PU21: A novel perceptually uniform encoding for adapting existing quality metrics for HDR*

Rafał K. Mantiuk
Computer Science and Technology
University of Cambridge
Cambridge, UK
rafal.mantiuk@cl.cam.ac.uk

Maryam Azimi
Computer Science and Technology
University of Cambridge
Cambridge, UK
maryam.azimih@gmail.com

***Abstract*—Standard image quality metrics, such as PSNR or SSIM, cannot be directly computed on linear high dynamic range colour values because such values non-linearly related to our perception of visible differences. In this work, we develop a new encoding function (PU21) to convert absolute high dynamic range (HDR) linear colour values into approximately perceptually uniform (PU) values, which can be used with standard quality metrics. The proposed PU21 function is based on a recent contrast sensitivity model, fitted to the measurements up to 10 000 cd/m^2. Unlike the conventional simplified approach of deriving PU functions based on peak sensitivities, we model realistic coding artefacts to find visibility thresholds for our derivation. Furthermore, the new PU accounts for the effect of glare on image quality. The proposed PU21 improves the accuracy of quality predictions for standard metrics of PSNR, VSI, FSIM, SSIM, and MS-SSIM in their correlation with subjective scores on HDR images included in UPIQ, one of the largest HDR image quality datasets.**

***Index Terms*—high dynamic range, image quality metric, contrast sensitivity, Perceptually Uniform encoding, banding artifacts, glare**

I. Introduction

Objective image and video quality metrics are vital components in development of many imaging applications such as acquisition, compression, communication, display systems, computer vision, and machine learning. While a large number of image and video quality metrics are designed for the assessment of standard dynamic range (SDR) content, far fewer metrics are available for high dynamic range (HDR) content. The two most commonly used HDR metrics are HDR-VQM [1] and HDR-VDP [2]. Although those metrics model many aspects of the human visual system (HVS), such as luminance and contrast masking, and can offer superior performance, they are too complex to be used in many applications. Because they are non-differentiable, they are unsuitable to be used as a loss function in optimization problems. When short execution times are essential, a simple arithmetic or structural metric, such as PSNR or SSIM [3], is often preferred.

Metrics such as PSNR and SSIM expect their input to be in an approximately perceptually uniform domain — Eucliden distances between just-distinguishable colours must not vary drastically across the colour space. Most SDR colour spaces, using gamma or similar pixel coding, are sufficiently perceptually uniform to directly use them with traditional quality metrics. However, HDR content is often represented as linear radiance values, which are strongly perceptually non-uniform. Therefore, HDR linear pixel values must be converted to an approximately perceptually uniform colour space before computing quality values [4].

Aydın et al. [5] proposed a perceptually uniform (PU) encoding for HDR quality assessment, converting absolute HDR linear RGB or luminance values into perceptually uniform values. The approach was shown to be an effective method for adapting existing SDR quality metrics to HDR content [6], [7]. Perceptual Quantizer (PQ) electro-optical transfer function (EOTF) [8], intended for encoding HDR pixel values, was also used for the similar purpose, even though it has not been designed for quality assessment [9].

In this work, we improve upon both PU and PQ functions and develop a new perceptually uniform encoding, specifically modelling the sensitivity to distortions in HDR content. Firstly, we make use of a contrast sensitivity function (CSF) that consists of measurements at background luminance values as low as 0.0002 cd/m^2 to as high as 10 000 cd/m^2 [10]. Secondly, we make more realistic assumptions about the spatial-frequency composition of typical image artefacts and use a recently proposed HDR *banding model* to predict sensitivity across HDR luminance range. Finally, we account for the effect of glare on detection thresholds. All those changes substantially improve the accuracy of popular quality metrics, including PSNR, SSIM, MS-SSIM, FSIM and VSI, when tested on one of the largest HDR image quality datasets.

II. Background and Related Work

Image quality has been traditionally evaluated disregarding the characteristic of the display on which the content is to be viewed. We discourage this practice, especially for HDR content, as the visibility of distortions can vary substantially with the display's peak luminance [11]. To assess image quality in a display-dependent manner, we will follow the processing shown in Fig 1. The HDR content, regardless of its representation in display-encoded colour space such as PQ and

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement N° 725253–EyeCode)

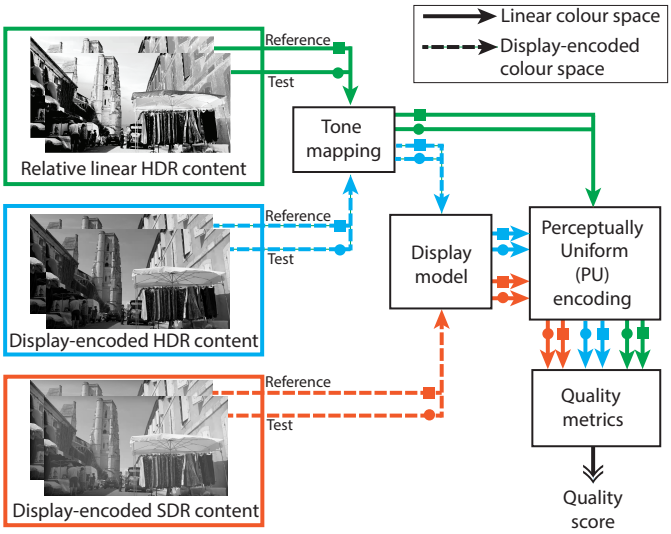

Fig. 1. Display-dependant HDR image quality assessment pipeline using the PU encoding with existing quality metrics for display-encoded and linear colour space input. Display-dependant SDR image quality assessment pipeline for display-encoded input is also shown.

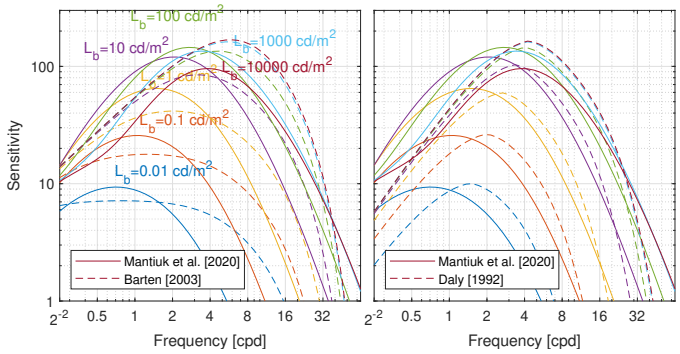

Fig. 2. Comparison of the new CSF model from [10] (solid lines) with Barten's [16] (dashed lines, left), and Daly's (dashed lines, right) CSF models. Barten model was used to derive the PQ EOTF and Daly's model was used for PU08. The new CSF function, based on the recent measurements [17], is significantly different from both models.

Hybrid-Log-Gamma (HLG) or in relative linear space, must be *tone mapped* for a particular HDR display. Tone mapping ensures that the displayed image does not exceed the peak display luminance, which could be as high as 4 000 cd/m^2 and as low as 400 cd/m^2 for some HDR displays. Next, a *display model* is used to simulate the amount of light that is emitted from the display in the physical units of luminance (cd/m^2). The luminance is then perceptually encoded and passed to a quality metric. Such a pipeline can be also used to evaluate the quality of SDR content in a manner that accounts for display brightness.

A. Display Model

Display models are used to predict light emitted from a display when driven with display encoded pixel values. The response of many displays can be modelled as:

$$L = (L_{\text{peak}} - L_{\text{black}}) f(V) + L_{\text{black}} + L_{\text{amb}} \quad (1)$$

where L is luminance emitted from the display, and V is display-encoded luma varying between 0 and 1. L_{peak} is the peak luminance, and L_{black} is the black level of the target display. L_{amb} is the ambient light that is reflected from the surface of a display. f is the EOTF, the inverse of the opto-electronic transfer function (OETF). In case of SDR, the EOTF represents the inverse of the gamma or sRGB non-linearity. For HDR it could be any of the HDR perceptual transfer functions such as PQ [8] or HLG [12]. The output from the display model represents the luminance values emitted from the target display.

B. Perceptually Uniform Encodings

Since linear HDR colour values are strongly perceptually non-uniform, they are unsuitable for coding or quality evaluation. The simplest encoding that improves perceptual uniformity is the logarithmic function, originally proposed in the context of encoding HDR pixels in [13]. While logarithmic function is a good first-order approximation of the sensitivity to light, it does not account for the lower sensitivity of the HVS at low luminance levels. This problem was addressed in [14] where encoding was derived from the psychophysical threshold-versus-intensity function. The derivation relied on the principle that a unit increment in the encoded space should correspond to the just-noticeable difference in the luminance space. In [5] the encoding was re-purposed to be used with quality metrics but it was derived from the peaks sensitivities of Daly's CSF [15]. The perceptual encoding in [5], which we will refer to as PU08 in this paper, was optimized to approximate gamma-encoding within the range from 0.1 to 80 cd/m^2, the typical range of luminance that SDR displays could reproduce. Based on the same principles, PQ function [8] was derived from Barten's CSF [16] and was standardized in ITU-R Recommendation BT.2100 as the EOTF for high dynamic range content.

III. New Perceptually Uniform Encoding — PU21

In this paper we extend and improve upon PU08 [5] and propose a new perceptually uniform encoding for HDR image and video quality assessment. The new encoding, which we will call PU21, is derived from the latest contrast sensitivity data (Sec. III-A), models realistic coding artefacts (Sec. III-B), and accounts for glare (Sec. III-D). The proposed PU21 improves the accuracy of quality predictions for a range of standard metrics on tested HDR images (Sec. IV).

A. New Contrast Sensitivity Function

PU21 is based on a new CSF [10] that was fitted to the data from five independent datasets, predicting contrast thresholds at luminance levels between 0.0002 cd/m^2 and 10 000 cd/m^2. One of the most recent of these datasets include contrast sensitivity measurements conducted on a very bright HDR display [17]. The main difference between the new CSF and the previous models [16], [18], used for PU08 and PQ, is the

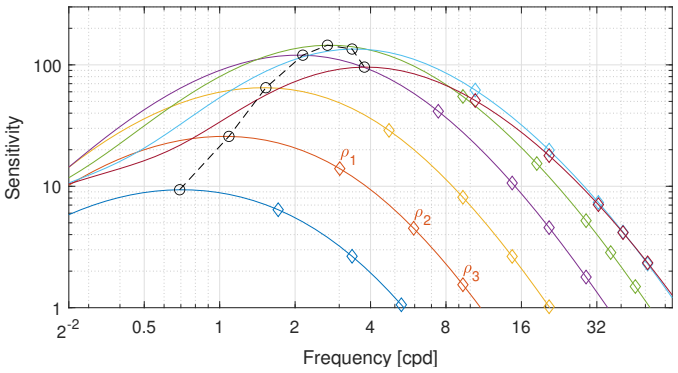

Fig. 3. PQ function has been derived from the peak sensitivities of the CSF (black circles). The new banding model relies on the probability summation of a number of frequency components, shown as diamond-shaped markers.

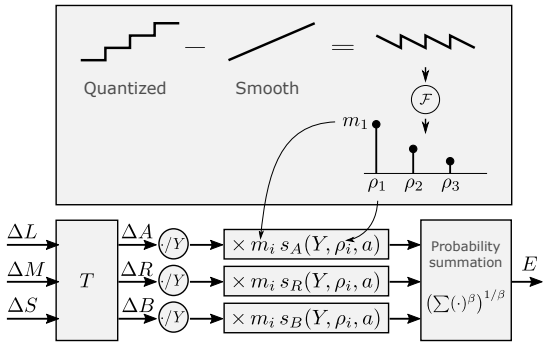

Fig. 4. HDR Banding detection model. The difference between the quantized and smooth (unquantized) signals is transformed into the Fourier domain to find the spatial frequencies (ρ_i) and amplitudes (m_i) of the banding artefacts (top box). Those are used to find a detection threshold using an energy model (bottom blocks) operating on a colour-opponent signals (ΔA, ΔR, ΔB), that are normalized by the luminance Y. ΔR refers to L-M difference (red-green) and ΔB refers to L-S difference (yellow-violet). The signal for each colour and frequency component is modulated by the spatio-chromatic CSF ($s_{\{A,R,B\}}(\cdot)$) from [10], which is a function of luminance Y, spatial frequency ρ, and stimulus size a.

drop of sensitivity at luminance values above 100 cd/m^2 (refer to 3). The previous CSF models assumed the sensitivity at high luminance level stays constant. Due to this difference, the PU21 derived from the new CSF model assumes that artifacts in very bright areas are less visible and therefore less severe.

B. HDR Banding Model

Derivation of an encoding function requires the assumption about the type of artefacts it should model. Both PQ and PU08 were derived assuming that the content contains artefacts whose fundamental frequency corresponds to the peak sensitivity of the CSF. Given this assumption, we can select the sensitivity values at the peaks and reduce a multi-dimensional CSF to single dimension of luminance (see Fig. 3). This assumption, however, is rather unrealistic as it entails that artefacts have a Gabor or wavelet-like shape with underlying frequency corresponding to the peak. In practice, most coding artefacts are caused by quantisation and result in banding or contouring artefacts. To model the sensitivity to banding artefacts, we rely on the banding detection model from [19], which extends the previous model [20] by accounting for a far greater luminance range.

The diagram of the banding model from [19] can be found in Fig. 4. The model predicts both luminance and chromatic banding artefacts, taking as input the background colour in LMS colour space and the direction in which the colours change within a smooth gradient of colours (ΔL, ΔM, ΔS). In this work, we only consider banding in luminance and assume the background chromacity corresponds to D65 white point. The model assumes the worst-case scenario of a smooth gradient that is quantised. It decomposes the quantisation error into its spatial frequency components using an analytical Fourier transform of a saw-tooth function. The transform results in a series of frequencies, ρ_1, .., ρ_n (shown as diamonds in Fig. 3) and amplitudes, m_1, ..., m_n. The amplitudes are modulated using the CSF (denoted as $s_{A/R/B}$ in Fig. 4) from [10] and then pooled together using the probability summation to find the probability of banding detection. If such probability exceeds a pre-determined threshold, the banding artefacts are predicted to be visible.

To determine the sensitivity to banding artefacts at each luminance level, we perform a binary search to find the contrast associated with the height of the saw-tooth function that results in visible banding. The sensitivity is the inverse of that contrast: $S = {}^{L}/_{\Delta L}$. The detection contrast for the banding model and peak sensitivities of the CSF are shown on the left of Fig. 5.

C. Numerical Integration

Once we find the sensitivity associated with each luminance level, we can derive PU encoding by integrating the inverse of detection thresholds, as demonstrated in [21, Sec. 2.4] and [14]. Here, we focus on a practical aspect of numerical integration. Because the thresholds have either complex form, or are found numerically, the integration used to find the PU encoding must be also numerical. Since the sensitivity to luminance is highly non-linear, the thresholds must be sampled on the logarithmic scale of luminance; i.e. we are integrating over $l = \log_{10}(L)$ rather than luminance L. Such a substitution of variables requires the introduction of a Jacobian determinant into our numerical integral:

$$P(L_k) = \sum_{i=1}^{k} \frac{S(L_i)}{L_i} L_i \log(10) \Delta = \sum_{i=1}^{k} S(L_i) \log(10) \Delta \,, \tag{2}$$

where $P(L_k)$ is the perceptually uniform value at luminance L_k, $S(L_i)$ is the sensitivity, $L_i \log(10)$ is the Jacobian determinant, and Δ is the difference between the logarithms of two luminance levels: $\Delta = \log_{10} L_{i+1} - \log_{10} L_i$. The values L_i are distributed on the logarithmic scale between 0.005 and 10 000 cd/m^2.

D. Glare-Adaptive Perceptually Uniform Encoding

The existing PU functions, derived from the threshold functions for fully adapted eye, assume the ideal viewing

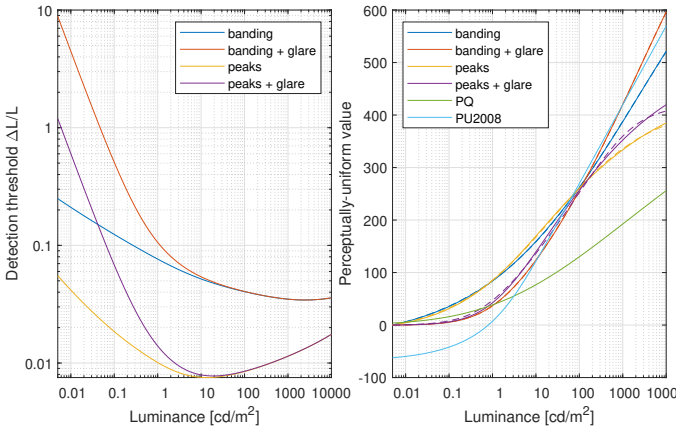

Fig. 5. *Left:* contrast detection thresholds for each variant of the PU21 curve. *Right:* The PU21 encoding functions, rescaled so that the values between 0.1 and 100 cd/m^2 correspond to 256 steps. *Banding* functions are derived from the banding model and *peaks* from the peak sensitivities of the CSF. PQ (scaled by 256) and PU08 are included for comparison. The dashed lines represent analytical functions fitted to the numerical derivation of the PU21 curves.

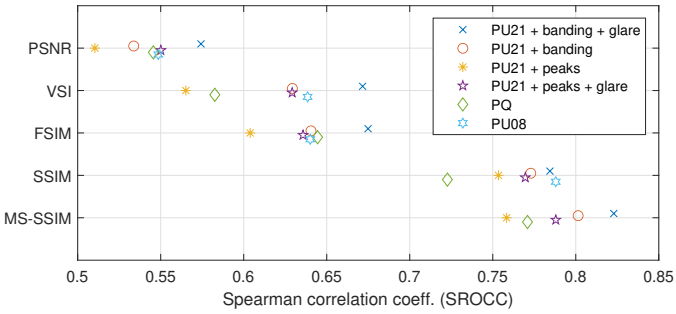

Fig. 6. Quality predictions for the HDR images of the UPIQ dataset. Different markers denote different PU21 functions used to encode HDR pixel values. The PU21 based on the new banding model that accounts for glare results in higher correlation with the subjective judgements for majority of the metrics. The result for PU08 and MS-SSIM is missing because MS-SSIM cannot handle negative pixel values, which are produced by PU08 encoding.

conditions: a dark room and a mostly uniform (low dynamic range) image. In practice, content is typically viewed at non-negligible ambient light levels and HDR content can contain very bright areas, which both introduce glare — scattering of the light in the eye's optics and on the retina. Such a glare can make details in darker image parts harder to see. The effect of glare in visible thresholds is ignored in both PQ and PU08 design and hence those functions do not reflect how artefacts are perceived when glare is present in an image.

To account for glare, we need to modify the sensitivity values. Assuming the amount of glare is uniform in the image and equal to L_g, the sensitivity corrected for glare can be computed as:

$$S_g(L) = \frac{S(L + L_g)\, L}{(L + L_g)}\,. \tag{3}$$

The above function assumes that the glare will increase the luminance on the retina by L_g, thus decreasing the physical contrast by a factor of $L/L+L_g$. The sensitivity values with and without glare-adaptive encoding are plotted on the left of Fig. 5. Note how glare affects sensitivity at lower luminance values the most. The assumption of the uniform amount of glare across the image is a simplification that let us use PU21 encoding as a spatially invariant operator. Modeling of spatially-varying glare requires convolution with large kernels [2], which could be too expensive for simple metrics.

E. PU Encoding Functions

We follow the steps above to compute four combinations of possible PU functions: using the peak CSF sensitivities (as PU08 and PQ) or the HDR banding model, and with or without the glare-adaptive encoding. Here, we assume the glare of $L_g = 0.5$ cd/m^2. Because we want the PU-encoded images to result in similar quality metric predictions as their SDR counterparts, we introduce one more step: the PU values are rescaled by a constant, so that the range of luminance between 0.1 cd/m^2 and 100 cd/m^2 is mapped to the 256 steps of the PU values. This luminance range is typical for many SDR displays and therefore should be mapped to 256 code values found in 8-bit SDR images. PU08 encoding contained a similar step. However, the values below 0.5 cd/m^2 were mapped to negative values, making PU08 usuitable for many quality metrics. As the final step, we fit analytical invertible functions to the numerical solutions so that the the encoding can be easily implemented with a single equation[1]. All the variants of the new PU21 functions are shown on the right of Fig. 5.

IV. Results

To evaluate the new PU21 encoding and to choose the best variant from the four developed combinations, we employ the new PU21 encodings to predict HDR image quality in UPIQ dataset [22]. This dataset consists of over 3779 SDR and 380 HDR image pairs. The dataset was created by aligning and rescaling quality scores from 2 SDR (TID2013 [23], LIVE [24]) and 2 HDR datasets [25], [26]. We use only the HDR portion of the dataset for our evaluation as this is the target application of PU21. UPIQ test images contain more than 25 types of distortions, such as noise, blur, and compression. We selected this dataset because it is the largest HDR image quality dataset available. Furthermore, the HDR images in UPIQ dataset are scaled in absolute colourimetric units, eliminating the need for tone mapping. We use four variants of the new PU21 encoding, PU08 and PQ to encode each image pair, following the procedures shown in the Fig. 1. We test the commonly used image quality metrics: PSNR, SSIM, VSI [27], FSIM [28], and MS-SSIM [29].

The correlations between metric predictions and subjective quality scores are shown in Fig. 6 in terms of Spearman Correlation Coefficient (SROCC). Out of the four variants of the new PU21 encoding tested, the one that combines the sensitivities derived from the HDR banding model and accounts for glare outperforms the other three. When comparing

[1]The PU21 code is available at: https://github.com/gfxdisp/pu21

glare/no-glare variants, accounting for glare clearly improves metrics performance. When comparing PU functions based on peak sensitives versus HDR banding model, using the latter, which better models the coding artefacts, clearly shows an advantage over using only peak sensitives.

The new PU21 based on HDR banding model and with glare adjustment clearly improves metrics predictions as compared to both PU08 and PQ EOTF. Overall, the new PU21 based on the HDR banding model and with glare encoding yields the highest correlation with subjective results. This is valid for all tested metrics, except for SSIM, for which its performance is comparable to that of the PU08 encoding. Applying PU21 does not impose any changes on the quality metric and can be coupled with any metric to improve its performance on HDR image quality prediction.

V. Conclusions

We proposed a new, more comprehensive perceptually uniform (PU) encoding function to adapt the existing quality metrics to HDR content. The proposed PU21 relies a new contrast sensitivity function that accounts for luminance levels well beyond the range supported by current HDR displays. It models the visibility of realistic coding artifacts, such as quantisation, which can be predicted using a new HDR banding model. The change of vision sensitivity and hence image quality in presence of glare is also considered in the PU21 by adding a glare-adaptive encoding. Performance evaluations on HDR images in UPIQ, one of the largest image quality datasets, showed that the proposed PU21 based on the HDR banding model and including glare-adaptive encoding improves the performance of the most commonly used image quality metrics of PSNR, VSI, FSIM, SSIM and MS-SSIM compared to the old PU08 and PQ.

References

[1] M. Narwaria, M. P. Da Silva, and P. Le Callet, "HDR-VQM: An objective quality measure for high dynamic range video," *Signal Processing: Image Communication*, vol. 35, pp. 46–60, 2015.

[2] R. K. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM T. Graphic*, vol. 30, no. 4, pp. 40:1–40:14, jul 2011.

[3] Z. Wang and A. C. Bovik, "Modern image quality assessment," *Synthesis Lectures on Image, Video, and Multimedia Processing*, vol. 2, no. 1, pp. 1–156, 2006.

[4] R. K. Mantiuk, "Practicalities of predicting quality of high dynamic range images and video," in *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2016, pp. 904–908.

[5] T. O. Aydın, R. Mantiuk, and H.-P. Seidel, "Extending quality metrics to full luminance range images," in *Human Vision and Electronic Imaging XIII*, vol. 6806. International Society for Optics and Photonics, 2008, p. 68060B.

[6] G. Valenzise, F. De Simone, P. Lauga, and F. Dufaux, "Performance evaluation of objective quality metrics for HDR image compression," in *Applications of Digital Image Processing XXXVII*, A. G. Tescher, Ed., sep 2014, p. 92170C.

[7] A. Artusi, R. K. Mantiuk, T. Richter, P. Hanhart, P. Korshunov, M. Agostinelli, A. Ten, and T. Ebrahimi, "Overview and evaluation of the JPEG XT HDR image compression standard," *Journal of Real-Time Image Processing*, vol. 16, no. 2, pp. 413–428, apr 2019.

[8] S. Miller, M. Nezamabadi, and S. Daly, "Perceptual signal coding for more efficient usage of bit codes," *SMPTE Motion Imaging Journal*, vol. 122, no. 4, pp. 52–59, 2013.

[9] A. Luthra, E. Francois, and W. Husak, "Call for evidence (CfE) for HDR and WCG video coding," *ISO/IEC JTC1/SC29/WG11 MPEG2014 N*, vol. 15083, 2015.

[10] R. K. Mantiuk, M. Kim, M. Ashraf, Q. Xu, M. R. Luo, J. Martinovic, and S. Wuerger, "Practical color contrast sensitivity functions for luminance levels up to $10\,000\,\mathrm{cd/m^2}$," in *Color Imaging Conference*, pp. 1–6(6).

[11] N. Y. Aliaksei Mikhailiuk and R. K. Mantiuk, "The effect of display brightness and viewing distance: a dataset for visually lossless image compression," in *Human Vision and Electronic Imaging*, 2021.

[12] T. Borer and A. Cotton, "A display-independent high dynamic range television system," *SMPTE Motion Imaging Journal*, vol. 125, no. 4, pp. 50–56, 2016.

[13] G. Ward-Larson, "LogLuv Encoding for Full-Gamut, High-Dynamic Range Images," *J. of Graphics Tools*, vol. 3, no. 1, pp. 15–31, 1998.

[14] R. Mantiuk, G. Krawczyk, K. Myszkowski, and H.-P. Seidel, "Perception-motivated high dynamic range video encoding," *ACM Transactions on Graphics*, vol. 23, no. 3, p. 733, 2004.

[15] S. J. Daly, "Visible differences predictor: an algorithm for the assessment of image fidelity," in *Human Vision, Visual Processing, and Digital Display III*, vol. 1666. International Society for Optics and Photonics, 1992, pp. 2–15.

[16] P. G. Barten, "Formula for the contrast sensitivity of the human eye," in *Image Quality and System Performance*, vol. 5294. International Society for Optics and Photonics, 2003, pp. 231–238.

[17] S. W. Wuerger, M. Ashraf, M. Kim, J. Martinovic, M. Pérez-Ortiz, and R. K. Mantiuk, "Spatio-chromatic contrast sensitivity under mesopic and photopic light levels," *J. Vision*, 2020.

[18] S. Daly, "Visible differences predictor: an algorithm for the assessment of image fidelity," in *Digital Images and Human Vision*. MIT Press, 1993, vol. 1666, pp. 179–206.

[19] M. Kim, M. Azimi, and R. K. Mantiuk, "Perceptually motivated model for predicting banding artefacts in high-dynamic range images," *Color and Imaging Conference*, 2020.

[20] G. Denes, G. Ash, and R. K. Mantiuk, "A visual model for predicting chromatic banding artifacts," in *Human Vision and Electronic Imaging*, 2019, pp. 212–1–212–8(8).

[21] R. K. Mantiuk, K. Myszkowski, and H.-p. Seidel, "High Dynamic Range Imaging," in *Wiley Encyclopedia of Electrical and Electronics Engineering*. Hoboken, NJ, USA: John Wiley & Sons, Inc., jun 2015, pp. 1–42. [Online]. Available: http://doi.wiley.com/10.1002/047134608X.W8265

[22] A. Mikhailiuk, M. Perez-Ortiz, D. Yue, W. Suen, and R. K. Mantiuk, "Consolidated dataset and metrics for high-dynamic-range image quality," *arXiv preprint arXiv:2012.10758*, 2020.

[23] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. Jay Kuo, "Image database TID2013: Peculiarities, results and perspectives," *Signal Processing: Image Communication*, vol. 30, pp. 57–77, jan 2015.

[24] H. Sheikh, M. Sabir, and A. Bovik, "A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3440–3451, 2006.

[25] P. Korshunov, P. Hanhart, T. Richter, A. Artusi, R. Mantiuk, and T. Ebrahimi, "Subjective quality assessment database of HDR images compressed with JPEG XT," in *International Workshop on Quality of Multimedia Experience*, 2015.

[26] M. Narwaria, M. P. Da Silva, P. Le Callet, and R. Pepion, "Tone mapping-based high-dynamic-range image compression: study of optimization criterion and perceptual quality," *Optical Engineering*, vol. 52, no. 10, p. 102008, oct 2013.

[27] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Transactions on Image processing*, vol. 23, no. 10, pp. 4270–4281, 2014.

[28] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.

[29] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.