

Perceptual display calibration

Rafał K. Mantiuk

May 7, 2017

This is a postprint version of the appendix included in the book *Displays: Fundamentals & Applications, Second Edition* by Rolf R. Hainich and Oliver Bimber, A K Peters/CRC Press 2016.

A good display design should make the inaccuracies and limitations of a display technology invisible to the human eye. However, this requires a good understanding of what is and what is not visible to the human eye. For that reason, display engineering is tightly linked to the understanding and modeling of visual perception.

Chapter 4 of this book introduced the basic limitations of visual perception that are commonly considered in display design. In this chapter we will take a closer look at selected perceptual models and present a few examples where they are used in practice.

As visual models operate on physical units of light, we first need to know how to convert the digital signals driving a display into photometric and colorimetric units. The display models used for such conversion are discussed in Sec. 1. The parameters for such display models can be found either by measuring display characteristics with a colorimeter, or with the help of visual calibration procedures if no measurement instruments are available, with the help of visual calibration procedures, discussed in Sec. 2. Once a displayed image is represented in photometric units, it is possible to predict visibility of small contrast patterns using a contrast sensitivity function (CSF). Sec. 3 explains when the CSF is the right visual model for the task and what its limitations are. The last section demonstrates how the CSF can be used to test for and reduce banding artefacts due to limited color channel bit-depth.

1 Display models

Whenever visual perception needs to be modeled for a display application, it is necessary to convert the displayed images or video into the right units. While displays operate in the digital signal domain, with pixels represented as triples of integers, most visual models expect an image to be represented in physical, photometric or colorimetric units. This is illustrated in Fig. 1: the display is driven by digital images to produce light, which is sensed by the human visual

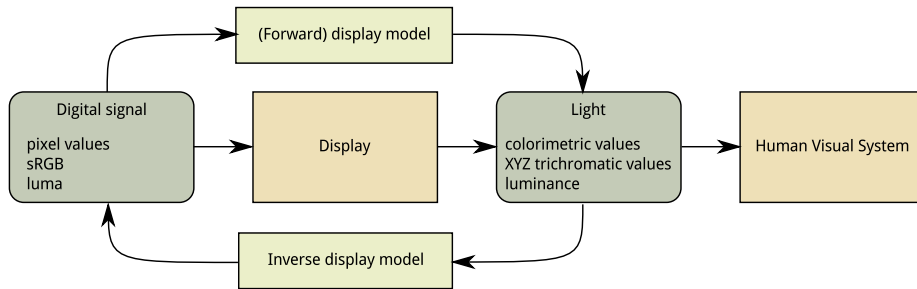


Figure 1: Forward display model simulates the process of transforming digital pixel values into physical light on a display. An inverse display model provides the inverse mapping — it can determine what combination of pixel values is needed to produce a certain color.

system. To simulate how the display transforms digital images into physical units of light, we need a forward display model. A forward display model is often referred as a display model without the word “forward”. To find the opposite transformation, from physical units into digital values, we need an inverse display model. This section discusses different variants of display models and how they account for display processing.

1.1 Gamma and sRGB

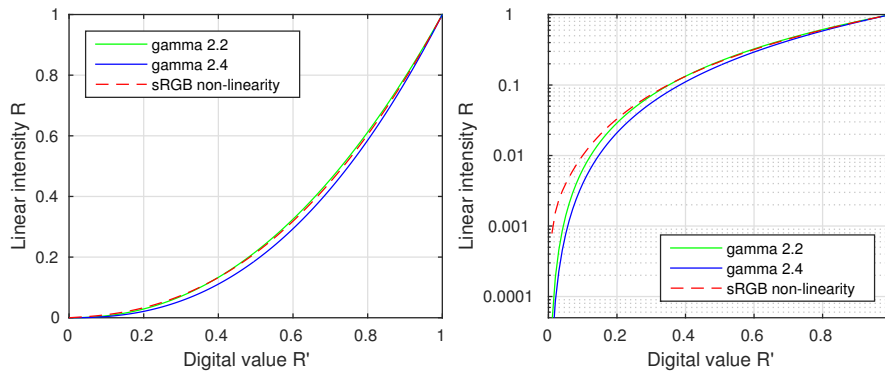


Figure 2: Comparison of gamma function with the sRGB non-linearity. The intensities are plotted using the linear scale on the left and the logarithmic scale on the right. The logarithmic scale is often used for plotting luminance as it better reflects perceived intensity of light.

The simplest forward display model, explaining the relation between digital and physical units, can be modeled with a so-called *gamma correction* function (refer to Section 6.2.7). For color images, the gamma transformation has the

form:

$$R = R'^{\gamma}, \quad G = G'^{\gamma}, \quad B = B'^{\gamma}, \quad (1)$$

where R' is a red pixel value in the range 0–1, and R is the corresponding red component in the physical space. The gamma value varies from 1.8 to 2.4 and its value is usually close to 2.2.

The R , G and B values in Eq. 1 are sometimes referred as *linear* RGB values because they are linearly related to colorimetric values and to luminance. This is to distinguish them from “gamma corrected” R' , G' and B' , which are the pixel values driving a display.

The gamma correction function is a reasonable approximation of a display model, especially for CRT monitors, but it tends to be inaccurate for low digital signal values. The sRGB standard refines the gamma correction function to better model a typical monitor’s response and is used for most computer monitors. The sRGB non-linearity is given by:

$$R = \begin{cases} \frac{R'}{12.92} & \text{if } R' \leq 0.04045 \\ \left(\frac{R'+0.055}{1.055}\right)^{2.4} & \text{otherwise} \end{cases}. \quad (2)$$

The equations for blue and green color components are analogous. Though the equation for the sRGB non-linearity seems very different from the gamma correction function in Eq. 1, both functions are in fact quite similar as shown in Fig. 2. Note that the sRGB follows more closely the shape of gamma 2.2, even though the sRGB equation contains exponent 2.4. The main difference between the gamma 2.2 and the sRGB non-linearity is best seen in the plot on the right, which uses the logarithmic scale for linear intensity: the intensities modeled by the sRGB non-linearity are larger. The difference between luminance values 0.001 and 0.0001 may seem negligible, however, the human eye can easily detect such small differences.

The gamma correction (Eq. 1) and sRGB (Eq. 2) functions are known as transfer functions, or more specifically electro-optical transfer functions (EOTFs). The name EOTF has historical roots in analog television, where the function described the relation between electrical signal used to drive a cathode ray tube (CRT) display and the light intensity produced by screen’s phosphors. The display technologies used today, such as liquid crystal displays (LCD) or organic light-emitting diode (OLED) displays, have a different response to electrical current than CRT displays. Yet, the display driving electronics of LCD or OLED emulates the behavior of a CRT display not only to remain backward compatible, but also because the gamma correction function well approximates the perceived brightness of light; the digital R' , G' and B' values result in more uniform perceived differences between different intensities of color than their linear counterparts, R , G and B .

1.2 Color transformation

A transfer function, discussed in the previous section, is only the first part of a display model. The second part, discussed in this section, explains rendering

of colors in terms of device independent CIE XYZ color values (refer to Section 4.3).

Most display models assume a linear relationship between the linear RGB values and absolute CIE XYZ trichromatic values:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = L_{peak} \cdot M_{RGB \rightarrow XYZ} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (3)$$

where L_{peak} is the peak luminance of a display given in cd/m^2 , and $M_{RGB \rightarrow XYZ}$ is a 3×3 color transformation matrix.

To create the color transformation matrix it is enough to know display primaries: CIE XYZ coordinates for pure red, green and blue colors emitted by a display. The matrix is constructed by concatenating XYZ values of the display primaries as columns of that matrix. This way the RGB triplet $[1 \ 0 \ 0]$ will result in the XYZ triplet corresponding to the red primary, and green and blue colors will be mapped similarly. For example, any HDTV display should have its primary colors closely matching ITU-R recommendation BT.709 and the chromaticity coordinates:

Chromaticity	Red	Green	Blue
x	0.6400	0.3000	0.1500
y	0.3300	0.6000	0.0600
Y	0.2126	0.7152	0.0722

The chromaticity coordinates can be transformed into XYZ trichromatic values using the formulas:

$$X = \frac{x}{y} \cdot Y, \quad Z = \frac{1 - x - y}{y} \cdot Y \quad (4)$$

From that, we get that the color transformation matrix for a BT.709-compliant display is:

$$M_{RGB \rightarrow XYZ} = \begin{bmatrix} X_{red} & X_{green} & X_{blue} \\ Y_{red} & Y_{green} & Y_{blue} \\ Z_{red} & Z_{green} & Z_{blue} \end{bmatrix} = \begin{bmatrix} 0.412424 & 0.357579 & 0.180464 \\ 0.212656 & 0.715158 & 0.072186 \\ 0.019332 & 0.119193 & 0.950444 \end{bmatrix}. \quad (5)$$

The matrix for inverse transformation, $M_{XYZ \rightarrow RGB}$, can be obtained by inverting the matrix $M_{RGB \rightarrow XYZ}$.

1.3 Gamma-offset-gain model

Both the gamma function and the sRGB non-linearity serve as good inverse display models (for driving the displays), but are not ideal as forward display models (for predicting display performance). This is because hardly any display can achieve intensity values as low as predicted by both models. For example, for pixel value 0, both sRGB and gamma transfer functions predict intensity 0. However, most display technologies emit some small amount of light even if the

pixel values are set to 0. Furthermore, unless a display is viewed in a pitch-dark room, some ambient light will be reflected from the screen thus elevating the amount of light that the user senses.

This inaccuracy of both transfer functions is not just an engineering approximation but it is an intentional part of the design. The primary role of the transfer functions is to specify how to drive the display given digital input. When the pixel value is 0, the display should emit the smallest possible amount of light, regardless of what that smallest amount is for a particular display. If a transfer function assumed a certain minimum quantity of light, no display would be allowed to produce deeper black level than specified by the transfer function.

To predict the light emitted by a display, better accuracy is offered by a gamma-offset-gain (GOG) display model [4] or one of its variations: GGO (gain, gamma, offset), or GOGO (gain, offset, gamma, offset) [8]. Those variations describe the same functional form of a model but using different parameters. The GOG model for gray-scale images is the relation between luma (gray-scale pixel) value and emitted luminance, and is modeled as

$$L = (L_{peak} - L_{black})V^\gamma + L_{black} + L_{refl} , \quad (6)$$

where L is luminance and V is luma, where luma varies between 0 and 1 (as opposed to 0–255). For color images, L and V could be replaced by linear R , G or B component and V by R' , G' , or B' pixel values. L_{peak} is the peak luminance of a display in a completely dark room, L_{black} is the luminance emitted from black pixels (*black level*), and L_{refl} is the ambient light that is reflected from the surface of a display, sometimes known as ambient flare. γ is a parameter that controls non-linearity of a display. For LCD displays L_{black} varies in the range from 0.1 to 1 cd/m² depending on the display brightness and contrast. L_{refl} depends on the ambient light in an environment and can be approximated in the case of non-glossy screens with:

$$L_{refl} = \frac{k}{\pi} E_{amb} , \quad (7)$$

where E_{amb} is the ambient illuminance in lux and k is the reflectivity for a display panel. The reflectivity is below 1% for modern LCD displays and can be slightly larger for OLEDs.

The inverse of the model takes the form:

$$V = \left[\left(\frac{L - L_{black} - L_{refl}}{L_{peak} - L_{black}} \right)^{\frac{1}{\gamma}} \right]_0^1 , \quad (8)$$

where the square brackets are used to denote clamping values to the range 0–1 (the values greater than 1 become 1, and less than 0 become 0).

Fig. 3 shows some examples of displays modeled by Eq. 6. Note that ambient light can strongly reduce the effective dynamic range of the display (top-left plot). The gamma parameter has no impact on the effective dynamic range,

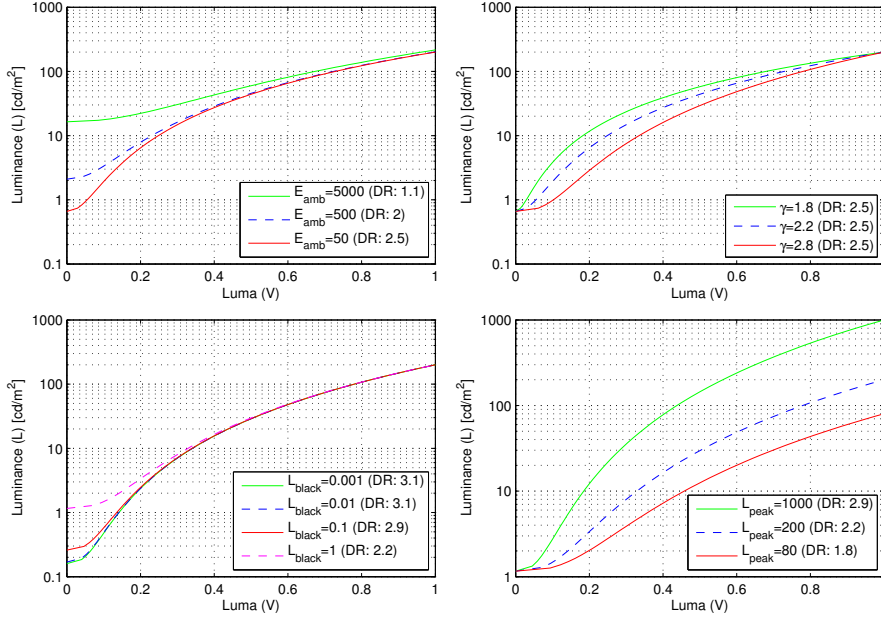


Figure 3: The relation between pixel values (luma — V) and emitted light (luminance — L) for several displays, as predicted by the GOG model from Eq. 6. The corresponding plots show the variation in ambient light, gamma, black level and peak luminance in the row-by-row order. The DR value in parenthesis is the display dynamic range as log-10 contrast ratio (equal to $\log_{10}(L_{max}/L_{min})$). The parameters not listed in the legend are as follows: $L_{peak}=200$ cd/m², $L_{black}=0.5$ cd/m², $\gamma=2.2$, $E_{amb} = 50$ lux, $k = 1\%$.

but higher γ values will increase image contrast and make the image appear darker (top-right plot). Lowering the black level increases effective dynamic range to a certain level, then has no effect (bottom-left). This is because the black in most situations will be “polluted” by ambient light reflected from the screen. Brighter display can offer higher dynamic range, provided that the black level of a display remains the same (bottom-right).

1.4 Other display models

GOG is a relatively simple display model, which was found to be adequate for CRT displays [27]. It can be classified as a physical model because it models the physical process of transforming a digital signal (or voltage) into light emitted from a CRT display [6]. The GOG model is also a good approximation for LCDs, as most LCDs try to emulate CRT displays. However, for some LCDs the GOG model introduces inaccuracy due to color channel interactions [27,33], known as cross-talk. Cross-talk is caused by the driving signal of one channel

affecting the other color channel as the electrical charge is applied.

Several display models address the issue of cross-talk between color channels in LCDs. One of them is known as a *masking model* [27]. The model relies on additivity of colors in the CIE XYZ color space. The method requires that a look-up table (LUT), mapping from digital values to XYZ, is stored for red, green, blue, cyan, magenta, yellow and white colors. The idea behind the model is best demonstrated in an example. Let us consider a mapping of an RGB triplet value [1 0.5 0.1] to CIE XYZ color values. First, the gray contribution of the color is extracted by looking up [0.1 0.1 0.1] value in the white LUT. Second, the yellow contribution is estimated by computing the difference in XYZ colors between pixel values [0.5 0.5 0] (*yellow* table) and [0.1 0.1 0.1] (*white* table). Then, the red contribution is computed as a difference between [1 0 0] (*red* table) and [0.5 0.5 0] (*yellow* table) XYZ values. Finally the gray, yellow and red contributions are added together in the XYZ space, resulting in the XYZ coordinates for the RGB pixel [1 0.5 0.1]. Although the procedure is relatively straightforward for the forward display model, it is much more difficult to apply in an inverse model as each step is conditional on the input RGB pixel values.

The physical display models, such as GOG or a masking model, need to make certain assumptions and thus are restricted to “well-behaved” displays, which adhere to those assumptions. If a display violates those assumption, it can be still modelled a 3D look-up table, mapping from RGB pixel values to CIE XYZ color values. This approach, however, requires large amounts of measurements and memory. For example, a relatively sparse grid of $10 \times 10 \times 10$ points, requires 1000 measurements. The accuracy of 3D LUT approach will depend on the interpolation method. For example, the accuracy is usually higher if the interpolation is performed in a uniform color space, such as CIE Lab [6]. The inverse display model usually requires a separate LUT, which cannot be directly measured with a photospectrometer or colorimeter. Such an inverse 3D LUT, from CIE XYZ to digital values, is usually obtained by extracting data from the forward display model.

If the cost of storing 3D LUT in memory is too high, the mapping can be approximated an analytic function. Such function may take a form of a polynomial [33], or a linear combination of radial basis functions [6]. Such representations are more compact than a 3D look-up table, offers higher flexibility than physical models, such as GOG, but can be more computationally demanding than the other two approaches.

The display models discussed so far can be used for a basic colorimetric or photometric calibration, however, they do not account for many other factors that affect the colors of displayed images. For example, the black level of a display is elevated by the luminance of neighboring pixels due to internal glare. Also, some display technologies (e.g. plasma displays, some HDR displays) need to manage a limited power budget and vary peak brightness with image content. In that case, a small white patch shown on a dark background can have much higher luminance than the same patch shown on a large bright-gray background. The models discussed above, however, account for most effects and are relatively accurate for modern LCD displays, which is the dominant display technology

at the moment.

2 Visual display calibration

Displays are normally calibrated by collecting sample colors using color measuring instruments, such as colorimeters or photospectrometers, and then fitting the collected data to one of the display models. A detailed description of display characterization procedures can be found in a freely available standard [26]¹. If, however, a measuring instrument is not available, there are several visual procedures that can help estimate the parameters of a display. Such procedures are discussed in this section.

2.1 Gamma calibration

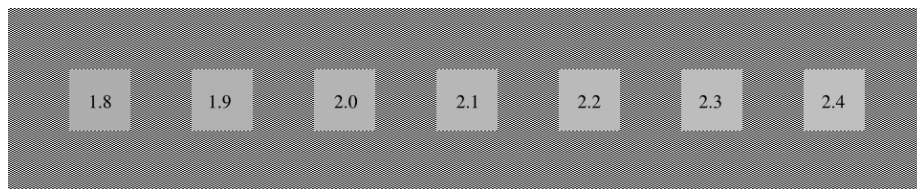


Figure 4: A chart that can be used to visually estimate the gamma exponent of a monitor. The task is to match the brightness of the patterned background to one of the uniform color patches with numbers. The numbers indicate the corresponding gamma values. The match is easier to make when a chart is seen from a large distance or when vision is not in focus (e.g. when looking on the thumb in front of a display plane instead of a display). Note that the chart needs to be enlarged so that the pixels in the pattern match the pixels on the screen, otherwise aliasing artefacts could make the match impossible.

A simple way to visually assess the value of a display gamma exponent is to match the brightness of a black and white pattern, containing the same amount of white and black pixels, with a uniform gray level. An example of a gamma chart for such matching is shown in Fig. 4. Gamma matching charts can be often found in software, in the display calibration settings of an operating system or a graphics card driver, but very few users know how to use those charts effectively. The goal is to see a blend of the white and black pattern so that it is possible to make a match with a uniform gray level. If the pattern appears sharp, most users are not able to make a brightness match. Therefore, it is advisable to look at those charts with defocused vision — without glasses, from a long distance, or by focusing the eye on a point in the front of a display instead of the display plane.

¹The standard documents and the test patterns can be downloaded from <http://www.icdm-sid.org/>.

When the white and black pattern is blended in the eye, it produces the luminance that is arithmetic mean of the luminance of black and white pixels. From the GOG display model (Sec. 1.3) we have that the luminance of a white pixel is:

$$L(1) = L_{peak} + L_{refl} , \quad (9)$$

and the luminance of a black pixel is:

$$L(0) = L_{black} + L_{refl} . \quad (10)$$

Now, we need to make the average luminance for the mixture of while and black lines equal to the luminance of the matching gray-level pixel value V_{match} , so that:

$$\begin{aligned} \frac{L(0) + L(1)}{2} &= L(V_{match}) \\ \frac{1}{2}(L_{peak} + L_{black}) + L_{refl} &= (L_{peak} - L_{black})V_{match}^\gamma + L_{black} + L_{refl} \quad (11) \\ \gamma &= \frac{\log(0.5)}{\log(V_{match})} . \end{aligned}$$

Therefore, having the pixel gray-value V_{match} that matches in brightness alternating white and bright lines let us easily estimate the value of the display gamma, even if we do not know anything else about the display, including black level, peak luminance and environment in which it is seen.

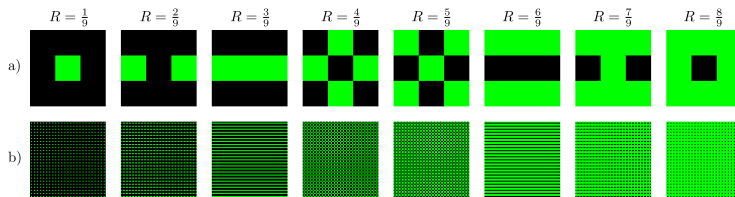


Figure 5: Half-tone patterns that could be used to recover the shape of a display transfer function. R is the ratio of “on” pixels to all pixels in a pattern. Row a) shows the pattern design and row b) shows how the pattern appears on a screen. The same pattern can be produced for red, green and blue primaries to recover individual transfer functions.

The same procedure may be used not only to find a single gamma parameter, but also to determine an arbitrary complex shape of a transfer function [25, 34] if it is substantially different from the GOG model. The matching must be done for 8–10 half-tone patterns, each containing different ratio of on and off pixels. An example of such half-tone patterns from [34] is shown in Fig. 5. If the ratio of “on” pixels to the total number of pixels is R , the perceived luminance of the pattern is R times the display peak luminance. The actual transfer function can be recovered by fitting a parametric display model [34], or by solving an optimization problem with a smoothness regularization term [25].

2.2 Color calibration

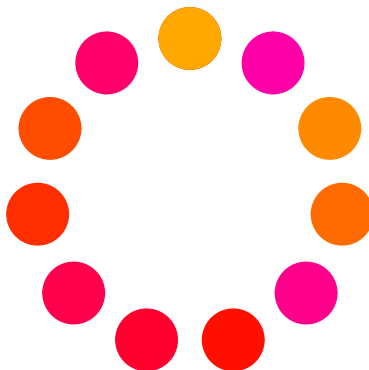


Figure 6: A single step from a unique-hue selection task, used to color-calibrate a display without measuring instruments. The task for the user is to select one color that has the least amount of yellow or blue tint. A sequence of such selections will indicate the unique red color.

The previous section explained how the first part of a physical display model, the transfer function, can be visually measured. This section explains the procedure for measured the second part of the model, a color transformation matrix.

Most human observers can adjust a color patch so it appears neither red nor green, or neither yellow nor blue. This ability to find unique hues can be used to color-calibrate the display without a measuring device [15].

A schematics of an experiment that could be used to find unique hues is shown in Fig. 6. The user is given a task to indicate a colored circle that has the least amount of blue or yellow tint. The colors are sampled from the HSV (hue-saturation-value) color space by varying hue component and keeping saturation and value fixed. The selection procedure can be repeated several times for the same color, each time narrowing the variation in hue to improve accuracy. The same procedure is run for red, green, blue and yellow colors, and also for the neutral gray point.

The experiment gives the RGB pixel values for four unique hues and the color of neutral gray. Those can be used to find a color transformation matrix from the measured display to another display, for which unique hues are already known [15]. The difficulty here is that the experiment reveals only the position of the unique hue planes but it does not indicate what was saturation or value (brightness) of the selected colors. Because of that, additional regularization terms and constraints are needed to find the transformation matrix.

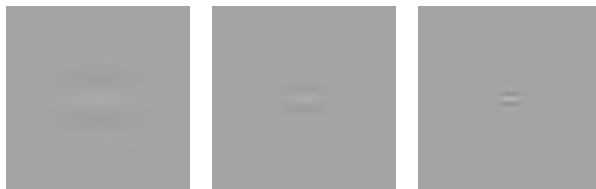


Figure 7: Gabor patches of different spatial frequency, typically used in detection experiments.

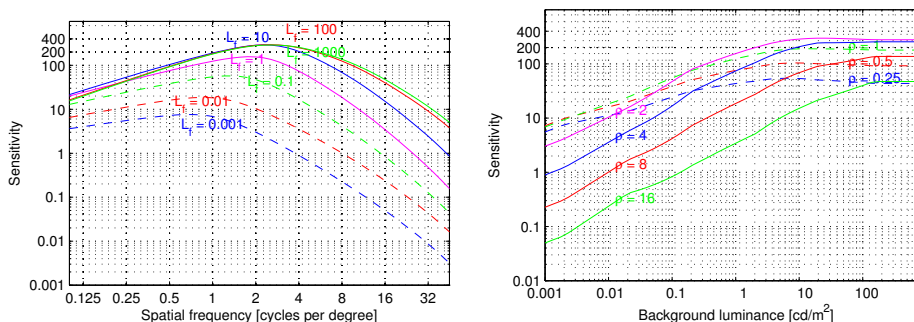


Figure 8: Contrast Sensitivity Function (CSF) plotted as the function of frequency (left) and luminance (right). Different line colors denote different background luminance (L_f), or spatial frequency (ρ). The plots are based on the model from [18] and data from [16].

3 Contrast sensitivity

The previous sections explained how to map digital pixel values into physical quantities of light, which can be used in visual models. This section will explain in more detail one particular visual model — a Contrast Sensitivity Function (CSF), which was briefly introduced in Section 4.2.4.

Contrast sensitivity explains how well a human observer can detect patterns of certain frequency on a uniform background. Such patterns could be for example sinusoidal waves modulated by a Gaussian envelope, so called Gabor patches, shown in Fig. 7. The sensitivity S is defined as the inverse of the detection threshold:

$$S = \left(\frac{\Delta L}{L} \right)^{-1} = \frac{L}{\Delta L}, \quad (12)$$

where L is the luminance of the background, and ΔL is the amplitude of the pattern (the difference between the maximum and mean luminance). The sensitivity predicted by the CSF, depends on the number of parameters, such as spatial frequency, temporal frequency, luminance of the background, orientation of the pattern, size of the pattern, and distance from the fovea. In this section we focus on the two parameters that influence sensitivity the most: spatial

frequency ρ and the luminance of the background L : $S = CSF(\rho, L)$.

The plot of the CSF as the function of those two parameters is shown in Fig. 8. The plot reveals that we are not very sensitive to low and high frequency patterns, and we are most sensitive to spatial frequencies between 1 and 5 cycles per visual degree (cpd) for high luminance. The peak sensitivity gradually shifts to about 1 cpd for low luminance. The CSF shows that the drop of sensitivity with luminance is more rapid for high frequencies. Indeed, small high frequency details become invisible at night when the visual system needs to rely on rod vision.

A large number of CSF measurements and models can be found in the literature [2, 3, 16, 24, 29, 30]. Most of the historical data was collected in order to gain insights into the visual system rather than provide practical models. Such data was often collected for artificial stimuli and conditions: monochromatic light (single wavelength) [30], corrected for aberrations [24], artificial pupils and monocular viewing. Consequently, such data may not explain how we perceive spatial patterns on actual displays seen in natural conditions. Therefore, it is important to ensure that the CSF models and data represent similar stimuli to those used in a particular application. For example, the data from [29] was collected for a white broadband light and the data from [16] was measured using side-by-side rather than sequential presentation of patterns, which is more relevant for the applications involving static images.

Barten derived a physically plausible model of the CSF [2], which is one of the most comprehensive works in this area, shown to well predict numerous sensitivity datasets. The original work was limited to photopic (daylight) vision but it was later extended to much lower luminance levels in [3]. Another comprehensive CSF model can be found in [7]. Those models were created by fitting analytical formulas to multiple datasets. In contrast to those, the model from [18] relies on a single dataset [16], measured using side-by-side comparisons.

It is important to recognize that the actual shape of the CSF can vary substantially between different models and measurements. This is due to multiple factors, such as differences in experimental procedures, stimuli and its presentation. Therefore, CSF should not be regarded as a rigid ground-truth model of the visual system sensitivity, but rather as a relative characteristic that often needs to be adapted to a particular application. The simplest form of such adaptation is a global change of sensitivity. The CSF curves, shown in Fig. 8 can be shifted along the vertical axis to account for different detection tasks. For example, detecting banding artefacts on a display can be more difficult and result in lower sensitivity than detecting Gabor patches in a well-controlled experiment. The peak sensitivity values varies between 300 and 500 for well controlled experiments, which correspond to the smallest detectable contrast between 0.2% and 0.33%. In most practical applications a suitable sensitivity is closer to 100 or less, which translates into detection contrast of 1% or higher.

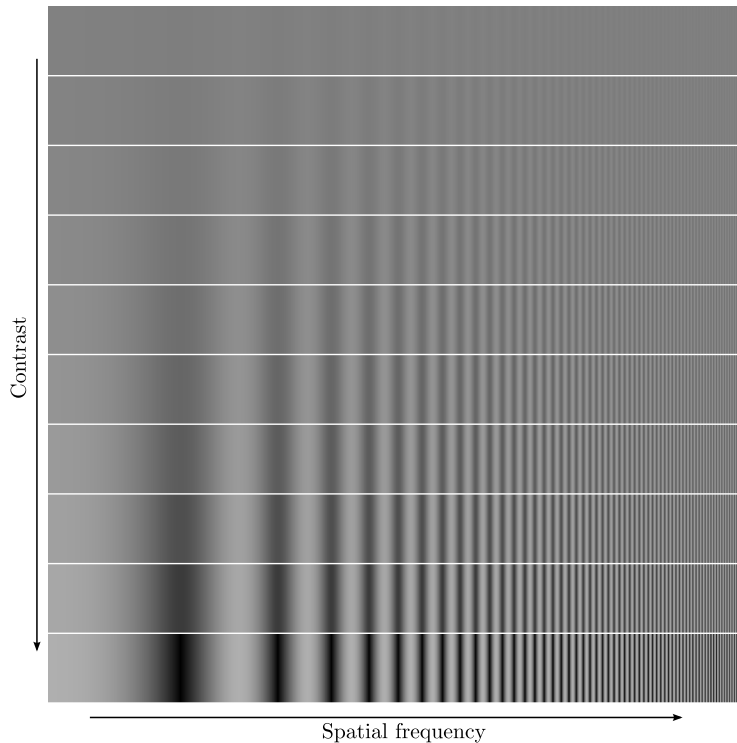


Figure 9: The spatial contrast appearance chart. For very low contrast in the top rows, the perceived magnitude of contrast varies with spatial frequency. But the perceived contrast magnitude remains mostly the same for bottom rows containing large contrast.

3.1 Contrast constancy

The CSF plotted as a function of frequency, such as the one on the left in Fig. 8, may resemble a modulation transfer function (MTF). An MTF is often used to describe the loss of resolution in optical systems as the function of spatial frequency. If we were to create a simplistic visual model, we could be tempted to use the CSF as a linear filter and weight each spatial frequency band in an input image by the values from the CSF to simulate a “perceived” image. This, however, would produce false results because the CSF is not a linear filter and it is valid only for very small, barely visible contrast.

The pattern shown in Fig. 9 demonstrates the CSF (also shown in Figure 4.4) but also its limitations. The perceived magnitude of contrast in the top rows varies with the frequency. This produces an illusory boundary between visible and invisible contrast, which defines the shape of the CSF. However, the perceived magnitude of contrast remains the same for the bottom rows, in which the physical contrast magnitude is much larger. Those rows do not contain the

same variation in perceived contrast as the top rows. This demonstrates *contrast constancy*, which is the ability of the visual system to perceive the same magnitude of contrast regardless of spatial frequency [12]. If the visual system could be explained by a linear filter, we would be able to see the same reduction of perceived contrast magnitude in top and bottom rows. An important consequence of contrast constancy is that CSF is applicable only to patterns of very low contrast, close to the detection thresholds, or close to the illusory CSF line in Fig. 9. If the contrast magnitude is much different from the detection contrast, the CSF should not be used to predict how perceived contrast changes.

Since the CSF only applies to a barely visible contrast, its usefulness may seem limited. In practice, however, many visual distortions found in display applications consists of contrast of low amplitudes, for which the CSF prediction are valid. Some examples of those distortions are display non-uniformity or banding due to quantization, discussed in Sec. 4. The CSF is also an important component of many supra-threshold² contrast perception models, such as models of contrast masking [17], or perceived contrast across the luminance range [31]. Therefore, CSF is a fundamental building block for modeling human vision.

3.2 Thresholds across the luminance range

Many visual artifacts, such as banding due to quantization discussed in the next section, can reveal themselves across a range of spatial frequencies, so it is impossible to select a single frequency that would be suitable for all cases. Instead, it is better to rely on a conservative estimate of the detection threshold T by selecting the peak of the CSF at a given background luminance L :

$$T(L) = \frac{L}{\max_{\rho} CSF(\rho, L)}. \quad (13)$$

The function $T(L)$ predicts the smallest detectable contrast at luminance L and it is plotted in Fig. 10. The contrast that is less than the function $T(L)$ (below the dashed line) is assumed to be invisible to the human eye. The plot shows that it is more difficult to see small contrast differences at low luminance levels.

The function $T(L)$ gives a conservative estimate, which means that actual visibility thresholds in images could be higher than predicted by the function. In rare cases they could be also lower because of the contrast masking phenomenon known as facilitation [17].

4 Quantization and bit-depth

The display models presented in Sec. 1 were missing an important processing step present in any digital display — quantization due to limited bit-depth. If the number of bits representing color is insufficient, a displayed image can reveal

²Supra-threshold refers to the stimuli whose contrast is much above the detection threshold.

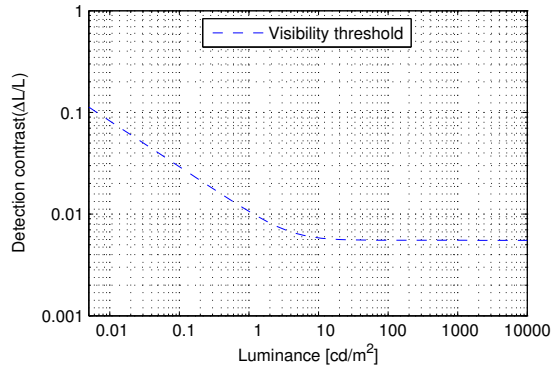


Figure 10: The smallest detectable difference in luminance across the luminance range. The plot is based on the data from [16].

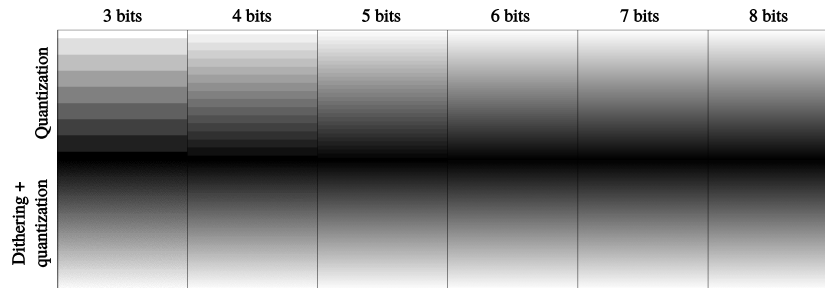


Figure 11: The effect of quantization on a gradient of pixel values. Each column contains an image of different bit-depth, from 3 to 8 bits. The lower part of the gradient was generated using dithering to simulate higher bit-depth.

banding artifacts, such as those shown in Fig. 11. Such artifacts can be masked at lower bit-depths by introducing dithering — the use of patterns that are the mixture of two consecutive pixel values, which give an illusion of intermediate intensity levels. An example of banding artifacts that has been masked by dithering is shown in the bottom of Fig. 11.

Dithering is so effective that many displays employ it to deliver smooth gradation of gray levels on low-bit-depth panels. For example, a display could be driven by a 8-bit signal from a PC but have only a 6-bit liquid crystal (LC) panel. In this case, the display control board converts input into a 6-bit driving signal for the LC panel and uses the remaining least significant bits to introduce dithering. The dithering can be added in both spatial and temporal domains. Temporal domain dithering is introduced by rapidly switching between pixel values in consecutive frames.

Bit-stealing is another technique for improving bit-depth resolution [28]. The technique involves a small adjustment between the three color components so

that intermediate luminance levels are produced at the cost of a small error in color. For example, a pixel color [11 10 10] (using 0–255 range) results in luminance that is higher than for pixel [10 10 10], but lower than for pixel [11 11 11]. As the red component contributes about 21% to luminance, the pixel [11 10 10] with elevated red component produces about 0.21 fractional step in luminance between [10 10 10] and [11 11 11]. In practice, an optimum pixel combinations can be found for every intermediate luminance level so that the distortion in chroma is minimized. The technique works well, especially for gray-scale images, because the visual system is more sensitive to luminance than to chroma variations.

Although both techniques can help extend the perceived bit-depth of a display, the display still requires an input signal of sufficient accuracy to generate dithering patterns or to jitter pixel values for bit-stealing. This raises the question of what bit-depth is sufficient to represent digital images so that no banding artifacts are visible.

4.1 Quantization errors

In this section we combine the display models introduced in Sec. 1 and the visibility threshold function from Sec. 3.2 to analyze the visibility of banding artefacts.

Fig. 12 shows the maximum quantization errors for a relatively bright display with the peak luminance of 500 cd/m² and the black level of 0.1 cd/m². The maximum quantization errors are computed as:

$$Q_e(L) = \frac{1}{L} \cdot \left(D \left(D^{-1}(L) + \frac{0.5}{2^b - 1} \right) - L \right), \quad (14)$$

where D is the transfer function transforming luma (0–1) into luminance (in cd/m²), and D^{-1} is the inverse transfer function transforming luminance into luma. The formula computes the difference between luminance affected by maximum quantization error (0.5 in the luma domain), and accurate luminance, L . The multiplication by $\frac{1}{L}$ brings the values into the same relative contrast units as used for the visibility thresholds in Fig. 10.

Fig. 12 compares four transfer functions: linear scaling into the 0–1 range (effectively no transfer function), the logarithmic function, the gamma 2.2 power function and the sRGB non-linearity. All transfer functions result in visible quantization errors for 8-bit encoding (plot on the left) and only logarithmic encoding brings errors below the visibility thresholds when 10-bit encoding is used (plot on the right). This result is consistent with everyday experience, where quantization errors are easily noticeable on regular 8-bit displays. The plots show that gamma and sRGB are well-aligned with the visibility thresholds for darker colors but they waste bits when encoding brighter colors. The logarithmic curve results in the opposite behavior — it allocates too many bits for darker colors. The lack of any transfer function (*Linear* label in Fig. 12) is the worst choice, with the most visible quantization errors.

Quantization errors are even more visible on a high dynamic range display, producing luminance in the range from 0.005 cd/m^2 to $10\,000 \text{ cd/m}^2$, as shown in Fig. 13. 10-bit logarithmic encoding results in banding at higher luminance levels and 2.2 gamma is clearly inadequate for all except the brightest pixels. The floating point encoding (black line) results in very small quantization errors but at the cost of high bit-depth.

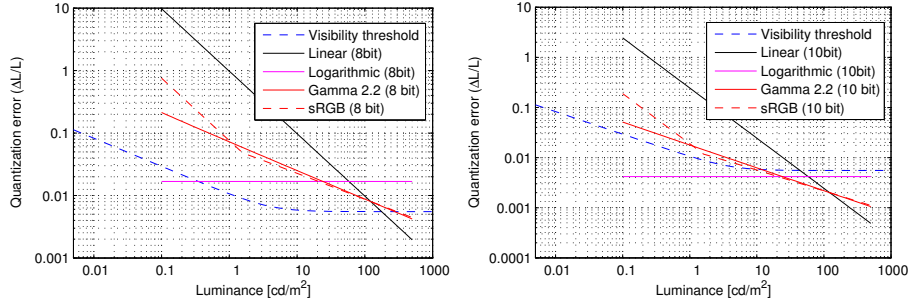


Figure 12: Quantization errors on low dynamic range display ($L_{peak}=500 \text{ cd/m}^2$, $L_{black}=0.1 \text{ cd/m}^2$) for selected transfer functions. The errors are shown for 8-bit encoding in the plot on the left and for 10-bit encoding in the plot on the right.

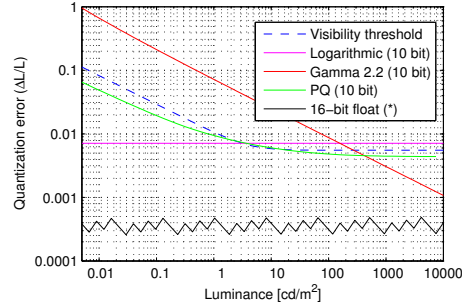


Figure 13: Quantization errors on high dynamic range display ($L_{peak}=10\,000 \text{ cd/m}^2$, $L_{black}=0.005 \text{ cd/m}^2$) for selected transfer functions and 10-bit encoding.

4.2 Perceptual transfer functions

To make the best use of available bit-rate, the transfer function should be aligned with the visibility threshold function $T(L)$. This idea is widely used in medical imaging where certified medical clinical displays have a transfer function scaled in just-noticeable-differences, so called DICOM gray-scale standard display function [1]. Such perceptual transfer function was later extended to high dynamic range (HDR) color images and used for video compression [19, 20]. A

modified version of such encoding, derived from Barten’s CSF model [3] and known as Perceptual Quantizer (PQ), was proposed as a transfer function for HDR content with a convenient analytical formula [23].

Perceptual transfer functions are straightforward to derive from the visibility threshold function $T(L)$. The goal is to find a function mapping luminance into an abstract response so that the increase in the response by 1 corresponds to the increase of luminance by the just-noticeable-difference ΔL . This condition is met when the derivative of the unknown response function R is specified as:

$$\frac{dR}{dL}(L) = \frac{1}{\Delta L(L)}. \quad (15)$$

Note that the detection threshold $\Delta L(L)$ is the function of luminance L . From Eq. 13 we get that $\Delta L(L) = T(L) \cdot L$ and hence the response function is:

$$R(L_x) = \int_{L_{min}}^{L_x} \frac{1}{T(L) L} dL, \quad (16)$$

where L_{min} is the lowest luminance encoded, mapped to the response value 0. The detailed derivation can be found in [21, sec. 2.4]. The resulting transfer function is shown in Fig. 14 and also in Figure 4.5 in Chapter 4.

Fig. 13 includes the quantization errors for the PQ perceptual transfer function [23]. Although different psychophysical data was used to derive the PQ function and to plot the dashed visibility threshold curve, both curves are relatively similar to each other. The study in [5] showed a slight advantage of the PQ over other perceptual transfer functions in terms of uniformity of banding artifacts across the luminance range.

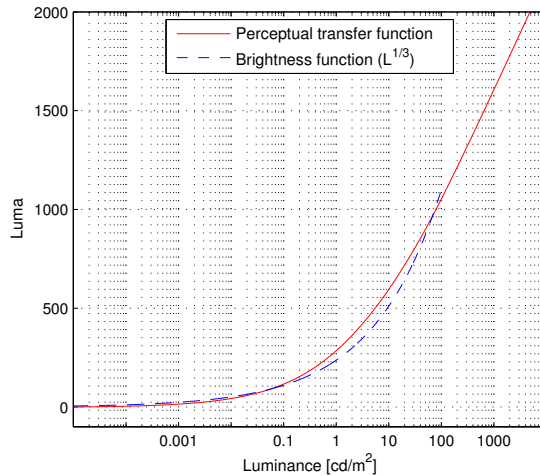


Figure 14: Perceptual transfer function compared with the brightness function (Stevens’ law for brightness).

Fig. 14 shows an interesting property of the perceptual transfer function. The shape of the transfer function is relatively close to the relation between luminance and perceived brightness (Steven’s law). This demonstrates, that the perceptual transfer function not only reduces visibility of quantization errors, but it also improves perceptual uniformity of the resulting luma values. Note that Steven’s power brightness function was demonstrated to hold only for lower luminance levels [13, p.101] and therefore the brightness function is plotted up to 100 cd/m^2 .

5 Summary

The quality of any display technology is ultimately judged by the human eye. Therefore, it is essential that visual perception is considered in any display-related application. This chapter introduced models and methods that help to characterize a display and its performance in perceptual terms.

First, we showed how most displays can be modelled using a gain-offset-gamma model (Sec. 1). A display model provides a link between digital signal driving a display and light emitted from the screen, which is eventually perceived by the display user. Selected parameters of such display model can be found using visual calibration procedures, discussed in Sec. 2. The contrast sensitivity function (CSF) is one of the most commonly used visual models when investigating perceptual aspects of a display design. Such function, however, is applicable only to very small contrast distortions, as discussed in Sec. 3. One successful application of CSF is testing for the visibility of banding artefacts, caused by limited bit-depth of the pixels (Sec. 4). The same visual characteristic can be also used to drive a perceptual transfer function, which minimizes the visibility of banding artefacts for a given bit-depth.

This chapter provided insight into a small portion of the wider field of perceptual display design and calibration. A general overview of perceptual considerations in display design can be found in Chapter 4 of this book and also in the survey paper [22]. The visibility of more complex spatial distortion, such as display non-uniformity, can be tested using perceptual image difference metrics, such as VDP [7] or HDR-VDP [18]. The code for the latter metric is freely available.³ When considering video, it is important to account for spatio-temporal contrast sensitivity of the visual system [32] and the temporal aspects of a display [14]. If assessment of color artefacts is important, some level of perceptual scaling can be achieved using uniform color spaces or color appearance models [10, 11]. Testing for distortions in binocular stereo displays may require disparity difference metrics [9]. This diversity of visual models, specialized in different aspects of vision, demonstrates the challenge of visual modeling. While specialized visual models, focused on a particular kind of distortion, have been successful used in many areas, there is no single general visual model, which could account for all aspects of vision and could be applicable to a wide range of problems.

³HDR-VDP-2 can be downloaded from <http://hdrvdp.sourceforge.net/>

References

- [1] PS 3.14-2003. DICOM Part 14: Grayscale Standard Display Function. Technical report, National Electrical Manufacturers Association, 2003.
- [2] Peter G. J. Barten. *Contrast sensitivity of the human eye and its effects on image quality*. SPIE Press, 1999.
- [3] Peter G. J. Barten. Formula for the contrast sensitivity of the human eye. In Yoichi Miyake and D. Rene Rasmussen, editors, *Proc. SPIE 5294, Image Quality and System Performance*, pages 231–238, dec 2004.
- [4] Roy S Berns. Methods for characterizing CRT displays. *Displays*, 16(4):173–182, 1996.
- [5] Ronan Boitard, Rafał K. Mantiuk, and Tania Pouli. Evaluation of color encodings for high dynamic range pixels. In *Human Vision and Electronic Imaging*, page 93941K, 2015.
- [6] Philippe Colantoni, Jean-baptiste Thomas, and Jon Y Hardeberg. High-end colorimetric display characterization using an adaptive training set. *Journal of the Society for Information Display*, 19(8):520, 2011.
- [7] S.J. Daly. Visible differences predictor: an algorithm for the assessment of image fidelity. In Andrew B. Watson, editor, *Digital Images and Human Vision*, volume 1666, pages 179–206. MIT Press, 1993.
- [8] Ellen A. Day, Lawrence Taplin, and Roy S. Berns. Colorimetric characterization of a computer-controlled liquid crystal display. *Color Research and Application*, 29(5):365–373, 2004.
- [9] Piotr Didyk, Tobias Ritschel, Elmar Eisemann, Karol Myszkowski, Hans-Peter Seidel, and Wojciech Matusik. A luminance-contrast-aware disparity model and applications. *ACM Transactions on Graphics*, 31(6):1, nov 2012.
- [10] Garrett Johnson Erik Reinhard, Erum Arif Khan, Ahmet Oguz Akyuz. *Color Imaging: Fundamentals and Applications*. CRC Press, 2008.
- [11] Mark D. Fairchild. *Color Appearance Models*. John Wiley & Sons, second edition, 2005.
- [12] M A Georgeson and G D Sullivan. Contrast constancy: deblurring in human vision by spatial frequency channels. *J. Physiol.*, 252(3):627–656, nov 1975.
- [13] Alan Gilchrist. *Seeing Black and White*. Oxford University Press, 2006.
- [14] Hao Pan, Xiao-Fan Feng, and S. Daly. LCD motion blur modeling and analysis. In *IEEE International Conference on Image Processing 2005*, pages II–21. IEEE, jul 2005.

- [15] D Karatzas and S M Wuerger. A Hardware-Independent Colour Calibration Technique. *Annals of the British Machine Vision Association*, 2007(3):1–10, 2007.
- [16] Kil Joong Kim, Rafał Mantiuk, and Kyoung Ho Lee. Measurements of achromatic and chromatic contrast sensitivity functions for an extended range of adaptation luminance. In Bernice E. Rogowitz, Thrasyvoulos N. Pappas, and Huib de Ridder, editors, *Human Vision and Electronic Imaging*, page 86511A, mar 2013.
- [17] G. E. Legge and J. M. Foley. Contrast masking in human vision. *Journal of the Optical Society of America*, 70(12):1458–71, dec 1980.
- [18] Rafał Mantiuk, Kil Joong Kim, Allan G. Rempel, and Wolfgang Heidrich. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Transactions on Graphics*, 30(4):40:1—40:14, jul 2011.
- [19] Rafał Mantiuk, Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. Perception-motivated high dynamic range video encoding. *ACM Transactions on Graphics (Proc. of SIGGRAPH)*, 23(3):733, aug 2004.
- [20] Rafał Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. Lossy Compression of High Dynamic Range Images and Video. In *Human Vision and Electronic Imaging*, page 60570V, 2006.
- [21] Rafał K. Mantiuk, Karol Myszkowski, and Hans-peter Seidel. High Dynamic Range Imaging. In *Wiley Encyclopedia of Electrical and Electronics Engineering*, pages 1–81. Wiley, 2015.
- [22] Belen Masia, Gordon Wetzstein, Piotr Didyk, and Diego Gutierrez. A survey on computational displays: Pushing the boundaries of optics, computation, and perception. *Computers & Graphics*, 37(8):1012–1038, dec 2013.
- [23] S. Miller, M. Nezamabadi, and S. Daly. Perceptual Signal Coding for More Efficient Usage of Bit Codes. *SMPTE Motion Imaging Journal*, 122(4):52–59, may 2013.
- [24] Kathy T Mullen. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *The Journal of Physiology*, 359(1):381–400, feb 1985.
- [25] Attila Neumann, Alessandro Artusi, Georg Zotti, Laszlo Neumann, and Werner Purgathofer. An interactive perception-based model for characterization of display devices. *Proceedings of SPIE*, 5293:232–241, 2003.
- [26] SID. *Information Display Measurement Standard*. Society for Information Display, 1.03b edition, 2012.

- [27] Nobuhiko Tamura, Norimichi Tsumura, and Yoichi Miyake. Masking model for accurate colorimetric characterization of LCD. *Journal of the Society for Information Display*, 11(2):333, 2003.
- [28] Christopher W. Tyler. Colour bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis. *Spatial Vision*, 10(4):369–377, jan 1997.
- [29] A. van Meeteren and J.J. Vos. Resolution and contrast sensitivity at low luminances. *Vision Research*, 12(5):825–IN2, may 1972.
- [30] Floris L. van Nes and Maarten A. Bouman. Spatial Modulation Transfer in the Human Eye. *Journal of the Optical Society of America*, 57(3):401, mar 1967.
- [31] Robert Wanat and Rafał K. Mantiuk. Simulating and compensating changes in appearance between day and night vision. *ACM Transactions on Graphics*, 33(4):147, 2014.
- [32] A. B. Watson. High Frame Rates and Human Vision: A View Through the Window of Visibility. *SMPTE Motion Imaging Journal*, 122(2):18–32, mar 2013.
- [33] Senfar Wen and Royce Wu. Two-primary crosstalk model for characterizing liquid crystal displays. *Color Research and Application*, 31(2):102–108, 2006.
- [34] Kaida Xiao, Chenyang Fu, Dimosthenis Karatzas, and Sophie Wuerger. Visual gamma correction for LCD displays. *Displays*, 32(1):17–23, 2011.