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-01436223
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.
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 ($10^3$ 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 layer and residue [5], such a mapping is characterized by the fact that the 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. 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 given in section 5.

The HDR video is first mapped into 8-bit by using a TMO. The resulting LDR video goes.

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. Finally, the conclusions are proposed methodology in terms of the rate-distortion curves and the quality of produced LDR images.

Section 4 gives the experimental results for the constrained trade-off between HDR image quality and LDR image quality during the coding for a given reference TMO. 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. Finally, the conclusions are proposed methodology in terms of the rate-distortion curves and the quality of produced LDR images.

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. 2b.). 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

\[
E((s(l) - h(l))^2) = \sum_{k} \left( \sum_{n} (s_n - h_n) \right)^2 \cdot p_k
\]

where \( s_n \) and \( h_n \) 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

\[
\begin{align*}
& \text{arg min}_{\delta} \sum_{k} \left( \sum_{n} (s_n - h_n) \right)^2 \cdot p_k \\
& \text{subject to } \sum_{k} s_k = \frac{v_{\text{max}}}{\delta} \\
& \quad \delta^3 \sum_{k} \left( \sum_{n} s_n - h_n \right)^2 \cdot p_k = \Delta
\end{align*}
\]

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:

\[
-2p_k \frac{s_k}{s_k^3} + \lambda_k + \lambda_2 \delta \sum_{n} (s_n - h_n) \cdot p_n = 0, \quad k = 1..N
\]
solution to the set of KKT conditions is finalized as:

\[ \sum_{k=1}^{N} s_k - \frac{v_{\text{max}}}{\delta} = 0 \]

and

\[ \delta^2 \sum_{k=1}^{N} (s_k - h_k)^2 \cdot p_k - \Delta = 0 \] (6)

where \( \delta_1 \) and \( \delta_2 \) 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 algorithms can be considered as an alternative, such an approach would also be computationally loaded for coding applications.

Our proposed solution is first to use the found optimum solution for \( g_x \) in (3) for the one-constraint optimization problem, and then to redefine the cost function as a MSE distance to the optimal solution, while keeping the MSE distance to the reference LDR image in a desired level. Let us assume that \( s_i \cdot h_i \), and \( g_i \) 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. 2c). Then, the new optimization problem in terms of the MSEs can be expressed as:

\[
\begin{align*}
\arg\min_{g_1, \ldots, g_N} & \quad \delta^2 \sum_{k=1}^{N} (s_k - g_k)^2 \cdot p_k \\
\text{subject to} & \quad \delta^2 \sum_{k=1}^{N} (g_k - h_k)^2 \cdot p_k = \Delta
\end{align*}
\]

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:

\[ \delta^2 \sum_{k=1}^{N} 2(g_k - s_k + \lambda_1 (g_k - h_k)) \cdot p_k = 0, \quad k = 1..N \]

\[ \delta^2 \sum_{k=1}^{N} (g_k - h_k)^2 \cdot p_k - \Delta = 0 \] (8)

where \( \lambda_1 \) is the Lagrangian coefficient for the constraint. The solution to the set of KKT conditions is finalized as:

\[
\sum_{k=1}^{N} s_k - \frac{v_{\text{max}}}{\delta} = 0
\]

\[
\delta^2 \sum_{k=1}^{N} (s_k - h_k)^2 \cdot p_k - \Delta = 0
\]

where \( \lambda_1 \) and \( \lambda_2 \) 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 algorithms can be considered as an alternative, such an approach would also be computationally loaded for coding applications.

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:

\[
\sum_{k=1}^{N} s_k - \frac{v_{\text{max}}}{\delta} = 0
\]

\[
\delta^2 \sum_{k=1}^{N} (s_k - h_k)^2 \cdot p_k - \Delta = 0
\]

where \( \lambda_1 \) is the Lagrangian coefficient for the constraint. The solution to the set of KKT conditions is finalized as:

\[ x = \arg\min_{x \in \mathbb{R}^n} F(x) \]

\[ \text{subject to} \quad G_i(x) = 0, \quad i = 1, \ldots, m \]

\[ h_j(x) \leq 0, \quad j = 1, \ldots, l \]

where \( F(x) \) is the objective function, \( G_i(x) \) are the equality constraints, and \( h_j(x) \) are the inequality constraints. The KKT conditions are given by:

\[ \nabla F(x) + \sum_{i=1}^{m} 
\]

\[ h_j(x) = 0, \quad j = 1, \ldots, l \]

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] to test our algorithms. 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 \( \lambda \) parameter is adjusted.

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).
such that the PSNR between the resulting and reference tone mapped images is 35 dB and 30 dB for two cases. While $\lambda$ 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 $\lambda$.

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 $\lambda$ 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 $\lambda$ increases, the rate-distortion performance is getting poor by shifting towards 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.

ACKNOWLEDGMENTS

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

![Fig. 4 Produced LDR images. Upper Line: Atrium Night; Bottom Line: Coby.](image1)

![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)](image2)
REFERENCES


