

Is There A Preference For Linearity When Viewing Natural Images?

David Kane and Marcelo Bertalmío.

Departament de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra,
Barcelona, Spain

ABSTRACT

The system gamma of the imaging pipeline, defined as the product of the encoding and decoding gammas, is typically greater than one and is stronger for images viewed with a dark background (e.g. cinema) than those viewed in lighter conditions (e.g. office displays).¹⁻³ However, for high dynamic range (HDR) images reproduced on a low dynamic range (LDR) monitor, subjects often prefer a system gamma of less than one,⁴ presumably reflecting the greater need for histogram equalization in HDR images. In this study we ask subjects to rate the perceived quality of images presented on a LDR monitor using various levels of system gamma. We reveal that the optimal system gamma is below one for images with a HDR and approaches or exceeds one for images with a LDR. Additionally, the highest quality scores occur for images where a system gamma of one is optimal, suggesting a preference for linearity (where possible). We find that subjective image quality scores can be predicted by computing the degree of histogram equalization of the lightness distribution. Accordingly, an optimal, image dependent system gamma can be computed that maximizes perceived image quality.

Keywords: System gamma, High Dynamic Range, Low Dynamic Range, Image Quality, Natural Images, Lightness Perception

1. INTRODUCTION

The encoding gamma of most image formats and the decoding gamma of displays can be approximated by a gamma function. The encoding gamma is a compressive non-linearity ($\gamma_{enc} < 1$) and the decoding gamma an expansive non-linearity ($\gamma_{dec} > 1$). System gamma is the final gamma value after the signal has passed through both the encoding and decoding gammas. Early display engineers assumed that the optimal system gamma was linear ($\gamma_{sys} = 1$),^{5,6} however in practice, display engineers set the system gamma to a value greater than one. The precise value appears to depend on the viewing conditions: In the early 60's it was observed that photographic prints, typically viewed with a light background used a system gamma of 1.1,² whilst transparencies (e.g. slides or movies), typically viewed with a dark surround, used a system gamma of 1.6. These numbers correspond well to the system gammas used in contemporary display pipelines; personal and office displays, again typically viewed with a light background, use a system gamma of 1.1, whilst modern cinemas use a system gamma of 1.5.¹ Bartleson and Brenman³ argue that people prefer to view a perceptually linear version of the original image, but that a non-linear system gamma ($\gamma \neq 1$) is needed to compensate for a mismatch between the viewing conditions of the original scene and the displayed image. Due to the strong correlations across space that exist in natural images, the area outside a camera's field of view is likely to have a similar mean luminance to the regions within the camera's field of view.⁷ In contrast the area surrounding an image often has a different luminance, particularly for self-illuminating displays. For instance, cinema screens are nearly always viewed with a surround that is darker than the display itself. This is an issue because the surround luminance condition is known to affect the relationship between onscreen luminance and perceived lightness; broadly speaking, when the surround luminance is dark, perceived lightness is a compressive function of onscreen luminance, but when a lighter surround is used the function becomes closer to linear, or even expansive.⁸⁻¹⁰ It was the observation that the non-linear perception of light is approximately the inverse of the system gammas used by display engineers that led Bartleson and Brenman³ to hypothesize that consumers prefer to view a perceptually linear version of the original image. This hypothesis has received support from psychophysical experiments that demonstrate that preferred system gamma does indeed increase with background luminance and vice-versa.^{3,4,11,12}

Further author information: (Send correspondence to David Kane)

David Kane: E-mail: kanepsychophysics@gmail.com

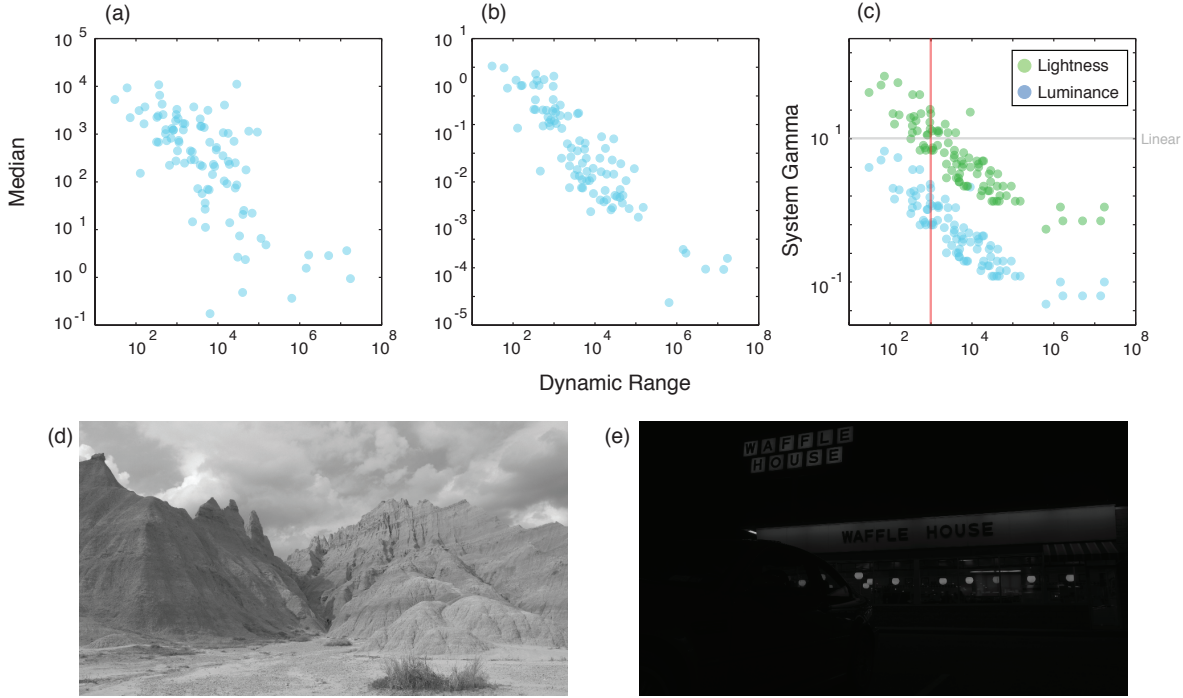


Figure 1. Statistics from the high dynamic range survey by Mark Fairchild.¹³ The dynamic range of images plotted against (a) the median absolute luminance of images and (b) the relative median luminance of images after they have been scaled to between 0 and 1. (c) plots the dynamic range of images against the value of system gamma that achieve the greatest histogram equalization of either the luminance distribution (blue dots) or the perceived lightness distribution (green dots). Note that when lightness perception is taken into account, the optimal system gamma may be greater than one for images with a dynamic range of less than three orders of magnitude. Two images are presented with a system gamma of one; (d) has dynamic range of 30.4 and (e) has a dynamic range of 17,313,520.

The motivation for this study comes from the twin observations that (a) the preferred system gamma is highly image dependent^{14,15} and (b) the recent finding that for high dynamic range images, the preferred system gamma may be less than one (when presented on a low dynamic range monitor),⁴ where dynamic range (DR) refers to the ratio between the maximum and minimum luminance values in a scene or image. The DR of natural scenes may span up to eight orders of magnitude, however the majority of imagery over the last century has been captured using low DR technology which can only capture between two and three orders of magnitude. This is also true for scientific databases of natural images.^{16,17} To emphasize the bias this may cause in Figure 1 we plot the DR of images from the high dynamic range survey of Mark Fairchild¹³ against the median luminance of the either (a) the absolute luminance values in the original scene or (b) the image after having been scaled to between zero and one. Figure 1 demonstrates a clear negative correlation between DR and the median luminance, particularly after normalization (Pearson's correlation $R = 0.9, p < 0.001$). Given that most commercially available cameras can only capture between two and three orders of magnitude, the majority of properly exposed images must exist within the left hand side of Figure 1. As a results neither scientific, personal, nor commercial images will encapsulate the full range of median luminance values that exist in the natural world. Recently, multi-exposure and fusion techniques have become more widespread. Such techniques can capture the full range of naturally occurring luminance values by combing multiple images of the same scene, captured using different exposure durations. With the rise of high DR capture techniques, a different problem has arisen. Consumer monitors can only display a DR of between two and three orders of magnitude, meaning a veridical representation of the relative luminance values of a high DR scene is impossible. When images with a high DR are presented with a system gamma of one the images are often dominated by dark, low contrast regions (Figure 1e), whilst images with a

low DR have excellent reproduction quality (Figure 1d). To solve this problem various tone-mapping operators (TMOs) have been developed. The goal of most TMOs is to recover the detail present in the original scene, whilst preserving the 'natural feel' of the scene.¹⁸ For instance many TMOs apply a compressive non-linearity to the image.¹⁹ This procedure achieves some degree of histogram equalization by stretching the representation of the relatively common, low luminance values and compressing the representation of the relatively sparse, high luminance values. System gamma can also perform an analogous histogram equalization when an appropriate value is selected. Accordingly, it may be hypothesized that subjects image dependent preference for system gamma is determined by the relative need for histogram equalization in an image.^{4, 14, 15}

1.1 Aim

The aim of this study is to investigate to what extent histogram equalization plays a part in subjects choice of preferred gamma. The importance of histogram equalization is likely to have been underestimated in previous studies because of the low DR images used.^{15, 20} The high dynamic range survey by Mark Fairchild¹³ makes for an excellent database of images to study image dependent preference in system gamma because it includes images with a broad range of DRs and thus luminance distributions (see Figure 1). In this paper we extend the research of Liu and Fairchild⁴ who investigated preference for system gamma using four images from the high dynamic range survey¹³ and a wide variety of background luminance conditions. The aim of the original study, was to examine the influence of background luminance on preferred gamma and the issue of dynamic range was not a primarily topic of investigation. However, the observation was made that subjects preference for system gamma was highly image dependent. In the original study subjects were asked to make pairwise comparative judgements of two simultaneously presented images and overall preference was inferred using Thurstone's law of comparative judgements.²¹ Although pairwise judgements are considered to allow precise comparisons between images, they suffer from a combinatorial explosion when too many images are used. In order to extend the study to a greater number of images, and to allow arbitrary comparisons between conditions, we ask subjects to make absolute image quality judgements. An analysis of subjects between and within trial-retrial reliability when making absolute image quality judgements is included in the results section. To allow a comparison between the two experimental paradigms, experiment one replicates a subset of the conditions used in Liu and Fairchild.⁴ In experiment two we extend the procedure to a greater number of images.

2. METHODS

2.1 Subjects

All subjects had normal to corrected vision. Subjects were informed they were free to leave the experiment at any time. Subjects were all members of the lab group and were not paid. Experiment one had 10 subjects. Experiment two had 8 subjects.

2.2 Apparatus

Stimuli were generated on an Apple MacBook running MATLAB (MathWorks) with functions from the Psychtoolbox.^{22, 23} The experiments took place in a purpose built laboratory in a darkened room. Subjects viewed stimuli on a Philips 109B CRT monitor running with a resolution of 1280 by 960 at 75 Hz. Images were viewed at a distance of 58cm so that 36 pixels subtended 1 degrees of visual field. The full display subtended 35.5 by 25.5 degrees. Monitor linearization was achieved by recording the relationship between the signal from the graphics card and the monitor luminance (using a Konica Minolta LS-100 photometer) and linearization took place used the internal lookup table of the graphics card. We used bit-stealing to increase the effective bit rate of the monitor from 8-bit per channel to a pseudo 10-bit grayscale channel.²⁴ The minimum luminance of the monitor is 0.6cdm^{-2} and the maximum 112cdm^{-2} .

2.3 Procedure

All procedures were approved by by the Informe del Comitè Ètico de Investigaci3n Clìnica and conformed with the Declaration of Helsinki. Subjects were asked to rate each image upon a sliding scale presented on the bottom of the display. The scale horizontally subtended 20 degrees and had three markers at the beginning, middle and end. Written from left to right and below the three markers were 'Terrible', 'Acceptable' and 'Excellent'.

Subjects were given no verbal or written instructions other than to rate each image along the given scale. Subjects interacted with the scale via a mouse. The position of the score marker was randomized on each trial. Subjects selected the marker by placing the cursor over the score marker and pressing (but not holding) the left mouse key. This action selected the score marker turned it from white to red. The position of the score marker could then be altered by moving the mouse left or right. The score marker was not allowed to go outside of the boundaries of the scale. Upon pressing the left mouse button for a second time the marker's position became fixed. The subjects were free to adjust the marker as many times as necessary. When the subject was satisfied with the given score, pressing the spacebar on the keyboard would initiate the next trial. To allow comparison with ongoing research in our laboratory the results are reported on a nine-point scale, where 0 corresponds to 'terrible', 5 to 'acceptable' and 9 to 'excellent'.

2.4 Stimulus

Images were .EXR files from the high dynamic range survey which can be obtained online. The full details of the image registration can be found in,¹³ but we shall review the main points here. Each image was constructed from between 8 to 18 (with a median of 9) exposures of the same scene separated by a one-stop. Images were fused into high dynamic range images using Adobe Photoshop's Merge-to-HDR function to generate a radiance map containing the relative luminance values in a scene. Absolute luminance information was obtained by recording from the scene using a photometer and computing an appropriate scaling factor.

Before presentation on our CRT monitor each image was rescaled to a quarter or an eighth of the original size using the Matlab function `imresize`. The base images provided online are not of equal size due to clipping for 'aesthetic' purposes.¹³ For the large condition the average image subtended 29.5 by 18.2 degrees, in the small condition the average image subtended 14.8 by 9.1 degrees. The `nearest` methodology was used to perform resizing. This methodology resizes via a simple subsampling the stimulus at the required resolution and was found to have the least impact on the shape of the luminance distribution: Natural scenes tend to be low key and the mode of the luminance histogram (when linear intensity scale is used) is frequently between the values of zero and one. Sampling methodologies that operate via interpolation shift the peak towards higher values. This is a particular problem for tone-mapping operators that are designed with natural image statistics in mind.

3. RESULTS

3.1 Results: Within and Between Subject Consistency

For all experiments reported in this paper subjects completed each trial twice. This allow us to compute the trial-retrial correlation and asses the reliability of our subjects. To do so, we combine the data across experiments. We find that the trail-retrial correlations of our subjects is strong (Pearson's correlation $R = 0.95, p < 0.0001$), this corresponds to a mean absolute difference between trials of 0.7 on a 9 point scale. This indicates that subjects make consistent and reliable absolute image quality judgments, consistent with the findings of other studies.^{15,20} The correlation between subjects is also strong (Pearson's correlation $R = 0.9, P < 0.0001$), although the mean absolute difference between subjects (when viewing the same stimulus) is greater at 1.5 on a 9 point scale. Thus subject's relative judgements of image quality demonstrate inter-subject constancy, although subjects absolute judgements may be scaled differently. In all the data reported in this study we report the image quality scores averaged across subjects.

3.2 Results: Experiment One

The aim of experiment one was to replicate the results of Liu and Fairchild⁴ using an absolute rating scale rather than pairwise comparative judgements. The results for the three images (Amikeus Beaver Dam 1, Hancock Kitchen Outside and Peak Lake) are shown in Figure 2 alongside a tone-mapped version of the image with the approximately correct system gamma applied (note the exact system value depends on the printer or monitor being used to view the stimulus). The fourth image, Lab Typewriter was not used in this study. The aim of the Liu and Fairchild⁴ paper was to investigate the impact of background luminance on the preferred system gamma. The experimental setup²⁵ consisted of an LCD monitor sat within a room with ceiling to floor lighting that could be manipulated independently of the display luminance. Images were displayed across the full field of the displays. We do not have the luxury of such a configuration. In our experiment, the image was displayed

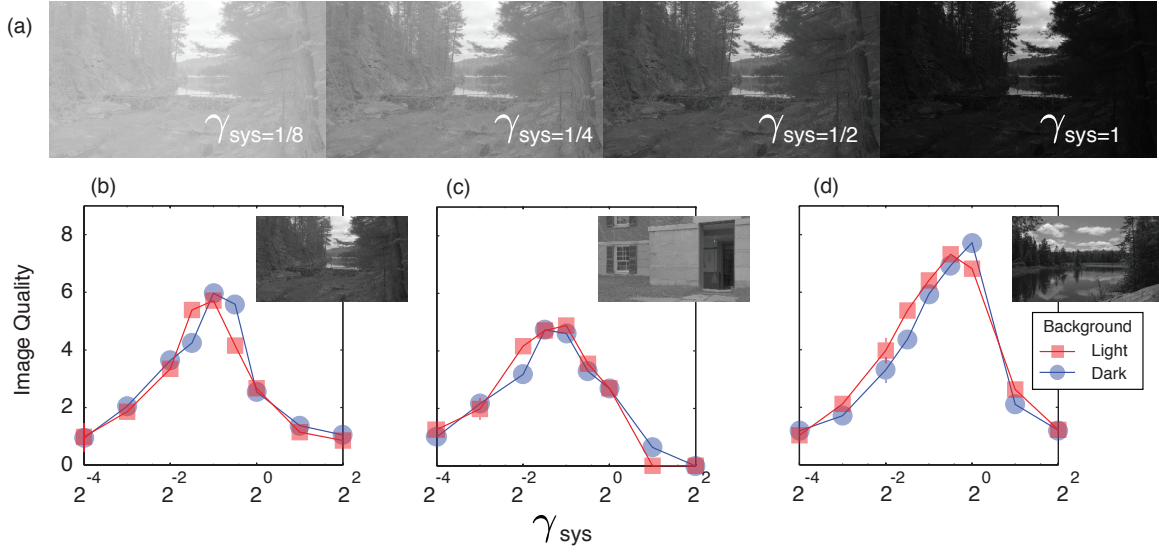


Figure 2. Results for experiment one, aimed at reproducing the results of Liu and Fairchild.⁴ The top row illustrates the impact of various system gamma levels upon the image Amikeus Beaver Dam 1. The bottom row shows the average image quality scores as a function system gamma, under two background luminance conditions; dark and light. For images (a) and (b) preferred gamma is below one, for image (c) preferred gamma is one. For all images the preferred system gamma is less when light background is used.

over a subset of the full monitor and two surround conditions were used: black and white corresponding to the minimum and maximum luminance of the monitor. In experiment one the images subtended 29.5 by 18.5 degree of the visual field on a display than subtends 35.5 by 25.5 degrees, corresponding to 60% of the total monitor area. We note that previous results suggest the impact of the background luminance is mostly limited to the immediate surround.^{26,27} To investigate the impact of the surround-image ratio an additional condition is reported in Figure 6 using images that subtend 14.8 by 9.1 degrees, just 14% of the total monitor area.

The results of experiment one are shown in Figure 2. Two main patterns are apparent: First, when a light background is used, subjects tend to prefer a lower system gamma. Second, the preferred system gamma is considerably less for the two images Amikeus Beaver Gam PM1 and Hancock Kitchen than for Peak Lake the preferred system gamma is one. Both findings are consistent with the findings noted in Liu and Fairchild.⁴ We note that for both Amikeus Beaver Gam PM1 and Hancock Kitchen the luminance distribution is low key, in contrast Peak Lake has a relatively flat histogram, consistent with the theory that histogram equalization determines image dependent preference for system gamma.

3.3 Results: Experiment Two

In experiment two we extended the experiment to an additional 13 images (16 in total). As DR, not background luminance was our primary objective we only ran the dark background condition. In Figure 3 we plot the preferred system gamma for each image against the image quality score at that preferred system gamma (i.e. the maximum reported image quality score). The results show a linear relationship (Pearson's correlation $R = 0.85, p < 0.001$). For the images and the dark background luminance condition used, subjects never reported a preferred system gamma of greater than one. The results indicate a preference for linearity, where no histogram equalization is necessary, a point we return to in the discussion.

3.4 Results: Predictive Model

Preferred gamma is affected by both the immediate surround of a display³ and the image in question.^{4,14,15} A unifying theory is that subjects choose the system gamma that produces the flattest (perceived) lightness

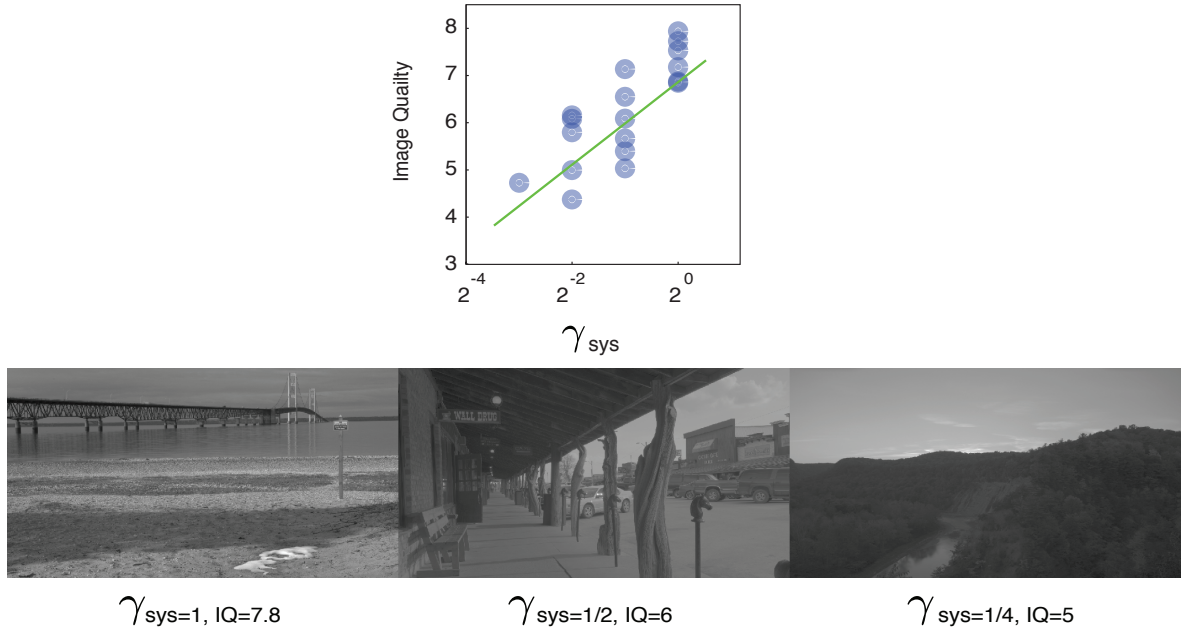


Figure 3. Preferred system gamma against the reported image quality scores for images displayed on a dark background. The results show a linear relationship between preferred system gamma and image quality scores with the highest image quality scores occurring for a gamma value of one. The bottom row illustrates the optimal system gamma for three images.

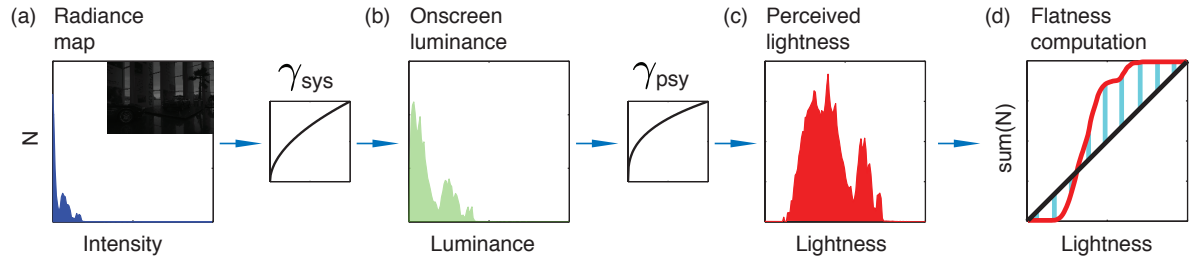


Figure 4. (a) The histogram of the 'raw' HDR image after normalization between the range 0 and 1. (b) The onscreen luminance histogram after being passed through the system gamma of the imaging pipeline. (c) The estimated perceived lightness distribution after being passed through an γ_{psy} . (d) We compute the flatness of a distribution as the root mean square difference between the cumulative lightness distribution and the identity line.

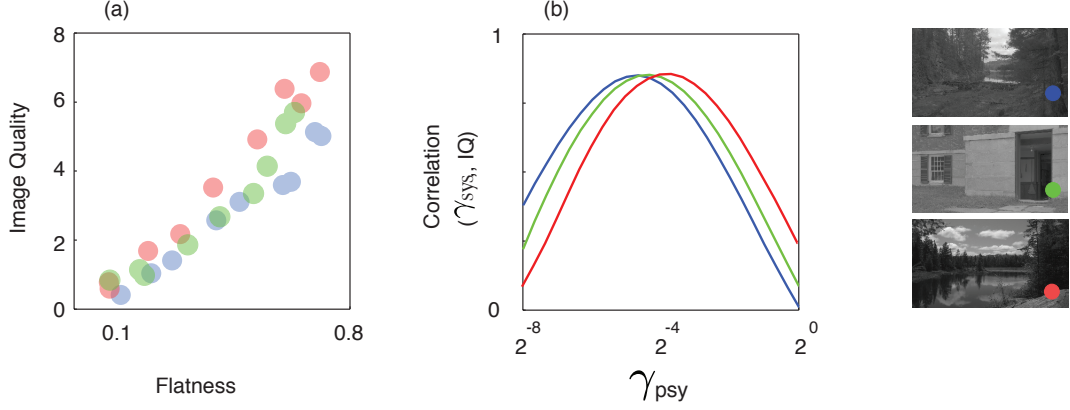


Figure 5. The results of the predictive model are illustrated using the images/data from experiment one. (a) Flatness as a predictor of image quality using the optimal γ_{psy} (b) Pearson's R-score as a function of γ_{psy} .

distribution. This theoretical framework incorporates both the impact of the immediate surround, which is known to modulate the manner in which lightness is perceived, and the statistics of the image in question, which will affect the degree of equalization required. To test this hypothesis we model perceived lightness as a gamma function of onscreen luminance. A gamma function has several advantages (a) it has mathematical simplicity, particularly considering that system gamma can also be modeled as a gamma function and (b) a gamma function can capture the continuum from compressive to expansive non-linearities. We acknowledge that a gamma function may not correctly capture lightness perception under all conditions,^{8–10} but note that lightness perception is highly sensitive to the exact experimental conditions and no research claims to have developed a function that can generalize to all experimental conditions; moreover, only a few studies have investigated lightness perception using natural scenes.^{10,28} Finally, as we shall infer the onscreen luminance to lightness function from our dataset, rather than directly measure the function, we consider it wise to use the most general formula available.

The model schematic is illustrated in Figure 4. The model operates on relative luminance values so the first step is to normalize the HDR radiance map to between 0 and 1. Second, the normalized image N is passed through the system gamma of the display pipeline to generate an estimate of I the relative onscreen intensity values.

$$I_i = N_i^{\gamma_{sys}} \quad (1)$$

Third, the onscreen intensity values I are then passed through an additional point-wise non-linearity to compute an estimate of the perceived lightness distribution L .

$$L_i = I_i^{\gamma_{psy}} \quad (2)$$

Note, stages two and three can mathematically be combined into one stage, but conceptually we prefer to keep perceptual and physical processes separate, particularly given that a more complex model of lightness perception may be needed in the future.

Finally, we compute the cumulative histogram h of the lightness distribution and compute the degree of flatness F as the root mean square difference between the cumulative lightness distribution and the unity line.

$$F = 1 - \frac{1}{n} \sqrt{\sum_{i=1}^n (h_i - i)^2} \quad (3)$$

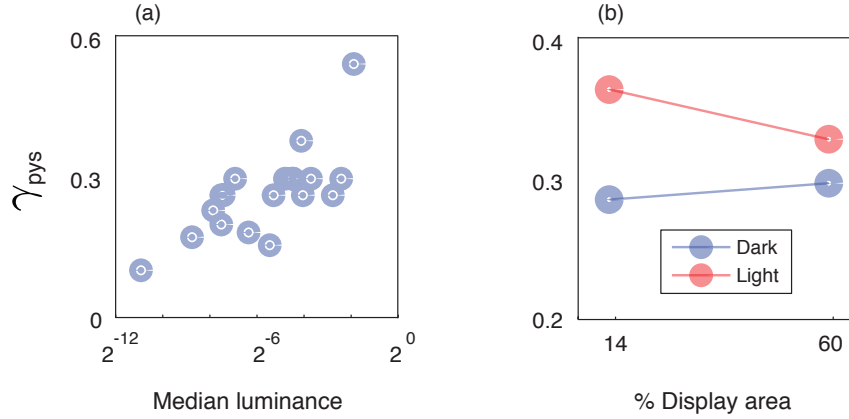


Figure 6. (a) The estimated value of γ_{psy} is positively correlated with the median luminance of the stimulus. (b) We plot how the estimated γ_{psy} varies with both the background luminance condition and the image size. As expected the estimated γ_{psy} is closer to one when a dark background is used. When a small surround luminance is used the estimated γ_{psy} is modulated upward in the light background condition, but there is only a limited impact on the black background condition.

To estimate the value of γ_{psy} that best models human perception, we find the value that gives the maximum correlation between our measure of flatness F and subjects image quality scores. As lightness perception is likely to vary both with the background luminance condition and the image in question, γ_{psy} is fit separately for each condition. This procedure is applied to all sixteen images used. In Figure 5 but we illustrate the model performance using the three images used in experiment one. Figure 5(a) plots the measured flatness for each image using the optimal value of γ_{psy} against subjects average image quality scores. Figure 5(b) plots γ_{psy} against the Pearson's correlation for the three images used. The results show image dependent variability in the selected γ_{psy} . To investigate this further in Figure 6 we plot the medium luminance of a stimulus against the estimated γ_{psy} for all 16 images tested. The results reveal a positive correlation (Pearson's correlation $R = 0.82, p < 0.001$) demonstrating that the luminance distribution of the image in question affects lightness perception. In order to investigate the role of background luminance Figure 6(b) plots the estimated gamma from experiment one and an additional experiment using smaller base images. The larger images used in experiment one covered 60% of the total screen area and the smaller images took up just 14% of the total screen area. The results demonstrate that both image size and background luminance affect the estimated γ_{psy} . As expected from previous research the estimated values of γ_{psy} is closer to zero (i.e. stronger) in the dark background condition. The difference between the two background conditions is greater in the small image condition. This is consistent with the overall luminance distribution of the display (image and background) affecting lightness perception. Consistent with this is the observation that changing the image size has a much greater impact on the estimated γ_{psy} in the light luminance condition. Note that the average luminance of the stimulus is closer to 0 (black) than to 1 (white), thus altering the image size makes a greater difference the mean luminance of the display configuration in the white background condition, than in the black background condition.

In order to generate a model that allows the prediction of optimal system gamma for any image we need to have a predictive model for γ_{psy} . To do so we compute the linear relationship between the log of the median luminance $u_{1/2}$.

$$\hat{\gamma}_{psy} = m \log_2(u_{1/2}) + c \quad (4)$$

We estimate $m = 0.02$ and $c = 0.46$. We acknowledge that these parameters may be influenced by the existence of two outliers and that the model needs to be tested on a new data set. Moreover, any predictive model must also include data for different background conditions and for different sizes of image and background.

However, we note that extending the current research paradigm has the potential to provide answers to these questions.

4. DISCUSSION

We demonstrate that subjects preference for system gamma can be predicted by the degree of histogram equalization in the perceived lightness distribution. This model can explain both how preferred system gamma varies with the background luminance condition and how preferred system gamma varies with the statistics of the image in question.

By extending our model to the full 105 images within the Fairchild high dynamic survey we can investigate the distribution of estimated preferred system gammas. In Figure 1(c) we plot dynamic range (DR) against the system gamma that most flattens the intensity distribution (blue dots) and the system gamma that best flattens the lightness distribution (green dots), where lightness is estimated using a γ_{psy} with a value of a 0.3. We find that for images with a DR of less than three orders of magnitude the estimated preferred gamma is greater than one, consistent with the findings of early studies using low DR images.^{3,14,15} Conversely, when the DR is greater than three orders of magnitude, the estimated preferred gamma is less than one.

The aim of tone-mapping algorithms is to ensure that high DR images look as good as possible on a low DR monitor. Our results demonstrate that a tone-mapping operator that simply searches for the optimal system gamma does not produce satisfactory images if the DR is high (see Figure 3). However the current analysis and image database provides an excellent way to test the performance of tone-mapping operators. If we treat system gamma as a dumb, yet optimized tone-mapping operator, then we can test more advanced algorithms against this baseline. Ideally, tone-mapping operators should be able to rotate the perceived image quality of all images shown in Figure 3 such that they have the same reported image quality as the low DR images. This could be done for a range of target displays with different DRs. Importantly, we suggest that future papers aimed at assessing the performance of tone-mapping operators use an image set that span the full range of dynamic ranges. Note, it is also important that tone-mapping operators preserve the appearance of low DR images.

Finally, a companion paper extends the perceptual theory developed in this paper to develop tone-mapping operators.²⁹

REFERENCES

- [1] Poynton, C., [*Digital video and HD: Algorithms and Interfaces*], Elsevier (2012).
- [2] Brenman, E. J., "The effect of level of illumination and relative surround luminance on the appearance of black and white photographs," *Photographic Science and Engineering* **6**, 172 (1962).
- [3] Bartleson, C. J. and Brenman, E. J., "Brightness reproduction in the photographic process," *Photographic Science and Engineering* **11**, 254–262 (1967).
- [4] Liu, C. and Fairchild, M. D., "Re-measuring and modeling perceived image contrast under different levels of surround illumination," in [*Color and Imaging Conference*], **2007**(1), 66–70, Society for Imaging Science and Technology (2007).
- [5] van der Poel, F. H. J. and Valetton, J. J. P., "The flying spot scanner," *Philips Technical Review* **15**, 221–232 (1954).
- [6] Poynton, C. A., "Rehabilitation of gamma," in [*Photonics West'98 Electronic Imaging*], 232–249, International Society for Optics and Photonics (1998).
- [7] Field, D. J., "Relations between the statistics of natural images and the response properties of cortical cells," *JOSA A* **4**(12), 2379–2394 (1987).
- [8] Whittle, P., "Brightness, discriminability and the crispening effect," *Vision research* **32**(8), 1493–1507 (1992).
- [9] Moroney, N., "Background and the perception of lightness," in [*9th Congress of the International Color Association*], 571–574, International Society for Optics and Photonics (2002).
- [10] Bartleson, C. and Breneman, E., "Brightness perception in complex fields," *Josa* **57**(7), 953–956 (1967).

- [11] Daniels, C. M., Giorgianni, E. J., and Fairchild, M. D., “The effect of surround on perceived lightness contrast of pictorial images,” in [*Color and Imaging Conference*], **1997**(1), 12–16, Society for Imaging Science and Technology (1997).
- [12] Fairchild, M. D., “Considering the surround in device-independent color imaging,” *Color Research and Applications* **20**(6), 352–363 (1995).
- [13] Fairchild, M. D., “The hdr photographic survey,” in [*Color and Imaging Conference*], **2007**(1), 233–238, Society for Imaging Science and Technology (2007).
- [14] Roufs, J. A., Koselka, V. J., and van Tongeren, A. A., “Global brightness contrast and the effect on perceptual image quality,” in [*IS&T/SPIE 1994 International Symposium on Electronic Imaging: Science and Technology*], 80–89, International Society for Optics and Photonics (1994).
- [15] Roufs, J. A. and Goossens, I., “The effect of gamma on perceived image quality,” in [*Display Research Conference, 1988., Conference Record of the 1988 International*], 27–31, IEEE (1988).
- [16] Geisler, W. S. and Perry, J. S., “Statistics for optimal point prediction in natural images,” *Journal of Vision* **11**(12), 14 (2011).
- [17] van Hateren, J. H. and van der Schaaf, A., “Independent component filters of natural images compared with simple cells in primary visual cortex,” *Proceedings of the Royal Society of London. Series B: Biological Sciences* **265**(1394), 359–366 (1998).
- [18] Larson, G. W., Rushmeier, H., and Piatko, C., “A visibility matching tone reproduction operator for high dynamic range scenes,” *Visualization and Computer Graphics, IEEE Transactions on* **3**(4), 291–306 (1997).
- [19] Devlin, K., “A review of tone reproduction techniques,” *Computer Science, University of Bristol, Tech. Rep. CSTR-02-005* (2002).
- [20] Roufs, J. A., “Brightness contrast and sharpness, interactive factors in perceptual image quality,” in [*SPIE 1989 Human Vision, Visual Processing, and Digital Display*], **1077**, 66–72.
- [21] Thurstone, L. L., “A law of comparative judgment,” *Psychological review* **34**(4), 273 (1927).
- [22] Pelli, D. G., “The videotoolbox software for visual psychophysics: Transforming numbers into movies,” *Spatial vision* **10**(4), 437–442 (1997).
- [23] Brainard, D. H., “The psychophysics toolbox,” *Spatial vision* **10**, 433–436 (1997).
- [24] Tyler, C. W., Chan, H., Liu, L., McBride, B., and Kontsevich, L. L., “Bit stealing: how to get 1786 or more gray levels from an 8-bit color monitor,” in [*SPIE/IS&T 1992 Symposium on Electronic Imaging: Science and Technology*], **1066**, 351–364 (1992).
- [25] Liu, C. and Fairchild, M. D., “Measuring the relationship between perceived image contrast and surround illumination,” in [*Color and Imaging Conference*], **2004**, 282–288, Society for Imaging Science and Technology (2004).
- [26] Rudd, M. E. and Zemach, I. K., “Quantitative properties of achromatic color induction: An edge integration analysis,” *Vision research* **44**(10), 971–981 (2004).
- [27] Allred, S. R., Radonjić, A., Gilchrist, A. L., and Brainard, D. H., “Lightness perception in high dynamic range images: Local and remote luminance effects,” *Journal of vision* **12**(2), 7 (2012).
- [28] Nundy, S. and Purves, D., “A probabilistic explanation of brightness scaling,” *Proceedings of the National Academy of Sciences* **99**(22), 14482–14487 (2002).
- [29] Cyriac, P., Bertalmio, M., Kane, D., and Vazquez-Corral, J., “A tone mapping operator based on neural and psychophysical models of visual perception,” in [*SPIE/IS&T 2015 Human vision and electronic imaging XX*], International Society for Optics and Photonics (2015).