https://doi.org/10.2352/ISSN.2470-1173.2018.14.HVEI-509

Pupillometry of HDR Video Viewing

Scott Daly[a], Evan Gitterman[a], and Grant Mulliken[b]
[a]Dolby Laboratories, Inc., San Francisco, CA, USA
[b]Apple, Cupertino, CA, USA

Abstract

Viewing of High Dynamic Range (HDR) video on capable displays poses many questions to our understanding of perceptual preference and vision science. One of the most fundamental aspects is the role played by light adaptation, as HDR content and displays allow for substantially increased light adaptation changes. In contrast, the traditional formats of standard dynamic range (SDR), being at best 3log10, kept the luminance ranges well within the steady state ranges of photoreceptor responses [14]. HDR video systems exceed the 3log10 luminance range, can be as high as 5log10 for professional displays, and be over a 6log10 range for laboratory research displays. In addition to the well-understood photoreceptor component of light adaptation is the pupillary component. While its light modulation is much smaller in range than the photoreceptor's adaptation range, it nevertheless has engineering consequences, and has been cited as a cause of putative discomfort with some HDR viewing. To better understand its role in light adaptation and discomfort, this study measured pupil behavior during naturalistic viewing of HDR video on a professional display, and performed various analyses.

Introduction

High dynamic range (HDR) displays and associated video signals offer the opportunity to display a much-increased degree of realism, as the standard dynamic range (SDR) systems have had to limit the scene and displayed dynamic range to around two log10 units. At best, SDR may allow for a range up to three log10. The increased in realism in HDR exceeding 5 log10 can be both in terms of the intra-scene dynamic range, as well as the changes in inter-scene range (i.e., from one scene to the next). It is generally understood that HDR allows for a larger range of light adaptation to occur in the visual system of the viewer, and thus questions about pupil behavior arise. Before delving into vision science, it is of interest to mention the common engineering expectations and speculations on pupil behavior for video.

It is generally known that the human pupil ranges from 2-3 mm in diameter for the brightest viewing conditions to 7mm for the darkest viewing conditions. The two main discussion points revolve around light adaptation and discomfort.

To complicate matters, there are often misunderstandings and over- simplifications of what constitutes HDR video. In the combination of the results and discussions, this paper will address both HVS and HDR misconceptions.

Adaptation

At the most simplified level of understanding, the pupil area changes to keep light flux on retina constant. With those who are not vision scientists, a common misunderstanding is to attribute all light adaptation to pupil area changes. In terms of application to video, the common expectation is that with SDR video, the viewer will have a larger pupil, and with HDR video, they will have a smaller pupil. There is a natural inclination to liken the visual system to camera engineering. The common misunderstandings also include assumptions of linearity. For example, in a comparison of viewing a 1000 cd/m^2 HDR with viewing a 100 cd/m^2 SDR display, a common expectation is that the viewers' pupil areas would compensate for the light level differences with the result that the $\text{Area}_{\text{SDR}} \sim 10\text{x } \text{Area}_{\text{HDR.}}$

This expectation is consistent with the common over-simplification that "HDR is all about brighter pictures". That is, the higher maximum luminance capability of HDR leads to pictures that are overall brighter, and using quantitative terminology, HDR results in a higher ADL (Average Displayed Luminance). Of course, that oversimplification doesn't consider distinctions between maximum diffuse white, peak luminance, the role of specular highlights in evoking realism, and interscene dynamic range.

Discomfort

The other oft-stated expectation is that HDR video will lead to discomfort due to increased pupil variability. The assumption is that the increased variability will result in more work for the muscles in the iris that control the pupil size. This will then lead to fatigue, which eventually leads to discomfort. This expectation is consistent with the other common over-simplification that "HDR is all about the explosions". That is, the main goal of HDR is to allow for higher dynamic transients and interscene ADL. Of course, that oversimplification overlooks the advantages of intrascene dynamic range, the role of specular highlights, and shadow detail.

Background

Known behavior of pupil sizes

Fortunately, there is a significant body of quantitative data on the behavior of the pupil. Consistent with the common assumption, light adaptation is a dominant factor [1,2,3,12]. As in camera engineering, the pupil is analogous to the aperture of a camera, generally described in log2 units of f-stops. However, another key factor in light adaptation is photoreceptor adaptation. Continuing the camera metaphor, the photoreceptor adaptation is analogous to the ~ ISO setting of a camera, which with digital sensors now occurs via an electronic gain after capture. A less well-known aspect of human vision is that the photoreceptors' temporal response changes with light adaptation level, and this is analogous to the exposure duration approach to adjust exposure.

However, unlike cameras, the pupil area changes only about 16x (1.2 log10), while the cone photoreceptors have a range greater than 5 log10. Due to these ratio differences, physiologists often regard pupil size as a minor effect, if not irrelevant, to light adaptation, and that the main attribute that should be considered/modelled/studied is receptor adaptation. However, for typical engineering ranges of cost factors as well as the light level changes for a single scene, the 16x factor from the pupil is still relevant, despite the pupil area changes being a minor effect in overall vision.

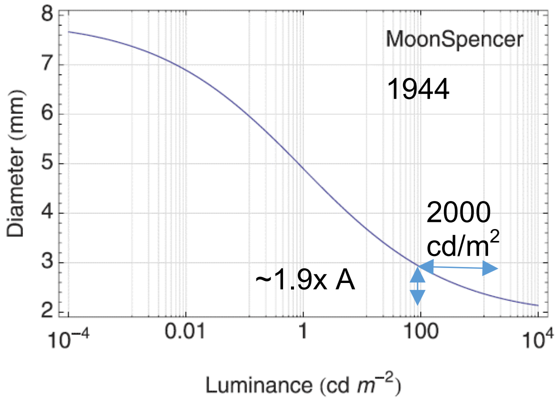

Figure 1: Watson & Yellot model [1] fit to Moon and Spencer data for pupil diameter as a function of luminance

The simplified expectation that a 10x maximum luminance difference from an HDR to an SDR display would lead to a 10x pupil area difference assumes a linearity that cameras have, but that the visual system does not. Figure 1 shows commonly cited data for pupil size as a function of light level in luminance, and it can readily be noticed that it is certainly not a linear relationship, even after converting diameters to area. For example, in considering an SDR display of 100 cd/m^2 against an HDR display of 2000 cd/m^2, we would expect a 2000/100= 20x ratio change in pupil area (A). But the figure shows that only a 1.9x change in area occurs for those two luminance levels. We use 2000 cd/m^2 here since that maximum luminance is currently available in consumer HDR displays (albeit only in the best performing displays).

In addition, there are other key factors that affect pupil diameter, which are secondary modulators to the main effects of light adaptation. One of these is cognitive load, found to control the pupil via the cingulate cortex [4,5]. Emotion is found to be another modulator, with control emanating from the locus ceruleus & cingulate cortex [5,6]. Attention is another effecter of pupil size, perhaps the smallest, with changes on the order of 0.2mm [11, 19]. Aging with its associated presbyopia leads to a smaller pupil size to increase depth of field to compensate for lens rigidity [2]. Lastly, one of the most well-known modulators of pupil size is inebriation.

Controlling Field

In Fig 1, the x-axis is labelled luminance, but does not go into specifics of the luminance. Pupil size adaptation is primarily affected by white and green light [2], and the portion of light in the visual field that most affects the pupil size is properly termed the controlling field (as opposed to adapting field). This generally refers to a brighter region than the rest of the visual field, and in most experiments, it is a luminous disc against a darker background. The controlling field size can be as small as 1 deg [3]. Stanley & Davies studied effects of the size of the controlling field from 25 to 0.4 deg [1]. Watson and Yellot modelled the Stanley and Davies data, as well as several other pupil adaptation data sets and found that the shape of curve stays the same but shifts along the log luminance axis such that the corneal flux density (luminance x area of controlling field) determines the position on the luminance axis [1]. This is illustrated in Fig 2.

Now let's consider this more advanced understanding in the context of complex (i.e., natural or synthetic) HDR imagery. Since the display is generally brighter than the surround, the pupil size will be controlled by the displayed imagery. This likelihood increases as the field of view (FOV) of the display has been increasing over time (technically at 65 deg for the new UHD formats). So, the controlling field will be within the image content. It is entirely within reason that content may have a bright region as small as 0.4 deg, and as large as 25 deg. So, in comparing the same 100 cd/m^2 versus 2000 cd/m^2 displays as we did in Fig. 1, we now see that a 10x difference in pupil area is possible considering unknown content-dependent controlling area effects as well as inter-scene changes.

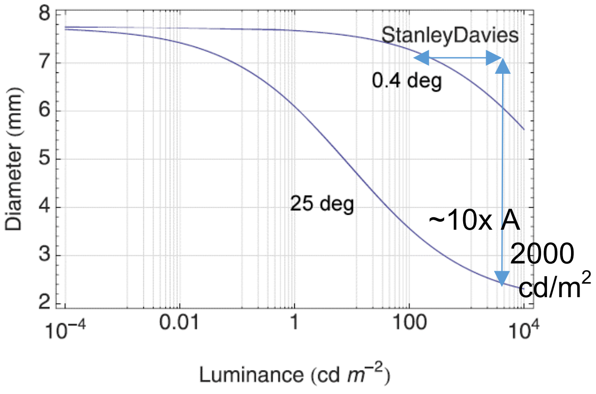

Figure 2: pupil size variations in consideration of possible controlling field size variations due to video content. Model from [1]

Pupil muscles and their enervation

In addition to a more advanced understanding of light adaptation, let us take a closer look at the possible sources of pupil variation related to discomfort: muscle fatigue and enervation. The pupil constricts as circular muscles (sphincter pupillae) of the iris contract, which is a parasympathetic enervation. The pupil dilates as the radial muscles (dilator iridis) contract, which is a sympathetic enervation.

The light component of pupil constriction is an afferent process through the optic nerve and pretectal nucleus going into the E.W. (Edinger-Westphal) nucleus as an excitation with increasing light. Factors that input the EW nucleus as inhibition include stress. The output combining these inputs acts through the parasympathetic pathway through the efferent path to the ciliary ganglions controlling the circular muscles of the iris. The dilation is mainly affected by signals through the superior cervical ganglion.

In both cases of constriction and dilation, the muscle activity acts radially or along the perimeter of the iris. This means both cases of constriction are proportional to diameter. So, for considerations of comfort, it is more important to look at the behavior of the diameter than the area of the pupil. While a common engineering expectation is that repeated flexing of these muscles would lead to fatigue, it must be remembered that due to the very small size of these muscles, the inertia is extremely small. Consequently, the forces involved are also very small. The situation is analogous to a common concern about the lifetime issues of DLP imagers, which involve rotation of small mirrors millions of times per second. This concern was found to be off-base since there was virtually zero inertia of these small mirrors. With these more detailed considerations, we have little expectation of discomfort due to increased variability of the circular or radial iris muscles. Still, it is worth investigating pupil variability for HDR because of other unknown possibilities.

Experiment

While the more advanced understanding of the visual system and HDR content as described in the background section dispel some of the speculations and misconceptions, it is still of interest to perform an actual study since the role of the spatial and luminance characteristics of HDR video content cannot be predicted from the existing pupil studies.

Goal

The primary goal is to compare pupil behavior for SDR and HDR video as displayed on an HDR-capable display. For this study, we will set the SDR parameters at 100 cd/m^2 maximum luminance and the HDR parameters will be set to 2000 cd/m^2 max luminance. While there are often other display parameters that vary across SDR and HDR, such as black level and color gamut, we will keep those fixed to focus on the variable most expected to impact the pupil behavior, which is the maximum luminance. So, all other parameters are the same across the SDR and HDR comparisons, which include a P3 color gamut, 12 bits RGB amplitude resolution, and a 0.005 black level, etc.

We used naturalistic viewing, having no task or GUI interactions. The viewing distance was 3H, and we used a 1920x1080 resolution display. We used dark ambient illumination conditions, and a comfortable fixed straight back chair. While not entirely consistent with naturalistic viewing, we used no audio in order to remove potential secondary modulations through audition.

Pupillometry system

The eye tracker, made by GazePoint, was placed underneath the TV and was focused on the subjects' face. It uses IR & near IR illumination. Its accompanying software outputs pupil size as well as gaze position, using the bright pupil tracking method, as shown in Fig. 3. Its camera has HD (1920x1080) resolution, and captures at 60Hz, which is sufficient considering the temporal response of pupil changes. Being black and placed directly under the TV, the eye tracker did not interfere with naturalistic viewing.

HDR and SDR test movie used in the study

For HDR video content we used a short demo movie from one of the major Hollywood studios titled Telescope, from 2014 and shown at NAB [7]. The movie was a demo in that HDR and Wide Color Gamut were intentionally considered in the cinematography, and it was one of the earliest HDR movies made with professional equipment. The production team consisted of Travis Labella (cinematographer), Matt Litwiller (producer), Eric Bodge (writer, producer), and Colin Davis (director, editor). It is widely used in the HDR engineering community for testing and demos. It has no dialogue, except for a short epilogue at the end, and therefore the story can easily be followed without sound.

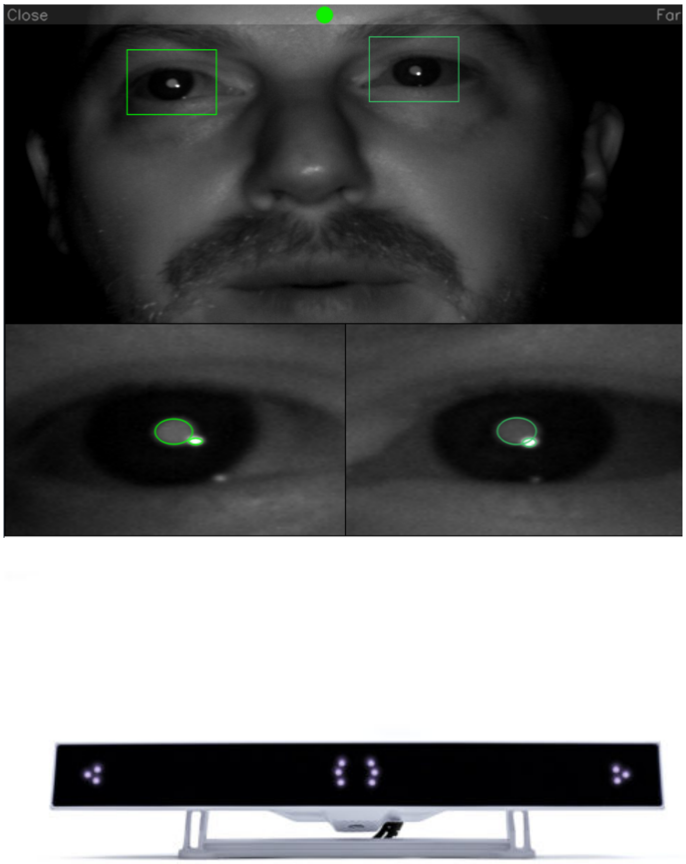

Figure 3: Pupillometry system and examples of pupil estimation

Technical details of the key equipment include the use of the Arri Alexa digital cinema camera, which is 12 bits RGB, with a 14-stop scene capture range (~4log10). Of course, such a capture range can be increased through contrast boosting performed in color grading. There were significant scenes of full-frame as well as composited computer graphics, performed in float EXR, and thus imposing no quantization or optical capture noise. The movie was color graded by the content creators to a 4000 cd/m^2 luminance maximum and a P3 gamut on a Pulsar Pro HDR reference monitor.

It was captured and displayed at 24 frames per second (fps). It is an ~ 9-minute short movie containing much content that must be clipped in SDR, but can be rendered fully in HDR. In addition, it also has content that is considered 'corner cases' or challenges to various points in the HDR ecosystem from capture to display. Examples of such 'HDR features' contained in the movie include star fields, fades to black, explosions, wide gamut nebula, a flashlight, emissive display panels, night scenes, shadow detail, skin colors, and specular reflections.

Display

The display used in the experiment was a single Maui HDR professional display, manufactured by Dolby. It is a Dual Modulation LCD with a 32" diagonal screen, and ~1500 BLE (backlight elements). It is 1920 x 1080 pixels, with a P3 color gamut, a maximum luminance of 2000 cd/m^2 max, a minimum luminance of 0.005 cd/m^2. The contrast as measured by the ICDM 5.13 metric (corner box) is 410,000:1.

Signal processing

As mentioned, a goal was to compare content at 2000 cd/m^2 max and 100 cd/m^2 max, for the HDR and SDR cases, respectively. The original HDR movie color graded at 4000 cd/m^2 was converted to a 2000 cd/m^2 max version and a 100 cd/m^2 version. Any study comparing displayed imagery with different luminance ranges requires some kind of tone-mapping process. One example is what was commonly done in NTSC and PAL video systems where displays of varying ranges from 100 to 500 cd/m^2 were used. The video signal's code values were left unchanged, and the physics of the displays were used to do the tone-mapping from the content (graded or shaded at 120 cd/m^2) to the display's max luminance which was often higher. The code values in the gamma corrected domain would be rescaled in the luminance domain. This process resulted in a shift of the mean luminance and the mode of the luminance histogram. This process is sometimes referred to as gamma scaling.

More advanced approaches use algorithms, referred to as tone-mappers, that use tone-scale nonlinearities that deviate from the gamma-corrected representation. There are two key types: those which map from a lower dynamic range to a higher range and vice versa. These are generally referred to as upmapping and downmapping, or SDR-to-HDR and HDR-to-SDR respectively. The term LDR (low dynamic range) is often used interchangeably with SDR, however, in some fields, LDR is a term reserved for displays with ranges lower than SDR (e.g., reflective displays). Although different algorithms have different goals, we used an algorithm that takes the philosophy that the key distinctions between HDR over SDR are that the ranges allocated to highlights and shadow detail are expanded, but the mean luminance level as well as midtones are relatively preserved. It is a commercial downmapping algorithm widely used in high-end consumer TVs. This algorithm was used in the downmapping mode. That is, the source video signal was downmapped from an HDR version (graded by a professional) to lower maximum luminance ranges. In this specific experiment, the source content was graded at 4000 cd/m^2, but our available HDR display was only 2000 cd/m^2, so the downmapping was used to map from 4000 to 2000 to create the HDR version. The same downmapping algorithm was used to map from the 4000 cd/m^2 version to a 100 cd/m^2 version to create the SDR version. In all cases, the black level (0.005 cd/m^2), the color gamut (P3), and the bit-depth (12 bits) were left unchanged. Consequently, the simulated SDR video imagery was slightly better than SDR imagery received by most viewers. However, the main aspect of SDR that relates to pupil changes would be the maximum luminance, the only parameter varied in this study.

The basic behavior of the downmapping algorithm used in this experiment is shown in Fig. 4, and it is also compared to the simple physics-based SDR-to-HDR process mentioned above. The histograms are from a single frame of an HDR video. The imagery contained a visible early evening sun, which shows up as the small spike at the rightmost of the histogram. The particular frame did not contain substantial pixels at the black level, so conclusions on the black level behavior cannot be made from these plots, as they are intended to focus on the highlight regions. The color lines are the histograms on a log-luminance axis for the downmapped imagery using the commercial algorithm. The source HDR video imagery was graded at 4000 cd/m^2 maximum luminance (green). Down mappings from 4000 to 1000 (purple), 400 (cyan), and 100 (orange) illustrate the key behavior. Note that most of the changes are in the highlights (above 30 cd/m^2). The mode of the histogram changes less than 4x, even though the maximums change by 40x, indicating the average mean luminance, as well as mid tone regions, are primarily preserved in luminances. Note: The average luminance level is often referred to as Average Displayed Luminance (ADL), as well as being incorrectly referred to as APL (which is intended to be an average in the picture, i.e., code value domain).

This behavior can be contrasted to the much larger mid-tone luminance changes that occur in the physical process (or any gamma-corrected scaling), which is shown as the black dashed curve. That curve shown an upmapping algorithm going from 100 to 400 as occurs commonly in the low-end TV market. Note the substantial differences in the mode and average luminance, and that the highlights have a shape very similar to the 100 cd/m2 version.

While the plot doesn't show the mapping for the 2000 cd/m^2 max luminance level, its behavior would lie between the 4000 and 1000 histograms, as shown by the arrow. The studio producing the content approved the various grades from the tone-mapping algorithm described here.

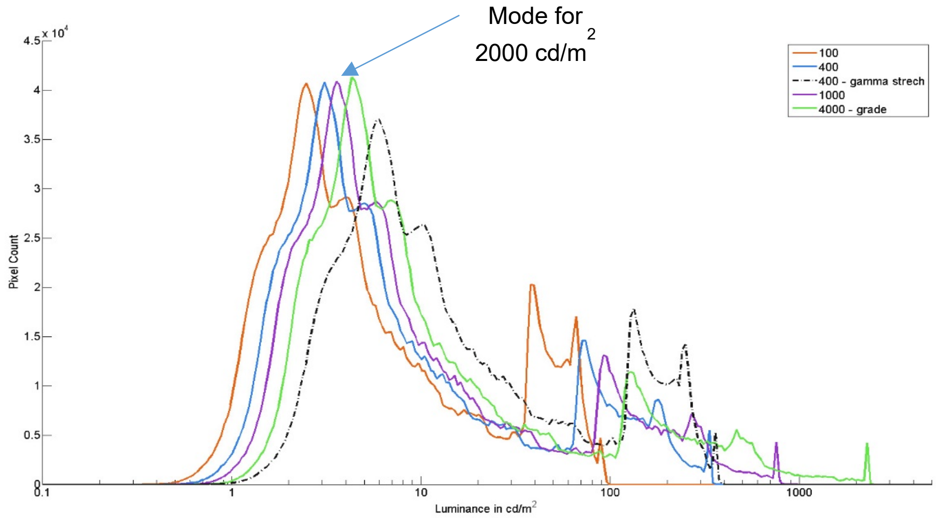

Figure 4: examples of signal processing using HDR tone mapping vs. default range scaling.

Results

The analysis started after ~1minute to avoid initial pupil transients, etc. Specifically, frame 0 in the analysis occurred right before the Telescope title image appears and the analysis ended around frame 475 right before the credits start. Example results from subject RG are shown in Figure 5. The red and blue solid lines show the pupil diameters in mm after temporal LPF filtering was applied to remove instrumentation noise, which was kept shorter than the pupil temporal response. The corresponding shaded regions are the envelope of raw data from the GazePoint pupil detector. The most obvious finding is that the pupil size is never constant (neither for the HDR nor the SDR).

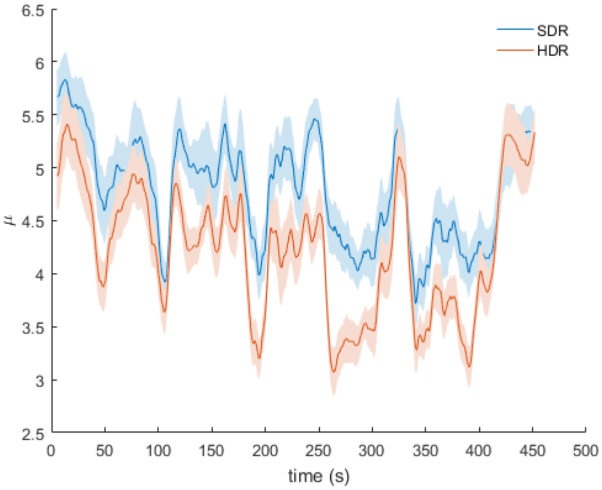

Figure 5: Results for subject RG (25<age<35).

Means and Extrema

From the data, we first discuss the means and extrema of the pupil size. As expected from the common understanding described in the introduction, the HDR viewing condition resulted in a smaller diameter. First, we look at the results in terms of diameters, which is the metric relevant to muscle action, and possible discomfort aspects. The mean for SDR = 4.81 mm, while the mean for HDR = 4.25 mm, giving a ratio of $D_{SDR}/D_{HDR} = 1.3$. In considering the extrema, $3.6 < SDR < 5.8$ while $3.2 < HDR < 5.45$.

Next, we look at the results in terms of pupil area, which is relevant to the light adaptation aspects, being proportional to light flux reaching the retina. The area mean for SDR = 5.8 mm^2 and the area mean for HDR = 4.5 mm^2. More usefully, the ratio of these means, $A_{SDR}/A_{HDR} = 1.3$, certainly less than the 20x factor expected from simple linear consideration of the max luminance differences of the SDR and HDR displays used in the experiment Next, we consider the extrema, where 3.2 mm^2 < SDR < 8.3 mm^2 and 2.6 mm^2 < HDR < 7.4 mm^2. Similarly, it is useful to look at the area extrema ratios for each of the SDR and HDR conditions. For SDR, $A_{SDRmax}/A_{SDRmin} = 2.55$, while the ratio for HDR, $A_{HDRmax}/A_{HDRmin} = 2.90$. While the extrema area ratio for HDR was higher than that for SDR, it was only slightly higher.

Variances of Diameters

In the interest of assessing possible discomfort, we plotted the histograms in Figure 6 and also calculated the variances, indicated in the figure. Despite the shift and minor shape change in the histograms, the variances are essentially equal. Of course, standard deviations would be even closer.

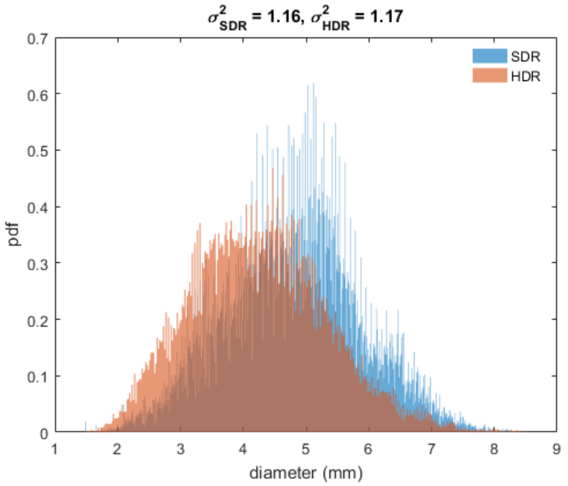

Figure 6: Histograms of pupil diameters, and variance calculations

Temporal Changes

Since the variances of the HDR and SDR cases were so similar we wanted to consider other factors. For example, the temporal changes in pupil diameter may be more important than overall variances. Figure 7 shows two types of temporal analysis. The upper plot shows the variance within a 10-sec moving window for both the SDR and HDR viewing cases. The lower plot shows the Δ diameter across a 2-sec moving window. As before, the solid line represents the data after temporal filtering to dampen instrument noise, and the shaded regions show the analysis based on the raw data from the GazePoint algorithm. With the exception of a few short intervals, such as around frame 105, there is no substantial difference between SDR and HDR temporal behavior for both types of temporal analysis tried.

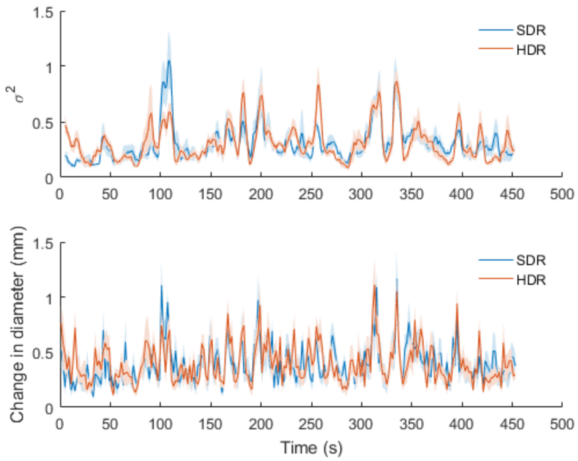

Figure 7: Temporal analysis

	SDR	HDR
DIAMETERS (mm)	------------------------------	------------------------------
MEAN	4.81	4.25
EXTREMA	3.6 < SDR < 5.8	3.2 < HDR < 5.45
SD	1.078	1.080
AREAS (mm^2)	------------------------------	------------------------------
MEAN	5.8	4.5
EXTREMA	3.24 < SDR < 8.27	2.56 < HDR < 7.43
EXTREMA AREA RATIO	2.55	2.90
SD	1.16*	1.17*

Table 1: summary of numerical pupil data (estimated from the diameters stats)*

Conclusions

One initial conclusion is that pupil size is not constant for SDR, contrary to the common engineering expectations. A second key conclusion is that HDR causes a smaller pupil diameter than SDR, by a small amount (HDR mean diameter is ~0.89x the SDR mean diameter). Another conclusion is that most of the changes as seen in Fig. 5 can be related to particular changes in the video content, which we invite the reader to do by accessing either SDR or HDR versions of the movie referenced [7]. In general, bright scenes decreased the pupil diameter, and dark scenes increased it. Momentary flashes caused a short transient decrease in pupil size, such as around frame 105. In most cases, the SDR and HDR pupil diameters track similarly, but in a few instances, they trend in opposite directions (frames 260-300). A hypothesis is that this section has a scene with a mix of bright and dark regions that could be foveated, and the subject might have foveated the dark region for the HDR and the bright region for the SDR viewing.

Comfort

Considering the comfort issues first, where diameter is more relevant to muscle action, the results show that HDR is similar to SDR based on pupil size variability as assessed by variance. In addition, other temporal analyses found similar pupil size changes for HDR and SDR. As mentioned, the main difference was that average pupil diameter decreased for HDR. This would not be expected to cause any discomfort as the diameters were still greater than 3mm, which is not close to the minimum pupil's minimum size where discomfort may begin. In fact, since it is known that a smaller pupil leads to higher acuity, it may actually be possible that the HDR case could be more comfortable. Discarding these untested speculations, our main conclusion is that we found no evidence of pupil-based discomfort. Of course, this conclusion is based on the type of tone mapping used in the experiment, which adds several stages of professional human interaction, such as in the initial grading of the content on an HDR display of 4000 cd/m^2 capability, and also the approval of the tone mapping algorithm used to make the 2000 cd/m^2 and 100 cd/m^2 versions shown. In the professional and research communities, it is known that uncomfortable HDR content can certainly be created. Examples include simple automatic SDR to HDR algorithms that do not attempt to control the ADL, as well as master grades that may intentionally cause discomfort, analogous to the use of flash pots in concerts [15].

Adaptation

The pupil behavior effects on adaptation, where area is more relevant due to retinal photon flux, are perhaps more interesting, as they have implications for color engineering. As mentioned, the results show that the simple engineering expectation that pupil areas would exactly compensate for max luminances is way off, where the found 1.3x ratio (of SDR/HDR means over the length of the movie) is substantially smaller than the 20x ratios expected from a simple consideration of the max luminances. In addition, the ratio is smaller than expected from use of the pupil response vs. light adaptation plots of Fig. 1 and 2 (Moon-Spencer, Stanley-Davies). However, this discrepancy is not the fault of those datasets, rather, it is in the oversimplification of the HDR range by merely describing its maximum luminance level, as is often done. A more advanced understanding of HDR that factors in the tone mapping and color grading will likely be able to explain the results, although that work has not yet been performed. As mentioned, the pupil variability for both SDR and HDR show greater than 2.5x variability during the course of video viewing.

Video Colorimetry

The effects on pupil variability have effects on light adaptation, and these in turn have implications for video colorimetry, in particular, the initial Y/Y_n calculation common to all CIE models [16]. These CIE models, such as CIELAB, have been very effective in many application areas, such as the graphics arts industry or product color characterization. Both cases have the commonality that they are modelling the appearance of reflective surface colors: objects and prints. In those applications, there is plenty of stable white reference, such as surrounding the product, or surrounding the image for WYSIWYG soft-proofing emissive displays. In those applications, there is generally calibrated illumination via the use of a viewing booth [17]. These scenarios help to keep the pupil size much more stable, as well as the photoreceptors' light adaptation state.

However, video is an application where the light adaptation is generally more controlled by the displayed content (either due to viewing in a dimmer ambient than the content, or because the display FOV is large enough that foveation stays on the display. A key factor is determining the Y_n value, which is often thought of as the light adaptation level, but is more carefully termed the white point. In engineering colorimetry, this value must be constant and signal independent. That requires that both the pupil area and the cone adaptation levels be fixed. This assumption works well for proofing, but has always posed problems for video applications. Fortunately, SDR limits those problems, and the main usage of CIE colorimetry applied to video was that Y_n would be set to the displays' max luminance. A problem arises because there is no white surround, and in most cases nowadays, the display drives adaptation, not the room illumination. Thus we have a signal dependent adaptation. As can be seen from our results, the pupil size fluctuations alone can cause Y_n to be off by 2.5x for SDR and 3x for HDR, if interpreted that Y_n should be related to retinal flux.

Various approaches in setting the Y_n term more dynamically for video have been explored, because $Y_n(\text{frame}) \mathrel{!=} Y_n(\text{display}_{max})$.

The simplest of the non-traditional approaches is to use adaptation to each individual frame, that is, $Y_n(\text{frame}) = Y_{max}(\text{frame})$. However, this is known to produce too much flickering. A more advanced non-traditional approach is to set the adaptation by scene, but this requires metadata to demarcate scene edits. Another of the more advanced non-traditional approaches is to have the temporal adaptation modelled as a temporal LPF [8]. Others have explored setting Y_n to the diffuse white in image, which in turn is $< Y_n(\text{display}_{max})$ [9]. This approach also requires image-based metadata, or system criteria on where to place the diffuse white. However, one key new feature of HDR is that it can allow for scene to scene overall luminance changes, so that daylight scenes can feel substantially different from indoor scenes, and night scenes [18]. Thus, constraining a system to set the diffuse white point to a constant luminance defeats some of the advantages of HDR. One of the more recent approaches is to consider multiple states of adaptation as a hull, based on considering video as a statistical ensemble of many possible images causing many adaptation states. A specific feature of that approach is to model the envelope of best possible adaptation [10] for maximum visibility.

Lastly, consideration of pupil effects as found in this study motivate questions about specific of the role of retinal illuminance (i.e., Trolands being the product of cd/m^2 and pupil area) and the pupil size variations: for example, considerations of the differing time constants, such as the pupil having faster dynamics than photoreceptor adaptation, which is generally assumed. In addition, there is the adaptation components being multiplicative such as in the pupil response, as opposed to having a sigmoidal shift as in the photoreceptors. Lastly, it brings up the specifics of whether the Y_n term is attempting to model pupil adaptation, photoreceptor adaptation, receptor and cortical effects [20], or all of these?

References

[1] A. B. Watson and J. I. Yellott (2012) A unified formula for light-adapted pupil size, JOV 12(10).
[2] Lobato-Rincón, et al (2014) "Pupillary behavior in relation to wavelength and age", Front Hum Neurosci. 2014; 8: 221.
[3] Atchison et al (2011) "Influence of field size on pupil diameter under photopic and mesopic light levels", Clinical and experimental optometry.
[4] J. Beatty, "Task-evoked pupillary responses, processing load, and the structure of processing resources," Psychol Bull., vol. 91, no. 2, pp. 276– 292, 1982.
[5] Siddhartha Joshi,Yin Li, Rishi Kalwani, and Joshua I. Gold, "Relationships between pupil diameter and neuronal activity in the locus ceruleus, colliculi, and cingulate cortex" Neuron. 2016 Jan 6; 89(1): 221–234.
[6] V. Sterpenich et al., "The locus ceruleus is involved in the successful retrieval of emotional memories in humans," J. Neurosci., vol. 26, no. 28, pp. 7416–7423, 2006.
[7] http://telescopemovie.com/
[8] Pattanaik, Tumblin, Yee, Greenberg (2000) Time-Dependent Visual Adaptation For Fast Realistic Image Display, Siggraph
[9] Fairchild and Chen (2011) "Brightness, Lightness, and Specifying Color in High-Dynamic-Range Scenes and Images" SPIE Electronic Imaging
[10] M. Nezamabadi, S. Miller, S. Daly, and R. Atkins (2014) Color signal encoding for high dynamic range and wide color gamut based on human

perception. Color imaging at SPIE' s Electronic Imaging. Proc. SPIE 9015, Color Imaging XIX: Displaying, Processing, Hardcopy, and Applications, 90150C (February 2014)
[11] Mathot et al (2013) The Pupillary Light Response Reveals the Focus of Covert Visual Attention. PLOS
[12] Spitschan et al (2017) The human visual cortex response to melanopsin-directed stimulation is accompanied by a distinct perceptual experience. PNAS
[13] ICDM Display Metrology manual
[14] T. Kunkel and E. Reinhard (2010) A reassessment of the simultaneous dynamic range of the human visual system" APGV 2010, p17-24.
[15] http://www.theatrefx.com/funfacts97.html
[16] R.W.G, Hunt (1987) Measuring Color, John Wiley and Sons., NYC, NY
[17] https://www.ebay.com/sch/sis.html?_nkw=Gretag+Macbeth+SPL+III+Spectralight+Light+Booth
[18] S. Ruggieri (2016) "Breaking out of the 100-nit box: a colorist's view of HDR grading", SMPTE 2016 Annual technical conference and exhibition.
[19] P. Bindah, M. Pereverzeva, S. Murray (2014) 'Pupil size reflects the focus of feature-based attention' J. Neurophys. 112, 3046-3052.
[20] E. H. Land and J.J. McCann (1971) 'Lightness and Retinex Theory', JOSA V61, 1-11.

Author Biography

Scott Daly is currently with Dolby Laboratories, and is working on High Dynamic Range (HDR), High Frame Rate (HFR), and VR imaging systems, with a focus on perceptual issues. He has degrees in Electrical Engineering and Bioengineering. His previous accomplishments include key contributions to the DICOM medical imaging standard, a technical Emmy for a video transceiver, and the Visible Differences Predictor (VDP) while at Kodak. He holds an Otto Schade award from SID from his accumulated work at Kodak and Sharp, and while at Dolby was a co-author of the cone-based Perceptual Quantizer nonlinearity (SMPTE 2084).

Evan Gitterman received his BS from Stanford University, where his research was primarily in neuromusic. He is currently at Dolby Laboratories, where his work includes EEG- and physiology-based assessment of multimedia technology experiences.

Grant Mulliken is currently with Apple, Inc. and is a research scientist studying neural coding principles and harnessing them to create innovative experiences. He has made contributions to basic science as well as neuro-engineering applications, including brain-machine interfaces. His specialties include systems and computational neuroscience: signal processing, machine learning, algorithm design, spiking neuron modeling, audio coding and acoustic, cortical brain-machine interfaces, multi-electrode neurophysiology, and medical device design, mixed signal VLSI circuit design, bioinstrumentation and neuromorphic engineering.