Volume xx (200y), Number z, pp. 1–3

Temporal Coherency in Video Tone Mapping, A Survey

Ronan Boitard[1,2] Dominique Thoreau[1] Rémi Cozot[2] and Kadi Bouatouch[2]

[1]Technicolor, 975 Av. des Champs Blancs, 35576 Cesson-Sevigne, France;
[2]IRISA, 263 Avenue du Général Leclerc, 35000 Rennes, France

Keywords: Video Tone Mapping, HDR Video, Temporal Coherency

1. INTRODUCTION

With the development of sensor technology, it is now possible to capture more information than can be displayed. While some emerging technologies are capable of displaying broader luminance ranges, they are still quite expensive and will not be available on the customer market soon. That is why some operations are still needed to convert High Dynamic Range (HDR) contents to Low Dynamic Range (LDR) ones. These operations are performed using Tone Mapping Operators (TMOs). Technology regarding the tone mapping of still pictures has been thoroughly explored and several satisfying solutions [DCWP02] have been designed. However, most of the existing operators are not designed to handle video sequences.

Applying a TMO separately to frames of a video sequence is source of temporal incoherency. In this article, we outline two types of temporal incoherency: flickering and temporal brightness incoherency. Over small time intervals, rapid changes in the scene introduce flickering artifacts. Several solutions exist to overcome the flickering problem [KUWS03, RIH04, MDK08, LK07, GKEE11]. These solutions use temporally close frames to smooth out abrupt changes of luminance. However, longer time ranges introduce temporal brightness incoherency which is not addressed by these solutions. One solution to solve this problem is the Brightness Coherency (BC) algorithm [BBC*12] that relies on a post-processing operation. This technique preserves the brightness ratio between each frame and the brightest frame of the scene.

This article is organized as follows. Section 2 details several solutions to preserve temporal coherency when tone mapping a HDR video sequence. Section 3 presents some results provided by these solutions. Finally, section 4 concludes this paper.

2. Temporal Coherency

Recall that the main issue in video tone mapping is to preserve the temporal coherency. two main types of temporal incoherency are usually considered: flickering and brightness incoherency. Flickering artifacts are due to rapid changes of the tone map curve, resulting in different mappings for similar HDR luminance levels in successive frames. Brightness incoherency occurs when the tone map curve of each frame of a video sequence is computed separately. Consequently, frames perceived as the brightest in the HDR sequence are not necessarily the brightest in the LDR sequence.

In this section, we present solutions that solve either of these two issues. Most of these solutions extend existing TMOs to video sequences. First we present the Photographic Tone Reproduction (PTR) [RSSF02] operator along with two techniques [KUWS03, RIH04] that reduce the flickering artifacts. Then, we detail the Display Adaptive [MDK08] operator and its temporal extension to reduce flickering artifacts. In section 2.3, we describe the Gradient Domain [FL02] operator. We show how this method has been extended to reduce flickering artifacts using motion estimation [LK07]. In section 2.4, we present a technique that adapts to any TMO and aims at reducing flickering artifacts using only on the LDR sequence [GKEE11]. Finally, we present the Brightness Coherency (BC) [BBC*12] that preserves temporal brightness coherency and adapts to any TMO.

2.1. Photographic Tone Reproduction

The PTR algorithm uses a system designed by Adams [Ada81] to rescale HDR frames at a defined exposure:

$$L_s = \frac{a}{\kappa} \cdot L_w \,, \tag{1}$$

$$\kappa = exp\left(\frac{1}{N} \cdot \sum_{n=1}^{N} log(\delta + L_w(n))\right) , \tag{2}$$

where a is the chosen exposure, L_w the HDR luminance and L_s the scaled luminance. The key value κ is a subjective indication of the image's overall brightness level. It is computed using Equation (2), where δ is a small value to avoid singularity and N the number of pixels in the image. The tone map curve is a sigmoid function given by Equation (3):

$$L_d = \frac{L_s}{1 + L_s} \cdot \left(1 + \frac{L_s}{L_{white}^2}\right) , \tag{3}$$

where L_{white} is used to burn out areas with high luminance value and L_d is the tone map LDR luma. Two parameters (a and L_{white}) are then necessary to compute the TMO results. In [RSSF02], a is set to 18% and L_{white} to L_max.

The main issue with this algorithm is that flickering artifacts appear for abrupt changes of the key value. To avoid these artifacts, Kang et al. [KUWS03] proposed to filter the current frame's key value using a set of the n previous frames key values. Ramsey et al. [RIH04] further refined this idea by making n adaptive. This adaptation discards outliers using a min/max threshold. We refer to this technique as $PTR + R$ hereinafter.

2.2. Display Adaptive Tone Mapping

Mantiuk et al. [MDK08] proposed a TMO that provides the least perceptually distorted LDR picture on a targeted display. Similarly to Tumblin and Rushmeier [TR93], whose operator's goal is to ensure that the scene and display brightnesses match, Mantiuk et al. compare the visual response of the Human Visual System (HVS) to the display-adapted LDR image with that of the original HDR image. The minimization of the residual error between these responses results in a piece-wise linear tone map curve.

However, this minimization is computed on a per frame basis and consequently inherits all the problems related to video tone mapping. To solve this issue, the authors propose a temporal adaptation using a low-pass filter applied to the nodes of the piece-wise tone curve. Similarly to the previous solutions [KUWS03, RIH04], the tone map curve is smoothed. One consequence is that the minimization of the perceptual distortion is not preserved during this temporal adaptation. We refer to this technique as DA_T hereinafter.

2.3. Gradient Domain Tone Mapping

Another TMO that has been adapted to video sequences is the Gradient Domain (GD) [FL02]. This technique compresses the dynamic range by attenuating the spatial gradient depending on its intensity. A spatially variable mapping function attenuates large gradients at various scales while preserving fine details. The spatial gradients are computed in the log domain. The final tone mapped image is obtained by solving a set of Poisson equations using the attenuated gradient field.

Similarly to the previous techniques, applying this TMO to a video sequence results in flickering artifacts. That is why, Lee et al. [LK07] proposed a technique to adapt this TMO to video tone mapping. They first perform a pixelwise motion estimation on the original HDR sequence. This motion field is used as a constraint for the temporal coherence between the LDR frames. We refer to this technique as GD_T hereinafter.

2.4. Flicker Reduction in Video Tone Mapping

Another technique [GKEE11] reduces flickering artifacts by post-processing the output of any TMO. This method uses a brightness threshold to detect flickering artifacts in successive frames of a video sequence. As soon as an artifact is located, it is reduced using an iterative brightness adjustment based on the brightness threshold.

2.5. Brightness Coherency for Video Tone Mapping

To better preserve the temporal brightness coherency, Boitard et al. [BBC*12] developed a method which adapts to any TMO through a post-processing operation. We refer to this technique as BC hereinafter. This technique uses the frame key value κ (computed with equation 2) to preserve the HDR brightness ratio in the tone map LDR sequence.

The HDR brightness ratio is equal to the LDR brightness ratio if

$$\frac{\kappa_F^{HDR}}{\kappa_V^{HDR}} = \frac{\kappa_F^{LDR'}}{\kappa_V^{LDR}} \,, \tag{4}$$

where κ_F^{HDR} is the current HDR frame key value and κ_V^{HDR} the key value of the brightest frame. To satisfy

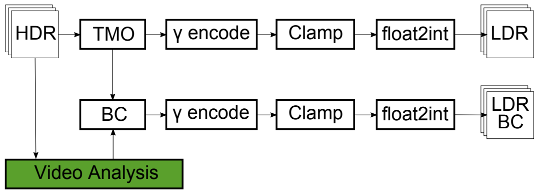

Figure 1: Tone mapping a video sequence with and without the BC algorithm. The video analysis (green box) is performed as a preprocessing step. "γ encode" is required to compensate for the use of the bits relative to how humans perceive light (usually $\gamma = 1/2.2$). "Clamp" removes the values outside the range [0, 1]. "float2int" converts floating point values into integers.

Equation (4), the tone map luma L_d is scaled according to Equation (5) to get the new tone map luma L'_d:

$$L'_d = L_d \cdot \left(Of + (1 - Of) \cdot \frac{\kappa_F^{HDR} \cdot \kappa_V^{LDR}}{\kappa_V^{HDR} \cdot \kappa_F^{LDR}} \right) = R \cdot L_d \quad (5)$$

where κ_F is computed using Equation (2), L_d and L'_d are respectively the tone map LDR luma and the post-processed LDR luma. Of is a user-defined parameter to prevent low scale ratio. In order to determine the anchor, i.e. the frame with the maximum HDR frame key value, a video analysis is performed prior to the tone mapping operation. Figure 1 depicts the workflow used to tone map a video with and without preserving the brightness coherency. This solution preserves the relative HDR brightness levels in the LDR tone map results.

3. Results

In this section, we present some results when applying TMOs to two HDR sequences. The first one, called Temple, is a computer generated sequence composed of 260 1920x1080 High Definition (HD) frames. Figure 2 presents some results on the reduction of flickering artifacts. Figure 2a and 2b show results with PTR and $PTR + R$ respectively. With PTR, each frame is tone mapped with a different brightness level which increases from frame to frame. As it uses temporal filter, the $PTR + R$ slows down this increase. Figure 2d shows more coherent frames (the statue and the wall) than the ones shown in figure 2c.

The second sequence is called UnderBridgeHigh and is composed of 250 HD frames at the frequency of 25 frames per second (fps). We generated these frames using a 5 pixel vertical traveling inside a high resolution HDR image. Such a sequence is particularly helpful to test the temporal brightness coherency as we have the same HDR values from frame to frame. Figure 3 presents some results with (fig. 2c) and without (fig. 2a and 2b) preserving temporal brightness coherency. Notice how the brightness of the downside of the bridge changes when the brightness coherency is not preserved. For more results see `http://people.irisa.fr/Ronan.Boitard/articles/2013/TemporalCoherencyVideos.zip`.

4. Conclusion

In this article, We first described the two main issues when tone mapping HDR video sequences. We then reviewed different techniques to extend TMOs to video tone mapping. Some techniques apply to only one TMO while others can adapt to any TMO. To conclude, the provided results show that preserving temporal coherency increases the quality of tone mapped sequences.

References

[Ada81] Adams A.: *The Print: The Ansel Adams Photography Series 3*. Little, Brown and Compagny, 1981.

[BBC*12] Boitard R., Bouatouch K., Cozot R., Thoreau D., Gruson A.: Temporal coherency for video tone mapping. In *Proc. SPIE 8499, Applications of Digital Image Processing XXXV* (Oct. 2012), van Eijk A. M. J., Davis C. C., Hammel S. M., Majumdar A. K., (Eds.), pp. 84990D–84990D–10.

[DCWP02] Devlin K., Chalmers A., Wilkie A., Purgathofer W.: Tone Reproduction and Physically Based Spectral Rendering. *Eurographics 2002* (2002).

[FL02] Fattal R., Lischinski D.: Gradient domain high dynamic range compression. *ACM Transactions on Graphics* (2002).

[GKEE11] Guthier B., Kopf S., Eble M., Effelsberg W.: Flicker reduction in tone mapped high dynamic range video. In *Proc. of IS&T/SPIE Electronic Imaging (EI) on Color Imaging XVI: Displaying, Processing, Hardcopy, and Applications* (2011), pp. 78660C–78660C–15.

[KUWS03] Kang S. B., Uyttendaele M., Winder S., Szeliski R.: High dynamic range video. *ACM Trans. Graph. 22*, 3 (July 2003), 319–325.

[LK07] Lee C., Kim C.-S.: Gradient Domain Tone Mapping of High Dynamic Range Videos. In *2007 IEEE International Conference on Image Processing* (2007), no. 2, IEEE, pp. III – 461–III – 464.

[MDK08] Mantiuk R., Daly S., Kerofsky L.: Display adaptive tone mapping. *ACM Transactions on Graphics 27*, 3 (Aug. 2008), 1.

[RIH04] Ramsey S., III J. J., Hansen C.: Adaptive temporal tone mapping. *Computer Graphics and Imaging - 2004*, 3 (2004), 3–7.

[RSSF02] Reinhard E., Stark M., Shirley P., Ferwerda J.: Photographic tone reproduction for digital images. *ACM Trans. Graph. 21*, 3 (July 2002), 267–276.

[TR93] Tumblin J., Rushmeier H.: Tone reproduction for realistic images. *Computer Graphics and Applications IEEE* (1993).

(a) Reinhard et al. PTR [RSSF02]

(b) Ramsey et al. $PTR+R$ [RIH04].

(c) Mantiuk et al. without flickering reduction DA [MDK08]

(d) Mantiuk et al. with flickering reduction DA_t [MDK08]

Figure 2: Tone mapped frames 108, 110 and 112 of the Temple sequence.

(a) Ramsey et al. $PTR+T$ [RIH04].

(b) Mantiuk et al. with flickering reduction DA_t [MDK08]

(c) Ramsey et al. and Boitard et al. $PTR+R+BC$ [RIH04, BBC*12].

Figure 3: Tone mapped frames 10, 50, 100, 150 and 200 of the UnderBridgeHigh sequence.