

Perceptual Visual Quality of UHD-2/8K on Consumer Display

Hosub Lee*, Bill Mandel* and Pawel Aichtel**

*Samsung Research America, Irvine, CA, USA

**Aichtel Pty Limited, Rockdale, Australia

Abstract

We performed a human-subject study on the perceptual visual quality of 8K resolution on consumer display devices. With the experimental setup simulating practical media consumption scenarios via TVs, we conducted subjective picture quality comparison tests between 8K and its predecessor, 4K. Here, we considered several factors about the user (e.g., visual acuity) and their underlying context (e.g., viewing distance). Statistical analysis of the collected experimental data confirmed that people could discern 8K against 4K after watching a short video and perceive 8K to have more clarity than 4K, even on a still image. We also found that human perception of 8K's visual quality depends on user characteristics (e.g., age) and contextual (e.g., content) factors. Last, we analyzed the test video clips to extract key features contributing to the most perceivable 8K content.

Author Keywords

UHD-2/8K; resolution; picture quality; clarity; perceptual testing.

1. Introduction

UHD-2/8K denotes an image or display resolution with 7,680 horizontal and 4,320 vertical pixels, resulting in about 33 million pixels. With four times more pixel density than its predecessor, UHD-1/4K (3840×2160), the 8K can present sharper and more detailed images. In reality, however, 8K is not widespread in the media and entertainment industry mainly due to the expensive cost of capturing, storing, and delivering 8K content, contrasting unclear customer needs or benefits of 8K.

Researchers had attempted to quantify the advantages of 8K and yielded contradicting results. Smith et al. conducted perceptual visual quality tests using uncompressed 8K videos in high dynamic range (HDR) and its 4K version downscaled via a cubic filter [1]. They utilized a single 88" 8K display sequentially presenting the test content in 4K and 8K and gathered how human viewers perceive the differences. By analyzing the experimental results based on characteristics of human vision, they concluded that 8K under typical viewing conditions did not significantly improve visual experience. Shishikui et al. also adopted a sequential (toggle) method using a 55" 8K display to collect and analyze people's subjective opinions towards the original 8K images in standard dynamic range (SDR) and its 4K variant downscaled from 8K through a low-pass filter [2]. Unlike [1], they stated that 8K indeed enhanced the viewer's high-level impressions (e.g., being-there) in a practical TV viewing distance.

Both abovementioned works used the same sequential method, visually amplifying a specific region with continuous shifts between 4K and 8K so the viewer can quickly notice it. As this method requires short clips to be displayed back and forth, the test subject must memorize the previous clip without having enough time to check the details under different resolutions. This aspect can prevent the researchers from getting reliable subjective responses about 8K perception [3]. In addition, prior research assumed ideal or impractical media consumption scenarios. For example, [1] conducted an experiment involving uncompressed 8K playback on a TV. This required a dedicated workstation with

a GPU to play native 8K content and transmit it to the TV using an 8K HDMI cable. While this setup would achieve true 8K resolution, in most cases, content is delivered in a compressed format through streaming apps like YouTube. Thus, previous research experiments may not accurately reflect real-world usage scenarios. Last, qualitative feedback from the subjects and the different characteristics of test content have yet to be fully investigated. Most works mainly focused on quantitative responses (i.e., numerical scores) collected during the experiment [4-6], limiting the explainability of their findings.

This paper aims to tackle the abovementioned difficulties by performing user-centric, perceptual visual quality comparison tests between 4K and 8K. To begin with, we adopted a simultaneous (side-by-side) method that displays the 4K content on the 4K display and 8K content on the 8K display separately. By simultaneously presenting the same content mastered in 4K and 8K on the corresponding resolution but same-sized (85") TVs, we intended to give test subjects more time to thoroughly and confidently evaluate their perceived image quality shown in 8K compared to 4K. We also prepared the test content in a HEVC-encoded format, a widely used video compression technique by streaming services, to simulate more practical media consumption scenarios through a TV at home. Moreover, we not only interviewed study participants to understand the reasoning behind their responses, but also analyzed the sharpness level of test content in a spatial frequency domain, thereby providing a clearer interpretation of the experimental results.

2. 4K/8K Comparison Test

Setup: Our study conducted a side-by-side comparison of 85" 4K and 8K TVs, both featuring the same quantum dot light-emitting diode (QLED) panel (4K: Samsung QN90B, 8K: Samsung QN900B) to accurately gauge how individuals perceive 8K. Previous studies employed the sequential method, in which different quality test content is displayed and evaluated successively on a single display. We tried the sequential method in our pilot study but discovered it to be unsuitable for comparing 4K and 8K quality, especially for videos, as participants struggled to remember previous video clips and make comparisons with the current clip. We then installed two TVs vertically on a stand, with the 4K TV at the bottom and the 8K TV at the top. We adjusted the perpendicular angle of each TV so that the top and bottom screens are slightly tilted towards the subject's eye to ensure proper vertical viewing angle (see Figure 1-(a)). Then, we configured Movie Picture Mode on both TVs as it is ideal for showing the original image quality without post-processing like AI upscaling. We calibrated the picture settings on both TVs to match their colors as much as possible, thereby ruling out the panel difference from factors impacting the experiment.

Next, we placed two vertically stacked TVs in a dedicated room. Here, we marked the viewing distance on the floor to guide viewers to compare 4K and 8K screens at a specific position (see green stickers in Figure 1-(b)). Referring to the literature [1-2], we adopted three viewing distances: 0.8 screen height (H), 1.6H,

and 3.2H. Considering H (42") of the test TV, we labeled these distances as 33.6" (Near), 67.2" (Medium), and 134.4" (Far). The horizontal viewing angles for Near, Medium, and Far distances were approximately 100, 60, and 30 degrees, respectively. No issues with viewing angles were identified during the pilot study. We also controlled lighting to experiment in a dark environment for which Movie Mode is optimized (see Figure 1-(c)).

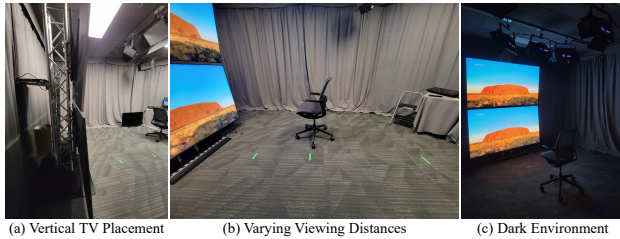


Figure 1. Experiment setup

Regarding the content for the experiment, we utilized the video clips captured by the 9×7 Digital Cinema Camera, which has a 9.3K horizontal resolution. To create 8K content, we cropped the original 9.3K content in 8K (7680×4320). For its 4K variant, we downsampled 8K content into 4K (3840×2160) using bicubic filters through DaVinci Resolve. Initially, we had the content in DNxHR-encoded MOV format (maximum content light level 1000 nits), which we converted to MP4 using HEVC encoding along with HDR10+ metadata. For the HEVC encoding, we used the hev1 codec with a maximum bitrate of 52.2 Mbps for 4K video files and 74.8 Mbps for 8K video files. We chose five video clips for the experiment after reviewing numerous options, guided by relevant literature. [7] indicated that human faces and nature scenes are effective in assessing human-perceived image quality. Additionally, [1] utilized animal scenes, which feature many high-frequency details (e.g., hair), to gauge human perception of 8K resolution. As a result, the selected contents are mainly categorized into scenery, animal fur, and human faces, and each clip is approximately 7 to 10 seconds long (see Figure 2).



Figure 2. Content

Procedure: The experiment comprises four parts: 1) participant information collection, 2) vision test, 3) 4K/8K comparison, and 4) post-interview. First, we informed each participant about the experimental procedures. To prevent preconceptions among the subjects, we intentionally refrained from using specific terms such as 8K or resolution to obscure the focus of the study. Instead, we opted for a more general term, visual quality. Once we obtained their informed consent, we collected their information: email address, background knowledge, age group, gender, and country of origin. Additionally, we measured their (corrected) visual acuity using the pocket-sized Snellen eyechart in a well-lit place. The test distance was 6 feet.

After moving to the experiment room, each subject sat at the chair at one designated viewing distance (e.g., 1.6H) and watched the same content displayed on 4K and 8K displays. For each content, we asked two questions: one for immediate impression after watching a short video and another for scored assessment on the paused video. With the first question, we forced them to choose the screen perceived to have more clarity than the other without

many cognitive procedures (namely, 8K perceptibility). For the second one, we allowed them more time to inspect the paused video, thereby more accurately quantifying their clarity perception towards 8K (namely, 8K perceived quality). Note that neither of the two TVs displayed any labeling or indication of whether the panel/content is in 4K or 8K, thus preventing any implicitly biased subject responses. The experiment was repeated for all available content at different viewing distances (e.g., 0.8H, 3.2H). Figure 3 shows the abovementioned protocol. Following this, we conducted a brief interview with the subject about their experience.

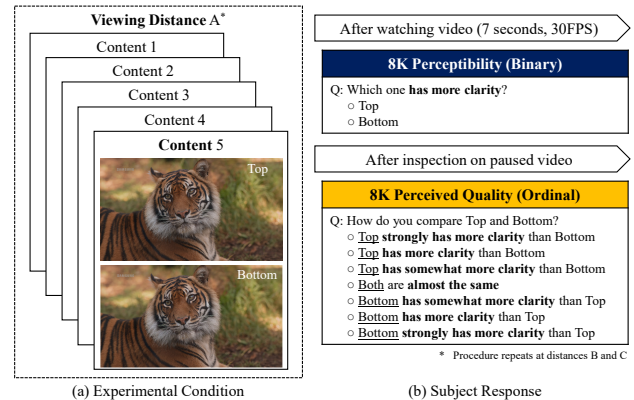


Figure 3. Experiment procedure

Experiment Data: We recruited 40 subjects in our research lab and experimented from 11/28/2023 to 02/01/2024 (two months). Most subjects were male (75%), aged 35–44 (40%), with normal 20/20 vision (87.5%), and their home country was in Asia (70%). Half of them were domain experts, while the other half were not. Considering these observer characteristics and experimental conditions (viewing distance, content), we have seven factors to consider when analyzing the experiment data. As each of the 40 subjects responded to two questions (8K perceptibility and perceived quality) towards five different content under three viewing distances, our data contains 1,200 responses in total.

3. Experiment Results

8K Perceptibility and Perceived Quality: We established a hypothesis about the human-perceived visual quality of 8K and tested whether this statement is true. As mentioned, 40 subjects participated in this experiment. Before validating our hypothesis, we performed a power analysis to check whether our sample size was large enough to detect the differences between the sample and theoretical means. For the binary response (8K perceptibility), we computed the required sample size (n) as 40.02, with the hypothesized effect size (d) of 0.4 and desired statistical power of 0.8, at a significance level of 0.05. We obtained similar results for the 7pt-scaled response (8K perceived quality); the required n was 29.78, assuming d was 0.46 and power was 0.8, at a significance level of 0.05. Thus, our sample size ($N = 40$) was sufficient for drawing valid implications from the experiment [8].

Next, we performed a series of one-sample, one-tailed (directional) t-tests. We configured the null and alternative hypotheses (H_0 and H_a , respectively) as below:

- H_0 : The sample mean response is at most the theoretical mean
- H_a : The sample mean response exceeds the theoretical mean

Regarding 8K perceptibility (0: 4K has more clarity, 1: 8K has more clarity), we set the theoretical baseline mean (μ) as 0.5,

assuming no perceivable differences between 4K and 8K; therefore, H_0 and H_a mean that 8K is not discernable with 4K regarding clarity and vice versa, respectively. A t-test suggested rejecting H_0 ($t = 19.34, p < 0.001$) in that the average clarity score given by the study subjects ($\bar{x} = 0.81, sd = 0.39$) was more significant than the baseline score ($\mu = 0.5$). For 8K perceived quality (1: 4K strongly has more clarity, ..., 4: 4K and 8K are almost the same, ..., 7: 8K strongly has more clarity), we configured the null and alternative hypotheses as same as abovementioned, except for the theoretical mean value (here, $\mu = 4.0$). A t-test also demonstrated a statistically reliable difference ($t = 18.38, p < 0.001$) between the collected clarity scores ($\bar{x} = 4.94, sd = 1.25$) and the hypothesized mean score ($\mu = 4.0$).

Considering the t-test results, we confirmed that human observers could immediately perceive 8K against 4K after watching a short video. After carefully reviewing the paused video, they could also evaluate 8K's visual quality to have more clarity than 4K. The qualitative feedback from the subjects also supported these findings (e.g., "Top is mostly better" [S15]). However, the mean difference in the 7pt-scaled clarity score between 4K and 8K was not huge ($|\bar{x} - \mu| = 0.94$). For example, a subject S22 mentioned, "Two screens are similar overall".

Factors impacting 8K Perceptual Quality: Next, we aimed to understand how two or more factors (e.g., content, age) affect the human-perceived visual quality of 8K and whether or not there are interaction effects between the factors. We adopted factorial analysis of variance (ANOVA) based on the general linear model (GLM) to estimate how the mean for the clarity scores deviates depending on the factors. To assess 8K perceptibility, we initially developed a GLM with the response variable being the binary scores provided by the subjects. The independent variables include all seven factors and their interactions (e.g., content and distance). After generating the ANOVA table for the fitted GLM, we identified the statistically significant factors that impact the results and accordingly constructed a simpler, more parsimonious model as below:

$$Y = \beta_0 + \beta_1(\text{age}) + \beta_2(\text{content}) + \beta_3(\text{distance} * \text{vision}) + \varepsilon \quad (1)$$

where Y is the dichotomous clarity score given by subjects (0: 4K, 1: 8K), age and vision are the subject's age group and measured visual acuity, content is the type of video clip presented to the subject, and distance is how far the subject was seated; β_0 is an intercept, β_p is a regression coefficient of the p^{th} independent variables, and ε is an unknown error term.

We then analyzed regression results to quantify the relationships between specific factors and the perceptibility of 8K resolution. Regarding content , for example, the Mountain and Tiger scenes yielded 38% and 28% higher clarity scores than the Meerkat scene (baseline), respectively ($p < 0.001$ for both). For age , older people aged 45–54 were 8% more positively sensitive to the clarity of the 8K screen over the 4K than younger people aged 25–34 ($p < 0.1$). Lastly, under the far viewing distance (3.2H), subjects with 20/20 and 20/25 vision could discern 8K from 4K 64% and 80% better than those with relatively inferior vision (20/40), respectively ($p < 0.01$ for both).

Regarding 8K perceived quality, we conducted the same analysis as for 8K perceptibility, except for the response variable Y . Y is now the ordinal clarity score (1: 4K, ..., 4: Same, ..., 7: 8K):

$$Y = \beta_0 + \beta_1(\text{age}) + \beta_2(\text{distance} * \text{content}) + \varepsilon \quad (2)$$

The regression results from the GLM (Equation 2) confirmed our previous findings about the content : test subjects gave about

26.5% higher clarity score on the Mountain scene than on the Meerkat scene ($p < 0.001$). We observed new aspects regarding age : people aged 35–44 perceived 8K's clarity about 4% less than the younger group aged 25–34 ($p < 0.05$). Regarding the distance , the subjects gave about 11.25% lower scores on the 8K screen when seated at the farthest (3.2H) position than when they did at the nearest (0.8H) ($p < 0.01$). However, the impact of viewing distance on the perceived quality of 8K was content -dependent. For the Tiger scene, for example, subjects indeed gave 12.9% and 15.8% higher clarity scores on the 8K screen at medium (1.6H) and far (3.2H) viewing distance compared to the baseline of 0.8H ($p < 0.05$ and $p < 0.01$, respectively).

The data analysis revealed that both perceptibility and perceived quality of 8K are highly context- and user-dependent. First, the content was critical when people perceived 8K's visual quality. The Mountain scene demonstrated an example of the most perceivable 8K content, while the Meerkat scene did for the least perceivable content (see Figure 4). Most subjects confirmed this (e.g., "The mountain image had an obvious clarity difference, especially in terms of color clarity and vibrance" [S02]). The viewing distance also mattered. In general, farther viewing distance lowers 8K perceptibility and perceived quality, but this also depends on the content. For some content (e.g., Tiger), people perceived 8K better at medium or far viewing distance than near distance (see Figure 5). A subject S18 stated, "Top looks better at far viewing distance".

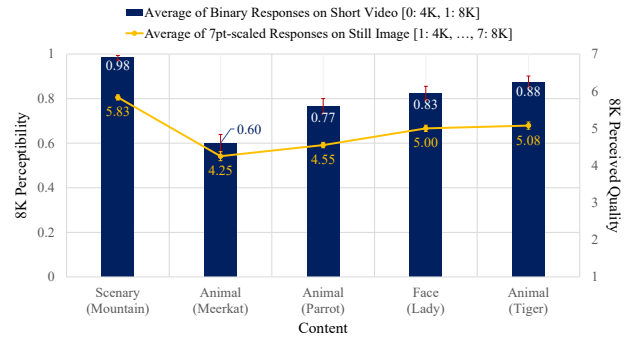


Figure 4. 4K/8K clarity comparison per content

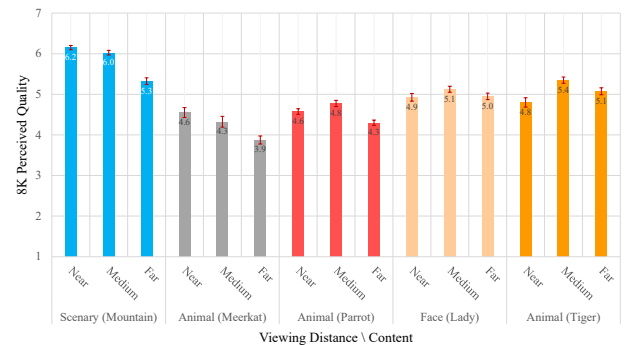


Figure 5. 4K/8K clarity comparison per content & distance

We also learned that viewers' characteristics, like age, could impact the perceived quality of 8K. The subjects aged 25–34 perceived 8K-level clarity better than those aged 35–44, but older people (45–54) were better than these younger groups. We suspect there are some physiological reasons, as addressed by previous researchers [9], for the increased sensitivity of older adults to screen sharpness. Visual acuity was a significant factor,

depending on the viewing distance. Subjects with normal vision (20/20) could perceive 8K much better than those with impaired vision at the far viewing distance. However, we noticed the opposite tendency at near or medium viewing distances.

4. 8K Content Analysis

Here, we analyzed the sharpness levels of 4K/8K content to better grasp the human perception of 8K visual quality. We performed modulation transfer function (MTF) analysis via the slanted edge method [10-11] upon select test content. This standardized method (ISO 12233) measures the MTF as a function of the spatial frequencies of image data, typically a rectangular region of interest (ROI) with a beveled edge shape. It first projects pixel intensities of the 2D edge onto the 1D line perpendicular to the target edge, building the edge spread function (ESF); the derivative of the ESF is the line spread function (LSF), which is then normalized and Fourier-transformed into the MTF curve.

We extracted a keyframe from each content in its original format. To compare 4K and 8K content on the same spatial frequency range, we upscaled the 4K content into 8K via the simplest interpolation method, Nearest Neighbor. After that, we selected and cropped the rectangular ROI from each content. We specified the ROI based on qualitative feedback collected after the experiment. For instance, study subjects stated that they focused on the edge of the main object (e.g., the boundary line between mountain and sky) when evaluating the content. We accordingly chose that area as ROI. Last, we draw MTF charts towards the ROI of 4K-upscaled-8K and original 8K content across spatial frequency (cycles/pixel) ranging from 0.0 to 1.0.

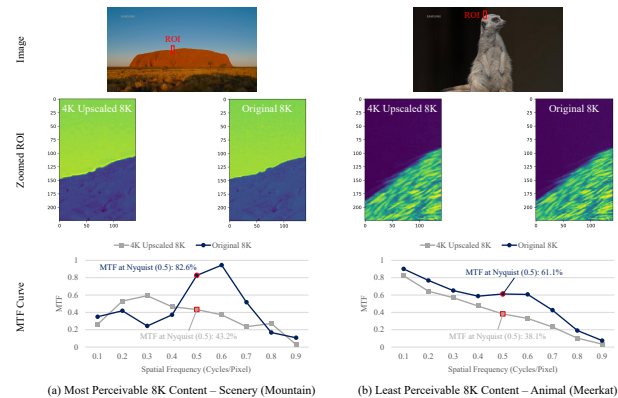


Figure 6. MTF analysis of select 4K/8K content

Figure 6 displays the results of MTF analysis on the most perceivable (Mountain) and least perceivable (Meerkat) 8K content. Regarding the Mountain content, 4K-upscaled-8K and the original 8K content showed very different MTF curves for the select ROI (see Figure 6-(a)-Zoomed ROI). Specifically, the 8K version displayed high MTF values ($> 82.6\%$) around the spatial frequency of 0.5–0.6, but it dropped afterward. The 4K variant, on the other hand, marked the highest MTF value at a low-frequency domain (0.3) and continuously degraded beyond that. Most importantly, a significant gap in MTF values (39.4% difference) between 4K and 8K exists at the Nyquist limit of 0.5 cycles per pixel (see Figure 6-(a)-MTF Curve). Considering this is the highest spatial frequency in a typical 8K content on an 8K screen [1], subjects could easily perceive the clarity differences between 4K and 8K at this point.

We observed the opposite toward the Meerkat scene. Now, the select ROI is the outline of the Meerkat head, which has a sloped

edge (see Figure 6-(b)-Zoomed ROI). Like the Mountain content, the 8K version demonstrated a high MTF value (61.1%) around Nyquist frequency, eventually converging to the lowest. However, the overall MTF curves derived from the ROI in 4K and 8K looked similar; they generally decrease as spatial frequency increases. In addition, the 4K/8K difference (23%) in MTF values at Nyquist was insignificant (see Figure 6-(b)-MTF Curve). Thereby, it would be more challenging for subjects to discern 8K from 4K on the meerkat with complex texture (hair) compared to the mountain with a plain surface (rock).

The MTF analysis revealed distinct sharpness differences in 8K content captured by the same camera, which could influence subjective clarity perceptions in each content. The sharpness of the boundary (edge) between the main object and the background seems especially critical. We presume that the complexity of the texture in an object (e.g., rock vs. hair) induces different sharpness measures, leading to content-dependent perceptions of visual quality. The Mountain content presents clear edges around the main object, resulting in a higher perception of 8K clarity. In contrast, the Meerkat scene has relatively fuzzy boundaries with fine hairs, leading to a significant portion of a neutral score.

However, sharpness (spatial frequency) is only one aspect of human visual perception. Human vision is also tuned to color contrast. In the Meerkat scene, for example, the fur is relatively low contrast (gray on gray) and, consequently, is not easily discernible by the human eye even though the actual details are there. Several subjects pointed this out (e.g., “For the scene with black/white (meerkat), it was hard to tell” [S19]). On the other hand, the Mountain and Tiger scenes have higher contrast, making it much easier to see the differences. A subject S01 said, “I see the top TV has more sharpness than the bottom one, especially on mountain edges, the tiger fur and bird leather”. Another aspect is the spatiotemporal dynamics of the content. The Mountain content was a time-lapse video, enabling test subjects to perceive its clarity without significant obstruction from the movement between frames. In contrast, the second video featured a meerkat that continuously moved its head and body, making it relatively more difficult to perceive the differences between 4K and 8K.

5. Discussion and Conclusion

Our study demonstrated that, from the user’s perspective, 8K is immediately discernable and perceived to present higher clarity levels than 4K in a living room situation. 8K resolution can increase people’s perception of clarity in images, possibly enhancing their overall quality of experience (e.g., realism [2]). We also noticed that the perceived visual quality of 8K depends on diverse factors such as content, viewing distance, and viewer characteristics. Regarding content, people can sense 8K visual quality more easily if clear boundaries exist between the main object and background on the presented image. The main object’s texture, color, and movement also mattered here. For the viewing distance, we obtained results similar to those in the literature (e.g., [12]); 8K quality is more visible at a closer distance. ITU also recommends the optimal viewing distance of an 8K display as 0.75H [13]. Yet, we found that it also depends on the content. For some content, our subjects perceived 8K the most at the medium (1.6H) distance. Therefore, the ITU could reevaluate the current guidelines, as also addressed by [1-2]. Lastly, each individual could perceive 8K differently, for example, per age. Hence, a personalized approach for 8K media consumption based on user characteristics and preferences (e.g., [14-15]) is worthwhile.

6. Acknowledgements

The authors thank McClain Nelson for preparing the experiment setup and all the subjects who faithfully participated in this study. We also thank Bonggil Bak and members of the 8K Association (8KA) for their constructive feedback on this work.

7. References

1. Smith MD, Zink M, Malik A. Tested Perceptual Difference Between UHD-1/4K and UHD-2/8K. *SMPTE Motion Imaging Journal*. 2020 Jul 8;129(6):43–51.
2. Shishikui Y, Sawahata Y. Quality of 8K Ultra-High-Definition Television Viewing Experience in Practical Viewing Conditions. *IEEE Transactions on Broadcasting*. 2021 Aug 30;68(1):2–12.
3. Carboni NL, Tombaugh JW, Dillon RF. The Effect of Simultaneous vs. Sequential Viewing of Digital Images on Comparative Judgment Performance. In *Medical Imaging III: Image Capture and Display 1989 May 8 (Vol. 1091, pp. 252–261)*. SPIE.
4. Park D, Kim Y, Park Y. Cognitive effect on image quality in full ultra-high definition (8K). *Journal of Information Display*. 2020 Apr 2;21(2):103–111.
5. Bonnineau C, Hamidouche W, Fournier J, Sidaty N, Travers JF, Déforges O. Perceptual Quality Assessment of HEVC and VVC Standards for 8K Video. *IEEE Transactions on Broadcasting*. 2022 Jan 10;68(1):246–253.
6. Keller D, Rao RR, Göring S, Raake A. The Effect of Viewing Distances on 4K and 8K HDR Video Quality Perception. In *2023 IEEE International Symposium on Multimedia (ISM) 2023 Dec 11 (pp. 123–130)*. IEEE.
7. Pei E, Lee H, Fedorovskaya E, Farnand S. Evaluation of Subjective Video Quality of Television Displays. *Electronic Imaging*. 2024 Jan 21;36:1–10.
8. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Routledge; 2013 May 13.
9. Park D, Park Y. The characteristic of image quality perception of elderly group: according to the comparison with adult group. *Journal of Information Display*. 2018 Jul 3;19(3):111–120.
10. Masaoka K, Yamashita T, Nishida Y, Sugawara M. Modified slanted-edge method and multidirectional modulation transfer function estimation. *Optics Express*. 2014 Mar 10;22(5):6040–6046.
11. Parulski K, Wueller D, Burns P, Yoshida H. Creation and Evolution of ISO 12233, the International Standard for Measuring Digital Camera Resolution. *Electronic Imaging*. 2022 Jan 16;34:1–7.
12. Keller D et al. Influence of Viewing Distances on 8K HDR Video Quality Perception. In *2023 15th International Conference on Quality of Multimedia Experience (QoMEX) 2023 Jun 20 (pp. 209–212)*. IEEE.
13. Series BT. The present state of ultra-high definition television. 2023. Available from: <https://www.itu.int/pub/R-REP-BT.2246-8-2023>
14. Lee H, Park Y. Data-driven Approaches for Discovery and Prediction of User-preferred Picture Settings on Smart TVs. In *Proceedings of the 2021 ACM International Conference on Interactive Media Experiences 2021 Jun 21 (pp. 134–143)*. ACM.
15. Lee H. Uncovering and Understanding Smart TV Users' Picture Quality Preferences via Big Data Analytics. In *2023 IEEE International Conference on Big Data (BigData) 2023 Dec 15 (pp. 1764–1773)*. IEEE.