

The Role of Lightness Perception in Determining the Perceived Contrast of Real World Scenes

David Kane & Marcelo Bertalmío, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain

Abstract

We investigate the impact of the background luminance upon the perceived image quality of real world scenes. To do so, we generate a set of small image patches that span the full range of mean luminance values and contrasts that may be displayed upon a monitor with a finite luminance range. Subjects viewed the images on a uniform black, grey or white surround and were asked to rate the perceived quality on a scale from 0 to 9. We find that the maximum image quality scores occur for images with a mean luminance of less than half, consistent with the image being passed through a compressive non-linearity before contrast is computed. Moreover, the maximum image quality scores occur at lower mean luminance levels when the background luminance is darker, a pattern consistent with investigations into lightness perception. We conclude that models of contrast perception require an adaptive model of lightness perception. However, we also note the considerable challenge of developing a model of lightness perception that can generalize to any given display configuration.

Introduction

When viewing real world scenes, the perceived lightness of a surface remains constant despite wide variations in the illuminant. However, the perceived lightness of self-illuminating displays appears much less stable. A number of studies investigate lightness perception using simple stimuli, typically composed of a few, small test patches viewed upon an extended, uniform surround. The key finding is that the background luminance level can dramatically alter the relationship between the onscreen luminance of a test patch and the perceived lightness of the patch [7, 17, 21, 24, 18, 4, 25, 19]; In some studies the effect is dramatic, with the luminance-to-lightness function being compressive when the stimuli is viewed upon a dark background, but closer to linear, or even expansive, when viewed upon a light background [17, 22, 3]. Similar findings are observed when the perceived lightness of real world scenes is investigated [4], although the effect of the background luminance is less dramatic and the function is always observed to be compressive [3].

There appears to be a close relationship between sensitivity to lightness variations and the suprathreshold, luminance-to-lightness function described above. Sensitivity to lightness variations is investigated by searching for the minimum luminance variation ΔI between a small test patch and an extended uniform background, with a luminance of I . It is observed that the cumulative of ΔI as a function of I closely matches the luminance-to-lightness functions, leading some to suggest a causal relationship [25, 5], although others note caution [23], see [8] for a recent overview.

ΔI is a measure of the physical contrast between two sur-

faces. Thus contrast and lightness appear closely related. Despite this, recent attempts to model the perceived contrast of real world scenes do not include an adaptive model of lightness perception that varies with background luminance level of a stimulus. This is true for both models in the computer science literature [2] and the vision science literature [1].

In this study we investigate the impact of lightness perception upon the perceived contrast of small patches of natural scenes. To do so, we generate a series of images that sample the full range of mean luminance values and contrasts that can be displayed on a monitor with a finite luminance range. This is achieved by scaling and shifting the luminance range of small image patches using equation 1, where I_o , σ_o and u_o denote the original intensity, standard deviation and mean luminance of the stimulus respectively, and σ_d and u_d , the desired standard deviation and mean luminance.

$$I = \frac{\sigma_d}{\sigma_o} (I_o - u_o) + u_d \quad (1)$$

The stimulus is centrally presented upon a uniform black, white or grey surround corresponding to 0, 50 and 100% of the monitor's maximum luminance. Subjects are asked to rate the perceived image quality on a scale from 0 to 9.

Methods

Observers

Ten observers took part in the experiment. Nine were naïve to the experimental objectives and one was the author. 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 on an Apple MacBook running MATLAB (MathWorks) with functions from the Psychtoolbox [6, 20]. Stimuli were displayed on a Philips 109B CRT monitor with spatial and temporal resolutions of 1280 by 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 by 25.5 degrees. 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), to create a linearization lookup table. The monitors minimum luminance was 0.6cdm^{-2} and the maximum 112cdm^{-2} . No technique was used to increase the effective bit-rate of the monitor.

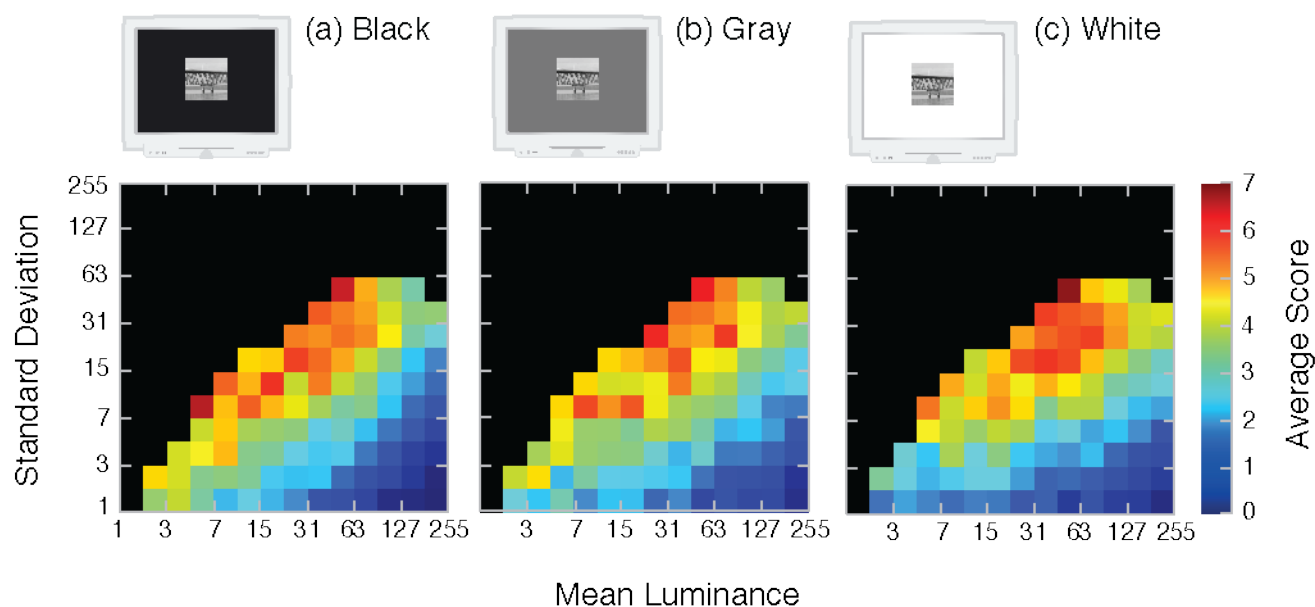


Figure 1. Image quality scores as a function of the physical mean luminance and standard deviation of the stimulus. The stimuli were presented on three background luminance conditions (a) black (b) grey or (c) white. Warmer colors denote higher image quality scores. Each region denotes the scores averaged across six observers and at least sixteen images

Stimuli

Images were .EXR files from the high dynamic range survey by Mark Fairchild. The images can be obtained online and full details of the image registration can be found in [10], but we shall review the main points here. The database consisted of 105 indoor and outdoor scenes. Each image was constructed from between 8 to 18 single exposure images of the same scene taken with one-stop separations in the exposure duration. Images were fused into HDR scenes using Adobe Photoshop Merge-to-HDR function to create the .EXR files that can be obtained online. The .EXR files are radiance maps that contain a linear estimate of the relative luminance values in a scene. Absolute information was obtained by recording from areas in the scene using a photometer and computing an appropriate scaling factor. The HDR radiance maps were resized to a quarter of their original area using bicubic interpolation. Images were then sub-divided into 256 by 256 non-overlapping patches, starting from the top left hand corner, to generate a total of 1440 image patches. Some patches were excluded if they were deemed to have no discernible content when viewed in isolation (e.g. a uniform patch of clear sky).

Procedure

On each trial subjects viewed a centrally presented image patch (7x7 degrees) on a uniform background. Three background conditions were used; black, grey or white, corresponding to the minimum, middle, and maximum luminance values of the monitor. Each background condition was tested in a different run. Subjects were asked to rate the perceived image quality on a scale of 0-9. The subject entered the score via a numeric keyboard and a key press initiated the next trial. No temporal off or on ramp was used; as such the subjects were in complete control of the viewing time. The reported score on the previous trial was shown onscreen to allow the observer to make a correction if a typo had

been made. Subjects were told that a rating of zero should be for stimuli that contained no visible detail (i.e. a uniform tone), otherwise subjects were free to scale their answers as they saw fit and apply their own criteria. A distribution of natural images was created that spanned the full range of mean luminance and contrast values that our CRT could display, as described in the previous section. Each image was presented to each observer twice and each subject viewed the same image set in each of the three background conditions, but each subject viewed a different, but overlapping, distribution of patches.

Results

Observer responses

Subjects were asked to make absolute judgments about the perceived image quality of small patches of natural scenes. A subjective procedure is useful in this case because it allows the experimenter to directly ask the subjective question of interest: how good does the image look? While very few psychophysical paradigms can be considered truly objective [16], the complexity of the stimulus and the potential for different subjects to apply different criteria when making quality judgments means it is important to ascertain how consistent subjects are in their responses. To do so, a double-pass procedure was employed, whereby each subject saw each unique stimulus twice. This allows us to compute the trial-retrial correlations. The results demonstrate that subjects responses are well correlated (Pearson's R-score 0.86, $p < 0.0001$, $N=436$).

To ascertain the variability between subjects, we calculate the mean scores of all subjects to a given stimulus and then correlate with the each subjects mean score for that stimulus. This revealed a strong between-subjects correlation (Pearson's $R=0.88$, $p < 0.0001$, $N=436$). Thus, despite the complex, real world scenes used and the open ended nature of the task, subjects' responses ap-

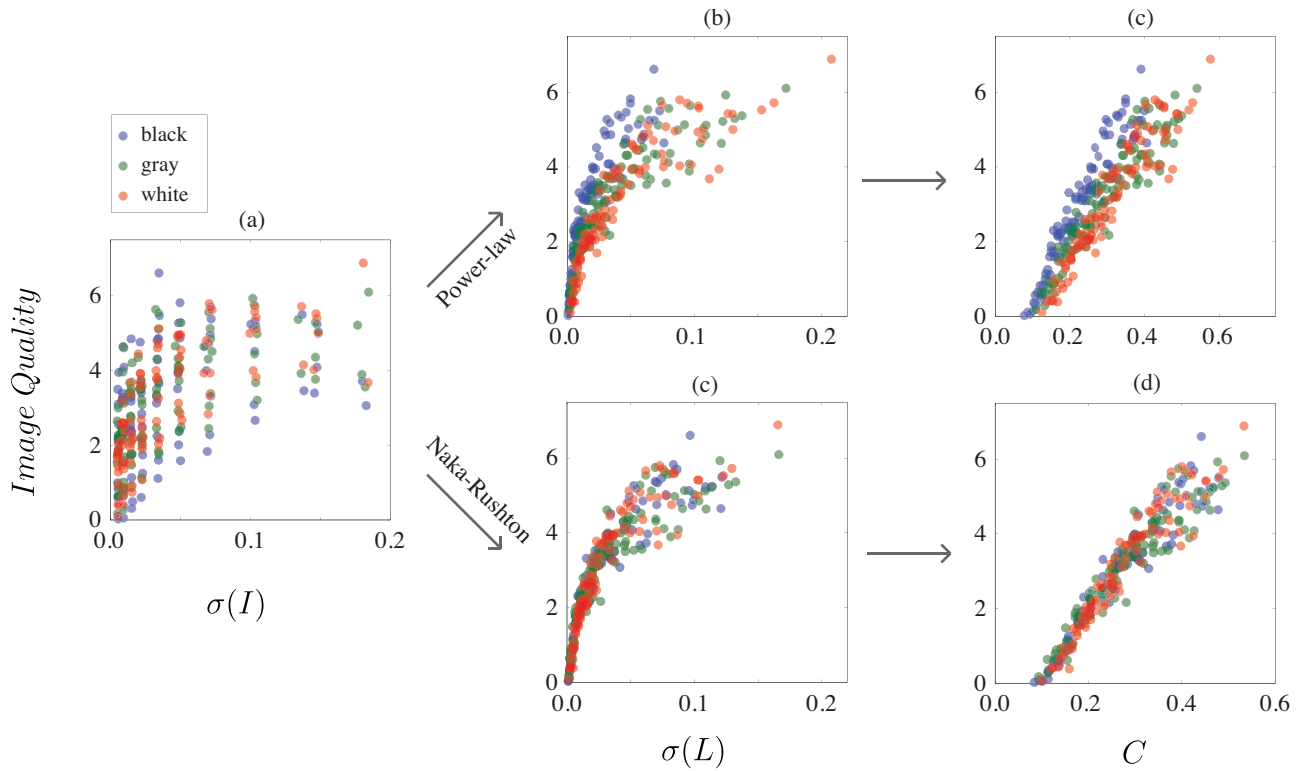


Figure 2. Image quality scores plotted as a function of (a) the physical standard deviation of the stimulus. (b, d) the standard deviation of the lightness image computer using either a power-law (b) or the Naka-Rushton equation (d). (d, e) The estimated contrast after having been passed through an addition power-law nonlinearity of 0.35.

pear to be reliably related to the stimulus content. In the following sections of the paper we demonstrate that the subjects responses can be predicted accurately by a simple contrast model.

Psychophysics

The results are shown in figure 1. Each square region denotes the average image quality score obtained for stimuli with a mean luminance and contrast within the bounds denoted by abscissa and ordinate. The scores are first averaged across the separate images each subject viewed in each luminance-contrast condition and then across subjects, with each subject contributing equally. Warmer colors denote higher scores. Overall, the peak image quality scores are obtained for images with a mean luminance of less than half the full range. This is consistent with the images being passed through a compressive non-linearity which would expand the representation of low luminance values and compress the representation of high luminance values.

The overall distribution of image quality scores shifts in the three background luminance conditions. In the the black background condition higher scores are obtained at low luminance levels, that in the white background condition. This is consistent with the use of a more compressive luminance nonlinearity in the black background conditions. This observation is broadly consistent with investigations of lightness perception in natural scenes [4, 3], as discussed in the introduction.

A simple model of contrast perception

We begin by exploring how well a simple three stage model can capture the data shown in Figure 1. The first stage putatively models lightness perception using a simple power-law:

$$L = N^\gamma \tag{2}$$

where N is the normalized luminance to between 0 and 1 and psychological gamma γ an exponent that adapts with the viewing condition. We term the output of equation 2, the 'lightness image'. The second stage computed the standard deviation of the so-called lightness image

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (L_i - \bar{L})^2}. \tag{3}$$

The third stage applies a second power law to the output of equation 3:

$$C = \sigma^{\gamma_c} \tag{4}$$

This model is optimized to the current data set by searching for the values of γ and γ_c that maximize the Pearson's correlation between the estimated contrast and the image quality scores. We optimize the functions independently for each viewing condition. The results show that the optimal γ increases monotonically as a function of the background luminance level, consistent with a number of studies directly measuring lightness perception [3],

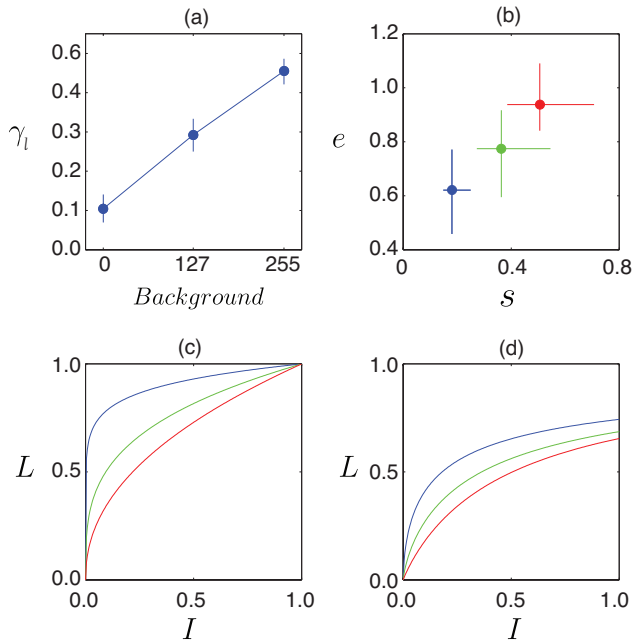


Figure 3. (a) The best fitting gamma-exponents from the power-law model as a function of the background luminance level. (b) A scatter plot of the best fitting parameter from the Naka-Rushton model. The blue, green and red correspond to the black, grey and white background luminance levels respectively. (c) The best fitting power functions (d) The best fitting Naka-Rushton functions.

but that the value of $\gamma_c = 0.35$ is constant for all three background conditions. The optimal Pearson's correlations are all strong ($R = 0.94$, $p < 0.0001$, $N = 106$) when each background condition is considered independently, however when considered as a whole, the correlation drops substantially ($R = 0.84$, $p < 0.0001$, $N = 106$). This is because the subjects' image quality scores span the same absolute range (0-8) for each background condition, but the estimated contrasts span different ranges. This shortcoming can be largely rectified through an addition of a multiplicative factor to the equations 2, 3 or 4. Doing so can lead to a high overall correlation for the pooled dataset ($R = 0.94$, $p < 0.0001$, $N = 106$). The estimated multiplicative factors can either be estimated directly from the data, or estimated by investigating the impact of γ_l upon the overall distribution of contrast upon the distribution of natural images patches.

The Naka-Rushton equation

In the previous section, we modeled lightness perception using a simple power-law. A power-law can approximate the major changes in lightness perception that occur as a function of the background luminance. However, a number of studies indicate the function is better approximated by a function that is non-linear of log-log axes [21, 24, 4]. One function that occurs frequently in the vision science literature is the two parameters Naka-Rushton function (Equation 5).

$$L = \frac{N^e}{N^e + s^e} \quad (5)$$

We repeat the procedure in the previous section to find the best fitting parameters e and s . The results reveal that the esti-

mated values of e and s are positively correlated as a function of the background luminance (Figure 3b). More remarkably however, the estimated functions, despite being optimized separately for each background luminance condition, produce an excellent linear predictor of the data pooled across the background luminance conditions ($R = 0.96$, $p < 0.0001$, $N = 106$). The finding is possible because unlike a simple power-law that estimates lightness perception in the range of zero to one, the Naka-Rushton function allows for different maximal values, as illustrated in Figure 3d.

The contrast sensitivity function

The contrast sensitivity function (CSF) describes the minimum contrast needed to detect a grating with a given probability. The CSF is critical to understanding many aspects of vision [9]. For instance, without considering the CSF a region of an image viewed at 1m or say, 100m, would contain the same visual detail. Accordingly, any general purpose model of contrast perception must include a model of the CSF. However, it is not clear to what extent the CSF determines the perceived image quality of a real world image viewed at a fixed distance. In particular, real world images contain content that is both spatial frequency broadband and phase aligned [15]. Thus information lost at high or low spatial frequencies (for which we are less sensitive), are likely to still be found at the intermediate spatial frequencies.

To investigate the impact of the CSF we repeat the modeling described above, but first pass the image through the CSF by Mannos and Sakrison [13] (Equation 6).

$$CSF = 2.6 \left(0.0192 + \frac{0.114}{SF} \right) e^{(-0.114/SF)^{1.1}} \quad (6)$$

The results demonstrate that the application of the CSF model leads to a small drop in the predictive power of the model ($R = 0.94$, $p < 0.0001$, $N = 106$). Additionally we note that the contrast estimates with and without the CSF are highly correlated ($R = 0.93$, $p < 0.0001$, $N = 106$). Thus the application non-flat CSF only introduces a modest effect. This compares to the impact of the two lightness functions derived using the Naka-Rushton equation in the white and black background luminance conditions which lead to less highly correlated estimates of contrast ($R = 0.82$, $p < 0.0001$, $N = 106$). We note that we are not making any conclusions about whether the form of the CSF estimated by Mannos and Sakrison [13] is correct, only that the impact of the CSF on the current dataset is limited relative to that of the lightness function.

Discussion

We investigated image quality scores for small patches of real world images presented on a uniform background. The results show a clear effect of the background luminance level upon the perceived contrast, and in turn, perceived image quality, of the small natural image patches. This pattern of results is broadly consistent with the established effect of the background luminance on the luminance-to-lightness function [3], however comparison with the literature on lightness perception is difficult because the exact nature of the luminance-to-lightness function appears highly sensitivity to the exact experimental conditions, as evidenced by the diversity of estimated functions

[7, 17, 21, 24, 18, 4, 25, 19, 22, 3]. Moreover, research on simple stimuli does not appear to generalize to the perception of real world scenes [3]. One possible explanation for this discrepancy is the manner in which the 'background', 'surround' and the 'test' of a stimulus is defined. For simple stimuli, the test stimuli normally cover a small region of the central visual field. Thus, the so-called background dominates both the central and peripheral visual fields. In contrast, in studies using real world scenes, the test stimulus dominates the central visual field. As noted by Bartleson and Breneman [4] the lightness function is likely to be determined by both the test stimulus and the surround luminance of a display. Without a model of lightness perception that operates on a general description of the stimulus, rather than the experimenter defined 'background', 'surround' and 'test', a satisfactory, general purpose model of lightness perception is out of reach. By extension, a general model of contrast perception is also arguably out of reach.

A second observation is that subjects report the same range of image quality scores regardless of the background luminance condition. This finding is not predicted by the model of contrast perception outlined in this paper when lightness is modeled as a simple power law. This is because a compressive non-linearity increases the average contrast (of the distribution of image patches used in this study), up until an exponent of around 0.3 at which point the contrast decreases again. Accordingly, the only way to predict the same pattern of results using the power-law model of lightness perception is to include an additional multiplicative factor to correct for this. Interestingly, and to the authors considerable surprise, the corrective factor was not required when lightness perception was modeled using the Naka-Rushton equation. This occurred despite the fact we fitted each function independently for the three background conditions. Interestingly the best fitting Naka-Rushton functions had different maximal outputs as shown in figure 3 (we stress that each function was fit independently for each background luminance condition). This result is not intuitive because it indicates that the maximal perceived lightness varies between the conditions. An alternative possibility is that the minimum luminance, rather than the maximal luminance is affected by the background luminance. This could result from the impact of scatter which has been demonstrated to affect the low luminance levels disproportionately and is more of an issue when the background luminance is lighter [24].

Conclusion and future work

We conclude that lightness perception is critical to the understanding of perceived contrast. However, as in this work, lightness perception is inferred, not directly measured, there remains considerable ambiguity in what the most appropriate luminance non-linearity is. Accordingly, in the absence of a suitable model of lightness perception future work should aim to both estimate lightness- and contrast- perception using similar experimental setups. This is the aim of a separate, but conceptually linked study, also submitted to the Electronic Imaging conference [11].

A note on image statistics

Although the mean luminance and contrast in the stimuli was obtained via artificial methods, the simplicity of the scaling and shifting operations meant that we were nonetheless constrained in the type of stimuli that could be presented, see [14] for a related

discussion on the luminance-contrast statistics of natural scenes. Moreover, we were constrained in the range of mean luminance and contrast values that could be presented due to the limited luminance range of the monitor. For instance, image patches with a low or high mean luminance are constrained in the highest contrast they can have due to floor and ceiling effects.

In this work it is assumed that image contrast is directly related to image quality and this assumption is born out by the high correlation coefficients obtained by the simple contrast model. However, we fully acknowledge that image quality is related to a number of other image factors. One issue we wish to stress is the concept of global and local contrast. The logical extension of our model would suggest that a binary stimulus, composed of zeros and ones (maximum and minimum luminance values) would receive the highest image quality scores as they maximize the standard deviation of a stimulus. This finding is at loggerheads with other findings on image quality, in particular, the finding that image quality scores may be related to the degree of histogram equalization in a stimulus [12]. Stimuli that maximize the standard deviation include a salt and pepper noise, or a simple step function. These are stimuli that maximize global contrast. In the former case the local contrast is the same as the global contrast, while in the latter case, the local contrast is zero aside from at the step boundary. This dichotomy is likely to affect image quality judgements and is the subject of ongoing experimentation where we study image quality scores for individual patches, as opposed to averaging across images as we do in this study. Early results suggest that images with high global contrast often receive low image quality scores, suggesting that the model will have to be adjusted accordingly.

One criticism of the current investigation is that each background condition is presented to the observer on different experimental runs. Thus there is the potential for subjects to rescale their answers to within a given range (e.g. 0 to 9) for each condition. Ongoing research reveals that when different images and background conditions are interleaved, or presented on separate runs, the overall pattern of results is similar.

References

- [1] HM Andrew and E Peli. Perceived contrast in complex images. *Journal of vision*, 13(13):3, 2013.
- [2] TO Aydin, R Mantiuk, K Myszkowski, and HP Seidel. Dynamic range independent image quality assessment. In *ACM Transactions on Graphics (TOG)*, volume 27, page 69. ACM, 2008.
- [3] CJ Bartleson. Optimum image tone reproduction. *Journal of the SMPTE*, 84(8):613–618, 1975.
- [4] CJ Bartleson and EJ Breneman. Brightness reproduction in the photographic process. *Photographic Science and Engineering*, 11:254–262, 1967.
- [5] EG Boring. *Sensation and perception in the history of experimental psychology*. D. Appleton-Century Company, 1942.
- [6] David H Brainard. The psychophysics toolbox. *Spatial vision*, 10:433–436, 1997.
- [7] EJ Breneman. 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.
- [8] MH Brill and RC Carter. Does lightness obey a log or a

power law? or is that the right question? *Color Research & Application*, 39(1):99–101, 2014.

- [9] FW Campbell and JG Robson. Application of fourier analysis to the visibility of gratings. *The Journal of physiology*, 197(3):551–566, 1968.
- [10] MD Fairchild. The hdr photographic survey. In *Color and Imaging Conference*, volume 2007, pages 233–238. Society for Imaging Science and Technology, 2007.
- [11] David Kane and Marcelo Bertalmío. The influence of lightness, and the crispening effect on the perceived contrast of textured images. In *Proc. IS&T International Symposium on Electronic Imaging*, 2016.
- [12] David Kane and Marcelo Bertalmío. System gamma as a function of image- and monitor- dynamic range. *Journal of Vision*, Under review.
- [13] JL Mannon and DJ Sakrison. The effects of a visual fidelity criterion of the encoding of images. *Information Theory, IEEE Transactions on*, 20(4):525–536, 1974.
- [14] Valerio Mante, Robert A Frazor, Vincent Bonin, Wilson S Geisler, and Matteo Carandini. Independence of luminance and contrast in natural scenes and in the early visual system. *Nature neuroscience*, 8(12):1690–1697, 2005.
- [15] D Marr and E Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London B: Biological Sciences*, 207(1167):187–217, 1980.
- [16] Michael J Morgan, D Melmoth, and Joshua A Solomon. Linking hypotheses underlying class a and class b methods. *Visual neuroscience*, 30(5-6):197–206, 2013.
- [17] N Moroney. Background and the perception of lightness. In *9th Congress of the International Color Association*, pages 571–574. International Society for Optics and Photonics, 2002.
- [18] AEO Munsell, LL Sloan, and IH Godlove. Neutral value scales. i. munsell neutral value scale. *Josa*, 23(11):394–402, 1933.
- [19] S Nundy and D Purves. A probabilistic explanation of brightness scaling. *Proceedings of the National Academy of Sciences*, 99(22):14482–14487, 2002.
- [20] Denis G Pelli. The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, 10(4):437–442, 1997.
- [21] A Radonjić, SR Allred, AL Gilchrist, and DH Brainard. The dynamic range of human lightness perception. *Current Biology*, 21(22):1931–1936, 2011.
- [22] JC Stevens and SS Stevens. Brightness function: Effects of adaptation. *JOSA*, 53(3):375–385, 1963.
- [23] SS Stevens. To honor fechner and repeal his law. *Science*, 1961.
- [24] WA Stiehl, JJ McCann, and RL Savoy. Influence of intraocular scattered light on lightness-scaling experiments. *JOSA*, 73(9):1143–1148, 1983.
- [25] P Whittle. Brightness, discriminability and the crispening effect. *Vision research*, 32(8):1493–1507, 1992.

Author Biography

David Kane received his PhD from University Collage London. From there David pursued an interest in working at the intersection of perception and technology and worked with Professor Marty Banks at The University of California Berkeley working on

issues around stereo 3D displays. Currently, David is working as in-house experimental psychologist in the image processing lab Marcelo Bertalmío at Universitat Pompeu Fabra.

Marcelo Bertalmío received the Ph.D. degree in electrical and computer engineering from the University of Minnesota in 2001. He is an Associate Professor at University Pompeu Fabra in Barcelona, Spain. His interests are Image Processing and Computer Vision for digital cinema applications, although he prefers the (analog) films of Preston Sturges and Luis Buuel.