HAL
open science

Optimized Tone Mapping With LDR Image Quality Constraint for Backward-Compatible High Dynamic Range Image and Video Coding

Alper Koz, Frederic Dufaux

▶ **To cite this version:**

Alper Koz, Frederic Dufaux. Optimized Tone Mapping With LDR Image Quality Constraint for Backward-Compatible High Dynamic Range Image and Video Coding. IEEE International Conference on Image Processing (ICIP'2013), IEEE, Sep 2013, Melbourne, Australia. hal-01436223

HAL Id: hal-01436223

https://hal.science/hal-01436223v1

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 LDR IMAGE QUALITY CONSTRAINT FOR BACKWARD-COMPATIBLE HIGH DYNAMIC RANGE IMAGE AND 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 is achieved in terms of mean square error (MSE) of the logarithm of luminance values of HDR image 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 not to take the quality of the resulting LDR images into account during the formulation. In this paper, we include the LDR image quality as a constraint to optimization problem and develop a methodology to compromise the trade-off between HDR image quality and LDR image quality during HDR image and video coding. The developed methodology is verified on HDR images by showing the increase (decrease) in the quality of generated LDR images while losing (gaining) from the rate-distortion performance of HDR image coding.

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

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 plasma and HDR displays, and would offer a much better perceived image quality in many applications ranging from digital cinema to medical imaging [1]-[2].

A major challenge for HDR imaging is however an efficient compression scheme, in particular for HDR video which requires significantly more storage size 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]-[9] 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 (TM) operation, 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 this pioneer approach provides a well settled mathematical framework for optimum tone mapping, one of its important shortcomings is not to take the quality of the resulting LDR images into account during the formulation. Consequently, the proposed method by Mai et al. [1] can produce some non-natural over illumination in the bright regions of the produced LDR images [10], which indicates that an optimized tone mapping with respect to the HDR metric (MSE) is not optimum in the LDR image quality sense. This pioneer study [1] is extended in a previously published work of the same researchers [11] by means of considering the LDR image quality and the bitrates of LDR layer and residue coding. However, they do not provide a practical methodology and a closed form solution to arrange the trade-off between HDR and LDR image qualities during the coding.

In this paper, we include the LDR image quality as a constraint to optimized tone mapping problem. We formulate the problem as finding the optimum tone mapping and inverse 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). Rather than

the selection of which TMO as a reference, our focus in this paper would be to develop a generic practical methodology to compromise the trade-off between the HDR image quality and LDR image quality during the coding for a given reference TMO.

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 LDR image quality constraint. In this section, we first give the direct formulation of the problem and discuss its applicability for coding applications. Then, we relax the cost function and derive a closed form solution to arrange the trade-off between HDR image quality and LDR image quality. Section 4 gives the experimental results for the proposed methodology in terms of the rate-distortion curves and the quality of produced LDR images. Finally, the conclusions are given in section 5.

2. OPTIMUM 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 TMO. 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 TMO 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 HDR frame and the pixel values of the 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 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

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

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

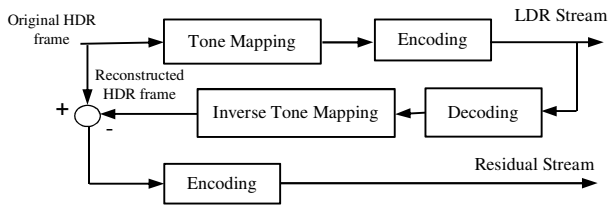


Fig. 1 General scheme for backward compatible HDR video compression [1].

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

A tone mapping characterized by the slopes in (3) minimizes the MSE between the original and reconstructed HDR frames. However, in contrary to the main goal of tone mapping research, such a formulation ignores the quality of produced LDR images.

3. TONE MAPPING WITH LDR IMAGE QUALITY CONSTRAINT

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.

Let us assume that h represents a reference global tone mapping operator and s is the optimized TMO we are searching (see Fig. 2.b.). 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

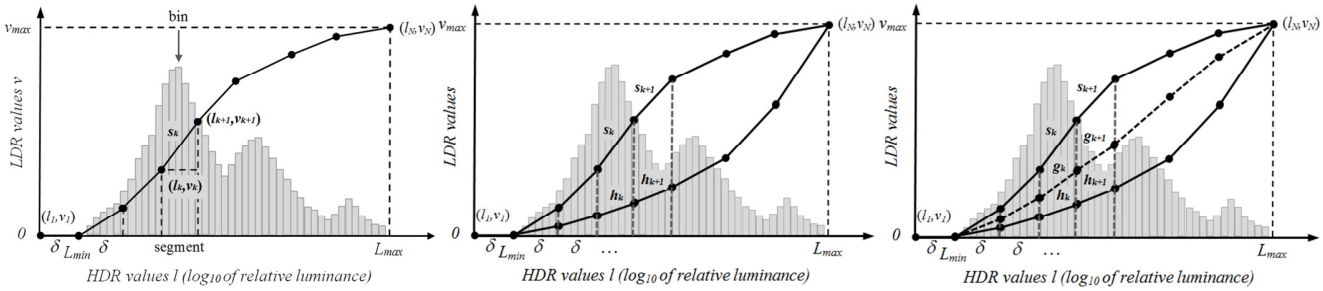
$$\begin{aligned} E\left((s(l)-h(l))^2\right) &\cong \sum_{k=1}^{N} \left(\sum_{n=1}^{k} (\delta . s_n) - \sum_{n=1}^{k} (\delta . h_n) \right)^2 . p_k \\ &= \delta^2 \sum_{k=1}^{N} \left(\sum_{n=1}^{k} (s_n - h_n) \right)^2 . p_k \end{aligned} \quad (4),$$

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 (2) returns into a double constraint problem as

$$\begin{aligned} &\underset{s_1 \ldots s_N}{\arg\min} \sum_{k=1}^{N} \frac{p_k}{s_k^2} \\ &\text{subject to } \sum_{k=1}^{N} s_k = \frac{v_{\max}}{\delta} \\ &\delta^2 \sum_{k=1}^{N} \left(\sum_{n=1}^{k} (s_n - h_n) \right)^2 . p_k = \Delta \end{aligned} \quad (5).$$

After defining the *Lagrangian* and performing some algebra on the derivative of the MSE in (4), the KKT conditions for the above optimization problem can be reduced to the following set of equations:

$$\frac{-2 p_k}{s_k^3} + \lambda_1 + \lambda_2 \delta^2 \sum_{a=k}^{N} \left(\sum_{n=1}^{a} 2(s_n - h_n) \right) . p_a = 0, \quad k = 1..N$$



(a) (b) (c)

Fig. 2 (a) Parameterization of the TM curve in [1]. The bar-plot is the luminance histogram of the HDR frame, (b) the searched optimum TM with the LDR constraint (s) and the reference TM operator (h) in (5), (c) 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 (7).

$$\sum_{k=1}^{N} s_k - \frac{v_{\max}}{\delta} = 0$$
$$\delta^2 \sum_{k=1}^{N} \left(\sum_{n=1}^{k} (s_n - h_n) \right)^2 .p_k - \Delta = 0 \qquad (6)$$

where λ_1 and λ_2 are the Lagrangian coefficients for the two constraints in (5). 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 loaded for coding applications. Considering these facts, we 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 (3) 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 (3), the reference global tone mapping operation, and the tone mapping operation we are looking for, respectively (see Fig. 2.c). Then, the new optimization problem in terms of the *MSE*s can be expressed as:

$$\underset{g_1 \cdots g_N}{\arg\min} \quad \delta^2 \sum_{k=1}^{N} \left(\sum_{n=1}^{k} (s_n - g_n) \right)^2 .p_k$$
$$\text{subject to} \quad \delta^2 \sum_{k=1}^{N} \left(\sum_{n=1}^{k} (g_n - h_n) \right)^2 .p_k = \Delta \qquad (7).$$

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:

$$\delta^2 \sum_{a=k}^{N} \left(\sum_{n=1}^{a} 2(g_n - s_n + \lambda(g_n - h_n)) \right) .p_a = 0, \quad k = 1..N$$
$$\delta^2 \sum_{k=1}^{N} \left(\sum_{n=1}^{k} (g_n - h_n) \right)^2 .p_k - \Delta = 0 \qquad (8)$$

where λ is the Lagrangian coefficient for the constraint. The solution to the set of KKT conditions is finalized as:

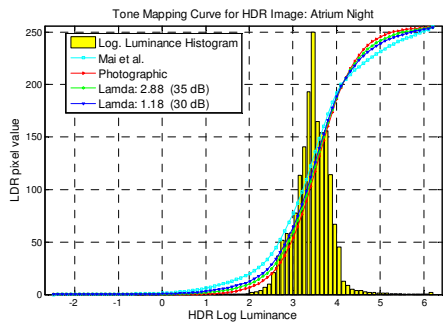


Fig. 3 TM curves for the optimized tone mapping of Mai [1], PTR [13], and the proposed methodology constraining the LDR quality. Yellow bar is the luminance ($\log_{10}$) histogram.

$$g_k = \frac{s_k + \lambda h_k}{1+\lambda} = h_k + \frac{(s_k - h_k)}{1+\lambda}, \quad k = 1..N$$
$$\lambda = \sqrt{\frac{\delta^2 \sum_{k=1}^{N} \left(\sum_{n=1}^{k} (s_n - h_n) \right)^2 .p_k}{\Delta}} - 1 = \sqrt{\frac{E\left((s(l) - h(l))^2\right)}{\Delta}} - 1 \qquad (9).$$

The λ parameter in such a closed form solution arranges the trade-off between the HDR and LDR image qualities 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 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 in the next section.

4. EXPERIMENTAL RESULTS

We select the global version of the photographic tone reproduction (PTR) in [13] as a reference TMO in our experiments. PTR [13] is a popular TMO with its better performance than others in terms of contrast loss [14] and psycho-visual experiments [15]. We use the HDR image database in [16]-[17] in our tests. Although we only present the results here for *Atrium Night* and *Coby* due to the lack of space, similar conclusions are also achieved for the other common images in these databases.

Figure 3 shows the resulting tone mapping curves for the HDR test image *Atrium Night* [16]-[17]. 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. 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 [13].

Figure 4 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 [13] 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.

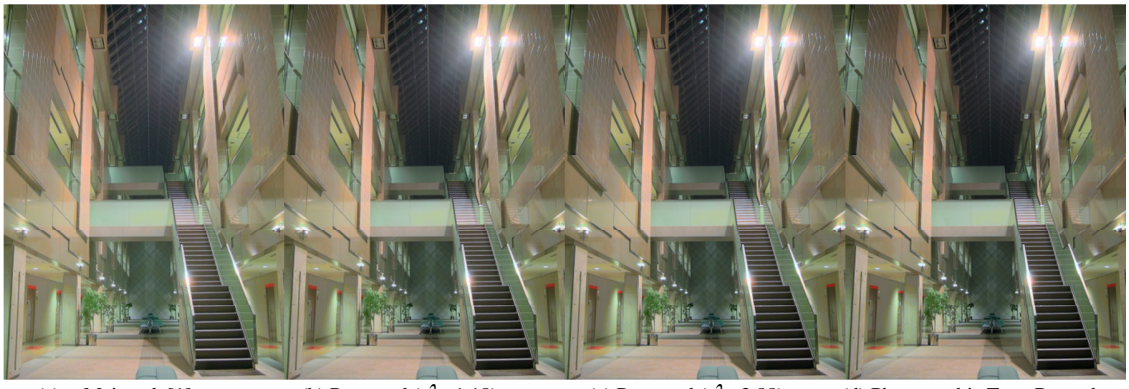
Figure 5 shows the rate-distortion curves in terms of coding bit rate of LDR image vs. the HDR-MSE, the mean square error between the logarithm of luminance values of original and reconstructed HDR image. H.264/AVC reference software JM 17.2 [18] is used for the encoding of LDR images. The supplemental information for the characterization of tone mapping curves is very negligible compared to the coding bitrate. As expected, 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. But, the quality of the generated tone mapped images is increasing on the other hand.

5. CONCLUSIONS

We reveal the basic trade-off between the LDR image quality and HDR image quality for HDR image and video coding. By putting the LDR image quality constraint into the optimized tone mapping problem formulated by Mai. *et al.* [1], we develop a methodology to arrange and verify this trade-off. Future work will focus on the extension of the problem for local TMOs and video.

ACKNOWLEDMENTS

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



(a) Mai et al. [1] (b) Proposed (λ : 1.18) (c) Proposed (λ : 2.88) (d) Photographic Tone Reproduct. [13]

(e) Mai et al. [1] (f) Proposed (λ : 1.02) (g) Proposed (λ : 2.60) (h) Photographic Tone Reproduct. [13]

Fig. 4 Produced LDR images. Upper Line: *Atrium Night;* Bottom Line: *Coby.*

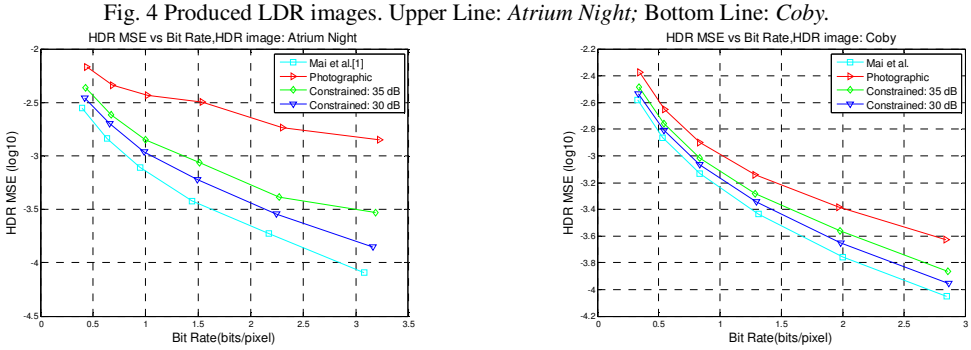


Fig. 5 HDR MSE (log10) vs. Coding Bitrate results for *Atrium Night* (left) and *Coby* (right). Some encoding parameters for the utilized H.264/AVC reference software JM 17.2 are as follows: ProfileIDC = 244; LevelIDC = 40; YUVformat=3 (i.e. 4:4:4)

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] 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.

[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] Erik Reinhard, Mike Stark, Peter Shirley and Jim Ferwerda,"Photographic Tone Reproduction for Digital Images', ACM Transactions on Graphics, 21(3), pp 267--276, July 2002.

[14] Patrick Ledda, Alan Chalmers, Tom Troscianko and Helge Seetzen, "Evaluation of tone mapping operators using a high dynamic range display", ACM Trans. Graph. 24, 3, 640-648 (2005).

[15] Erik Reinhard, Mike Stark, Peter Shirley and Jim Ferwerda,"Photographic Tone Reproduction for Digital Images', ACM Transactions on Graphics, 21(3), pp 267--276, July 2002.

[16] HDR image Gallery [online]. Available: http://www.mpi-inf.mpg.de/resources/hdr/gallery.html

[17] HDR image Gallery [online]. Available: http://www.anyhere.com/gward/hdrenc/pages/originals.html

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