A DUAL MODULATION ALGORITHM FOR ACCURATE REPRODUCTION OF HIGH DYNAMIC RANGE VIDEO

Emin Zerman, Giuseppe Valenzise, Frederic Dufaux

LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, 75013 Paris, France

ABSTRACT

In this paper, we consider the problem of accurately reproducing high dynamic range (HDR) images and video on HDR displays. Differently from conventional image reproduction devices, HDR displays require dual modulated signals to drive the back LED panel and the front LCD screen. In this work, we present an iterative scaling approach to estimate these two signals, which is able to reproduce accurately luminance values stored in HDR files. To avoid temporal artifacts due to frame-by-frame processing, we also describe a temporal smoothing mechanism to compute backlight illumination. Our results demonstrate higher accuracy than simpler approaches such as built-in display rendering. Furthermore, the proposed method offers the possibility to estimate precisely the luminance emitted by the display.

Index Terms— High dynamic range, rendering, dual modulation, luminance estimation, temporal variation.

1. INTRODUCTION

High dynamic range (HDR) displays enable to reproduce a much broader span of brightness, contrast and colors than conventional low dynamic range (LDR) screens. As a result, they are expected to rapidly spread in the television market as a means to provide more immersive and realistic multimedia services. A popular paradigm for displaying HDR content consists in coupling a locally dimmed light source, such as a panel of LED's, with a front LCD screen. This allows to achieve high peak brightness values, while keeping the black level of the display very low [1]. However, the reproduction of an HDR image in this framework entails significantly more processing than in conventional LDR devices, as both LED and LCD driving signals need to be computed. Given an HDR picture, the problem of estimating the corresponding LED/LCD panel values is known as *dual modulation* [2].

Because of current technology limitations of HDR displays, dual modulation requires a global optimization approach, since the overall rendering of an HDR picture might change due to perturbations in the value of a single LED and can be influenced by the rendering of previous video frames. In practice, many dual modulation algorithms relax this globality constraint and act locally, trading reproduction accuracy for computational complexity. For example, the rendering algorithms built into HDR displays generally give up peak brightness and dynamic range in order to render HDR video in real time, at high frame rates. Indeed, built-in rendering is often a common choice in many applications, e.g., it has been used in the subjective evaluation of HDR compression performance [3]. Nevertheless, in some psycho-visual experiments it could be desirable to reproduce as accurately as possible the luminance levels stored in the HDR content, or at least to know with a sufficient precision the actual per pixel luminance emitted by the display. The accurate reproduction and estimation of luminance values, both for HDR images and videos, constitute the goal and the contribution of this work.

In this paper, we propose a dual modulation algorithm for HDR image and video content, which has the following three characteristics: *i)* it can accurately reproduce HDR luminance; *ii)* it enables to estimate precisely and with pixel granularity the luminance emitted by an image/video displayed using the proposed method; *iii)* it takes into account temporal dependencies in HDR video, reducing the impact of reproduction artifacts such as flickering. We have tested the proposed algorithm on a SIM2 HDR47-4K display [4], comparing it with the built-in rendering provided by the manufacturer. We show that our method is systematically more precise in reproducing HDR content, and that we can estimate accurately the emitted luminance, which is unfeasible with the built-in rendering. At the same time, our results on HDR video are encouraging, showing that temporal fluctuations can be substantially reduced by smoothing LED values across time.

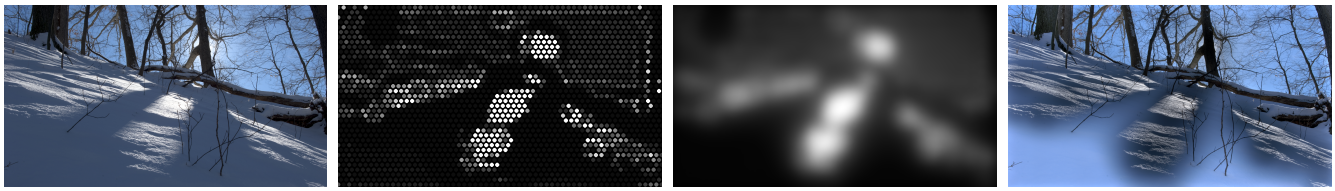
The rest of the paper is structured as follows: Section 2 discusses related work; the details of the proposed algorithm are given in Section 3; the accuracy of the proposed algorithm is discussed in Section 4; finally, Section 5 concludes the paper.

2. RELATED WORK

Several dual-modulation algorithms for LED/LCD displays have been proposed in the literature [5–12]. Most of these works solve the dual modulation problem using some kind of approximation, e.g., they find LED backlight illumination by taking the maximum, average, or weighted average of pixel values [5, 8], or use local block-based approaches [6, 7]. LCD values are generally obtained after the computation of the LED values, by dividing the HDR image luminance by the backlight. Dual modulation can be formulated more rigorously as an optimization problem [9, 10, 12]. However, these approaches have been mainly targeting low-resolution backlight panels, with tens of LED's, in the context of low-consumption, locally dimmed LCD displays. Current HDR technology, instead, requires dealing with thousands of LED's, as well as with their large point spread function.

Rendering HDR video requires additional care compared to HDR image, as frame-by-frame rendering might lead to temporal flickering due to high-frequency changes in the backlight. Even though it is possible to reduce the flickering using post-processing [13], it is preferable to directly handle it during rendering. Burini et al. [11] considered temporal variation in video sequences and, in order to reduce the flickering effect, implemented an infinite impulse response (IIR) filter integrated to their dual modulation

The work presented in this document was supported by BPIFrance and Région Ile de France, in the framework of the FUI 18 Plein Phare project.



(a) HDR image (Tonemapped) (b) LED values (c) Backlight (d) LCD values

Fig. 1. Algorithm results for HDR image named *DevilsBathtub*

method. They propose a block-based gradient descent algorithm, and try to jointly minimize the reproduction error and the power consumption required by the LED's. However, their work was effective for an LCD display with only 16 LED's, but it is computationally too demanding to be extended to configurations with thousands of LED's.

An important feature of dual modulation algorithms is the ability to estimate the actual luminance emitted by the display. In our previous work [14], we have proposed a dual modulation method able to provide an accurate estimation of the real displayed luminance. In that work, we focused on the impact this knowledge may have in the computation of HDR image fidelity metrics. In this paper, we extend the rendering method presented in [14] to the case of video, and we focus on the performance in terms of reproduction and luminance estimation precision of the proposed method.

3. RENDERING AND LUMINANCE ESTIMATION

In this section we describe the proposed dual modulation algorithm for rendering HDR images and video. We design and test our algorithm using as target the SIM2 HDR47-4K display, which has a measured peak luminance of about 4250 cd/m^2, and is composed by a back panel of 2202 independently controllable LED's and a 1920×1080 pixels LCD panel. The display can be driven using the automatic built-in rendering, or through a custom dual modulation input provided by users. Therefore, we use the latter mode to transmit dual-modulated HDR pictures and video to the display for rendering. We notice that the proposed approach is still valid with other display models, provided that some parameters of the device are known or previously measured. Notably, these include the maximum power consumption of the display, as well as an estimation of LED point spread function (PSF).

3.1. HDR Image Rendering

The proposed dual modulation rendering of an HDR picture is based on the following steps:

1. Preprocessing: the HDR image is scaled in order to have display-referred values, i.e., the target desired luminance in the limits of the HDR display physical capabilities. For instance, values above the maximum displayable brightness are clipped.

2. Finding target backlight: We next find the target optimal backlight, BL_{target}, that will ensure both minimum power consumption and maximum fidelity to the original luminance values. First, we compute local maxima of the target luminance over 30-pixel radius windows (corresponding to the area of a LED). Then, the maximum luminance values allowed for each pixel are found by dividing the target luminance of that pixel by the estimated LCD leakage factor $\epsilon = 0.005$. The resulting image is median filtered to avoid spurious peaks, due to single bright spots or noise. After the computation of the target backlight, the LEDs and backlight are initialized by sampling BL_{target} on LED locations and taking the convolution with the Point Spread Function (PSF), respectively. That is,:

$$BL_t = LED_t * PSF \tag{1}$$

where t is the iteration number and $t = 0$ for initialization.

3. Iterative scaling: The LED values are iteratively scaled as follows:

$$LED_{t+1} = LED_t \times \left(\frac{BL_{target}}{BL_t} \right). \tag{2}$$

LED_{t+1} is then clipped to take values in $[0, 1]$, i.e., it is projected onto the set of feasible LED's values. The operations in (1) and (2) are carried out consecutively by increasing the iteration number until $\sum ||PU(BL_t) - PU(BL_{t-1})||^2$ falls below a threshold. Taking perceptually uniform (PU) [15] encoded backlight makes the computation of the cost function perceptually meaningful and speeds up convergence. When the iterative scaling converges, the resulting LED_{final} values are possibly further scaled to meet the power constraints of the display.

4. Finding LCD values: LCD pixel values are found by dividing (pixel-wise) each color channel of the original image by the final backlight estimate, and by applying gamma correction, i.e.:

$$LCD_k = \left(\frac{I_k}{LED_{final} * PSF} \right)^{1/\gamma_k} \tag{3}$$

where I is the HDR image, $k \in \{R, G, B\}$ is the RGB channel indicator, and γ_k is the gamma correction factor, determined experimentally for each channel.

Our Matlab implementation of this algorithm takes an average of 19 seconds (corresponding to about 24 iterations) on an Intel i7-3630QM 2.40 GHz 8 GB RAM PC for rendering a 1920×1080 pixels image. However, a faster and more optimized implementation is possible and is part of our current work. As an example, the final LED values, backlight, and LCD image of the *DevilsBathtub* content from Fairchild's HDR dataset [16] are presented in Figure 1.

3.1.1. Estimation of emitted luminance

Knowing the values of LED's and LCD panel, the HDR image pixels produced by the display are the product of backlight and LCD values. That is, for each color channel k, the rendered image I'_k is:

$$I'_k = (LED_{final} * PSF) \cdot LCD_k. \tag{4}$$

The emitted luminance is then computed as:

$$L = 0.2126 \cdot I'_{Red} + 0.7152 \cdot I'_{Green} + 0.0722 \cdot I'_{Blue}, \tag{5}$$

assuming ITU-R BT.709 primaries [17].

3.2. HDR Video Rendering

The HDR dual modulation algorithm described in the previous section provides accurate HDR reproduction. However, small perturbations in the original HDR pixel values may lead to overall changes in the produced backlight. As a result, this can produce temporal flickering when rendering HDR video. In order to reduce the impact of flickering, we consider two solutions. First, we initialize the LED values for the current frame f using those of the previous frame, i.e., $LED_0^f = LED_{final}^{f-1}$. Second, we smooth the target backlights across time, over consecutive overlapping windows, as described in the following.

Given a video frame f, its initial target backlight BL_{target}^f is computed as explained in Section 3.1.Then, for each frame, we consider a look-ahead window of N frames, and we arrange their corresponding backlights in a stack A^f, i.e.:

$$A^f = \left[BL_{target}^f BL_{target}^{f+1} ... BL_{target}^{f+N-1}\right]. \quad (6)$$

Afterwards, we aim at smoothing the trajectory of backlight pixel values over the window, by computing their upper envelope. To this end, we extract the backlight pixel signal across time for each pixel location (i, j), i.e., we obtain the N-dimensional column vector $A_{j,k}^f$. In order to compute the envelope of this signal, one cannot employ simply a low-pass filter, as averaging may produce lower target backlight than necessary, thus reducing peak brightness and reproduction fidelity. Instead, we adopt a simple approach, that consists in convolving each sample independently by a Gaussian window, and taking the maximum at each time instant.

More precisely, let S_l be a $N \times N$ matrix such that $S_f(a, b) = 1$ if $a = b = l$ and 0 otherwise. Multiplying $A_{j,k}^f$ by S_l yields:

$$T_{j,k,l}^f = [0 \dots A_{j,k}^f(l) \dots 0]^T, \quad (7)$$

i.e., a vector with all zeros but the l-th element, which is the l-th entry of $A_{j,k}^f$. Now, let $W_{j,k,l}^f = T_{j,k,l}^f * w_\sigma$ be a low-pass version of $T_{j,k,l}^f$ obtained by convolution with a Gaussian smoothing filter w_σ of variance σ^2. The envelope signal $M_{j,k}^f$ is then obtained by stacking the vectors $\{W_{j,k,l}^f\}$ for $l = 1, \dots N$ into an $N \times N$ matrix

$$B_{i,j}^f = [W_{j,k,1}^f \dots W_{j,k,N}^f] \quad (8)$$

and taking the maximum across columns.

The procedure described through equations (6) - (8) is repeated using a sliding window approach, i.e., the backlight target is updated as $BL_{target}^f = M^f$, and the frame index f is increased by one. An example of filtered target backlight for a given pixel position of the *ChristmassTree* [18, 19] video sequence is shown in Figure 2. Once the smoothed target backlight has been computed, the rest of the rendering part follows the algorithm described in Section 3.1.

4. EXPERIMENTAL VALIDATION

In order to assess the performance of the proposed dual modulation algorithm, we have matched the target luminance stored in the HDR files with the emitted and estimated luminance of the display. Differently from previous work [14], we evaluate the performance on complex images rather than simple test patterns. In the following we refer to these three quantities, which are all expressed in cd/m^2:

- *Expected luminance*, corresponding to the display-referred luminance stored in the HDR file. Values higher than display peak brightness are clipped.

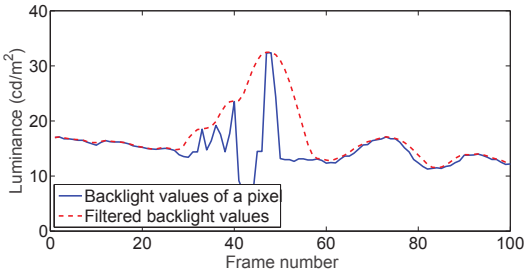
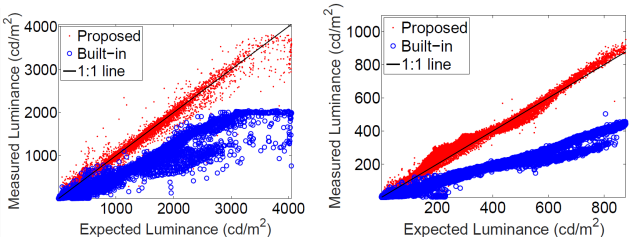


Fig. 2. Example of temporal smoothing of backlight pixel trajectories for the *ChristmassTree* video sequence.

- *Estimated luminance*, i.e., pixel-wise luminance values estimated as in (5).
- *Measured luminance*, obtained using a Canon EOS700D DSLR camera to acquire an HDR picture of the display. The generated images have been photometrically calibrated using a Minolta LS-100 luminance meter to store the real emitted luminance. This operation has been done both for built-in rendering and for the proposed dual modulation algorithm.

4.1. Fidelity of Reproduction

To gauge the fidelity of HDR reproduction, the measured luminance values are plotted against expected luminance values for 7 different HDR contents from Fairchild database [16]; namely, *AirBellowsGap*, *DevilsBathtub*, *HancockKitchenOutside*, *MasonLake(1)*, *PaulBunyan*, *RedwoodSunset*, and *UpheavalDome*. Example plots can be seen in Fig. 3. Notice that the built-in rendering method provides lower luminance and higher scatter compared to the proposed rendering algorithm. The fact that the luminance emitted with the proposed algorithm matches the expected one (points clustered on the 45° line) demonstrates the higher accuracy of our rendering. In order to have a more quantitative evaluation, Pearson Correlation Coefficient (PCC) and Root Mean Squared Error (RMSE) indexes, measuring the linear dependence between two variables and the variance of estimates, respectively, have been computed and reported in Table 1. As it can be seen, the results of the proposed algorithm are more precise than built-in rendering for each content.



(a) AirBellowsGap (b) RedwoodSunset

Fig. 3. Plots of measured luminance vs. expected luminance

4.2. Fidelity of Estimation

In order to assess the emitted luminance estimation accuracy, the measured luminance values have also been plotted against the estimated luminance values. We report as an example, the scatter plot

Table 1. Correlation results for luminance measurement for expected luminance and measured luminance.

	Built-in		Proposed	
	PCC	RMSE	PCC	RMSE
AirBellowsGap	0.9807	311.16	0.9924	65.61
DevilsBathtub	0.8692	96.17	0.9089	59.61
HancockKitchenOutside	0.9400	103.88	0.9572	36.89
MasonLake(1)	0.9159	188.41	0.9312	76.10
PaulBunyan	0.9633	143.73	0.9703	37.83
RedwoodSunset	0.9933	107.69	0.9936	20.83
UpheavalDome	0.9782	142.11	0.9798	37.17

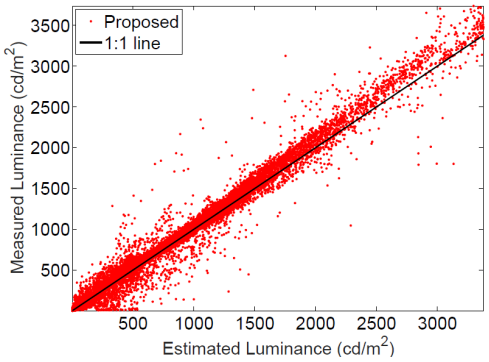


Fig. 4. Plots of measured luminance vs. estimated luminance. These results are presented only for proposed rendering.

for the *AirBellowGap* image in Fig. 4, and quantitative results using the same indexes as in Section **??** in Table 2. As the correlation scores and example figure indicate, the estimated luminance values are very close to 45° line and provide very high correlation rates.

4.3. Temporal variation

As discussed above, small high-frequency variations in the backlight across time may produce flickering in the displayed HDR video. In order to measure temporal variation in the backlight produced our rendering method, we compute the temporal perceptual information index (TI) defined in ITU-T Recommendation P.910 [20] on the backlight images. TI is computed as:

$$TI = \max_f \big(\mathrm{std}(BL_f - BL_{f-1})\big), \tag{9}$$

where std denotes standard deviation computed over space, and f denotes frame number.

Table 2. Correlation results for luminance measurement for estimated luminance and measured luminance, using the proposed rendering method.

	PCC	RMSE
AirBellowsGap	0.9937	62.83
DevilsBathtub	0.9089	59.61
HancockKitchenOutside	0.9575	36.40
MasonLake(1)	0.9312	76.10
PaulBunyan	0.9695	38.29
RedwoodSunset	0.9947	23.86
UpheavalDome	0.9815	40.00

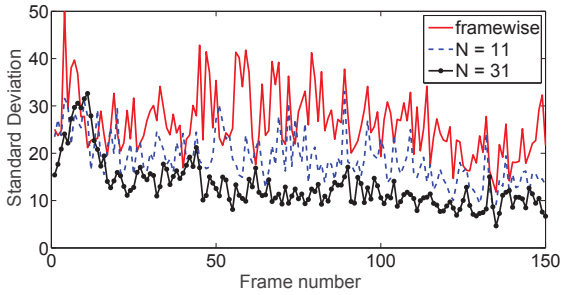


Fig. 5. Standard deviation change with respect to frame number for *Market3* sequence

Table 3. TI for different video contents. (*BL*: Balloon, *CT*: ChristmassTree, *FE*: FireEater2, *MK*: Market3, *TB*: Tibul2)

Rendering	BL	CT	FE	MK	TB
framewise	30.68	184.79	85.37	49.92	142.61
$N = 1$	27.32	184.28	63.38	49.18	127.43
$N = 11$	27.00	115.13	67.41	33.05	76.29
$N = 21$	18.44	74.20	48.65	35.72	54.07
$N = 31$	18.92	44.51	39.66	32.63	55.53

We compute TI for 5 different video sequences, namely *Balloon*, *FireEater2*, *Market3*, *Tibul2* [21], and *ChristmassTree* [18,19], rendered using both the frame-by-frame algorithm described in Section **??** and the video algorithm described in Section **??**. For the latter, we consider different windows size $N = 1, 11, 21, 31$, where $N = 1$ corresponds to frame-by-frame processing with the only difference that LED's are initialized using previous LED values. The frame-wise standard deviation for *Market3* sequence is illustrated in Fig. 5. As expected, increasing the window length N reduces the standard deviation of frame difference. To compare the results, the TI of each video sequence are reported in Table 3. Again, TI values are dropping with increasing window length. Interestingly, temporal variation is already significantly reduced by only constraining the initialization for each frame LED's to be the same as the previous frame ($N = 1$). These results show that the proposed HDR video rendering algorithm can effectively reduce the effects caused by temporal variation.

5. CONCLUSION AND FUTURE WORK

In this work, an HDR image and video frame reproduction algorithm is presented. The proposed dual modulation method can reproduce HDR content and estimate the emitted luminance accurately. The algorithm employs a sliding window-based filter to avoid flickering. The fidelity of reproduction of HDR luminance and the accuracy of emitted luminance estimation have been tested, and our results show that the proposed algorithm can reproduce HDR content with high precision. As future work, it will be interesting to extend our method to color as well, and to assess its performance by means of a subjective test campaign.

6. REFERENCES

[1] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, "High

dynamic range display systems," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 760–768, Aug. 2004.

[2] M. Narwaria, M. P. Da Silva, and P. Le Callet, "Dual modulation for LED-backlit HDR displays," in *High Dynamic Range Video*. Academic Press, 1st edition, 2016.

[3] P. Korsunov, P. Hanhart, T. Richter, A. Artusi, R. Mantiuk, and T. Ebrahimi, "Subjective quality assessment database of HDR images compressed with JPEG XT," in *7th International Workshop on Quality of Multimedia Experience (QoMEX)*, 2015, number EPFL-CONF-207497.

[4] SIM2, "http://www.sim2.com/hdr/," June 2014.

[5] T. Funamoto, T. Kobayashi, and T. Murao, "High-picture-quality technique for LCD television: Lcd-ai," *AVC Products Development Center, Matsushita Electric Industrial, Co., Ltd*, pp. 1–1, 2000.

[6] C.-C. Lai and C.-C. Tsai, "Backlight power reduction and image contrast enhancement using adaptive dimming for global backlight applications," *Consumer Electronics, IEEE Transactions on*, vol. 54, no. 2, pp. 669–674, 2008.

[7] H. Cho and O.-K. Kwon, "A backlight dimming algorithm for low power and high image quality LCD applications," *Consumer Electronics, IEEE Transactions on*, vol. 55, no. 2, pp. 839–844, 2009.

[8] N. Burini, E. Nadernejad, J. Korhonen, S. Forchhammer, and X. Wu, "Image dependent energy-constrained local backlight dimming," in *IEEE ICIP 2012*. IEEE, 2012.

[9] N. Burini, E. Nadernejad, J. Korhonen, S. Forchhammer, and X. Wu, "Modeling power-constrained optimal backlight dimmingfor color displays," *J. Display Technol.*, vol. 9, no. 8, pp. 656–665, Aug 2013.

[10] J. Korhonen, C. Mantel, N. Burini, and S. Forchhammer, "Modeling the color image and video quality on liquid crystal displays with backlight dimming," in *Visual Communications and Image Processing (VCIP), 2013*. IEEE, 2013, pp. 1–6.

[11] N. Burini, C. Mantel, E. Nadernejad, J. Korhonen, S. Forchhammer, and J. Pedersen, "Block-based gradient descent for local backlight dimming and flicker reduction," *Display Technology, Journal of*, vol. 10, no. 1, pp. 71–79, Jan 2014.

[12] S. Cha, T. Choi, H. Lee, and S. Sull, "An optimized backlight local dimming algorithm for edge-lit LED backlight LCDs," *Journal of Display Technology*, vol. 11, no. 4, pp. 378–385, 2015.

[13] E. Nadernejad, C. Mantel, N. Burini, and S. Forchhammer, "Flicker reduction in LED-LCDs with local backlight," in *Multimedia Signal Processing (MMSP), 2013 IEEE 15th International Workshop on*. IEEE, 2013, pp. 312–316.

[14] E. Zerman, G. Valenzise, F. De Simone, F. Banterle, and F. Dufaux, "Effects of display rendering on HDR image quality assessment," in *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 2015.

[15] T. O. Aydın, R. Mantiuk, and H.-P. Seidel, "Extending quality metrics to full luminance range images," in *Electronic Imaging 2008*. International Society for Optics and Photonics, 2008, pp. 68060B–68060B.

[16] M. D. Fairchild, "The HDR photographic survey," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2007, vol. 2007, pp. 233–238.

[17] ITU-T, "Parameter values for the HDTV standards for the studio and for international programme exchange.," Recommendation BT.709, 1993.

[18] M. Azimi, A. Banitalebi-Dehkordi, Y. Dong, M. T. Pourazad, and P. Nasiopoulos, "Evaluating the performance of existing full-reference quality metrics on high dynamic range (HDR) video content," in *International Conference on Multimedia Signal Processing (ICMSP)*, 2014.

[19] A. Banitalebi-Dehkordi, M. Azimi, M. T. Pourazad, and P. Nasiopoulos, "Compression of high dynamic range video using the HEVC and H.264/AVC standards," in *Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine), 2014 10th International Conference on*. IEEE, 2014, pp. 8–12.

[20] "Recommendation ITU-T P.910:, subjective video quality assessment methods for multimedia applications," 1999.

[21] S. Lasserre, F. LeLéannec, and E. Francois, "Description of HDR sequences proposed by technicolor," *ISO/IEC JTC1/SC29/WG11 JCTVC-P0228, IEEE, San Jose, USA*, 2013.