



Methods for Improving the Tone Mapping for Backward Compatible High Dynamic Range Image and Video Coding

Alper Koz, Frederic Dufaux

► To cite this version:

Alper Koz, Frederic Dufaux. Methods for Improving the Tone Mapping for Backward Compatible High Dynamic Range Image and Video Coding. Signal Processing: Image Communication, 2014, 29 (2), pp.274-292. hal-01433782

HAL Id: hal-01433782

<https://hal.science/hal-01433782>

Submitted on 10 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Methods for Improving the Tone Mapping for Backward Compatible High Dynamic Range Image and Video Coding

Alper Koz and Frederic Dufaux

Multimedia Group, Signal and Image Processing Department (TSI), Télécom ParisTech,
75014 Paris, France.

e-mail: {alper.koz, frederic.dufaux}@telecom-paristech.fr

Abstract

Backward compatibility for high dynamic range image and video compression forms one of the essential requirements in the transition phase from low dynamic range (LDR) displays to high dynamic range (HDR) displays. In a recent work [1], the problems of tone mapping and HDR video coding are originally fused together in the same mathematical framework, and an optimized solution for tone mapping is achieved in terms of the mean square error (MSE) of the logarithm of luminance values. In this paper, we improve this pioneer study in three aspects by considering its three shortcomings. First, the proposed method [1] works over the logarithms of luminance values which are not uniform with respect to Human Visual System (HVS) sensitivity. We propose to use the perceptually uniform luminance values as an alternative for the optimization of tone mapping curve. Second, the proposed method [1] does not take the quality of the resulting tone mapped images into account during the formulation in contrary to the main goal of tone mapping research. We include the LDR image quality as a constraint to the optimization problem and develop a generic methodology to compromise the trade-off between HDR and LDR image qualities for coding. Third, the proposed method [1] simply applies a low-pass filter to the generated tone curves for video frames to avoid flickering during the adaptation of the method to the video. We instead include an HVS based flickering constraint to the optimization and derive a methodology to compromise the trade-off between the rate-distortion performance and flickering distortion. The superiority of the proposed methodologies is verified with experiments on HDR images and video sequences.

Index Terms— High Dynamic Range Video Compression, Tone Mapping, Perceptually Uniform Luminance Values, Backward Compatibility, H.264/AVC, Constrained Optimization, Flickering

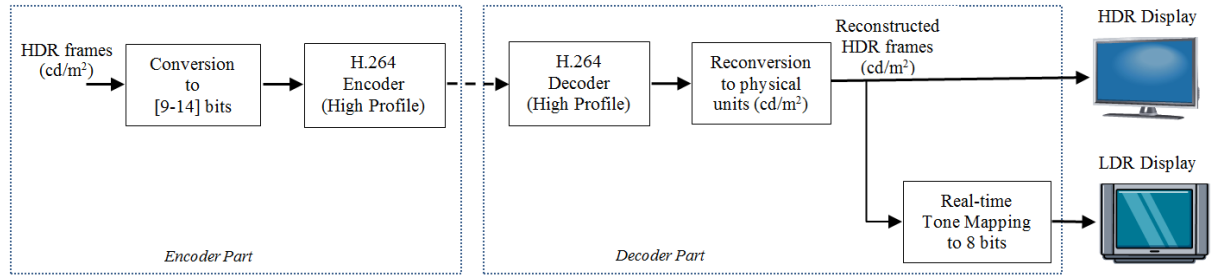
1. INTRODUCTION

Video content represented by eight bits per pixel has been widely accepted in many applications such as internet/video streaming, DVD, HDTV, camcorders, personal computers, televisions and other consumer electronics products. The reasons of such a wide acceptance of eight-bit representation can be argued as the compact representation of one pixel value as a *byte* for the storage in memory chips and the sufficiency of 256 levels to cover the luminance range of commonly used low dynamic range (LDR) displays ($0.1\text{-}80\text{ cd/m}^2$) [2]-[5]. However, image capturing and display technologies have improved to span a wider dynamic range and true color representations in recent years. LCD and plasma displays with a peak luminance of $500\text{-}1000\text{ cd/m}^2$ have dominated the market. HDR prototype displays are now available with a dynamic range of $1,000,000:1$ and a peak luminance of 4000 cd/m^2 [6]. The cinema industries have also begun to capture movies as digital images, rather than film, by using modern digital cameras that have extended dynamic range feature [7]. Even more, the smart phones, which have reached 55 % of usage penetration in EU5 countries (UK, DE, FR, ES, and IT) by October 2012 [8], now possess HDR photo utilities to generate more realistic views. Accordingly, standard eight-bit representation has become insufficient for these new technologies and related applications in digital cinema, medical imaging, and post-production.

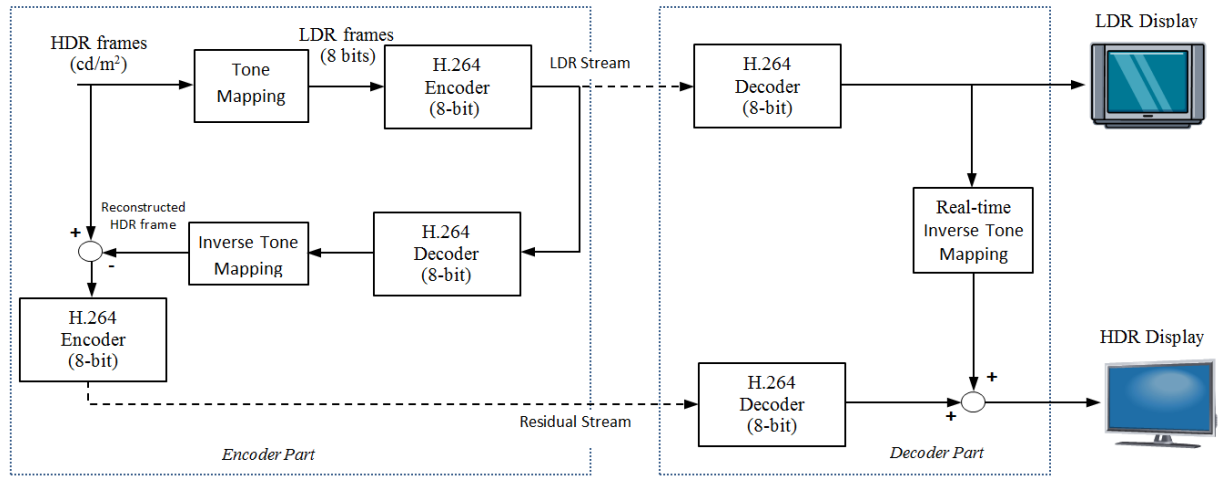
HDR image and video formats are proposed to overcome the limitations of standard eight-bit representation. The main goal with these formats is to encode, transmit, and display the accurate physical luminance values (cd/m^2) in a real-world scene rather than the eight-bit intensity values and hence to generate a scene referred image independent of the display technology [1]-[5]. The proposed formats for such a goal obviously require higher bit depths than standard eight-bits when the entire luminance range of natural scenes ranging from extreme darkness (10^{-6} cd/ m^2) to bright sunshine (10^8 cd/ m^2) is considered. Current formats such as Radiance RGBE (*.hdr*), OpenEXR (*.exr*), and LogLuv TIFF (*.tiff*) use 16-bit floating point, 32-bit floating point or 32-bit integer pixels [1].

A major challenge for such a high-bit depth representation is an efficient compression, in particular for HDR video which requires significantly more storage size and transmission bandwidth than standard 8-bit LDR video. The size of the HDR video captured by the first HDR camera demonstrated at the annual ACM SIGGRAPH conference in 2009 was reported as 42 GB for a footage of one minute, compared with just 9 GB for its LDR counterpart [2]. The previous work has mainly handled this challenge in two ways (Figure 1). The first approach [9]-[11] is to take the advantage of high profile of existing state-of-the-art H.264/AVC (Advanced Video Coding) codec, which can support the bit-depths more than 8 bits up to 14 bits. These methods first convert the real luminance values (cd/m^2) of the HDR video pixels into a bit-depth between [9-14] bits and apply the H.264/AVC high profile encoder. After the decoding and reversion to luminance values (cd/m^2), the reconstructed HDR video is shown on a HDR display, or on a LDR display by performing a real-time tone mapping operation, as illustrated in Figure 1-a.

The second approach [1], [5], [12]-[13] aims to provide the backward compatibility with existing standard 8-bit displays in order to enable a successful transition to HDR technology (Figure 1-b). The methods in this approach first map the HDR video into 8-bit to produce an LDR video. The LDR video goes through the video encoding and decoding process by using a standard 8-bit (e.g. H.264/AVC) encoder. Then, the high dynamic range video is reconstructed from LDR video by applying a real-time inverse tone mapping operation. The parameters regarding the tone mapping and inverse tone mapping can be sent as a look-up table [1] or can be encoded using supplemental enhancement information [15]. The residue of the reconstructed HDR video can also be further encoded as an enhancement layer in the bit-stream. As this approach includes the LDR stream separately in the resulting bit-stream, it provides a direct compatibility to LDR displays. However, it requires real-time inverse-tone mapping operation to show the video content on HDR displays.



(a)



(b)

Figure 1: A general scheme for the two approaches in HDR video compression, (a) using high profile H.264/AVC codec, and (b) using 8-bit H.264/AVC codec for backward compatibility to LDR displays.

In a recent work [1], the optimum tone mapping for backward compatible HDR video compression which minimizes the mean square error between the logarithm of the luminance values of the original and reconstructed HDR frames (HDR MSE) (Figure 1-b) is formulated in a detailed and complete manner. In order to find a computationally efficient solution, Mai *et al.* [1] have estimated the distortion in the chain of tone-mapping, encoding, decoding, and inverse-tone mapping with a statistical distortion model and found a closed-form solution for tone mapping based on the luminance histogram of the HDR video frames. The pioneer aspects of this performed research were to consider the tone mapping and HDR

video coding problems jointly in the same mathematical framework and to derive a simple closed-form solution. However, it also includes the following shortcomings on the downside.

First, the proposed method by Mai *et al.* [1] works on the logarithm of actual physical luminance values (cd/m^2) in accordance with the classical *Weber* law to produce uniform values with respect to the HVS sensitivity. However, Aydin *et al.* in [16] has indicated that the luminance sensitivities of HVS cannot be just modeled with the *Weber* law for the whole luminance range and has instead developed a perceptual mapping as an alternative to the logarithm operation to convert the luminance values into perceptually uniform encoded values. Second, the proposed approach [1] does not take the quality of the resulting LDR images into account during the formulation of the optimization. Such an approach contradicts with one of the main goals of tone mapping research, which aims to produce more pleasant LDR views for a given HDR image. Consequently, the proposed method [1] can produce some non-natural over illumination in the bright regions of the produced LDR images, although it minimizes the HDR-MSE for the reconstructed HDR images. Although the authors extend this pioneer study in [17] by means of considering the LDR image quality and the bitrates of LDR layer and residue coding, they do not provide a practical methodology and a closed form solution to set the trade-off between HDR and LDR image qualities. Third, the adaptation of the proposed tone mapping [1] to video sequences is performed by applying a low-pass filter to the generated tone curves for video frames. Such an approach does not consider any HVS measure for the flickering kind of distortions which emerges as sudden changes in the average brightness of the consecutive frames, when a high luminance object or region enters or leaves the camera view [18]. It is therefore not an optimum solution giving the best rate distortion performance which avoids the flickering distortions at the same time.

In this paper, we improve the optimized tone mapping for backward compatible HDR image and video compression in three aspects by considering the mentioned three shortcomings. Our contribution is three-folded:

- First, we propose to use perceptually uniform encoded values rather than the logarithm of the absolute luminance values for the optimization of tone mapping curve. We verify that the

proposed approach gives better results in terms of the perceptual metric, perceptually uniform peak signal to noise ratio (PU-PSNR) [16], and produces more realistic LDR images.

- Second, we include the LDR image quality as a constraint to optimized tone mapping problem. We reformulate the problem with a more general description, namely, finding the optimum tone mapping giving the best rate-distortion performance for HDR images while keeping the quality of produced LDR images in a desired level. The quality of the produced LDR image is determined with respect to the MSE distance to a reference LDR image, which is produced by a reference global tone mapping operator (TMO). With such a constrained formulation, we develop a generic methodology to compromise the trade-off between the HDR image quality and LDR image quality for HDR video coding.
- Third, we include a flickering constraint to the optimized tone mapping problem for the adaptation of the method to the video. We use a perceptual measure based on the classical *Weber* law [19] developed in [18] to control the level of flickering between the average intensities of consecutive frames. By using such a measure as a constraint, we derive a methodology to compromise the trade-off between the rate-distortion performance and flickering distortion.

The next section gives an overview of the HDR video coding methods. We present the optimized tone mapping for backward compatible HDR video coding developed in [1] in more detail as it forms a base for our proposed methodologies. Section 3 first describes the perceptually uniform (PU) encoding of luminance values [16] and then presents the proposed optimized tone mapping based on PU values. Section 4 presents the tone mapping optimization problem with the LDR image quality constraint. Section 5 follows with the tone mapping optimization problem with the flickering constraint. The experimental results and comparisons for the proposed methods are presented in Section 6. Finally, we draw our conclusions in Section 7.

2. HDR VIDEO COMPRESSION CLASSIFICATION

The methods on HDR video compression can be classified in two categories as the methods for native encoding of HDR video by using high profile H.264/AVC encoder, and the methods for backward compatible encoding of HDR video by using standard 8-bit encoders.

A. Methods for native encoding of HDR video using high profile H.264/AVC encoder and decoder

The core part of the methods in this approach [9]-[11] is to convert the real luminance values of HDR video pixels into the bit-depths that the high profile can support. The pioneer work by Thoma *et al.* [9] performs this conversion by applying the modified versions of LogLuv transform [20], which maps the luminance range of each HDR frame to the range $[0, 2^n-1]$ with respect to the maximum and minimum luminance values of the frame, where n is the bit-depth of the resulting video.

The *LogLuv* transform [20] is originally proposed by Larson to provide a compact encoding suitable for the transfer, manipulation, and storage of high dynamic range color images. This format is a replacement for conventional RGB images and encodes color pixels as log luminance values and chromaticity components (u, v) [20] as follows:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.497 & 0.339 & 0.164 \\ 0.256 & 0.678 & 0.066 \\ 0.023 & 0.113 & 0.864 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

$$x = \frac{X}{X + Y + Z} \quad (2)$$

$$y = \frac{Y}{X + Y + Z} \quad (3)$$

$$u_e = \left\lfloor 410 \frac{4x}{-2x + 12y + 3} \right\rfloor \quad (4)$$

$$v_e = \left\lfloor 410 \frac{4y}{-2x + 12y + 3} \right\rfloor \quad (5)$$

$$L = \lfloor 256(\log_2 Y + 64) \rfloor \quad (6).$$

Such a transform maps the real luminance values (Y) in the interval $[5.44 \times 10^{-20}, 1.84 \times 10^{19}]$ to 15 bit integer luma values (L) in the range $[0, 2^{15}-1]$ and covers over 38 orders of magnitude in 0.27 % relative step size. As this step size is much below the visible quantization threshold of 1 % [19], Thoma *et al.* [9] modify the *LogLuv* mapping to exploit the full range of luma code values for a given bit depth as:

$$L_n = \left\lfloor 410 \frac{2^n - 1}{\log_2(Y_{\max}/Y_{\min})} (\log_2(Y) - \log_2(Y_{\min})) \right\rfloor \quad (7)$$

$$Y = 2^{(L_n + 0.5) \left(\frac{\log_2(Y_{\max}/Y_{\min})}{2^n - 1} \right) + \log_2(Y_{\min})} \quad (8)$$

where Y_{\max} and Y_{\min} are the maximum and minimum luminance values of a HDR video frame. The transformation is applied to each frame of the HDR video, and Y_{\max} and Y_{\min} for each frame are also transferred with the bit-stream to perform the reconversion (Figure 1-a).

In the extended version of this work [10], temporal motion compensation problems generated by individual mapping of frames is solved by the weighted prediction (WP) tool of H.264/AVC. In a similar work [11], such a mapping is enhanced by smoothing the HDR frames with respect to the Human Visual System (HVS) sensitivities in different wavelet resolutions. In [14], Mantiuk *et al.* have perceptually optimized the luminance quantization levels for the contrast threshold perception of HVS. The optimized quantization scheme can support 10-11 bits to encode 12 orders of magnitude of visible luminance range without causing any visible distortion. The advantage of these methods was mostly their conceptual simplicity. On the down side, they do not support the backward compatibility with 8-bit displays which is required for a smooth transition to HDR technology.

B. Methods for backward compatible encoding of HDR video

The methods in this approach [1], [5], [12]-[13] are proposed to provide the backward compatibility with existing standard 8-bit displays. The basic strategy for such an aim is to include an inter-layer prediction stage in order to estimate the HDR video from LDR video. The residue of the estimation and the prediction coefficients are then encoded and written to the output bit-stream with the independently

encoded LDR video. In [12], a block wise approach is proposed for interlayer prediction. A linear multiplication and an addition of offset are applied to each block of LDR video frame to estimate the corresponding block of HDR video frame. This method is improved in [5] by differentially encoding the offsets with respect to the previously encoded left and upper blocks.

Mantiuk et al. [13] has used a frame-wise reconstruction function instead of a block-wise prediction. Their approach first applies a color transformation [14] to both the LDR and corresponding HDR frames to end up with a compatible domain where two contents are comparable. Then, they apply a reconstruction function, which is based on the statistics (i.e. arithmetic mean, median or midrange) of the HDR pixel values corresponding to each LDR pixel value, to the LDR frame. The invisible noise in the residual frames is also filtered after the prediction. Some other similar approaches [21] are also proposed for backward compatible compression of HDR textures to enable real-time applications on existing GPU-hardware.

In a recent work [1], tone mapping and inverse tone mapping are considered jointly as the parts of coding and decoding processes. By statistically modeling the error in the chain of tone-mapping, encoding, decoding, and inverse-tone mapping (Figure 1-b), Mai *et al.* [1] has achieved to find the optimum tone mapping and inverse tone mapping which minimizes the mean square error between the logarithms of the luminance values of the original and reconstructed HDR frames (HDR-MSE). As this method forms a base for our proposed improvements, we would give its description in detail.

Optimum Tone Mapping for backward compatible HDR video compression

Given the notations l and v , corresponding to the *logarithm* of the luminance of the HDR frame and the pixel values of tone mapped LDR version respectively, tone mapping curve in [1] is first parameterized as a piece-wise linear function with the nodes (l_k, v_k) as shown in Figure 2. Each segment k between two nodes (l_k, v_k) and (l_{k+1}, v_{k+1}) has a constant width in HDR values equal to δ (selected as 0.1). The tone mapping operation is then characterized by a set of slopes

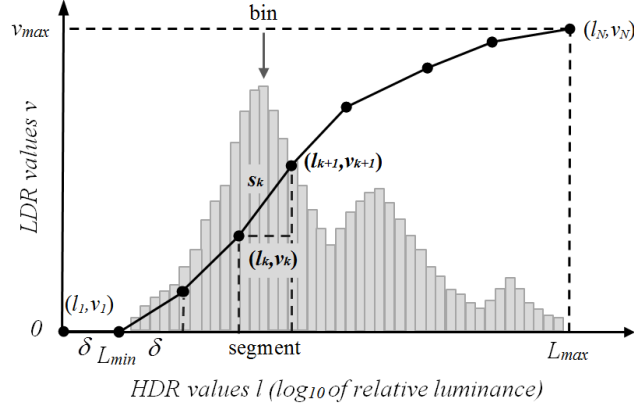


Figure 2: Parameterization of the TM curve in Mai et al. [1]. The bar-plot is the luminance histogram of the HDR frame.

$$s_k = \frac{v_{k+1} - v_k}{\delta} \quad (9),$$

which forms a vector of tone-mapping parameters θ . Using such a parameterization and the statistical modeling of the H.264/AVC coding error, the optimized tone mapping problem is given as

$$\begin{aligned} \arg \min_{s_1 \dots s_N} \quad & \sum_{k=1}^N \frac{p_k}{s_k^2} \\ \text{subject to} \quad & \sum_{k=1}^N s_k = \frac{v_{\max}}{\delta} \end{aligned} \quad (10),$$

where p_k is the summation of the normalized histogram of luminance values for the k 'th bin, N is the total number of bins in the histogram, and v_{\max} is the maximum LDR value [1]. By computing the first order Karush-Kahn-Tucker (KKT) optimality conditions of corresponding Lagrangian, their ultimate closed form solution is derived as

$$s_k = \frac{v_{\max} \cdot p_k^{1/3}}{\delta \cdot \sum_{k=1}^N p_k^{1/3}} \quad (11).$$

A tone mapping characterized by the slopes in (11) minimizes the MSE between the luminance values of the original and reconstructed HDR frames (Figure 1-b).

3. OPTIMUM TONE MAPPING WITH PERCEPTUALLY UNIFORM ENCODING

The proposed tone mapping method by Mai *et al.* [1] operates on the logarithm of luminance values to comply with the sensitivity of HVS to different luminance levels in accordance with *Weber* law [19]. However, *Weber* law is not an exact model of HVS contrast sensitivity for the whole luminance range [16]. Therefore, such an optimization procedure is not optimum in the perceptual sense although it minimizes the mean square error between the logarithms of the luminance values of the original and reconstructed HDR frames. By considering this fact, we propose to use the perceptually uniform (PU) encoded luminance values developed in [16] instead of the logarithm operation, and to minimize the mean square error between the PU values to maximize the perceptual metric, *PU-PSNR*. We first briefly summarize the PU encoding in the next subsection and then present the proposed tone mapping method.

A. Perceptually Uniform Encoding of Luminance Values

Commonly used quality metrics such as PSNR and SSIM assume that pixel intensities are perceptually uniform regardless of their luminance values. Although such an assumption applies well for CRT displays with the luminance range of 0.1-80 cd/m^2 , it does not exactly fit to much brighter LDR (500 cd/m^2), plasma (1000 cd/m^2) and HDR (4000 cd/m^2) displays as the distortions in image pixels become more visible with higher luminance. Aydin *et al.* [16] has therefore proposed a mapping which converts luminance values into approximately perceptually uniform values in order to enable an extension of the well-known PSNR metric to all luminance levels visible to HVS.

Their motivation to this end is mainly based on two observations. First, classical *Weber* law [19] stating that the ratio of the perceivable luminance difference to the background luminance (i.e. *contrast threshold*) is constant is only valid for the background luminance greater than 500 cd/m^2 , as indicated with the contrast versus intensity (*cvi*) curve in Figure 3. It is therefore not a very accurate model for the range of luminance shown in typical displays (0.1 cd/m^2 to 100-1000 cd/m^2). Second, the compressive transform function of sRGB color space used in typical displays to convert luma values (pixel intensities) to luminance values is not very appropriate for brighter displays. This is verified in Figure 3-b by

showing the strong deviations in the slopes of the quantization error between the luminance values corresponding to consecutive luma values from *cvi* curve for brighter displays.

In order to derive a perceptually uniform (PU) encoding that is applicable to whole luminance range, contrast thresholds should be properly estimated as a function of the pixel luminance. The thresholds are however dependent on many aspects of complex images such as spatial frequency, orientation, image size, and adapting luminance. In order to end up with such a function which is only dependent on the pixel luminance, the authors use the contrast sensitivity function (CSF) in [22] and choose the value that corresponds to the maximum sensitivity for each factor in the CSF function except the adapting luminance. Then, they also reduce the dependence of the contrast thresholds to adapting luminance by assuming that it is equal to pixel luminance. Such choices result in a conservative estimation of the contrast thresholds. The estimated thresholds correspond to the differentials of the PU encoding curve that is searched for. At the same time, such an encoding should also be backward compatible with the sRGB transform resulting in similar values to be able to use in LDR displays. Based on these constraints, the authors in [16] first find the forward mapping from luminance to luma values by a recursive process. Then, they perform an approximation to make the PU encoding curve as close as possible to the sRGB curve for the LDR range of 0.1-100 cd/m². The mathematical details of this recursive process and the approximation are given in [14]. The resulting PU encoding is shown in Figure 4. While there is a closer match to sRGB color transform for the LDR range of 1-100 cd/m², the difference for the higher luminance is quite significant.

PU encoding curve is used to define the *PU-PSNR* metric between the HDR images where the luminance values are converted to PU values and the PSNR in terms of the PU values is computed between the HDR images. Such a metric is tested by means of psychovisual tests performed on regular and bright displays and validated as giving better results than the standard PSNR. We will use the perceptually uniform encoded values in this paper to derive the optimum tone mapping which maximizes the PU-PSNR value between the original and reconstructed images for backward-compatible HDR image and video compression.

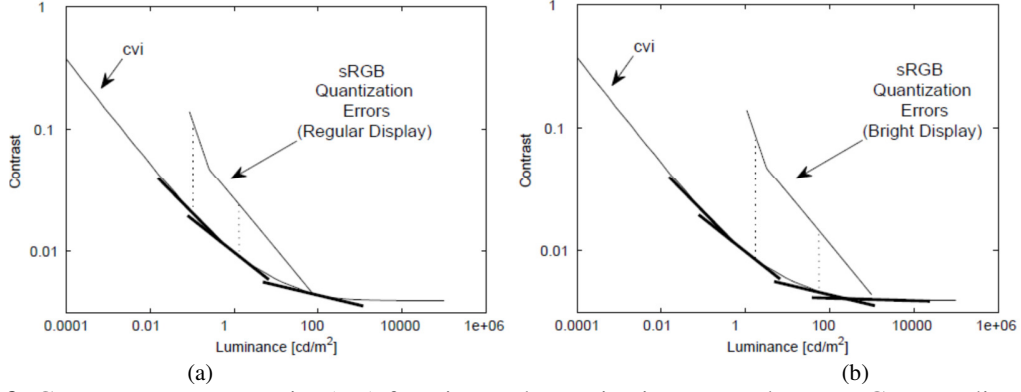


Figure 3: Contrast versus Intensity (*cvi*) function and quantization errors due to sRGB encoding, (a) in a regular display (80 cd/m²) and, (b) in a bright display (1000 cd/m²) [16].

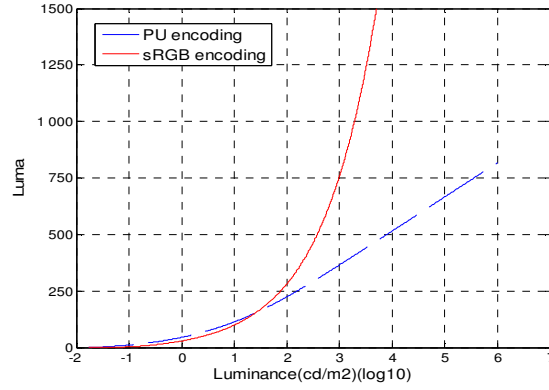


Figure 4: Perceptually Uniform Encoding and sRGB encoding as a function of luminance (cd/m²).

B. Proposed Method with Perceptually Uniform Luminance Values

Figure 5 illustrates the main stages of the proposed method in parallel with Figure 1-b. First, the luminance values of HDR frames (represented as floating point real numbers in cd/m²) are converted into perceptually uniform values by PU encoding (Figure 4). The resulting values are represented as l_u in floating points. Then, the optimized tone mapping is applied to l_u values to obtain the LDR video in 8-bit representation. After the LDR video goes through the coding and decoding, inverse tone mapping is applied to the resulting sequence to obtain the perceptually uniform values. Finally, the actual luminance values are found by using the inverse PU encoding.

As the solution derived in [1] is only based on the histogram of the input image and the statistical modeling of H.264/AVC coding error, such a formulation for the optimum solution can also be applied to

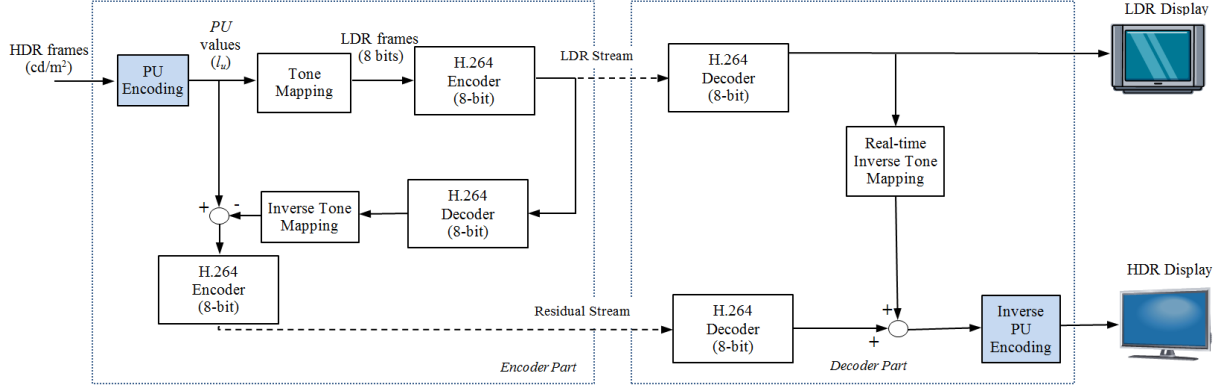


Figure 5: Proposed scheme for backward compatible HDR video compression with PU encoding.

perceptually uniform values without loss of generality. For the given frame, the histogram over l_u values is computed and the slopes of the tone mapping operator is found similar to Equation (11). Then, the resulting tone mapping is applied to the frames.

4. OPTIMUM TONE MAPPING WITH LDR IMAGE QUALITY CONSTRAINT

One of the important shortcomings of the optimized tone mapping developed by Mai et al. [1] was not to take the quality of the resulting LDR images into account during the formulation of optimization. The proposed method [1] can therefore produce some non-natural over illumination in the bright regions of the produced LDR images, although it minimizes the HDR-MSE for the reconstructed HDR images. Considering that the global tone mapping operations are widely studied to obtain a pleasant LDR view in HDR research, we would search in this section to find the optimized tone mapping while preserving the LDR quality in a desired level. The quality of the LDR images is determined with respect to the MSE distance to a reference LDR image, which is produced by a reference global tone mapping operator. *MSE* or equivalently *PSNR* is one of the most common metrics used in image and video processing to present the level of distortion on the content. In our case, it also provides a tractable optimization problem where the derivatives of the expressions can be analytically computed.

Let us assume that h represents a reference global tone mapping operator and s is the optimized TMO we are searching (see Figure 6.a.). Using the same piece-wise linear parameterization also for h , the MSE between the produced LDR image and the reference LDR image can be approximated as

$$E((s(l) - h(l))^2) \cong \sum_{k=1}^N \left(\sum_{n=1}^k (\delta \cdot s_n) - \sum_{n=1}^k (\delta \cdot h_n) \right)^2 \cdot p_k = \delta^2 \sum_{k=1}^N \left(\sum_{n=1}^k (s_n - h_n) \right)^2 \cdot p_k \quad (12)$$

where s_k and h_k correspond to the slopes for s and h , respectively. By constraining also the MSE distance to the reference LDR image, the minimization problem in (10) returns into a double constraint problem as

$$\begin{aligned} & \arg \min_{s_1 \dots s_N} \sum_{k=1}^N \frac{p_k}{s_k^2} \\ & \text{subject to } \sum_{k=1}^N s_k = \frac{v_{\max}}{\delta} \\ & \delta^2 \sum_{k=1}^N \left(\sum_{n=1}^k (s_n - h_n) \right)^2 \cdot p_k = \Delta \end{aligned} \quad (13)$$

After defining the *Lagrangian* and performing some algebra on the derivative of the MSE in (12), the KKT conditions for the above optimization problem can be reduced to the following set of equations:

$$\begin{aligned} & \frac{-2p_k}{s_k^3} + \lambda_1 + \lambda_2 \delta^2 \sum_{a=k}^N \left(\sum_{n=1}^a 2(s_n - h_n) \right) \cdot p_a = 0, \quad k = 1..N \\ & \sum_{k=1}^N s_k - \frac{v_{\max}}{\delta} = 0 \\ & \delta^2 \sum_{k=1}^N \left(\sum_{n=1}^k (s_n - h_n) \right)^2 \cdot p_k - \Delta = 0 \end{aligned} \quad (14)$$

where λ_1 and λ_2 are the *Lagrangian* coefficients for the two constraints in (13). Finding a closed form solution seems not feasible for such a set of equations as it ultimately ends up with a fourth order polynomials of s_k 's. Although adaptive search algorithms can be considered as an alternative, such an approach would also be computationally heavy for coding applications. Considering these facts, we

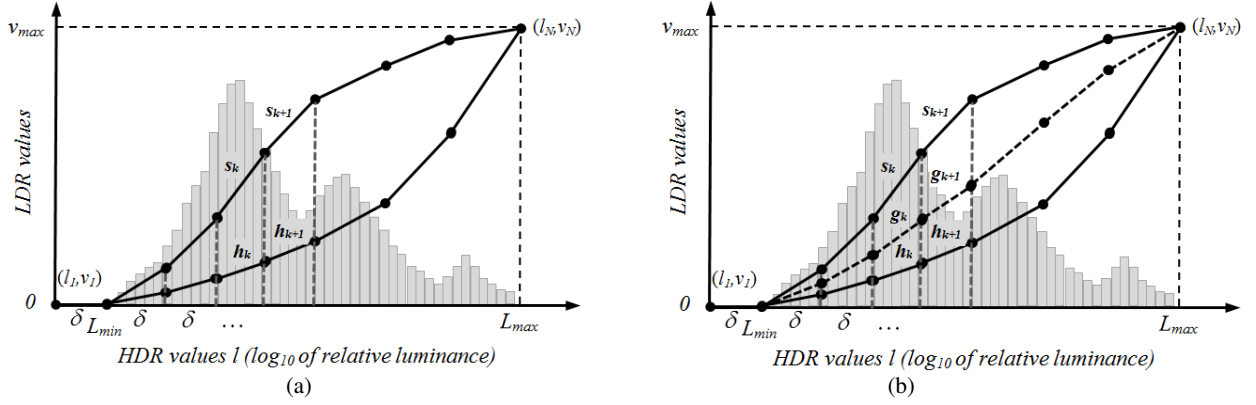


Figure 6 (a): the searched optimum TM with the LDR constraint (s) and the reference TM operator (h) in (13), **(b)** the optimized TM minimizing HDR MSE (s), the reference TM operator (h), and the searched optimum TM with the LDR constraint (g) for the modified problem in (15).

propose to follow a two-step strategy to arrive a sub-optimal solution to arrange the trade-off between the LDR quality and HDR quality.

Our proposed solution is first to use the found optimum solution for s_k in (11) for the one-constraint optimization problem, and then to redefine the cost function as a MSE distance to this optimal solution, while keeping the MSE distance to the reference LDR image in a desired level. Let us assume that s_k , h_k , and g_k correspond to the given optimal solution in (11), the reference global tone mapping operation, and the tone mapping operation we are looking for, respectively (see Figure 6.b). Then, the new optimization problem in terms of the MSEs can be expressed as:

$$\begin{aligned} \arg \min_{g_1 \dots g_N} \quad & \delta^2 \sum_{k=1}^N \left(\sum_{n=1}^k (s_n - g_n) \right)^2 \cdot p_k \\ \text{subject to} \quad & \delta^2 \sum_{k=1}^N \left(\sum_{n=1}^k (g_n - h_n) \right)^2 \cdot p_k = \Delta \end{aligned} \quad (15).$$

The solution of the above constrained optimization problem with second order expressions can be found analytically by calculating the KKT conditions of its Lagrangian function:

$$\begin{aligned}
& \delta^2 \sum_{a=k}^N \left(\sum_{n=1}^a 2(g_n - s_n + \lambda(g_n - h_n)) \right) \cdot p_a = 0, \quad k=1..N \\
& \delta^2 \sum_{k=1}^N \left(\sum_{n=1}^k (g_n - h_n) \right)^2 \cdot p_k - \Delta = 0
\end{aligned} \tag{16}$$

where λ is the *Lagrangian* coefficient for the constraint. By using also the formulation for MSE in (12), the solution to (16) can be finalized as:

$$\begin{aligned}
g_k &= \frac{s_k + \lambda h_k}{1 + \lambda} = h_k + \frac{(s_k - h_k)}{1 + \lambda}, \quad k=1..N \\
\lambda &= \sqrt{\frac{\delta^2 \sum_{k=1}^N \left(\sum_{n=1}^k (s_n - h_n) \right)^2 \cdot p_k}{\Delta}} - 1 = \sqrt{\frac{E((s(l) - h(l))^2)}{\Delta}} - 1
\end{aligned} \tag{17}$$

The λ parameter in such a closed form solution sets the trade-off between the HDR image quality and LDR image quality by adjusting the distance of the resulting tone mapping curve to the corresponding two tone mapping curves. When λ is equal to 0, g becomes equal to s . The MSE between the original and reconstructed HDR frames after the coding is minimized, but the MSE distance of the resulting LDR image to the one produced with reference TMO (h) is maximized. As λ increases (or Δ decreases), g moves away from s and results in an increase in HDR MSE, but gets closer to the reference TMO (h) on the other hand. Our experiments will focus on the verification of this trade-off.

5. OPTIMUM TONE MAPPING WITH FLICKERING CONSTRAINT

Another shortcoming of the optimized method developed by Mai et al. [1] was not to handle the adaptation of the derived tone mapping to the video in detail. As a simple remedy solution to the flickering kind of distortions that can occur on the generated LDR video, they propose to apply a low pass filtering to the generated tone curves for video frames without taking explicitly any HVS aspect into account. In this section, we reformulate the optimized tone mapping problem with an HVS based flickering constraint to achieve its application to video.

In vision research, the minimum amount of perceptible change in the luminance of a stimulus is defined as *just noticeable difference* (JND). The relation of the JND to the stimulus luminance is stated by *Weber* law as:

$$k_w = \frac{\Delta L}{L} \quad (18).$$

Such a relation says that the size of the JND (ΔL) is a constant proportion (k_w) of the initial stimulus luminance (L) [19].

Guthier *et al.* [18] in a recent work propose to use the *Weber* constant (k_w) as a perceptual measure to avoid the flicker distortion in the resulting tone mapped video. Their proposed solution imposes the average brightness of the consecutive frame to be in the JND (ΔL) range of the average brightness of previous frame while performing the tone mapping operation. The performance of such a perceptual metric to detect the flickering on LDR frames is in fact equivalent of the problem of binary hypothesis in decision theory. A metric can give false positives if there is not a flickering effect in the real HDR sequence, or can miss the flickering detection if there is really a flickering effect on the LDR video. Such a performance analysis over the false positive and missed detection rates is performed in [18] for the *Weber* measure by using a ground truth obtained by the psychovisual experiments over a large set of video sequences. The authors [18] statistically validate the *Weber* measure as a good indicator of the flickering distortion with these tests. In this paper, we will use such a measure to represent the flickering distortion on the LDR video and derive the flickering constrained optimized tone mapping.

Without loss of generality, let us first define the following variables and parameters:

- s Optimized tone mapping of Mai et al. [1] in (11) for the n -th frame of the video (Figure 7).
- g Optimized tone mapping we are searching for the n -th frame with the flickering constraint (Figure 7).
- L_s The average intensity of the produced LDR frame with s .

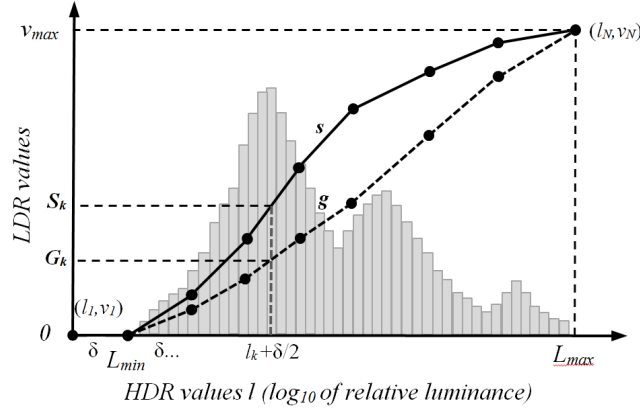


Figure 7: Illustration for the optimum tone mapping of Mai et al. (s) and the flickering constrained optimized tone mapping (g).

L_g The average intensity of the produced LDR frame with g .

L_{n-1} The average intensity of the produced $(n-1)$ -th frame.

ΔL_{n-1} The JND level for the average brightness of the $(n-1)$ -th frame, which is $k_w L_{n-1}$.

In order to avoid a flicker effect between the $(n-1)$ -th and n -th frames, the average intensity of the resulting LDR frame (L_g) should be in the JND (ΔL_{n-1}) interval of L_{n-1} :

$$L_g = \begin{cases} L_{n-1} + k_w L_{n-1}, & \text{if } L_s > (L_{n-1} + k_w L_{n-1}) \\ L_{n-1} - k_w L_{n-1}, & \text{if } L_s < (L_{n-1} - k_w L_{n-1}) \\ L_s, & \text{otherwise} \end{cases} \quad (19).$$

In terms of the histogram values, the average brightness L_g can be expressed as

$$L_g \cong \sum_{k=1}^N p_k g(l_k + \delta/2) = \sum_{k=1}^N p_k G_k \quad (20)$$

where G_k corresponds to the tone mapped value for the average luminance of the k -th bin of the histogram (Figure 7).

With such a constraint on L_g value, we will find the optimum tone mapping curve g , whose produced LDR image has the minimum mean square error (MSE) distance to the one produced with the tone mapping s :

$$\begin{aligned}
& \arg \min_{G_1 \dots G_N} \sum_{k=1}^N p_k (G_k - S_k)^2 \\
& \text{subject to } \sum_{k=1}^N p_k G_k = L_g
\end{aligned} \tag{21}$$

where $S_k = s(l_k + \delta/2)$ (Figure 7).

After defining the *Lagrangian* and solving the *KKT* conditions, the solution to the above problem is found as:

$$\begin{aligned}
G_k &= S_k - \frac{\lambda}{2}, \quad k = 1..N \\
\lambda &= 2(L_s - L_g)
\end{aligned} \tag{22}$$

where λ is the *Lagrangian* coefficient for the constraint. The above equation states that the optimized tone mapping (s) should be shifted vertically to upwards or downwards with an amount equal to the difference of the average values L_s and L_g to obtain the optimized tone mapping with the flickering constraint.

The algorithm to avoid flickering with such an equation is applied iteratively to each frame of the HDR video sequence as follows:

- Find and apply the optimum tone mapping s in (11) to the first frame of the sequence.
- For each frame beginning from the second frame,
 - find the optimum tone mapping s and compute the average intensity, L_s ,
 - compute the required average intensity, L_g , in (20) for a given k_w value by using the average intensity of the previous frame, L_{n-1} ,
 - find the constrained optimized solution by using (22) and apply to the frame.

The Weber measure k_w in such an algorithm controls the level of the flickering effect on the resulting LDR video. As k_w is decreased, the flickering between the average values of consecutive frames decreases. However, this results in a deviation from the optimized tone mapping s and therefore, the rate distortion performance of HDR video coding is expected to decline. If k_w is increased, L_s mostly falls in the JND range of L_{n-1} (Eq. (19)) and the constrained tone mapping g gets closer or equal to the optimized

tone mapping s . In this case, the rate-distortion is expected to be better but the resulting video contains more flickering distortions. Our experiments will focus on the verification of this trade-off.

6. EXPERIMENTS AND COMPARISONS

We have utilized the HDR images and video sequences from the databases given in [23]-[25] for the experiments. H.264/AVC reference software JM 17.2 [26] is used for the encoding of LDR images and video sequences. The experiments on the HDR images are performed by using the intra coding of the H.264/AVC whereas the video sequences are coded with a GOP structure IPPP and a GOP size 15. The supplemental information for the characterization of tone mapping curves is ignored in the given bitrates as it is very negligible (in byte levels) compared to the coding bitrate. In addition, the complexities of the performed PU encoding and the weighting or shifting operations on the tone mapping curves for the LDR quality constraint and flickering constraint optimizations are also negligible compared to the complexity of LDR video coding, which is common in all the compared methods. The results and comparisons for each of the proposed methods are presented in the following subsections.

A. *Experiments on PU Encoding based Tone Mapping*

We have utilized a look-up table for PU encoding and inverse PU encoding, which gives the absolute luminance value and its corresponding PU encoded value as a pair, rather than fitting a model, during the experiments. A look-up table offers better accuracy and is usually faster to compute than power or logarithmic functions approximating such encodings [16]. The size of the table for the whole luminance range from 10^{-2} to 10^7 cd /m² is about 96 kB, which is quite small for storage both at the encoder and decoder.

The performance of the proposed tone mapping based on perceptually uniform encoded values is compared with the tone mapping based on luminance values [1]. The comparisons are performed with respect to the quality of the produced LDR images and with respect to the quality of reconstructed HDR image after the coding of LDR images at different bitrates.

Figure 8 illustrates the resulting tone mapping curves and the histograms for the logarithm of the luminance values (l) and PU values (l_u), respectively, for the HDR test images *Atrium Night*, *Coby*, *Memorial*, and *Forest Path*. The number of the histogram bins in the proposed method is held equal to the one in the compared method for a fair comparison. The resulting tone mapping curves for the first three images are different for both approaches. Tone mapping curves for *Forest path* are similar as the luminance range of *Forest Path* is mostly higher than the 100 cd/m^2 , where PU encoding is mostly linear with respect to the logarithm of the luminance (Fig. 4).

Figure 9, 10 and 11 shows the produced LDR images for both approaches. Some over-illumination spread over the whole image is observable for the compared method [1] in particular for *Atrium Night* and *Coby*. Such an affect can also be seen in the high luminance regions (e.g. the windows and the dome) of *Memorial*. The proposed method eliminates these effects by using the perceptually uniform luminance values and produces more natural LDR images. A drawback is seen however for the darker region of *Memorial* which is zoomed in Figure 12. Although the high luminance regions are more realistic for the proposed approach, the details in the darker regions (e.g. the ceiling in the middle-left of *Memorial*) are less visible in contrast to the compared method [1]. The proposed PU encoding is performing more compression compared to the logarithmic transformation as can be deduced from Figure 4. The produced LDR images for *Forest Path* are very similar as the tone mapping curves are very close.

Figure 13 gives the rate distortion curves in terms of the coding bit rate of LDR image vs. the quality of the reconstructed HDR image. We have used three different quality metrics for the comparisons, namely, HDR MSE (i.e. mean square error between the logarithm of luminance values of original and reconstructed HDR images) [1], PU-PSNR [16], and the HDR visible difference predictor (HDR VDP) [24]. HDR-VDP [27] is proposed as an extension of the VDP metric [22] to HDR images and forms one of the few perceptual metrics with PU-PSNR that is specially tailored for HDR images. The metric takes two HDR images as the input and produces a probability map which shows the probability of the visibility of the difference between two images at each pixel location. The percentage of the pixels that exceeds 0.95 or 0.75 probability is given as the output of the metric. As expected, the compared method

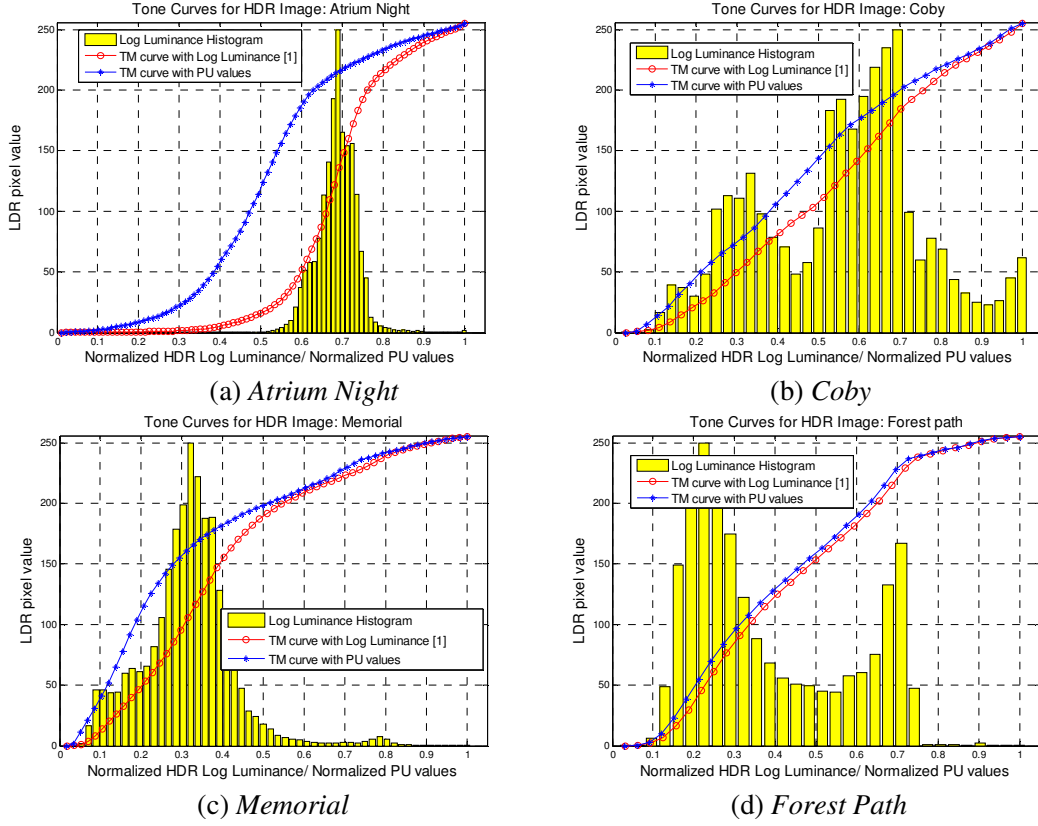


Fig. 8: Tone curves with log. luminance [1] and with the proposed PU encoding. (a) *Atrium Night*, $l_{max} = 3.95$; $l_{min} = -1.77$; $l_{u, max} = 506.62$; $l_{u, min} = 0.32$, (b) *Coby*, $l_{max} = 2.93$; $l_{min} = -0.59$; $l_{u, max} = 355.47$; $l_{u, min} = 22.16$, (c) *Memorial*, $l_{max} = 5.05$; $l_{min} = -0.48$; $l_{u, max} = 673.89$; $l_{u, min} = 25.88$, (d) *Forest Path*, $l_{max} = 4.38$; $l_{min} = 1.36$; $l_{u, max} = 572.81$; $l_{u, min} = 149.04$. (Horizontal axes of the graphs are normalized with respect to the max. and min. values of log. luminance and PU values.)

[1] gives lower MSE results than the proposed method, as this method is optimally designed to give minimum MSE for the logarithm of luminance values. However, the proposed method gives better PU-PSNR results as it minimizes the MSE between the perceptually uniform values. While the two methods give very similar HDR VDP results for *Atrium Night* and *Memorial*, the compared method [1] is better for *Coby*. It is not easy to judge a conclusion regarding the superiority of one method to the other with respect to HDR VDP results.



(a) with Log. Luminance [1]



(b) with proposed PU Encoding

Fig. 9: Produced LDR images for *Atrium Night*.



(a) with Log. Luminance [1]



(b) with proposed PU Encoding

Fig. 10: Produced LDR images for *Coby*.



(a) with Log. Luminance [1] (b) with proposed PU Encoding
Fig. 11: Produced LDR images for *Memorial*.



(a) with Log. Luminance [1] (b) with proposed PU Encoding
Fig. 12: Zoom-in of the darker regions for *Memorial*.

B. Experiments on LDR Image Quality Constraint Optimization

We select the global version of the photographic tone reproduction (PTR) in [28] as a reference TMO in our experiments. PTR [28] is a popular TMO with its better performance than others in terms of contrast loss and psycho-visual experiments [29]. Our main focus in the experiments is however to verify the methodology to compromise the trade-off between the HDR quality and LDR quality, rather than the selection of which TMO as a reference.

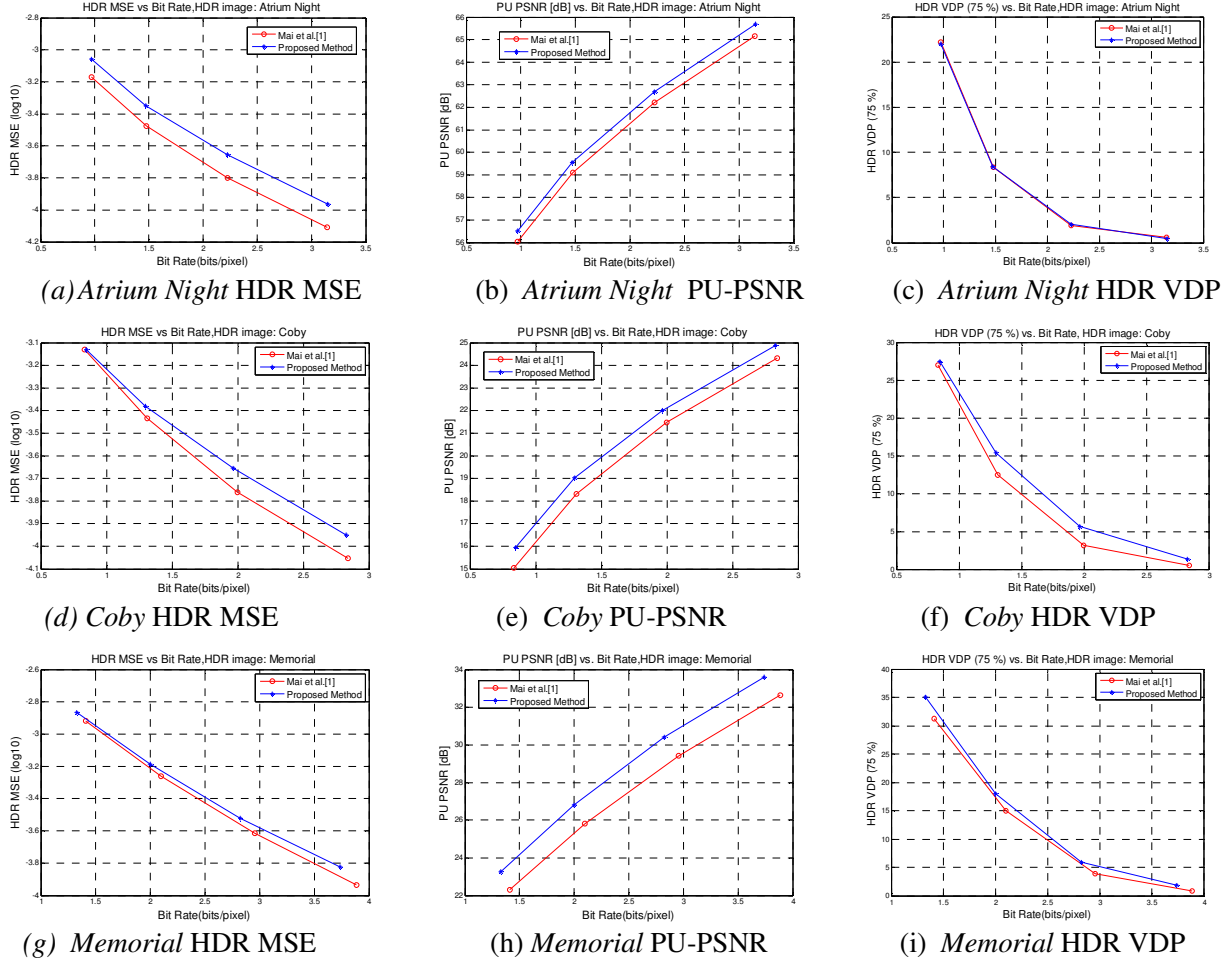


Fig. 13: HDR MSE (log10), PU-PSNR (dB) and HDR VDP (75 %) results for *Atrium Night*, *Memorial* and *Bristol Bridge* for the proposed method and compared method [1]. Some encoding parameters for the utilized H.264/AVC reference software JM 17.2 are as follows: ProfileIDC = 244; LevelIDC = 40; YUVformat=1 (i.e. 4:2:0).

Figure 14 shows the resulting tone mapping curves for the HDR test image *Atrium Night* and *Coby*. The λ parameter is adjusted such that the PSNR between the resulting and reference tone mapped images is 35 dB and 30 dB for two cases to be able to show the trade-off between the rate-distortion performance and the LDR image quality. Without loss of generality, other selections of PSNR values in a decreasing manner are also possible to validate such a trade-off. While λ is smaller, the resulting tone mapping curve for the LDR constrained case is closer to optimum tone mapping of Mai *et al.*[1]. When it is higher, the curve is becoming nearer to the reference TMO [28]. It should be noted that the λ value to

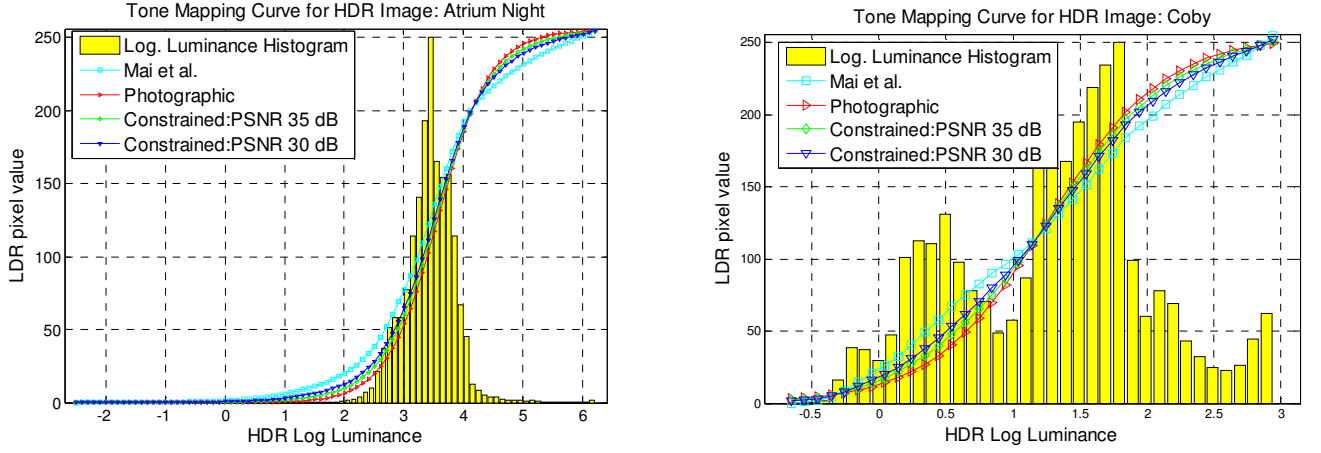


Fig. 14: TM curves for the optimized tone mapping of Mai [1], PTR [25], and the proposed methodology constraining the LDR quality. Yellow bar is the luminance (\log_{10}) histogram.

adjust a specific PSNR value depends on the content of the HDR image as the tone mapping of Mai et. al [1] and the reference photographic tone reproduction [28] are dependent on the luminance histogram and the maximum luminance of the HDR image. For instance, while 30 dB of PSNR is obtained when λ is equal to 1.02 for Coby, the same PSNR is obtained when λ is equal to 1.18 for Atrium Night.

Figure 15 and 16 illustrates the resulting tone mapped images for the mentioned four cases for *Atrium Night* and *Coby*. Some over-illumination spread over the whole image is observable for the Mai *et al.* [1] for *Atrium Night*. Some blurring kind of distortion is also noticeable for *Coby*. On the other hand, reference TMO [25] is providing more natural images with better contrast. The resulting tone mapped images with the proposed methodology provides a changing intermediate quality with respect to the value of λ .

Figure 17 shows the rate-distortion curves in terms of the coding bit rate of LDR image vs. the HDR-MSE. Mai *et al.* [1] gives the lower HDR-MSE results among all, as this method is optimally designed to give the minimum MSE for the logarithm of luminance values. For smaller λ value, the rate-distortion performance of the proposed methodology is closer to the optimized tone mapping [1]. However, the LDR quality is getting worse as a trade-off. As λ increases, the rate-distortion performance is getting poor by shifting towards to the one of reference TMO [28]. But, the quality of the generated tone mapped

images is increasing on the other hand as the produced LDR image is getting closer to the reference tone mapped image which provides a more natural view with a better contrast as indicated in Figure 15 and 16.

C. Experiments on Flickering Constraint Optimization

We use the publicly available *Sun* sequence [25] which is suitable for our flickering tests with its regions where a bright object (e.g. the Sun) is entering and leaving the camera view. Figure 18 illustrates examples for the flickering effect between the consecutive frames when the optimized tone mapping in [1] is individually applied to each frame. When the Sun is entering to the scene in the consecutive frame, the average brightness in the regions other than the Sun is decreasing significantly in contrary to the actual values. The proposed method is constraining this effect by holding the change in the average values of the consecutive frames in the JND interval.



(a) Mai et al. [1]



(b) Proposed (λ : 1.02, PSNR: 30 dB)



(c) Proposed (λ : 2.60, PSNR: 35 dB)



(d) Photographic Tone Reproduct. [28]

Fig. 15: Produced LDR images for Coby.



(a) Mai et al. [1]



(b) Proposed (λ : 1.18, PSNR: 30 dB)



(c) Proposed (λ : 2.88, PSNR: 35 dB)



(d) Photographic Tone Reprod. [28]

Fig. 16: Produced LDR images for Atrium Night.

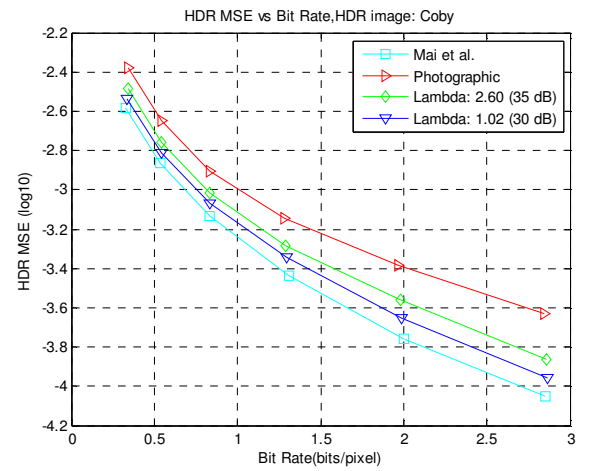
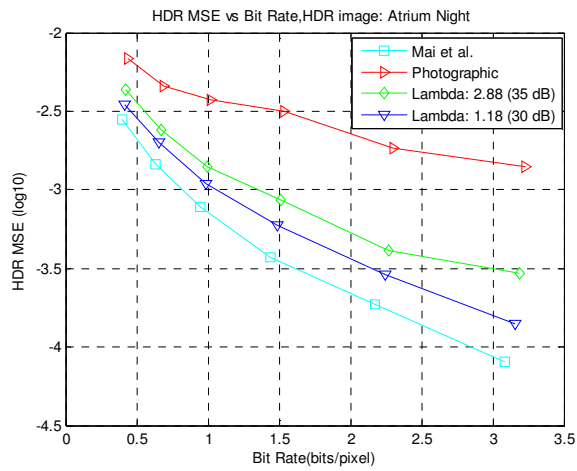


Fig. 17: HDR MSE (log10) vs. Coding Bitrate results for *Atrium Night* (left) and *Coby* (right)

Figure 19 shows the changes in the average intensity of the frames for the LDR video generated for the three clips from *Sun* sequence by Mai et al [1], which is applying the tone mapping operation separately to each frame of the HDR video without any flickering compensation. The figure also shows the average intensity changes after the proposed flickering compensation with different Weber measures for the same three clips. The proposed method is successfully controlling the change in the average values with respect to the given *Weber* constant k_w and generating a smooth transition between frames.

Figure 20 gives the rate-distortion curves in terms of coding bit rate of generated LDR video vs. the HDR-MSE. As expected, optimized tone mapping of Mai *et al.* [1] without any flickering correction gives the lowest HDR-MSE results among all. When the flickering is constrained, the rate-distortion performance is decreasing as a trade-off as the tone mapping is shifting away from the optimized tone mapping of Mai et al. [1] as indicated in (22). As the constraint is becoming stronger with smaller Weber measure k_w , the decline in the rate-distortion performance is getting higher. Note that the flickering effect that the proposed method handles mainly emerges as DC shifts. Therefore, the proposed method which is having less flickering effect does not have a better compression efficiency as the existing coding standards including also the H.264 are efficiently adaptable to DC shifts without a significant bitrate loss.

7. CONCLUSIONS

We have first revealed in this paper that an optimized tone mapping [1] which is designed to minimize the MSE between the logarithms of luminance values is not optimum in perceptual sense in two aspects. First, the produced LDR images can show some over illumination in the bright regions. Second, the quality of the reconstructed images can be further improved with respect to the perceptual metrics. The proposed method based on the perceptually uniform encoding of luminance values eliminates such a disadvantage and gives better results in terms of the perceptual metric, PU-PSNR. However, the proposed method shows also some limitations in revealing the contrast in dark regions in contrast to the compared method [1]. As a second improvement, the basic trade-off between the LDR image quality and HDR image quality for HDR image and video coding is verified by experiments. By putting the LDR image

quality constraint into the optimized tone mapping problem formulated by Mai. *et al.* [1], we develop an methodology to set this trade-off. Finally, we reveal the compromise between the rate distortion performance of HDR video coding and the flickering distortion that can occur on the resulting video due to the sudden changes in the average values of consecutive tone mapped frames. By putting a flickering constraint into the problem, we develop a methodology to solve the flickering problem in an optimized way. Future work will focus on the LDR image quality constraint optimization problem for local TMOs and pixel-wise flickering distortions that can occur on the tone mapped video sequences.

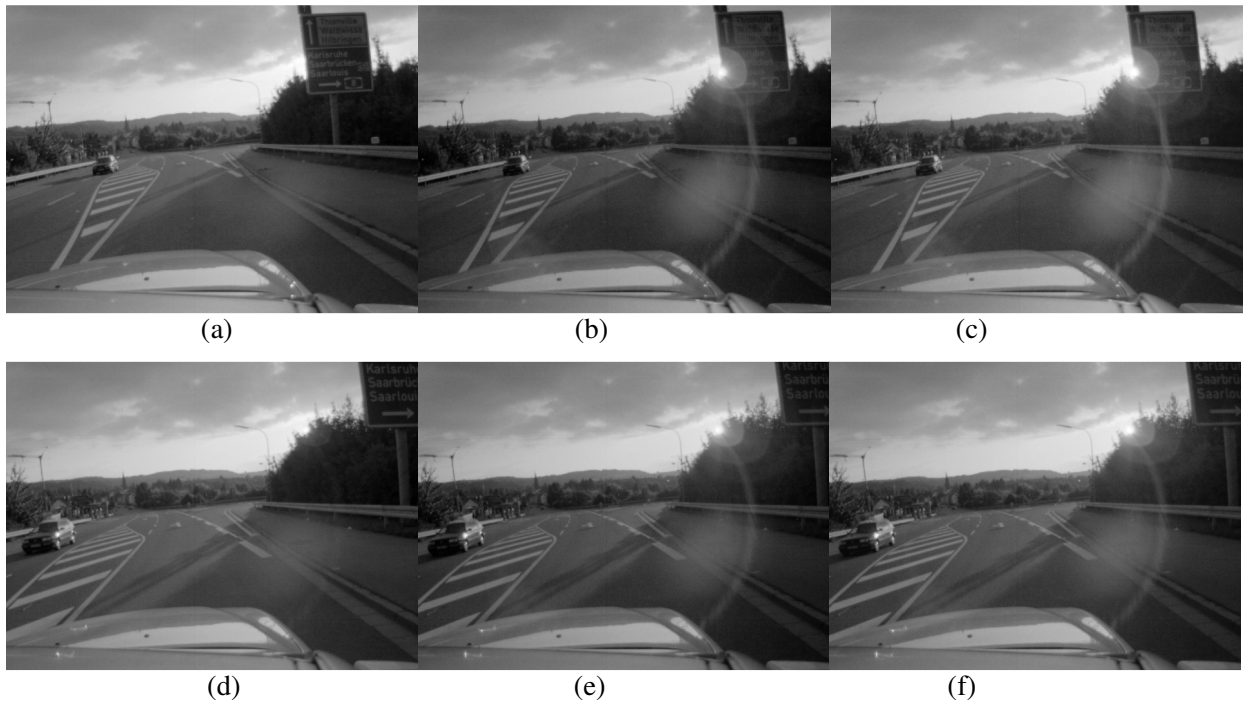


Fig. 18: Tone mapped frames from *Sun* sequence: a) Frame 222 (Average Intensity (AI): 104.43), b) Frame 223 before flickering correction (AI: 90.64), c) Frame 223 after flickering correction (AI: 103.64), d) Frame 242 (AI: 101.57), e) Frame 243 before flickering correction (AI: 93.18), f) Frame 243 after flickering correction (AI: 100.18). ($k_w = 0.01$)

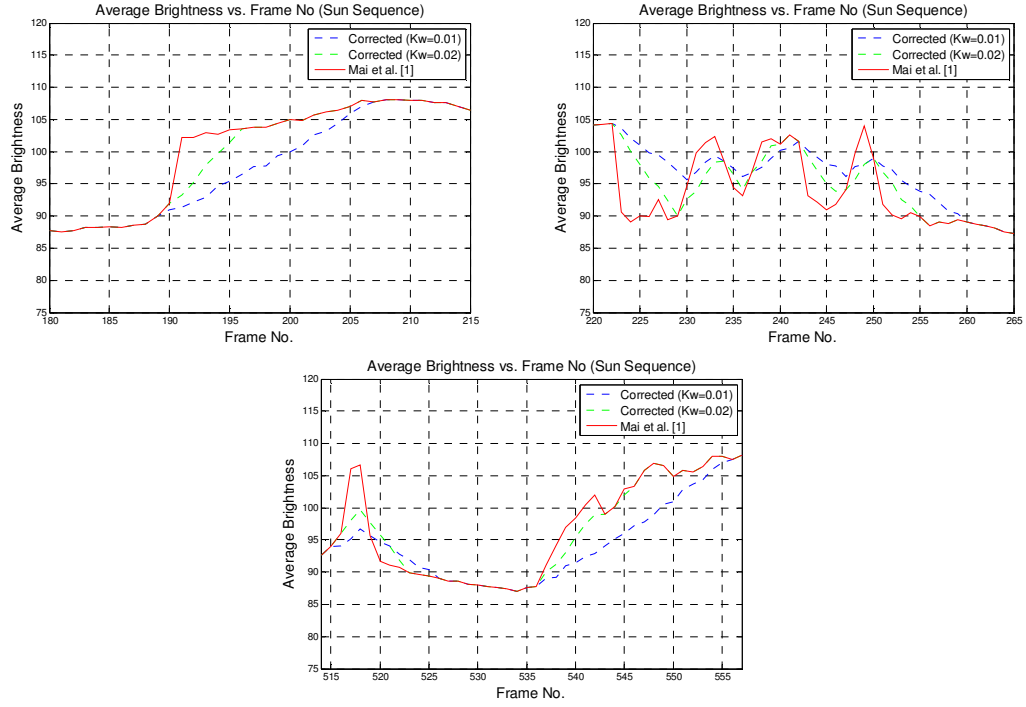


Fig. 19: Average intensity changes for three clips from *Sun* sequence before and after flickering correction ($k_w = 0.01$ and $k_w = 0.02$).

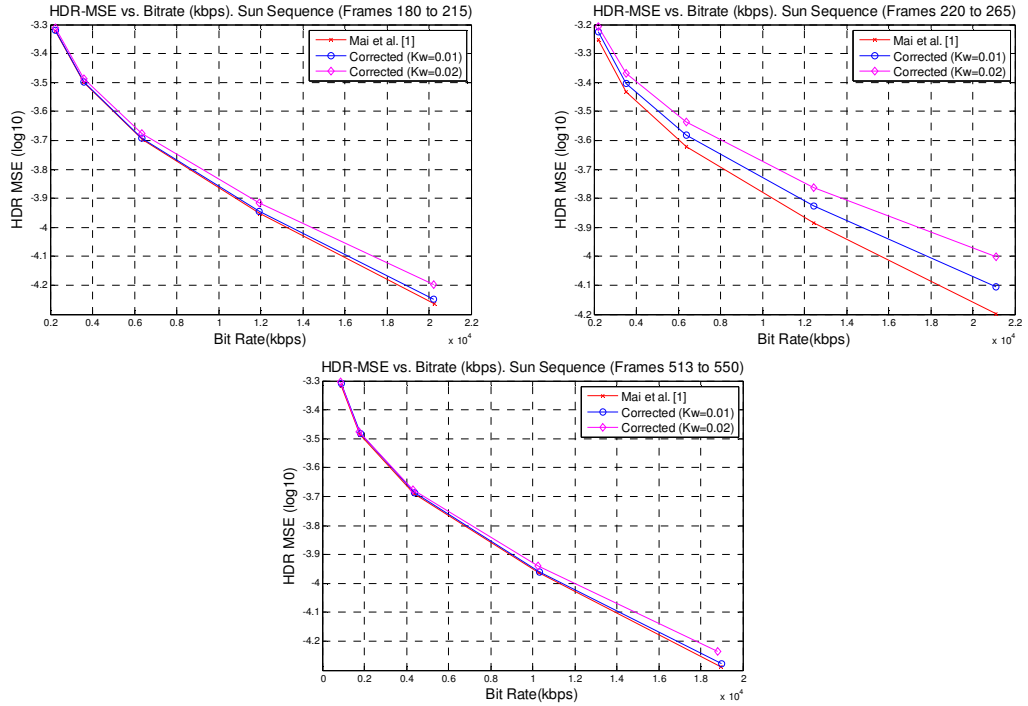


Fig. 20 HDR MSE (log10) vs. Coding Bitrate (kbps) graphs for three clips from *Sun* sequence before and after flickering correction. Some encoding parameters for the utilized H.264/AVC reference software JM 17.2 are as follows: ProfileIDC = 244; LevelIDC = 40; YUVformat=1 (i.e. 4:2:0).

ACKNOWLEDGMENTS

This work was partly funded by the *Région Ile de France* in the framework of the *NEVE*x project.

REFERENCES

- [1] Z. Mai, H. Mansour, R. Mantiuk, P. Nasiopoulos, R. Ward, and W. Heidrich, “Optimizing a Tone Curve for Backward-Compatible High Dynamic Range Image and Video Compression”, *IEEE Transactions on Image Processing*, vol. 20, no. 6, pp. 1558-1571, June 2011.
- [2] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced High Dynamic Range Imaging*, 278 pages by AK Peters (now CRC Press), 2011.
- [3] R. Mantiuk, G. Krawczyk, R. Mantiuk and H.-P. Seidel, “High dynamic range imaging pipeline: perception-motivated representation of visual content”, in *SPIE Proc. Human Vision and Electronic Imaging XII*, vol. 6492, 2007.
- [4] A. Koz and F. Dufaux, “A Comparative Survey on High Dynamic Range Video Compression”, in *SPIE Proc. Applications of Digital Image Processing XXXV*, vol. 8499, San Diego, CA, August 2012.
- [5] S. Liu, W.-S. Kim, and A. Vetro, “Bit-depth scalable coding for high dynamic range video,” in *Proc. SPIE Vis. Commun. Image Process.*, vol. 6822, 2008.
- [6] SIM2 Engineering Display series [Online]. Available: http://www.sim2.com/HDR/hdrdisplay/hdr47e_s_4k
- [7] Digital cinematography [Online]. Available: http://en.wikipedia.org/wiki/Digital_cinematography
- [8] EU5 Smartphone Penetration [Online]. Available: http://www.comscore.com/Insights/Press_Releases/2012/12/EU5_Smartphone_Penetration_Reaches_55_Percent_in_October_2012
- [9] A. Motra and H. Thoma, “An Adaptive LogLuv Transform for High Dynamic Range Video Compression”, pp. 2061-2064, in *Proc. IEEE ICIP*, 2010.
- [10] J.-U. Garbas and H. Tahoma, “Temporally Coherent Luminance-to-Luma Mapping for High Dynamic Range Video Coding with H.264/AVC”, pp. 829 – 832, in *Proc. IEEE ICASSP*, 2011.
- [11] Y. Zhang, E. Reinhard and D. Bull “Perception-based High Dynamic Range Video Compression with Optimal Bit Depth Transformation”, pp. 1321-1324, in *Proc. IEEE ICIP*, 2011.
- [12] A. Segall, “Scalable Coding of High Dynamic Range Video”, vol. 1, pp. 1-4, in *Proc. IEEE ICIP*, 2007.
- [13] R. Mantiuk, A. Efremov, K. Myszkowski, and H.-P. Seidel, “Backward compatible high dynamic range MPEG video compression,” *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 25, no. 3, pp. 713–723, 2006.
- [14] R. Mantiuk, G. Krawczyk, K. Myszkowski, and H.-P. Seidel, “Perception-motivated high dynamic range video encoding,” *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 23, no. 3, pp. 730–738, 2004.
- [15] A. Segall, L. Kerofsky, and S. Lei, “Tone Mapping SEI Message,” Jul. 2006, Tech. Rep., ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT-T060.
- [16] T. O. Aydin, R. Mantiuk, and H.-P. Seidel, “Extending Quality Metrics to Full Luminance Range Images”, *Human Vision and Electronic Imaging XIII (Proceedings of SPIE)*, pp. 6806-10, Jan. 2008.

- [17] Z. Mai, H. Mansour, P. Nasiopoulos, and R. Ward, "Visually-Favorable Tone-Mapping with High Compression Performance", pp.1285-1288, in Proc. IEEE ICIP, 2010.
- [18] B. Guthier et al., "Flicker Reduction in Tone Mapped High Dynamic Range Video". Proc. SPIE 7866, Color Imaging XVI: Displaying, Processing, Hardcopy, and Applications, January 2011.
- [19] Weber Law [Online]. Available:
<http://sunburst.usd.edu/~schieber/coglab/WebersLaw.html>
- [20] G. W. Larson, "Logluv encoding for full-gamut, high-dynamic range images," *J. Graph. Tools*, vol. 3, no. 1, pp. 15–31, 1998.
- [21] F. Banterle, K. Debattista, P. Ledda, and A. Chalmers, "A GPU-friendly method for high dynamic range texture compression using inverse tone mapping", pp. 41-48, in Proc. of graphics interface, 2008.
- [22] S. Daly, "The visible differences predictor: an algorithm for the assessment of image fidelity", Digital images and human vision, pp. 179-206, MIT Press Cambridge, MA, USA, 1993.
- [23] HDR image Gallery [online]. Available: <http://www.mpi-inf.mpg.de/resources/hdr/gallery.html>
- [24] HDR image Gallery [online]. Available: <http://www.anywhere.com/gward/hdrenc/pages/originals.html>
- [25] HDR Video Samples [Online]. Available: <http://www.mpi-inf.mpg.de/resources/hdr/video/>
- [26] H.264/AVC JM17.2 Reference Software [Online]. Available: <http://iphome.hhi.de/suehring/tml/>
- [27] R. Mantiuk, S. J. Daly, K. Myszkowski, and H.-P. Seidel, "Predicting visible differences in high dynamic range images – model and its calibration", in SPIE Human Vision and Electronic Imaging X, vol. 5666, pp. 204-214, San Jose, USA, 2005.
- [28] E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, "Photographic Tone Reproduction for Digital Images", ACM Transactions on Graphics, vol. 21, issue 3, pp 267-276, July 2002.
- [29] P. Ledda, A. Chalmers, T. Troscianko and H. Seetzen, "Evaluation of tone mapping operators using a high dynamic range display", ACM Transactions on Graphics, vol. 24, issue 3, pp. 640-648, July 2005.