Received: 24 November 2018 | Revised: 7 February 2019 | Accepted: 26 February 2019
DOI: 10.1002/jsid.762

REGULAR CONTRIBUTED PAPER

WILEY

Angularly continuous light-field format: Concept, implementation, and evaluation

Aron Cserkaszky[1,2] | Peter A. Kara[2,3,4] | Roopak R. Tamboli[2] | Attila Barsi[2] | Maria G. Martini[3] | Laszlo Bokor[4] | Tibor Balogh[2]

[1]Faculty of Information Technology and Bionics, Pazmany Peter Catholic University, Budapest, Hungary

[2]Holografika Ltd., Budapest, Hungary

[3]Wireless Multimedia Networking Research Group, Kingston University, London, UK

[4]Department of Networked Systems and Services, Budapest University of Technology and Economics, Budapest, Hungary

Correspondence
Aron Cserkaszky, Faculty of Information Technology and Bionics, Pazmany Peter Catholic University, Budapest, Hungary.
Email: cserkaszky.aron@itk.ppke.hu

Funding information
European Unions Horizon 2020 Research and Innovation Program, Grant/Award Number: 643072 QoE-Net and 676401 ETN-FPI; KFI, 16-1-2017-0015

Abstract

In this paper, we present the concept, the implementation, and the evaluation of our novel angularly continuous light-field format. We compared the subjective visual performance of our format with the perspective camera format through a series of subjective and objective tests. In our extensive subjective study, we used multiple absolute and comparative rating scales and various visual stimuli with different contents and angular resolutions. The perceived quality was assessed by a total of 36 test participants, who were separated based on their scientific expertise. The objective quality was evaluated through the degradations caused by three well-known compression methods. The obtained results indicate that our light-field format may outperform the conventional format, and it generally can provide at least equivalent visual quality. Furthermore, these findings open the way for data size optimization, without compromising the achieved level of perceived quality.

KEYWORDS

angular resolution, light-field format, light-field visualization, objective quality assessment, perceived quality

1 | INTRODUCTION

At the time of writing this paper, multiple advanced visualization systems are continuously emerging. Some of them have already begun to spread on the consumer market (eg, ultra-high definition [UHD] displays and virtual reality [VR] head gears), while others are still in earlier phases (eg, high dynamic range [HDR] displays and augmented reality [AR] head gears). When it comes to advanced multimedia systems—for multimedia consumption specifically—recent years have demonstrated the painful fall of stereoscopic 3D technology, at least within the use case of home entertainment. Stereoscopic 3D cinema is still able to provide entertainment to considerably large numbers of viewers; however, its television counterpart never managed to reach any significant dominance on the market, and now the major manufacturers are halting development as a reaction to the choices of the users.

At the end of the day, researchers, developers, content providers, practically everyone involved in the creative and contributive processes towards multimedia technologies must face the cold, hard fact that user experience determines the true value of a given system or a service. With this in mind, the importance of perceived quality simply cannot be overestimated; we cannot go wrong by allocating the greatest levels of relevance and significance to what the users see and how satisfied they become during multimedia consumption, while using the investigated platforms. This notion is often expressed as Quality of Experience (QoE), which can be defined as "the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state."[1]

The demise of stereoscopic 3D home entertainment originated from its cumbersome nature. First of all, similarly to stereoscopic 3D cinema, special 3D glasses had to be worn during multimedia consumption. Although this could actually be tolerated, this also meant a significant limitation in the number of simultaneous viewers: the number of available glasses determined the maximum number of viewers, as stereoscopic 3D technology necessitated such viewing equipment. Of course it is slightly unfair to address this form of visualization in past tense, as if it was completely absent in present days; however, it should be admitted that while it still lingers in the present, it is definitely not the future.

It also needs to be noted that while 3D cinema typically provides entertaining visuals for limited periods of time (based on the duration of the given movie), home multimedia consumption now commonly accommodates the phenomenon of "binge-watching." While marathon-viewing movies (eg, trilogies) were already present in the past, it is entertainment companies with video-on-demand (VoD) services (ie, Netflix) that made binge-watching a common piece of our cultural reality. The fact that all episodes of a series become available for viewing on the date of release also evidently contributes to the "just one more episode" behavior.[2]

Let us now imagine that we wish to binge-watch a given television series of our choice. Regardless of genre and no matter whether we are talking about a multitude of shorter episodes or fewer longer ones, this activity highly suggests an extended period of time directly dedicated to the content. Although "content is king" indeed, the visualization technology plays its own role in this scenario (and generally in every scenario of multimedia consumption). When we imagine this scenario, do we include the possibility of having viewing equipment on our heads for several hours? Would that be compatible with our idea of comfort? After all, such activity is meant to be enjoyed under comfortable circumstances.

This easily highlights that not only stereoscopic 3D but also VR is a suboptimal advanced multimedia system for frequently recurring and extended forms of home entertainment and media consumption, especially if we take into consideration the list of known issues that rise even during short-time equipment usage.* We, the authors, however, need to state that we do not intend to question or doubt the technological and also artistic importance of the cinematic efforts in the world of stereoscopic 3D and VR. They do have their relevance and potential, but in the long run, it would be desirable to view 3D multimedia content without additional viewing devices.

Such glasses-free 3D technologies already exist today. Multiview displays—which are not to be confused with the Picture-by-Picture (PbP) solution where multiple sources are shown on one screen simultaneously—can visualize 3D content in a viewing-position-dependent manner. This means that if you view the screen of the display from a different angle, you will actually see the content from the corresponding angle. The major limitation here is that this technology is designed for so-called "sweet-spots" within the field of view (FOV) of the display. The FOV here is an angle measured from the display in which the viewer can move and can experience the parallax effect. The value of the FOV is typically small, and the content repeats itself horizontally, thus supporting multiple simultaneous viewers. This means that the number of those who can enjoy the glasses-free 3D experience is limited by the number of sweet spots supported by the display and the content.

Light-field displays—sometimes mentioned in the literature as super multiview displays—do not have such limitations, as they utilize the entire FOV to provide virtually continuous glasses-free 3D visualization. The currently available devices support horizontal parallax only (HPO) and are thus known as HPO displays. This means that if the viewer moves left or right within the FOV, the view changes accordingly, but this does not work along the vertical axis. Those displays that shall enable both horizontal and vertical parallax will be known as full parallax (FP) displays, which are yet to be developed.

Light-field displays have massive potentials in a vast array of fields, including medical applications, 3D design, gaming, traffic control, military control interfaces, and many more. In the scope of this paper, we particularly focus on the use cases for multimedia visualization (including light-field cinema[3]), as light-field displays are viewed as one of the most promising advanced multimedia systems. Yet with great potentials come great challenges and difficulties.[4] There are multiple reasons why light-field displays have not entered the consumer market yet (particularly for multimedia consumption), even though they started appearing in the industry. First of all, the current costs of such end-user devices could only be affordable for a very thin slice of global society, denying market entry. Second, multimedia content creation is still in an

*Oculus Rift and Touch health and safety warnings https://www.oculus.com/legal/health-and-safety-warnings/

experimental phase, and most of the research efforts turn towards static visualization. Third, light-field data can be absolutely immense, posing significant difficulties for both storage and transmission. Summa summarum, the displays are too expensive, and even if they were affordable, there would be no content to watch, and also, the real-time transmission of the content would not be feasible.

Light-field content—if captured or rendered in high quality—can indeed have vast requirements towards data sizes. It can be stored in many different ways. The easiest way to imagine such content is an array of 2D views, which is a one-dimensional array if the data are HPO and two-dimensional if it supports the parallax effect along both axes. While multiview displays visualize in the order of 10 views, we deal with hundreds of views in case of light-field displays. It is important to note here that light-field displays do not actually project an array of 2D views, as such data are converted into light rays before visualization. Evidently, it is possible to store the content directly as a set of rays, and there are many other methods available as well. The most vital fact here is that the chosen method of encoding affects the quality of the visualized content through the potentially different stored light-field rays and thus also affects QoE.

A light-field file format is a standard way for the rays of a 4D light-field to be encoded for storage. The currently available light-field formats, discussed in detail later, have certain shortcomings and restrictions that need to be tackled and overcome in order to provide suitable formats for future applications. For example, they do not provide efficient data interchange between arbitrary capture and display systems. New light-field formats need to satisfy multiple difficult criteria simultaneously, while prioritizing for the perceived quality of the visualized content at the side of the display.

In this paper, we introduce a research on the perceived quality of our novel light-field format. We carried out a series of subjective and objective tests to compare the visual performance of our own format with the conventional perspective camera format. The theoretical concept of the angularly continuous, display-independent light-field format, and the preliminary testing of the implementation were disseminated in earlier works.[5–7] This work reports an extensive visual quality assessment, which used multiple subjective rating scales, source content, and test conditions (with varying angular resolution) to investigate the capabilities of our proposed format. Visual performance was also investigated using various objective quality metrics and compression methods.

The remainder of this paper is structured as follows: section 2 reviews the related work in light-field QoE, particularly regarding perceived visual quality. The currently available light-field formats are comprehensively analyzed in section 3, focusing on their capabilities, use cases, and shortcomings. Our novel light-field format is introduced in section 4, including considerations regarding its implementation. The experimental setup of the subjective tests on perceived quality is detailed in section 5, and the obtained results are discussed in section 6. Objective quality is assessed in section 7. The paper is concluded in section 8.

2 | THE PERCEIVED QUALITY OF LIGHT-FIELD VISUALIZATION

Light-field QoE has started rapidly gaining attraction in recent years. In the past decade, various light-field displays have emerged, many of which are suitable for multimedia consumption scenarios, such as the television-like HoloVizio 80WLT[†] or the C80 light-field cinema system.[‡] As more and more researchers now have access to such displays in their institutions, studies on light-field QoE are highly supported.

The published works of Kovács et al[8–13] address light-field visualization from both the angle of perceived quality and the measured objective capabilities of systems. Spatial and angular resolution are particularly highlighted in these works, as they are key indicators of light-field visualization quality. It is important to differentiate these technical terms for display and content. In case of display capabilities, angular resolution refers to the "minimal angle of change that rays can reproduce with respect to a single point on the screen."[11] For content, it is practically the density of visual data. If we consider an HPO system that visualizes HPO content, and the raw data are a series of 2D images, then angular resolution is the ratio of the number of these images and the FOV. In such a format, content spatial resolution is defined by the 2D dimensions of these images.

The research on perceived quality carried out by Kara et al[14–16] exhaustively investigated content spatial and angular resolution. Their findings point out the possible compromises with spatial resolution, as lower values come with no pixelation that is uniform across the entire scene in the plane of the display. Instead, the leading visual phenomenon in this case is blur, which has different thresholds of toleration. On the other hand, angular resolution has been proven to be more critical and thus more difficult to reduce without significant negative impacts on QoE, which is also dependent on the parameters of observation, such as static or mobile observers.[17–19] The authors have also addressed the dependencies between spatial and

[†]http://holografika.com/80wlt/
[‡]http://holografika.com/c80-glasses-free-3d-cinema/

angular resolution and found that certain levels of spatial resolution reduction can actually visually compensate the disturbances in the smoothness of horizontal motion parallax, for static content[20] and light-field video[21] as well. Other results demonstrate that low amounts of spatial degradations can actually appeal to the viewer, as they may smoothen object edges in the scene or increase contrast.[22]

Low content angular resolution may be greatly compensated by view interpolation. The work of Cserkaszky et al[23] shows that even though interpolation techniques are "estimations" and are thus bound to decrease visual quality, such degradation may not always be visible and tend to have smaller negative effects on the perceived quality compared with representations with low angular resolution.

Tamboli et al[24–28] investigated the perceived quality of light-field view synthesis, created a high-angular-resolution dataset for objective and subjective assessments, evaluated content features, and developed an objective light-field quality metric with an angular component. The works recognize content angular resolution as one of the most vital quality factors, for both objective and subjective visual quality. In this paper, we involve the novel metric of Tamboli et al[25] in our analysis of objective quality.

Perra[29] also proposed an objective metric in his study on decompressed light fields. Other works of Perra et al[30–32] address the QoE of light-field subsampling, investigate the reconstruction of point clouds based on light fields, and study the use of nonoverlapping tiles for generating pseudo-temporal sequences of light-field images in an attempt to efficiently encode the data.

Because of the apparent sheer importance, spatial and angular resolution enhancement efforts are spreading in the field. As a recent example, the work authored by Gul et al[33] introduces a method for this purpose, which was trained through supervised learning. The results are promising, as the proposed method may provide significant improvements compared with certain conventional interpolation methods.

The perceived quality of light-field visualization evidently depends on data compression as well. Coding will play a significant role in the delivery of future light-field multimedia, which will need to balance between the extent of data reduction and the possible changes in visual quality. Results of scientific effort can already be observed in the works of Viola et al[34,35] and Paudyal et al.[36] Other works of Viola et al[37] address subjective test methodology, which was also investigated by Darukumalli et al.[38] A different work of Paudyal et al[39–43] demonstrates the importance of light-field content and displays system selection for subjective tests on perceived quality, consider watermarking, and present a light-field dataset captured by a Lytro camera. A database particularly created for QoE studies on the perceived quality of light-field visualization was also presented by Murgia et al.[44]

Shi et al[45] proposed a database as well and carried out subjective and objective evaluations on their static contents. The experiment used a stereoscopic 3D TV for the subjective tests, and the test participants had to interact with the visualized content by changing perspective with the help of a computer mouse (clicking and dragging). Light-field databases were also presented by Rerabek et al[46] and Wang et al[47] and the so-called "classic" datasets—such as the Stanford Archive§—were reviewed by Wanner et al.[48]

The work of Wang et al[49] investigates the QoE of multi-layer light-field displays (MLLFD). For such displays, special considerations regarding the perceived quality are required, as quality factors (ie, perceived resolution) may differ based on the implementation. Other important components of user experience are also measured, such as naturalness.

Agus et al[50] investigated the visualization of 3D medical (radiology) data on a light-field display. The authors state that their subjective tests—involving physicians and radiologists—have proven that such visualization method is "clearly superior" to conventional 2D displays. The work of Cserkaszky et al[51] also highlights the potential for nuclear medicine and points out research synergies. Furthermore, the paper of Lévêque et al[52] considers light field in their analysis on the perceived quality of medical contents.

The recent work of Wijnants et al[53] proposed HTTP adaptive streaming to transmit the data of static light fields. The core idea of the concept is approaching the source views of the model or the scene as the segments of a video sequence.

Interactive features were addressed by the contribution of Adhikarla et al,[54] describing a research in which free-hand gestures were tracked to carry out tasks (touching highlighted red sections) on a projected surface. Although perceived quality was not the primary research focus, the subjective test carried out is quite noteworthy, as they used the User Experience Questionnaire (UEQ)[55] to measure attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. The work of Marton et al[56] used multiple methods to navigate through large-scale static models visualized on a light-field display. For instance, hand gestures were tracked using depth sensors, enabling actions such as zooming or panning. The subjective assessment in the study also used a 3D mouse as the controller, with dedicated buttons for zooming in and out. Similarly to the previous work, several quality aspects were investigated, including ease of learning, ease of reach-

§http://lightfield.stanford.edu/lfs.html ing desired positions, and the perceived 3D image quality. The results indicate that the 3D mouse was preferred by the test participants over hand gestures.

The previously introduced researches on perceived quality mainly focused on the quality of static content visualization. There are works on light-field video as well, such as the video service feasibility research of Dricot et al,[57] the live capture system of Adhikarla et al,[58] the video frame quality analysis of Tamboli et al,[59] the proposed light-field-based 3D telemedicine system of Xiang et al,[60] the position statement of Kovács on light-field video,[61] the works of Kara et al[62,63] on dynamic adaptive light-field video streaming or the real-time telepresence system of Cserkaszky et al.[64]

The work presented in this paper addresses static content quality; still models were used to evaluate the perceived quality of our novel light-field visualization format. We decided to use such models in order to exclude visual factors that are present in videos and to allow test participants to extensively observe the stimuli before assessing the quality. The currently available formats and our own format are introduced in the following sections.

3 | LIGHT-FIELD VISUALIZATION FORMATS

The main purpose of light-field visualization formats is to efficiently store the measured or rendered rays of the light-field, as illustrated on Figure 1, with respect to quality and computing requirements of visualizing the stored light field. This enables the user to view the visual content encoded by the light-field format on a light-field display or a subset of the light field on a traditional 2D display. Light-field visualization formats behave more or less the same way as 2D movie and picture formats; we expect them to offer a good visual quality with respect to the specific use case, quality requirements, and available resources, like bandwidth, computing power, or storage space.

General properties of a useful light-field visualization format include backward compatibility, efficient data size, reasonable processing requirements, independence from any specific light-field capture or display system, and applicability to a wide range of use cases. Backward compatibility is obviously important as there are not many light-field displays. Efficient data size is also self-explanatory; HPO light fields generally contain hundred times as much information as traditional 2D content. But of course, the size efficiency cannot go against the visual quality of the representation.

A reasonable processing requirement is analogous to the nowadays hardware-accelerated media codecs of 2D content. However, in case of light field, we cannot yet expect hardware acceleration because of the very early stage of the adoption of the technology. We have to cope with the general purpose processing power of CPUs and GPUs available today. The capture and display system independence means that the format should not rely on any specifics of a particular system and should facilitate the interchange of data between any capture system to any display system. Last but not least, the format should not restrict the use cases by assuming unwarranted limitations of the extent of the stored light field.

Levoy and Hanrahan,[65] already in 1996, proposed a light-field format that characterized the ray structure of the 4D light field with the intersection of rays with two parallel planes, essentially creating a light-field slab. In this format, each ray is described by four spatial indices as the rays pass through a pixel of the first and the second plane. The authors also proposed covering the full solid angle with multiple of these slabs if the FOV of the content necessitated it. This solution is similar to the cube maps that are popular in VR applications nowadays and would have similar sampling issues at the boundaries of slabs. Hence, this format is not efficient for a light-field display with a wide FOV, and such displays already exist on the market today.

With the short-lived hype of hand-held light-field cameras necessarily came formats that accommodated them. Lytro created the *.lfp* light-field format to capture the full-parallax narrow-baseline light field.[66] The format can store the raw image of the microlens-based optics of the camera and can be processed into a color plus depth (RGBD) image by the provided software. The inherent focus on narrow-baseline light fields of this format makes it inefficient to store wide-baseline HPO light fields.

Another class of light-field formats[67,68] is using existing video compression methods to code static light-field images, treating the adjacent views as consecutive frames. This essentially handles the view dimension by treating it as a temporal dimension, while assuming a relatively large correlation of the neighboring views. These formats are not able to efficiently encode light-field video, and due to the correlation assumption, they work best with lenslet-based narrow-baseline light fields.

Since light fields are often captured with pinhole cameras placed at different locations, an obvious way to store their light field is to take the camera images with the respective parameters of the capture setup. These parameters include the extrinsic, intrinsic, and distortion properties of each camera, along with their differing color capturing properties. This format is relatively efficient in terms of storage size because of the already existing high-quality image compression methods, but, due to its generality, a large amount of processing is needed before it can be used for light-field visualization. The processing steps could include reconstruction, interpolation, or conversion to the light field of the display.

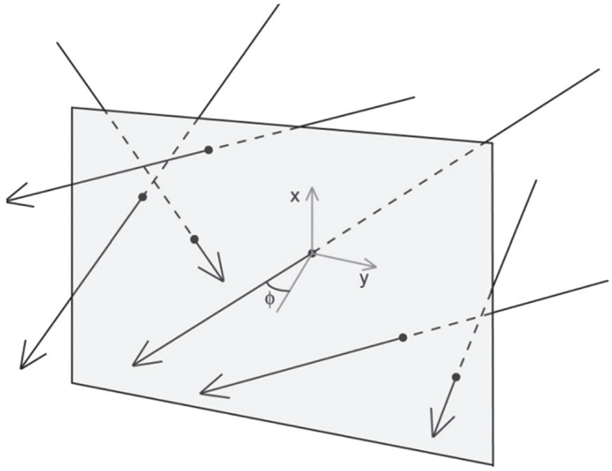

FIGURE 1 Illustration of light-field rays creating the "window of light"

Both JPEG[69] and MPEG[70] are working towards standardized light-field formats. However, current efforts of JPEG Pleno mainly focus on full-parallax and narrow-baseline light fields, while the MPEG-I is initially geared towards 360-video, free-viewpoint television and point-cloud-based rendering.

Probably, the most challenging use case for any light-field format is teleconferencing. In that scenario, the capturing of the light-field happens in real time, so the light-field information needs to be processed, compressed, and sent through the network with minimal delay. Then on the display side, this is decompressed and converted to the specific light-field ray structure it can emit and then shown the other party of the teleconference. The reasonable computation, small bandwidth, and low delay requirements make existing light-field formats inadequate.

4 | THE PROPOSED ANGULARLY CONTINUOUS LIGHT-FIELD FORMAT

Our main reason to create a novel light-field format is that existing formats did not prove to be efficient in interchanging data between arbitrary light-field capture systems and arbitrary light-field display systems without introducing prohibitive performance or visual quality degradations. In the design of the new format, we only assumed a relatively flat light-field display screen to enable efficient processing of the light field for existing displays. This led us to propose a format that parametrizes the 4D light field with two spatial coordinates on a plane, that approximately matches the screen of the display and with two angular coordinates that describe the direction of the emitted light ray.

Building upon the format of Levoy and Hanrahan,[65] we call the spatial dimensions s, t, and the angular dimensions φ, θ, as shown on Figure 2. We named our contribution the angularly continuous format due to the focus on the angularly uniform sampling of the light field. In the currently prevalent HPO light fields, this second angular dimension is reduced to one, so we also refer to the format as s-t-φ.

Let us now review how this format satisfies the requirements we proposed for a light-field visualization format. Backward compatibility with traditional displays is straightforward. A light-field viewing software on a 2D display can show the central view of the s, t dimensions of our format, or it can render views of other directions. This property has a corollary benefit that users viewing a light field with certain s, t, φ dimensions have inherent quality expectations when viewing such content on a given display, analogously to a 2D content and display case. Also, light-field device manufacturers could compare the capabilities and resolution of their displays measured along these dimensions.

Regarding efficient storage, our current implementation of the coding of the format uses existing 2D compression methods along the s, t image dimensions of the different directional views: PNG for each view for static scenes and HEVC for the views of a light-field video. The format includes a small clear text header file describing the extents along the s, t, φ, θ dimensions of the content, the specific φ, θ directions of each view and meta-data of the content. Hence, in the current implementation of the format, a specific light-field content is practically a folder containing a *header.txt* file along the *view_i_j.{png,avi}* files for each φ, θ directions with s × t pixels. This results in reasonable file sizes but can obviously be improved with exploiting the correlation of the adjacent views. This improvement would only affect the specific coding of the format but not the structure of the light field it stores. As future work, we will improve storage requirements by using codecs like MV-HEVC that compresses along both the temporal and angular dimensions of a light-field video.

One of the key features of our format is the low amount of processing required for light-field streaming between capture and display systems. A specific light-field capturing setup has inherent expectations of the resolution of the light field it is capable recording, hence precomputed look-up-tables (LUT) can efficiently convert the captured rays into the s, t, φ parameters. On the other side, based on the dimensions of the light field, a similarly precomputed LUT is used to convert the rays stored in the format to the rays the display is able to emit. Our GPU-shader-based implementation is able to do this task in real time for the HoloVizio light-field displays of Holografika in full resolution.

The display and capture system independence is inherently satisfied by our general format, with the sole assumption that the screen of the display should be approximately flat. Although this also means that the capturing system should only view the scene from one side, which is reasonable if we do not want the cameras to capture themselves.

The format also enables a wide range of light-field visualization use cases by providing a general interchange of light-field content between capture and display systems, with various properties. In a light-field VoD streaming scenario, various resolutions of the content can be precomputed and then served to the viewer with full QoE integration.[63]

The implementation of the format was straightforward considering the available high-quality 2D compression technologies and our four chosen dimensions. The compression could only be done efficiently in the spatial dimensions that are basically the same as the image domain that the compression algorithms were designed for.

The preliminary verification of the implementation was carried out via rendering and viewing the format on a 2D display. We rendered multiple static scenes with varying spatial and angular parameters and repeated these processes with the conventional format. After the rendered contents were successfully converted and shown on HPO light-field displays, we devised a series of subjective tests, where test participants were separated by related expertise. In the following section, the configuration and the relevant parameters of these tests are described in detail.

5 | EXPERIMENTAL SETUP OF PERCEIVED QUALITY EVALUATION

The assessment of the perceived quality was carried out in two subjective tests. First, we made an expert evaluation, with 12 experts in the field participating in our test. This was followed by an extensive quality assessment, with 24 "naïve" test participants (without any prior experience with light-field visualization). In both tests, we compared the perceived quality of visual stimuli created and stored with the conventional linear camera array technique (or perspective camera format) and with our angularly continuous format.

We used two static contents (complex mathematical models) and applied our test conditions to them to render the visual stimuli. One model (model 1) was a 972-faced polyhedron, and the other one (model 2) was a structure of 120 regular dodecahedra.[¶] The reason why we selected these models as source stimuli is that we have already utilized them in previous researches on perceived quality,[18,20,23] and the findings have shown that their visual quality is highly affected by degradations.

TABLE 1 Investigated input types

Input Type	Views per Degree	Number of Views
A	2	101
B	1	51
C	0.667	33
D	0.5	25

From the start of the research, we had visual degradations in mind, as formats need to perform well even if the input is insufficient in quality. As seen in related work and also in our own prior research, one of the most vital attributes of light-field data is angular resolution. Therefore, we chose this parameter as the variable for the test conditions.

While angular resolution varied, spatial resolution was constant for each and every source stimulus. The stimuli were rendered in 1440 × 1080, which means that the original 2D views had this given resolution. In earlier works, we concluded that this resolution is capable of providing excellent perceived quality on the selected light-field display.[14,20] As it has been defined earlier, content angular resolution is the ratio of FOV and the number of views. We defined four different input types (A, B, C, and D) for the two formats, listed in Table 1, whose values are reported in views per degree and the corresponding number of views.

These values were selected based on prior research. While two views per degree can provide excellent visualization quality, one view per degree is a typical borderline; for some, it provides adequate quality, whereas for others, it is simply unacceptable.[15,16,20] We also targeted lower angular resolutions, in order to investigate performance at critical levels (eg, in case of type D, the details of the models became barely recognizable because of the immense extents of the crosstalk effect).

The visual stimuli of the conventional format were captured using a linear camera array, arrangement of which is demonstrated in Figure 3, where the cameras cover the entire FOV of the display and were pointed at the center of the scene. The arrangement of visual information in our novel format is provided in Figure 4, illustrating pixels emitting different light rays along the φ dimension. Note that these two cases were calibrated in a way to match each other in the visualized FOV and the number of cameras and φ view directions were equal in all compared pairs.

The rendered images were then directly converted to the light field of the selected display system. The system at hand was the HoloVizio C80 light-field cinema

[¶]George W. Hart's Rapid Prototyping Web Page: www.georgehart.com/rp/rp.html

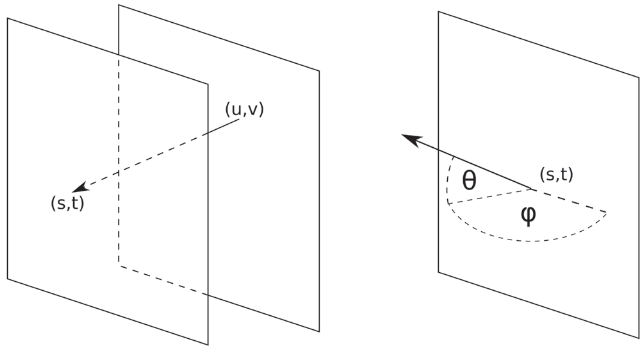

FIGURE 2 The bi-planar (s, t, u, v) parametrization and the proposed angularly-continuous (s, t, φ, θ) parametrization

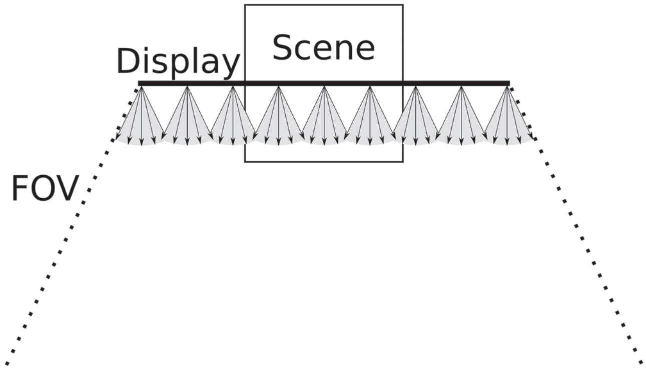

FIGURE 4 Arrangement of visual information in the angularly-continuous format

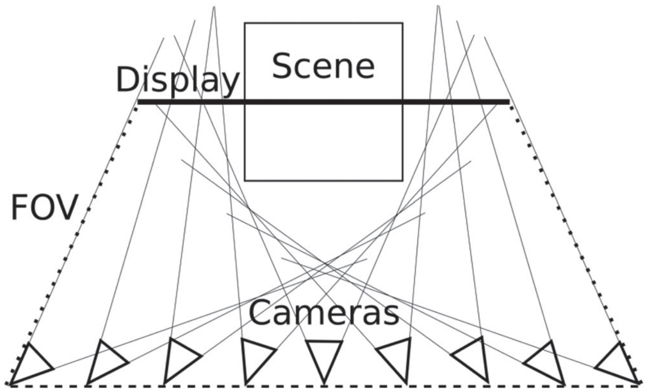

FIGURE 3 Arrangement of the linear camera array

system, which was chosen based on its high-end visualization capabilities, in order to support the accuracy of the subjective quality assessment task. The display was calibrated for a 50-degree FOV, which is also reflected in the value correspondence of Table 1. The C80 is a front-projection light-field display, which means that the optical engine array—which consists of 80 projectors, hence the name—is located on the same side of the screen as the observer.

The fact of the C80 being a front-projection system on its own creates a restriction for the utilization in the experiment: observers must not be located between the optical engine array and the screen, as it may cause the loss of visual information (ie, shadow cast on the screen) due to the occlusion of the projected light rays with the observer. We thus defined the closest viewing distance to be at the arc of the array, which is 4.6 m away from the screen. As viewing distance is typically given based on the height of the screen, this corresponded to 2.5 H. The default position of observation was at 2.5 H distance in the middle for each test stimuli, but test participants were encouraged to change the distance and viewing angle during stimulus visualization, in order to increase the precision of visual examination. The distance could be increased up to 3.5 H (6.5 m), and sideways movement was bound by the width of the system (4 m).

The tests were carried out in an isolated laboratory environment; test participants were not subject to external distractions. The lighting condition of the location was 20 lx. The entire screen was used to visualize the model, and only one visual stimulus was projected at a time. Before participating in the tests, each test participant was screened for vision using the Snellen charts and Ishihara plates.

Examples of the visualized test stimuli are shown in Figures 5, 6, 7 and 8. The examples only include types A and D, as they were the extremes of the test conditions. The images were captured by a single DSLR camera, viewing the screen from the center, at 2.5 H distance.

The above parameters applied to both the expert evaluation and the extensive quality assessment. The subjective test of course differed in the type of test participants and also in the subjective assessment tasks. Furthermore, in the extensive quality assessment, two more source models were used, with the same test conditions.

5.1 | Expert evaluation

The first test was a paired comparison, between the corresponding stimuli. Corresponding in this context refers to the different formats. This means that eg, type A of model 1 in perspective camera format was compared with type A of model 1 in the angularly continuous format. As the four test conditions (input types) were applied to two source stimuli, there were in total eight pairs. The comparison task was carried out using a three-point scale, with options "Worse," "Same," and "Better." So again, the experts had to assess the presence of observable differences between

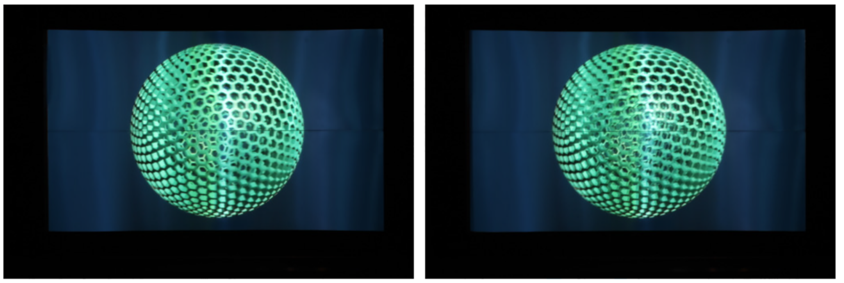

FIGURE 5 Types A (left) and D (right) of model 1 in perspective camera format

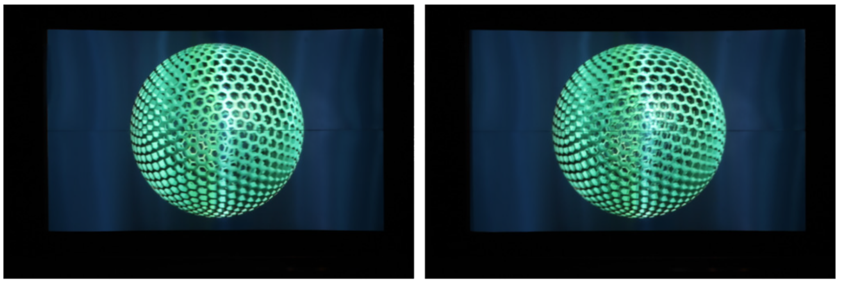

FIGURE 6 Types A (left) and D (right) of model 1 in the angularly continuous format

the formats, and if the stimuli could be distinguished, they had to choose the visually preferred one. They also had the option to provide written feedback about the differences, to help us understand the obtained results.

The visual stimuli were displayed in pairs without separation. This means that stimuli changed directly from one to the other inside a pair. The pairs themselves were separated by 5-second blank screens. It is important to note here that switching inside a pair was partially controlled by the test participant. Instead of having one single switch, multiple switches were enabled, and the expert could switch back and forth between the format until reaching verdict. The comparison task was performed right after each pair.

5.2 | Extensive quality assessment

As mentioned earlier, the second test contained two more source models, beyond the complex mathematical bodies. Models 3 and 4 were laser-scanned statues,[#] rendered with the same parameters as models 1 and 2. Compared with models 1 and 2, these models utilized the depth budget of the display much less. Figure 9 shows the models as seen on the rendering machine from central view.

The assessment of perceived quality consisted of four evaluation tasks: (1) each test stimulus was subject to the well-known five-point Absolute Category Rating (ACR), which means that integer scores were provided between 1 (bad) and 5 (excellent). (2) Test participants also had to provide binary scores of acceptance for these stimuli. This means that each stimulus had to be rated with either 0 (unacceptable) or 1 (acceptable) regarding the overall visualization quality. (3) Similarly to the expert evaluation, stimulus pairs were formed with the different formats, and they were to be compared. However, instead of comparing the overall visualization quality, the task was separated into the comparison of two aspects. Both aspects were compared using the ITU-R Rec. BT-500.13 seven-point ("Much Worse," "Worse," "Slightly worse," "Same," "Slightly better," "Better," "Much better") scale. One aspect was the so-called "3D continuity" of visualization. This aimed to reflect the perceived disturbances in the smoothness of the horizontal motion parallax. (4) The other aspect was "image quality." It focused separately on the quality of the representation, regardless how smooth the horizontal motion parallax was. Reductions in angular resolution affected both aspects of quality perception but differently.

We chose to separately investigate these aspects, as the formats behave differently when subject to angular resolution reduction. Both representation techniques suffer degradations, but these degradations are far from being identical when visualized on the screen of a light-field display.

[#] Jotero.com 3D scan and 3D measurement: forum.jotero.com/viewtopic.php?t=3

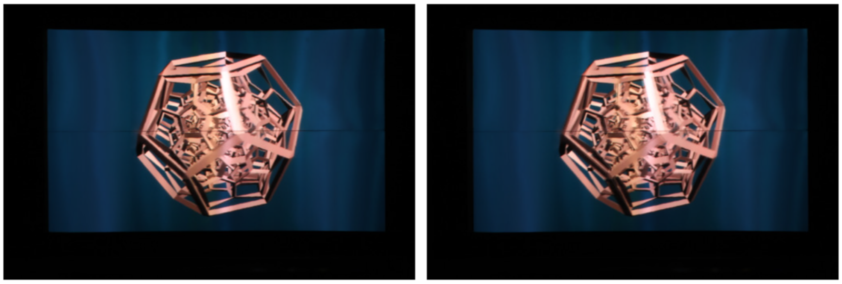

FIGURE 7 Types A (left) and D (right) of model 2 in perspective camera format

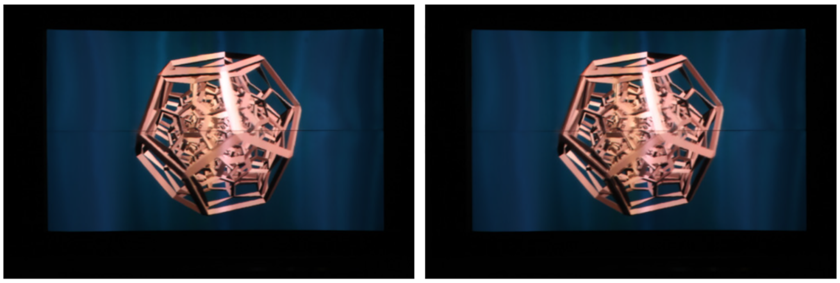

FIGURE 8 Types A (left) and D (right) of model 2 in the angularly continuous format

For the paired comparisons, stimulus separation was identical to the expert evaluation. However, the main difference here was that switching was limited to two instances; this means that if we refer to the stimuli in a given pair as X and Y, the order of visualization was X–Y–X–Y.

Prior to each subjective test, a training phase was conducted for the test participant. On the basis of the findings of an earlier study,[15] particularly great attention was dedicated to the disturbance of the smoothness of horizontal motion parallax in the training phase. Generally, it can be stated that test participants without experience and expertise with such visualization are most likely to be unaware of the visual sensation of disturbed parallax. In the real world, the parallax effect is evidently as smooth as it can be, which does not apply to light-field imaging with insufficient angular resolution. Static scenes with multiple simple 3D objects (eg, cubes) were shown to the test participants with different angular resolutions. These values matched the range in which the experiment was conducted (0.5 to 2 views per degree). Using values outside the interval could have biased the results through unnecessary rating option reservation, compressing scores towards the middle of the absolute scale. Furthermore, the objects were placed in such a manner that the positive and negative depth budget of the display were evenly utilized. This was required to show the test participant that the parallax effect is the smoothest in the plane of the screen, and thus, the highest levels of degradation through reduced angular resolution apply to those portions of the content, which stand out from this plane the most; in essence, the greater the depth, the greater the degradation. Moreover, through these example stimuli, the test participant was instructed in detail how to differentiate between 3D continuity and image quality.

6 | OBTAINED SUBJECTIVE TEST RESULTS

In this section, we introduce the subjective test results obtained during the two experiments. In the figures, we denote our angularly continuous format as "AC," and the perspective camera format as "PC." In the analysis of the results, no outliers were detected, and therefore, the presented mean scores and rating distributions contain the quality assessments of all test participants. The prevention of outliers was supported by the training phase, during which rating scale usage was not only practiced, but it was also highly emphasized for the test participants that the distances between the options of the ACR scale were uniform. Regarding the comparison scale, test subjects were instructed to only indicate a difference if (a) they notice a perceptible difference between the two stimuli, and (b) they actually have a preference. Forcing a preference can result in a random selection, which can bias results. Clear instructions on quality assessment and on scale usage can also combat preconception-based cognitive bias (eg, "there should be a difference"). Furthermore, the extensive train-

FIGURE 9 Models 3 (left) and 4 (right)

ing phase also aided the reduction of the apparent "wow effect," which is probably the most common bias for the evaluation of emerging display technologies. Finally, the scores of the different test participants were not weighted; they were equally taken into account during the calculations. Every test participant passed the screening for normal vision, thus this required no consideration in the analysis of the results.

6.1 | Expert evaluation

Out of the 12 experts, 11 were male, and 1 was female. The average age was 38, and the test participants fit into an age interval from 27 to 58.

A total of 96 comparison scores were collected, as 12 experts rated four test conditions implemented for two source stimuli. From these scores, 52 declared AC to be better, 43 did not differentiate the formats, and one favored PC.

The distribution of comparison scores for the two models is shown in Figure 10. The scores reflect that inputs with insufficiently low angular resolution (C and D) resulted in a clear favoring of AC. Types A and B formats were completely indistinguishable for model 2. Model 1 at B received scores very similar to C and D, and A was more on the side of the formats being indistinguishable in perceived quality.

These results are also reinforced by the comments received. The majority of the experts stated that visualization with AC was greatly preferable for C and D. Yet it needs to be noted that most of them also pointed out that it was rather challenging to rate A for model 1; there was a serious doubt regarding the perception of quality, whether there was an actual perceivable difference or not. Several experts found themselves on the borderline between the two scores, and this fact was observable during the test itself as well since in these cases, the number of switches between the stimuli was the highest.

Although in total, the scores of the experts do favor AC, we need to keep in mind that formats are created for everyone and not for experts only. The reason why we stress this statement is because an expert evaluation on its own can indeed provide us useful information, but it is far from being sufficient. Experts perceive visual quality differently because of their vast prior technical experiences with the visualization technology at hand, and their judgment is not necessarily in alignment with the assessment of those who lack such expertise. Therefore, we cannot jump to any conclusion without knowing how the 24 "naïve" test participants rated the light-field representations.

6.2 | Extensive quality assessment

Out of the 24 test participants, 14 were female and 10 were male. The average age was 21, and they fit into an age interval from 18 to 38. Compared with the group of experts, these test participants were much younger, typically university students.

Let us first take a look at the binary scores on acceptability. A total of 768 scores were gathered, as 24 test participants rated the visual acceptability of two formats implemented on four source contents and four input types. For AC, 311 scores out of 384 (81%) were "acceptable," and for PC, this number was 315 (82%). This suggests insignificant difference in acceptability. The acceptance rate of the formats can be further investigated, if analyzed per input type and per source content; these are shown on Figure 11 and Figure 12, respectively. We can see that neither the input type nor the source content produced a difference between the acceptance rates of the two formats.

The rating tendencies are also in alignment with what can be expected. For input type, the lower the density of the input, the lower the acceptance. As for the content, model 1 received the lowest acceptance rate, as it is highly vulnerable to disturbances in the horizontal parallax because of its extensive utilization of the depth budget and its fine-grained structure. Model 2 is similarly a sensitive source content but endures more as the components closest to the viewer are less complex. Results show that the laser-scanned statues of models 3 and 4 were deemed acceptable in perceived quality, even at input type D. This is mostly because of the low spatial volumes and depth variations of the models.

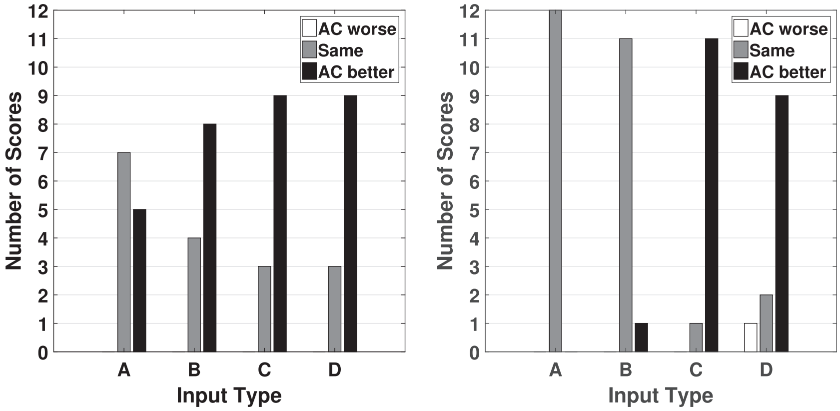

FIGURE 10 Distribution of comparison scores for Model 1 (left) and 2 (right)

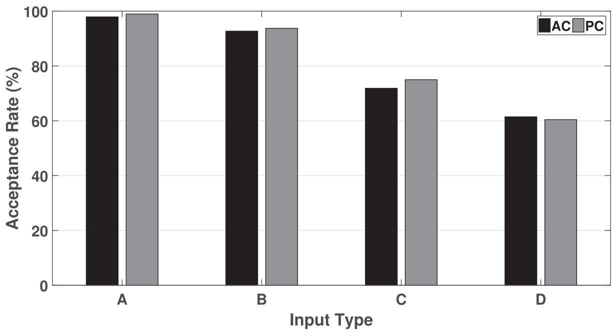

FIGURE 11 Acceptance rate of the formats per input type

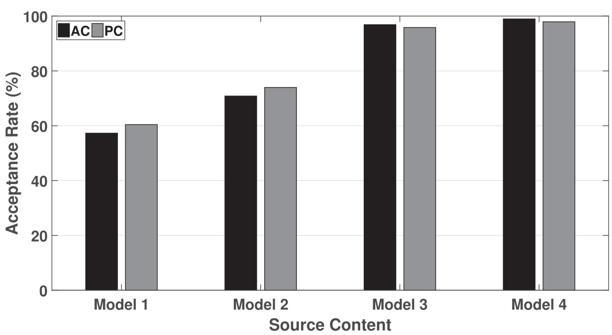

FIGURE 12 Acceptance rate of the formats per model

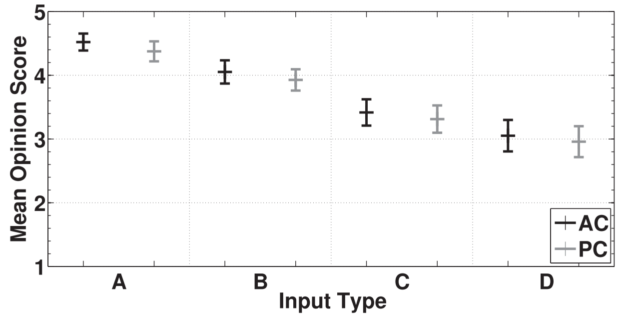

FIGURE 13 Mean opinion score (MOS) of the formats per input type

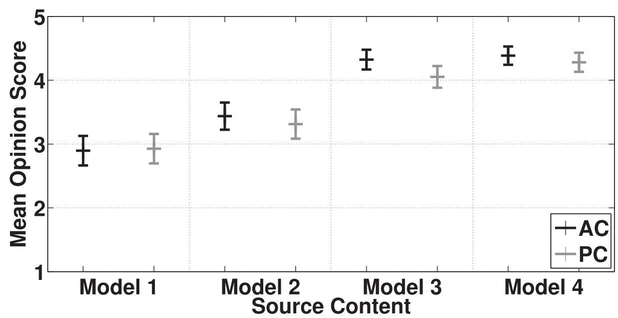

FIGURE 14 Mean opinion score (MOS) of the formats per model

Figure 11 indicates an acceptance rate around 60% for input type D. However, Figure 12 shows that both models 3 and 4 received rates around 100%. This means that the corresponding values for models 1 and 2 must be around 20% to 30%. Indeed, for input type D, model 1 is 25% for AC and 20% for PC, and model 2 is 29% for AC and 33% for PC.

The other evaluation scale of perceived overall quality was the five-point ACR. The mean opinion scores (MOS) per input type with 0.95 confidence intervals are shown on Figure 13, which indicates that there is no statistically significant difference between the perceived quality of the formats. In each and every comparison of the ACR scores, AC performed slightly better, but these differences are negligible.

Figure 14 shows the MOS per source content. Both these ACR scoring clusters and those based on input type correlate well with the binary scores of acceptance. So to summarize the results so far, the binary acceptance scores and the ACR scores both indicate the lack of significant difference between the investigated light-field formats regarding perceived quality. However, these were indirect comparisons, as the stimuli were rated individually.

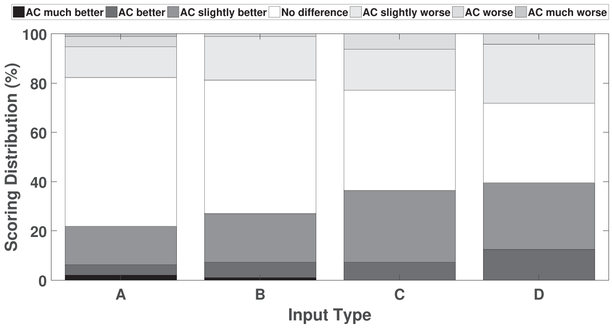

FIGURE 15 3D continuity comparison scores per input type

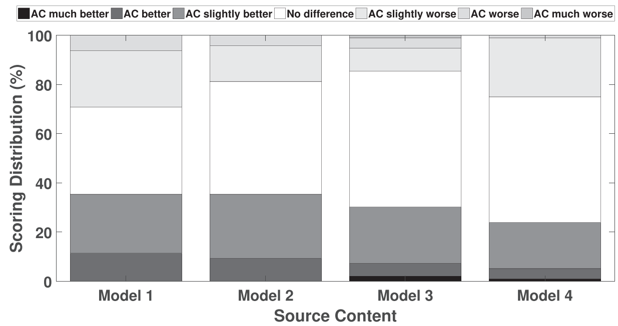

FIGURE 17 3D continuity comparison scores per model

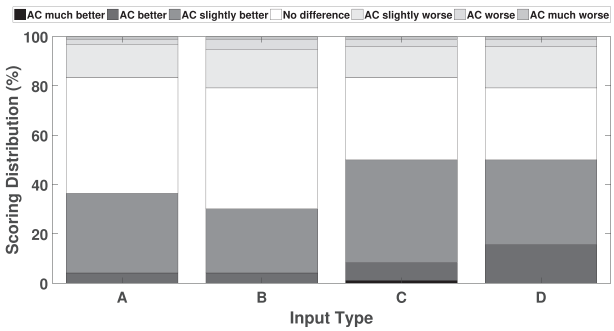

FIGURE 16 Image quality comparison scores per input type

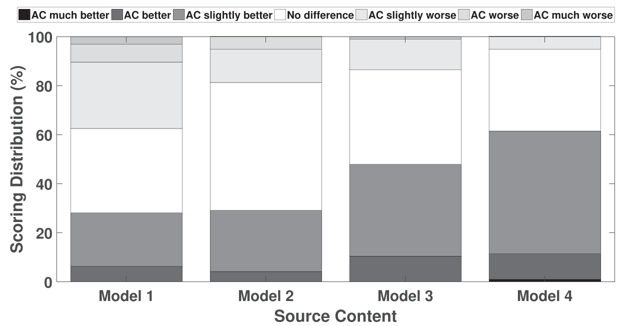

FIGURE 18 Image quality comparison scores per model

The paired comparisons were not indirect comparisons, as test participants had to provide a single score after observing two visual stimuli, and the task was to decide whether they differed in quality or not, and if they did, the better one was to be selected, also indicating the extent of difference on the seven-point scale. The obtained comparison scores are shown on Figure 15 —17, and 18. The qualitative tags on the displayed scores are not identical to the ones that were seen by the test participants; the tags here are focused on AC, as its performance is addressed in this paper.

3D continuity is more or less balanced between the two formats, when considering the inputs with different angular resolutions. There is a tendency that the lower angular resolution becomes, the more scores distinguish the performance of the formats in this aspect. This is also present for perceived image quality, however, here, there is a clear preference of AC (half of the scores for C and D).

When clustering by source content, we can see that less sensitive models are more difficult to distinguish for 3D continuity. Apart from model 4, AC is preferred, although the difference is not significant. The same does not apply to image quality. While PC is chosen model 1, and model 2 is nearly balanced out, AC dominates models 3 and 4. In fact, over 60% of the scores favor AC for model 4, and the amount of PC preference is below 10%.

Generally, we can see that the majority of the scores reported either indifference or slight difference. For 3D continuity, 336 out of 384 (87.5%) were these scores, and this number is 337 for image quality. This means that if difference is measurable in the aspects of perceived quality, it is practically not straightforward.

The subjective tests of extensive quality assessment did not include a written feedback option like the expert evaluation. From those reports, it is understandable that observers were not always fully convinced about the perceived difference between the formats. A more fine-grained scale allows smaller differences to be reported, which was the goal of the scale selection. Yet it needs to be noted that is also enables the registration of hesitant scores, in the form of “slightly different” scores, that sometimes lack strong visual proof.

The scores obtained from the seven-point scales for 3D continuity and image quality are in alignment with the binary and the ACR scores, which report the absence of a statistically significant difference between the performance of the two formats. Summa summarum, after completing this series of subjective tests on perceived quality, we do not have a hard proof that clearly shows that one of these formats is universally better or worse than the other, although in certain cases, AC did outperform PC. Yet this is not an issue at all, as it was not our goal to necessarily achieve this. The AC light-field format we designed and implemented can have further optimization for data size, which means that at the end of the day, we can represent light-field content with equivalent perceived quality but with smaller data requirements.

7 | OBJECTIVE QUALITY ASSESSMENT

In this section, we compare the AC and PC light field formats in terms of objective quality. The aim of such quality assessment methods is to cost-efficiently predict the QoE without directly involving human observers.[71] However, subjective tests on visual quality are required to evaluate the performance of objective methods, as the correlation between the two is desired; the more objective results correlate with subjective scores, the better. In our work, objective quality was measured via the resilience of the investigated formats towards lossy compression schemes. The exact same 3D models were used in this part of the research as introduced earlier for the subjective tests. As there are multiple ways of calculating objective quality, we used the four most common metrics of the scientific literature, plus the novel metric of Tamboli et al[25] specially created to assess light-field quality.

7.1 | Compression methods

Each source stimulus—rendered in both AC format and PC format—was compressed using JPEG,[‖] JPEG2000,[**] and WebP[††] compression methods at 20 quality/compression levels. Thus, the four source stimuli compressed using the three compression methods at 20 quality levels resulted in 240 sets of 101 images (angular resolution was identical to input type A). Including the four uncompressed sets, the total number of sets amounted to 244. The value of "Quality Parameter" (or "Compression Ratio" in case of JPEG2000) was varied from 5 to 100 in steps of five. The JPEG and JPEG2000 compressions were achieved in MATLAB. For WebP compression, we used the "`cwebp`" and "`dwebp`" tools from the WebP codec v0.6.1.

Each compressed image was compared with the corresponding original image using different quality metrics. A static 3D view on the light-field display was composed of a set of images in our experiment. Therefore, an average of the per-image quality values was considered as the quality of the 3D view, computed in a full-reference (FR) setting.

‖ https://jpeg.org/jpeg/index.html
** https://jpeg.org/jpeg2000/index.html
†† https://developers.google.com/speed/webp/

7.2 | Quality metrics

The FR 2D image quality metrics used in this experiment were from the class of pixel-based, structure-based, and scene-statistics-based metrics. Specifically, we used peak signal-to-noise ratio (PSNR), multiscale structural similarity (MS-SSIM),[72] feature similarity index measure (FSIM),[73] and Information Fidelity Criterion (IFC).[74]

The structure-based metrics—such as MS-SSIM—assume that the human visual system (HVS) is highly adaptive for extracting structural information from a scene and assess image quality based on the degradation in structural information. FSIM uses two low-level features for image quality assessment, namely, significance of a local structure and the contrast information. IFC relies on natural scene statistics models and assesses the perceptual quality by quantifying the mutual information between the reference and the distorted (compressed) images.

We now briefly explain the FR 3D objective quality metric used in this paper.[25] The metric—considering the spatio-angular nature of the light-field content—evaluates the spatial and angular quality scores of a 3D perspective visualized on a light-field display and then pools them into a 3D quality score using a pooling parameter.

The spatial quality score Q_{2D} involves steerable pyramid decomposition of each of the constituent images of a 3D view, followed by fitting an univariate generalized Gaussian distribution (UGGD) on the coefficients. A feature vector corresponding to a 3D view is formed by stacking the parameters of UGGD for all the constituent images. Then, the spatial quality score Q_{2D} is the distance between a feature vector of a pristine 3D view and a feature vector of a distorted 3D view, where each constituent image is distorted (compressed).

The angular quality score Q_θ evaluates pairwise structural similarity between the optical flow arrays computed for a pristine 3D view and a corresponding distorted 3D view. Optical flow values are calculated between successive constituent images of a 3D view. The key idea is that the difference between an optical flow array for a distorted 3D view and the corresponding reference optical flow array—as measured by any structural similarity metric—indicates disturbances in angular continuity.

The Q_{2D} and Q_θ scores are then pooled into the 3D quality score as $Q_{3D} = Q_{2D}^{1-\alpha} \times Q_\theta^{\alpha}$. In this work, the specific metrics used to compute Q_{2D} and Q_θ were "WaveHedges"[75] and "MS-SSIM," respectively. The value of the pooling parameter α was "0.89." These parameters were found to maximize the correlation between subjective and objective scores, obtained in an earlier study conducted by Tamboli et al[25] In their experiment, spatial distortions were added to the content while angular resolution was

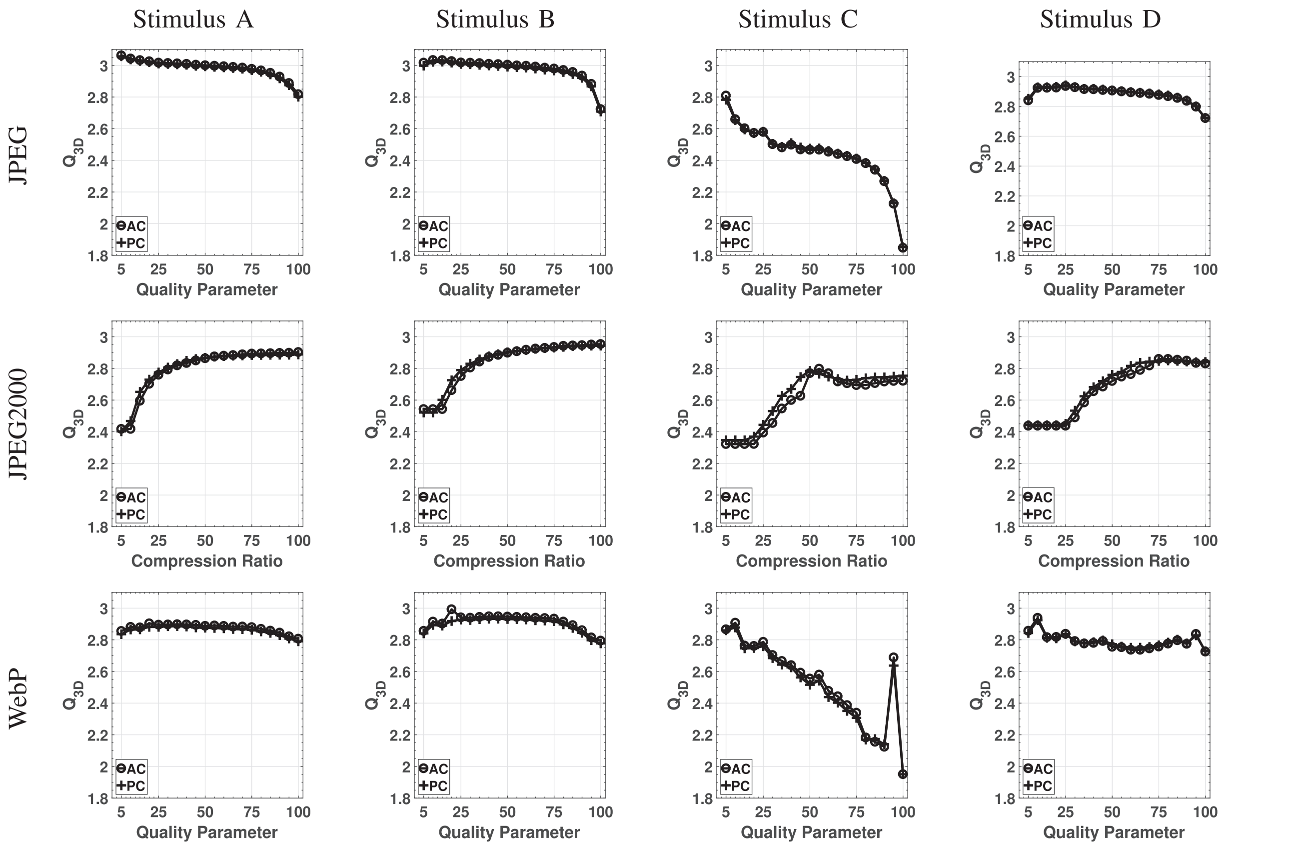

FIGURE 19 3D quality assessment using Q_{3D} metric

fixed. Similarly, in this research, spatial distortions (compression) were added to the content, and angular resolution was fixed. Therefore, we used the Q_{3D} metric in its original settings, without any optimization specific to the light-field content or distortions.

7.3 | Obtained objective quality results

The results of the 2D objective quality assessment indicate that in general, our AC format achieved marginally higher objective scores across the stimuli and the compression methods. This finding particularly applies to higher compression ratios and lower quality levels. In a few cases—such as IFC scores for stimuli A and B—PC format achieved slightly higher objective values. The obtained data regarding these 2D quality metrics are analyzed in detail in an earlier publication.[7]

Results for the 3D quality score Q_{3D} are shown in Figure 19. No significant difference was observed between two formats in terms of objective quality. For stimuli A, B, and D, variation in Q_{3D} scores was very small, whereas for stimulus C, Q_{3D} scores varied significantly. This can be attributed to the fact that stimulus C has large depth variations compared with the other three stimuli.[23] Also, the absence of explicit angular distortions may have resulted in low variations in Q_{3D}, as the angular quality score has higher weight during the pooling operation.

The ranges of values taken by Q_{2D}, Q_{θ}, and Q_{3D} were found to be (16.01, 1668.31), (1.36, 1.41), and (1.85, 3.06), respectively. It was found that the minor differences that exist in some cases were due to large differences in the corresponding spatial quality scores. Indeed, computing the element-wise absolute differences in the Q_{2D} values, followed by computing variances across the quality parameter values or compression ratios, revealed that the variances were of the order 10^5. Among the three compression methods, variances were high for the JPEG2000 method. Across the stimuli—although no clear pattern was observed—variances of Q_{2D} values for stimuli C and D were high in general.

On the other hand, the angular quality scores Q_{θ} were not very different. The variances of the differences in angular quality scores Q_{θ} of the two light-field formats—calculated separately across the stimuli, across the compression methods, as well as across the quality parameter values or compression ratios—were of the order 10^{-3}. Since the angular quality scores received higher weight ($\alpha = 0.89$), minuscule differences were observed in the overall 3D quality score.

The spike in Q_{3D} value at quality parameter value of 95 for stimulus C under WebP compression arose due to the artifacts introduced by the "dwebp" tool. The said tool was used to convert "webp" images to "png" format images for computations in MATLAB. This anomaly was observed in all images generated for stimuli C and D, in both representation formats, even with the newer versions of the dwebp tool.

In the earlier work of Tamboli et al,[25] the aforementioned FR 3D objective quality metric was shown to be a good indicator of perceived quality on a large light-field display. Specifically, certain spatial distortions were added to multi-camera datasets before the display-specific light-field conversion and the objective quality assessment was performed. The objective scores were found to correlate well with subjective score obtained through a test conducted on Holografika's HV721RC display.[‡‡] Similarly, in this paper, only spatial distortions (compression artifacts) were introduced to the content without any display-specific light-field conversion. Therefore, we believe that the objective quality score Q_{3D} provides a good estimate of perceptual experience if the contents used in this experiment were to be visualized on a real light-field display. Indeed, the results presented here corroborate with the observations made in the expert evaluation study. These subjective scores indicate that for low angular resolution, the stimuli rendered in AC format were found to be better than those rendered in PC format. On the other hand, for sufficient angular resolutions, both formats were found to provide similar perceptual experience. As the objective quality assessment targeted high angular resolution, we can conclude that neither of the utilized metrics measured a significant difference between the visual representations created by the two formats, which is reinforced by the corresponding subjective scores.

8 | CONCLUSION

In this paper, we presented the concept, the implementation, and the quality assessment of a novel angularly continuous light-field format. The evaluation of perceived quality was performed with multiple rating scales, source contents, and values of input angular resolution, separately with experts and with "naïve" test participants. While the results of the expert evaluation suggested that contents with low angular resolution very clearly benefit from our format, this was not as apparent in the extensive quality assessment. Practically speaking, nearly 90% of the scores on the seven-point scale reported a lack of visual difference or only indicated slight differences. Also, there was a considerable amount of opposing opinions as well. We can, however, conclude that our light-field format is at least as good as the conventional perspective camera format. Our findings are also supported by different metrics of objective quality assessment. With this in mind, we can proceed to address our future work of data size optimization, without compromising the visual quality of the format, which may provide a more resource-efficient visualization alternative. Furthermore, such optimized format may significantly benefit future services, such as light-field video transmission.

[‡‡]HoloVizio 721RC, http://www.archive.holografika.com/Products/HoloVizio-721RC.html

ACKNOWLEDGMENTS

The work presented in this paper was supported by the European Union's Horizon 2020 Research and Innovation Program through the Marie Sklodowska Curie Grants 643072 QoE-Net and 676401 ETN-FPI and the KFI_16-1-2017-0015 HoloDiagnozis project via the NRDI Fund of Hungary. The authors would like to thank Aniko Simon for her editorial services and all the test participants for making this research possible.

ORCID

Aron Cserkaszky https://orcid.org/0000-0001-5708-9999
Maria G. Martini https://orcid.org/0000-0002-8710-7550

REFERENCES

1. Brunnström K, Beker SA, De Moor K, et al. Qualinet white paper on definitions of quality of experience; 2013.
2. Trouleau W, Ashkan A, Ding W, Eriksson B. Just one more: modeling binge watching behavior. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM; 2016. p. 1215–1224.
3. Kara PA, Nagy Z, Martini MG, Barsi A. Cinema as large as life: large-scale light field cinema system. *2017 International Conference on 3D Immersion (IC3D)*. IEEE; Brussels, Belgium; 2017. p. 1–8.
4. Cserkaszky A, Kara PA, Tamboli RR, Barsi A, Martini MG, Balogh T. Light-field capture and display systems: limitations, challenges, and potentials. *SPIE novel optical systems design and optimization XXI*; San Diego, USA; 2018. p. 1–9.
5. Cserkaszky A, Kara PA, Barsi A, Martini MG. Towards display-independent light-field formats. *International Conference on 3D Immersion (IC3D)*. Brussels: IEEE; 2017. p. 1–7.
6. Cserkaszky A, Kara PA, Barsi A, Martini MG. Expert evaluation of a novel light-field visualization format. *3DTV Conference*; Stockholm, Sweden; 2018. p. 1–4.
7. Tamboli RR, Cserkaszky A, Kara PA, Barsi A, Martini MG. Objective quality evaluation of an angularly-continuous light-field format. *nternational Conference on 3D Immersion (IC3D)*; Brussels, Belgium; 2018. p. 1–8.

8. Kovács PT, Balogh T. 3D visual experience. High Qual Visual Experience. 2010:391–410.
9. Kovács PT, Balogh T. 3D display technologies and effects on the human vision system. *2011 2nd International Conference on Cognitive Infocommunications (COGINFOCOM)*. Budapest, Hungary: IEEE; 2011. p. 1–33.
10. Kovács PT, Balogh T. 3D light-field display technologies. Emerging Technologies for 3D Video: Creation, Coding, Transmission and Rendering. 2013:336–345.
11. Kovács PT, Boev A, Bregović R, Gotchev A. Quality measurements of 3D light-field displays. *Proc. Eighth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*; Chandler, Arizona; 2014. p. 1–6.
12. Kovács PT, Lackner K, Barsi A, et al. Measurement of perceived spatial resolution in 3D light-field displays. *International Conference on Image Processing (ICIP)*. Paris, France: IEEE; 2014. p. 768–772.
13. Kovács PT, Bregović R, Boev A, Barsi A, Gotchev A. Quantifying spatial and angular resolution of light-field 3-D displays. IEEE J Sel Top Signal Process. 2017;11(7): 1213–1222.
14. Kara PA, Kovács PT, Martini MG, Barsi A, Lackner K, Balogh T. Viva la resolution: the perceivable differences between image resolutions for light field displays. *5th ISCA/DEGA Workshop on Perceptual Quality of Systems (PQS 2016)*; Berlin; 2016.
15. Kara PA, Martini MG, Kovács PT, et al. Perceived quality of angular resolution for light field displays and the validity of subjective assessment. *International Conference on 3D Imaging (IC3D)*; Liege; 2016. p. 1–7.
16. Kara PA, Kovács PT, Vagharshakyan S, et al. The effect of light field reconstruction and angular resolution reduction on the quality of experience. *12th International Conference on Signal Image Technology & Internet Based Systems (SITIS) 3rd International Workshop on Quality of Multimedia Services (QUAMUS)*; Naples; 2016. p. 781–786.
17. Kara PA, Tamboli RR, Cserkaszky A, Martini MG, Barsi A, Bokor L. The viewing conditions of light-field video for subjective quality assessment. *International Conference on 3D Immersion (IC3D)*. Brussels, Belgium: IEEE; 2018. p. 1–8.
18. Kara PA, Cserkaszky A, Darukumalli S, Barsi A, Martini MG. On the edge of the seat: reduced angular resolution of a light field cinema with fixed observer positions. *9th International Conference on Quality of Multimedia Experience (QOMEX)*; Erfurt; 2017. p. 1–6.
19. Kara PA, Cserkaszky A, Barsi A, Martini MG. The couch, the sofa, and everything in between: discussion on the use case scenarios for light field video streaming services. *International Young Researcher Summit on Quality of Experience in Emerging Multimedia Services (QEEMS)*; Erfurt; 2017. p. 1–4.
20. Kara PA, Cserkaszky A, Barsi A, Papp T, Martini MG, Bokor L. The interdependence of spatial and angular resolution in the quality of experience of light field visualization. *International Conference on 3D Immersion (IC3D)*. Brussels: IEEE; 2017. p. 1–8.
21. Kara PA, Tamboli RR, Cserkaszky A, Martini MG, Barsi A, Bokor L. The perceived quality of light-field video services. *SPIE Applications of Digital Image Processing XLI*; San Diego, USA; 2018. p. 1–10.
22. Kara PA, Kovács PT, Vagharshakyan S, et al. Perceptual quality of reconstructed medical images on projection-based light field displays. eHealth 360° - International Summit on eHealth, Revised Selected Papers, Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering. Budapest, Hungary: Springer, 2017;181:476–483.
23. Cserkaszky A, Kara PA, Barsi A, Martini MG. To interpolate or not to interpolate: subjective assessment of interpolation performance on a light field display. *IEEE International Conference on Multimedia and Expo (ICME) 8th Workshop on Hot Topics in 3D Multimedia (HOT3D)*; Hong Kong; 2017. p. 55–60.
24. Tamboli RR, Vupparaboina KK, Ready J, Jana S, Channappayya S. A subjective evaluation of true 3D images. *2014 International Conference on 3D Imaging (IC3D)*. Brussels, Belgium: IEEE; 2014. p. 1–8.
25. Tamboli RR, Appina B, Channappayya S, Jana S. Super-multiview content with high angular resolution: 3D quality assessment on horizontal-parallax lightfield display. Signal Process Image Commun. 2016;47:42–55.
26. Tamboli RR, Appina B, Channappayya S, Jana S. Achieving high angular resolution via view synthesis: quality assessment of 3D content on super multiview lightfield display. *2017 International Conference on 3D Immersion (IC3D)*. Brussels, Belgium: IEEE; 2017. p. 1–8.
27. Tamboli RR, Kara PA, Appina B, Martini MG, Channappayya SS, Jana S. Effect of primitive features of content on perceived quality of light field visualization. *10th International Conference on Quality of Multimedia Experience (QOMEX)*; Sardinia, Italy; 2018. p. 1–3.
28. Tamboli RR, Reddy MS, Kara PA, Martini MG, Channappayya SS, Jana S. A High-angular-resolution turntable data-set for experiments on light field visualization quality. *10th International Conference on Quality of Multimedia Experience (QOMEX)*; Sardinia, Italy; 2018. p. 1–3.
29. Perra C. Assessing the quality of experience in viewing rendered decompressed light fields. Multimedia Tools Appl. 2018;77(16): 1–20.
30. Perra C, Song W, Liotta A. Effects of light field subsampling on the quality of experience in refocusing applications. *2018 Tenth International Conference on Quality of Multimedia Experience (QOMEX)*. Pula, Italy: IEEE; 2018. p. 1–3.
31. Perra C, Murgia F, Giusto DD. An analysis of 3D point cloud reconstruction from light field images. *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*; Oulu, Finland; 2016. p. 1–6.
32. Perra C, Assuncao P. High efficiency coding of light field images based on tiling and pseudo-temporal data arrangement. *2016 IEEE International Conference on Multimedia & Expo workshops (ICMEW)*, IEEE; Seattle, WA, USA; 2016. p. 1–4.
33. Gul MSK, Gunturk BK. Spatial and angular resolution enhancement of light fields using convolutional neural networks. IEEE Trans Image Process. 2018;27(5): 2146–2159.
34. Viola I, Rerabek M, Bruylants T, Schelkens P, Pereira F, Ebrahimi T. Objective and subjective evaluation of light field image compression algorithms. *2016 Picture Coding Symposium (PCS)*. IEEE; Nuremberg, Germany; 2016. p. 1–5.
35. Viola I, Ebrahimi T. A new framework for interactive quality assessment with application to light field coding. Proc SPIE. 2017;10396: 103961F 17 pp.
36. Paudyal P, Battisti F, Sjöström M, Olsson R, Carli M. Towards the perceptual quality evaluation of compressed light field images. IEEE Trans Broadcast. 2017;63(3):507–522.
37. Viola I, Rerabek M, Ebrahimi T. A new approach to subjectively assess quality of plenoptic content. Appl Digital Image Process XXXIX. 2016;9971: 99710X–1. SPIE.
38. Darukumalli S, Kara PA, Barsi A, Martini MG, Balogh T, Chehaibi A. Performance comparison of subjective assessment methodologies for light field displays. *2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. Limassol, Cyprus: IEEE; 2016. p. 28–33.

39. Paudyal P, Gutierrez J, Le Callet P, Carli M, Battisti F. Characterization and selection of light field content for perceptual assessment. *9th International Conference on Quality of Multimedia Experience (QOMEX 2017)*; Erfurt, Germany; 2017. p. 1–6.
40. Gutiérrez J, Paudyal P, Carli M, Battisti F, Le P. Perceptual analysis and characterization of light field content. Issue Overview, p. 49.
41. Paudyal P, Battisti F, Carli M. Effect of visualization techniques on subjective quality of light field images. *2017 IEEE International Conference on Image Processing (ICIP)*. Beijing, China: IEEE; 2017. p. 196–200.
42. Paudyal P, Battisti F, Neri A, Carli M. A study of the impact of light fields watermarking on the perceived quality of the refocused data. *2015 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*. IEEE; Lisbon, Portugal; 2015. p. 1–4.
43. Paudyal P, Olsson R, Sjöström M, Battisti F, Carli M. Smart: a light field image quality dataset. *Proceedings of the 7th International Conference on Multimedia Systems*. Klagenfurt, Austria: ACM; 2016. p. 49.
44. Murgia F, Giusto D. A database for evaluating the quality of experience in light field applications. *2016 24th Telecommunications Forum (TELFOR)*. IEEE; Belgrade, Serbia; 2016. p. 1–4.
45. Shi L, Zhao S, Zhou W, Chen Z. Perceptual evaluation of light field image. *2018 25th IEEE International Conference on Image Processing (ICIP)*. Athens, Greece: IEEE; 2018. p. 41–45.
46. Rerabek M, Ebrahimi T. New light field image dataset. *8th International Conference on Quality of Multimedia Experience (QOMEX)*; Lisbon, Portugal; 2016. p. 1–2.
47. Wang T-C, Zhu J-Y, Hiroaki E, Chandraker M, Efros AA, Ramamoorthi R. A 4D light-field dataset and CNN architectures for material recognition. *European conference on computer vision*. Amsterdam, The Netherlands: Springer; 2016. p. 121–138.
48. Wanner S, Meister S, Goldluecke B. Datasets and benchmarks for densely sampled 4D light fields. *VMV*, Citeseer; Lugano, Switzerland; 2013. p. 225–226.
49. Wang S, Ong KS, Surman P, Yuan J, Zheng Y, Sun XW. Quality of experience measurement for light field 3D displays on multilayer LCDs. J Soc Info Display. 2016;24(12): 726–740.
50. Agus M, Bettio F, Giachetti A, et al. An interactive 3D medical visualization system based on a light field display. Visual Comput. 2009;25(9): 883–893.
51. Cserkaszky A, Kara PA, Barsi A, Martini MG. The potential synergies of visual scene reconstruction and medical image reconstruction. *SPIE Novel Optical Systems Design and Optimization XXI*; San Diego, California; 2018. p. 1–7.
52. Lévêque L, Liu H, Baraković S, et al. On the subjective assessment of the perceived quality of medical images and videos. *2018 Tenth International Conference on Quality of Multimedia Experience (QOMEX)*. Pula, Italy: IEEE; 2018. p. 1–6.
53. Wijnants M, Lievens H, Michiels N, Put J, Quax P, Lamotte W. Standards-compliant HTTP adaptive streaming of static light fields. *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. Tokyo, Japan: ACM; 2018. p. 4.
54. Adhikarla VK, Sodnik J, Szolgay P, Jakus G. Exploring direct 3D interaction for full horizontal parallax light field displays using leap motion controller. Sensors. 2015;15(4): 8642–8663.
55. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. *Symposium of the Austrian HCI and Usability Engineering Group*. Graz, Austria: Springer; 2008. p. 63–76.
56. Marton F, Agus M, Gobbetti E, Pintore G, Rodriguez MB. Natural exploration of 3D massive models on large-scale light field displays using the FOX proximal navigation technique. Comput Graphics. 2012;36(8): 893–903.
57. Dricot A, Jung J, Cagnazzo M, et al. Subjective evaluation of super multi-view compressed contents on high-end light-field 3D displays. Signal Process Image Commun. 2015;39: 369–385.
58. Adhikarla VK, Marton F, Balogh T, Gobbetti E. Real-time adaptive content retargeting for live multi-view capture and light field display. Visual Comput. 2015;31(6–8): 1023–1032.
59. Tamboli RR, Kara PA, Cserkaszky A, et al. 3D objective quality assessment of light field video frames. *3DTV Conference*; Stockholm, Sweden; 2018. p. 1–4.
60. Xiang W, Wang G, Pickering M, Zhang Y. Big video data for light-field-based 3D telemedicine. IEEE Network. 2016;30(3): 30–38.
61. Kovács PT. 3D Video on light-field displays; 2010.
62. Kara PA, Cserkaszky A, Barsi A, Martini MG, Balogh T. Towards adaptive light field video streaming. IEEE COMSOC MMTC Communications-Frontiers. 2017;12(4):50–55.
63. Kara PA, Cserkaszky A, Martini MG, Barsi A, Bokor L, Balogh T. Evaluation of the concept of dynamic adaptive streaming of light field video. IEEE Trans Broadcast. June 2018;64(2): 407–421.
64. Cserkaszky A, Barsi A, Nagy Z, Puhr G, Balogh T, Kara PA. Real-time light-field 3D telepresence. *7th European Workshop on Visual Information Processing (EUVIP)*; Tampere, Finland; 2018. p. 1–5.
65. Levoy M, Hanrahan P. Light field rendering. *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*. New Orleans, Louisiana: ACM; 1996. p. 31–42.
66. Georgiev T, Yu Z, Lumsdaine A, Goma S. Lytro camera technology: theory, algorithms, performance analysis. *Conference, Multimedia Content and Mobile Devices*, vol. 8667, International Society for Optics and Photonics; Burlingame, CA; 2013. p. 86671J.
67. Liu D, Wang L, Li L, Xiong Z, Wu F, Zeng W. Pseudo-sequence-based light field image compression. *2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. Seattle, Washington: IEEE; 2016. p. 1–4.
68. Vieira A, Duarte H, Perra C, Tavora L, Assuncao P. Data formats for high efficiency coding of Lytro-Illum light fields. *2015 International Conference on Image Processing Theory, Tools and Applications (IPTA)*. Orleans, France: IEEE; 2015. p. 494–497.
69. Ebrahimi T, Foessel S, Pereira F, Schelkens P. JPEG Pleno: toward an efficient representation of visual reality. IEEE Multimedia. 2016;23(4): 14–20.
70. Domański M., Stankiewicz O, Wegner K, Grajek T. Immersive visual mediampeg-i: 360 video, virtual navigation and beyond. *2017 International Conference on Systems, Signals and Image Processing (IWSSIP)*. Poznan, Poland: IEEE; 2017. p. 1–9.
71. Chikkerur S, Sundaram V, Reisslein M, Karam LJ. Objective video quality assessment methods: a classification, review, and performance comparison. IEEE Trans Broadcast. 2011;57(2): 165–182.
72. Wang Z, Simoncelli EP, Bovik AC. Multiscale structural similarity for image quality assessment. *The thirty-seventh asilomar conference on signals, systems and computers*; Pacific Grove, California; 2003. p. 1398–1402.
73. Zhang L, Zhang L, Mou X, Zhang D. FSIM: a feature similarity index for image quality assessment. IEEE Trans Image Process. 2011;20(8).
74. Sheikh HR, Bovik AC, de Veciana G. An information fidelity criterion for image quality assessment using natural scene statistics. IEEE Trans Image Process. 2005;14(12).
75. Deza MM, Deza E. Encyclopedia of Distances: Springer; 2009.

AUTHOR BIOGRAPHIES

Aron Cserkaszky received his MSc degree in physics from the Budapest University of Technology and Economics in Hungary in 2011. He is currently a researcher at Holografika Ltd. in the area of full parallax imaging, working on novel light-field formats and light-field capture systems. He previously worked on GPU photon simulations and image reconstructions of PET and SPECT systems.

Peter A. Kara received his MSc degree in Computer Engineering from the Budapest University of Technology and Economics in Hungary in 2013. He participated in the CONCERTO project of the European Union's 7th Framework Programme and in the EU H2020 MSCA-ITN QoE-Net. He was a research associate of the Wireless Multimedia Networking Research Group at Kingston University, and he is currently with Holografika Ltd. His research interests include multimedia streaming, quality assessment of services, light fields, and cognitive bias.

Roopak R. Tamboli is a senior research fellow at the Indian Institute of Technology Hyderabad (IITH). He works in the areas of 3D reconstruction from multiple views and quality assessment of light-field content. He has been closely involved in setting up the Immersive Multimedia and Telepresence Lab at IITH. Pertaining to QoE, his work included the generation of super-multiview content, as well as subjective and objective methods for evaluating the quality of such content visualized on light-field displays. Prior to this, he received his Master's degree (M.Tech) in 2015 from IITH. For his Master's thesis, he explored sparse representations and compressed sensing of electrocardiogram signals. During an internship with Yokohama Research Laboratories, Hitachi Ltd., Japan, he explored compressed sensing for MR images. He has recently joined QoE-Net in 2018 and presently is with Holografika Ltd., where he works on the QoE of light-field content.

Attila Barsi received his MSc degree in Informatics Engineering from Budapest University of Technology and Economics in 2004. From 2004 to 2007, he was attending the PhD program of Budapest University of Technology and Economics, researching real-time graphics. Since 2006, he is in employment at the light-field display manufacturer company, Holografika Ltd., where he is researching real-time light-field capture and rendering. Since 2009, he is employed as the lead software developer for the company. He was working on several EU research projects and is currently participating as a supervisor in the QoE-NET and ETN-FPI European training networks. His research interest includes real-time rendering, light fields, multi-camera systems, and5-35 video streaming on high speed networks.

Maria G. Martini is (full) professor in the Faculty of Science, Engineering and Computing in Kingston University, London, where she also leads the Wireless Multimedia Networking Research Group, and she is course director for the MSc in "Networking and Data Communications." She received the Laurea in electronic engineering (summa cum laude) from the University of Perugia (Italy) in 1998 and the PhD in Electronics and Computer Science from the University of Bologna (Italy) in 2002. She is a Fellow of The Higher Education Academy (HEA). She has led the KU team in a number of national and international research projects, funded by the European Commission (eg, OPTIMIX, CONCERTO, QoE-NET, Qualinet), UK research councils (EPSRC, British Council, Royal Society), UK Technology Strategy Board/InnovateUK, and international industries. Her research interests include wireless multimedia networks, cross-layer design, joint source and channel coding, 2D/3D error resilient video, 2D/3D video quality assessment, and medical applications. She is the author of approximately 150 international scientific articles and book chapters and the inventor of several patents on wireless video.

Laszlo Bokor graduated in 2004 with MSc degree in computer engineering from the Budapest University of Technology and Economics (BME) at the Department of Telecommunicatons. In 2006, he got an MSc+ degree in bank informatics from the same university's Faculty of Economic and Social Sciences. After years of research work in multiple EU funded and national R&D projects, he received his PhD degree from the BME Doctoral School of Informatics. Currently, he is with the Department of Networked Systems and Services as assistant professor and leads the Commsignia - BME HIT Automotive Communications Research Group. He is a member of the IEEE and HTE, member of the Hungarian Standards Institution's Technical Committee for Intelligent Transport Systems (MSZT/MB 911) and the BME's Multimedia Networks and Services Laboratory (MEDIANETS) where he participates in researches of wireless communications and works on projects related to advanced mobile technologies. His research interests include IPv6 mobility, SDN/NFV-based mobile networks, mobile broadband architectures, network performance analyzing, network simulation, heterogeneous future mobile networks, mobile healthcare infrastructures, and V2X communication in cooperative intelligent transportation systems. He received the UNKP-16-4-I. Postdoctoral Fellowship (2016) from the New National Excellence Program of the Ministry of Human Capacities of Hungary.

Tibor Balogh CEO and Founder of Holografika, has extensive experience in the field of holography and optical engineering. He graduated as an electric engineer from the Technical University of Budapest and worked for SZKI (one of the major software houses in Hungary). Later, he was an assistant professor at the Eotvos Lorand University. Today, he is responsible for the overall management of his company, supervises the work of the research and development team and forms the overall business strategy. He was awarded the Joseph Petzval medal, the Kalmar prize, and the Dennis Gabor Award for his work, was World Technology Award finalist in 2006. He has several patents and publications and actively follows the developments of 3D display technologies around the world.

How to cite this article: Cserkaszky A, Kara PA, Tamboli RR, et al. Angularly continuous light-field format: Concept, implementation, and evaluation. *J Soc Inf Display*. 2019;27:442–461. https://doi.org/10.1002/jsid.762