HAL open science

Optimized Tone Mapping with Flickering Constraint for Backward-Compatible High Dynamic Range Video Coding

Alper Koz, Frederic Dufaux

To cite this version:

Alper Koz, Frederic Dufaux. Optimized Tone Mapping with Flickering Constraint for Backward-Compatible High Dynamic Range Video Coding. Int. Workshop on Image and Audio Analysis for Multimedia Interactive Services (WIAMIS'2013), Jul 2013, Paris, France. hal-01436226

HAL Id: hal-01436226

https://hal.science/hal-01436226v1

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.

OPTIMIZED TONE MAPPING WITH FLICKERING CONSTRAINT FOR BACKWARD-COMPATIBLE HIGH DYNAMIC RANGE VIDEO CODING

Alper Koz and *Frederic Dufaux*

Télécom ParisTech, Dept. Traitement du Signal et des Images, 75014 Paris, France.
{alper.koz, frederic.dufaux}@telecom-paristech.fr

ABSTRACT

Backward compatibility to low dynamic range (LDR) displays is an important requirement for high dynamic range (HDR) image and video coding in order to enable a successful transition to HDR technology. In a recent work [1], an optimized solution for tone mapping and inverse tone mapping of HDR images and video frames is achieved in terms of the mean square error (MSE) of the logarithm of luminance values of HDR pixels for backward-compatible compression. Although this pioneer optimization approach provides a well settled mathematical framework for tone mapping, one of its important shortcomings is the flickering distortions that can occur on the resulting video due to the individual tone mapping applied to each frame. In this paper, we include the flickering distortion level as a constraint to optimization problem and develop a methodology to compromise the trade-off between the rate-distortion performance and flickering effect. The developed methodology is validated on HDR video sequences by showing the changes in rate-distortion performances with respect to the different flickering levels.

***Index Terms*—** High Dynamic Range Video Compression, Tone Mapping, Constrained Optimization, Flickering

1. INTRODUCTION

High dynamic range (HDR) imaging has been developed in the last decade to encode the entire luminance range of real world scenes ranging from extreme darkness (10^{-6} cd/ m^2) to bright sunshine (10^8 cd/ m^2). The main goal of this technology is to capture and store the accurate physical luminance values rather than 8-bit intensities and to generate a scene referred image independent of display devices. Such a technology is expected to eliminate the limitations of standard 8-bit representations for the new generation HDR displays, and would offer a much better perceived image quality in many applications ranging from digital cinema to medical imaging [2].

A major challenge for HDR imaging is however an efficient compression scheme, in particular for HDR video which requires significantly more storage size and transmission bandwidth than standard 8-bit LDR video. The previous work has mainly handled this challenge in two ways. The first approach [3]-[5] is to take the advantage of high profile of existing state-of-the-art H.264/AVC (Advanced Video Coding) codec. As the maximum bit-depth supported in high-profile mode is 14 bits, the core part of this approach was to convert 16 bits HDR video into 14 bits. In [3], this conversion is performed by applying the modified version of LogLuv transform [6], which maps the luminance range of each HDR frame to the range [0, 2^{14}-1] with respect to the max. and min. luminance values of the frame. In the extended version of this work [4], temporal motion compensation problems generated by individual mapping of frames is solved by weighted prediction (WP) tool of H.264/AVC [7]. In a similar work [5], such a mapping is enhanced by smoothing the HDR frames with respect to the Human Visual System (HVS) sensitivities in different wavelet resolutions. The advantage of these methods was their conceptual simplicity. On the down side, they do not support the backward compatibility with 8-bit displays and are far from satisfying the real time requirement as the encoding complexity of the H.264/AVC high profile is still high.

The second approach [1],[8]-[11] aims to provide the backward compatibility with existing standard 8-bit displays in order to enable a successful transition to HDR technology. 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 [8], a block wise approach is proposed for interlayer prediction. A linear multiplication and an addition of offset are applied to each block of LDR frame to estimate the corresponding block of HDR frame. This method is improved in [9] by differentially encoding the offsets with respect to the previously encoded left and upper blocks. In a recent work [1], block wise prediction is replaced by a global (inverse) tone mapping operation (TMO), and the optimum prediction with respect to the mean square error (MSE) of the logarithm of luminance values is formulated in a detailed and complete manner.

Although the approach in [1] provides a well settled mathematical framework for optimum tone mapping, one of its important shortcomings is the direct adaptation of the method to the video by means of applying an individual tone mapping to each frame of the HDR video sequence. Such an approach results in sudden changes in the average brightness of the consecutive frames, when a high luminance object or region enters or leaves the camera view. In this paper, we include such a flicker effect in the average brightness of the LDR frames as a constraint to the optimized tone mapping problem and derive a methodology to compromise the trade-off between the rate-distortion performance and flickering distortion.

In the next section, we have given the overview of the optimized tone mapping developed in [1]. Section 3 presents the tone mapping optimization problem with the flickering constraint and derives a closed form solution to set the trade-off between the rate-distortion performance of HDR video coding and flickering distortion. Section 4 gives the experimental results for the proposed methodology and Section 5 concludes the paper.

2. OPTIMIZED TONE MAPPING FOR BACKWARD COMPATIBLE HDR VIDEO COMPRESSION

A general scheme for backward compatible HDR video compression is illustrated in Figure 1 [1]. The HDR video is first mapped into 8-bit by using a tone mapping operator. The resulting LDR video goes through the standard video encoding and decoding process (H.264/AVC). Then, an HDR video is reconstructed from the LDR video by applying an inverse TMO. The parameters of the TMO can be stored as a look-up table [1] or can be encoded using supplemental enhancement information [12]. The residue can also be further encoded as an enhancement layer in the bitstream.

Mai *et al.* [1] in their recent work have developed the optimum tone mapping and inverse tone mapping which minimizes the MSE between the logarithmic ($\log_{10}$) values of the luminance of original HDR content and its reconstructed version. In order to find a computationally efficient solution, they 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 the tone mapping curve based on the luminance histogram of the HDR video frames.

Given the notations l and v corresponding to the logarithm of the luminance of an HDR frame and the pixel values of its tone mapped LDR version respectively, the tone mapping curve in [1] is first parameterized as a piece-wise linear function with the nodes (l_k, v_k) as shown in Fig. 2.a. 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

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

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

$$\arg\min_{s_1 \ldots s_N} \sum_{k=1}^{N} \frac{p_k}{s_k^2}$$

$$\text{subject to} \sum_{k=1}^{N} s_k = \frac{v_{\max}}{\delta} \quad (2),$$

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

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

A tone mapping characterized by the slopes in (3) minimizes the MSE between the original and reconstructed HDR frames. However, a direct application of such a tone mapping separately to each frame of an HDR video sequence can produce sudden changes in the average brightness of the produced LDR frames, which results in a flickering distortion on the produced LDR video. In the next section, we will reformulate this optimization problem with such a flickering constraint.

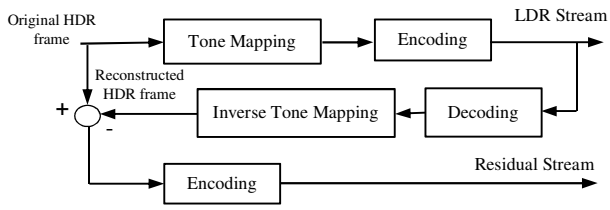


Fig. 1 General scheme for backward compatible HDR video compression [1].

3. OPTIMIZED TONE MAPPING WITH FLICKERING CONSTRAINT

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:

$$k_w = \frac{\Delta L}{L} \quad (4).$$

Such a relation says that the size of the JND (ΔL) is a constant proportion (k_w) of the initial stimulus luminance (L) [13].

Guthier *et al.* [14] in a recent work use the *Weber* law as a perceptual basis 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. Their research validate that the Weber constant (k_w) is a good measure for the flickering distortion on the resulting LDR video [14]. We will use such a measure to 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 (3) for the n-th frame of the video (Fig. 2. b).
- g Optimized tone mapping we are searching for the n-th frame with the flickering constraint (Fig. 2. b).
- L^s The average intensity of the produced LDR frame with s.
- 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 $k_w L_{n-1}$.

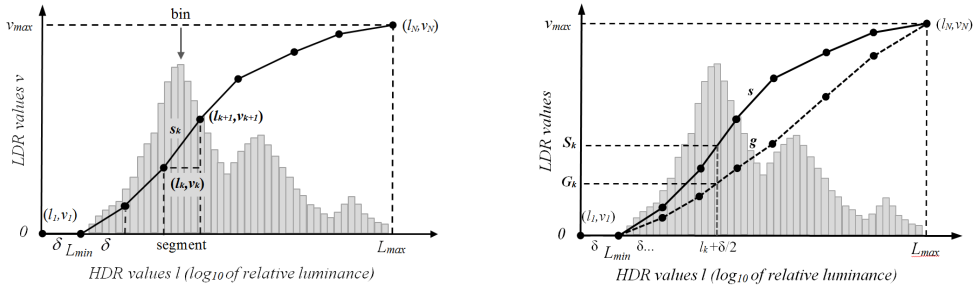
In order to avoid a flicker effect between the *(n-1)*-th and n-th frames, the average intensity of the resulting LDR frame (L^g) should be in the JND (ΔL_{n-1}) interval of L_{n-1}:

$$L^g = \begin{cases} L_{n-1} + k_w L_{n-1}, \text{ if } L^s > (L_{n-1} + k_w L_{n-1}) \\ L_{n-1} - k_w L_{n-1}, \text{ if } L^s < (L_{n-1} - k_w L_{n-1}) \\ L^s \qquad\qquad \text{, otherwise} \end{cases} \quad (5).$$

In terms of the histogram values, the average brightness L^g can be expressed as

$$L^g \cong \sum_{k=1}^{N} p_k g(l_k + \delta/2) = \sum_{k=1}^{N} p_k G_k \quad (6)$$

where G_k corresponds to the tone mapped value for the average luminance of the k-th bin of the histogram (Fig. 2.b).



(a) (b)

Fig. 2 (a) Parameterization of the TMO curve in [1]. The bar-plot is the luminance histogram of the HDR frame, (b) illustration for the optimum tone mapping of Mai et al. (*s*) and the flicker constrained optimized tone mapping (*g*).

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*:

$$\underset{G_1 \ldots G_N}{\arg\min} \sum_{k=1}^{N} p_k (G_k - S_k)^2$$
$$\text{subject to} \sum_{k=1}^{N} p_k . G_k = L^g \tag{7}$$

where $S_k = s(l_k + \delta/2)$ (Fig. 2.b).

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

$$G_k = S_k - \frac{\lambda}{2}, \quad k = 1..N$$
$$\lambda = 2(L^s - L^g) \tag{8}$$

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 (3) 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 (5) 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 (8) 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. (5)) 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 validation of this trade-off in the next section.

4. EXPERIMENTAL RESULTS

We use the publicly available *Sun* sequence [15] which is suitable for our tests with its regions where a bright object (e.g. the Sun) is entering and leaving the camera view. Figure 3 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.

Figure 4 shows the changes in the average intensity of the frames before and after flickering correction for the three clips from *Sun* sequence. 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 5 shows the rate-distortion curves in terms of coding bit rate of generated LDR video vs. the HDR-MSE, the average MSE between the logarithm of the luminance values of original and reconstructed HDR frames. H.264/AVC reference software JM 17.2 [16] is used for the encoding of produced LDR video. The supplemental information for the characterization of tone mapping curves is very negligible compared to the coding bitrate. 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 constraint is becoming stronger with smaller Weber measure k_w, the decline in the rate-distortion performance is getting higher.

5. CONCLUSIONS

We reveal the basic trade-off between the rate distortion performance of backward compatible HDR video coding and the flickering distortion that can occur on the tone mapped video. By putting a flickering constraint into the optimized tone mapping problem described in [1], we develop a methodology to arrange and validate this trade-off. As a secondary conclusion, the research also indicates that the AVC standard is not sufficient for the coding of HDR video sequences and a revised standard description would be required in the long term. Future work will focus on the extension of the problem to pixel-wise flickering distortions.

ACKNOWLEDMENTS

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

(a) (b) (c)

Fig. 3 Tone mapped frames from *Sun* sequence: a) Frame 222 (Average Intensity: 104.43), b) Frame 223 before flickering correction (Average Intensity: 90.64), c) Frame 223 after flickering correction (Average Intensity: 103.64, k_w= 0. 01).

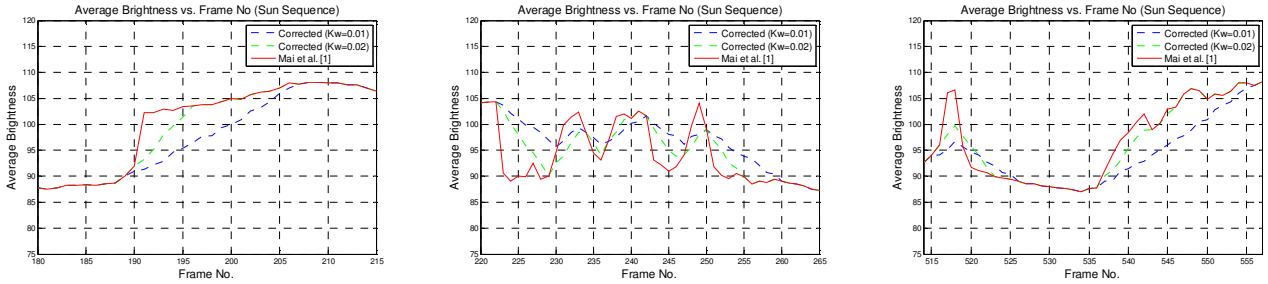


Fig. 4 Average intensity changes for three clips from *Sun* sequence before and after flickering correction (k_w= 0. 01 and k_w= 0. 02).

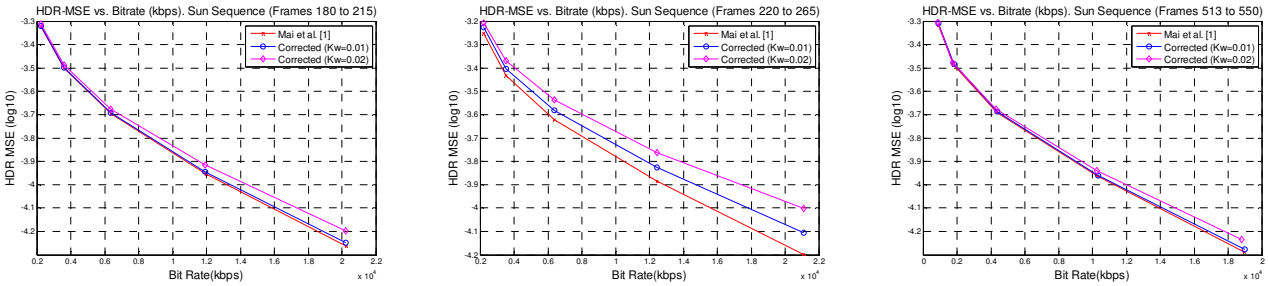


Fig. 5 HDR MSE (log10) vs. Coding Bitrate (kbps) graphs for three clips from *Sun* sequence before and after flickering correction. Some encoding parameters for the utilized H.264/AVC reference software JM 17.2 are as follows: ProfileIDC = 244; LevelIDC = 40; YUVformat=1 (i.e. 4:2:0).

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, pp. 1558-1571, Vol. 20, No. 6, June 2011.

[2] Francesco Banterle, Alessandro Artusi, Kurt Debattista, and Alan Chalmers, *Advanced High Dynamic Range Imaging,* 278 pages, by AK Peters (now CRC Press), 2011.

[3] A. Motra and H. Thoma, "An Adaptive LogLuv Transform for High Dynamic Range Video Compression", pp. 2061-2064, in Proc. IEEE ICIP, 2010.

[4] 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.

[5] 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.

[6] G. W. Larson, "Logluv encoding for full-gamut, high-dynamic range images," *J. Graph. Tools*, vol. 3, no. 1, pp. 15–31, 1998.

[7] J.M. Boyce, "Weighted prediction in the H.264/MPEG AVC video coding standard," pp. 789-792, in Proc. Int. Symposium on Circuits and Systems (ISCAS), May 2004.

[8] A. Segall, "Scalable Coding of High Dynamic Range Video", vol. 1, in Proc. IEEE ICIP, 2007.

[9] 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.

[10] A. Koz and F. Dufaux, "A Comparative Survey on High Dynamic Range Video Compression", in SPIE Proc. Applications of Digital Image Processing XXXV, August 2012.

[11] A. Koz and F. Dufaux, "Optimized Tone Mapping with Perceptually Uniform Luminance Values for backward-compatible High Dynamic Range Video Compression", in IEEE VCIP, 2012.

[12] 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.

[13] Weber Law [Online]. Available: http://sunburst.usd.edu/~schieber/coglab/WebersLaw.html

[14] 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).

[15] HDR Video Samples [online]. Available: http://www.mpi-inf.mpg.de/resources/hdr/video/

[16] H.264/AVC JM17.2 Reference Software [Online]. Available: http://iphome.hhi.de/suehring/tml/