

Optimizing a Tone Curve for Backward-Compatible High Dynamic Range Image and Video Compression

S. Aravind Kumar¹

¹ Anna University of technology Chennai

Anand Institute of Higher Technology, Kazhipattur, Tamil Nadu, India.

Abstract. High dynamic range (HDR) is a technique that allow a great dynamic range of luminance between the lightest and darkest area of an image. For video compression, the HDR sequence is reconstructed by inverse tone-mapping a compressed Standard dynamic range (SDR) version of the original HDR content. In this paper, we show that the appropriate choice of a Tone-mapping operator (TMO) can significantly improve the reconstructed HDR quality. It is used to compress a large range of pixel luminance in to smaller range that is suitable for display on devices with limited dynamic range. we formulate a numerical optimization problem to find the tone-curve that minimizes the expected mean square error (MSE) in the reconstructed HDR sequence. We also develop a simplified model that reduces the computational complexity of the optimization problem to a closed-form solution. It is also shown that the LDR image quality resulting from the proposed methods matches that produced by perceptually-based TMOs.

Keywords: Bit-depth scalable, High dynamic range, video compression, tone-mapping.

1. Introduction

Natural scene contain far more visible information that can be captured by the majority of digital imagery and video devices High dynamic range (HDR) video encoding goes beyond the typical color space restrictions and attempts to encode all colors that are visible and distinguishable to the human eye. The main motivation is to create a video format that would be future-proof, independent of a display technology, and limited only by the performance of the human visual system (HVS). HDR images preserve colorimetric or photometric pixel values (such as CIE XYZ) within the visible color gamut and allows for intra-frame contrast exceeding 5–6 orders of magnitude ($10^6:1$), without introducing contouring, banding or posterization artifacts caused by excessive quantization. High dynamic range video formats are unlikely to be broadly accepted without the backward-compatibility with these devices. Such backward-compatibility can be achieved if the HDR video stream contains 1) a backward-compatible 8-bit video layer which could be directly displayed on existing devices, and 2) additional information which along with this 8-bit layer can yield a good quality reconstructed version of the original HDR content. Finally, a HDR residual signal can also be extracted and encoded in the bit stream as an enhancement layer.

To compute the tone-curve, we propose a method that minimizes the difference in the video quality between the original and the reconstructed HDR video. This difference results in quality loss and is due to tone-mapping, encoding, decoding and inverse tone-mapping the original video. Minimizing this difference would reduce the size of the HDR residual signal in the enhancement layer. We also achieve the primary goal of tone-mapping which is to produce an LDR image with a visual response as similar as possible to the original HDR image. Although the initial assumptions used in our approach pose a difficult optimization problem, we demonstrate that for typical compression distortions there exists a closed-form solution that approaches the optimum.

This paper is organized as follows: an overview of related work is presented in Section 2. In Section 3, the proposed tone-mapping approach that considers tone-mapping together with compression is discussed in detail. Section 4 demonstrates and analyzes the performance of the proposed methods. Finally, we draw our conclusions in Section 5.

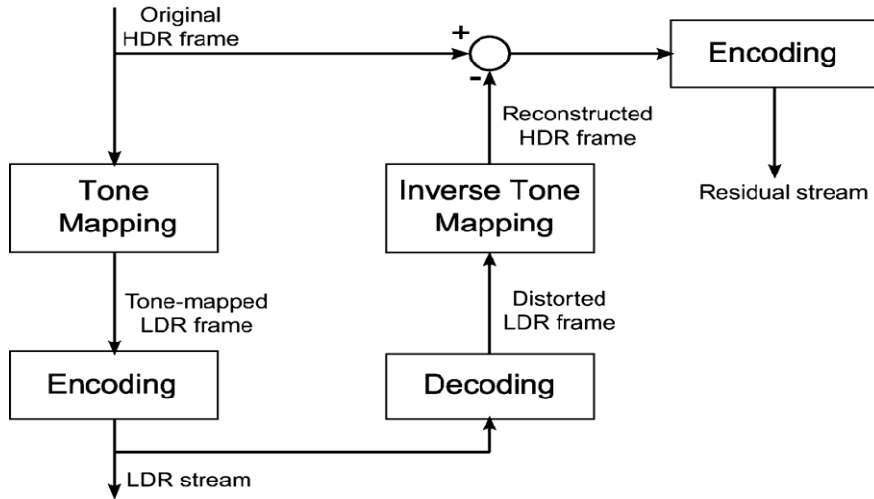


Fig. 1: General structure of the scalable approach used for backward-compatible HDR video encoding. The base layer encodes an 8-bit LDR representation of the HDR input. The enhancement layer encodes the difference (residual) between the inverse tone-mapped base layer and the original HDR source.

2. Related Work

Backward compatible HDR video encoding has received significant interest recently. A color space of encoding HDR content based on the luminance threshold sensitivity of the human visual system. They concluded that 10–12 bit luma encoding is sufficient to encode the full range of visible and physically plausible luminance levels. Their encoding, however, is not backward-compatible with the existing video decoding hardware. For still image compression, backward compatibility can be achieved by encoding a tone-mapped copy of the HDR image together with a residual or a ratio image that allows the reconstruction of the original HDR image. A tone-mapping curve was encoded together with the tone-mapped and residual video sequences. The residual video sequence was additionally filtered to remove the information that is not visible to the human eye.

Recently, several proposals for bit-depth scalability have been introduced to provide backward-compatible HDR video bit streams. These proposals incorporate backward-compatible encoding of high fidelity video as an extension to the H.264/AVC video encoding standard. The extension includes tone-mapping SEI messages, which encode the shape of the tone-mapping curve. The scalable video coding (SVC) extension is used to encode an additional residual stream needed to reconstruct the information lost due to tone-mapping, or to provide bit-depth scalability. The proposed tone mapping is meant to be used in combination with the scalable video coding. In that case the backward compatible (tone-mapped) sequence is generated by our operator, instead of being provided by a user, which results in better compression efficiency. we compare the results of our study with the two tone-mapping operators that performed the best in this study: the photographic TMO and the adaptive logarithmic TMO .

3. Problem Statement And Proposed Solution

In this section, we present the challenges of obtaining a good quality reconstructed HDR representation in a backward-compatible HDR video encoding system and describe in detail the approach we propose towards overcoming these challenges. The performance of a backward-compatible HDR video and image encoding system depends on the coding efficiency of the LDR base layer and the HDR enhancement layer. Performance gains can be achieved by finding a TMO that preserves the necessary information in the LDR base layer so that after it passes through the inverse TMO process, the resultant HDR reconstructed signal is of high quality. The coding efficiency of the base layer does not depend much on the TMO used as most TMOs attain a similar level of contrast in the LDR representation.

Our proposed approach attempts to find the best global (spatially invariant) tone-mapping curve that minimizes the mean square error (MSE)¹ between the original HDR content and the reconstructed version obtained after tone-mapping, compression, decompression, and inverse tone-mapping. This process is illustrated in Fig. 2. Let I denote the input HDR image/frame, and the tone mapped LDR version as shown in Fig. 2(a). Let \hat{I} be the decoded LDR frame, and the reconstructed HDR frame produced after inverse-tone-mapping. Also let θ be the set of parameters that control the tone-mapping operator, then our goal is to find the tone-mapping parameters that minimize MSE, which we denote $\|I - \hat{I}\|_2^2$ as using the norm notation. We

choose the mean square error as the HDR quality metric for its simplicity, despite its shortcomings in reflecting the perceptual quality of images.

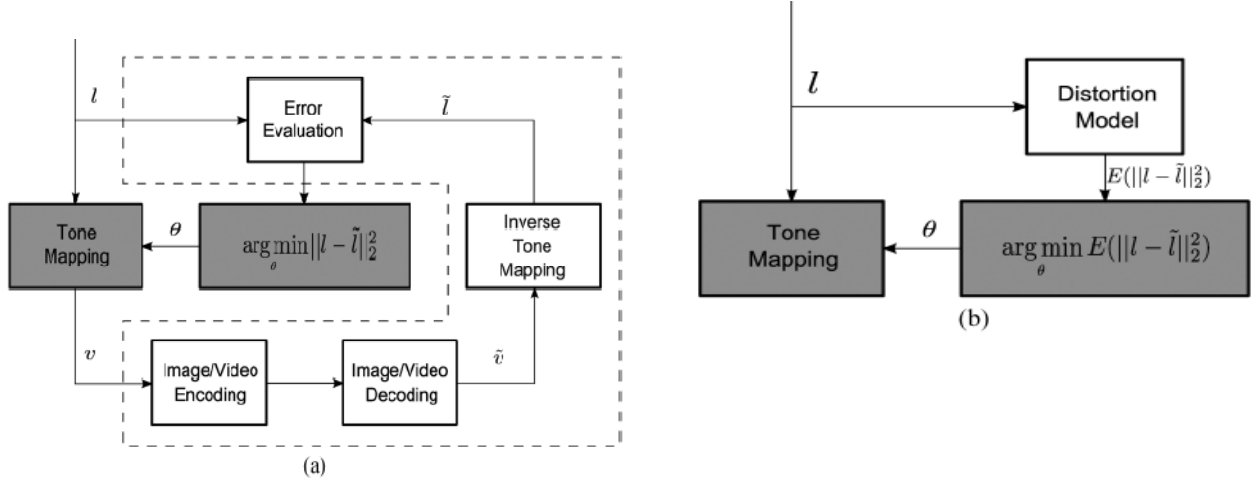


Fig.2: System overview of the proposed tone-mapping method. (a) demonstrates the ideal scenario where the actual H.264/AVC encoding is employed. (b) shows the practical scenario which is addressed by this paper.

4. Experimental Results And Discussion

In this section we first validate the proposed methods: optimization using the statistical model and the closed-form solution. Then, our models are further analyzed based on the generated tone curve and the distortion of the reconstructed HDR content. The performance of our models is also evaluated by comparing it with existing tone-mapping methods. We use H.264/AVC encoding as an example to demonstrate the results. In the experiments below, all tone-mapped images are compressed/decompressed using the intra mode of the H.264/AVC reference software, where inter-frame mode is also used.

4.1. Tone-Mapping curve

The global tone-mapping curve is a function that maps HDR luminance values to either the display's luminance range, or directly to LDR pixel values. The tone-mapping curve is usually continuous and non-decreasing. The two most common shapes for the tone curves are the sigmoidal ("S-shaped") or a compressive power function with an exponent 1 (gamma correction). According to the Weber-Fechner law, which forms a vector of tone-mapping parameters. Using this parameterization, the forward tone-mapping function is defined as

$$S_k = \frac{V_{k+1} - V_k}{\delta} \quad (1)$$

$$V(l) = (l - l_k) \cdot S_k + V_k \quad (2)$$

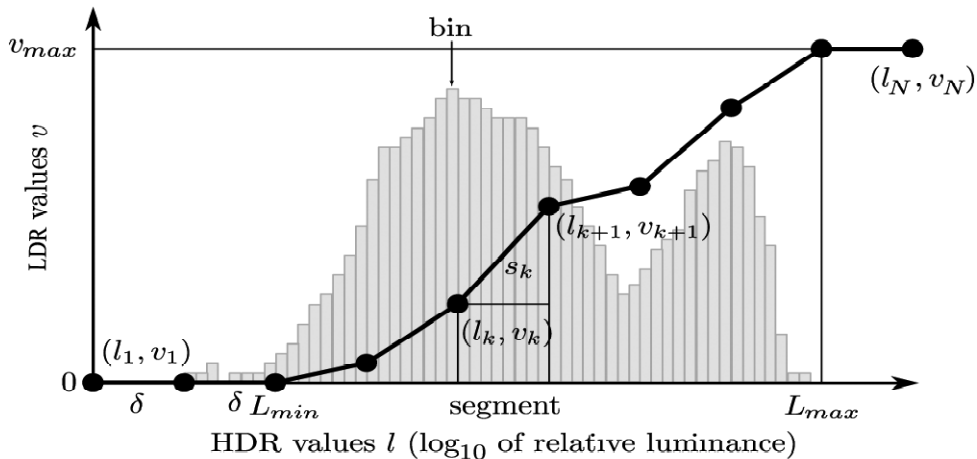


Fig. 3: Parameterization of a tone-mapping curve and the notation. The bar-plot in the background represents an image histogram used to compute $p(l)$.

The inverse mapping function is then

$$\bar{l}(v; S_k) = \begin{cases} \sum_{l \in S_0}^{v-v_k} l \cdot pL(l) \\ \text{for } S_k < 0, S_k = 0 \text{ where } S_k \in \{S_1, \dots, S_N\} \end{cases} \quad (3)$$

4.2. Statistical distortion model

We estimate the error $\|\bar{l} - l\|_2^2$ assuming that the compression distortions follow a known probability distribution P_C . Under this assumption, the expected value of the error $\|\bar{l} - l\|_2^2$ is

$$E \left[\|\bar{l} - l\|_2^2 \right] = \sum_{l=l_{\min}}^{l_{\max}} \sum_{\bar{v}=0}^{v_{\max}} \left(\bar{l}(v; S_k) - l \right)^2 \cdot pc(v(l) - \bar{v} / v(l)) \cdot p_L(l) \quad (4)$$

Where $pc(v - \bar{v} / v)$ is the probability that the encoding error equals $v - \bar{v}$.

4.3. Optimization problem

The optimum tone curve can be found by minimizing the $\varepsilon(S_k)$ function with respect to the segment slopes

$$\begin{aligned} S_k \arg \min \varepsilon(S_k) \quad \text{where } S_1, \dots, S_N \quad \text{Subject to:} \\ S_{\min} \leq S_k \leq S_{\max} \quad \text{for } k=1, \dots, N \\ \sum_{k=1}^N S_k \cdot \delta = v_{\max}. \end{aligned} \quad (5)$$

The first constraint restricts slopes to the allowable range, while the second ensures that the tone curve spans exactly the range of pixel values from 0 to v_{\max} .

4.4. Closed-Form solution

After reorganizing we get

$$\varepsilon(S_k) \approx \sum_{l=l_{\min}}^{l_{\max}} \sum_{\bar{v}=0}^{v_{\max}} pc(v - \bar{v}) \cdot \sum_{l=l_{\min}}^{l_{\max}} \frac{pL(l)}{S_k^2} \cdot \text{Var}(v - \bar{v}) \quad (6)$$

Since the variance of $(v - \bar{v})$ does not depend on the slopes, it does not affect the location of the global minimum of $\varepsilon(S_k)$ and thus can be omitted when searching for the minimum. our method can lower the total bit rate for the bit-depth scalable coding. Although the primary goal of tone-mapping algorithms is to optimize the visual quality of the displayed LDR image, we instead designed a tone-mapping operator that optimizes the compression performance in the backward-compatible encoding scheme. our study shows that the overall quality of images tone-mapped with our method is comparable to other tone-mapping algorithms and none of the tone-mapped images.

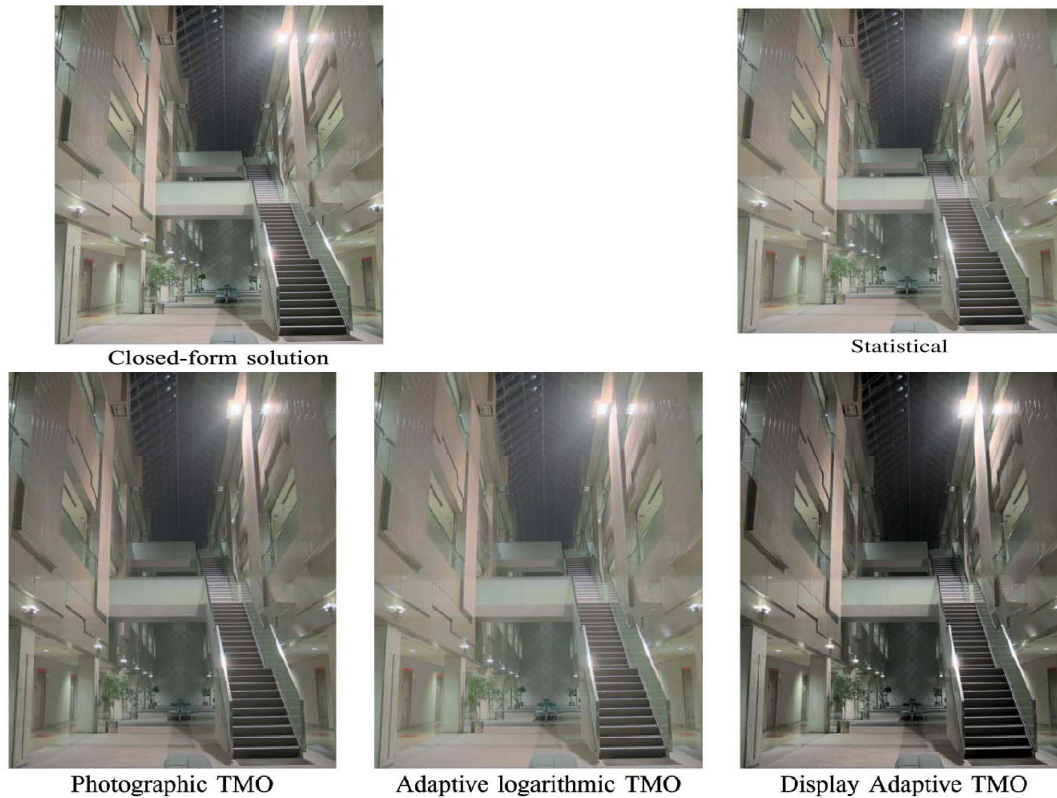


Fig. 4: shows the tone-mapped images using the existing tone-mapping methods. All the tone-mapped images shown are compressed. The compression quantization parameters used for “AtriumNight” is 10. The number of segments used for the histogram is 87.

5. Conclusion

In this paper, we showed that the appropriate choice of a tone-mapping operator (TMO) can significantly improve the reconstructed HDR quality. We developed a statistical model that approximates the distortion resulting from the combined processes of tone-mapping and compression. Using this model, we formulated a constrained optimization problem that finds the tone-curve which minimizes the expected HDR MSE. The resulting optimization problem, however, suffers from high computational complexity. Therefore, we presented a few simplifying assumptions that allowed us to reduce the optimization problem to an analytically tractable form with a closed-form solution. The closed-form solution is computationally efficient and has a performance compatible to our developed statistical model. Moreover, the closed-form solution does not require the knowledge of QP, which makes it suitable for cases where the compression strength is unknown. Although our models are designed to minimize HDR MSE, the extensive performance evaluations show that the proposed methods provide excellent performance in terms of SSIM and the LDR image quality, in addition to an outstanding performance in MSE.

6. References

- [1] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, “High dynamic range display systems,” *ACM Trans. Graph. (Proc. SIGGRAPH)*, vol. 23, no. 3, pp. 757–765, 2004.
- [2] IEC61966-2-4, Colour Measurement and Management—Part 2-4: Colour Management—Extended-Gamut YCC Colour Space for Video Applications— xvYCC 2006.
- [3] R. Mantiuk, G. Krawczyk, K. Myszkowski, and H.-P. Seidel, “Perception- motivated high dynamic range video encoding,” *ACM Trans. Graph. (Proc. SIGGRAPH)*, vol. 23, no. 3, pp. 730–738, 2004.
- [4] G. Ward, “Real pixels,” *Graphics Gems II*, pp. 80–83, 1991.
- [5] R. Bogart, F. Kainz, and D. Hess, “Openexr image file format,” in *Proc. ACM SIGGRAPH 2003, Sketches Applicat.*, 2003.
- [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] Y. Gao and Y. Wu, “Applications and Requirement for Color Bit Depth Scalability,” Oct. 2006, Tech. Rep., ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT-U049.
- [8] G. J. Sullivan, H. Yu, S.-I. Sekiguchi, H. Sun, T. Wedi, S. Wittmann, Y.-L. Lee, A. Segall, and T. Suzuki, “New standardized extensions of mpeg4-avc/h.264 for professional-quality video applications,” in *Proc. ICIP*, 2007, vol. 1, pp. I-13–I-16.
- [9] M. Winken, D. Marpe, H. Schwarz, and T. Wiegand, “Bit-depth scalable video coding,” in *Proc. ICIP*, 2007, vol. 1, pp. I-5–I-8.