Simon, Aniko, Guindy, Mary, Kara, Peter Andras, Balogh, Tibor and Szy, Laszlo "Through a different lens : the perceived quality of light field visualization assessed by test participants with imperfect visual acuity and color blindness." Proc. SPIE Volume 12097, Big Data IV: Learning, Analytics, and Applications, 120970P (2022). https://doi.org/10.1117/12.2618638

# Through a different lens: the perceived quality of light field visualization assessed by test participants with imperfect visual acuity and color blindness

Aniko Simon[a], Mary Guindy[b,c], Peter A. Kara[d,e], Tibor Balogh[b], and Laszlo Szy[f]

[a]Sigma Technology, Budapest, Hungary
[b]Holografika, Budapest, Hungary
[c]Pazmany Peter Catholic University, Budapest, Hungary
[d]Budapest University of Technology and Economics, Budapest, Hungary
[e]Kingston University, London, UK
[f]Radicon, Budapest, Hungary

## ABSTRACT

Light field visualization technology has progressed significantly in the past two decades. With the emergence of commercially-available devices, both industry and academia have begun research on the potential use cases of future society, including medical visualization, 3D digital signage, telepresence, military applications and many more. During the evaluation and quality assessment of such usage contexts and display types, test participants are typically screened for visual acuity via the Snellen chart and color vision via the Ishihara plates. However, there is an unfortunate global trend that the eyesight of the new generations is getting notably worse, and other sight-related issues, such as color vision deficiency, are becoming more common as well. Therefore, while medical technologies do relentlessly combat the diseases and disorders of the human eye, long-term innovations of visualization must also account for such users. Yet at the time of this paper, those with imperfect vision are underrepresented in light field research. In this paper, we present the results of the series of subjective tests carried out on light field displays, exclusively with test participants that otherwise would not qualify to assess visualization quality in a regular study. The experiments aim at investigating the most relevant research questions of light field visualization quality, such as spatial resolution, angular resolution and viewing distance. Test participants with imperfect visual acuity are classified by the diopters of their corrective lenses, and correlations between diopters and subjective ratings are addressed. Similar analyses were performed for color-blind test participants as well.

**Keywords:** Light field visualization, visual acuity, color vision, spatial resolution, angular resolution, viewing conditions

## 1. INTRODUCTION

As light field displays started appearing in research institutions, they opened up the scientific opportunity to conduct experiments related to visualization quality. Although the availability of such displays is still rather scarce, quite a few subjective studies have been carried out and published over the past decade. Typically in a series of subjective tests, test participants either rate visual quality, report on perceptual thresholds, select the preferred representation or perform interactive tasks. The collected data may support the development of systems and services based on light field technology in many ways. For example, exploring perceptual thresholds pave the way for perceptual coding, which means that two representations are perceptually indistinguishable, and

Further author information: (Send correspondence to Aniko Simon)
Aniko Simon: E-mail: aniko.simon@sigmatechnology.se
Mary Guindy: E-mail: m.guindy@holografika.com
Peter A. Kara: E-mail: kara@hit.bme.hu
Tibor Balogh: E-mail: t.balogh@holografika.com
Laszlo Szy: E-mail: laszlo.szy@radicon.hu

the representation that requires fewer resources is selected for use. Another example can be personal preference. For instance, if both representations are degraded in terms of quality, but they suffer different types and extents of quality degradation, then subjective tests on personal preference may identify the representation that is more tolerable. Subjective tests are also the ground truth for the training and evaluation of objective quality metrics. They either necessitate the entire reference stimulus (full-reference metrics) or a portion of it (reduced-reference metrics), or assess the quality of the impaired stimulus without any information related to the reference stimulus (no-reference metrics).

In order to achieve reliable results and to make sure that the data that may serve as ground truth is accurate, subjective tests usually exclude potential test participants with imperfect visual capabilities. Visual acuity is commonly measured by the Snellen chart, and color blindness is detected by Ishihara test, particularly regarding red-green color deficiencies. Since 3D visualization is at hand here, the random-dot stereo butterfly test is relevant as well, but it is not as widespread as the other two methods for screening test participants.

The tendencies of our current century show that more and more individuals of the younger generations suffer from sight-related issues. Therefore, during the optimization of future light field systems and services, the perception and preferences of such individuals should be taken into account as well.

It is, of course, perfectly meaningful to align objective metrics with the perceptual capabilities of individuals who do not suffer from sight-related issues. However, due to the total lack of subjective tests that involve test participants with imperfect visual capabilities, we are completely unaware of how they perceive light field visualization.

In this paper, we present our first work on light field visualization quality assessed by test participants with imperfect visual capabilities; those who normally would not qualify to take part in such a subjective study. As this is the initial work in a long series of studies, first of all, the extents of spatial and angular resolutions are investigated, which are two of the most important Key Performance Indicators (KPIs) of light field visualization.[1] Since viewing distance may particularly affect such test participants, multiple viewing distances are considered. Observer movement is also included in the experiment, but it is to be studied in greater detail in subsequent works.

The remainder of the paper is structured as follows: Section 2 reviews the related scientific literature. The details of the subjective tests are introduced in Section 3, the results of which are presented in Section 4. The paper is concluded in Section 5.

## 2. RELATED WORK

Numerous subjective tests on light field contents were published in recent years. However, not all of them were performed using light field displays. For example, quite a few subjective studies used 2D displays to represent the light field contents. In this review of the scientific literature, we solely focus on subjective tests that were carried out on real light field displays.

The works of Kovacs *et al.*[2–4] investigate spatial and angular resolutions. The experimental setups relied on symbol recognition (i.e., the test participants had to determine the orientation of the visualized symbol) and the discrimination of spatially close contents. These works constitute some of the earliest subjective tests on light field displays, focusing on perceptual thresholds.

Tamboli *et al.*[5–7] studied the effects of particular spatial distortions on perceived quality and how the synthesis of source views (i.e., replacing original views with synthesized ones, the quality of which may be lower due to inaccurate estimations) affect what the viewers perceive. In case of the latter, the more views were replaced, the more light field visualization quality was degraded. The choice of viewing conditions is quite interesting, as these seated test participants did not necessarily view the display from the center. Such experimental setup is specifically addressed in the work of Kara *et al.*[8] Later, Tamboli *et al.*[9] also investigated the personal preference regarding default content orientation.

Cserkaszky *et al.*[10] carried out tests on the subjectively perceived performance of interpolation techniques. In the research, test participants compared interpolated visual stimuli to each other, and reported their personal preferences. In a different work of the authors, test participants assessed a novel light field format.[11]

Zhang *et al.*[12] proposed a small-scale, mobile (i.e., flying) light field telepresence system, and evaluated the performance of the prototype. Cserkaszky *et al.*[13] also introduced a solution for light field telepresence, but the prototype was not tested via subjective tests.

The works of Darukumalli *et al.*[14,15] study the the impact of different levels of content zoom and the Region of Interest (RoI). The authors utilize various rating scales to collect the data, including the 5-point Absolute Category Rating (ACR) and Degradation Category Rating (DCR) scale*, a 5-point visual comfort scale† and a 7-point pair comparison scale‡.

The works of Kara *et al.*[16–18] address the Field of View (FOV), the viewing conditions and the connection between the degradation of spatial and angular resolutions. The authors also proposed and evaluated the concept of dynamic adaptive light field video streaming.[19]

Again, all the presented works of the scientific literature screen test participants for visual acuity and color vision in order to ensure assessment accuracy. In the following section, we introduce the experimental setup of the series of subjective tests that were carried out to study how imperfect visual acuity and color deficiencies affect the perception of light field visualization. According to the best knowledge of the author, this experiment is the first of its kind in this regard.

## 3. EXPERIMENTAL SETUP

### 3.1 Light field display and environment

The visual stimuli of the subjective tests were visualized on the HoloVizio HV640RC large-scale back-projection light field display. The tests were carried out in a laboratory environment, isolated from audiovisual distractions and external light sources.

### 3.2 Test variables

Based on the research question, three test variables were used: angular resolution, spatial resolution and viewing distance, for which 3, 2 and 2 values were selected, respectively. Regarding angular resolution, light field was created from either 56, 84 or 112 source views, and since the display has a 56-degree FOV, these values correspond to 1 degree, 0.66 degrees and 0.5 degrees, respectively. In an alternate format, they correspond to 1, 1.5 and 2 source views per degree, respectively. Spatial resolution was either $640 \times 480$ or $1024 \times 768$. The viewing distance was either 1.86 meters or 3.72 meters. These two distances were calculated based on the following equation,[20,21]

$$Viewing\ distance \leq \frac{Interpupillary\ distance}{tan(Angular\ resolution)} \tag{1}$$

which is derived from the average interpupillary distance (6.5 cm) and angular resolution. It defines the maximum viewing distance from which light field visualization – at a given angular resolution – should be observed. Since the lowest angular resolution in the experiment was 1 degree, solving this equation results 372 cm. The other distance selected for the tests was the half of the recommended maximum.

### 3.3 Source contents

A visual stimuli for the subjective tests were created from 8 source contents. They were all static light field scenes (i.e., 3D objects with a plain background). They were chosen based on their performance in prior research.

They were diverse in terms of structure, complexity, depth, colors and textures. As shown on Figure 1, the source contents included simple shapes, complex mathematical bodies, various objects (including textures) and laser-scanned statues. In the analysis, we refer to the source contents as contents A, B, C, D, E, F, G and H (left to right, top to bottom).

*https://www.itu.int/rec/T-REC-P.910

†https://www.itu.int/rec/R-REC-BT.2021

‡https://www.itu.int/rec/R-REC-BT.500

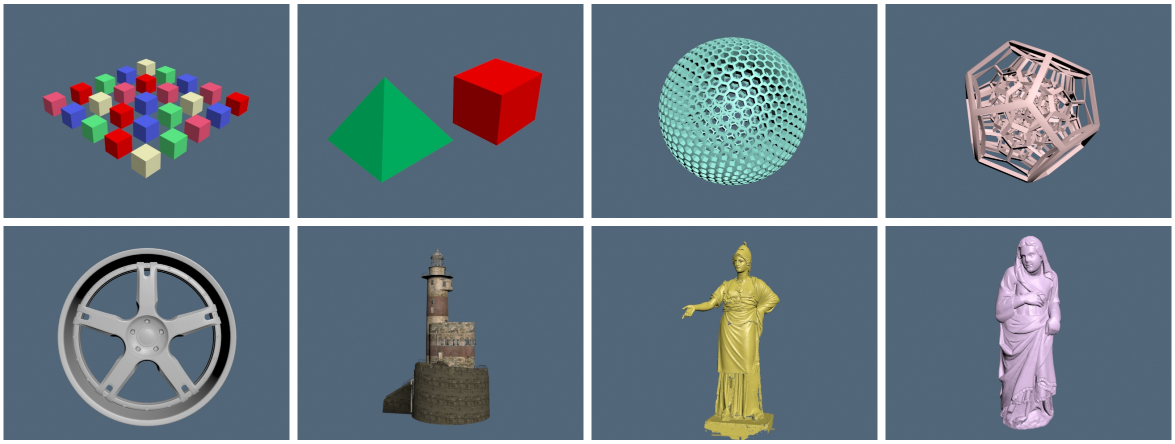

Figure 1. Source contents used to create the visual stimuli.

### 3.4 Test conditions and test stimuli

Every single test variable was combined with every other test condition, resulting a total of twelve test conditions. Two of these – the highest resolutions at the different distances – were selected as reference quality, and the remaining ten were to be assessed by the test participants. As there was a total of 8 source contents, to which every combination of resolution was applied, there were 48 different visual stimuli in total.

### 3.5 Rating scale

The tests used an ACR scale for evaluation. However, instead of using the conventional 5-point scale, a 10-point scale – ranging from 1 to 10 – was used instead, in order to provide more assessment space to differentiate the 48 test stimuli. 10 corresponded to the reference quality, and 1 was the worst possible rating.

### 3.6 Test procedure

The test stimuli were shown to the test participants in a randomized order, clustered by source and viewing distance. First, the reference stimulus was shown, which was to be carefully observed – but not to be rated – by the test participant. Then, the 5 degraded stimuli were shown, separated by plain separation screens. However, there were actually 6 stimuli to be rated, since we employed the methodology of hidden reference – the reference was to be assessed as well, without the awareness of the test participant.

Only one individual participated in the test at a time. The viewing distance was strictly set, indicated by highly-visible markers on the floor. The default horizontal position was the center view. Sideways movement within the FOV was permitted, limited to one step left and one step right, with respect to the center view.

### 3.7 Test participants

A total of 15 individuals participated in the subjective tests. 10 were male and 5 were female. The youngest test participant was 26 years old, and the oldest was 72. The average age was 40. All of the test participants wore either glasses or contact lenses, diopters ranging between $-6$ and $+2$. 8 of the test participants were color-blind; dominantly the red and green colors of the spectrum were affected. Generally, test participants who were not color-blind had higher diopters. During the training phase, test participants were not only familiarized with the specific assessment task and the test procedure, but the possible visual phenomena of the experiment – particularly the disruption of the smoothness of the parallax effect – was also demonstrated on other contents to support rating consistency. In the experiment, color-blind test participants are referred to as Group 1, and the remaining 7 test participants constitute Group 2.

## 4. RESULTS

A single test participant provided 96 quality ratings in total. As 15 test participants completed the experiment, we obtained 1,440 subjective scores.

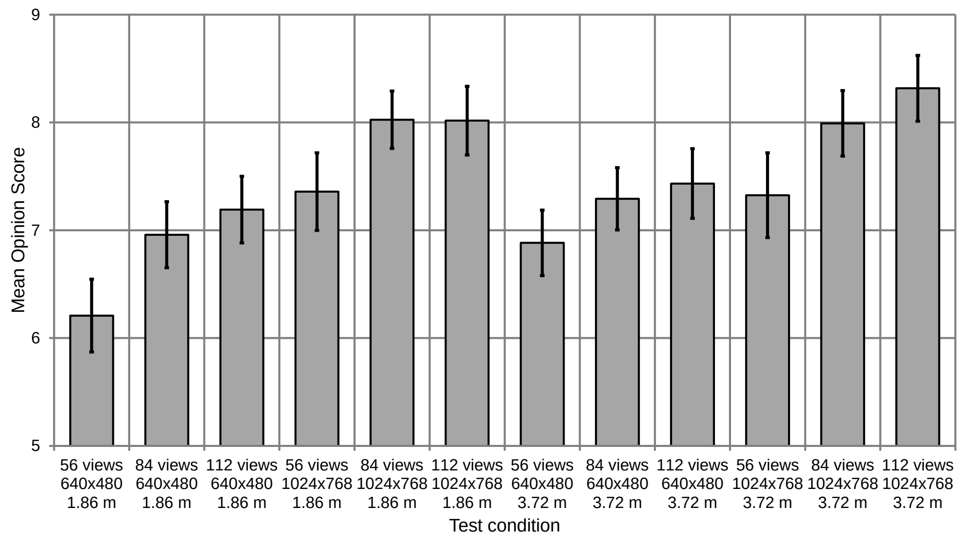


Figure 2. Mean Opinion Scores of the test conditions for all test participants.

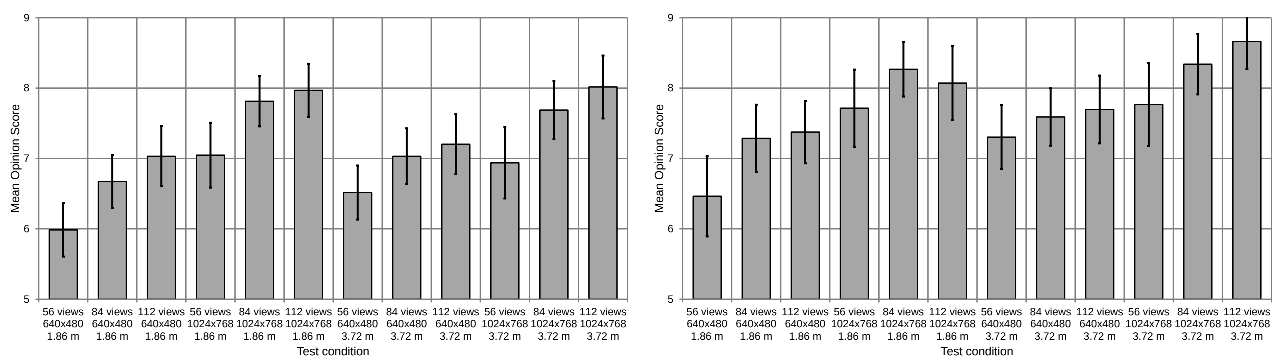


Figure 3. Mean Opinion Scores of the test conditions for Group 1 (left) and Group 2 (right).

The Mean Opinions Scores (MOS) of the test conditions, shown on Figure 2, highlight the importance of the hidden reference methodology. As stated earlier, on the rating scale, 10 represented the quality of the reference stimulus. Yet the MOS for the hidden reference was 8.016 at 1.86 m and 8.316 at 3.72 m viewing distance. Out of the 240 ratings collected for the reference stimulus at the two viewing distances, only 59 (less than 25%) of them were 10. Without the inclusion of the hidden reference, one could assume a generally lower level of perceived quality.

The results also indicate that test participants could not distinguish the stimulus rendered from 84 views at the higher spatial resolution and the reference stimulus. In fact, at the closer viewing distance, the degraded stimulus received a slightly higher MOS (8.025) than the reference.

The impact of the viewing distance is quite apparent from these results. When the test participants observed the degraded stimuli from the closer viewing distance, more indicators of impairments were perceivable. This is particularly applicable to the blur caused by low spatial resolution. If we consider the data collected at the closer distance, every comparison of spatial resolution (i.e., at all three angular resolutions) shows a difference at a statistically significant level.

The rating tendencies are similar for Group 1 and Group 2, as shown on Figure 3. However, globally, the color-blind test participants provided lower scores. It is also interesting to see that in case of Group 2, at the closer viewing distance, there is an MOS difference of 0.2 between the aforementioned stimuli rendered from 84 views at the higher spatial resolution and the reference stimulus

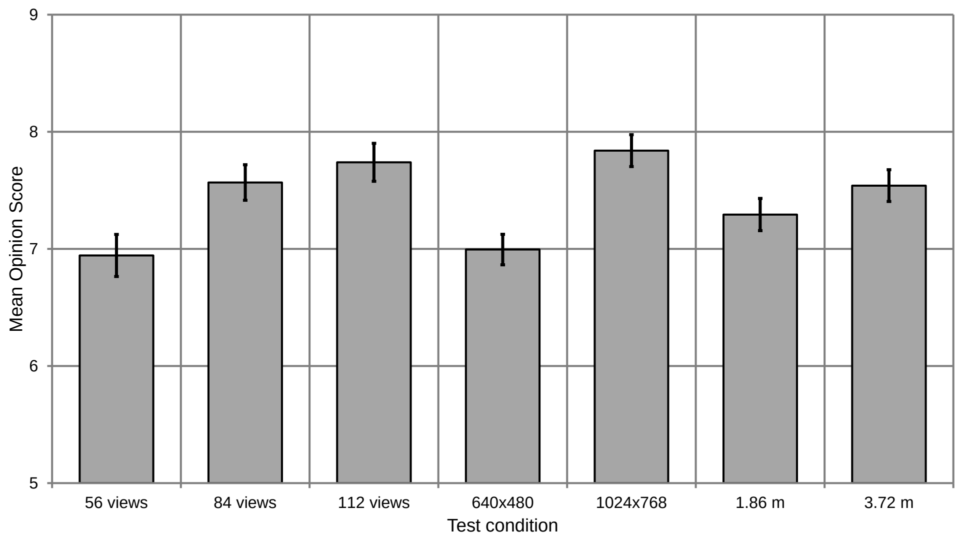


Figure 4. Mean Opinion Scores of the test conditions (collapsed for test variables) for all test participants.

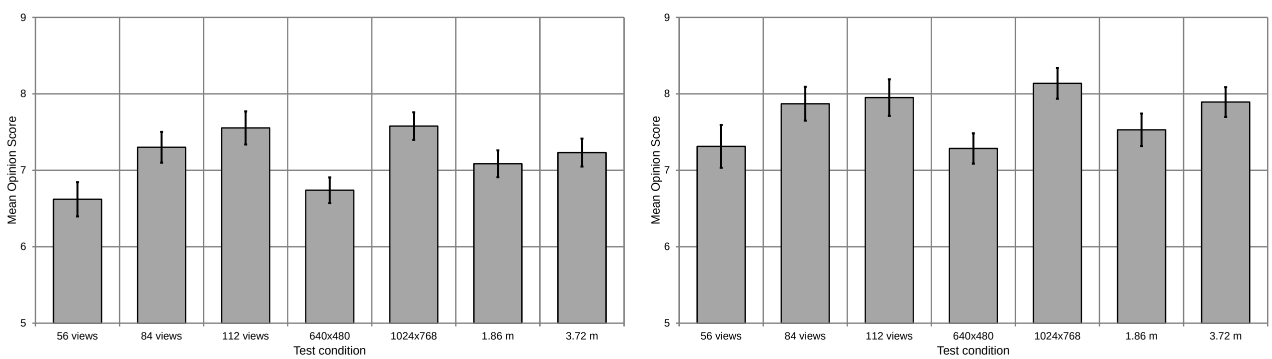


Figure 5. Mean Opinion Scores of the test conditions (collapsed for test variables) for Group 1 (left) and Group 2 (right).

The impact of each test variable is shown on Figure 4. The lowest and the highest angular resolution values show similar rating tendencies as the two values of spatial resolution. Both exhibit a statistically significant difference. There is an MOS difference of around 0.25 between the two viewing distances, which, as seen earlier, primarily originates from the scoring related to the change in spatial resolution.

The separation of these results based on grouping is shown on Figure 5. The difference between the corresponding MOS values is rather apparent. For the three angular resolutions, the two spatial resolutions and the two distances, the differences are 0.69, 0.57, 0.4, 0.55, 0.56, 0.44 and 0.66, respectively. Regarding the MOS differences among the values of a given test variable, in case of Group 1, the viewing distance had a smaller impact, and the two higher angular resolution values were distinguished more. In case of Group 2, the opposite applies.

One major limitation of the MOS analysis is that it does not report rating inconsistencies. Let us define specific ratings inconsistent, if the objectively worse representation receives a higher rating. According to this definition, if they received the same rating, then they are still consistent.

Rating inconsistency in the context of experimental validity has already been addressed in the scientific literature.[22] However, that was related to insufficient training procedures. As stated in the experimental setup, the test participants were introduced to the different visual phenomena – particularly the degradation of the parallax effect and thus the emergence of the crosstalk effect – prior to the subjective test.

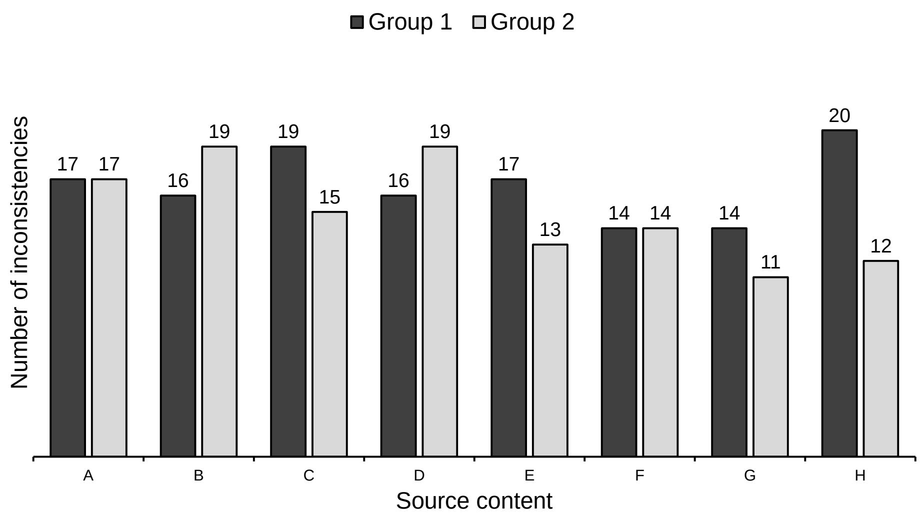


Figure 6. Number of rating inconsistencies related to angular resolution per source content.

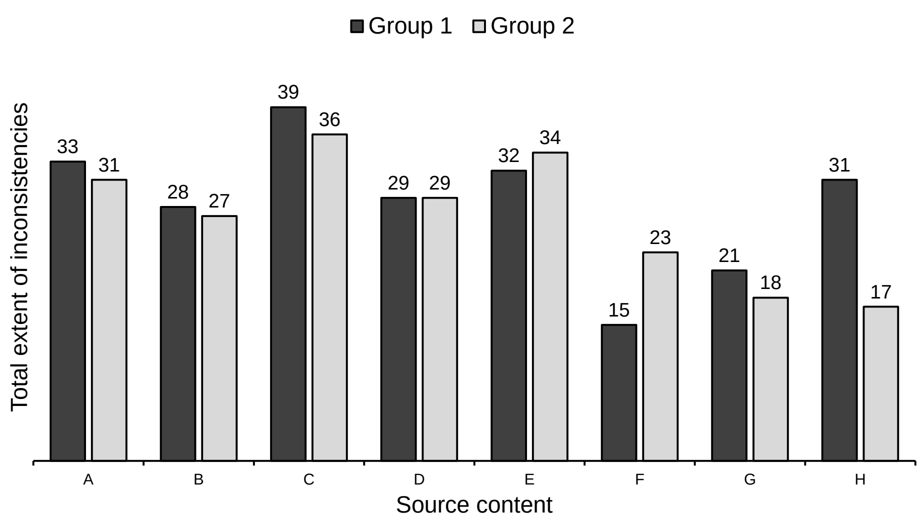


Figure 7. Total extent of rating inconsistencies related to angular resolution per source content.

In this experiment, the highest portion of rating inconsistencies applied to angular resolution. As a single test participant provided 12 quality ratings for each of the 8 source contents and there were 3 values of angular resolution, one test participant assessed $(4 \times 8 =)$ 32 stimulus triplets of angular resolution, and each source content can be characterized by $(4 \times 15 =)$ 60 triplets. For example, one triplet is 56 views / $640 \times 480$ / 1.86 m; 84 views / $640 \times 480$ / 1.86 m; 112 views / $640 \times 480$ / 1.86 m. A triplet is considered inconsistent, if a lower angular resolution receives a higher quality rating. The extent of inconsistency is the greatest inconsistent difference. For example, if the previously mentioned triplet receives the scores of 8, 7 and 8, respectively, then the extent is 1. If the scores are 8, 7 and 6, then the extent is 2 (not 3, as only the first and the third elements of the triplet are considered, since that difference is the greatest).

The number and the total extent of inconsistencies related to angular resolution per source content are shown in Figures 6 and 7, respectively. For most contents, the inconsistencies do not differ between the two groups. The only major difference is for content H, which is a laser-scanned statue. The most likely explanation is that content H has the least amount of differences in depth. The other last-scanned statue, content G, has an arm reaching out, which may provide a solid visual indication for the changes in the smoothness of the parallax effect.

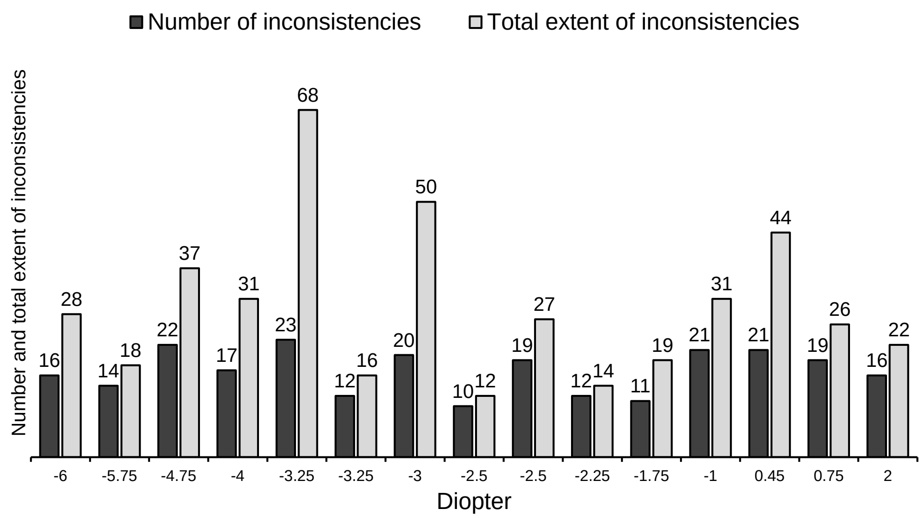


Figure 8. Number and total extent of rating inconsistencies related to angular resolution per diopter.

It should be noted that for most contents, the number of inconsistencies is around 15–20. Again, each source content can be characterized by 60 triplets, which means that for many contents, nearly a third of the assessment of the stimulus triplets of angular resolution was inconsistent.

Inconsistency can also be approached based on the diopter values of the test participants. The results of the analysis is shown in Figure 8. The distribution of the inconsistencies indicate that the diopter values of the test participants did not play a notable factor regarding the number and the total extent of the rating inconsistencies. The number of inconsistencies per test participant varied between 10 and 23, and the average was about 17. This average is more than 50% of all the investigated rating triplets. Additionally, the results show that most of the test participants had inconsistency extents around 1 and 2, and that there are outliers who amassed significant extents. The average extent was 1.75, which, for the most outstanding outlier, was nearly 2.96.

## 5. CONCLUSION

In this paper, we introduced the first of our experiments on the quality of light field visualization, as perceived by individuals with imperfect visual acuity and color deficiencies. The results generally indicate that color-blind test participants assessed the visual stimuli with lower subjective scores. The impact of viewing distance and spatial resolution is highlighted, along with the high numbers and extents of rating inconsistency regarding angular resolution.

There are many research questions to address in future work. First of all, other KPIs should be studied as well, and the ones covered by this paper should be investigated in a more fine-grained approach, separately addressing interdependencies. Due to the several test variables, one variable only had two or three values in this experiment, in order to maintain a reasonable test duration. However, the usage of light field displays over extended periods of time should be studied as well, particularly regarding perceptual fatigue. Observer movement during passive use cases should be addressed, as well as the task performance of active use cases.

## ACKNOWLEDGMENTS

The research reported in this paper is part of project no. BME-NVA-02, implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021 funding scheme. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813170. Also received funding by 2018-2.1.3-EUREKA-2018-00007 and KFI 16-1-2017-0015, NRDI Fund, Hungary. The authors would like to thank all the individuals who participated in the experiment and thus made this work possible.

## REFERENCES


[1] Kara, P. A., Tamboli, R. R., Doronin, O., Cserkaszky, A., Barsi, A., Nagy, Z., Martini, M. G., and Simon, A., "The key performance indicators of projection-based light field visualization," *Journal of Information Display* (2019).

[2] Kovács, P. T., Lackner, K., Barsi, A., Balázs, Á., Boev, A., Bregović, R., and Gotchev, A., "Measurement of perceived spatial resolution in 3D light-field displays," in [*International Conference on Image Processing*], 768–772, IEEE (2014).

[3] Kovács, P. T., Boev, A., Bregović, R., and Gotchev, A., "Quality measurements of 3D light-field displays," in [*Eighth International Workshop on Video Processing and Quality Metrics for Consumer Electronics*], 1–6 (2014).

[4] Kovács, P. T., Bregović, R., Boev, A., Barsi, A., and Gotchev, A., "Quantifying spatial and angular resolution of light-field 3-D displays," *IEEE Journal of Selected Topics in Signal Processing* **11**(7), 1213–1222 (2017).

[5] Tamboli, R., Vupparaboina, K. K., Ready, J., Jana, S., and Channappayya, S., "A subjective evaluation of true 3d images," in [*International Conference on 3D Imaging (IC3D)*], 1–8 (2014).

[6] Tamboli, R. R., Appina, B., Channappayya, S., and Jana, S., "Super-multiview content with high angular resolution: 3D quality assessment on horizontal-parallax lightfield display," *Signal Processing: Image Communication* **47**, 42–55 (2016).

[7] Tamboli, R. R., Appina, B., Channappayya, S. S., and Jana, S., "Achieving high angular resolution via view synthesis: quality assessment of 3d content on super multiview lightfield display," in [*International Conference on 3D Immersion (IC3D)*], 1–8 (2017).

[8] Kara, P. A., Cserkaszky, A., Darukumalli, S., Barsi, A., and Martini, M. G., "On the edge of the seat: Reduced angular resolution of a light field cinema with fixed observer positions," in [*Ninth International Conference on Quality of Multimedia Experience (QoMEX)*], 1–6 (2017).

[9] Tamboli, R. R., Kara, P. A., Cserkaszky, A., Barsi, A., Martini, M. G., and Jana, S., "Canonical 3d object orientation for interactive light-field visualization," in [*Applications of Digital Image Processing XLI*], **10752**, 77–83, International Society for Optics and Photonics (2018).

[10] Cserkaszky, A., Barsi, A., Kara, P. A., and Martini, M. G., "To interpolate or not to interpolate: Subjective assessment of interpolation performance on a light field display," in [*IEEE International Conference on Multimedia & Expo (ICME) Workshops*], 55–60 (2017).

[11] Cserkaszky, A., Kara, P. A., Tamboli, R. R., Barsi, A., Martini, M. G., Bokor, L., and Balogh, T., "Angularly continuous light-field format: Concept, implementation, and evaluation," *Journal of the Society for Information Display* **27**(7), 442–461 (2019).

[12] Zhang, X., Braley, S., Rubens, C., Merritt, T., and Vertegaal, R., "Lightbee: A self-levitating light field display for hologrammatic telepresence," in [*CHI Conference on Human Factors in Computing Systems*], 1–10 (2019).

[13] Cserkaszky, A., Barsi, A., Nagy, Z., Puhr, G., Balogh, T., and Kara, P. A., "Real-time light-field 3d telepresence," in [*7th European Workshop on Visual Information Processing (EUVIP)*], 1–5, IEEE (2018).

[14] Darukumalli, S., Kara, P. A., Barsi, A., Martini, M. G., and Balogh, T., "Subjective quality assessment of zooming levels and image reconstructions based on region of interest for light field displays," in [*International Conference on 3D Imaging (IC3D)*], 1–6, IEEE (2016).

[15] Darukumalli, S., Kara, P. A., Barsi, A., Martini, M. G., Balogh, T., and Chehaibi, A., "Performance comparison of subjective assessment methodologies for light field displays," in [*International Symposium on Signal Processing and Information Technology*], 28–33, IEEE (2016).

[16] Kara, P. A., Kovács, P. T., Martini, M. G., Barsi, A., Lackner, K., and Balogh, T., "From a different point of view: How the field of view of light field displays affects the willingness to pay and to use," in [*8th International Conference on Quality of Multimedia Experience (QoMEX)*], IEEE (2016).

[17] Kara, P. A., Tamboli, R. R., Cserkaszky, A., Martini, M. G., Barsi, A., and Bokor, L., "The viewing conditions of light-field video for subjective quality assessment," in [*International Conference on 3D Immersion (IC3D)*], 1–8 (2018).

[18] Kara, P. A., Cserkaszky, A., Barsi, A., Papp, T., Martini, M. G., and Bokor, L., "The interdependence of spatial and angular resolution in the quality of experience of light field visualization," in [*International Conference on 3D Immersion (IC3D)*], 1–8 (2017).

[19] Kara, P. A., Cserkaszky, A., Martini, M. G., Barsi, A., Bokor, L., and Balogh, T., "Evaluation of the concept of dynamic adaptive streaming of light field video," *IEEE Transactions on Broadcasting* **64**(2), 407–421 (2018).

[20] Kara, P. A., Tamboli, R. R., Cserkaszky, A., Barsi, A., Simon, A., Kusz, A., Bokor, L., and Martini, M. G., "Objective and subjective assessment of binocular disparity for projection-based light field displays," in [*International Conference on 3D Immersion (IC3D)*], 1–8 (2019).

[21] Kara, P. A., Barsi, A., Tamboli, R. R., Guindy, M., Martini, M. G., Balogh, T., and Simon, A., "Recommendations on the viewing distance of light field displays," in [*Digital Optical Technologies 2021*], **11788**, International Society for Optics and Photonics (2021).

[22] Kara, P. A., Martini, M. G., Kovács, P. T., Imre, S., Barsi, A., Lackner, K., and Balogh, T., "Perceived quality of angular resolution for light field displays and the validity of subjective assessment," in [*International Conference on 3D Imaging (IC3D)*], 1–7 (2016).