https://doi.org/10.2352/lim.2024.5.1.12

Gamma Maps: Non-linear Gain Maps for HDR Reconstruction

Trevor D. Canham,[1] SaiKiran Tedla,[1] Michael J. Murdoch,[2] and Michael S. Brown[1]
[1] York University, Canada; [2] Rochester Institute of Technology, USA

Abstract

To accommodate displays with varying dynamic ranges, image encoding frameworks are emerging that propose to include metadata within a standard dynamic range (SDR) image to encode an arbitrary, user-defined residual which allows the SDR image's pixel values to be transformed into its intended high dynamic range (HDR) version. The suggested metadata is a compressed version of the gain map computed as the pixel-wise ratio between the HDR and SDR image. Multiplying the gain map with the SDR image reconstructs the HDR image. This paper proposes an effective alternative for HDR recovery in the form of a pixel-wise exponent map instead of the multiplicative gain map. We demonstrate experimentally that the exponent map approach produces higher quality HDR reconstructions over the gain map strategy according to several metrics.

Introduction

While most smartphone cameras can capture images with high dynamic range (HDR) tonal ranges, the final image is typically encoded in SDR formats like 8-bit sRGB (BT.709) [1] to ensure interoperability with displays and software applications. However, with the prevalence of HDR displays on smartphones and tablets, there is now a need to provide flexible encoding strategies that accommodate image content for different display dynamic ranges.

Leading technology developers and standards organizations are presently formulating encoding guidelines for adaptive dynamic range formats. Notable examples are Adobe's gain map specification [2], Apple's Extended Dynamic Range (EDR) [3], Android's UltraHDR [4], Samsung's SuperHDR [5], and the ISO/WD 21496 recommendation for digital photography [6].

These encoding strategies have all converged on a common framework for computing a pixel-wise gain map to enable the translation or interpolation between image variants intended for Standard Dynamic Range (SDR) and High Dynamic Range (HDR) displays. Specifically, these gain maps, computed as the ratio between SDR and HDR pixel values, are downsampled, encoded as an 8-bit JPEG file, and embedded as metadata within a multi-image format (such as HEIC, AVIF, or JPEG-R) alongside the primary 8-bit SDR image. This embedded gain map can be decompressed and upsampled to reconstruct the HDR rendition of the image through multiplication with the SDR image.

In this paper, we propose a subtle algebraic variation on the gain-map framework, which improves the resulting reconstructions quantitatively and qualitatively when tested on a dataset of native HDR images processed through a collection of tone-mapping operators. Instead of encoding a pixel-wise multiplicative residual (gain map), we propose to encode a pixel-wise exponential residual (gamma map). While a subtle change, we show that gamma maps give significantly better HDR recovery measured by three perceptual metrics than gain maps when encoded as JPEG metadata of the same size (~25KB). We demonstrate results on a dataset of 112 SDR/HDR pairs computed using several tone mapping operators.

Related Work

The human visual system uses gain and non-linear response within its photoreceptors to maintain visibility in the face of varying dynamic range [7]. Early imaging engineers designed systems that took advantage of this property of human vision to render realistic images in media with limited dynamic range [8]. Over the next century, these design elements were formalized computationally, culminating in the proposal of several effective methods. In earlier models, all pixels are processed through the same non-linear compression function, derived from the global dynamic ranges of the source scene and target display [9, 10]. Later, locally varying operators to mimic the independent response of photoreceptors were proposed [7].

Modern professional and consumer-grade cameras can now capture images with high dynamic tonal ranges [11]. However, HDR images need to be converted to SDR using tone mapping methods to maintain compatibility with SDR displays. While photographers and cinema colorists often prefer manual tone mapping, there remains a high demand for automatic methods. A recent review by Ou et al. [12] details the dominant tone-mapping strategies implemented in hardware. In the experimental section, we choose a foundational representative from three categories of tone mapping methods: exposure/filtering-based [7], log-based [13], and histogram equalization-based [14]. While the gain map framework is intended to be agnostic to the tone mapping operator, the operator used does impact the distribution of the gain map, which creates a differential in the difficulty of compression.

Our work is also related to methods that aim to up-covert legacy SDR images to HDR. Mantiuk et al. [16], Cyriac et al. [17], and several industry groups (such as Dolby, Colorfront, HDR10) offer analytical solutions for converting between SDR to HDR. These methods perform well where quantization, compression, and clipping errors are minimal but are not capable of recovering missing HDR details. While these methods target a similar purpose of SDR to HDR conversion, they are forced to do this blindly. The problem addressed in this paper assumes that HDR and SDR versions are already defined at encoding time and stores metadata in the SDR file such that the HDR image can be reconstructed.

Method

We first describe the existing gain map framework and then our modification in the following section.

Existing Framework

Figure 1 provides an overview of the adaptive dynamic range encoding framework. Given an HDR scene captured by camera sensors typically using 10-12 bits, the image is tone mapped to a standard dynamic range and quantized to 8 bits (N-bit quantization is denoted as $Q_N(\cdot)$), and stored as a JPEG file for general interchange. The framework will compute a pixel-wise gain map and embed this with the primary JPEG as metadata.

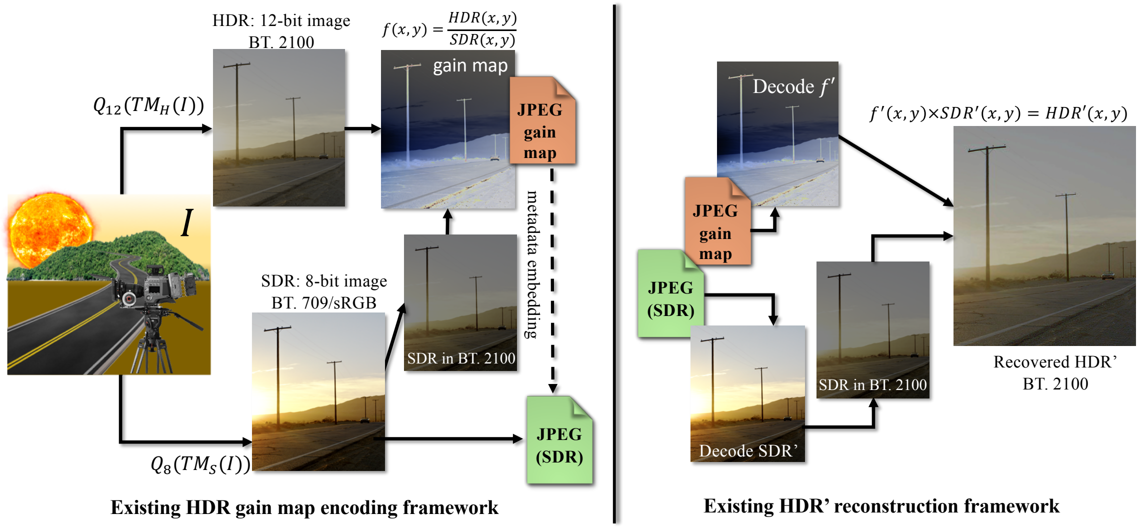

Figure 1. *An image is captured and different renderings for HDR and SDR are derived through quantization and tone-mapping. Since 8-bit SDR images are more interoperable, it is advantageous to encode this as a JPEG along with metadata from which the HDR version can be reconstructed. This is accomplished by converting the SDR image into the HDR display representation and taking SDR to HDR ratio to produce a gain map. Then, once normalized and encoded with JPEG compression, this can be embedded alongside the SDR image in a multi-image format. Finally, the HDR version can be reconstructed after the fact by decoding and de-normalizing the gain map and multiplying it by the SDR image.*

Adobe's [2] preliminary version of the under-development ISO gain map standard (ISO/WD 21496 [6]) begins with three-channel color images S and H, encoded for SDR and HDR display, respectively. First, the SDR image is transformed into the display space of the HDR image as in Figure 1, such that they share a common representation. The gain map is computed as

$$f(x,y) = \frac{(H+\varepsilon)}{(S+\varepsilon)}, \tag{1}$$

where ε is a small offset applied to each image to avoid divide-by-zero errors.

Then, $f(x,y)$ is normalized to range $[0,1]$ for JPEG-encoding. This is accomplished by computing the maximum and minimum gain map values, $f_{\min}$ and $f_{\max}$, and then normalizing the gain map in log_2 space like so:

$$f_{\text{norm}}(x,y) = \frac{(log_2(f(x,y)) - log_2(f_{\min}))}{(log_2(f_{\max}) - log_2(f_{\min}))}. \tag{2}$$

Finally, a power γ is applied after the normalization process to redistribute gain map values for quantization. Both Android [4] and Adobe [2] documents recommend resizing the map.

Specifically, these documents recommend quantizing the gain map to 8-bit precision, down-sampling to $1/4$ resolution or lower, and encoding with JPEG compression at a quality setting from 85 to 90 out of 100 [4]. Using a multi-picture format (e.g. HEIC, AVIF, JPEG-XL, JPEG-R) the map can be stored alongside S. The parameters $log_2(f_{\max})$ and $log_2(f_{\min}))$ values, ε, and γ are then stored as XMP metadata. This strategy allows for legacy 8-bit applications to process S as normal, while applications that have implemented gain map decoding can reconstruct its corresponding HDR representation.

However, the decoding process returns a slightly different function $f'_{\text{norm}}(x,y)$, due to quantization, downsampling and compression. This process starts by converting the 8-bit gain map back to a floating point representation in the range $[0,1]$ and the normalization is inverted as follows:

$$f'(x,y) = 2^{(f_{\text{norm}}(x,y))*(log_2(f_{\max})-log_2(f_{\min}))+log_2(f_{\min})}. \tag{3}$$

Finally, the SDR image is converted back to a floating point representation and then to the HDR display space (as above) and the HDR rendition, H' is reconstructed:

$$H' = (S+\varepsilon) \odot f'(x,y) - \varepsilon. \tag{4}$$

Proposed Variation

Instead of finding a gain map as in equation 1 of the existing framework, we advocate the use of a pixel-wise exponent such that $(H+\varepsilon) = (S+\varepsilon)^{g(x,y)}$. We can solve for $g(x,y)$ as follows:

$$g(x,y) = log_{S+\varepsilon}(H+\varepsilon), \tag{5}$$

or alternatively,

$$g(x,y) = \frac{log(H+\varepsilon)}{log(S+\varepsilon)}. \tag{6}$$

After encoding, we apply compression and decoding as in the existing framework (equations 2 and 3). The resulting gamma map $g'(x,y)$ is applied as a pixel-wise exponent in place of equation 4 of the existing framework.

The final recovered HDR image, H' is given by:

$$H' = (S+\varepsilon)^{g'(x,y)} - \varepsilon. \tag{7}$$

Experiments

This section tests the proposed gamma map variation against the existing gain map framework.

Dataset

In the work of Cyriac et al. [17], a dataset of 28 HD (1920×1080) SDR and HDR mastered images provided by Froehlich et al. [21] and the ARRI camera group [22] (encoded in BT.709 [1] and BT. 2100 [23], respectively) were used as references to optimize the method. Starting from their native format (HDR or SDR), a professional colorist derived an ideal visual match in the alternative format using DaVinci Resolve as

Table 1. Our proposed gamma map is compared against the standard gain maps for approximating manual and automatic tone mapping transforms representative of the most prominent approaches implemented in hardware (log [13], filtering [7], histogram equalization [14]) using Cyriac et al.'s dataset of cinema images as a source [17]. Best results in bold.

TM	Method	HDR-VDP3 [26] ↓	Mean Δ_{E00} ↓	PQ VIF [27] ↑	Size (KB) ↓
Manual	Gain	9.85	2.01	0.75	**22**
	Gamma	**8.39**	**1.63**	**0.76**	25
Drago et al. [13]	Gain	9.17	1.89	0.61	**25**
	Gamma	**8.00**	**1.67**	**0.68**	**25**
Reinhard et al. [7]	Gain	7.91	2.02	**0.81**	29
	Gamma	**7.68**	**1.62**	0.78	**20**
Larson et al. [14]	Gain	9.72	1.97	0.77	**22**
	Gamma	**8.10**	**1.60**	**0.81**	24
Average	Gain	9.16	1.97	0.73	25
	Gamma	**8.04**	**1.63**	**0.75**	**24**

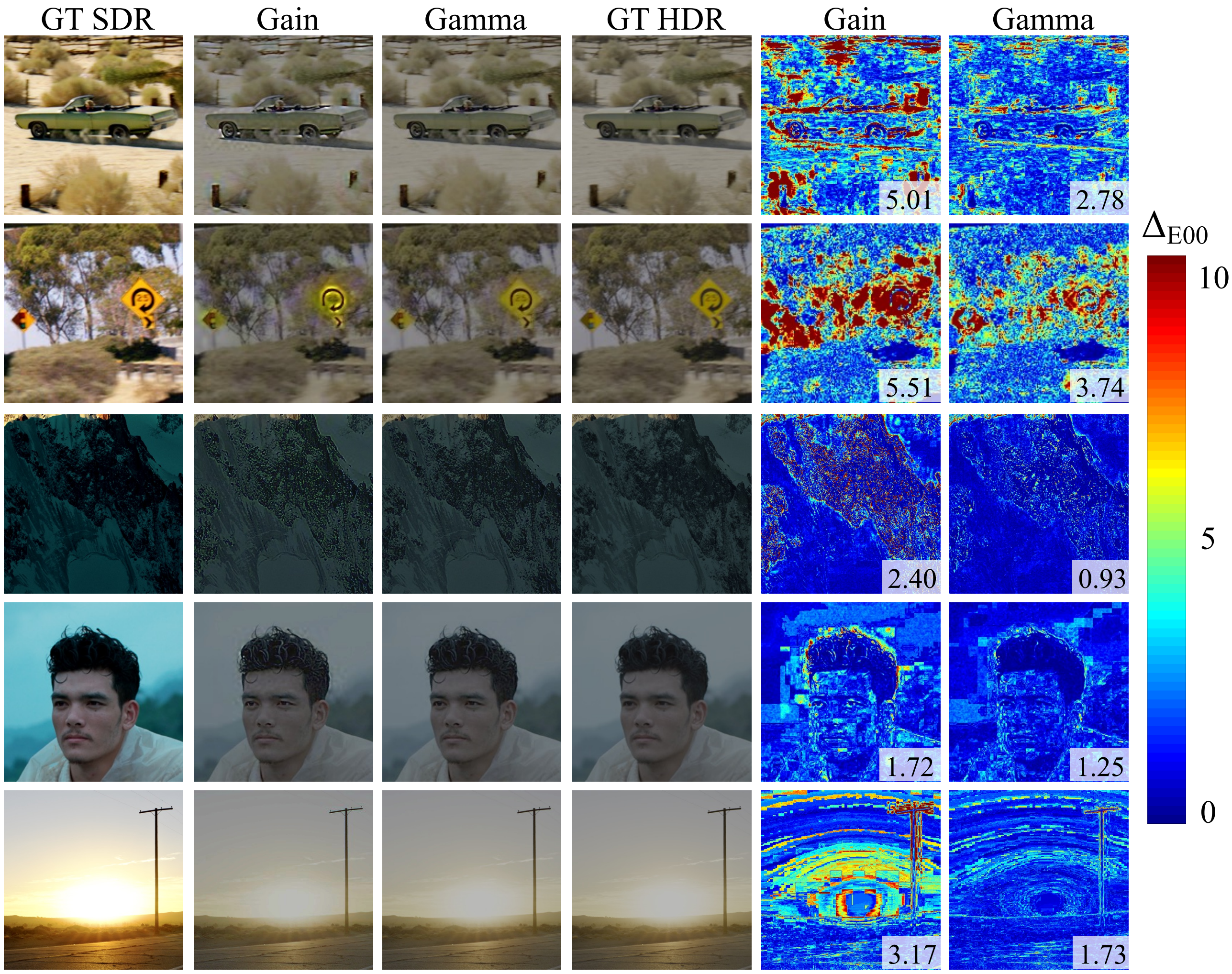

Figure 2. *Qualitative comparison between the proposed gamma maps and the existing gain map framework on cropped regions from the dataset of Cyriac et al. [17] with manually tone-mapped SDR images. PDF assumes all images are in SDR/sRGB format. Thus, the HDR images appear desaturated. Still, the gain map results exhibit spurious colors, compression artifacts, and aberrations in the standard approach, while the gamma map alleviates these artifacts. Especially noticeable are the overshot highlights visible at high-contrast edges in the Gain images.* Δ_{E00} *heat maps are included to aid in localizing artifacts, where the maximum value is set to ten* Δ_{E00}*. The mean* Δ_{E00} *for each map is shown in the bottom right corner.*

a color grading platform. To accomplish this task, the colorist employed global non-linear mappings and local contrast adjustments via image segmentation. A Sony BVM-X300 with HDR and SDR calibration profiles was employed as a reference monitor. The HDR profile was set to a peak white luminance of 1,000 cd/m^2 and ST. 2084 [24] decoding function, while the SDR profile was calibrated to BT.709 (100 cd/m^2, 2.4 gamma). The colorist switched profiles to properly compare the two renderings. In this way, the SDR/HDR image pairs were derived with an authentic cinema workflow.

To account for alternative mappings between SDR and HDR, the HDR images were processed by automatic tone mapping methods representative of the three major categories from the review of Ou et al. [12]. To do this, HDR images were transformed to a display linear representation with BT.709 primaries (without clipping or quantization) such that they were out of range and required tone mapping. Then, the methods of Drago et al. [13] (log-based), Reinhard et al. [7] (exposure-based) and Larson et al. [14] (histogram equalization) were employed via their MATLAB HDR toolbox implementation from Banterle et al. [25] with default parameters.

Comparisons

The existing gain and proposed gamma map frameworks were implemented in MATLAB as described above. Bicubic resizing was employed to obtain the $1/4$ resolution version, and the gamma parameter was not used. JPEG gain map images were stored with the quality parameter set to 90 out of 100.

We employ an HDR native metric (HDR-VDP3 [26]) which has been designed to predict visual sensitivity to image differences in HDR luminance regimes via psychophysical experimentation. To account for color differences, we employ the CIE Delta E 2000 color difference metric (Δ_{E00}) on a pixel-wise basis. Following the recommendation from Sugito et al. [27] on HDR metrics, we also apply the Visual Information Fidelity (VIF) metric to reconstructed HDR signals encoded as described in the previous section. This work demonstrated that applying this standard SDR metric to HDR encoded images achieves a higher correlation to observer opinion scores of block-based compression artifacts than native HDR metrics.

Quantitative Results

Table 1 shows the comparison of the proposed gamma map variation versus the standard gain map framework for encoding a variety of tone mapping differences. The gamma map framework shows consistent improvements with respect to HDR-VDP3, mean Δ_{E00}, and PQ VIF while maintaining the resulting metadata size.

Qualitative Results

Figure 2 compares the proposed and standard approaches for several images from the dataset tested in Table 1. It can be observed that the spurious JPEG artifacts are alleviated by the gamma map framework in a number of examples. Pixel-wise differences are visualized in Δ_{E00} heat maps, where the maximum value is set to ten Δ_{E00}.

Discussion

Using the simple algebraic variation of gamma maps, we show in Table 1 that we outperform the existing gain map framework with the same metadata memory footprint. The qualitative results of Figure 2 further emphasize the benefits of the proposed variation which alleviates acute JPEG artifacts.

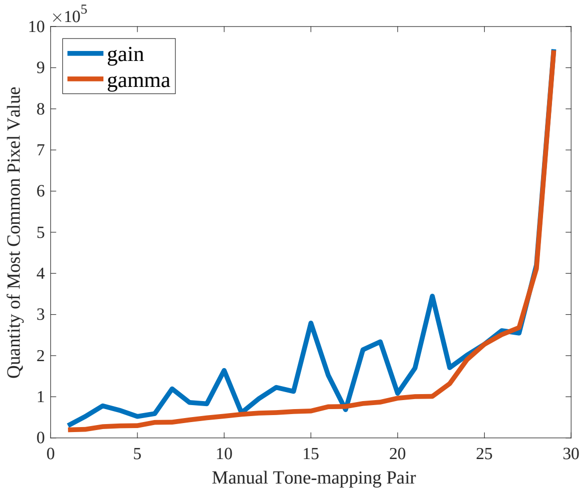

Figure 3. Compression artifacts are alleviated in gamma maps because their values are better distributed than those of gain maps. This is demonstrated by plotting their maximum histogram values (the quantity of the most common pixel) for each manual tone-mapping pair in ascending order of gamma map value.

Analysis of the residuals showed that gain maps resulted in a distribution of values that were more likely to cluster tightly around the max value, which is a less efficient use of the quantization space. Gamma maps, on the other hand, resulted in distributions which occupied a wider range with a lower max value. This is demonstrated in Figure 3, which shows the maximum histogram value $max(max(r), max(g), max(b))$ for gain and gamma maps of each image pair in the manual dataset in ascending order across the x-axis.

There is a relevant distinction between the logarithmic encoding of the proposed variation and that of the existing framework. While our proposed variation works in the image intensity domain, the existing framework's logarithmic and exponential processing steps occur during normalization in the domain of pixel-wise gain. Here, small values are not necessarily correlated to low intensity, losing the perceptual alignment of the proposed variation. The result is that variations in the non-linear encoding parameters of the existing framework trade quality for metadata size, while we improve quality and maintain metadata size.

Conclusion

The current display ecosystem needs a dynamic encoding scheme to facilitate transformations between displays with varying reproduction capabilities, that are viewed in varying environments with varying brightness settings. Specifically, a standard for encoding arbitrary local mapping transforms in the form of a gain map of pixel-wise linear scalar values between images rendered for SDR and HDR display has been proposed by major technology developers. While effective for encoding the transforms in a manner which is not restrictive to mastering artists, the suggested technique is heavy and prone to artifacts related to its compression. We propose a variation to this standard where pixel-wise exponents are calculated. We demonstrate through qualitative and quantitative experiments that this encoding improves results by alleviating compression artifacts. For future work we are interested in conducting psychophysical experimentation into the ideal intermediate representation between dynamic range renderings.

Acknowledgments

The authors would like to acknowledge Quique Cadañas for generating the manually tone mapped dataset.

References

[1] Recommendation ITU-R BT.709, Parameter values for the HDTV standards for production and international programme exchange (1993)
[2] E. Chan, Adobe gain map specification v1.0 (2023) online
[3] K. Greenebaum, High Dynamic Range rendering with EDR (2020) online
[4] Android, Ultra-hdr image format v1.0 (2023) online
[5] Samsung, Galaxy S24 Ultra (2024) online
[6] ISO/WD 21496-1, Gain map metadata for image conversion (Under development)
[7] E. Reinhard, K. Devlin, Dynamic range reduction inspired by photoreceptor physiology, IEEE TVCG 11(1), 13–24 (2005)
[8] E.J. Giorgianni and T.E. Madden, Digital Color Management: Encoding Solutions, vol. 13. Wiley (2009)
[9] J. Tumblin, H. Rushmeier, Tone reproduction for realistic images, IEEE CGA 13(6), 42–48 (1993)
[10] E. Reinhard, M. Stark, P. Shirley, J. Ferwerda, Photographic tone reproduction for digital images, ACM TOG 21(3), 267–276 (2002)
[11] S. Liu, X. Zhang, L. Sun, Z. Liang, H. Zeng, L. Zhang, Joint HDR Denoising and Fusion: A Real-World Mobile HDR Image Dataset, CVPR (2023)
[12] Y. Ou, P. Ambalathankandy, S. Takamaeda, M. Motomura, T. Asai, M. Ikebe, Real-time tone mapping: A survey and cross-implementation hardware benchmark, IEEE TCSVT 32(5), 2666–2686 (2022)
[13] F. Drago, K. Myszkowski, T. Annen, N. Chiba, Adaptive logarithmic mapping for displaying high contrast scenes, Computer graphics forum. vol. 22, 419–426 (2003)
[14] G.W. Larson, H. Rushmeier, C. Piatko, A visibility matching tone reproduction operator for high dynamic range scenes, IEEE TVCG 3(4), 291–306 (1997)
[15] M. Ashikhmin, J. Goyal, A reality check for tone-mapping operators, ACM TAP 3(4), 399–411 (2006)
[16] R. Mantiuk, S. Daly, L. Kerofsky, Display adaptive tone mapping, SIGGRAPH, pp. 1–10 (2008)
[17] P. Cyriac, T. Canham, D. Kane, M. Bertalmío, Vision models fine-tuned by cinema professionals for high dynamic range imaging in movies, Multimedia Tools and Applications 80, 2537–2563 (2021)
[18] S.Y. Kim, J. Oh, M. Kim, Deep sr-itm: Joint learning of super-resolution and inverse tone-mapping for 4K UHD HDR applications, CVPR (2019)
[19] S.Y. Kim, J. Oh, M. Kim, Jsi-gan: Gan-based joint super-resolution and inverse tone-mapping with pixel-wise task-specific filters for uhd hdr video. AAAI (2020)
[20] X. Chen, Z. Zhang, J.S. Ren, L. Tian, Y. Qiao, C. Dong, A new journey from sdrtv to hdrtv, CVPR (2021)
[21] J. Froehlich, S. Grandinetti, B. Eberhardt, S. Walter, A. Schilling, H. Brendel, Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays, Electronic Imaging (2014)
[22] ARRI Camera Group, Enhanced capture material for hdr4eu project d2.2
[23] Recommendation ITU-R BT. 2100, Image parameter values for high dynamic range television for use in production and international programme exchange (2016)
[24] SMPTE ST. 2084, High dynamic range electro-optical transfer function of mastering reference displays (2014)
[25] F. Banterle, A. Artusi, K. Debattista, A. Chalmers, Advanced High Dynamic Range Imaging (2nd Edition), CRC press (July 2017)
[26] R.K. Mantiuk, D. Hammou, P. Hanji, HDR-VDP-3: A multi-metric for predicting image differences, quality and contrast distortions in high dynamic range and regular content. arXiv (2023)
[27] Y. Sugito, J. Vazquez-Corral, T. Canham, M. Bertalmío, Image Quality Evaluation in Professional HDR/WCG Production Questions the Need for HDR Metrics, IEEE TIP (31), 5163-5177 (2022)

Author Biography

__Trevor Canham__ is studying color imaging under the supervision of Michael Brown at York University in Toronto. He received the BSc in Motion Picture Science from the Rochester Institute of Technology, and spent several years working in Marcelo Bertalmío's Image Processing for Enhanced Cinematography lab in Barcelona, Spain. His interests lie in the interaction between color phenomenology and imaging systems. He was recently awarded best student paper at the 31st Color & Imaging Conference and the Color Research Society of Canada's graduate student award.

__SaiKiran (Sai) Tedla__ is a PhD student studying camera pipelines at York University. Sai received his Masters from York University and undergraduate from University of Colorado Colorado Springs. He has previously interned at Samsung AI Center and is a current intern at Adobe. His interests are varying and he was worked on exposure, color correction, focus, compression, and inverse pipeline problems.

__Michael J. Murdoch__ is an Associate Professor and Director of the Munsell Color Science Laboratory at the Rochester Institute of Technology, where his research focuses on color in advanced displays and LED lighting. He holds a BS in Chemical Engineering from Cornell, MS in Computer Science from RIT, and PhD in Human-Technology Interaction from Eindhoven University of Technology.

__Michael S. Brown__ is a professor and Canada Research Chair at York University. His research focuses on in-camera color processing algorithms for photographic and scientific applications.