093 | -0.41638 | 1.06515 | 0.9342 | 0.9357 | 0.9968 | 0.9998 | 0.0259 |
| 2D → OPT            | 0.02748 | -0.31877 | 1.12589 | -0.21345 | 0.47994 | 0.9257 | 0.9405 | 0.9958 | 0.9998 | 0.0221 |
| 2D → LF             | 0.02725 | -0.30645 | 1.12026 | -0.54832 | 0.78461 | 0.9531 | 0.9538 | 0.9936 | 0.9999 | 0.0349 |

Apart from the overall inter-setup comparisons, it is interesting to analyse the influence of the tested hologram on the scoring behaviour of the test subjects: do responses to the test display systems differ for different holograms? For reasons of brevity, only the difference range (dashed line), the 25%-75% inter-quartile interval, and the median difference are provided in Fig. 13 and Fig. 14. The former figure represents the inter-setup MOS differences per test hologram (for all test conditions), separately for the center Fig. 13 (a,b,c) and right-corner Fig. 13 (d,e,f) views. Here, we consider the median difference between the MOS of each pair of setups as an indicator of the difference between the scores (shown as red line inside each box). The smaller the inter-quartile range for each box plot, the higher the certainty on the difference-level (red-line) of the opinions for that object. For example, in Fig. 13 (a), the MOS of the holographic setup for all distorted versions of "Mermaid" is ≈ 0.75 larger than the MOS for the light field setup. Considering the small size of the inter-quartile-range ≈ 0.25, one may conclude that the shown difference almost equally persists across all the distortion types and distortion levels. When comparing Fig. 13 (a,b,c) with each others, the general trend of the MOS differences for the different setups (shown in Fig. 11), persists for each of the tested holograms. Thereby the MOS values among non-holographic displays are spread less compared to the MOS of the holographic display. The MOS results for "Chess" are the only results that show a rather stable behaviour across all setups. In Fig. 14, another categorization of the MOS differences between the three setups and two perspectives are shown. Here, the MOS obtained from the holograms with the same compression level (quality range) are compared across setups. When considering the median differences (red lines), a rather similar trend is recognizable in Fig. 14 (a,b,d,e), where the score gaps between the holographic display and 2D or LF displays for the bit-depths 0.5 and 0.75 bpp are larger; while the level of disagreement is smaller for the lowest and the highest bit-depths. The level of certainty for these results (referring to the size of the inter-quartile range) increases for higher bit-depths. In the case of a direct comparison of LF and 2D display the differences are statistically not relevant.
Fig. 13. Box plots of the MOS difference per hologram. Top: Box plots corresponding to the MOS differences for center view. Bottom: Shows MOS differences collected from right-corner view. Column-wise, the MOS differences between holographic (holographic setup) - light field display, holographic - 2D regular display, and light field-2D display are shown, respectively.

Fig. 14. Box plots of the MOS difference per compression level. The first and second rows shows the MOS differences for the center and right-corner view, respectively.
6. Conclusions

In this paper, we reported the results of a series of comprehensive subjective experiments where, for the first time, a set of digital holograms was designed and produced to evaluate various aspects of macroscopic holography. For each hologram 12 distorted versions were generated through compression at different bit-rates using state-of-the-art holographic encoders. Three separate subjective experiments were designed and implemented utilizing a holographic display, a light field, and a regular 2D monitor.

For the subjective tests, a double stimulus, multi-perspective, multi-depth subjective testing methodology was designed and adapted to the characteristics of the utilized displays. A total of 120 human subjects participated in the experiments. The acquired quality scores of the reconstructed holograms were compared based on perspective and focal distance. Our results showed no explicit distinction between the scores of holograms when different parts of the encoded objects were in focus. However, with a change of perspective there was a consistent gap between the rated visual qualities. The corner view generally scored lower than the center view, especially for mid-range and high distortion levels.

Further, we compared the scores obtained from a holographic display with the scores from the light field and 2D displays and identified another rather consistent and distinctive gap. Our results show that the same distorted holograms rendered on holographic displays appear less distorted to the human eye in comparison to the reference hologram than as it would be the case for a light field or a 2D display. Non-holographic displays depicted a higher sensitivity to artifacts than the holographic display, particularly due to (1) numerical reconstruction process that is required and the associated speckle noise introduced in this process, and (2) the higher fidelity of these displays. Nonetheless, it was demonstrated that the scores on different displays are highly correlated and follow a consistent trend through the quality range. This indicates that under the given test conditions numerically reconstructed holograms displayed on light field or 2D displays allow for appropriate predictions on the perceptual visual quality of holographic displays.

It is important to stress that the presented research has its limitations with respect to available display technology and issued test material, which implicates that extrapolation of these results has to be performed with care. Nonetheless, this paper provides a first in-depth described methodology to facilitate quality assessment procedures on holographic data and a base for further extensions, which can be potentially tuned to particularities of the issued display devices.

Finally, we are hoping that the provided results plus the annotated database of our holograms, which are publicly available, facilitate a reliable test-bed for designing or improving available holographic processing methods and plenoptic quality metrics, as well as systematic benchmarking operations for digital holograms.

Funding

Politechnika Warszawska; Cross-Ministry Giga KOREA Project (GK19D0100); FP7 Ideas: European Research Council (617779).

Acknowledgments

The models "Perforated Ball" and "Biplane" are courtesy of Manuel Piñeiro from GrabCad.com and ScanLAB Projects, respectively. Authors thank Dr. Saeed Mahmoudpour for providing comments to improve the quality of the manuscript.

Disclosures

The authors declare no conflicts of interest.
References

1. B. Lee, “Three-dimensional displays, past and present,” Phys. Today 66(4), 36–41 (2013).
2. 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, 114–130 (2019).
3. P. Schelkens, T. Ebrahimi, A. Gilles, P. Gioia, K.-J. Oh, F. Pereira, C. Perra, and A. Pinheiro, “JEPG Pleno: Providing Representation Interoperability for Holographic Applications and Devices,” ETRI J. 41(1), 93–108 (2019).
4. A. Symeonidou, D. Blinder, A. Munteanu, and P. Schelkens, “Computer-generated holograms by multiple wavefront recording plane method with occlusion culling,” Opt. Express 23(17), 22149–22161 (2015).
5. J.-H. Park, “Recent progress in computer-generated holography for three-dimensional scenes,” J. Inf. Disp. 18(1), 1–12 (2017).
6. Y. Pan, J. Liu, X. Li, and Y. Wang, “A Review of Dynamic Holographic Three-Dimensional Display: Algorithms, Devices, and Systems,” IEEE Trans. Ind. Inf. 12(4), 1599–1610 (2016).
7. T. Sugie, T. Akamatsu, T. Nishitsuji, R. Hirayama, N. Masuda, H. Nakayama, Y. Ichihashi, A. Shiraki, M. Oikawa, N. Takada, Y. Endo, T. Kakue, T. Shimohaba, and T. Ito, “High-performance parallel computing for next-generation holographic imaging,” Nat. Electron. 1(4), 254–259 (2018).
8. T. Shimohaba and T. Ito, Computer Holography: Acceleration Algorithms and Hardware Implementations (CRC Press, 2019).
9. T. Nishitsuji, T. Shimohaba, T. Kakue, and T. Ito, “Review of fast calculation techniques for computer-generated holograms with the point light-source-based model,” IEEE Trans. Ind. Inf. 13(5), 2447–2454 (2017).
10. A. El Rhammad, P. Gioia, A. Gilles, A. Cagnazzo, and B. Pesquet-Popescu, “Color digital hologram compression based on matching pursuit,” Appl. Opt. 57(17), 4930–4942 (2018). WOS:000434872300027.
11. J. P. Peixeiro, C. Brites, J. Ascenso, and F. Pereira, “Holographic Data Coding: Benchmarking and Extending HEVC With Adapted Transforms,” IEEE Trans. Multimedia 20(2), 282–297 (2018).
12. M. V. Bernardo, P. Fernandes, A. Arrifano, M. Antonini, E. Fonseca, P. T. Fiadeiro, A. M. G. Pinheiro, and M. Pereira, “Holographic representation: Hologram plane vs. object plane,” Signal Process. Image Commun. 68, 193–206 (2018).
13. T. Kim, J. Kim, S. Kim, S. Cho, and S. Lee, “Perceptual crosstalk prediction on autostereoscopic 3d display,” IEEE Trans. Circuits Syst. Video Technol. 27(7), 1450–1463 (2017).
14. S. Delis, I. Mademlis, N. Nikolaids, and I. Pitas, “Automatic detection of 3d quality defects in stereoscopic videos using binocular disparity,” IEEE Trans. Circuits Syst. Video Technol. 27(5), 977–991 (2017).
15. A. Gilles, P. Gioia, R. Cozot, and L. Morin, “Hybrid approach for fast occlusion processing in computer-generated hologram calculation,” Appl. Opt. 55(20), 5459–5470 (2016).
16. A. Gilles, P. Gioia, R. Cozot, and L. Morin, “Computer generated hologram from multi-view-plus-depth data considering specular reflections,” in 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), (IEEE, 2016), pp. 1–6.
17. D. Blinder, A. Ahar, A. Symeonidou, Y. Xing, T. Bruylants, C. Schretter, B. Pesquet-Popescu, F. Dufaux, A. Munteanu, and P. Schelkens, “Open Access Database for Experimental Validations of Holographic Compression Engines,” in 2015 Seventh International Workshop on Quality of Multimedia Experience (qomex), (Ieee, New York, 2015). WOS:000375091800066.
18. A. Symeonidou, D. Blinder, A. Ahar, C. Schretter, A. Munteanu, and P. Schelkens, “Speckle noise reduction for computer generated holograms of objects with diffuse surfaces,” Proc. SPIE 9896, 98960F (2016).
19. A. Symeonidou, D. Blinder, and P. Schelkens, “Colour computer-generated holography for point clouds utilizing the phong illumination model,” Opt. Express 26(8), 10282–10298 (2018).
20. E. H. Adelson and J. R. Bergen, “The Plenoptic Function and the Elements of Early Vision,” in Computational Models of Visual Processing, (MIT Press, 1991), pp. 3–20.
21. I. Viola, M. Řefábek, T. Bruylants, P. Schelkens, F. Pereira, and T. Ebrahimi, “Objective and subjective evaluation of light field image compression algorithms,” in 2016 Picture Coding Symposium (PCS), (2016), pp. 1–5.
22. “Recommendation ITU-R BT.500-13 - Methodology for the subjective assessment of the quality of television pictures,” (2012).
23. I. Viola and T. Ebrahimi, “Quality Assessment Of Compression Solutions for Icip 2017 Grand Challenge on Light Field Image Coding,” in 2018 IEEE International Conference on Multimedia Expo Workshops (ICMEW), (2018), pp. 1–6.
24. ISO/IEC JTC1/SC29/WG1, “JEPG Pleno Call for Proposals on Light Field Coding,” (2017). WG1N74014, 74th JPEG Meeting, Geneva, Switzerland.
25. I. Viola, M. Řefábek, and T. Ebrahimi, “Impact of interactivity on the assessment of quality of experience for light field content,” in 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), (2017), pp. 1–6.
26. I. Viola, M. Řefábek, and T. Ebrahimi, “Comparison and Evaluation of Light Field Image Coding Approaches,” IEEE J. Sel. Top. Signal Process. 11(7), 1092–1106 (2017).
27. E. Alexiou and T. Ebrahimi, “On the performance of metrics to predict quality in point cloud representations,” in Applications of Digital Image Processing XL, vol. 10396 (International Society for Optics and Photonics, 2017), p. 103961H.
28. A. Javaheer, C. Brites, F. Pereira, and J. Ascenso, “Subjective and objective quality evaluation of compressed point clouds,” in 2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP), (2017), pp. 1–6.
29. S. Perry, “JPEG Pleno Point Cloud Coding Common Test Conditions v3.2,” (2020). WG1N87037, 87th JPEG Meeting, Online (Erlangen, Germany).
30. A. Ahar, D. Blinder, T. Bruylants, C. Schretter, A. Munteanu, and P. Schelkens, “Subjective quality assessment of numerically reconstructed compressed holograms,” in Applications of Digital Image Processing XXXVIII, vol. 9599 (SPIE, United States, 2015), pp. 95990K.
31. H. Amirpourazarian, E. Fonseca, M. Pereira, and A. Pinheiro, “Quality Assessment of Holographic Data,” Doc. ISO/IEC JTC 1/SC 29/WG1 M84117, 84th Meeting, Brussels, Belgium (2019).
32. R. Kizhakkumkara Muhamad, A. Ahar, T. Birnbaum, A. Gilles, S. Mahmoudpour, and P. Schelkens, “JPEG Pleno Holography Common Test Conditions V1.0,” (2020). WG1N88040, 88th JPEG Meeting, Online (Geneva, Switzerland).
33. P. W. M. Tsang and T.-C. Poon, “Review on the state-of-the-art technologies for acquisition and display of digital holograms,” IEEE Trans. Ind. Inf. 12(3), 886–901 (2016).
34. T. Kreis, “3d display by referenceless phase holography,” IEEE Trans. Ind. Inf. 12(2), 685–693 (2016).
35. H. Bjelkhagen, “Ultra-realistic 3-d imaging based on colour holography,” in Journal of Physics: Conference Series, vol. 415 (IOP Publishing, 2013), p. 012023.
36. J. Su, X. Yan, Y. Huang, X. Jiang, Y. Chen, and T. Zhang, “Progress in the synthetic holographic stereogram printing technique,” Appl. Sci. 8(6), 851 (2018).
37. I.-B. Sohn, H.-K. Choi, D. Yoo, Y.-C. Noh, J. Noh, and M. S. Ahsan, “Three-dimensional hologram printing by single beam femtosecond laser direct writing,” Appl. Surf. Sci. 427, 396–400 (2018).
38. M. A. Klug, “Display applications of large-scale digital holography,” in Holography: A Tribute to Yuri Denisyuk and Emmett Leith, vol. 4737 H. J. Caulfield, ed., International Society for Optics and Photonics (SPIE, 2002), pp. 142–149.
39. X. Yang, F. Xu, H. Zhang, H. Wang, Y. Li, and J. Zhang, “High-resolution fresnel hologram information simplification and color 3d display,” Optik 216, 164919 (2020).
40. H. Jeon, B. Kim, M. Jun, H. Kim, and J. Hahn, “High-resolution binary hologram printing methods,” in Practical Holography XXXV: Displays, Materials, and Applications, vol. 11306 H. I. Bjelkhagen, ed., International Society for Optics and Photonics (SPIE, 2020), pp. 122–127.
41. A. Symeonidou, D. Blinder, B. Ceulemans, A. Munteanu, and P. Schelkens, “Three-dimensional rendering of computer-generated holograms acquired from point-clouds on light field displays,” in SPIE Optics + Photonics, Applications of Digital Image Processing Xxxx, vol. 9971 A. G. Tescher, ed. (Spie-Int Soc Optical Engineering, Bellingham, 2016), p. 99710S. WOS:000390023100024.
42. A. Ahar, M. Chlipala, T. Birnbaum, W. Zaperty, A. Symeonidou, T. Kozacki, M. Kujawinska, and P. Schelkens, “Digital holography database for multi-perspective multi-display subjective experiment (holodb),” http://ds.ecore-interfere.eu/holodb/ (2019 accessed August 20,2020).
43. A. Golos, W. Zaperty, G. Fink, P. Makowski, and T. Kozacki, “Fourier rgb synthetic aperture color holographic capture for wide angle holographic display,” Proc. SPIE 9970, 99701E (2016).
44. D. Claus, D. Biescu, and P. Bryanston-Cross, “Quantitative space-bandwidth product analysis in digital holography,” Appl. Opt. 50(34), H116–H127 (2011).
45. G. Stroke, “Lensless fourier-transform method for optical holography,” Appl. Phys. Lett. 6(10), 201–203 (1965).
46. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, “Efficient subpixel image registration algorithms,” Opt. Lett. 33(2), 156–158 (2008).
47. J. W. Goodman, Introduction to Fourier optics (Roberts and Company Publishers, 2005).
48. K. Matsushima, “Shifted angular spectrum method for off-axis numerical propagation,” Opt. Express 18(17), 18453–18463 (2010).
49. D. S. Taubman and M. W. Marcellin, “JPEG2000: standard for interactive imaging,” Proc. IEEE 90(8), 1336–1357 (2002).
50. P. Schelkens, A. Skodras, and T. Ebrahimi, The JPEG 2000 suite, vol. 15 (John Wiley & Sons, 2009).
51. J. G. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) Standard,” IEEE Trans. Circuits Syst. Video Technol. 22(12), 1649–1668 (2012).
52. T. Birnbaum, A. Ahar, D. Blinder, C. Schretter, T. Kozacki, and P. Schelkens, “Wave Atoms for Lossy Compression of Digital Holograms,” in 2019 Data Compression Conference (DCC), (2019), pp. 398–407.
53. T. Bruylants, A. Munteanu, and P. Schelkens, “Wavelet based volumetric medical image compression,” Signal Process. Image Commun. 31, 112–133 (2015). WOS:000350078800010.
54. Fraunhofer, “HEVC reference repository,” https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/branches/HM-16.18-dev (2018 accessed July 1, 2018).
55. D. Taubman, “High performance scalable image compression with EBCOT,” IEEE Trans. on Image Process. 9(7), 1158–1170 (2000).
56. T. Kozacki, M. Chlipala, and P. L. Makowski, “Color fourier orthoscopic holography with laser capture and an led display,” Opt. Express 26(9), 12144–12158 (2018).
57. X. Li, J. Liu, J. Jia, Y. Pan, and Y. Wang, “3d dynamic holographic display by modulating complex amplitude experimentally,” Opt. Express 21(18), 20577–20587 (2013).
58. A. W. Lohmann, R. G. Dorsch, D. Mendlovic, Z. Zalevsky, and C. Ferreira, “Space–bandwidth product of optical signals and systems,” J. Opt. Soc. Am. A 13(3), 470–473 (1996).
59. J. W. Goodman, Introduction to Fourier Optics (Roberts & Company, 2004).
60. EIZO, “ColorEdge CG318-4K,” https://www.eizoglobal.com/products/coloredge/cg318-4k/index.html (2019 (accessed Feb 15, 2019)).
61. Holografika, “Holovizio 722RC,” http://holografika.com/722rc/ (2019 (accessed Feb 15, 2019)).
62. P. H. Westfall, “Kurtosis as peakedness, 1905–2014. rip,” The Am. Stat. 68(3), 191–195 (2014).
63. P. Hanhart, M. Rerabek, F. De Simone, and T. Ebrahimi, “Subjective quality evaluation of the upcoming hevc video compression standard,” in SPIE Optical Engineering+ Applications, (International Society for Optics and Photonics, 2012), pp. 84990V.
64. J. P. Peixeiro, C. Brites, J. Ascenso, and F. Pereira, “Holographic data coding: Benchmarking and extending hevc with adapted transforms,” IEEE Trans. Multimedia 20(2), 282–297 (2018).
65. T. Birnbaum, A. Ahar, D. Blinder, C. Schretter, T. Kozacki, and P. Schelkens, “Wave atoms for digital hologram compression,” Appl. Opt. 58(22), 6193–6203 (2019).