How to Rate the New Glasses-free Experience: Subjective Quality Assessment Methodologies for Experiments on Light Field Displays

Subbareddy Darukumalli*†, Peter A. Kara‡, Attila Barsi*, Maria G. Martini‡
*Holografika, Budapest, Hungary
Email: {s.darukumalli, a.barsi}@holografika.com
†Pazmany Peter Catholic University, Budapest, Hungary
Email: darukumalli.subbareddy@itk.ppke.hu
‡WMN Research Group, Kingston University, London, UK
Email: {p.kara, m.martini}@kingston.ac.uk

***Abstract*—Novel forms of 3D visualization are rising, where users no longer need to rely on additional equipment, such as special glasses or head gears, in order to properly perceive the visualized content. One particularly notable type of these systems are light field displays. User experience on such glasses-free displays is just as important as on any visualization technology, and thus researchers conduct experiments to evaluate it. While there are standardized techniques for conventional 2D and stereoscopic 3D displays, at the time of this paper, there are no standards for measuring the user's Quality of Experience on these displays. In this paper, we discuss the current state of experimental methodologies used to assess subjective perceived quality on light field displays. The discussion also calls attention to the special considerations which are typical for such 3D visualizations but not present in different systems.**

Keywords*—Quality of Experience; Light Field Display; Perceived quality; Subjective quality assessment methodologies; Experimental test conditions; Subjective rating scales; Viewing distance; Parallax effect

I. Introduction

Quality of Experience (QoE) in the area of telecommunications and multimedia has been intensely studied in the past decades, as it has become clear that the delight of the user and the fulfillment of the user's expectations are vital measures of quality and value. The countless studies performed and papers published on this topic enabled the scientific community to understand many of the reasons behind subjective evaluations of quality, and to model and predict QoE, making quality assessment effective and cost-efficient. Such experiments use standardized subjective quality assessment methods, serving several purposes, e.g., the reduction of experimental biases.

As new technologies emerge, new standards are developed as well, taking into consideration the properties of the visualization technologies – in case we talk about visual quality – and every major element that can affect the experience of the user. Standardized techniques are also beneficial because they support the reproducibility of researches, they make different experimental results easily comparable, increase the reliability of conclusions on contributions and thus the overall "value" of the research.

As an example, the stereoscopic 3D experience has become commercially available in the recent years, and researches targeting this technology can rely on the corresponding standard [1]. 3D visualization is also possible on glasses-free systems – such as light field displays – and there is a notable research effort in the community, which shall enable this technology to enter the consumer market.

However, at the time of this paper, there are no standardized techniques to measure QoE on light field displays. Current experiments tend to use the methodologies of conventional 2D displays and/or stereoscopic 3D ones, while glasses-free 3D visualization has certain attributes and visual phenomena that are not present on different systems.

This position paper intends to start a discussion on the subjective quality assessment methodologies for the experiments carried out on light field displays. As the continuous incoming stream of technological researches and developments on such displays are advancing this form of 3D visualization forward, more and more experiments are being carried out on their perceived quality. However, the methodologies for evaluating these advances are not standardized yet and the topic of methodologies is not a frequent theme of scientific disseminations of knowledge regarding light field visualization. The discussion in this paper is built around the different challenges and open questions regarding the subjective rating techniques, and focuses on considerations that are unique to such systems.

The remainder of the paper is structured as follows: An overview of the related standards and recommendations currently in force are introduced in Section II, along with the methodologies used in subjective quality evaluations on light field systems. Section III is the discussion on the topic of methodologies, pointing out current and future issues. The paper is concluded in Section IV.

II. Related Work

A. Standards and Recommendations

The first standard for stereoscopic television picture [2] was proposed in the year 2000. In addition to the known factors of monoscopic quality assessment, the depth resolution, the depth motion, the puppet theater effect, and the cardboard effect were considered as well. However, perceptual quality related issues – such as general viewing conditions, selection of test materials and visual fatigue – were not addressed. This was followed by a standard in 2015 [1], that in addition to the previously included aspects of quality, also addressed the perceptional dimensions, such as picture quality, depth quality and visual comfort. The ITU-R BT.500 recommendation [3] was extended with methodologies that were developed to evaluate the perceptual quality of stereoscopic televisions.

The rapid growth of 3D visualization technologies in the recent years resulted in standards proposed in 2016 [4]–[6], which mainly focus on 3D video quality assessment. The selection of the appropriate 3D displays and environmental constraints for the subjective evaluation of stereoscopic video content is proposed in ITU-T P.914 [4]. ITU-T P.915 [5] is the recommendation on the subjective assessment of video quality, depth quality, naturalness, visual discomfort, Quality of Experience (QoE), viewing experience and presence. The main objective of ITU-T P.916 [6] is to address the visual fatigue symptoms during the subjective assessment of stereoscopic video content, such as accommodation-vergence conflict [7], binocular disparity, binocular depth cues, cardboard effect, depth cues, depth distortion, depth motion, depth resolution, frame effect, medical signs, medical symptoms, planar motion, puppet theater effect, size distortion, monocular depth cues, visual comfort zone, visual discomfort, visual fatigue and window violation. This recommendation does not include objective assessment methods of 3D visual fatigue.

At the time of this paper, all the available 3D standards on subjective quality assessment address stereoscopic display technologies. Currently there is no ITU standard that addresses the subjective evaluation of autostereoscopic light field displays. The Information Display Measurements Standard (IDMS) produced by the International Committee for Display Metrology (ICDM) [8] includes considerations regarding two-view and multiview autostereoscopic displays and light field displays, such as the crosstalk effect, viewing distance, viewing angle, valid viewing area, 3D geometry distortion, image resolution and angular resolution. However, the primary approach of the standard is rather objective quality metrics than perceived quality.

B. *Subjective Quality Experiments*

The work of Kovacs *et al.* [9] introduces measurement methods for spatial, angular and depth resolution for light field displays. Spatial resolution is measured in circles per degree, angular resolution represents the number of unique directions that can be emitted from a point and measured in circles per degree, and depth resolution shows the minimum distinguishable depth that can be reproduced by the display used. The measurement of spatial resolution involves displaying sinusoidal black and white patterns of increasing frequencies on the screen. Photos are taken of the resulting images and these photos are analyzed in the frequency domain to determine the limits of visibility. The angular resolution of the display can be measured in number of views the display can generate throughout its Field of View (FOV). The perceptible depth resolution is measured in terms of absolute distance values as provided to the rendering engine, and these values are converted to relative values (relative to the total depth range provided by the display).

The impact of compression on the perceived quality of light field video is presented in the work of Dricot *et al.* [10]. Objective quality metrics are also proposed in the paper, which were validated through an experiment of subjective quality, involving 16 test participants. The tests were carried out on a HoloVizio C80 light field cinema system [11], displaying content particularly rendered in an angular resolution of 2 images per degree for this experiment. The Double-Stimulus Impairment Scale (DSIS) subjective evaluation method was used to rate the visual quality. The observer was first shown an unimpaired (uncompressed) reference content, followed by the same content with impairments (compression artifacts). The viewing distance was approximately 6 meters, which is around 3.3H, given the 1.8 meters display height. The display was calibrated for a 40-degree FOV. The total duration was around 30 minutes (as recommended by ITU-R BT.500 [3]), and the stimulus length was 5 to 6 seconds (in accordance with Chen *et al.* [12]). The findings indicate that the DSIS evaluation results were not affected by artifacts that were not related to compression or synthesis.

A 3D full-reference objective quality metric is proposed by combining the spatial information from each constituent image and the angular information from consecutive images for each 3D view in the work of Tamboli *et al.* [13]. The subjective evaluation was designed for a Holovizio HV721RC light field display [14] and the content was created using the turntable approach, with an angular resolution of 1 view per degree. The content was presented in five viewing locations separated by 12.5 degrees in the FOV and each of the 20 test participants scored the content from these five locations for each visual stimulus. In total, 168 stimuli were rated in a random order on a 5-point Absolute Category Rating (ACR) scale [15]. Rating was done in either 2 or 3 sessions, as the total duration was roughly 90 minutes. The viewing distance was around 2.5 meters, which corresponded to 2.8H, as the selected light field display had a screen height of 90 centimeters.

The work of Kara *et al.* [16] investigates the perceived quality of angular resolution. Stimuli were rendered in various angular resolutions, including below 1 degree, where the crosstalk effect and disturbances of horizontal parallax degraded the user experience. A total of 20 test participants rated the visual quality on an absolute scale from 1 to 10, and the test lasted 11-12 minutes, with 10 seconds of stimulus visualization duration. The experiment used a HoloVizio C80 light field cinema system [11], and the viewing distance was 4.6 meters (2.5H). In a different work of the authors [17], spatial resolutions were compared in a pair comparison, using a 5-point Degradation Category Rating (DCR) scale [15]. The total duration was 25 minutes, and the other parameters were identical to the previously introduced work.

III. Discussion

In this section of the paper, we discuss the relevant methodological properties of subjective quality assessment experiments on light field displays. We first detail the considerations regarding the assessment task itself, followed by the parameters of viewing conditions. The discussion also covers miscellaneous topics, such as visual comfort.

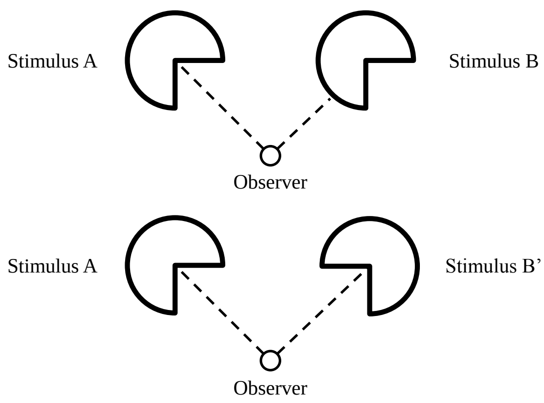

Fig. 1. Side-by-side comparison in conventional visualization (top) and mirrored (bottom). Stimulus A and B have the same content material, but they may differ in quality. The observer is represented by a pinhole camera. The dashed lines are the corresponding light rays for a forward-facing observer.

A. *Quality Assessment Task and Rating Scale*

In general, the methods of subjective data collection have always been a topic of discussion among the experts of the field of QoE. Different display technologies may demand the consideration of certain aspects that are less relevant for others – e.g., in case of stereoscopic 3D displays, the visual fatigue symptoms [6] – as among many experimental properties, the way of observation, and the physiological and cognitive demand may differ. Usually these additional considerations mean inclusion and addition, but exclusion may occur as well.

In case of light field visualization [18], there is a certain subjective test type that cannot be applied in its traditional form during the assessment of visual quality. It is the side-by-side comparison task, where two stimuli are present simultaneously; one side of the screen displays one stimulus, and the other side shows another. The core of this issue is that unlike conventional 2D and stereoscopic 3D displays, the observation of visualization depends on the position and orientation of the viewing person. Because of this, the stimuli are seen from different angles, which renders the comparison unusable.

One may state the observer movement allows the collection of complete visual data of both stimuli. That is indeed true, however, at a given point in time, when the observer is located in a certain position and looks in a certain direction, so the stimuli are still seen from different angles, and the main idea of such comparison task is to see both stimuli simultaneously under the identical circumstances.

One way to enable the observer to see the stimuli from the same angle while facing forward in a middle position it to horizontally mirror (flip) one of the stimuli (see Figure 1). The biggest problem with this approach is that one does not simply flip a stimulus. It may actually work for simple geometrical shapes, or even complex mathematical bodies, however, the perception of most visual content is affected by mirroring. The most apparent example would be a text or the logo of a brand, but i.e., even the human face is perceived differently if mirrored (unless it is perfectly symmetric). A simple way to overcome the entire issue and use the scale is by utilizing vertical (top/bottom) separation instead of horizontal. However, this can only work on horizontal-only parallax (HOP) displays, and will not be suitable for vertical and full parallax displays, that are currently being developed.

On the assessment task itself, the rating scales that are collecting scores regarding the overall user experience also need to integrate the assessment of perceptual phenomena which are not present in the evaluation of other visualization technologies, such as the smoothness of horizontal parallax. Also, certain visual quality degradations occur differently and are thus perceived differently, e.g., lower source content image (spatial) resolution materializes as blur instead of pixellation, as light rays hit irregular positions.

B. *Viewing Distance*

According to the standards of 3D visual quality [5], the viewing distance should be around 3H in general for television-like displays. For monitors, it is an interval between 1H and 3H, and 6H to 10H for handheld devices. It is not defined for displays larger than the typical television-like ones, such as the HoloVizio C80 light field cinema system [11], where the screen is 3 meters wide.

The viewing distance for projection-based light field displays can be a rather tricky topic. The first issue is that being too far from the screen of the display significantly affects the perception of content depth and the observer no longer experiences the horizontal parallax without motion, as the movements of the head and the eyes become too small compared to the angular resolution of the system. On the other hand, being too close to the screen will result in the fact that only a portion of the light rays will be correctly projected onto the eyes of the observer; here the collection of correctly projected rays also depend on the orientation of the eyes. There is also a quite trivial, yet characteristic issue with being to close to a frontal-projection-based light field display. If the distance between the observer and the screen is too small, this distance might actually be smaller than what is between the screen and the projector array. If the observer is located between the screen and the projector, then – depending on the height of the observer – the body of the observer can block the light rays coming from the projector, which then cannot be reflected from the screen, leaving the affected segment of visual information empty. As an example, the projector array of a C80 cinema system is approximately 4.6 meters away from the screen.

C. *Observer Position and Movement*

Beside the viewing distance, the position of the observer is also a quite important parameter in subjective experiments for light field displays. As it was stated earlier, the observation of the content is position-dependent. In related scientific work, there are examples when test participants have multiple discrete viewing positions [13] apart from the center view, or even a continuous viewing interval [16]. Continuous intervals evidently indicate movement, which means that the observer experiences horizontal motion parallax. Requirements regarding content angular resolution are affected by the movement of the user, as one may become more sensitive to disturbances in the smoothness of the parallax effect, such as discrete image borders between adjacent views.

The presence of observer movement also needs to be taken into consideration during the calculation of the total test duration. Walking left and right in a given interval during content observation adds an additional factor of exhaustion, which may take its toll on the performance of the test participant.

D. Miscellaneous Considerations

The visual comfort of the user is a pressing issue when it comes to 3D visualization, especially due to the short-term health effects discovered in the recent years of commercial 3D equipment usage (e.g., Oculus Rift [19]). Even though light field displays visualize a more natural 3D sensation – due to the lack of additional viewing gears and thanks to the smooth parallax effect – such displays may also result short-term and long-term medical issues. The feeling of discomfort may originate from the disturbances of parallax effect (i.e., the crosstalk effect), visualization blur due to the low spatial resolution of the content, inappropriate display calibration, and even the lack of background information can cause discomfort when the content is zoomed too far out [20]. It is possible that many other factors may cause observer discomfort as well, however, at the time of this paper, the experiments on the discomfort caused by light field display are sparse and medical studies are lacking.

The ambient light and light sources in general can also fundamentally affect subjective quality experiments performed on light field displays. It is more than enough to think about frontal-projection-based light field systems, and how their visual performance may be degraded by natural and artificial light sources. The final effect can be imaged to be similar to a conventional single projector in a room where the curtains are not shut – during daylight of course – and/or the ceiling lights are turned on.

Finally it needs to be mentioned that light field displays can be vertically parallax as well. At the time of this paper, there are no commercially available full parallax light field displays, however, it is within reasonable reach. The evaluation of the smoothness of vertical parallax would necessitate an extent of vertical movement of the human observer. Even without vertical parallax and vertical movement, the correction height of observation is critical in order to properly experience the capabilities of light field displays. As soon as vertical parallax and thus full-parallax displays emerge in subjective researches, its correct assessment will require standardized methods.

IV. Conclusion

In this paper, we presented a discussion on the quality assessment methodologies for subjective experiments on light field displays. The discussion emphasizes that while there are already subjective studies in this field of research, no standard of recommendation determines the proper parametrization of such experiments, while some of these conditions may fundamentally affect the subjective results. Future scientific work shall aim to support the discovery and analysis of experimental parameters and test conditions that may influence the new glasses-free user experience.

Acknowledgment

The work in this paper was funded from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 643072, Network QoE-Net.

References

[1] ITU-T Rec., "BT.2021: Subjective methods for the assessment of stereoscopic 3DTV systems," 2015.

[2] ITU-T Rec., "BT.1438: Subjective assessment of stereoscopic television," 2000.

[3] ITU-R Rec., "BT.500-13: Methodology for the subjective assessment of the quality of television pictures," 2012.

[4] ITU-T Rec., "P.914: Display requirements for 3D video quality assessment," 2016.

[5] ITU-T Rec., "P.915: Subjective assessment methods for 3D video quality ," 2016.

[6] ITU-T Rec., "P.916: Information and guidelines for assessing and minimizing visual discomfort and visual fatigue from 3D video," 2016.

[7] J. Kim, D. Kane, and M. S. Banks, "The rate of change of vergence–accommodation conflict affects visual discomfort," *Vision research*, vol. 105, pp. 159–165, 2014.

[8] International Committee for Display Metrology (ICDM), "Information Display Measurements Standard (IDMS)," 2012.

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

[10] A. Dricot, J. Jung, M. Cagnazzo, B. Pesquet, F. Dufaux, P. T. Kovács, and V. K. Adhikarla, "Subjective evaluation of super multi-view compressed contents on high-end light-field 3D displays," *Signal Processing: Image Communication*, vol. 39, pp. 369–385, 2015.

[11] Holovizio C80 light field cinema system, www.holografika.com/Documents/HoloVizio_C80.pdf (retrieved March 2017).

[12] W. Chen, J. Fournier, M. Barkowsky, and P. Le Callet, "New requirements of subjective video quality assessment methodologies for 3dtv," in *Video Processing and Quality Metrics 2010 (VPQM)*, 2010.

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

[14] Holovizio 721RC light field display, www.holografika.com/Documents/Holovizio_721RC-emailsize-nomovie.pdf (retrieved March 2017).

[15] ITU-T Rec., "P.910: Subjective video quality assessment methods for multimedia applications," 2008.

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

[17] P. A. Kara, P. T. Kovács, M. G. Martini, A. Barsi, K. Lackner, and T. Balogh, "Viva la resolution: The perceivable differences between image resolutions for light field displays," in *5th ISCA/DEGA Workshop on Perceptual Quality of Systems (PQS)*, 2016, pp. 107–111.

[18] I. Viola, M. Řeřábek, and T. Ebrahimi, "A new approach to subjectively assess quality of plenoptic content," in *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 2016, pp. 99 710X–99 710X.

[19] Health and safety documentation of the Oculus Rift, https://static.oculus.com/documents/310-30023-01_Rift_HealthSafety_English.pdf (retrieved March 2017).

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