High Dynamic Range imaging: is the game worth the candle?

Dennis Kuepper, Zofia Barańczuk, Ursina Caluori, Iris Sprow, Matthias Scheller Lichtenauer, Klaus Simon and Peter Zolliker, EMPA, Laboratory for Media Technology, Switzerland.

Abstract

High dynamic range imaging (HDRi) is a technology concerned with representing a range of luminances larger than state of the art displays and closer to luminance ranges occuring in natural scenes. We investigate whether the approach of evaluating a dynamic range by the number of just noticeable differences contained within does make sense in a HDRi workflow.
We found that disturbing effects of neighboring luminances were hardly perceivable on a standard dynamic range display if the background luminance exceeds $5\,\mathrm{cd/m^2}$.

Introduction

Representation in imaging media is limited by the spatial resolution, the temporal resolution and the sets of available colors, named gamut. Since the human visual system has a higher ability to distinguish luminance differences than chromatic differences on a spatial scale [1, 2, 3], the luminance information transfers the structure of an image representing a scene [4, 5]. The luminance range of an image is defined by the difference of the highest to the lowest luminance contained in it, while luminance contrast is defined by ratios of them. Contemporary standard displays achieve peak luminances between 80 and $400\,\mathrm{cd/m^2}$. In print, the reflectance of the paper can be reduced to 0.5% to 10% , depending on the paper type. Within this *dynamic range*, a number of luminance steps can be made out by the human visual system. The number of these luminance steps is the *dynamic resolution* of a visual system. The number of representable levels in an imaging system should be higher than the dynamic resolution, levels are often represented in the binary system, e.g. 256 representable levels are expressed as $\log_2(256) = 8\,\mathrm{bit}$ (Figure 1).

Distinction of gray levels is of particular interest in medical applications, since many medical imaging methods (tomography, mammography) deliver achromatic pictures only [6]. Taking the data on which the DICOM standard is based and assuming a peak luminance of $4000\,\mathrm{cd/m^2}$, approximately $2^{10} = 1024$ gray levels can be simultaneously differentiated under optimal conditions [7].

However, in images, neighboring high intensity regions can deteriorate the contrast detection performance. This is due to glare effects in the eye [8, 9]. Reflections of ambient light at a display surface can significantly influence contrasts in low intensity regions [10]. Since contrasts in low intensity regions of an image tend to be of importance for the diagnostic purpose, the studies of medical displays therefore often use very low luminance levels of background intensity and room illumination [9].

Glare effects have been quantitatively investigated in the first half of the 20th century with the equivalent background technique [11, pg. 578]. The Stiles-Holladay equation states that the influence of a point glare source E at eccentricity θ on an increment contrast threshold for a background luminance L leads to an equivalent increment contrast threshold as a background luminance β:

$$\beta = L + 10\frac{E}{\theta^2},\ \theta > 0.5\,\mathrm{deg}. \tag{1}$$

Assuming independent contributions, glare sources are additive. The Stiles–Holladay equation has later been extended for age effects, influence of ocular pigmentation and eccentricity angles larger than 30° [8, 12].

We were not interested in adding to this theory, but to test whether glare effects can be reproduced on displays when background luminance exceeds $5\,\mathrm{cd/m^2}$.
For this, we simulated a standard dynamic range display with a peak luminance below $200\,\mathrm{cd/m^2}$ on a high dynamic range display with a peak luminance exceeding $2000\,\mathrm{cd/m^2}$.

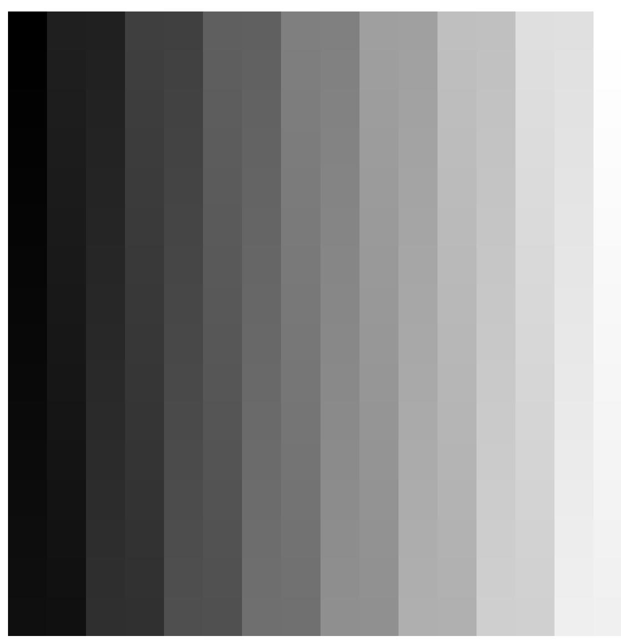
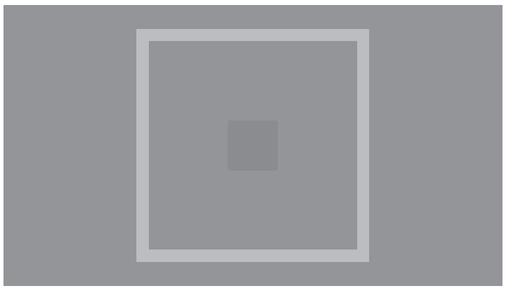
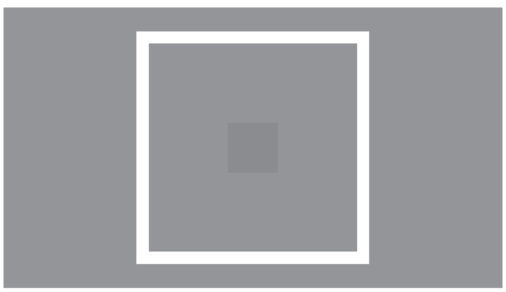


Figure 1. *A particular spatial arrangement of the 256 luminance steps available in 8bit sRGB. While one should hardly see any horizontal borders between adjacent squares represented in print or on standard displays, almost each border is clearly visible on a display with a high luminance range when the maximal luminance is chosen as a white point.*

(a) phase 1



(b) phase 2



(c) phase 3

(d) phase 4

(e) phase 5

Figure 2. *The interfaces in the five phases of the experiment. Images had the native resolution of the display (1920x1080px).*

Methods

We used a commercially available SIM2[1] high dynamic range display [13]. To investigate the influence of the background luminance and glare, we designed a decremental contrast detection experiment in five phases with varying interfaces (see Figure 2). The task of the observers was to detect whether the luminance of the central square was darker than the luminance of the background, i.e. whether the central square was visible at all. Seven observers with normal or corrected to normal visual acuity made a total of 2075 observations.

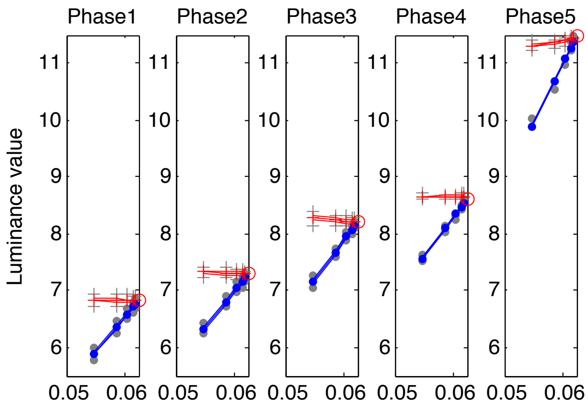
The experiment took place in a dark laboratory. Black curtains hung behind the observer and on both sides of the display to minimize backscattering of the light emitted by the display. Observers were seated in a car seat with the back of their head at 240 cm distance to fix their position relative to the screen. When calculating with a distance range of 200–240 cm, the central square covered a *visual angle* of more than 2°, while the square ring structures were at least at an *eccentricity* of 3° respectively 5° (screen width 105 cm). We used the 2° observer model in all calculations.

As stimuli, we used input images in the Radiance format ('.hdr') generated with Matlab. We conducted the experiment with three background values of 0.0625, 0.25 and 0.5, corresponding to luminances above 5, 25 and 50 $\mathrm{cd/m^2}$ respectively. The luminances of the square rings were chosen as values 2.0 and 32.0 in the input images, corresponding to approximately 200 $\mathrm{cd/m^2}$ in phases 2 and 4 and 3000 $\mathrm{cd/m^2}$ in phases 3 and 5. We measured the stimuli with a KonicaMinolta CS2000 spectroradiometer. Each measurement was immediately repeated three times in a row and averaged. Before and after each decrement measurement, we measured the background luminance at the same position. We did so since in the experiment, we displayed the background and the ring before and after decrements, either. As values tended to drift, we measured background luminance immediately preceding and following the luminance measurement of the decrement patch. The amount of decrement is defined as difference between the luminance measurement of the decrement patch and the average luminance of the two background measurements. We started with phase 1 stimuli and ended with phase 5. Between measurements of stimuli, we measured a patch with one of the ring luminances. The whole procedure was repeated a second time (Figure 3).

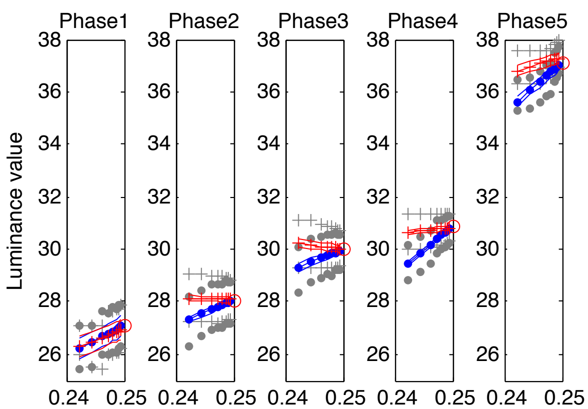
LCD displays generate an image by using a backlight and a front panel with locally varying transmittance. In the display we used, the backlight is generated by an array of white LED, each of which covers many pixels of the front panel. The high dynamic range of the display can be achieved by this locally varying illumination. The driver has to compensate the different background luminance by adapting the transmissivities in the front panel accordingly. As our measurements reveal, the luminance resulting from a particular input value is influenced by the luminance in a neighborhood of approximately half the display height.

To reduce possible resulting banding, we modulated the luminance signal of the input images with a sequence of $-1^n \frac{b}{16}$ along a Hilbert curve, where n is a running index along the curve and b is the input value to the background pixels.

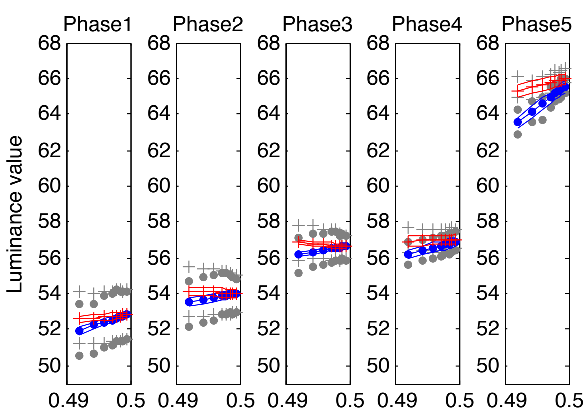
[1] http://www.sim2.com/HDR/



(a) background 0.0625



(b) background 0.25



(c) background 0.5

Figure 3. *Luminance measurements of background and decrements in units of* $\mathrm{cd/m^2}$. *The abscissa depicts the gray levels we stored in the Radiance image format used as input. Mean repeated background measurements are in red crosses at the abscissa position of the decrement they enclosed in the measurement sequence (see text), decrement measurements are in blue. Solid lines are at 1 e.s.d. around the position of the means.*

Results

We define the contrast detection threshold as the value for which the logistic regression curve from relative input values to proportion correct detections is equal 0.5. The detection thresholds did not significantly change when we divided the input value to the square by the input value to the background – with the exception of phases 4 and 5 (Figure 4). In phase 5, a larger decrement was needed for a reliable detection. The effect is most pronounced on the darkest background and lesser on the two brighter backgrounds. This is in accordance with the background effects postulated in [9] that were experimentally tested here. The same general picture remains when performing the regression against measured luminance ratios between background and decrement. Please note that the measured variance in stimulus production is not taken into account in the regression from measured values.

Let ΔL_t be the luminance contrast decrement threshold to the background and let L_{max} denominate the peak luminance in the scene. In our experiments, $10^{-5} < \Delta L_t / L_{max} < 10^{-4}$. Observers were able to reliably discriminate luminances three magnitudes lower than the peak luminance in the image.

Discussion

We simulated a standard dynamic range display on a high dynamic range display in phases 1, 2 and 4, but used higher luminance ranges in phases 3 and 5. The physical measurements of the display luminance with a KonicaMinolta CS2000 spectroradiometer reveal some drift and instabilities in stimulus production for the two brighter backgrounds (Figure 3). This variance in stimulus production can explain some of the variability for the two brighter backgrounds in Figure 4. We note as well that there is a significant amount of cross talk within the display, so the luminance of the stimuli in phase 5 is well above all other phases for each background. The luminances were measured with a tube from the spectroradiometer to the display [13], thus excluding any reflections of ambient light at the display surface (flare).

Within the precision of our experimental data, there is no significant influence of the surrounding ring on the contrast detection threshold in phases 1 to 4 for the darkest background where the effect should be most pronounced and stimulus production is most reliable (see top panel in Figure 3). It takes a significantly higher luminance relatively close to the stimulus as in phase 5 to significantly alter contrast detection thresholds.

The maximal perceptual volume of a grayscale can be estimated by the number of contrast difference thresholds or just noticeable differences contained within. The maximal perceptual volume assumes that the observer's visual system is optimally adapted to each stimulus within the volume. Using the method described in Figure 1, we can state that 256 grey levels are not enough to represent the just noticeable differences when peak luminance exceeds $2000\,\mathrm{cd/m^2}$ and the minimal luminance is around $5\,\mathrm{cd/m^2}$. Assuming the Weber-Fechner law holds for phases 1 to 4, one can estimate with a Fechnerian integration technique that there would be between 600 and 800 just noticeable differences with a peak luminance as we used in phase 3. Therefore, one has to use a bit depth exceeding 8bit in the luminance channel for high dynamic range imaging.

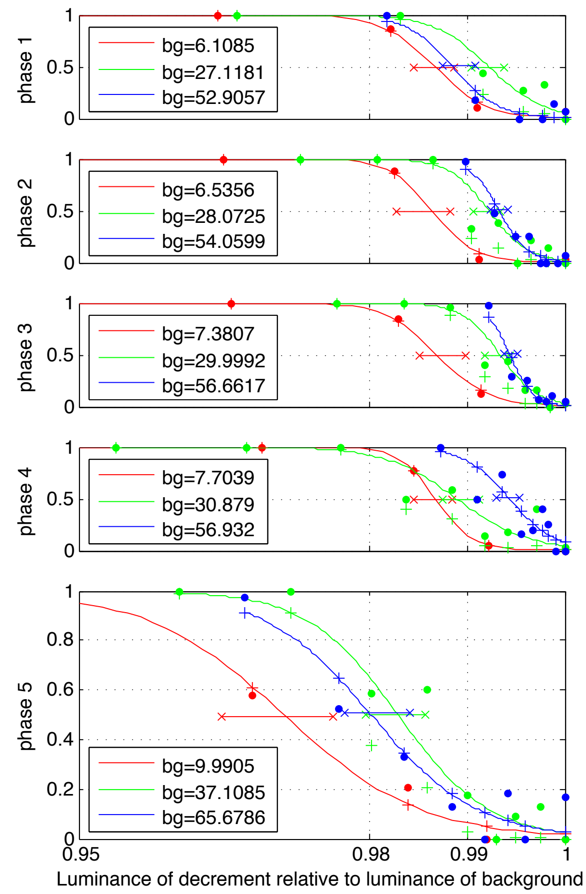


Figure 4. *Contrast detection thresholds for different phases and background input values. Threshold confidence intervals are bootstrapped 1000 times. Regression is from measured luminance values to proportion correct.*

If we assume that an observer remains adapted to a middle gray background of $50\,\mathrm{cd/m^2}$, luminance contrast detection thresholds significantly rise relative to a variable adaptation [14, Figure 43, pg. 81] and the perceptual volume of the gamut shrinks accordingly. But an observer's state of adaptation is hard to estimate, particularly if the images displayed are part of a stream, e.g. a movie. Comparing phase 5 on middle background with phase 4 on the brightest background in our experiments, we can assume that the effects of glare in our setup can be as important as the effects of adaptation in the cited source. Observers in our experiments are adapted in a mesopic to photopic range, since we could measure $< 1\,\mathrm{lx}$ at the position of the observer's head in phase 1 and about $40\,\mathrm{lx}$ in phase 5 on the darkest background (Sekonic i-346 illuminometer).
As our experiments reveal, one will need higher peak luminances than $200\,\mathrm{cd/m^2}$ or almost scotopically adapted observers to reproduce disability glare effects. The parallel processing of luminance differences, however, can be significantly affected with less extreme differences (discomfort glare) [15].

To reproduce effects of contrast reduction appearing in real world scenes in simulations on displays, minimizing crosstalk between neighboring pixels is necessary. Combining a spatially variable light source (projector) with a spatially variable reflectance of an achromatic screen might achieve this. A static version would project an RGB image on a printed grayscale version of the image.
Although not identical, a Kodak patent [16] and the transflective display [17] of the OLPC project use a similar idea.

Conclusions

In our experiments, disability glare on a background exceeding $5\,\mathrm{cd/m^2}$ is negligible in a standard dynamic range below $200\,\mathrm{cd/m^2}$ peak luminance. As these effects are of particular interest in simulation and design, using high dynamic ranges in these environments is a necessity.
On the other hand, glare effects lower perceivable details and this was not preferred by observers in our experiments on image quality resumed in [4]. As long as a higher resolution can improve perception of detail, we can therefore expect that consumers will prefer it to a higher peak luminance. Whether a high bit depth and a high dynamic range workflow is worth while therefore depends on the application.

Acknowledgments

The project was partially funded by the Swiss National Science Foundation Project 143226 and supported by the HDRi COST action IC1005.

References

[1] J.G. Robson, Spatial and temporal contrast-sensitivity functions of the visual system, J. Opt. Soc. Am., 56(8), pg. 1141–1142. (1966).
[2] F.W. Campbell and J.G. Robson, Application of Fourier analysis to the visibility of gratings, J. of Physiology, 197, pg. 551–566. (1968).
[3] K.T. Mullen. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings, J. of Physiology, 359, pg. 381–400. (1985).
[4] I. Lissner, J. Preiss, P. Urban, M. Scheller Lichtenauer and P. Zolliker, Image-Difference Prediction: From Grayscale to Color, IEEE Trans. on Image Processing, 22(2), pg. 435–446. (2013).
[5] S. Le Moan and P. Urban, Evaluating the perceived quality of spectral images, Proc. IEEE Int. Conf. on Image Processing, pg. 2024–2028, (2013).
[6] K.A. Fetterly, H. R.Blume, M.J. Flynn and E. Samei, Introduction to grayscale calibration and related aspects of medical imaging grade liquid crystal displays, J. of digital imaging, 21(2), pg. 193–207. (2008).
[7] T. Kimpe and T. Tuytschaever, Increasing the number of gray shades in medical display systems — how much is enough?, J. of Digital Imaging, 20(4), pg. 422–432. (2007).
[8] G. Spencer, P. Shirley, K.Zimmerman and D.P. Greenber, Physically-based glare effects for digital images, Proc. conf. on Computer graphics and interactive techniques, pp. 325-334. (1995).
[9] M. Choi, D. Sharma, F. Zafar, W.C. Cheng, L. Albani and A. Badano,

Does veiling glare hinder detection tasks in high-dynamic-range displays?, J. of Display Technology, 8(5), pg. 273–282. (2012)
[10] M. Scheller Lichtenauer, I. Sprow and P. Zolliker, Choice based experiments in multiple dimensions. Color Research & Application. 38(5), pg. 334–343. (2013).
[11] G. Wyszecki, W.S. Stiles, Color Science, Wiley, New York. (1967).
[12] J.J. Vos, On the cause of disability glare and its dependence on glare angle, age and ocular pigmentation, Clinical and Experimental Optometry, 86(6), pg. 363-370. (2003).
[13] I. Sprow, D. Kuepper, Z. Barańczuk and P. Zolliker, Image quality assessment using a high dynamic range display, Proc. AIC (pp. 307310). (2013).
[14] American Association of Physicists in Medicine (AAPM), Task Group 18, Assessment of Display Performance for Medical Imaging Systems, available at http://deckard.mc.duke.edu/ samei/tg18_files/tg18.pdf
[15] G. Bargary and J.L. Barbur, Parallel processing under conditions of discomfort glare, Proc. ECVP, pg. 24. (2013).
[16] W.Y. Fowlkes and L.W. Tutt, Image display for displaying a reflection image and a luminous image, U.S. Patent No. 6,573,661. (2003).
[17] J.J. Romero, The take-anywhere, do-anything display, Spectrum, IEEE 47.1, pg. 46-51. (2010).

Author Biography

Dennis Kuepper received his MSc in Computer Science from the ETH Zurich in 2008, where he had been focused on computational geometry, color science and scientific visualization. In his master thesis, he investigated algorithms for non-photorealistic rendering and the use of symmetries in 3D models to improve the automatic generation of line-drawings. He joined the Media Technology Lab in 2009 and has developed complete color workflows including device characterization, separation, tone mapping and gamut mapping.

Zofia Barańczuk earned her BSc in Computer Science and her MSc in Mathematics from the University of Warsaw. She worked on her doctorate in our group, researching psychophysical testing methods and gamut mapping. She received her PhD from the University of Jena. She is currently working on tone mapping algorithms for HDR images.

Ursina Caluori received her MSc in Computer Science from the ETH Zurich in 2008 where she specialized in Computer Graphics. In her master thesis she developed an alternative to current-established color management systems. Ursina is mainly concerned with OCR, color management and gamut mapping. She also supports the work in the psychophysical area.

Iris Sprow received her BSc in Imaging and Photographic Technology from the Rochester Institute of Technology in 2005. Originally trained as a photographer, she is currently involved with the visual evaluation of images. In 2009 she received her MSc from the University of the Arts London for her work in the area of online psychovisual testing.

Matthias Scheller Lichtenauer received his MSc in Computer Science from the ETH Zurich in 2008 with a focus on software engineering. In addition, his studies included courses in human factor design. His research work includes analytic image processing and the application of machine learning methods in psychometry. In parallel to his work as an engineer at Empa, he pursues a PhD at the University of Jena in Germany.

Klaus Simon studied Information Technology at the University of Saarland where, in 1987, he received his PhD in efficient data structures and algorithms. Until 1988 he was working as an R&D engineer at the Bodenseewerk Gerätetechnik GmbH with Perkin Elmer. In 1988 he joined the ETH (Swiss Federal Institute of Technology Zurich) as a senior research associate where he was appointed assistent Professor in 1992. Since 1999 he is the head of the Media Technology Lab at EMPA where he is devoted to efficient algorithms, probabilistic theory, image processing and teaching.

Peter Zolliker studied Physics at the ETH Zurich and received his PhD in Crystallography from the University of Geneva in 1987. After his postdoc position at the Brookhaven National Laboratory in New York, he joined the R&D team at Gretag Imaging in 1988 where he worked on image analysis, image quality as well as color management for digital and analogue printers. Since 2003 he is working at EMPA where he is engaged in color management and statistical analysis. He leads our activities in high dynamic range imaging.