System gamma as a function of image- and monitor-dynamic range

David Kane

Department of Information and Communication Technologies, Universitat Pompeu Fabra, Roc Boronat, Barcelona, Spain

Marcelo Bertalmío

Department of Information and Communication Technologies, Universitat Pompeu Fabra, Roc Boronat, Barcelona, Spain

System gamma is the end-to-end exponent that describes the relationship between the relative luminance values at capture and the reproduced image. The system gamma preferred by subjects is known to vary with the background luminance condition and the image in question. We confirm the previous two findings using an image database with both high and low dynamic range images (from 10^2 to 10^7), but also find that the preferred system gamma varies with the dynamic range of the monitor (CRT, LCD, or OLED). We find that the preferred system gamma can be predicted in all conditions and for all images by a simple model that searches for the value that best flattens the lightness distribution, where lightness is modeled as a power law of onscreen luminance. To account for the data, the exponent must vary with the viewing conditions. The method presented allows the inference of lightness perception in natural scenes without direct measurement and makes testable predictions for how lightness perception varies with the viewing condition and the distribution of luminance values in a scene. The data from this paper has been made available online.

Introduction

A simplified characterization of the imaging pipeline is presented in Figure 1 and consists of two gamma nonlinearities: the encoding gamma (γ_{enc}) and the decoding gamma (γ_{dec}). The encoding gamma is applied to the image recorded by a camera prior to quantization and storage (Figure 1a, b) and is compressive ($\gamma_{dec} < 1$) for all low bit-depth image formats. The decoding gamma is a property of the media (monitor/projector/print) in question and is expansive ($\gamma_{enc} > 1$). As quantization occurs after the encoding gamma and before the decoding gamma, the effective quantization rate is finer at low luminance levels, than at high (Figure 1c). This approximately aligns the quantization rate to human sensitivity, a process known as perceptual linearization.

The product of the encoding and decoding gammas is known as the syst\em gamma (γ_{sys}) and describes the relationship between an image recorded by a camera's sensor and the image displayed on a monitor. Early display engineers assumed that the system gamma should be linear ($\gamma_{sys} = 1$; e.g., van de Poel & Valeton, 1954). However, it has been known for over 60 years that the system gamma is often set to a value of greater than 1, as illustrated in Figure 1d. In the 1940s, the recommended system gamma for photographic projections, typically viewed with a dark surround, was 1.6 (Jones, 1944), but between 1 and 1.1 for printed images, typically viewed with a light surround (Breneman, 1962; Hunt, 1987). Today, the recommended system gamma for a cinema display (typically viewed with a the dark surround) is 1.5, and the system gamma for an office display (typically viewed with a light surround) is 1.1 (Poynton, 2012). Thus the preferred system gamma (as a function of viewing condition) has remained constant over the last 60 years.

Histogram equalization is achieved when the intensity values in an image are equally distributed over the full range. Achieving some degree of histogram equalization in an image is seen as desirable for some applications (Gonzalez & Woods, 2002). However, as described in the preceding paragraph, the standard imaging pipeline for consumer imagery applies a system gamma of greater than 1, when a value of less than 1 is needed to achieve histogram equalization for real-world scenes that have a mean luminance of less than half of the total range (Olshausen & Field, 2000).

Citation: Kane, D., & Bertalmío, M. (2016). System gamma as a function of image- and monitor-dynamic range. *Journal of Vision*, *16*(6):4, 1–13, doi:10.1167/16.6.4.

doi: 10.1167/16.6.4 Received July 31, 2015; published April 12, 2016 ISSN 1534-7362

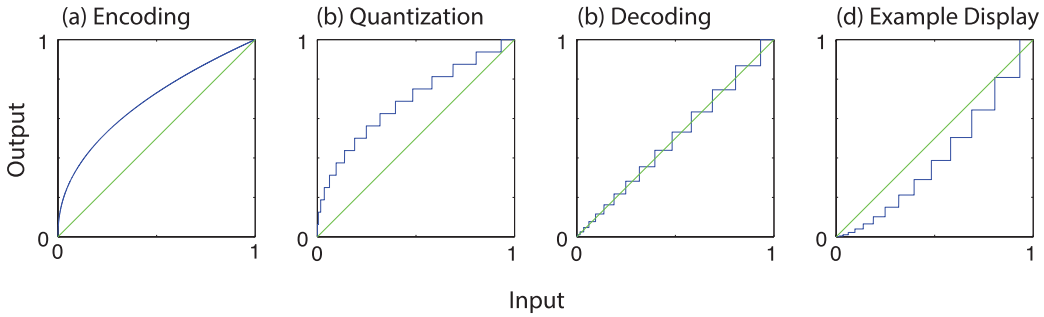

Figure 1. The grayscale imaging pipeline from capture to presentation. (a) The image recorded on a camera sensor is passed though a compressive encoding gamma. (b) Quantization is applied (for illustration purposes we quantize at 4 bits). (c) An expansive decoding gamma relinearizes the image. (d) However, the decoding gamma used in consumer displays is typically greater than the inverse of the encoding gamma, thus the end-to-end exponent, known as the system gamma, is expansive. In (d) we illustrate the standard system gamma used for cinema displays, which is 1.5 (Poynton, 2012).

Dynamic range refers to the ratio between the highest and lowest luminance of an image, monitor, or scene. In this paper we use images from the high dynamic range survey by Fairchild (2007). These images are radiance maps that describe the relative luminance values in a scene and are obtained by combining images taken with multiple exposure durations, thus avoiding the problems of over- and underexposure associated with single-exposure photography. This allows the capture of scenes with an arbitrary dynamic range, thus overcoming one of the major disadvantages of existing databases of natural images (e.g., Geisler, Perry, Super, & Gallogly, 2001), which avoid high dynamic range scenes, such as an image taken directly into sunlight. The images also contain calibration data for absolute luminance values in the original scene that were obtained with a photometer. Using this information we can compute the dynamic range of each image, and in Figure 2a, we plot the normalized median luminance of each image against the dynamic range of each image (see Appendix 1 for details). The results demonstrate that the median luminance of an image is highly correlated with the dynamic range of an image (Pearson's $R = 0.92$, $p < 0.001$, $N = 91$). Thus as the dynamic range of an image increases, the image becomes increasingly dark, and in turn, an increasingly compressive nonlinearity is required for histogram equalization. This is illustrated in Figure 2 using an image with a low dynamic range (Figure 2b) and an image with a high dynamic range (Figure 2c). Interestingly, studies investigating the preferred system gamma using low dynamic range images demonstrate that subjects do indeed prefer a system gamma of greater than 1 (Bartleson & Breneman, 1967b; Hunt, 2005; Roufs & Goossens, 1988) consistent with the established imaging pipeline (Poynton, 2012), while a recent study demonstrates that subjects may prefer a system gamma of less than 1 for some high dynamic range images (Liu & Fairchild, 2007). Overall, this suggests that histogram equalization does play a role, but not as an absolute predictor of the preferred system gamma. Consistent with this, Singnoo and Finlayson (2010) showed that the preferred system gamma is correlated with the system gamma that achieves the greatest histogram equalization, but that a corrective factor is needed to obtain absolute estimates of system gamma. We shall return to this point in the later stages of this paper.

A number of authors have suggested that lightness perception plays a role in our choice of system gamma (Breneman, 1962; Fairchild, 1995; Jones, 1944). Lightness is the nonlinear perception of luminance values in the world or on a display. The exact nature of the function is sensitive to the experimental conditions (Bartleson & Breneman, 1967a; Radonjić, Allred, Gilchrist, & Brainard, 2011; Stevens & Stevens, 1963; Whittle, 1992; Wyszecki & Stiles, 1982). When natural scenes are used, the luminance-to-lightness function may be approximated by a simple power law (Bartleson, 1975; Clark, 1967). In this model, the exponent varies with the surround luminance of a monitor. The exponent required to model lightness perception is more compressive when the surround luminance of a monitor is dark than when the surround is light (Bartleson, 1975; Clark, 1967). This pattern of lightness perception is in the opposite direction of the preferred system gamma, leading a number of authors to propose a causal relationship (Bartleson & Breneman, 1967b; Jones, 1944). This hypothesis is supported by direct experimental evidence demonstrating that the preferred system gamma is lower for images presented with a bright surround than that obtained with a dark surround (Bartleson & Breneman, 1967b; Daniels, Giorgianni, & Fairchild, 1997; Liu & Fairchild, 2007; Roufs & Goossens, 1988).

Aim/hypothesis

The literature indicates that the preferred system gamma is both viewing-condition dependent and image dependent. The change in the preferred system gamma

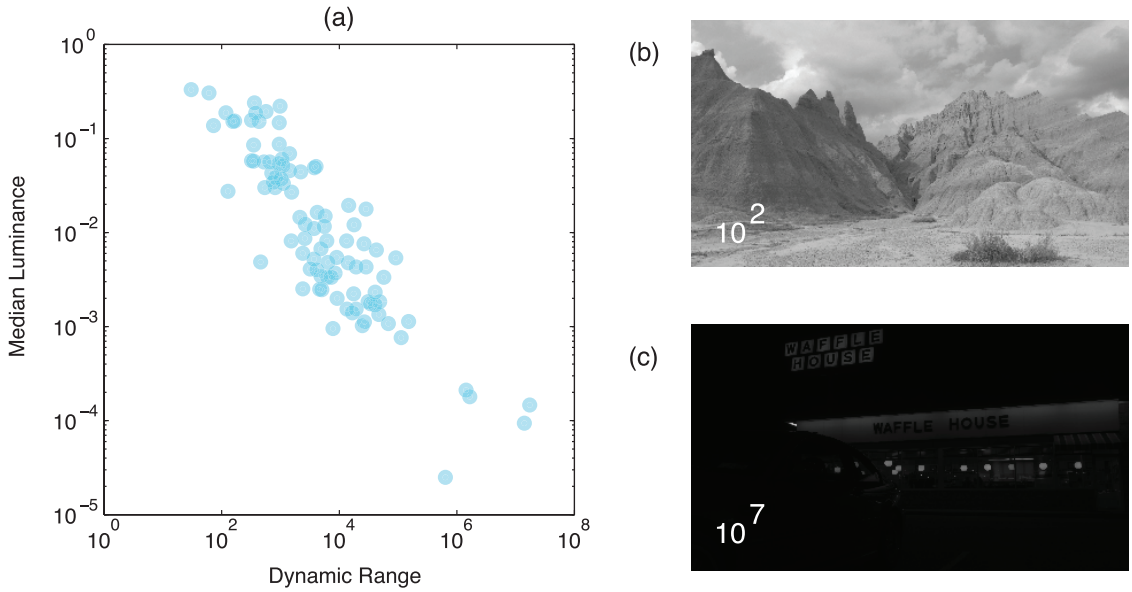

Figure 2. Image statistics from a database of high dynamic range images (Fairchild, 2007). (a) There is a negative correlation between the median luminance and the dynamic range of images in the database. Note, the median luminance is calculated after the luminance range has been normalized to between 0 and 1. On the right-hand side are two images from the high dynamic range survey by Fairchild (2007) presented with a system gamma of 1: (b) An image with a low dynamic range of two orders of magnitude and (c) an image with a high dynamic range of six orders of magnitude.

as a function of the viewing condition appears to be related to lightness perception (Breneman, 1962; Jones, 1944), whereas image dependent changes appear to be related to the need for histogram equalization (Singnoo & Finlayson, 2010). In this paper we investigate whether these two findings can be unified by the simple hypothesis that the preferred system gamma is the one that best equalizes the lightness distribution. More formally, we describe the model in Equations 1 through 4. First, the normalized (0-1) image N is raised to the power of system gamma (γ_{sys}) to produce the image that will be displayed on a given monitor.

$$\mathbf{I}_i = \mathbf{N}_i^{\gamma_{sys}} \quad (1)$$

Second, we estimate the perceived lightness L of the image by raising the image I to the power of γ_{psy}, which we shall term the psychological gamma. A value for γ_{psy} can either be obtained from the literature or from the experimental data, as described in later sections.

$$\mathbf{L}_i = \mathbf{I}_i^{\gamma_{psy}} \quad (2)$$

Third, the degree of histogram equalization (F) is computed as 1 minus the root-mean-square (RMS) difference between the normalized cumulative histogram of the lightness image (H) and the identity line.

$$\mathbf{F} = 1 - \frac{1}{(n+1)}\sqrt{\sum_{k=0}^{n}\left(\mathbf{H}\left(\frac{k}{n}\right) - \frac{k}{n}\right)^2} \quad (3)$$

Finally, for a given value of psychological gamma, we find the value of system gamma that maximizes F.

$$\gamma_{est} = \underset{\gamma_{sys}}{\operatorname{argmax}}\left(\mathbf{F}(\gamma_{sys})\right) \quad (4)$$

In this model the predicted value of system gamma will depend on two factors: the value of psychological gamma (γ_{psy}) and the luminance distribution of the image.

To investigate a subject's preference for system gamma (which we term the preferred system gamma), we select images from the aforementioned high dynamic range image survey by Fairchild (2007), which contain images with dynamic ranges from two to seven orders of magnitude, and thus images that will require different levels of system gamma to achieve histogram equalization (see Figure 2). To investigate viewing-condition dependency we investigate the preferred system gamma for images viewed on a black or white background and upon monitors; a cathode ray tube (CRT), a liquid crystal display (LCD), and an organic light-emitting diode (OLED) that have different dynamic ranges. We choose these monitors/conditions as both the background luminance condition (Moroney, 2002; Nundy & Purves, 2002; Stevens & Stevens, 1963; Whittle, 1992) and the dynamic range of the scene/monitor (Hoffman, Johnson, Kim, Vargas, & Banks, 2015; Nezamabadi, Miller, Daly, & Atkins, 2014; Radonjić et al., 2011) are known to affect the shape of the luminance-to-lightness function.

Methods

Observers

Eight observers took part in the experiments. Seven were naive to the experimental objectives and one was the author. In Experiments 1, 4, and 5, only six of the eight subjects were used as we were unable to obtain data on the LCD display for two subjects. All subjects had corrected or normal vision. All procedures complied with the Declaration of Helsinki and were approved by the Comité Ético de Investigación Clínica, Parc de Salut MAR, Barcelona, Spain.

Apparatus

Stimuli were generated using MATLAB (MathWorks) with functions from the PsychToolbox (Brainard, 1997; Pelli, 1997). Stimuli were either displayed on a Philips 109B CRT monitor, an OLED Sony Trimaster PVM, or a MacBook Pro LCD display. The CRT was run with spatial and temporal resolutions of 1280×960 pixels and 75 Hz. The display was viewed at a distance of 58 cm so that 36 pixels subtended 1 degree of visual angle. The full display subtended $35.5° \times 25.5°$. The Sony Trimaster PVM was run at a resolution of 1920×1080 pixels and 60 Hz. The display was viewed at 74 cm such that 36 pixels subtended 1 degree of visual field. The full display subtended $53° \times 30°$. The MacBook display has a resolution of 2560×1600 and was viewed at a distance of 46 cm such that 72 pixels spanned 1 degree of visual field. The full display subtended $36.5° \times 22°$. Monitor linearization was achieved by recording the relationship between the signal from the graphics card and the monitor luminance (measured using a Konica Minolta LS 100 photometer). An encoding nonlinearity that was the inverse of the decoding nonlinearity was applied to all images before 8-bit quantization and presentation. The CRT's minimum luminance was 1.0 cd/m^2 and the maximum was 112 cd/m^2. The LCD display of the MacBook Pro has a minimum luminance of 0.65 cd/m^2 and a maximum luminance of 383 cd/m^2. The Sony Trimaster PVM OLED-B monitor has a minimum luminance of below the minimum of our photometer at 0.001 to 100 cd/m^2 (see Cooper, Jiang, Vildavski, Farrell, & Norcia, 2013; Ito, Ogawa, & Sunaga, 2013 for a detailed examination of the Sony Trimaster PVM).

Procedure

On each trial subjects viewed a centrally presented image and were asked to rate the quality of the image on a sliding scale. The scale was marked *Terrible*, *Satisfactory*, and *Excellent* at the left, middle, and right extremes of the scale, respectively. These values correspond to 0, 5, and 9 on the presented figures. During each run, subjects were presented with a variety of images displayed with different values of system gamma, where system gamma describes the power-law relationship between the original, linearly captured and stored image and the final image as presented on the display (Equation 1). Values of system gamma varied between 2^{-4} to 2^2 at half log two intervals. On each run we ensured that subjects viewed both high and low quality images. Subjects were given a trial run before the full experiment to ensure they viewed the range of images available and could scale their answers accordingly. No subjects reported any difficulty rating the images. Each run consisted of a number of base images as described in the text. Each run only tested one background luminance condition to avoid cross adaption and every run was conducted in a dark room. The background luminance values were either black or white corresponding to 0% to 100% of the total monitor's luminance range. The original images had a high but variable spatial resolution of greater than 4000×3000 pixels. To reduce the spatial resolution we applied MATLAB's "imresize" command using the nearest setting to reduce images to a quarter of their original area. Images were then displayed without any additional resizing by the PsychToolbox's internal procedures. The justification for this procedure is included in the Appendix 2.

Experiment 1

Results

The results from Experiment 1 are shown in Figure 3. Each column shows the results for one base image. The first row shows the original image presented with a system gamma of 1 (assuming a decoding gamma of 2.2^{-1} in the reader's monitor or printer). The second row shows the images presented with the preferred system gamma chosen by the average subject. In the third row we plot image quality scores as a function of system gamma. The data is averaged across subjects with each subject contributing equally to the final score. All error bars in the manuscript are 95% error bars computed by bootstrapping with replacement. Blue dots denote data from the black background condition and green dots data from the white background condition. The results show that the image quality functions tend to be shifted toward the left in the light background condition, indicating that subjects prefer a lower, more compressive system gamma to be applied to the image when a white background is used. We estimate the preferred system gamma by fitting a fourth-order polynomial to the data

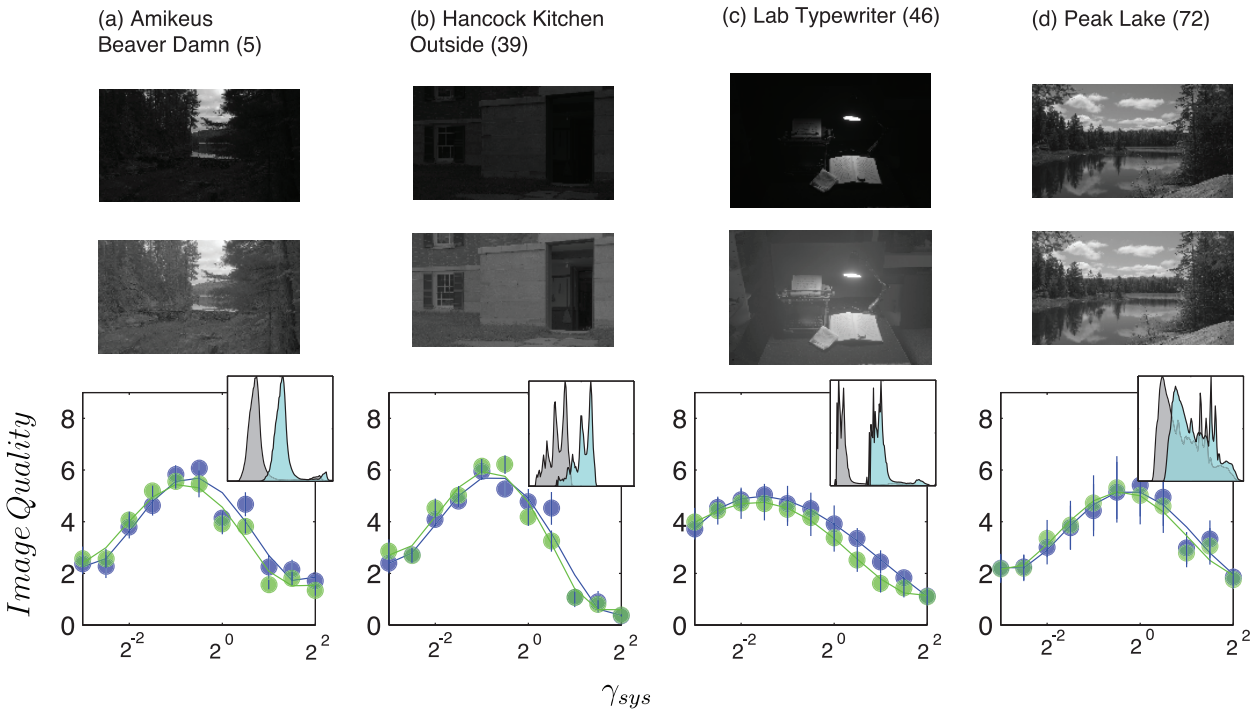

Figure 3. Top row: Images presented with a system gamma of 1. Middle row: Images presented with the preferred system gamma obtained in the white background condition. Bottom row: Image quality as a function of system gamma ($N = 6$). Blue data is for the black background condition; green data is for the white background condition. The insets show the histograms of lightness values for an image with a system gamma of 1 (gray) and the preferred system gamma inferred from the observer data (turquoise). For clarity, on the ordinate, we plot the square root of N (the number of pixels with a given luminance or lightness) and on the abscissa we plot lightness, determined using the value of psychological gamma shown in Figure 4b.

and identifying the value of system gamma that corresponds to the peak of the function (within the tested response range). The solid lines in Figure 3 illustrate the best-fitting polynomial. Consistent with the original Liu and Fairchild (2007) study, we find the preferred system gamma is higher when a dark background is used. Moreover, we find the ratio between the preferred system gammas in the black and white background conditions to be 1.2, close to the value of 1.14 obtained by Liu and Fairchild (2007) who used the same four images, but a different research paradigm to estimate this ratio.

Model: Equalization of the lightness histogram

The hypothesis detailed in Equations 1 through 4 states that the optimal value of system gamma is the one that best flattens the lightness histogram. A subprediction of this hypothesis is that image quality scores should vary monotonically with the degree of flatness in the lightness histogram (Equation 3). This property (if validated) can be used to estimate the parameter γ_{psy} that models lightness perception (Equation 2). To do so, we search (via brute force) for the value of γ_{psy} that maximizes the Pearson's correlation between flatness and image quality for all values of system gamma applied to a given image (i.e., each column of Figure 3). Encouragingly the search procedure obtains high correlation coefficients for all conditions (Figure 4a) indicating that image quality scores are linearly related to the degree of equalization in the lightness histogram. The resulting estimates of psychological gamma are plotted as a function of the preferred system gamma in Figure 4b. Circles denote data from the black background condition and squares data from the white background condition. When the background luminance level is increased from black to white, the preferred system gamma decreases and this change is matched by an equal and opposite change in the estimated psychological gamma. A discussion relating the values of psychological gamma obtained in this study to those in the literature is saved for the Discussion; however, it is worth noting that the estimated value of psychological gamma is image dependent. This is largely due to Images 46 and 72. Image 46 has a low median luminance, while Image 72 has a high median luminance. We formally test the theory that the estimated psychological gamma is positively correlated with the median luminance of the stimulus in the next experiment using a greater number of images.

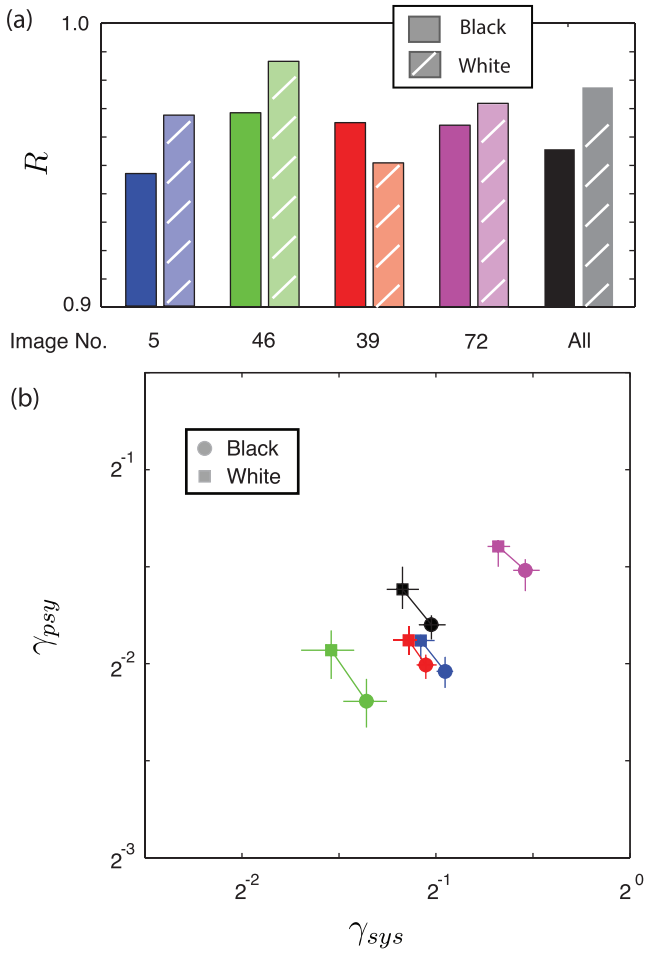

Figure 4. Model results from Experiment 1. (a) The Pearson's correlations of the image quality scores against the degree of flatness for the four images in the black and white background conditions using the values of psychological gamma described below. (b) Scatter plot for the preferred system gamma (corresponding to the peak of the best-fitting fourth-order polynomial) against psychological gamma (estimated by finding the value that correlates best with the degree of flatness; see Equations 1 through 4). Circles denote the black background condition and squares denote the white background condition. Both the preferred system gamma and the psychological gamma are image dependent. Changes in the preferred system gamma are matched by an equal and opposite change in the estimated psychological gamma (i.e., lie upon a line with a slope of minus one).

Experiment 2

Rationale

In this section, we explore the issue of image-dependent variability in the preference for system gamma. To do so, we extend the experimental paradigm to a greater number of images ($N = 26$), but only investigate system gamma scores in the black background condition. Initial testing used 16 randomly selected images, but an extra 10 were selected to ensure a better sampling of the dynamic ranges found in the Fairchild image database.

Psychophysical results

In this section, we plot a summary of the results across the full set of images tested ($N = 26$). For each image we compute the maximum image quality score and the corresponding preferred system gamma. In Figure 5 we plot the preferred system gamma against the median luminance of each image (each dot denotes a separate image). The results demonstrate that lower values of system gamma are required for images with a lower median luminance (Spearman's correlation; $R = 0.89$, $p < 0.0001$, $N = 26$). This finding is consistent with the notion that the preferred system gamma is, at least in part, determined by the need for histogram equalization. In Figure 5b we plot the maximum image quality score against the corresponding preferred system gamma. The results demonstrate that the lower the preferred system gamma, the worse the overall image appearance (Spearman's correlation; $R = 0.8$, $p < 0.001$, $N = 26$). It is not possible to tell whether the image quality scores remain flat or decrease for values of system gamma greater than 1; however, it is clear that image quality scores do not increase further when a system gamma of greater than 1 is required.

We note that the degree of histogram equalization that can be achieved via a simple power law is greater for low dynamic range images than for high dynamic range images (see Cyriac, Bertalmio, Kane, & Vazquez-Corral, 2015, for an explanation). The correlation between the degree of histogram equalization achieved using the preferred system gamma (Equation 3) is negative (log-log Pearson's $R = -0.82$, $p < 0.0001$, $N = 26$) and thus may account for why the image quality scores achieved with the preferred system gamma are lower for high dynamic range images.

Finally, in Figure 5c we test the hypothesis that the estimated psychological gamma is positively correlated with u, which is the median luminance of the image after it has been normalized to between 0 and 1. We fit Equation 5 to the data and find that the best-fitting line on the logarithms (base two) of the data has a gradient of $a = 0.0833$ and a constant of $b = -1.26$. The Pearson's correlation is significant at $R = 0.71$, $p < 0.001$, $N = 26$.

$$\log_2(\gamma_{\text{psy}}) = a\log_2(u) + b \quad (5)$$

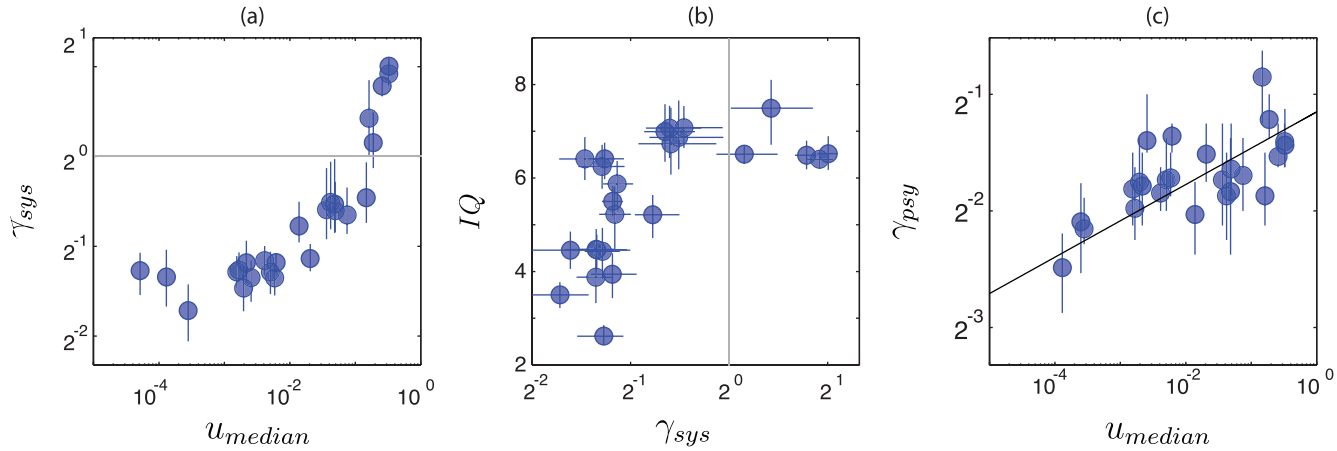

Figure 5. Results from Experiment 2. (a) Scatter plot for the estimated preferred system gamma as a function of the median luminance of an image. The median luminance was computed after the images had been scaled to between 0 and 1. (b) The maximum image quality score as a function of the preferred system gamma. (c) The estimated psychological gamma as a function of median luminance.

Results: Model comparisons

Singnoo and Finlayson (2010) investigated whether preference for system gamma could be explained by a preference for maximizing image entropy, which they noted is the equivalent to histogram equalization (Singnoo & Finlayson, 2010). The study asked subjects to select the preferred system gamma using a method of adjustment task for 46 images, including 15 low dynamic range images from the Kodak image database and 27 from their own database that were selected to have a broad range of mean luminance values and thus presumably dynamic ranges. Singnoo and Finlayson (2010) reported that the preferred system gamma (γ_{pref}) is linearly correlated with the gamma that achieves the maximum image entropy (γ_{eq}); however, to achieve absolute estimates of the preferred system gamma, a corrective factor is needed. The correction estimated by Singnoo and Finlayson was composed of a multiplicative factor $m = 0.32$ and an additive factor $c = 0.64$ (Equation 6).

$$\gamma_{\text{pref}} = m\gamma_{\text{eq}} + c \quad (6)$$

In Figure 6a we compare the performance of the two models on log-log (base 2) axes. The red dots indicate the predictions of the model of Singnoo and Finlayson after the correction (Equation 6) has been applied. The blue data are predictions of the model developed in this paper (Equations 1 through 4). For this computation we use a psychological gamma of 0.34, close to that of the value used in a number of color models, which is commonly 1/3 (Fairchild, 1995).

The axes of Figure 6a are equal, and perfect predictors should fall upon the identity line. To estimate the performance of the models we compute the RMS error (Figure 6c) and find the best-fitting straight line to the log of subjects' and the model's system gamma values (Figure 6b). The results show that the model of Singnoo and Finlayson exhibited a significantly greater RMS error and a pattern of estimates that is biased by both a multiplicative and a constant factor in log space.

Experiment 3: OLED versus CRT

Rationale

The CRT display used in the above experiment has a dynamic range of around 100. In contrast, modern OLED monitors achieve much darker blacks in the order of 0.0001 cd/m^2 (Ito et al., 2013) while both monitors have a maximal luminance of around 100 cd/m^2. Thus the difference in dynamic range is from 10^2 to 10^6. Sensitivity as a function of the background luminance has been found to be more compressive over the range of luminance values displayed on an OLED monitor than on a CRT display (Hoffman et al., 2015; Nezamabadi et al., 2014). Thus it is reasonable to expect that the preferred system gamma shall also be affected. Accordingly, we repeat Experiment 1 using a Sony Trimaster PVM OLED monitor. For clarity, we note that the displayed images always spanned the full dynamic range of the monitor in question.

Results

In Figure 7a we plot the preferred system gamma on the CRT display against the preferred system gamma on the OLED display. The results reveal that the preferred

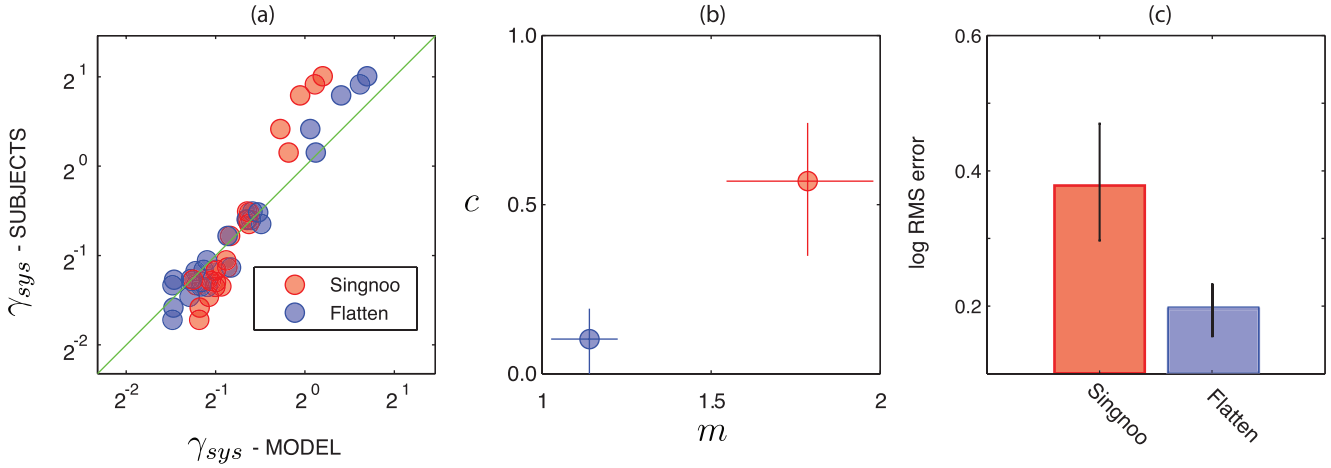

Figure 6. (a) Subjects' preferred system gamma as a function of the system gamma predicted by two models: The lightness equalization model developed in this paper and that of Singnoo and Finlayson (2010). The green line is the identity line and perfect estimators should fall upon this line. The best-fitting parameters *m* and *c* of a straight line fit to the data in (a). (c) The log RMS error of the two models tested.

system gamma is always higher when viewed on an OLED display. In turn, the estimated values of psychological gamma are lower on the OLED display than on the CRT (Figure 7b). In Figure 7c we test the hypothesis that the subjects are choosing a value of system gamma that matched the perceived distributions of lightness values on the CRT and OLED monitors (i.e., the value of system gamma that makes both images appear as similar as possible). To do so, we plot the preferred system gamma against the estimated values of psychological gammas on the CRT (yellow circles) and OLED (green squares) monitors. As in Experiment 1 (Figure 4b), we find that changes in the estimated preferred system gamma are matched by an equal and opposite change in the estimated psychological gamma (i.e., the changes fall upon a line with a slope of −1). Thus under the assumption that the estimated value of psychological gamma accurately reflects lightness perception, we can conclude the perceived distribution of lightness values will be the same on the two monitor types when the preferred system gamma is used.

Finally, in Figure 7d we demonstrate that the maximum image quality scores obtained on the OLED display are greater than those obtained on the CRT display; however, this effect appears to converge for high image quality scores. Thus the advantage of using an OLED display over a CRT display appears to be confined to images with a high dynamic range (which tend to receive low image quality scores), perhaps because these images contain more information in the dark regions of the image. It is possible that optical scatter may also play a role by reducing the effective dynamic range of the OLED display when bright scenes are displayed (see Stiehl, McCann, & Savoy, 1983).

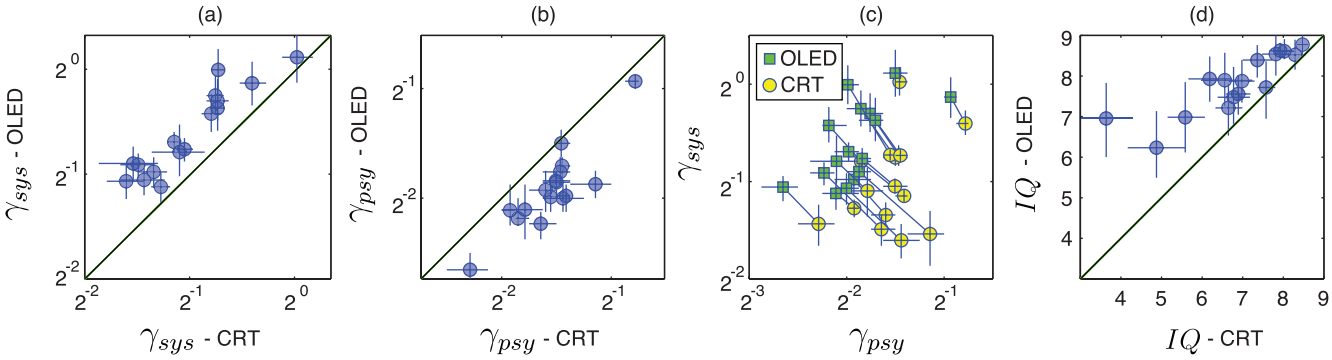

Figure 7. (a) The preferred system gamma for images displayed on the CRT and OLED monitors. (b) Psychological gamma estimated on the CRT and OLED monitors. (c) The preferred system gamma as a function of the estimated psychological gamma. Squares denote values estimated on the OLED monitor and circles values estimated on the CRT monitor. As in Figure 4, changes lie upon a line with a slope of −1. (d) The greatest image quality scores estimated on an OLED monitor against the greatest image quality scores estimated on a CRT monitor.

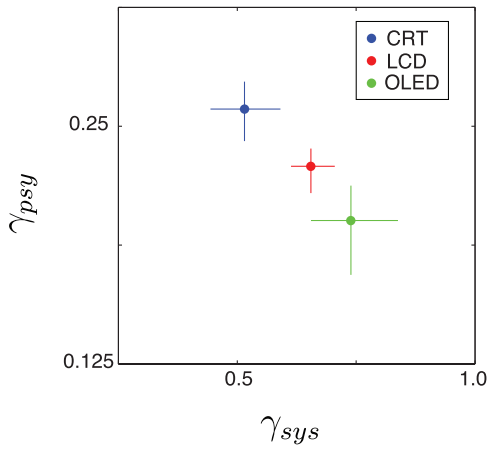

Figure 8. The estimated psychological gamma as a function of the estimated preferred system gamma on the CRT, LCD, and OLED monitors, but only for the black background condition.

Experiment 4: CRT, LCD, and OLED monitors

LCD monitors have a dynamic range between those of a CRT and an OLED monitor. Moreover, LCD monitors are in more widespread usage than either of the two monitors evaluated thus far. As such, we repeat Experiment 1 on an LCD monitor and repeat the analysis of Experiment 1 for the three monitor types. In Figure 8a we plot the estimated psychological gammas as a function of the estimated preferred system gamma for the three monitor types, but only in the black background condition. The results show that the estimated gamma value for the LCD fall in between that of the CRT and OLED monitors. This strengthens the argument that dynamic range plays a role in determining the lightness function and in turn the preferred system gamma.

Experiment 5: The impact of background luminance

The data for both the white and black background conditions and all three monitors are presented in Figure 9. As can be seen, the use of the white background lowers the preferred system gamma and raises the estimated psychological gamma for all three monitors. However, the impact of changing the background varies on the three monitor types. We note that a direct comparison is hard here for two reasons: First, the relative area of the background luminance is different between the different monitors as illustrated in Figure 9. Second, the ambient light conditions change when a lighter background luminance is used. This is

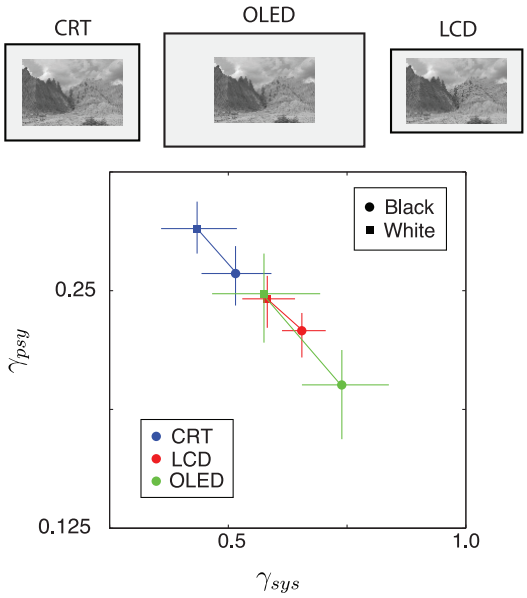

Figure 9. The estimated psychological gamma as a function of the estimated preferred system gamma on the CRT, LCD, and OLED monitors in the black and white background luminance conditions.

because the monitors are not perfect absorbers of light and thus ambient light may be reflected from the screen towards the observer's eye, reducing the effective dynamic range of the displayed image. To investigate this we recorded the light present at the observers' location under the experimental conditions. To do so, we repeat the conditions of Experiment 5 but present a uniform black or white luminance values in place of the real-world images. We then record the light emitted from the uniform test images in both the black and white background luminance conditions. In the case of the CRT and LCD monitors, the background luminance condition does not affect the minimum recorded luminance; however, for the OLED monitor the minimum luminance increases from 0.001 with a black background to 0.2 cd/m^2 when a white background is used. Thus in our experimental setup, the dynamic range of the OLED monitor is similar to that of the CRT when a white surround is used. We note that the impact of the white background luminance is similar to that of turning the laboratory lights on, which leads to a minimum luminance of between 0.2 and 0.5 cd/m^2, depending on what surface (white or black) is immediately behind the monitor.

Discussion

In this paper we set out to develop an explanation for how the preferred system gamma is affected both by the displayed image and the viewing conditions. We

present a simple model to estimate the preferred system gamma (Equations 1 through 4) and we find that this model makes different predictions to an existing model for estimating the preferred system gamma (Singnoo & Finlayson, 2010). We have published the data on FigShare.com (Kane & Bertalmío, 2016).

The model's success is based on the finding that subjective image quality scores are linearly correlated with the degree of flatness in the perceived lightness of a stimulus (Equations 1 through 4). In this work we model lightness perception as a simple power law. We manipulate two factors (dynamic range and background luminance) that are known to affect lightness perception and find that our simple model can account for the data via an adaption of the exponent.

Further work and lightness perception

In this study, a model of lightness perception is inferred from the data, but not directly measured. As such there is ambiguity as to what the values of psychological gamma estimated in this study represents. Encouragingly, the average value of psychological gamma obtained in the black background condition on the CRT (when averaged across images) is 0.31, which is close to the value of 1/3 common to many color models (Fairchild, 1995), but somewhat below those obtained by Stevens and Stevens (1963, table III), who estimated values that span from 0.33 (for the dark adapted eye), to 0.49 (adaption to 10,000-ft lamberts). Nonetheless the work would be advanced further by a direct estimation of the lightness in the viewed image, perhaps via a magnitude estimation task (e.g., Bartleson & Breneman, 1967a). For instance, the value of psychological gamma obtained is correlated with the normalized median luminance of the tested image (Figure 5c). This may reflect an adaptation to the luminance distribution of the stimulus (Radonjić et al., 2011) and/or the adapting luminance level (Stevens & Stevens, 1963), or alternatively some unknown aesthetic criteria, unrelated to any perceptual adaptation. This ambiguity can only be resolved by a direct estimation of perceived lightness within the image itself.

Model extrapolation: Viewing conditions and lightness perception

The current model is optimized only for the monitors, image set, and background-surround luminance combinations used in this paper. While the impact of the background or surround luminance of a display on the preferred system gamma (Experiment 1) has been investigated previously (Bartleson, 1975; Bartleson & Breneman, 1967b; Daniels et al., 1997; Hunt, 2005; Liu & Fairchild, 2007), the impact of a monitor's dynamic range has not (Experiments 3, 4, and 5), nor has the interaction between the background luminance and the dynamic range of a monitor (Experiment 5). The importance of dynamic range has not been directly investigated previously; however, it may have affected the estimates produced by other studies. For instance, the work by Bartleson and Breneman (1967b) computed the compensation ratio for photographic prints (viewed with a light background) against photographic projections (viewed with a dark background). The dynamic range of prints is less than two orders of magnitude, while the dynamic range of projections is at least two orders of magnitude (Poynton, 2012; Stiehl et al., 1983). Our data would predict that the preferred system gamma would be lower for lower dynamic range displays, and thus the high compensation ratio of 1.4 obtained by Bartleson and Breneman (1967b) relative to those found by this study (1.21) and others (1.14 and 1.16, respectively, in Daniels et al., 1997 and Liu & Fairchild, 2007) could potentially be accounted for by the different dynamic range of the two media.

Extending the model to an arbitrary display configuration will require a model of lightness perception that can both generalize to the variety of display configurations, image sets, and viewing conditions that exist. This is problematic for two reasons: First, as noted in Experiment 5, the luminance values sent to a monitor do not always correspond in a simple way to the light received at the eye, due to a monitor's reflectance and optical scatter (Stiehl et al., 1983). Second, lightness perception has been investigated using both simple and real-world stimuli, but unfortunately there is no clear correspondence between the two sets of results (see Bartleson, 1975, for a detailed discussion on this topic). In laboratory studies small test stimuli are typically viewed upon a uniform background and the relationship between the background luminance and lightness perception is investigated. In these studies the surround luminance of a monitor is fixed and is typically dark. Thus when the background luminance is altered, there is a large change in the mean luminance of the central visual field. In contrast, when real-world scenes are used, the surround luminance of a display is adjusted and this manipulation primarily affects the peripheral vision field. Lightness perception estimated using real-world scenes tends to show a relatively small modulation when the surround luminance changes, compared to that observed using simple stimuli, where the function may change more dramatically (Bartleson, 1975). Importantly, the ratio of exponents estimated using natural scenes predict the ratio of preferred system gamma in the black and white background conditions, while the exponents derived using simple stimuli do not correspond well (Bartleson, 1975). The

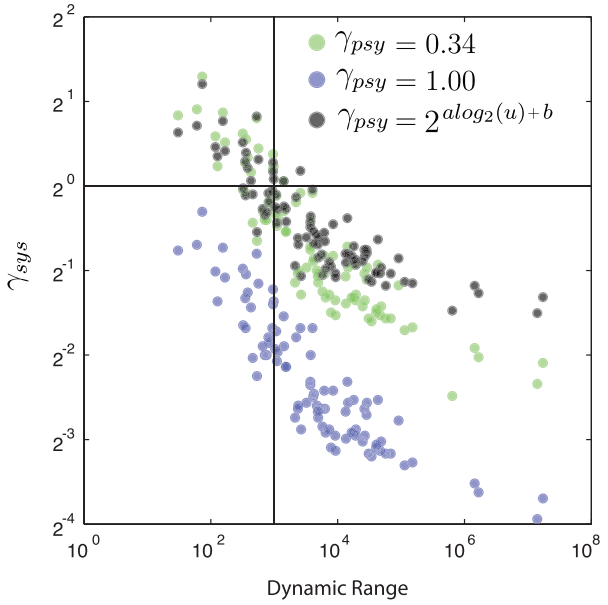

Figure 10. Application of the model developed in this paper (Equations 1 through 4) to the high dynamic range image database of Fairchild (2007) for different levels of psychological gamma. The black data is for a model of psychological gamma that adapts with the normalized median luminance of the stimulus (Equation 5).

authors suspect that the discrepancy between studies using real and artificial stimuli will only be resolved by using a more holistic, less arbitrary description of the stimulus, perhaps in terms of the global luminance distribution across the full visual field. This conclusion is consistent with that proposed by Bartleson and Breneman (1967b) who coined the term "inducting field," referring to both the stimulus and surround. This has implications for comparisons between studies using real-world scenes too, as the relative ratio of between the surround area and the image area will vary between studies (e.g., cinema vs. mobile phone). Other potential factors affecting lightness perception include the absolute luminance level of the stimulus (Bartleson, 1975; Stevens & Stevens, 1963), the possible impact of crispening (Whittle, 1992), and more generally, that lightness perception is only approximated by a power law (Clark, 1967).

Model extrapolation: Image dependency

The main goal of this study was to develop a model that could account for image-dependent variations in the preferred system gamma. In Figure 10 we examine the prediction of the model developed in this paper, when applied to all images in the Fairchild (2007) database. To do so, we plot the dynamic range of an image as a function of the system gamma required to flatten either the intensity distribution of an image (blue) or the lightness distribution estimated using an exponent of 0.34 (green) or using a variable exponent (black) that adapts to the normalized median luminance of the stimulus (Equation 5). The results show that flattening the intensity distribution never predicts a gamma of greater than 1. This is because all the real-world scenes tested have a mean luminance of less than half. However, with the fixed and variable exponent we find that images with a dynamic range of less than three orders of magnitude require an exponent of greater than 1, a finding in keeping with the established imaging pipeline as discussed in the Introduction (Poynton, 2012), which is designed to work with low dynamic range images that are the mainstay of consumer and commercial content.

The model estimates depicted in Figure 10 can be used to estimate the preferred system gamma for an individual image, or a set of images, given an accurate value of psychological gamma and the median luminance of the image or images. This analysis may prove useful particularly as high dynamic range capture-and-display technologies become more widespread and the standard display setting may need to be reevaluated.

Keywords: system gamma, image quality, natural scenes, tone-mapping operators, high dynamic range, low dynamic range

Acknowledgments

This work was supported by the European Research Council, Starting (Grant ref. 306337), the Spanish government (grant ref. TIN2015-71537-P), and by the Icrea Academia Award.

Commercial relationships: None.
Corresponding author: David Kane.
Email: kanepsychophysics@gmail.com.
Address: Department of Information and Communication Technologies, Universitat Pompeu Fabra, Roc Boronat, Barcelona, Spain.

References

Bartleson, C. J. (1975). Optimum image tone reproduction. *Journal of the SMPTE*, *84*(8), 613–618.

Bartleson, C. J., & Breneman, E. J. (1967a). Brightness perception in complex fields. *Journal of the Optical Society of America*, *57*(7), 953–956.

Bartleson, C. J., & Breneman, E. J. (1967b). Brightness

reproduction in the photographic process. *Photographic Science and Engineering*, *11*(4), 254–262.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436, doi:10.1163/156856897X00357.

Breneman, E. J. (1962). The effect of level of illuminance and relative surround luminance on the appearance of black-and-white photographs. *Photographic Science and Engineering*, *6*(3), 172–179.

Clark, L. D. (1967). Mathematical prediction of photographic picture quality from tone-reproduction data. *Photographic Science and Engineering*, *11*(5), 306–315.

Cooper, E. A., Jiang, H., Vildavski, V., Farrell, J. E., & Norcia, A. M. (2013). Assessment of OLED displays for vision research. *Journal of Vision*, *13*(12):16, 1–13, doi:10.1167/13.12.16. [PubMed] [Article]

Cyriac, P., Bertalmio, M., Kane, D., & Vazquez-Corral, J. (2015). A tone mapping operator based on neural and psychophysical models of visual perception. In B. E. Rogowitz, T. N. Pappas, & H. de Ridder (Eds.), *IS&T/SPIE Electronic Imaging*. Bellingham, WA: International Society for Optics and Photonics.

Daniels, C. M., Giorgianni, E. J., & Fairchild, M. D. (1997). The effect of surround on perceived lightness contrast of pictorial images. In M. R. V. Sahyun (Ed.), *Color and Imaging Conference* Vol. 1997 (pp. 12–16). Scottsdale, AZ: Society for Imaging Science and Technology.

Fairchild, M. D. (1995). Considering the surround in device-independent color imaging. *Color Research & Application*, *20*(6), 352–363.

Fairchild, M. D. (2007). The HDR photographic survey. In J. Larimer & M. Moroney (Eds.), *15th Color and Imaging Conference* (pp. 233–238). Albuquerque, NM: Society for Imaging Science and Technology.

Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, *41*(6), 711–724.

Gonzalez, R. C., & Woods, R. E. (2002). *Digital image processing* (pp. 299–300). New York: Pearson Higher Education.

Hoffman, D. M., Johnson, P. V., Kim, J. S., Vargas, A. D., & Banks, M. S. (2015). 240 Hz OLED technology properties that can enable improved image quality. *Journal of the Society for Information Display*, *22*(7), 346–356.

Hunt, R. W. G. (1987). *The reproduction of colour in photography, printing and television*. Tolworth, England: Fountain Press.

Hunt, R. W. G. (2005). *The reproduction of colour* (6th ed.). Oxford, UK: Fountain Press.

Ito, H., Ogawa, M., & Sunaga, S. (2013). Evaluation of an organic light-emitting diode display for precise visual stimulation. *Journal of Vision*, *13*(7):6, 1–21, doi:10.1167/13.7.6. [PubMed] [Article]

Jones, L. A. (1944). Psychophysics and photography. *Journal of Optical Society of America*, *34*(2), 66–85.

Kane, D., & Bertalmío, M. (2016). Database: Image quality as a function of system gamma and dynamic range, doi:10.6084/m9.figshare.2071048. Available at https://figshare.com/s/1bf3329b6cda69662cbd

Liu, C., & Fairchild, M. D. (2007). Re-measuring and modeling perceived image contrast under different levels of surround illumination. In *15th Color and Imaging Conference* (Vol. 2007, pp. 66–70). Albuquerque, NM: Society for Imaging Science and Technology.

Moroney, N. (2002). Background and the perception of lightness. In R. Chung & A. Rodrigues (Eds.), *9th congress of the international colour association*, Vol. 4421 (pp. 571–574). Rochester, NY: Society for Imaging Science and Technology.

Nezamabadi, M., Miller, S., Daly, S., & Atkins, R. (2014). Color signal encoding for high dynamic range and wide color gamut based on human perception. In *Proceedings of SPIE 9015, color imaging XIX: Displaying, processing, hardcopy, and applications, 9015C*, doi:10.1117/12.2042893.

Nundy, S., & Purves, D. (2002). A probabilistic explanation of brightness scaling. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(22), 14482–14487, doi:10.1073/pnas.172520399.

Olshausen, B. A., & Field, D. J. (2000). Vision and the coding of natural images. *American Scientist*, *88*, 238–245, doi:10.1511/2000.3.238.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Poynton, C. A. (2012). *Digital video and HD: Algorithms and interfaces*. Waltham, MA: Morgan Kaufmann, Elsevier.

Radonjić, A., Allred, S. R., Gilchrist, A. L., & Brainard, D. H. (2011). The dynamic range of human lightness perception. *Current Biology*, *21*(22), 1931–1936.

Roufs, J. A. J., & Goossens, I. M. J. (1988). The effect of gamma on perceived image quality. In *Confer-*

ence Record of the 1988 International Display Research Conference (pp. 27–31). San Diego, CA: IEEE.

Singnoo, J., & Finlayson, G. D. (2010). Understanding the gamma adjustment of images. In J. A. Ferwerda & G. J. Woolfe (Eds.), *18th color and imaging conference*, Vol. 2010 (pp. 134–139). San Antonio, TX: Society for Imaging Science and Technology.

Stevens, J. C., & Stevens, S. S. (1963). Brightness function: Effects of adaptation. *Journal of the Optical Society of America*, *53*(3), 375–385.

Stiehl, W. A., McCann, J. J., & Savoy, R. L. (1983). Influence of intraocular scattered light on lightness-scaling experiments. *Journal of the Optical Society of America*, *73*(9), 1143–1148.

van de Poel, F. H. J., & Valeton, J. J. P. (1954). The flying spot scanner. *Philips Technical Review*, *15*, 221–232.

Whittle, P. (1992). Brightness, discriminability and the "crispening effect." *Vision Research*, *32*(8), 1493–1507.

Wyszecki, G., & Stiles, W. S. (1982). *Color science* (Vol. 8). New York: Wiley.

Appendix 1

Image statistics

The statistics in Figure 2 are computed from the high dynamic range survey by Fairchild (2007). The database consists of 105 high dynamic range images produced by a multi-exposure and fusion technique. Each image contains a radiance map, which is an estimate of the relative luminance values in a scene. A photometer was used to measure the absolute luminance values within a 1° portion of the original scene (for 91 of the original images). These values were used to create a multiplicative factor that could be used to obtain an estimate of the absolute luminance values in a scene. The dynamic range of each image was computed from the 0.1% and 99% percentiles to avoid the influence of noise that could lead to negative values. The median luminance was computed after the image had been normalized to between 0 and 1. A Pearson's correlation estimates a strong negative correlation between the absolute luminance in a scene and the median luminance (Pearson's $R = 0.92$, $p < 0.001$, $N = 91$).

Appendix 2

Resampling methodology

In this paper we wanted to explore the preference for system gamma using images with luminance distributions found in the real world. To present each image on the various monitors we had to downsample the spatial resolution. We observe in Figure 11 that the resampling methodology bicubic interpolation can substantially alter the natural image statistics. In Figure 11a we plot the original median luminance against the median luminance found after rescaling the image to a 16th of the original area. We find that bicubic interpolation substantially raises the median luminance, particularly for higher dynamic range scenes, and in turn reduces the system gamma required for equalization of the lightness histogram (Figure 11b). This effect is not observed when nearest neighbor resampling is used.

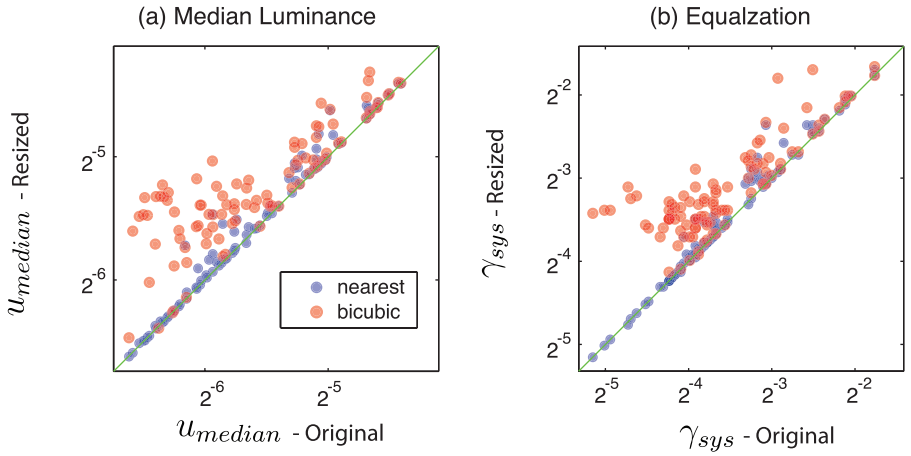

Figure 11. The effect of bicubic and nearest neighbor interpolation methods on the statistics of real-world images from the high dynamic range survey by Fairchild (2007).