Video Viewing Preferences for HDR Displays Under Varying Ambient Illumination

Allan G. Rempel[1,2*] Wolfgang Heidrich[1†] Hiroe Li[2‡] Rafał Mantiuk[1§]
1) The University of British Columbia, 2) Dolby Canada

Abstract

Recent high dynamic range (HDR) display devices provide significantly greater output brightness and dynamic range compared to conventional display technology. A possible concern for the extended use of HDR displays is the potential to cause visual fatigue. Furthermore, ambient illumination has a significant effect on the perception of the imagery displayed, and its impact on user preferences for brightness and contrast must be understood.

In our work we examine these issues by conducting two user studies. In each study, subjects watched video content on an HDR display in several different ambient illumination environments, and were asked to adjust the brightness and black level of the display to their preference. Subjects were also given questionnaires to document their observations and subjective preferences as well as any visual fatigue they may have experienced. We found that subjects experienced minimal visual fatigue, and also found statistically significant differences in preferred display settings under different ambient lighting conditions.

CR Categories: I.3.6 [COMPUTER GRAPHICS]: Methodology and Techniques—Interaction techniques.

1 Introduction

Conventionally available "low dynamic range" (LDR) display technologies have a very limited dynamic range compared to the abilities of the human visual system (HVS). This has led to the development of "high dynamic range" (HDR) technologies such as image sensors (e.g. [Acosta-Serafini et al. 2000]), video standards such as the 10-bit log H.264 (AVC), and HDR file formats (e.g. [Mantiuk et al. 2004; Ward and Simmons 2005; Mantiuk et al. 2006]). HDR displays were first demonstrated by Seetzen and co-workers [2003; 2004].

Adapting existing content for display on HDR devices is an underconstrained problem, and one that has received increased attention in recent years. Much of this work has centered around the transformation of imagery between HDR and LDR formats through tone mapping (e.g. [Pattanaik et al. 1998; Ward and Simmons 2005]) and reverse tone mapping (e.g. [Banterle et al. 2006; Rempel et al. 2007]). However, user studies are often helpful to provide validation of techniques as well as other information about how displays should be set up for optimal viewing. Recently, HDR displays have been the focus of a number of user studies (e.g. [Ledda et al. 2005; Seetzen et al. 2006; Yoshida et al. 2006; Akyüz et al. 2007], but these have only taken into consideration still images, not video, and have not considered the possibility of visual fatigue under long term use, the adaptation of the eyes to ambient light or the relationship between ambient light and user preferences for the presentation of HDR imagery.

The research presented in this paper analyzes these factors. In order for HDR displays to emerge as viable alternatives to standard televisions, visual fatigue, such as double vision or headaches, need to be explicitly ruled out even for extended viewing periods. Similarly, it is necessary to understand if higher brightness and dynamic range are preferred by users in standard television environments, and how preferences for these parameters might depend on the surroundings. To that end, we constructed the viewing environment pictured in Figure 1 to determine how users' preferences are affected by dark and bright ambient surroundings.

Figure 1: *Viewer in dark (left) and bright (right) ambient light.*

Many consumer-grade devices such as displays and camcorders have special "demo modes", in which color saturation, sound volume, and similar parameters are artificially enhanced. Such modes give the impression of better image or sound quality in side-by-side comparisons with other devices on a showroom floor, but they are ill-suited for extended use. A key goal of our work is to analyze whether enhanced brightness and contrast produce a similar "demo effect", or whether they yield sustained improvements in perceived image quality. A brief summary of our findings is that

- users tend to minimize the black level settings within the physical limits of the HDR display we used in our experiments. This behavior is independent of the ambient illumination.
- there were no signs visual fatigue in any of the subjects, even with high contrast settings and in dark environments.
- the majority of subjects prefer lower display brightness for darker environments. This dependency is sub-linear. A minority of subjects preferred close-to-maximum brightness independent of the ambient illumination levels.

The remainder of this paper is structured as follows. In Section 2 we briefly summarize the related work. We then describe the experimental design (Section 3), and discuss the results of our experiments (Section 4). We conclude with a discussion of our findings in Section 5.

---

[*]email: agr@cs.ubc.ca
[†]email: heidrich@cs.ubc.ca
[‡]email: hiroe.li@dolby.com
[§]email: mantiuk@cs.ubc.ca

2 Related Work

LDR to HDR expansion. Following the introduction of HDR displays [Seetzen et al. 2004], researchers started developing methods to improve the contrast of legacy low dynamic range images and video to make use of the improved dynamic range of HDR displays [Banterle et al. 2006; Meylan et al. 2006; Meylan et al. 2007; Rempel et al. 2007; Didyk et al. 2008]. This process can be seen as an inverse to conventional (forward) tone mapping operators (TMOs, e.g. [Reinhard et al. 2002; Pattanaik et al. 1998]). While the technical details of the reverse or inverse TMOs differ, and only some of them aim to provide a true mathematical inverse to a (forward) TMO, they share the common goal of improving image contrast without introducing disturbing artifacts. Akyüz et al. [2007] demonstrated in a user study that even a simple linear scaling of the contrast can be successful in achieving this goal.

Perceptual Studies. Other researchers have also turned their attention to user studies to evaluate some of the forward operators and to generally learn about viewers' responses to HDR imagery. Ledda et al. [2005] used an HDR display in combination with two LDR displays to determine which TMOs produced images that better represented a reference HDR image. Ashikhmin and Goyal [2006] evaluated five TMOs using real physical scenes to determine which operators produced images that were considered real, pleasing, or representative of the original scenes. More recently, Yoshida et al. [2007] and Čadík et al. [2008] each evaluated a number of TMOs using perceptual attributes such as brightness, contrast, and detail reproduction. Akyüz et al. [2007] evaluated six reverse operators on a number of images to determine which produced the most pleasing results. In their study, Seetzen et al. [2006] obtained preferred parameters for luminance, contrast, and amplitude ratio for HDR displays, while Yoshida et al. [2006] studied the preferred brightness, contrast, and color saturation characteristics of a TMO. However, many issues have remained unexplored, such as viewer preferences while watching video rather than still images, adaptation to the ambient illumination, visual fatigue from extended HDR viewing, and the relationship between ambient illumination and display brightness and contrast.

Ambient Illumination. Ambient illumination plays an important role in the the perception of visual imagery. This relationship has been analyzed in research dating back decades, notably as one of Stevens' famous psychophysical observations [1961; 1963]. Novick [1969] determined observers' preferred tone-reproduction curve gamma values on cathode-ray tube television monitors under several levels of surround luminance, using both tungsten and artificial daylight ambient illumination. He observed that as ambient illumination increases, viewers' preferred gamma value markedly decreases, but that gamma is largely independent of either the color of the ambient illumination or even whether the image is displayed in full color or green monochrome. This agreed with earlier work by Bartleson and Breneman [1967] that used slide projection. De Marsh [1972] built upon these results and determined approximate gamma values of 1.0 for a bright surround (20 fL (foot-lambert) or 69 $\mathrm{cd/m^2}$), 1.2 for a dim surround (4 fL or 14 $\mathrm{cd/m^2}$), and 1.5 for a dark surround. More recently, Devlin et al. [2006] considered the effect of ambient illumination on contrast perception for still images being shown on a cathode-ray tube monitor, and developed a function to correct for that.

With the popularization of mobile devices, renewed attention has been given to the automatic adjustment of display brightness under varying ambient illumination. Merrifield and Silverstein [1988] described a general model for the relationship between display brightness and ambient illumination. More recently, Swinkels et al. [2008] described a technique for adjusting the display brightness of a mobile device in a rapidly fluctuating ambient light environment without introducing flicker.

Adaptive Display. Several patents describe devices that modulate the output of a display (usually LCD) in response to variation in ambient illumination. Some (e.g. [Pittman 1976; Kalmanash 1993]) simply modulate the display brightness in response to ambient illumination for day or night viewing, while others have a more complex mechanism to adjust the output to compensate for reflected ambient light (e.g. [Cappels and Hernandez 1997]) or to correct for spatial non-uniformities that can be caused by a number of factors including ambient light (e.g. [Katyl et al. 2001]). However, none of these discuss high dynamic range displays.

Mantiuk et al. [2008] have developed a tone mapping operator that modulates its output based on the capabilities of the display and the ambient lighting in which the output is to be displayed. Ghosh et al. [2005] took the opposite approach by developing a system in which the ambient lighting of the room was computer-controlled and varied with the scene being displayed. An informal survey showed that participants who played a racing video game strongly preferred the dynamic lighting over static lighting. Another similar technology is the line of "Ambilight" flat-panel television displays which create varying ambient lighting by illuminating the wall behind the display based on the intensity of the displayed imagery [Philips 2008].

3 Experimental Design

We conducted two experiments to explore visual fatigue, viewer preferences, and the relationship between ambient lighting and preferred brightness and contrast. The first study was primarily designed to obtain information about visual fatigue in standard low-light settings, similar to the environments commonly used for watching television. This experiment used longer (1.5 hour) movies.

The second experiment tested a wider range of lighting configurations, including one comparable to bright office lighting. The goal of this experiment was to analyze the dependency of viewer preferences on ambient light levels. This experiment used shorter (22 minute) TV shows.

3.1 Setup

Our goal was to construct a viewing environment as realistic and free from distractions as possible. We wanted participants to come as close as possible to the experience of having a high dynamic range home theater environment in which to make their adjustments and comments.

Display. We used a Dolby DR-37P 37" prototype high dynamic range display with an LCD panel and LED backlight. We measured the maximum and minimum intensities and contrast of the display with a Minolta LS-100 luminance meter. Maximum luminance was at 4000 $\mathrm{cd/m^2}$, while the minimum was below the detection threshold of the luminance meter (0.001 $\mathrm{cd/m^2}$). The maximum simultaneous contrast (ratio between the lightest and darkest regions displayable at the same time) was approximately 90,000:1. The response curves of the individual color channels were measured and accounted for in all experiments.

Viewers sat in an armchair at a viewing distance of about 1.5 m, which is approximately the recommended HDTV viewing distance of three times the height of the display (3H) [ITU 1990]. The viewer's eyes were positioned at the center of the display in both height and width.

Acoustics. Since the prototype HDR display produces a significant amount of fan noise, acoustic damping was required for our experiments with video viewing. In the first experiment, we enclosed the display in a sound dampening box and conducted the experiment in a room designed for a Noise Criterion level of NC-25 to NC-30 [Beranek 1957]. In the (shorter) second experiment, sealed headphones were used.

Illumination. In the first experiment, ambient illumination in the room was provided by six black incandescent torchiere floor lamps standing 184 cm high. Each was fitted with a Philips "Natural Light" 150 W incandescent bulb. The lamps were positioned as shown on the left of Figure 2. The light from each of the torchiere cones shone upward and outward, providing a diffuse ambient light that did not directly illuminate either the subject or the display.

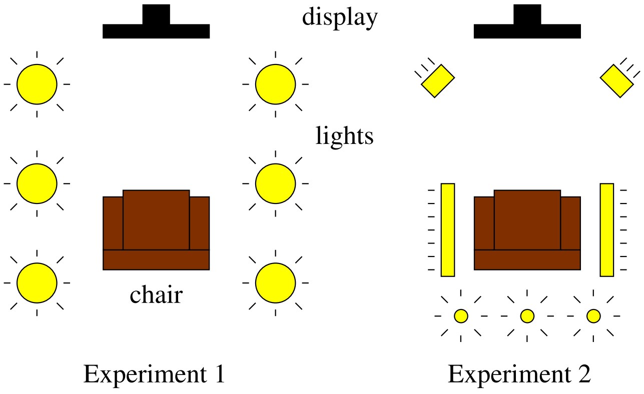

Figure 2: *Layouts of Experiments 1 (left) and 2 (right).*

Our first experiment focused on the issue of potential fatigue, and therefore we designed that experiment to use low ambient illumination levels (see Table 1). As we shall see in Section 4, these relatively small variations in ambient illumination did not result in statistically significant differences in user preferences for brightness or contrast.

In the second experiment, we increased the ambient illumination range by a factor of 10. To this end, we redesigned the lighting environment as shown on the right side of Figure 2. In addition to the torchieres from the first experiment, we used two 40 W fluorescent bulbs were positioned on the floor on either side of the armchair, and two 500 W directional halogen lamps.

Amb. light, exp. 1	<0.01(*)	0.75	8.5	28	74	
Amb. light, exp. 2	<0.01(*)				70	700

Table 1: *Ambient light levels in lux. (*) Ambient lighting was off entirely for the lowest setting.*

We measured the five different ambient light settings of the first experiment and the three settings of the second experiment. The measured levels (in lux) are shown in Table 1. For comparison, typical office lighting is approximately 500 lux [Karwowski 2001]. All measurements were taken with the HDR display switched off. The display itself contributes an average of 4-7 lux to the overall ambient illumination, depending on the specific content being shown. Figure 1 shows a viewer in the setup of Experiment 2 in both the lowest and highest ambient lighting conditions.

Video Content. All video content originated from commercial DVDs, and was adjusted for viewing on the HDR displays. For the first experiment, we used 5 feature-length movies ("Blades of Glory", "Cats & Dogs", "Charlotte's Web", "The Princess Bride", and "Zoolander"), while the second experiment used shorter TV episodes ("Friends" Season 5, episodes #11, 16, and 19).

Currently, there are only two methods for automatically and efficiently applying reverse tone mapping to large numbers of video sequences: the simple linear-luminance scaling proposed by Akyüz et al. [2007], and the more sophisticated method by Rempel et al. [2007]. Of these two methods, we chose the simple linear scaling for our experiments, for two reasons. First, Rempel et al.'s method is designed to be very conservative, and thus it rarely uses the full display intensity. This makes it difficult to use that method for analyzing brightness preferences. Second, linear scaling provides a worst case scenario for the amplification in contrast. The gradients in natural images follow a heavy tail distribution [Ruderman 1994; Dror et al. 2004]. Linear scaling of contrast shifts the whole curve, while more sophisticated methods [Rempel et al. 2007; Meylan et al. 2007; Banterle et al. 2006] have a more localized effect. For our experiment, Equation 1 shows the output intensity I, for a linearized pixel value $x \in [0 \ldots 1]$ in terms of a user selected peak brightness and black level:

$$I(x) = (\text{peak} - \text{black}) \cdot x + \text{black}. \tag{1}$$

User Controls. Instead of directly exposing the peak brightness and black level as user controlled parameters, we opted to remap these two parameters to controls that the subjects would be more familiar with as they are similar to the controls on normal televisions. The two controls we used were a "brightness" parameter b and a "contrast" parameter c, from which the peak luminance and black level were derived as follows:

$$\text{peak} = I_{max} \cdot b \qquad \text{black} = \text{peak} \cdot (1 - c). \tag{2}$$

The test subjects used a wireless USB input device to adjust brightness and contrast settings. After each adjustment, the system briefly displayed a line of text indicating one of "BRIGHTNESS" or "CONTRAST" and one of "++", "−−", "MIN", or "MAX". To avoid subjects remembering positions, subjects did not receive any absolute indication of their locations on the brightness or contrast continua.

A possible concern with a study like ours was that subjects might be distracted by the content itself, and might neglect to focus on the given task of optimizing the parameter settings. In order to counter this effect, and help subjects to find optimal settings quickly, we implemented a *staircase* procedure, as follows. At any point in time we kept track of a *minimum* and a *maximum* value for each of the two parameters (brightness and contrast), which represent the range of acceptable values for those parameters. The range would be stretched if the viewer selected values outside the range for an extended period. After 3 minutes of inactivity, the ranges were reduced in size by one half and the parameter settings were automatically changed to the opposite end of the new range, prompting the subject to make further adjustments while at the same time converging to a desired "optimal" parameter setting. An example run for a typical subject is shown in Figure 3. The squares show the automatic settings made by the staircase procedure after 3 minutes of inactivity, while the diamonds show the last value the viewer selected (i.e. the viewer's preferred value) prior to the next 3-minute period of inactivity. The sequence shows the viewer consistently returning to the same levels after the automatic settings changes.

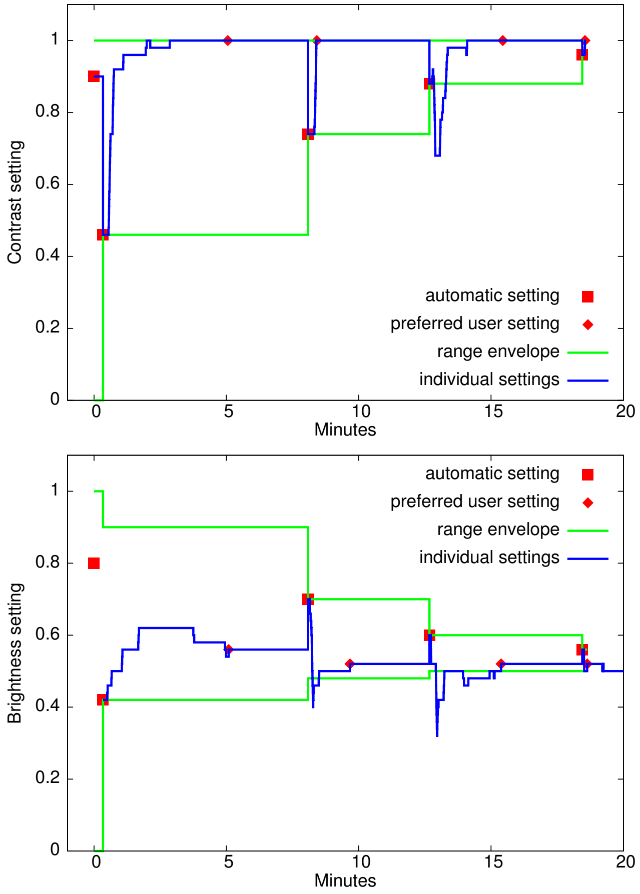

Figure 3: *Each square is an automatic setting change; each diamond is the subject's last selected point before quiescence. The darker line shows each change made by the subject.*

3.2 User Studies

Ten subjects (4 female, 6 male, aged 19–71) participated in the first experiment, and seventeen subjects (10 female, 7 male, aged 19–79) in the second. Each of them had normal (20/20) or corrected-to-normal vision and normal color vision, which was confirmed through the administration of a Snellen visual acuity test and the Ishihara [2007] color deficiency tests.

Prior to the viewings, subjects were briefed on the task to be performed and were instructed to adjust the brightness and contrast of the HDR display to a level they found pleasing. Subjects were encouraged to continue to adjust settings throughout the experiment as desired. Subjects were also told that they might occasionally find brightness and contrast changing spontaneously, in which case they were to continue to adjust brightness and contrast to their desired levels.

In the first experiment, each subject watched the same five movies of approximately 1.5 hours duration each. Each movie was seen on a different day, rather than all on the same day, to avoid introducing fatigue due to extremely long sessions. In the second experiment, each subject watched the same three 22-minute television episodes in a single session. The shows were arranged in a random order for each subject, although in the second experiment, where each ordering occurred multiple times, we stratified the samples to obtain a more uniform distribution of the orderings. For each subject, each show was seen at a different light level chosen from the set described above. The ordering of the light levels was also random, and independent of the order of the shows. No two subjects had both the same ambient light ordering and the same show ordering.

The data collected consisted of a time-stamped log of all brightness and contrast changes made during a session as well as a visual fatigue questionnaire and a more general questionnaire. The visual fatigue questionnaire asked users to identify on an 11-point Likert [1934] scale whether they experienced any of 10 symptoms of visual fatigue. These symptoms, listed in Table 2, are widely used in the human factors literature to self-report visual fatigue arising from the use of video display units [Dillon and Emurian 1995]. Likert scales are commonly used in user studies to determine the strength of a subject's perception or opinion of some factor, and they may measure intensity from "none" to "extreme" or from "strongly disagree" to "neutral" to "strongly agree." The general questionnaire consisted of both subjective questions about their experiences during the experiment and questions which asked the viewer to compare the HDR display to other types of displays on a 7-point Likert scale.

Possible symptoms of visual fatigue	Worst case score (out of 10)
Double vision	0
Problems in focusing	1
Burning/pricking sensation in the eyes	2
Blurred vision	1
Tearing/Watery eyes	2
Pain around the eyes	0
Headache	0
Image break-up	0
Image floating	0
Color change	0

Table 2: *Symptoms listed on visual fatigue questionnaire for subjects after each viewing.*

4 Analysis and Results

Visual Fatigue. Our Likert questionnaires showed that viewers experienced remarkably little visual fatigue even after the longer (1.5 hour) movie sessions. Over half the subjects selected 0 ("none") for all symptoms, and over 95% of individual responses, including those of the older participants, indicated 0. In the remaining responses where subjects indicated more than 0 visual fatigue, the average level was 1.18 out of 10.0. The total reported level of visual fatigue across all questions was 0.0325 out of 10.0. Five one-sample t-tests were conducted across the five different ambient illumination settings. The mean fatigue score was not statistically different from 0 in each case ($t(9) = 1.5$, $t(9) = 1.809$, $t(9) = 1.5$, $t(9) = 1.809$, $t(9) = 1.406$) and the low scores showed no correlation with the ambient condition. Only four of the symptoms received any scores greater than 0. Table 2 shows the highest score received for each symptom. For each of the four symptoms with scores above 0, the worst-case visual fatigue score (1 or 2 out of 10) was only reported by 1 out of 40 responses. Another set of visual fatigue scores obtained after the shorter sessions with higher ambient lighting in experiment 2 was also not statistically different from 0 ($t(16) = 1.3237$).

These visual fatigue results were obtained at the viewers' preferred brightness settings. However, we do not expect significant visual fatigue at other settings or in other viewing environments either. The factors that have been found to contribute to aesthenopia (eye strain) in the human factors literature include viewing distance, duration of use, age, gender, and lighting and glare [Dillon and Emurian 1996]. Our visual fatigue results are uniformly low in both high and low ambient illumination environments, among a demographically diverse group of subjects, for both shorter and longer session durations, and at a relatively short viewing distance that would tend to exhibit high rather than low levels of aesthenopic complaints.

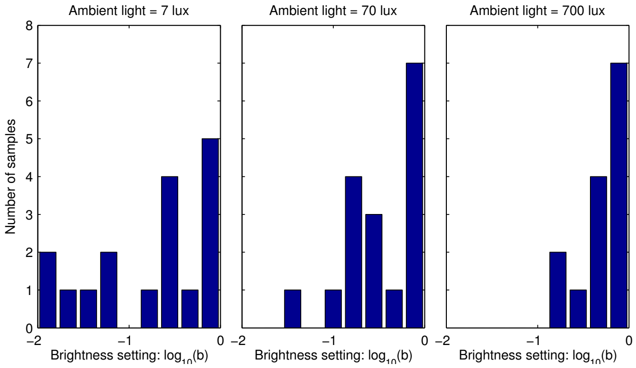

Figure 4: *Histograms of the preferred brightness settings for the three ambient light levels.*

Brightness and Black Level Preference. The adjustment patterns depicted in Figure 3, as well as similar graphs for other subjects (provided in the supplemental material) demonstrate that the staircasing procedure is effective in helping the subjects zero in on a preferred brightness setting over the course of a session. Further evidence on such a convergence is provided by the decreasing frequency of adjustments over time within each session (see below, and Figure 8). In our analysis, we therefore use the last setting for each user in each session as the preferred setting for that user in the respective ambient conditions. Two subjects had substantially different behavior patterns than the other subjects in how they set brightness during the session, so we discarded those subjects as outliers. Specifically, one subject set the brightness repeatedly to levels near the maximum under all three ambient conditions, while the other meandered significantly and even made a comment on that after the experiment. The adjustment plots for all subjects, including the outliers, are included in the supplementary materials.

The differences in the preferred brightness settings for the three ambient light levels are shown as histograms in Figure 4. As the ambient light increased, subjects tended to avoid the lower brightness settings. The histograms show significant individual differences in user preference for low ambient illumination, but a strong preference for bright screens in bright, office-like environment. The contrast preference is unaffected by the ambient illumination, with all subjects adjusting contrast to near maximal levels in all settings, effectively lowering the display black level to its minimum physically achievable value.

We also analyze aggregate preference over all subjects. Because our data fails the assumption on normal distribution and equal variance, we use the non-parametric Kruskal-Wallis test rather than ANOVA to test for the factor significance. The test indicated a significant difference in brightness settings between the three ambient light levels ($\chi^2 = 6.25$, $p < 0.05$). The episodes were selected to vary little in their overall brightness, but some differences could not be avoided (compare cumulative histograms in Figure 6). We analyzed whether adjustments may be affected by the video content. This effect was found to be just below the significance level $p = 0.05$ ($\chi^2 = 5.69$, $p = 0.058$). This suggests that the subjects could compensate with the brightness setting for the overall video brightness. We did not find any effect of gender ($\chi^2 = 0.4$, $p < 0.05$). But we found that the older group of subjects (36–79 years) choose the preferred brightness to be on average 0.3 log-10 units higher than for the younger group of subjects (19–35 years, $\chi^2 = 5.76$, $p < 0.05$).

The difference in the median brightness settings for the three ambient light levels, shown in Figure 7, was surprisingly low. If the subjects tried to fully compensate for the difference in the mean luminance of the surround, the slope of the curves in Figure 7 would be 1. This is because increasing the illuminance level by 1 log unit, results in roughly 1-log unit higher luminance of all diffuse surfaces. Meanwhile, the slope of the curve between the 70 and 700 lux setups was ≈0.26. This suggests that the large contrast between the luminance of the surround and the display content was not distracting and had moderate effect on the brightness settings.

The difference in the brightness settings can be partly explained by the increased reflectance of light from the screen and therefore loss of contrast. The light setup was designed in such a way that direct reflections of the light sources on the screen were avoided. For estimating the indirect illumination bouncing off the display, the LCD panel can be modeled as a diffuse reflector with a reflectance of about 1%. In a 700 lux environment, the luminance reflected off the screen is therefore approximately

$$0.01 \cdot \frac{700}{\pi} \frac{\text{lux}}{\text{sr}} \approx 2.2\,\text{cd}/\text{m}^2. \tag{3}$$

If the original contrast shown on the screen is $\Delta L/L$, the contrast reduced by the reflected light, $L_{reflected}$, is

$$\frac{\Delta L}{L + L_{reflected}} = \alpha \cdot \frac{\Delta L}{L}. \tag{4}$$

In Figure 5 we plot how much contrast is lost $(1-\alpha)$ due to ambient light reflection for a range of luminance values produced on the screen. The vertical lines indicate 10th and 90th percentile of the video content at the preferred brightness levels for a particular ambient light setup. Setting the display to higher brightness levels reduces the contrast loss in darker image regions. The histograms in Figure 4 also show that subjects entirely avoided lower brightness settings for the high ambient light setup. Therefore, the loss of contrast in the dark regions could prompt subjects to elevate brightness settings for higher ambient light levels.

Since our prototype display could show luminance levels almost an order of magnitude higher than a standard TV display, we could check whether current consumer-grade TV displays are bright enough to meet viewers' preferences. The peak luminance of a typical LCD-TV display usually does not exceed 500 cd/m^2. Since the median pixel value for all video frames was 47 (see Figure 6), the corresponding luminance values on a display with the 500 cd/m^2peak luminance is ≈12 cd/m^2($L_{peak} \cdot (47/255)^{2.2}$). This value is lower than the median luminance of the video content shown on our prototype display (peak 4000 cd/m^2), which was 37.3 $\text{cd}/\text{m}^2$for the darkest and 67.7 $\text{cd}/\text{m}^2$for the brightest ambient light level, as shown in Figure 7. It must be noticed, however, that there were a number of subjects selecting the highest possible brightness setting, especially for the two brighter ambient light setups (see histograms in Figure 4). We did not allow the brightness to be further increased to avoid clipping of the video content.

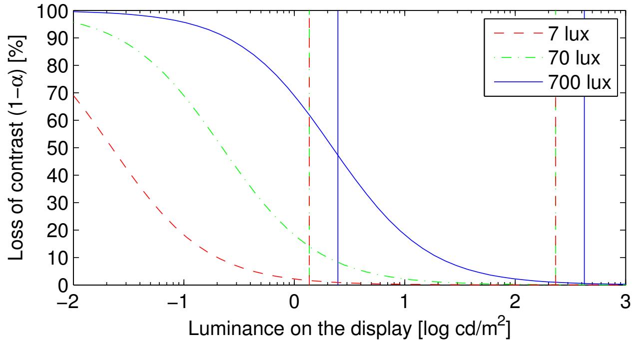

Figure 5: *Contrast loss on the display due to reflections from the panel at the three ambient illumination levels used in the second experiment. The dashed lines represent 10th and 90th percentiles of the video content for the preferred display brightness levels at a particular ambient light level (refer to Figure 7).*

Nevertheless, we can expect that a brighter display could result in preferences for even higher luminance levels. This suggests that viewers may prefer displays offering higher peak luminance levels than those offered by the majority of displays, even for relatively dim environments.

This matches the observation by Seetzen et al. [2006] that perceived image quality increases with higher display peak luminance, as long as the contrast ratio is sufficiently high. In our experiment, all subjects maximized the contrast and most selected a brightness significantly higher than what is possible on conventional displays. Seetzen et al. had high dynamic range still-image content. We can postulate that subjects in our study might have selected even higher luminance levels if the study had used true HDR content rather than adjusted LDR footage.

Both the slope and the shape of the brightness adjustment curve are very similar to that found in a different application area: cockpit displays. Merrifield and Silverstein [1988] measured preferred display brightness adjustment for an aircraft cockpit display viewed under a range of ambient illumination levels. They reported the slope of the manual brightness adjustment to be 0.276 over the photopic range of ambient illumination, but found no changes in the brightness settings over the scotopic range of ambient illumination. This remarkably resembles our results shown in Figure 7, even though our task did not involve directing viewers' attention outside the display, as was the case for the cockpit display measurements.

Adjustment Frequencies. To further analyze the reliability of our results, we also considered the frequencies of adjustments both across sessions and within a single session. Figure 8 shows the resulting statistics for our first experiment. The second experiment produced similar results. As one might expect, subjects tended to experiment more with the parameter settings in the first session than in the remaining ones. However, our experimental design forced subjects to make a significant number of adjustments (200 on average) even for subsequent sessions. Within each session, subjects required time to find suitable settings from the initial random starting point, but then quickly settled on parameters they found most suitable. This convergence shows that final settings are reliable and not scene dependent.

Comparison to Other Types of Displays. In addition to the formal experiments described above, we also asked subjects to rate the experience of HDR displays in comparison to other display technologies they had previously experienced. In cases where subjects did not know the name of the other technology, we clarified the types and their characteristics for those subjects. Subjects showed a strong preference for watching movies on the HDR display over CRT displays, and slightly weaker preferences for the HDR display over rear projection televisions and LCD displays. Six subjects were indifferent between the HDR display and a movie theater and only three preferred the movie theater.

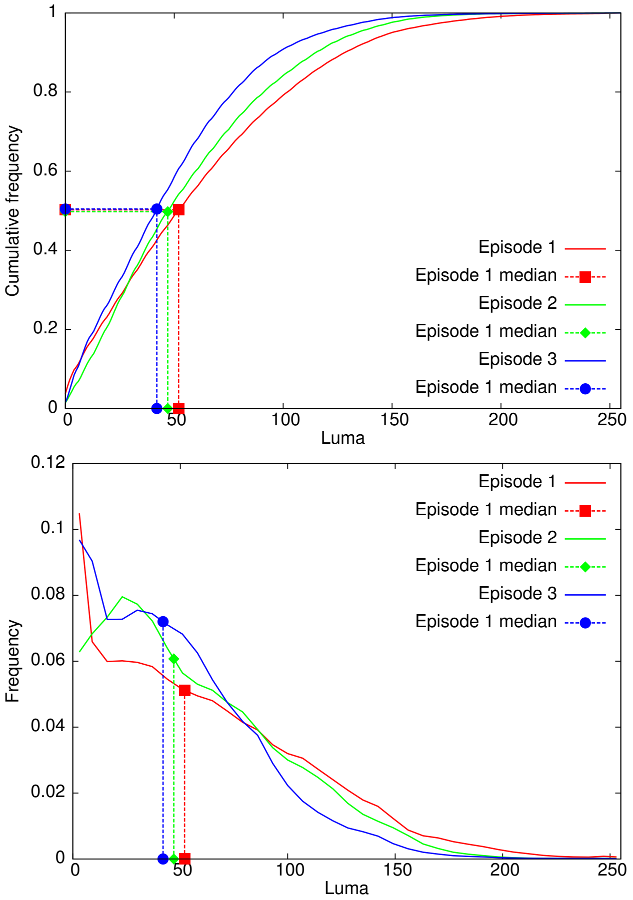

Figure 6: *Cumulative (above) and non-cumulative (below) histograms of the luma values for the three episodes. The dashed lines indicate the median luma value for each episode.*

5 Conclusions

We conducted two user studies in which we analyzed viewing preferences and the potential for visual fatigue on next generation HDR displays. Our experiments show that visual fatigue is not a serious concern even in dark environments. While subjects tended to always maximize the available display contrast, we found a sublinear relationship between the preferred display brightness and the level of ambient illumination. These results are consistent across a wide demographic spectrum.

The results of this study could be the first step in designing HDR television sets with automatic brightness controls to provide a more pleasurable viewing environment under a variety of ambient viewing conditions.

In the current study we have focused on live action content without drastic illumination effects. This choice was based on the need to

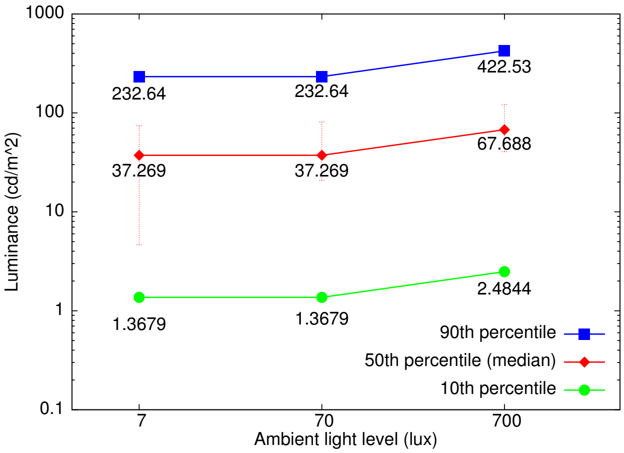

Figure 7: *Preferred display brightness levels by ambient lighting level. The lines represent luminance of the 10th, 50th and 90th percentiles of the video content (computed for all the frames). The error bars show the 25th and 75th percentiles of the brightness settings. The preferred brightness setting was found by computing the median of all subject settings.*

limit the number of parameters in the study. Now that visual fatigue has been ruled out and we have gained a better understanding of brightness and contrast preferences under different ambient light levels, we would like to conduct further studies that analyze the effect of more extreme content, such as animated content and very dark or highly stylized live action footage.

6 Acknowledgments

This work was supported by Dolby under the Dolby Research Chair in Computer Science at the University of British Columbia. We would also like to thank Dolby Canada for use of their facilities and help in organizing the studies.

References

ACOSTA-SERAFINI, P., MASAKI, I., AND SODINI, C., 2000. Single-chip imager system with programmable dynamic range. U.S Patent 6,977,685, Feb.

AKYÜZ, A., REINHARD, E., FLEMING, R., RIECKE, B., AND BUELTHOFF, H. 2007. Do HDR displays support LDR content? A psychophysical evaluation. *ACM Trans. Graph. (Proc. ACM Siggraph) 26*, 3, Article 38.

ASHIKHMIN, M., AND GOYAL, J. 2006. A reality check for tone-mapping operators. *ACM Trans. Applied Perception 3*, 4, 399–411.

BANTERLE, F., LEDDA, P., DEBATTISTA, K., AND CHALMERS, A. 2006. Inverse tone mapping. In *Proc. of GRAPHITE '06*, 349–356.

BARTLESON, C. J., AND BRENEMAN, E. J. 1967. Brightness reproduction in the photographic process. *Photographic Science and Engineering 11*, 4, 254–262.

BERANEK, L. L. 1957. Criteria for noise in buildings. *NOISE Control 3*, 1, 19–27.

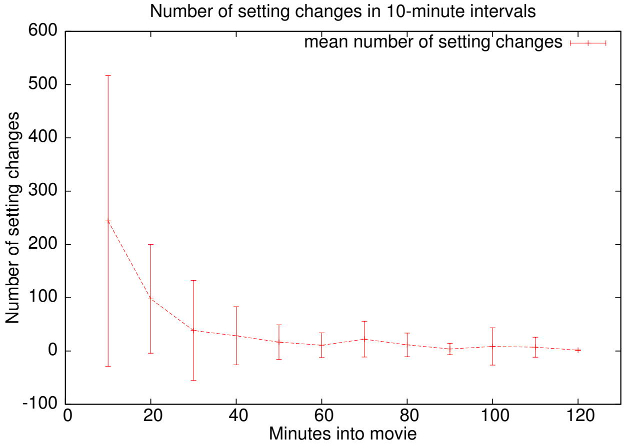

Figure 8: *Average setting changes made during a first session. Subsequent sessions showed similar patterns with approximately half the overall number of changes per session.*

CAPPELS, R. D., AND HERNANDEZ, M., 1997. System and method for adjusting the output of an output device to compensate for ambient illumination. U.S Patent 5,670,985, Sept.

DE MARSH, L. E. 1972. Optimum telecine transfer characteristics. *Journal of the Society of Motion Picture and Television Engineers 81*, 784–787.

DEVLIN, K., CHALMERS, A., AND REINHARD, E. 2006. Visual calibration and correction for ambient illumination. *ACM Trans. Applied Perception 3*, 4, 429–452.

DIDYK, P., MANTIUK, R., HEIN, M., AND SEIDEL, H. 2008. Enhancement of bright video features for hdr displays. In *Eurographics Symposium on Rendering*, 1265–1274.

DILLON, T. W., AND EMURIAN, H. H. 1995. Reports of visual fatigue resulting from use of a video display unit. *Computers in Human Behavior 11*, 1, 77–84.

DILLON, T. W., AND EMURIAN, H. H. 1996. Some factors affecting reports of visual fatigue resulting from use of a VDU. *Computers in Human Behavior 12*, 1, 49–59.

DROR, R., WILLSKY, A., AND ADELSON, E. 2004. Statistical characterization of real-world illumination. *Journal of Vision 4*, 9, 821–837.

GHOSH, A., TRENTACOSTE, M., SEETZEN, H., AND HEIDRICH, W. 2005. Real illumination from virtual environments. In *Eurographics Symposium on Rendering*, 243–252.

ISHIHARA, S. 2007. *Ishihara's Tests for Colour Deficiency*. Kanehara Trading Inc.

ITU. 1990. ITU-R BT.709, basic parameter values for the HDTV standard for the studio and for international programme exchange. Standard Recommendation 709, International Telecommunication Union.

KALMANASH, M. H., 1993. Backlight for liquid crystal devices. U.S Patent 5,211,463, May.

KARWOWSKI, W. 2001. *International encyclopedia of ergonomics and human factors*. Taylor & Francis.

KATYL, R. H., GREENE, R. G., YOST, B., AND KRUSIUS, J. P., 2001. Correction methods for brightness in electronic display. U.S Patent 6,275,825, Aug.

LEDDA, P., CHALMERS, A., TROSCIANKO, T., AND SEETZEN, H. 2005. Evaluation of tone mapping operators using a high dynamic range display. *ACM Trans. Graph. (Proc. ACM Siggraph) 24*, 3, 640–648.

LIKERT, R., ROSLOW, S., AND MURPHY, G. 1934. A simple and reliable method of scoring the Thurstone attitude scales. *Journal of Social Psychology 5*, 228–238.

MANTIUK, R., KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2004. Perception-motivated high dynamic range video encoding. *ACM Trans. Graph. (Proc. of ACM Siggraph) 23*, 3, 733–741.

MANTIUK, R., EFREMOV, A., KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2006. Backward compatible high dynamic range MPEG video compression. *ACM Trans. Graph. (Proc. of ACM Siggraph) 25*, 3, 713–723.

MANTIUK, R., DALY, S., AND KEROFSKY, L. 2008. Display adaptive tone mapping. *ACM Trans. Graph. (Proc. of ACM Siggraph) 27*, 3, Article 68.

MERRIFIELD, R., AND SILVERSTEIN, L. D. 1988. The ABC's of automatic brightness control. In *Society for Information Display Internatiational Symposium Digest of Technical Papers*, 178–181.

MEYLAN, L., DALY, S., AND SÜSSTRUNK, S. 2006. The reproduction of specular highlights on high dynamic range displays. In *Proc. of the 14th Color Imaging Conference*.

MEYLAN, L., DALY, S., AND SÜSSTRUNK, S. 2007. Tone mapping for high dynamic range displays. In *Proc. of Human Vision and Electronic Imaging XII*, vol. 6492.

NOVICK, S. B. 1969. Tone reproduction from colour telecine systems. *British Kinematography Sound And Television 51*, 10, 342–347.

PATTANAIK, S. N., FERWERDA, J. A., FAIRCHILD, M. D., AND GREENBERG, D. 1998. A multiscale model of adaptation and spatial vision for realistic image display. In *Proc. of ACM Siggraph '98*, 287–298.

PHILIPS, 2008. Philips ambilight. http://www.misc.philips.com/ambilight/uk/index2.html, accessed Nov. 19, 2008.

PITTMAN, C. R., 1976. Ambient light responsive illumination brightness control circuit. U.S Patent 3,962,600, June.

REINHARD, E., STARK, M., SHIRLEY, P., AND FERWERDA, J. 2002. Photographic tone reproduction for digital images. *ACM Trans. Graph. (Proc. of ACM Siggraph) 21*, 3, 267–276.

REMPEL, A. G., TRENTACOSTE, M., SEETZEN, H., YOUNG, H. D., HEIDRICH, W., WHITEHEAD, L., AND WARD, G. 2007. Ldr2hdr: On-the-fly reverse tone mapping of legacy video and photographs. *ACM Trans. Graph. (Proc. of ACM Siggraph) 26*, 3, Article 39.

RUDERMAN, D. L. 1994. The statistics of natural images. *Network: Computation in Neural Systems 5*, 598–605.

SEETZEN, H., WHITEHEAD, L., AND WARD, G. 2003. A high dynamic range display using low and high resolution modulators. In *Society for Information Display Internatiational Symposium Digest of Technical Papers*, 1450–1453.

SEETZEN, H., HEIDRICH, W., STUERZLINGER, W., WARD, G., WHITEHEAD, L., TRENTACOSTE, M., GHOSH, A., AND VOROZCOVS, A. 2004. High dynamic range display systems. *ACM Trans. Graph. (Proc. ACM Siggraph) 23*, 3, 760–768.

SEETZEN, H., LI, H., YE, L., HEIDRICH, W., WHITEHEAD, L., AND WARD, G. 2006. Observations of luminance, contrast, and amplitude resolution of displays. In *Society for Information Display (SID) Digest*, 1229–1233.

STEVENS, J. C., AND STEVENS, S. S. 1963. Brightness function: Effects of adaptation. *Journal of the Optical Society of America 53*, 3, 375–385.

STEVENS, S. S. 1961. To honor Fechner and repeal his law. *Science 133*, 3446, 80–86.

SWINKELS, S., HEYNDERICKX, I., YEATES, D., AND ESSERS, M. 2008. Ambient light control for mobile displays. In *Society for Information Display Internatiational Symposium Digest of Technical Papers*, 1006–1009.

ČADÍK, M., WIMMER, M., NEUMANN, L., AND ARTUSI, A. 2008. Evaluation of HDR tone mapping methods using essential perceptual attributes. *Computers & Graphics 32*, 330–349.

WARD, G., AND SIMMONS, M. 2005. JPEG-HDR: A backwards-compatible, high dynamic range extension to JPEG. In *Proc. of Color Imaging Conference (CIC) '05*.

YOSHIDA, A., MANTIUK, R., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2006. Analysis of reproducing real-world appearance on displays of varying dynamic range. In *Proc. of Eurographics 2006*, 415–426.

YOSHIDA, A., BLANZ, V., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2007. Testing tone mapping operators with human-perceived reality. *Journal of Electronic Imaging 16*, 1, 013004–1 – 013004–14.