GRADIENT DOMAIN TONE MAPPING OF HIGH DYNAMIC RANGE VIDEOS

Chul Lee and Chang-Su Kim

School of Electrical Engineering, Korea University, Seoul, Korea
E-mails: kayne@korea.ac.kr, cskim@ieee.org

ABSTRACT

A gradient domain tone mapping algorithm is proposed to display high dynamic range (HDR) video sequences in low dynamic range (LDR) devices in this work. The proposed algorithm obtains a pixelwise motion vector field and incorporates the motion information into the Poisson equation. Then, by attenuating large spatial gradients, the proposed algorithm can yield a high-quality tone-mapped result without flickering artifacts. Simulation results show that the proposed algorithm provides a better performance than the frame-based method, which processes each frame independently.

***Index Terms*—** high dynamic range imaging, tone mapping, dynamic rage compression, and video processing.

1. INTRODUCTION

The dynamic range of a digital image represents the ratio of the intensities between the brightest pixel and the darkest pixel. Whereas the dynamic ranges of conventional display devices are typically less than two orders of magnitude, real world scenes have much higher dynamic ranges. Therefore, the LDR images, taken by digital cameras and shown on conventional displays, cannot express the gamut of real scenes faithfully.

HDR images can be constructed with multiple photographs of a scene, which are differently exposed [1]. Furthermore, the recent advance of sensor technology makes it possible to record the full dynamic range of a scene in a single shot [2]. It is expected that general still or video cameras will be able to capture HDR scenes directly in near future. However, the problem arises on how to display HDR images or videos on LDR display devices. Thus, various algorithms, called tone mapping or dynamic range compression schemes, have been developed to solve this issue. One of the most effective schemes is the Fattal *et al*.'s tone mapping scheme [3], which processes images in the gradient domain.

Compared with the tone mapping of still images, relatively little effort has been made for the tone mapping of video sequences. In [4], a compression algorithm for HDR videos was proposed, which represents the tone-mapped LDR information as a base layer and the HDR refinement information as an enhancement layer. Thus, it can supports conventional LDR display devices as well as HDR devices. Alternatively, a tone mapping algorithm can be applied to the frames of a HDR video sequence at the receiver side. In this work, we focus on this second approach.

A high quality tone-mapped LDR video cannot be obtained by simply compressing the dynamic ranges of successive frames independently. This simple approach leads to flickering artifacts, which are caused by the lack of temporal coherence. Only a few methods have been proposed to overcome this problem. In [5], Wang *et al*. treated a HDR video as a 3D block of pixels and extended the Fattal *et al*.'s tone mapping scheme [3] straightforwardly to the 3D case. However, they noticed the temporal blurring of video sequences, when the temporal gradients are also attenuated. Therefore, only the spatial gradients are attenuated for dynamic range compression. The main disadvantage of the Wang *et al*.'s algorithm is that it does not take the motion information into account during the tone mapping.

In this work, we propose a new approach to the dynamic range compression of HDR videos, which exploits the motion information to achieve a higher-quality tone mapping result. By incorporating a temporal coherence term into the Poisson equation, the proposed algorithm can suppress flickering artifacts, while compressing the dynamic ranges effectively. Simulation results are presented to validate the performance of the proposed algorithm. Moreover, it is shown that the proposed algorithm can enhance the quality of LDR sequences with very low contrast.

The remainder of this paper is organized as follows. Section 2 reviews the gradient domain tone mapping. Section 3 describes the proposed tone mapping algorithm for HDR videos, and Section 4 discusses the simulation results. Finally, Section 5 concludes the paper.

2. GRADIENT DOMAIN TONE MAPPING

For the sake of completeness, we briefly review the gradient domain tone mapping for still images, which was proposed in [3] by Fattal *et al*. The gradient domain tone mapping is based on the assumption that human visual system is more sensitive to local intensity changes rather than to absolute luminance values and that the intensity changes are well represented by gradients. Intuitively speaking, large gradients correspond to the boundaries between objects. Therefore, by attenuating the large gradients, we can compress the dynamic range of an input image, while preserving the details within each object.

Specifically, a modified gradient map $G(x, y)$ is obtained by multiplying the gradient map $\nabla H(x, y)$ of the original image by an attenuation function $\Phi(x, y)$ as follows.

$$G(x, y) = \nabla H(x, y)\Phi(x, y). \tag{1}$$

The modified map $G(x, y)$ may not be integrable. Instead, [3] attempts to obtain the output image I, whose gradient is closest to G in the least square sense. After the discrete approximation of gradient operators, the cost function to be minimized is given by

$$\begin{aligned} C_1 &= \sum_{x,y} \left\|\nabla I(x, y) - G(x, y)\right\|^2 \\ &= \sum_{x,y} \left\{[I(x+1, y) - I(x, y) - G_x(x, y)]^2 \right. \\ &\quad \left. + [I(x, y+1) - I(x, y) - G_y(x, y)]^2\right\}, \end{aligned} \tag{2}$$

where G_x and G_y stand for the x and y components of the gradient, respectively.

To minimize C_1, we differentiate it respect to $I(x,y)$ and set it to 0, which yields the following equation

$$\begin{aligned} & I(x+1,y)+I(x-1,y) \\ & \quad +I(x,y+1)+I(x,y-1)-4I(x,y) \\ = \quad & G_x(x,y)-G_x(x-1,y) \\ & \quad +G_y(x,y)-G_y(x,y-1). \end{aligned} \tag{3}$$

This is the discrete approximation of the Poisson equation

$$\nabla^2 I = \operatorname{div} G, \tag{4}$$

whose numerical solution can be obtained efficiently due to its sparsity.

For the attenuation function $\Phi(x,y)$ in (1), [3] uses a Gaussian pyramid and defines the function at level k by

$$\Phi_k(x,y) = L(\Phi_{k+1})(x,y)\varphi_k(x,y) \tag{5}$$

where L is an upsampling operator. The scaling factor $\varphi_k(x,y)$ is defined as

$$\varphi_k(x,y) = \frac{\alpha}{\|\nabla H_k(x,y)\|}\left(\frac{\|\nabla H_k(x,y)\|}{\alpha}\right)^{\beta}, \tag{6}$$

where α and β are user-controllable parameters. Then, $\Phi_0(x,y)$ at level 0 becomes the final attenuation function $\Phi(x,y)$.

3. PROPOSED TONE MAPPING ALGORITHM FOR HDR VIDEOS

The tone mapping scheme in the previous section can be applied to each frame of a video sequence independently. However, in such a case, the temporal coherence can be broken, and the resulting LDR video often exhibits severe flickering artifacts. In this work, we exploit the temporal motion information to obtain a high-quality, temporally coherent LDR video.

We first estimate a pixelwise motion vector field between two successive HDR frames $H(x,y,t)$ and $H(x,y,t-1)$. For the motion estimation, the block matching algorithm is employed using a square window around each pixel. As shown in Fig. 1, let (v_x, v_y) denote the motion vector for a pixel $H(x,y,t)$. Then, the same motion vector is used to estimate the LDR pixel value $I(x,y,t)$ at the spatial coordinate (x,y) and the time instance t. Let $P(x,y)$ denote the pixel in the previous frame, specified by the motion vector (v_x, v_y), *i.e.*,

$$P(x,y) = I(x-v_x, y-v_y, t-1).$$

We expect that, after the tone mapping, $I(x,y,t)$ and $P(x,y)$ should have similar values, since they represent the same object point at successive time instances. Therefore, in addition to the original cost function C_1 in (2), we introduce a new cost for the temporal coherence, given by

$$C_2 = \sum_{x,y}\{I(x,y,t)-P(x,y)\}^2. \tag{7}$$

Then, the overall cost is defined as a weighted sum of the two costs, given by

$$C = C_1 + \lambda C_2 \tag{8}$$

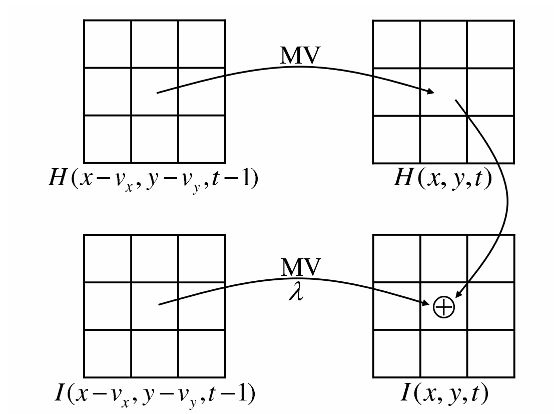


Fig. 1. A pixelwise motion vector field is obtained between the HDR frames $H(x,y,t-1)$ and $H(x,y,t)$. It is then used as a constraint for the temporal coherence between the LDR frames $I(x,y,t-1)$ and $I(x,y,t)$.

where λ is a weighting coefficient. As λ gets larger, flickering artifacts can be more effectively reduced. However, too large a λ can degrade the quality of each tone-mapped frames.

We differentiate C respect to $I(x,y,t)$ and set it to 0. Then, after rearranging terms, we have the modified Poisson equation

$$\begin{aligned} & I(x+1,y)+I(x-1,y) \\ & \quad +I(x,y+1)+I(x,y-1)-(4+\lambda)I(x,y) \\ = \quad & G_x(x,y)-G_x(x-1,y) \\ & \quad +G_y(x,y)-G_y(x,y-1)-\lambda P(x,y) \end{aligned} \tag{9}$$

where the time index t in $I(x,y,t)$ is omitted for short notations. We solve the set of equations for all pixels using a modified version of the full multigrid algorithm [6] with the Gauss-Seidel smoothing iteration.

Finally, we obtain the luminance of each frame $I(x,y)$. In order to reconstruct the color components for each pixel, we use the following system of equations [2].

$$\begin{bmatrix} R_{\text{out}} \\ G_{\text{out}} \\ B_{\text{out}} \end{bmatrix} = \begin{bmatrix} L_{\text{out}}\left(\frac{R_{\text{in}}}{L_{\text{in}}}\right)^s \\ L_{\text{out}}\left(\frac{G_{\text{in}}}{L_{\text{in}}}\right)^s \\ L_{\text{out}}\left(\frac{B_{\text{in}}}{L_{\text{in}}}\right)^s \end{bmatrix} \tag{10}$$

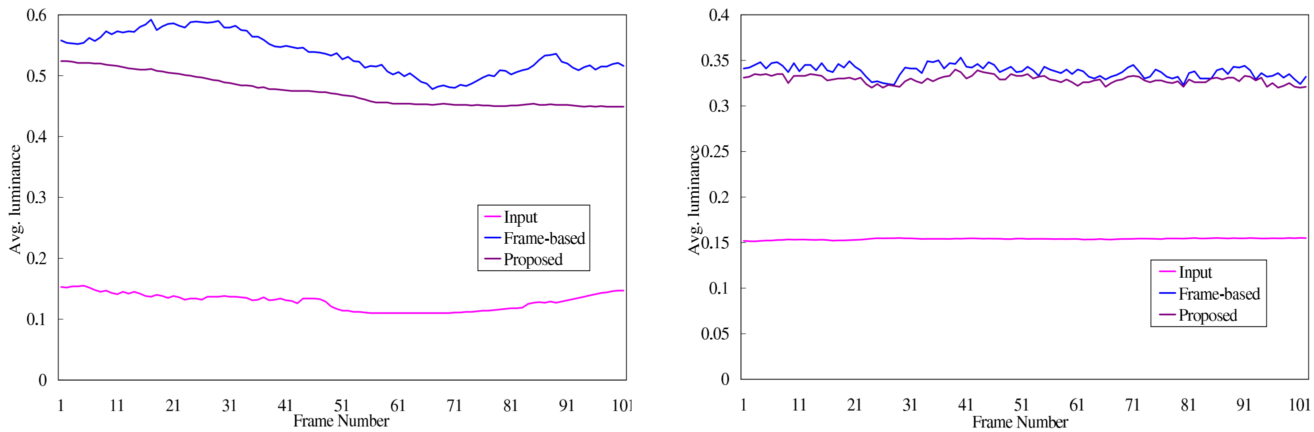
where L_{in} and L_{out} are the luminance values before and after the HDR compression, respectively. R_{in}, G_{in}, and B_{in} are the three color components of an HDR pixel, while R_{out}, G_{out}, and B_{out} are the tone-mapped color components. The exponent s is the saturation parameter.

4. EXPERIMENTAL RESULTS

We investigated the performance of the proposed algorithm on various HDR and LDR video sequences, and confirmed that it provides satisfactory results. In this section, we give a few examples to illustrate the validity of the proposed algorithm.

4.1. High Dynamic Range Video

An HDR sequence, whose tone mapped results are shown in Fig. 3, is used in this test. The parameters α and β in (6) are set to one tenth



(a) (b)

Fig. 2. Comparison of normalized average luminance values on (a) an HDR sequence and (b) an LDR sequence. The normalization is performed with respect to the maximum luminance in the input sequence. The proposed algorithm provides about 24.1% and 31.2% lower average deviation than the frame-based method on the HDR sequence and the LDR sequence, respectively.

of the average gradient and 0.92, respectively. Also, λ in (8) is 0.3, and s in (10) is 0.6.

Fig. 3(a) compares the normalized average luminance of each frame, obtained by the frame-based method and the proposed algorithm. The normalization is carried out with respect to the maximum luminance in the sequence. The frame-based method applies the tone mapping scheme in [3] independently to each frame, without considering the temporal coherence. Thus, the average luminances vary unpredictably from frame to frame. On the other hand, the proposed algorithm provides a flatter curve, which indicates that flickering artifacts are effectively reduced. More specifically, the proposed algorithm provides about 24.1% lower standard deviation of the average luminances than the frame-based method.

It is observed in Fig. 3 that the proposed algorithm provides a better visual quality than the frame-based method, by employing the temporal motion information.

4.2. Low Dynamic Range Video

We also apply the proposed algorithm to enhance an LDR sequence, which has very low contrast. In this test, α is also set to one tenth of the average gradient, $\beta = 0.92$, $\lambda = 0.25$ and $s = 1.0$.

By attenuating relative large gradients and rescaling the result, the details in dark regions can be faithfully reconstructed. Fig. 3(b) shows the average luminance variation, and Fig. 4 shows selected frames. The proposed algorithm yields about 31.2% lower standard deviation of the average luminances, and provides a high quality output. On the other hand, the frame-based method does not take into account the motion information, and causes motion artifacts around the walking man.

5. CONCLUSIONS

In this work, we extended the Fattal *et al.*'s gradient domain tone mapping scheme to the processing of HDR video sequences. The proposed algorithm incorporates the temporal motion information into the Poisson equation, and thus can provide a tone-mapped sequence without flickering artifacts. The simulation results demonstrated that the proposed algorithm is capable of enhancing LDR sequences with very low contrast, as well as providing satisfactory tone-mapping results for HDR sequences.

6. ACKNOWLEDGEMENTS

This research was supported by the MIC (Ministry of Information and Communication), Korea, under the ITFSIP (IT Foreign Specialist Inviting Program) supervised by the IITA (Institute of Information Technology Assessment).

We would like to thank Grzegorz Krawczyk for making the HDR sequence in Fig. 3 available.

7. REFERENCES

[1] P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *SIGGRAPH 97*, Aug. 1997, pp. 369–378.

[2] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging*, Morgan Kaufmann Publishers, 2005.

[3] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. Graphics*, vol. 21, no. 3, pp. 249–256, July 2002.

[4] R. Mantiuk, A. Efremov, K. Myszkowski, and H. Seidel, "Backward compatible high dynamic range MPEG video compression," *ACM Trans. Graphics*, vol. 25, no. 3, pp. 713–723, July 2006.

[5] H. Wang, R. Raskar, and N. Ahuja, "High dynamic range video using split aperture camera," in *IEEE 6th Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras*, Oct. 2005.

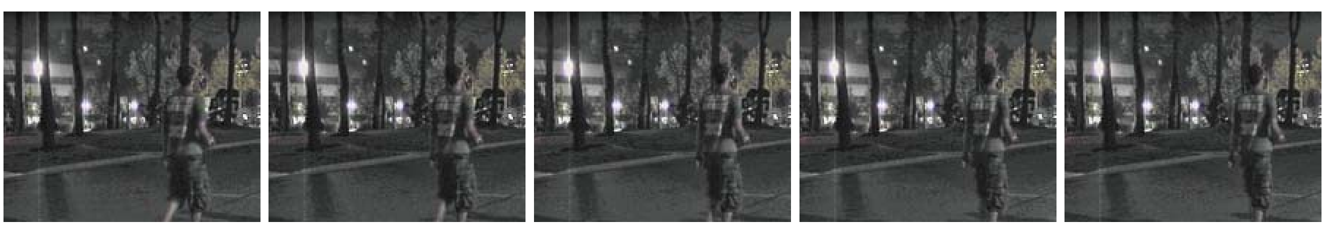
[6] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C : The Art of Scientific Computing*, Cambridge University Press, 2nd edition, 1992.

(a)

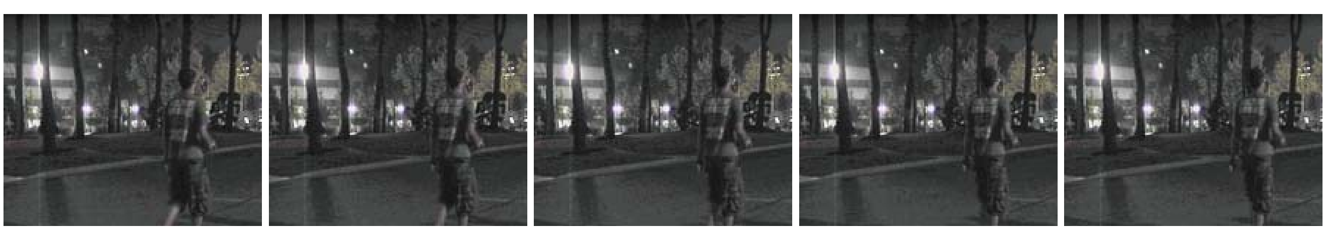
(b)

Fig. 3. Tone-mapping results of (a) the frame-based method and (b) the proposed algorithm on an HDR video sequence. The HDR sequence is the courtesy of Grzegorz Krawczyk.

(a)



(b)



(c)

Fig. 4. (a) An LDR sequence. The tone-mapping results of (b) the frame-based method and (c) the proposed algorithm. The frame number ranges from 31 to 35.