Perceived Intra-Frame Dynamic Range in Cinema Environments

Tine Vyvey[1], dr. Elena Nunez Castellar[1,2], Dirk Maes[3], Bruno Vandevelde[3], Prof. dr. Jan Van Looy[1]

[1] Research Group for Media and ICT (iMinds-MICT), Ghent University, Ghent, Belgium
[2] Department of Data Analysis, Faculty of Psychology and Educational Sciences, Ghent University
[3] Barco Technology Centre, Barco n.v., Kortrijk, Belgium

Tine.Vyvey@UGent.be, ElenaPatricia.NunezCastellar@UGent.be, Dirk.Maes@barco.com,
Bruno.Vandevelde@barco.com, Jan.VanLooy@UGent.be

Abstract

Due to rapid technological progress in High Dynamic Range (HDR) video capture and display, perceived contrast by users of HDR technology is becoming increasingly important in the visual signal processing chain. More specifically, when considering a higher contrast and brightness range in a cinema context, it is essential to understand the limits of visual perception under specific cinema conditions and brightness levels. An experiment ($N = 36$) was conducted to test the perceivable dynamic range in a cinema environment. We investigated how the perceivable dynamic range evolves with various background levels and different strengths and shapes of local highlights. Results showed a significant effect of image composition and the average picture level on the perceivable near black dynamic range.

Index Terms: HDR, projection, cinema, contrast, measurement, visual perception

1. Introduction

High Dynamic Range (HDR) video is taking center stage as a next generation visualization technology. Various television makers announced HDR capable sets. Over-the-top (OTT) content providers, such as Netflix, are considering HDR as being potentially more important than 4K and are consequently planning HDR content delivery [1]. There is also increased interest in HDR imaging for cinema applications [2]. HDR display devices are capable of reproducing a broader dynamic range of luminosity than conventional display technology [3]. The intention is to present the human eye with a range of luminance more similar to that of everyday life. The dynamic range of the Human Visual System (HVS) is an essential parameter in the design of HDR displays. In the real world, people are capable of perceiving daylight levels of around 10^8 Cd/m² to night luminance of approximately 10^{-6} Cd/m² [4]. However, very long light to dark adaptation times of the HVS and considerations regarding viewer comfort limit the perceivable dynamic range for displays. Concerns have been raised about visual discomfort when using such extreme brightness levels in a dark cinema environment as well as the practical constraints of projector technology [5].

Other studies to date on HDR content focus principally on the performance of traditional "television monitor" products. Daly, Kunkel, Sun, Farrell, and Crum [6] conducted a small-screen viewer study of luminance limits preferences, finding a dynamic range to span six orders of magnitude. The findings of the research of Kunkel and Reinhard [7] showed that in a steady state of adaptation, the HVS is capable of distinguishing contrasts within a range of 3.7 log units, equivalent to a contrast ratio of 5000:1. However, little is currently known as to the role of environmental light. The presence of ambient light has a direct impact on the perceived dynamic range as reflections of the ambient light (direct or indirect) by the screen or display are superimposed on the actual image. Previous research [2] has pointed to the effect of scattering in the projector optics and port-window as well as light reflections by the theatre walls, ceiling and floor on the black level of the projected image. Several questions regarding the role of context in relation to user experience of HDR technology remain. Hence, we conducted a psychophysical experiment to obtain an improved understanding of perception limits of the HVS, in particular, the dynamic range in a single frame and across different frames, and the influences of environmental factors in the context of HDR cinema projection.

2. Method

2.1. Design

First, a pilot study of the cinema experiment ($N = 9$, 18-30 years old) was conducted. After the pilot study, improvements in the test setup were made and signal levels were adapted. Next, 36 subjects (18-34 years old, 54% female) participated in the main series of experiments using a within-subjects design. A special setup was installed in a cinema theatre to enable the generation of test patterns with an extended dynamic range. In front of a normal projection screen (12.6 m wide) a matt black painted plate and an outdoor LED display were installed. Both the black plate and the LED display measured 0.8 x 0.8 m and were symmetrically installed at 1.4m from the center of the screen (see Figure 1). In Study 1 and 2, subjects had to count six to ten squares with different near-black levels (each square measured 6 x 6 cm) that were randomly positioned on the black painted screen, while a white LED display was manipulated by using different strengths [0% (.04 Cd/m²) - 50% (2200 Cd/m²) - 100% (3880 Cd/m²)] and shapes [left – right – full]. In addition, we investigated whether the brightness of the projected background [0% (Study 1), 1% or .265 Cd/m², and 10% or 2.65 Cd/m² (Study 2)] influenced the number of perceived squares.

In Study 3, respondents had to count the squares with different near white levels against a white background on the LED display, which was manipulated by using different background levels (38.5 Cd/m², 190 Cd/m², 733 Cd/m², 2410 Cd/m², 4080 Cd/m²). Six to ten squares were randomly displayed on the white LED background. In addition, we investigated whether the brightness of the projected background [0% (Study 1), 1% or .265 Cd/m², and 10% or 2.65 Cd/m²]
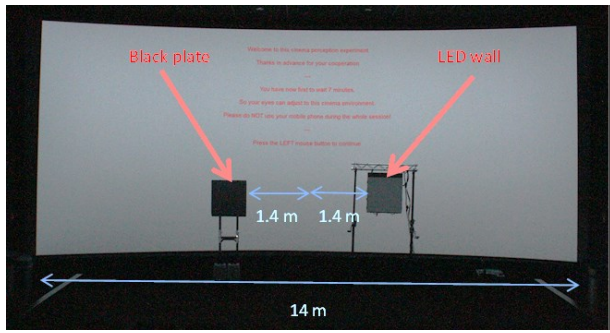
influenced the perceived number of squares.

Figure 2: *Set-up of the experiment*

2.2. Procedure

In all three studies, participants were seated in a theatre hall with a viewing distance equal to .75 times the screen width (9.5 m). The ambient light of the theatre hall was reduced to a minimum (.0002 Cd/m² measured on the projection screen with the projector off). After adapting for 7 minutes to the cinema light conditions, the participants had to go through a number of trials in which they were asked to count the squares on the black painted screen with 0% background (Study 1) and 1% and 10% background (Study 2), which together lasted around 16 minutes. Participants used the scroll bar and left button to select the number of squares they had counted whereby both selection and reaction time were logged. With a similar method, the third study was conducted were the participants had to count the squares on the white LED display. To ensure the accuracy of the answers, the experiments were counterbalanced. To exclude people with visual impairment, participants had to fill out a short online survey about their vision health and other questions about demographics, and media usage. In addition, participants had to complete a visual discomfort questionnaire before and after the experiment.

3. Results

3.1. Studies 1 and 2

3.1.1. HVS

A Repeated Measures ANOVA test was conducted to test if the different gray background levels has an influence on the perceived number of squares. The results show that there is a significant main effect of the projected background level, $F_{(72.45, 16.01)} = 40.73$, $p < .001$, $r = .48$. With a background level of 0%, the results indicated that the participants' perceivable dynamic range stretches to .0048 Cd/m². The participants reached a tipping point at 7 displayed squares, or .016% white added ($M_{6\ squares\ (sq)} = 3.02$, $SD_{6sq} = .48$, $M_{7sq} = 3.91$, $SD_{7sq} = .8$, $p < .001$, $r_{6sq\ vs.\ 7sq} = .54$). The differences starting from 7 displayed squares were not significant ($p = .576$). With a background percentage of 1%, the participants reached a tipping point at 9 displayed squares, or .033% white added to the stimuli (.009 Cd/m², $M_{6sq} = 4.23$, $SD_{6sq} = .48$, $M_{7sq} = 5.11$, $SD_{7sq} = .6$, $M_{8sq} = 5.75$, $SD_{8sq} = .71$, $M_{9sq} = 6.25$, $SD_{9sq} = .89$, $p < .001$, $r_{6sq\ vs.\ 7sq} = .88$, $r_{7sq\ vs.\ 8sq} = .78$, $r_{8sq\ vs.\ 9sq} = .66$).

With a background percentage of 10%, participants reached a tipping point at 8 squares, or .057% white added to the stimuli (.015 Cd/m², $M_{6sq} = 3.67$, $SD_{6sq} = .53$, $M_{7sq} = 4.25$, $SD_{7sq} = .55$,

$M_{8sq} = 4.77$, $SD_{8sq} = .96$, $p < .001$, $r_{6sq\ vs.\ 7sq} = .97$, $r_{7sq\ vs.\ 8sq} = .78$). Differences above 8 displayed squares, although in the right direction, were not significant, $p = .458$ (see Figure 2).
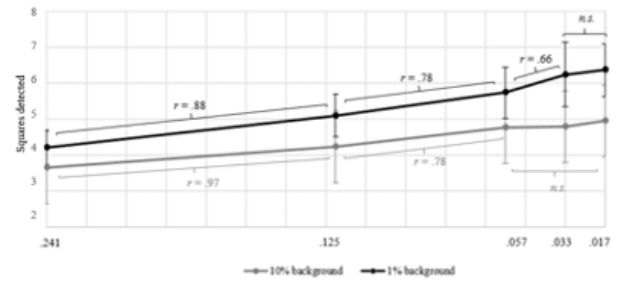

Figure 1: *Detected squares for Study 2*

3.1.2. LED shape

Mauchly's test indicated that the assumption of sphericity for a repeated measure test was met $\chi^2(2) = 1.59$, $p = .452$. A significant main effect of the size of the activated area of the LED display on the number of squares the participants detected was found, $F_{(2, 42)} = 55.28$, $p < .001$. The participants detected the least number of squares with a full LED display ($M = 4.06$, $SD = .83$, $r_{right\ vs.\ full} = .7$). The effect of the position of the activated area on the LED wall, left ($M = 4.66$, $SD = .65$) and right half ($M = 4.59$, $SD = .85$), was not significant, $p = .198$.

3.1.3. LED strength

A Greenhouse-Geiser correction ($\varepsilon = .7$), was used to adjust the degrees of freedom of the significant Mauchly's test, $\chi^2(2) = 10.99$, $p < .01$. A Repeated Measures ANOVA indicates that there is a significant main effect of the strength of the LED display on the number of detected squares. The results show that the participants detected more squares when the LED display was off ($M = 5.24$, $SD = .74$, $p < .001$, $r_{0\%\ vs.\ 50\%} = .95$), followed by 50% ($M = 4.25$, $SD = .79$, $p < .001$), and 100% ($M = 3.82$, $SD = .82$, $p < .001$, $r_{50\%\ vs.\ 100\%} = .85$), $F_{(1.41, 29.52)} = 261.49$, $p < .001$.

3.1.4. Interaction effects

The Repeated Measures ANOVA revealed a significant interaction effect between the different near black levels of the squares and the LED strengths for the different backgrounds. For the 1% background, participants reached a tipping point at 9 squares with a LED strength of 50% and a tipping point at 7 squares with a LED strength of 100%. No tipping point was found when the LED wall was turned off, $F_{(8, 120)} = 15.4$, $p < .001$ (see Figure 3). This means that on average people can detect the 4 highest gray levels or black levels down to .009 Cd/m² against a 1% gray background with the LED wall off. For a 10% background, results showed that participants reached a tipping point at 8 squares for 0%, 50% and 100% LED strength, $F_{(8, 168)} = 8.81$, $p < .001$ (see Figure 4). This means that on average, people are able to detect the three highest gray levels, or black levels down to .015 Cd/m² can be reliably detected against a 10% gray background with the LED wall off.
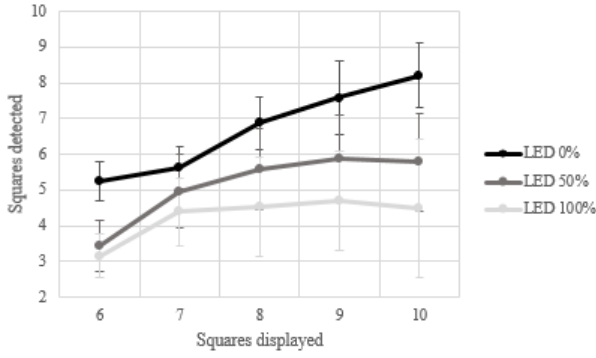
Figure 3: *Black detail with background 1%, interaction effect of LED strength and displayed squares*
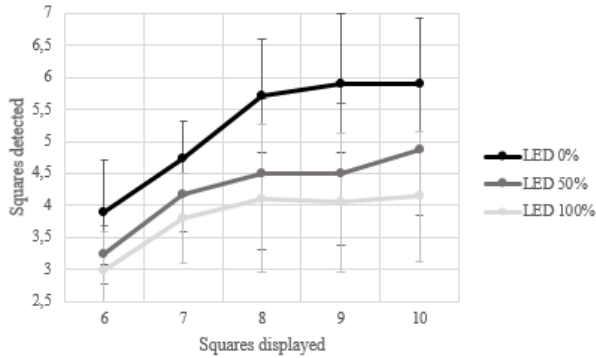

Figure 4: *Black detail with background 10%, interaction effect of squares displayed and LED strength*

For backgrounds 1% and 10%, a significant interaction effect was found between the different near black levels of the squares and the shape of the LED, $F(4, 60) = 10.61$, $p < .001$ (1% background), $F(8, 168) = 3.83$, $p < .001$ (10% background). In both cases, participants detected fewer squares with a LED wall that was fully turned on, than with a partially lit LED wall. No significant differences between the left and right LED shape were found (*infra*).

For both background percentages, results indicated a significant interaction effect between LED strength and shape $F(2, 30) = 29.15$, $p < .001$ (1% background), $F(4, 84) = 6.65$, $p < .001$ (10% background). In general, we see a decrease in detected squares when LED strength increases.

3.2. Study 3

3.2.1. HVS

A Repeated Measures ANOVA test was conducted to test whether the different white background strengths of the LED wall had an influence on the perceived number of squares. Mauchly's test indicated that the assumption of sphericity for a repeated measure test was not met $\chi^2(5) = 13.57$, $p = .017$. Therefore, we used a corrected value (Greenhouse-Geisser correction, $\varepsilon = .7$) of F. In the near white dynamic range, we find a tipping point related to the projected background level. With a background percentage of 1%, the results indicated that the participants' perceivable dynamic range spans to 8 displayed squares, or 92% intensity versus a white (100%)

background, $F(2.09, 50.2) = 345.86$, $p < .001$ ($M_{6sq} = 3.8$, $SD_{6sq} = .39$, $M_{7sq} = 4.7$, $SD_{7sq} = .29$, $M_{8sq} = 5.66$, $SD_{8sq} = .59$, $r_{6sq\ vs.\ 7sq} = .94$, $r_{7sq\ vs.\ 8sq} = .94$). This was also the case for a background percentage of 10%, $F(2.91, 58.2) = 294.64$, $p < .001$ ($M_{6sq} = 3.97$, $SD_{6sq} = .34$, $M_{7sq} = 4.91$, $SD_{7sq} = .21$, $r_{6sq\ vs.\ 7sq} = .97$, $r_{7sq\ vs.\ 8sq} = .94$).

3.2.2. LED strength

A Repeated Measures ANOVA test was conducted to assess whether different intensities of the white background and the near white levels have an influence on the perceived number of squares. Mauchly's test indicated that the assumption of sphericity for a repeated measure test was met $\chi^2(9) = 14.21$, $p = .12$. With a projected gray background of 1%, the results show that there was a significant main effect of LED strength on the perceived number of squares $F(4, 96) = 10.95$, $p < .001$. The participants saw the most squares at LED strength half ($M = 5.23$, $SD = .43$), followed by 4% ($M = 5.07$, $SD = .41$), 16 % ($M = 4.98$, $SD = .61$), 100% ($M = 4.87$, $SD = .47$), and 1% ($M = 4.75$, $SD = .52$). However, the post hoc test revealed that the differences between 1% and 100%, and between 4% and 16% LED strength were not significant, $p > .05$.

With a projected gray background of 10%, the results indicated that there was also a significant main effect of LED strength on the perceived number of squares $F(4, 80) = 9.12$, $p < .001$. As with a 1% projected background, a post hoc test indicated that participants saw the most squares at a 50% LED strength ($M = 5.48$, $SD = .4$), followed by 16% ($M = 5.43$, $SD = .29$), 100% ($M = 5.41$, $SD = .44$), 4% ($M = 5.27$, $SD = .34$), 1% ($M = 5.1$, $SD = .39$). The results revealed that only the differences between 1% and all other cases except 4% ($p < .001$), between 4% and 16% ($p < .05$), and between 4% and 50% ($p < .01$) were significant (see Figure 5).
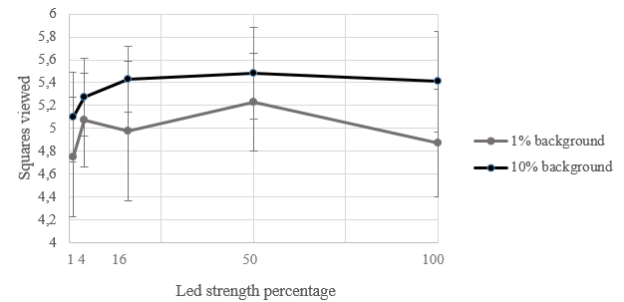

Figure 5: White detail with background 1% and 10% and LED strength

A Spearman correlation was run to assess the relationship between the relative contrast (LED strength minus the different near white levels) and the relative scores of the participants (perceived squares/displayed squares). There was a strong, positive correlation between the relative contrast and the relative scores of the participants $r(23) = .49$, $p = .012$.

3.2.3. Interaction effects

The Repeated Measures ANOVA revealed an interaction effect for background 1% and 10%. In the interaction between the near white dynamic range and LED strength, Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(77) = 141.38$, $p < .001$ (1% background), $\chi^2(135) = 191.5$, $p < .001$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity [$\varepsilon = .52$

(background 1%), $\varepsilon = .46$ (background 10%)]. As expected based on the increasing relative differences in contrast with increasing number of squares, an interaction effect was found between the different near white levels of the squares and the LED strength, $F(6.26, 150.26) = 3.88$, $p < .001$ (1% background), $F(7.36, 147.12) = 7.96$, $p < .001$ (10% background).

For the three studies, no differences in gender or the experiment order were found.

3.3. Visual discomfort

The visual discomfort scale consists of 20 items [8]. For internal consistency, Cronbach's alpha was shown to be $\alpha = .89$ for the pretest and $\alpha = .92$ for the posttest. The results of the t-test indicate that there are no significant differences between the pre- ($M = 11.72$, $SD = 2.3$) and posttest ($M = 12.72$, $SD = 2.83$), in terms of general visual discomfort, $t(35) = -1.71$, $p = .091$. However, participants did report having dry or tired eyes after the experiment due to the brightness of the LED wall.

4. Discussion

Emerging HDR technology will potentially transform the digital media industry. With the goal of informing the development of HDR imaging technology, this research presents a series of experiments to assess and understand the perception limits of the HVS in a cinema context.

The present study found that the average person can detect a black level as low as .005 Cd/m² against a background of .0008 Cd/m². Furthermore, we found that no compression of the perceivable near white dynamic range occurs up to 4000 Cd/m². This confirms the 6 orders of magnitude dynamic range found in previous studies when dark and bright elements are presented sequentially [4]. When the eye is adapted to a 1% background level, the detection threshold for the average observer increases to .009 Cd/m². When the eye is adapted to a 10% gray background level this further rises to .015 Cd/m².

In Study 1 and 2, when the LED wall is fully on, we observed that only two near black levels can be detected reliably by the average observer (with 1% background, slightly worse with a 10% background). Thus, the minimal detectable black level increases to about .032 Cd/m² for the average observer. The elevation of the background black level by reflections in the room is believed to be a major contributing factor to the increased black level detection threshold, in spite of black screen attenuation.

In Study 3, we did not see a clear effect of increasing the white intensity of the LED display from 38.5 Cd/m² to 4080 Cd/m². However, with a projected gray background 10%, we see a small gradual increase in the detected number of squares as the LED brightness increases for 1% to 50%, for 100% LED brightness we see a minor decline. This behavior is consistent with an s-shaped eye sensitivity curve that has the steepest slope in the middle of the dynamic range. With a projected gray background of 1% we see a similar trend, however with a deviation at 16%. This deviation might be explained by the fact that, because of the use of a diffuser in front of the LED wall, the white uniformity was no longer perfect, and therefore the random positioning of the squares could have had some effect on the visibility. With the 1% gray projected background, the decline in the number of detected squares at 100% LED brightness appears to be stronger, as would be expected when the eye adaptation is shifted to lower brightness levels. This

could be a first indication that when the eye is adapted to a 1% gray level, the near white dynamic range starts to compress at 4000 Cd/m². Future studies could test this with higher peak white levels and with a perfect white uniformity to confirm this.

A remarkable result is that, regardless of the concerns about the visual discomfort caused by extreme brightness levels in a dark cinema environment [5], [9], our results did not indicate that the participants experienced visual discomfort during the study. However, in the questionnaire, participants did report having dry or tired eyes after the experiment due to the brightness of the LED wall.

A weakness of the present study that needs to be mentioned is that in previous research it has been shown that the veiling glare in the HVS hinders the visibility of stimuli [10, 11]. In this study, the distance between the bright LED wall and the black screen is relatively large and therefore veiling glare is expected to have had limited impact on the results. This is confirmed by the identical detection thresholds found for left-half and right-half of the LED wall activated. The results should thus be interpreted as addressing the perceivable intra-frame dynamic range on the condition that bright and dark elements are sufficiently spatially separated. Further research that looks into the impact of the veiling glare in the human eye and its effect on the perception of HDR images is needed to learn about the perceivable intra-frame dynamic range when bright and dark elements are located in closer proximity. Finally, we only used static images in this psychophysical experiment. Other studies could try to recreate the same set-up to analyze animated content in a cinema environment.

5. Conclusions

This study conducted a psychophysical experiment to assess the impact of various background levels and different strengths and shapes of local highlights on the perceivable dynamic range. Furthermore, we studied the potential for visual fatigue on next generation HDR displays in a cinema environment. The present study found that in a movie theatre, on average, people can detect black details down to .005 Cd/m². Limited compression of the perceived near white dynamic range occurs up to 4000 Cd/m². However, the perceivable near black dynamic range is reduced substantially when dark and bright elements are combined in the same image frame. When the eye is adapted to a 1% background level, the detection threshold for the average observer increases to .009 Cd/m². When the eye is adapted to a 10% gray background level this further rises to .015 Cd/m². The presence of bright highlights adds again to this threshold. Despite our expectations, participants did report experiencing some visual discomfort during the study. The results of this study indicate that the scores in the visual discomfort scale in the posttest did not significantly differ from the scores in the pretest. With these results, this study provides new insights to support the future transition of cinema theatres to HDR projection.

6. Acknowledgements

7. References

[1] E. Alvarez. (2016, February 23). *Netflix is going all in on HDR and more original content* [Online]. Available: http://www.engadget.com/2016/02/23/netflix-hdr-original-content-mwc-2016/

[2] C. Tydtgat, D. Maes, G. Stojmenovik, and A. Grillet, "Modelling of achievable contrast and its impact on HDR projection in commercial cinema environments," *Annual Technical Conference and Exhibition, SMPTE 2015*, Hollywood, CA, 2015, pp. 1-14, Oct. 2015.

[3] A. G. Rempel, W. Heidrich, H. Li, and R. Mantiuk, "Video viewing preferences for HDR displays under varying ambient illumination," *Proc. 6th Symp. on Applied Perception in Graphics and Visualization*, New York, NY, USA, pp. 45-52, Sep. 2009.

[4] P. Ledda, L. P Santos, and A. Chalmers, "A local model of eye adaptation for high dynamic range images," *Proc. 3rd Int. Conf. on Computer graphics, virtual reality, visualisation and interaction in Africa,* New York, NY, USA, pp. 151-160, Nov. 2004.

[5] J. Houston and W. Beck, "Design considerations for cinema exhibition using RGB laser illumination," *SMPTE Motion Imaging J*, vol. 124, no. 4, pp. 26-41, May 2015.

[6] S. Daly, T. Kunkel, X. Sun, S. Farrell, and P. Crum, "Preference limits of the visual dynamic range for ultra high quality and aesthetic conveyance*," Proc. SPIE, Human Vision and Electronic Imaging XVIII*, Burlingame, CA, USA, pp. 86510J-86510J, Mar. 2013.

[7] T. Kunkel and E. Reinhard, "A reassessment of the simultaneous dynamic range of the human visual system," *Proc. 7th Symp. on Applied Perception in Graphics and Visualization*, New York, NY, USA, pp. 17-24, Jul. 2010.

[8] D. Watson, L. A Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: the PANAS scales," *J. Pers Soc Psychol.*, vol. 54, no. 6, pp. 1063. June 1988.

[9] D. Wang, T. Wang, and Y. Gong, "Stereoscopic visual fatigue assessment and modeling," *IS&T/SPIE Electronic Imaging*, pp. 901122-901122, Mar. 2014.

[10] M. Choi, D. Sharma, F. Zafar, W. C. Cheng, L. Albani, and A. Badano. "Does veiling glare in the human eye hinder detection in high-dynamic-range displays?" *J. Display Technol.*, vol. 8, no. 5, pp. 273-282, May 2012.

[11] M. J. Murdoch, I. E. Heynderickx, "Veiling glare and perceived black in high dynamic range displays," *JOSA A.,* vol. 29, no. 4, pp. 559-66, Apr. 2012.