HIGH DYNAMIC RANGE, VISUAL ADAPTATION, AND CONSUMER PERCEPTION: APPLYING VISUAL SCIENCE IN THE SEARCH FOR CONSENSUS ON HDR

Sean T. McCarthy, Ph.D.
ARRIS

INTRODUCTION

As momentum builds in the market transition to 4k Ultra HD, the emergence of High Dynamic Range (HDR) technology is drawing video engineers' attention to new issues that must be resolved if the full potential of a next-generation viewing experience is to be realized.

While 4k UHD provides the pixel density essential to enabling an immersive, big-screen TV viewing experience, HDR represents an opportunity to deliver a level of realism on video displays of all sizes that is far beyond what has been possible until now. By delivering greater contrast, increased luminance and an expanded color gamut, HDR vastly surpasses the SDR (Standard Dynamic Range) parameters that were standardized with ratification of Recommendation ITU-R BT. 709[1] 25 years ago.

Unlike 4k UHD, which has been a fairly straight-forward step in the evolution of display resolution built on the SDR foundation, HDR introduces a new paradigm where the dimensions of the new viewing experience must be defined in keeping with basic principles of the human visual response system. This will impact everything that's done in the creation and dissemination of video content from initial capture through production, post production, and processing for distribution.

Presently there are a number of approaches to HDR vying for traction in the marketplace. Much of the modeling that has gone into setting their parameters has focused on human contrast sensitivity, often in the context of non-broadcast use cases such as cinematic and episodic TV programming.

However, as various organizations pursue HDR standardization initiatives to facilitate market adoption, the priority must be on choosing solutions from a much broader perspective on visual response processes. Along with responsiveness to degrees of contrast, developers must consider human perceptual factors such as:

- Light and dark adaptation
- Brightness sensitivity
- Reaction to ambient light
- How color is perceived under different conditions
- Responses to frame-rate flicker that might be introduced with expansions in dynamic range

Moreover, planning for HDR must take into account the live TV broadcast environment. Content producers and distributors will have to determine how human visual response in the HDR environment will impact quality parameters for advertising, channel changes between HDR and SDR programming and the presentation of user interfaces, captioning and other textual and graphic elements. The impact of various HDR modes on bitrate and bandwidth requirements will also be an important consideration, especially for MVPDs.

In this paper, we look at the approaches to HDR and review the characteristics of human perception that will impact industry players' selection of HDR modes and their implementation in the multichannel viewing environment. Our goal is to provide basic information that will help manufacturers, producers and distributors answer key questions such as “How bright a screen is bright enough?”; “How bright is too bright?”; “How quickly can a person adapt to changes in brightness, and what does that mean in terms of commercial insertion?”; and “How does the experience of HDR TV change with room lighting?”

THE EMERGENCE OF HIGH DYNAMIC RANGE IN TELEVISION

High contrast ratios between the darkest and brightest elements of images within a single frame and across sequential frames have long been applied in the cinematic viewing experience and to improve accuracy in the reading of medical images (Figure 1). In the case of theatrical displays, high levels of contrast are achieved in a dark-room environment where the highest luminance levels are well below that of even cathode ray tube TV displays but nonetheless deliver a much better viewing experience than can be achieved with the lower contrast levels used in traditional television programming.

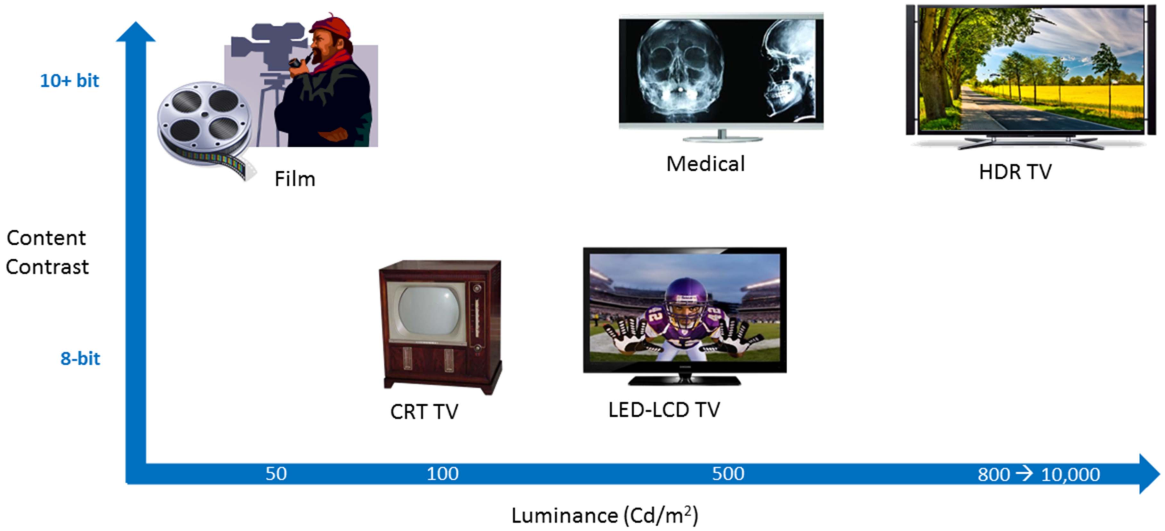


Figure 1. HDR television is part of the evolution of display and coding technologies and services.

Today these contrast limitations are not determined by the capabilities of display devices. Instead, they’re a function of industry adherence to the parameters set by the ITU-R BT. 709, which were based on the fact that the typical luminance of CRTs used as reference displays for HDTV production is about 100 cd/m^2.

No such limitations were present with the introduction of HDR in medical imaging, which, by virtue of raising the peak luminance level up to 4,000 cd/m^2, more closely approached what we know today as HDR in terms of contrast range and luminance. (In actual practice, the Digital Imaging and Communications in Medicine[2] (DICOM) standard is employed in controlled lighting environments where sufficient contrast is achieved with a top luminance level around 600 cd/m^2.)

The television industry is now developing standards for HDR by specifying electro-optic transfer functions (EOTF) and opto-electric transfer functions (OETF). [An EOTF specifies the non-linear mapping of illumination at the camera to digital code values. An OETF specifies the inverse non-linear mapping of digital code values to display luminance.] SMPTE ST-2084[3] specifies an EOTF and OETF that uses a technique of perceptual quantization[4] that is based on the same sophisticated model of human spatial contrast sensitivity[5] that is the basis of the DICOM standard. SMPTE ST-2084 supports peak luminance up to 10,000 cd/m^2 and is intended for displays used primarily for mastering non-broadcast content. An alternative EOTF/OETF has been proposed[6] that builds on the current SDR standards, ITU-R BT.18867 (EOTF) and ITU-R BT.709 (OETF), by adding a "knee" to handle luminance greater than 100 cd/m^2.

So far, the commercial introduction of HDR has been in conjunction with 4k UHD sets, most of which employ various iterations of LED LCD display technology. However, HDR is not intrinsically tied to 4k UHD and therefore could be used to greatly enhance the viewing experience with HDTV displays and, eventually, personal devices such as tablets and smartphones as well.

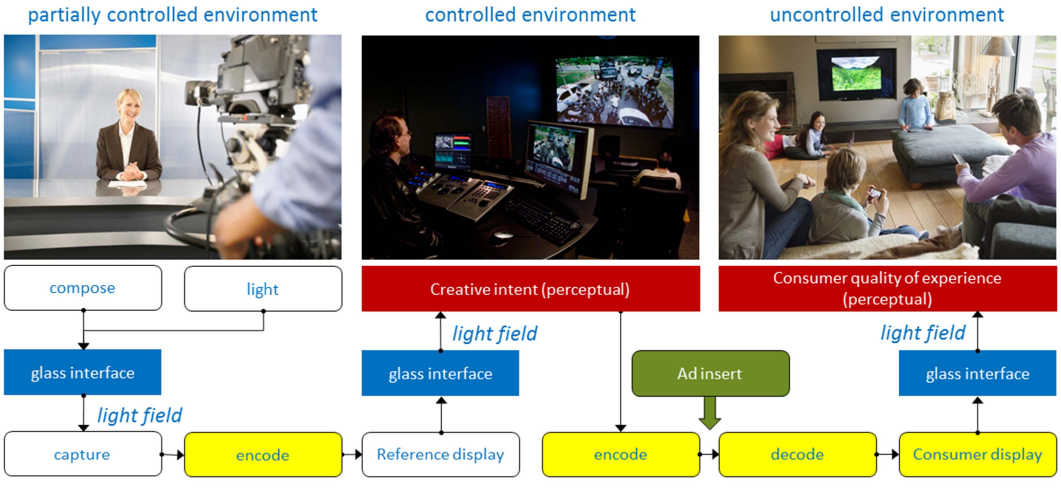
Along with higher luminance and contrast ratios, HDR goes hand-in-hand with an industry trend to broaden the color gamut beyond the 16.78 million colors supported by ITU-R BT.709. ITU-R BT.2020[8] has standardized a palette reaching 1.07 billion colors with 10-bit encoding and 68.7 billion colors with 12-bit encoding. In between these two levels is the DCI P3[9] color gamut, representing about 700 million colors, which has been in use with cinematic projections since 2007. Most providers of first-generation HDR technology are delivering 90 percent or more of the colors in the DCI P3 space.

STANDARDIZING HDR PARAMETERS

With a proliferation of HDR modes in play, the industry has made standardization of HDR a top priority, as evidenced by the efforts of the ITU, MPEG, Blu-ray Disc Association, SMPTE, ATSC, Ultra HD Forum, and UHD Alliance. As these initiatives progress toward what hopefully will be a unified approach to bringing this technology to market, it's essential that these groups look beyond the impact degrees of contrast have on human visual response to how the full range of perceptual processes are impacted by any given set of contrast, luminance and color parameters.

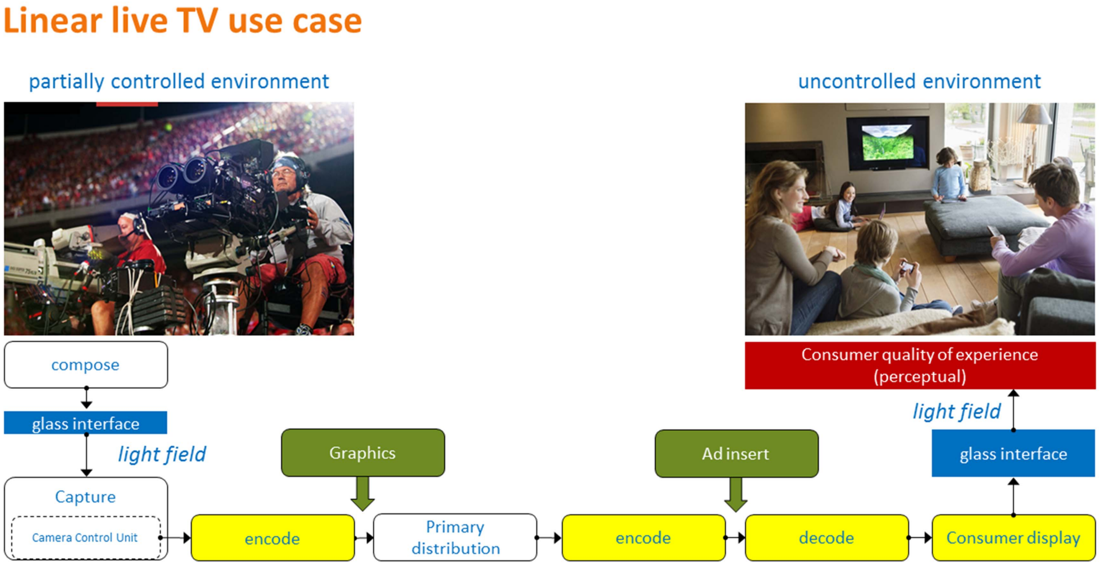
Addressing the Live HDR Imperative

It's also important to note that, as illustrated by Figures 2 and 3, the test case for setting specifications and best practices should be live broadcast TV. In the case of non-live, cinematic or episodic programming (Figure 1), the workflow for processing content entails capturing and encoding from a composed setting, an entire program, for delivery into post production processing.



Fidelity in the light field would create the intended perceptual experience

Figure 2. Non-live cinema and episodic programming enable experts to craft content to deliver the creative intent. Nonetheless, television is viewed in many different lighting environments that could impact the effectiveness of HDR. Moreover, advertisements and other interstitial content is usually inserted into programming. Those interstitials could be SDR, which could impact the overall consumer quality of experience.



Fidelity in the light field would create the intended perceptual experience

Figure 3. Live HDR television poses special challenges including management of real time graphics and logo overlays, management of venue lighting conditions, and management of HDR and SDR interstitials.

Post production editors have an opportunity to use reference monitors in non-real time to ensure the creative intent of the producers is captured in the formatting of the program for HDR. With the fidelity of the light field maintained from capture through post production, the intended viewing experience is assured so long as that fidelity is maintained by the display in an uncontrolled viewing environment.

In contrast, with live broadcast the correspondence in fidelity between capture and viewing is dependent on decisions made in an uncontrolled environment at the point of capture. These decisions can produce unintended or intended consequences in the viewing experience, depending on how well the cameramen and the camera control unit operators understand the unique characteristics of HDR and to what extent those characteristics have been defined and adopted by the production community.

Production practices need to be adopted to ensure maximum creative benefits in the use of HDR extend all the way to make-up artists and set designers. Old ways of doing things can produce unrealistic results under the probing realism of HDR cameras.

Conforming Out-of-Program Dynamics to HDR

The standardization of HDR and its production practices also needs to be factored into advertising, graphic overlays, picture-in-picture applications and user interfaces to ensure a consistent, non-jarring viewing experience. Principles need to be developed based on the nuances of opto-physiological processes discussed in the ensuing section as well as basic common sense. For example, the industry does not want to repeat the experiences that led to the CALM (Commercial Advertisement Loudness Mitigation) Act by inserting HDR-enabled ads that don't conform to the parameters used with HDR programming.

Applying the Principles of Visual Science

There is a wide range of information to be applied from visual science in industry efforts to establish standards and practices relating to delivering the best possible HDR viewing experience to consumers. As discussed in the following section, these include:

- Processes of light and dark adaptation
- Effects of luminance and screen size on flicker perception
- The impact of luminance on color perception
- The impact of speed of adaptation on scene changes, program changes and commercials
- How light field and context effect perceptions of brightness and color

PROCESSES OF LIGHT AND DARK ADAPTATION

The human visual system adapts to 9 log units of light intensity[10]. Several physiological processes contribute to the human visual system's ability to adapt to this enormous range of light intensities. These include:

- Changes in pupil diameter
- The impact of luminance on retinal illuminance
- The impact on light absorption from the bleaching of photoreceptor photo pigments
- Speed of adaptation
- Contrast sensitivity

An understanding of these processes can be useful in addressing potential issues such as: sequential viewing of HDR and non-HDR content, maximum luminance and preservation of creative intent.

Changes in Pupil Diameter

Changes in the size of the pupil regulate the area through which light enters the eye by a factor of approximately 16. In fully dark-adapted conditions, the pupil of the eye has a diameter of approximately 8 mm. In fully light-adapted conditions, the pupil diameter reduces to approximately 2 mm.

The impact of a typical HD viewing environment, corresponding to a 30-degree stimulus on the size of the retinal image, is plotted in Figure 4 in accord with the Watson-Yellot[11] model. In this calculation a display having a luminance of 100 cd/m^2 may be expected to have a diameter of ~4.5mm. At 1000 cd/m^2, the pupil diameter can be expected to be ~2.5 mm, which would correspond to an approximately 3-fold reduction in pupil area.

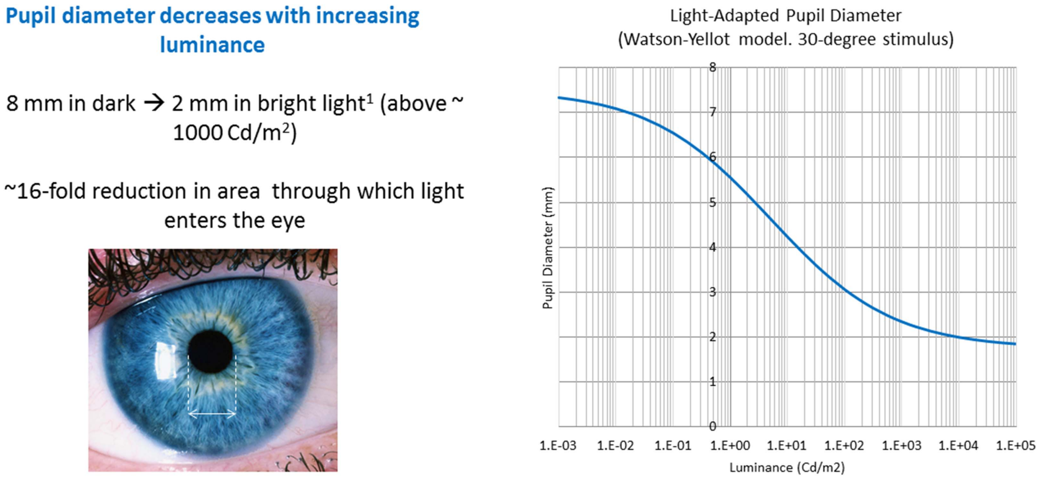


Figure 4. Illustration of the change in pupil size with luminance.

Between fully dark-adapted and light-adapted conditions, the pupil diameter changes in a graded manner. Exposure to bright light can reduce the area of the pupil, and consequently reduce retinal illumination, by approximately a log unit in a half second.

Pupil diameter can also be affected by ambient background light or lack thereof, which should be taken into account with the differences between cinematic and home viewing. Other factors affecting pupil diameter include the size of the visual stimulus, the position of the stimulus relative to the center of vision and the use of monocular or binocular stimuli, also all affect pupil diameter, all of which are considered in the Watson-Yellot model.

The Impact of Luminance on Retinal Illuminance

Pupil diameter as impacted by display luminance regulates the illuminance of the retina. Vision scientists use the term “troland” (Td), equal to luminance multiplied by pupil diameter, as a measure of retinal illuminance. Table 1 provides troland values for various luminance levels calculated using the Watson-Yellot model of pupil diameter for a visual stimulus of 30-degrees. (Note that between 10 and 1000 cd/m^2, luminance changes by a factor of 100 but retinal illuminance changes by a factor of approximately 30.)

Luminance (cd/m^2)	0.001	0.01	0.1	1	10	100	1000	10000
Pupil Diameter (mm)	7.3	7.1	6.5	5.6	4.2	3.1	2.3	2.0
Pupil Area (mm^2)	42.2	39.4	33.6	24.2	14.1	7.4	4.3	3.1
Retinal Illuminance (troland)	0.042	0.39	3.4	24	141	736	4309	31286

Table 1. Luminance, Pupil Size and Retinal Illuminance (troland)
(Pupil sizes predicted from Watson-Yellot model for 30-degree stimulus)

Both screen luminance and ambient background should also be considered when predicting pupil size and retinal illuminance. Although changes in pupil size account for a relatively small portion of visual adaptation, it is worth noting that reduction in pupil size also increases depth of focus and reduces glare.

Settings for HDR must take into account the impact of luminance on the retinal photoreceptors that determine illuminance in both bright and dark home environments and even outdoor situations in cases where the viewing experience is extended to handheld devices. The level of sensitivity and the speed of adaptation can vary considerably depending on those conditions.

The retina has four photoreceptor classes: rods; short-wavelength sensitive S-cones (“blue”); medium-wavelength sensitive M-cones (“green”); and long-wavelength sensitive L-cones (“red”). It is well known that rods are responsible for night vision and cones for daylight and color perception, but it is more accurate to think of visual adaptation in three categories that better reflect the gradual shift from rod-dominated vision to cone-dominated vision as light conditions brighten:

- Scotopic (below 0.001 cd/m^2) dominated by rods
- Mesopic (0.001 to 10 cd/m^2) mix of rods and cones
- Photopic (above 10 cd/m^2) dominated by cones

In darkened cinema theaters and home environments, mesopic-level adaptation might be a significant consideration. In bright home and mobile environments, photopic-level adaptation would be more typical.

Bleaching Adaptation

With the current generation of displays, we are just entering into a point where what is known as “bleaching adaptation” could become significant and impact a viewer’s experience. This is the phenomenon that occurs when the eye adjusts to sustained periods of illumination. The bleaching impact will be an important consideration in determining what average and peak levels of HDR brightness should be in the context of temporal shifts in luminance.

Vision is initiated by activation of light-sensitive photopigments. Photoreceptors contain a high concentration of light-sensitive biological pigment called rhodopsin in rods and cone-opsin in cones. The concentration of excitable photopigments is reduced during sustained illumination in bleaching adaptation.

The photopigment is made up of a protein (opsin) and a vitamin-A based chromophore. The chromophore absorbs a photon causing it to change shape, which in turn changes the shape of the surrounding protein. The altered protein ignites a powerful cascade of biological reactions that results in the light-induced neural response.

Once the chromophore absorbs a photon, it can no longer respond to light and needs to be recycled, which takes some time. During illumination, photopigments are used up and the optical density of the photoreceptors decreases; i.e., photoreceptors absorb less light because there are fewer chromophores available to catch photons.

This is when bleaching adaptation occurs (Figure 5). Over a period of steady illumination, the photopigment concentration will reach a new steady-state level[12] when the rate at which chromophores absorb photons is balanced by the rate at which they are replenished through the retina's recycling program.

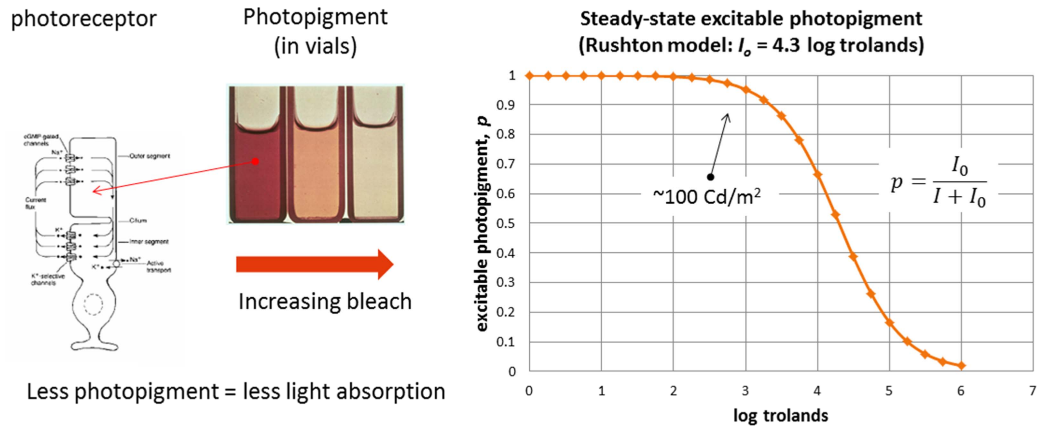


Figure 5. Illustration of the effect of luminance on photoexcitable photopigment if retinal photoreceptors.

<u>Speed of Response and Adaptation</u>

Rod-driven vision differs from cone-driven vision not only in terms of absolute sensitivity but also in terms of the speed at which light variations are signaled. The peak of the response of a dark-adapted rod to a flash of light occurs at approximately 120 msec[13].

The peak of a light-adapted rod response occurs at approximately 75 msec[13]. The peak of cones response occurs at approximately 20 msec[14]. The change in the time-scale of neural responses can be expected to play an important role in sensitivity to flicker and to motion judder.

Photoreceptors adapt to moderate non-bleaching step changes in increased illumination on the time scale of seconds[15, 16]. During this light adaptation process, the absolute sensitivity of photoreceptors decreases and the kinetics of responses increase in speed.

The time course of adaptation to decreased illumination (dark adaptation) depends on the intensity and duration of preceding stimuli[10]. If the preceding exposure was low enough that no significant photopigment bleaching occurred, then dark adaptation may also be measured on the time scale of seconds but slightly slower than light adaptation. For beaching adaptation, dark adaptation may take longer, on the order of tens of seconds to minutes. Dark adaptation also has two distinct phases: one is driven by recovery of cone sensitivity; the other even slower phase is driven by rod sensitivity.

In bright home and mobile viewing environments, both light and dark adaptation to modulations in illumination may be expected to proceed on a time scale measured in seconds. In dark home and theater environments, rapid changes going back and forth from mesopic-level to photopic-level luminance might result in slower dark adaptation.

EFFECTS OF LUMINANCE AND SCREEN SIZE ON FLICKER PERCEPTION

Flicker perception, or the sensitivity to temporal changes across video frame sequences, has not been an issue at the various frame rates used with video mapped to the dynamic ranges of ITU-R BT.709 on traditional screens. Yet with the onset of HDR on large-screen displays, flicker sensitivity could become an issue affecting the rate at which content is captured as well as the display frame and refresh rates.

The same is true of judder, the perception of uneven or jerky video playback that arises from movement of objects, edges or detail from one frame to the next. Increases in contrast, sharpness of detail and motion speed can cause judder, especially with increased screen sizes, which, for any given viewing distance, have the effect of bringing the images closer to the viewer.

Figures 6 and 7 illustrate the impact of luminance and screen size, respectively, on flicker. The critical flicker frequency (CFF) is the temporal frequency at which flicker is perceived as a steady light. CFF may be predicted from luminance with the Ferry-Porter Law[10], which states that the CFF increases in proportion to the logarithm of luminance.

Flicker sensitivity also depends on the location of the stimulus on the retina and the size of the stimulus. The temporal frequency threshold for flicker in the center of vision is higher than in the peripheral at all luminance levels. And the temporal frequency threshold for large stimuli (such as from a UHD display or cinema screen) is higher than for smaller stimuli at all luminance levels. (Granit-Harper law[10])

Given that TV manufacturers have increased frame rates in newer models to 120 and even 240 Hz, there's no reason to expect flicker or judder perception to be caused by HDR display systems. However, content captured at the 24 frame-per-second (fps) rate used with films and episodic series could be problematic. If the need to accommodate HDR leads the motion picture industry to raise the capture frame rates to, say, 48 fps, distributors will have to make adjustments based on the impact higher rates will have on bandwidth requirements for HDR content.

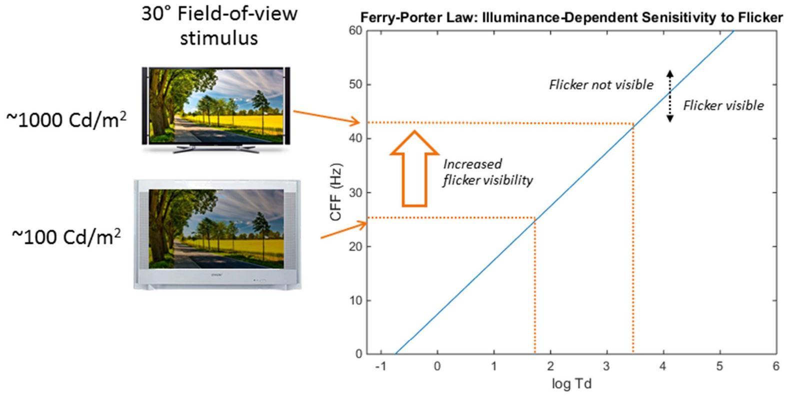


Figure 6. Illustration of the Ferry-Porter law. Brighter HDR displays may be expected to make non-smooth motion and flicker more noticeable.

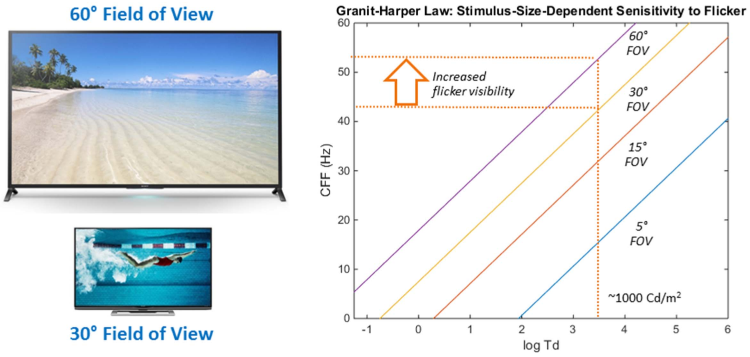


Figure 7. Illustration of the Granit-Harper law. Larger displays may be expected to make non-smooth motion and flicker more noticeable.

THE EFFECT OF LUMINANCE ON COLOR PERCEPTION

As luminance increases so does the ability of the human visual system to discriminate between colors at ever smaller gradations. As shown in Figure 8, color discrimination may be described in terms of MacAdam ellipses[17], a measure of the just-noticeable difference (JND) between colors. The size of MacAdam ellipses shrinks with increased luminance[18], which means that the JND decreases. Consequently, more bits would be needed to code color without introducing noticeable errors, particularly at high luminance. Thus, 10-bit encoding may be expected as a minimum bit depth for HDR for any color space (ITU-T BT.709, ITU-T 2020, or DCI P3).

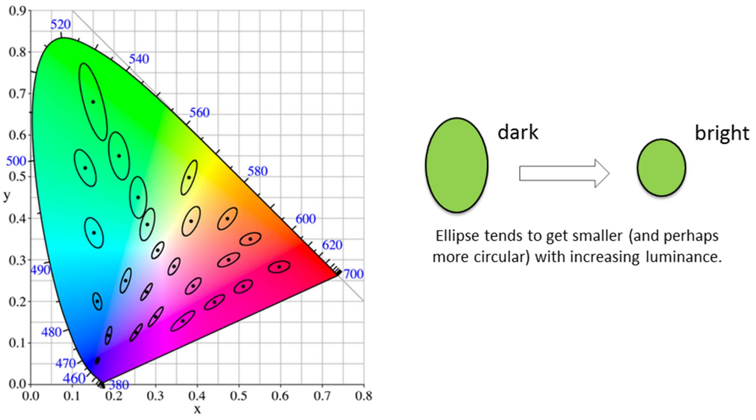
Luminance also affects the perception of the hue (Figure 9). The Bezold-Brücke[19] effect describes the perception of two stimuli having the same wavelength but different luminance as different hues.

EFFECTS OF ADAPTATION SPEED ON PROGRAM CHANGES AND COMMERCIALS

As detailed in the preceding discussion, the speed of light and dark adaptation depends on the level of retinal illuminance, the duration of illumination, the kinetics of the changes in pupil size, the rate of change of photoreceptor sensitivity, the rate of change of excitable photopigment, and the state of overall visual adaptation (scotopic, mesopic, or photopic).

All of this has important implication for the impact of rapid local or global luminance changes in television programming when HDR and non-HDR content is presented sequentially to viewers.

Here, it’s important that HDR-formatted commercials be pegged to the same parameter set for HDR programming. Conversely, non-HDR compliant commercials should not be placed with HDR programming.



Ellipse tends to get smaller (and perhaps more circular) with increasing luminance.

Figure 8. Illustration MacAdam ellipses that may be used to describe color discrimination.

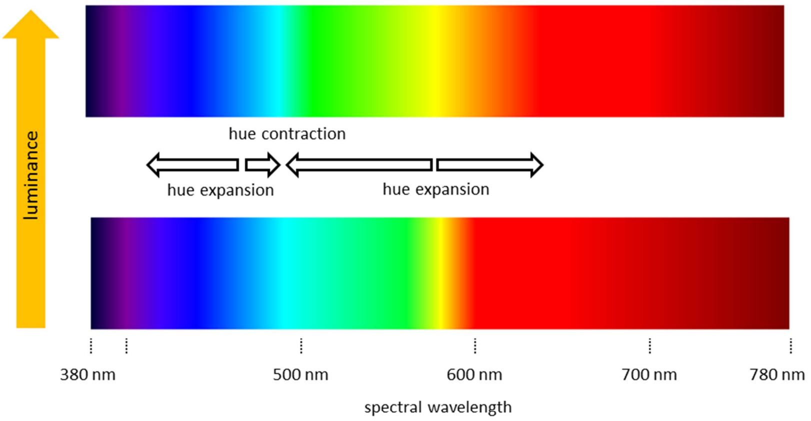
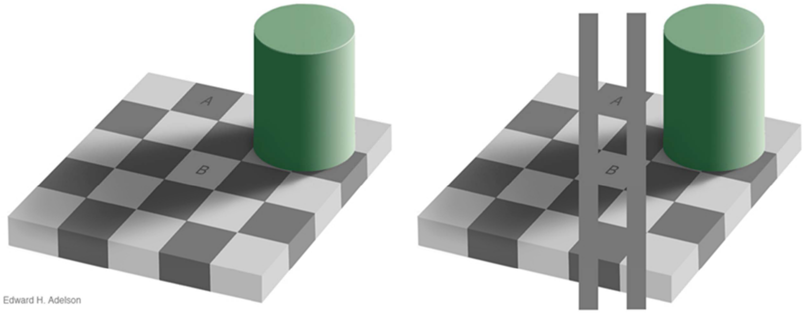


Figure 9. Illustration of the Bezold-Brucke effect in which luminance affects perception of hue.

THE EFFECT OF THE LIGHT FIELD AND CONTEXT ON BRIGHTNESS AND COLOR

Perceptions of brightness and color are not just a matter of luminance intensity. Variations in the light field across the frame can profoundly influence visual responses[20], as shown in Figure 10. (See also the highly compelling color illusion created by R. Beau Lotto[21]).



©1995, Edward H. Adelson. These checkershadow images may be reproduced and distributed freely.
http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html

Figure 10. Illustration that the visual system is more than simply a logarithmic light meter. The squares A & B are identical though they a perceived to be very different.

Such nuances will come into play as content producers move to using HDR in the creation process. They will need to avoid relying on luminance specifications alone in ensuring their creative intent is conveyed to viewers.

CONCLUSION

HDR is bringing a welcome transformation in the television viewing experience to the benefit of producers, distributors and consumers, provided the industry is careful to map parameters to the realities of human perception. Specifically, there needs to be general understanding with respect to the following points:

The requirements of live broadcast programming should be the threshold for setting HDR parameters:

- Dynamic contrast and color ranges will have to deliver an optimal viewing experience without reliance on intervening post-production processes.
- On-site production teams will have to incorporate understanding of HDR to ensure fidelity to creative intent.

The different rates of light and dark adaptation could play a significant role in QoE

- Light adaptation is fast. Dark adaptation is slower
- Mixed SDR and HDR could impact QoE for ad insertion, scene changes and program selection.
- Bleaching adaptation could significantly slow adaptation to changes in program luminance.

Bleach fractions might begin to be significant at 1000 cd/m2.

- They definitely will have an impact at multiple thousands of cd/m2

- It might be beneficial to limit average scene luminance to low bleaching fractions
- It might be beneficial to limit the duration of high-luminance highlights to minimize after images

Adaptation could impact user interaction:

- Graphic overlays might need to be tailored to the expected light/dark adaption state
- GUI could easily be impacted

HDR should not be considered in isolation. It impacts:

- Color perception with implications for bit depth and color gamut
- Sensitivity to flicker as a function of screen size, frame rate and display refresh rate

REFERENCES

1) ITU-R BT.709-5, "Parameter values for the HDTV standards for production and international programme exchange", Recommendation ITU-R, BT Series Broadcasting service (television), 2002.
2) NEMA PS3.14-2011 "Digital Imaging and Communications in Medicine (DICOM) Part 14: Grayscale Standard Display Function", NEMA 2011
3) SMPTE ST-2084 "High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays" SMPTE. 2014
4) Miller, S., Daly, S., and M. Nezamabadi. "Perceptual signal coding for more efficient usage of bit codes", SMPTE Motion Imaging Journal 122, 52{59 (May 2013)
5) Barten, P.G.J., Contrast Sensitivity of the Human Eye and its Effects on Image Quality. SPIE Optical Engineering Press: Bellingham, WA, 1999.
6) Borer, T. "Non-linear Opto-Electrical Transfer Functions for High Dynamic Range Television" BBC R&D White Paper. WHP 283. July 2014
7) ITU-R BT.1886 "Reference electro-optical transfer function for flat panel displays used in HDTV studio production" ITU-R, BT Series Broadcasting service (television), 2011.
8) ITU-R BT.2020. "Parameter values for ultra-high definition television systems for production and international programme exchange." Recommendation ITU-R, BT Series Broadcasting service (television), 2014.
9) SMPTE ST-431-1 (DCI P3). "D-Cinema Quality – Screen Luminance Level, Chromaticity and Uniformity" SMPTE 2006.
10) Levin, L.A., Nilsson, S.F.E., Ver Hoeve, J., Wu, S., Kaufman, P.L., and Alm, A. (2011) Alder's Physiology of the Eye, 11th edition. Saunders – Elsevier.
11) Watson, A.B. and Yellot, J.I. (2012) "A unified formula for light-adapted pupil size." J. Vis 12(10):12
12) Rushton, W.A.H. & Henry, G.H. "Bleaching and regeneration of cone pigments in man." Vision Research 8, 617-6311968
13) Friedburg, C., Thomas, M.M. and Lamb, T.D.(2001) "Time course of the flash response of dark- and light-adapted human rod photoreceptors derived from the electroretinogram." J. Physiol. 534.1, pp. 217-242.
14) Van Hateren, J.H. and Lamb T.D. (2005) "The photocurrent of human cones is fast and monophasic." BMC Neuroscience, 7:34.
15) Cao, L.-H., Luo D.-G., and Yau K.-W. (2014) "Light responses of primate and other mammalian cones." PNAS. Vol. 11 No. 7

16) Korenbrot, J.I. (2011) "Speed, adaptation, and stability of the response to light in cone photoreceptors: The functional role of Ca-dependent modulation of ligand sensitivity in cGMP-gated ion channels "JGP vol. 139, No. 1, 31-56.
17) MacAdam, D.L. "Speed, adaptation, and stability of the response to lightOpt Soc Amer. Vol 32, No. 5. 1942
18) Pridmore, R.W. and Melgosa, M. (2005) "Effect of luminance on color discrimination ellipses: Analysis and prediction data." Color Res. & App. Vol. 30, No. 3, 186-197
19) Pridmore, R.W. (1999) "Bezold–Brucke hue-shift as functions of luminance level, luminance ratio, interstimulus interval and adapting white for aperture and object colors" Vision Research 39, 3873-3891
20) Adelsen, E.H. "Checkershadow illusion." 1995 (http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html)
21) Lotto, R.B. www.lottolab.com University College London