REAL-TIME TEMPORALLY COHERENT LOCAL HDR TONE MAPPING

Simone Croci+, Tunç Ozan Aydın*, Nikolce Stefanoski*, Markus Gross*+, Aljoscha Smolic**

Disney Research*, ETH Zurich+

ABSTRACT

Subjective studies showed that most HDR video tone mapping operators either produce disturbing temporal artifacts, or are limited in their local contrast reproduction capability. Recently, both these issues have been addressed by a novel temporally coherent local HDR tone mapping method, which has been shown, both qualitatively and through a subjective study, to be advantageous compared to previous methods. However, this method's high-quality results came at the cost of a computationally expensive workflow that could only be executed offline. In this paper, we present a modified algorithm which builds upon the previous work by redesigning key components to achieve real-time performance. We accomplish this by replacing the optical flow based per-pixel temporal coherency with a tone-curve-space alternative. This way we eliminate the main computational burden of the original method with little sacrifice in visual quality.

***Index Terms*—** Tone Mapping, Real-time, Temporal Coherency

1. INTRODUCTION

Temporal coherency has been a chronic problem of High Dynamic Range (HDR) video tone mapping methods. The lack of temporal coherency in tone mapped content often manifests itself as abrupt brightness changes between consecutive frames. To make matters worse the human visual system is highly sensitive to such temporal artifacts, which had severely restricted the use of video tone mapping operators in practice. To alleviate this problem, a number of video tone mapping operators restricted themselves to using *global* tone curves. However these global video tone mapping operators offer very limited artistic control over the results, and tend to loose a significant amount of scene detail. Due to these limitations, the outcome of global tone mapping methods often appear as though they were shot in Standard Dynamic Range (SDR) to begin with[1]. Such limitations have been documented by a recent subjective study [1].

More recently, a new HDR video tone mapping operator method has enabled *local* video tone mapping in a temporally coherent manner [2]. This method overcame the limitations of previous techniques, in that it has the capability of reproducing a significant amount of scene details while still producing a temporally coherent outcome. As such, Aydın et al.'s [2] method allowed a great amount of artistic control over the tone mapping process, and has been shown to produce unique visual styles.

While effective, the temporal coherency in Aydin et al. [2] comes at the cost of computation time. Their method relies on per-pixel temporal filtering over motion paths that are estimated by dense optical flow fields and image domain warping. Since all these components are computed on a user-defined temporal window (usually around 20 frames), Aydin et al. [2] is not suitable for real-time applications. In fact, the authors report the per-frame processing time is 1.5 seconds at 720p resolution without factoring in the time required for optical flow computation. While the computational cost would significantly vary depending on the selected algorithm, it is safe to assume that the processing time should be doubled when accounting for optical flow computation.

In this work, we revisit Aydin et al. [2] (which we henceforth refer to as the *offline method*) and make a number of key modifications to ease the original method's computational burden while retaining the visual quality as much as possible. The resulting *Real-time Video Tone Mapper (rtvtm)* enables local HDR video tone mapping in real-time (24fps) at 720p resolution on a current desktop PC.

In the remainder of this paper we first briefly review relevant prior work (Section 2), summarize the offline method's temporally coherent video tone mapping approach (Section 3), and discuss the proposed modifications to the original algorithm for achieving real-time performance with minimal loss of image quality (Section 4). Section 5 and the supplemental video[2] show a fully functional prototype implementation of our method, as well as side-by-side comparisons between the offline method and the proposed real-time method, whereas in Section 6 we discuss cases in which the real-time method would fail.

2. PREVIOUS WORK

HDR image tone mapping has been extensively studied in the literature. For a comprehensive overview of the field we refer

[1]This is not surprising as the camera response function is essentially a global tone mapping operator

[2]http://zurich.disneyresearch.com/~taydin/rttmo.mov

the reader to Reinhard et al. [3]. Video tone mapping, on the other hand has received far less interest when compared to image tone mapping. One of the main road blocks of video tone mapping research has been the difficulty of obtaining high-quality HDR footage. This issue has been alleviated with the availability of various commercial camera systems that can capture HDR natively (such as Red Epic Dragon, Sony 40 F55 and F65, and ARRI Alexa XT), as well as the emergence of research prototypes that demonstrated similar capabilities [4].

Within the body of tone mapping research, a number of tone mapping operators (TMOs) presented temporal components that enable processing HDR video. Among these TMOs, global operators [5, 6, 7, 8, 9] as well as S-shaped camera response curves are known to have decent temporal coherence, but they lack in contrast reproduction due to the absence of local tone curve adjustment. Local operators [10, 11, 12, 13, 14] often maintain high contrast at the cost of noticeable temporal artifacts. The advantages and shortcoming of these operators have been discussed and subjectively evaluated by Eilertsen et al. [1].

Other work in video tone mapping [15, 16, 17] focused on extending the photographic TMO's tone curve [18] with temporal filtering for computing a temporally coherent key value. In their case, temporal coherence is achieved by temporally filtering tone mapping parameters and ensuring that they vary smoothly over time. In this work we also take a similar approach, but adopt it to be used with a local tone mapping technique. Another temporal extension has been proposed for the gradient domain image TMO [19]. However, the results of the methods are demonstrated on a limited number of examples, and they have not been tested extensively on public HDR video data sets.

Subjective studies and surveys that evaluated the HDR video tone mapping [1, 20, 21] suggested utilizing temporal processing for avoiding visual artifacts, and local processing for maintaining a high level of detail and contrast as two ares for improvement. Since then, Aydın et al. [2] proposed a method that addressed temporal artifacts using a per-pixel spatio-temporal filtering technique, and Eilertsen et al. [22] focused on real-time processing and reducing the visibility of camera noise by carefully adjusting the tone curve.

3. OFFLINE METHOD

Figure 1 shows the main processing steps of the offline method [2], which we briefly discuss in this section. The input to this method is an HDR video sequence whose luminance channel is assumed to be stored using at least 10 bits. In fact the authors used footage obtained from an Arri Alexa XT captured in 12 bit logarithmic uncompressed ARRIRAW format. The first step in the offline tone mapping pipeline is optical flow computation, both in forward and backward directions for the entire video sequence. While fast optical flow estimation methods do exist, high quality optical flow methods (such as Zimmer et al. [23] that has been used in Aydin et al.'s workflow [2]) are often computationally intensive. Next, the input video is filtered using the optical flow estimates both temporally and spatio-temporally in separate threads. The spatio-temporal filtering is performed on a temporal neighborhood of user-selected length (which was selected as 21 in the original method). Optical flow estima-

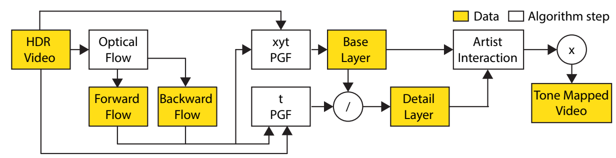

Fig. 1. Offline HDR video tone mapping workflow proposed by prior work. See text for details.

tion and per-pixel spatio-temporal filtering together create the main performance bottleneck of the original tone mapping workflow. While the spatio-temporal filtering operations are performed efficiently using the *permeability guided filtering (PGF)* framework, the size of the temporal neighborhood and the necessary per-pixel warping operations degrade the overall filtering performance. As we show later in Section 4, replacing these components with a more efficient alternative makes a real-time tone mapping workflow possible with minimal loss in visual quality.

In the remaining steps of the original offline tone mapping workflow, first the detail layer is computed by dividing the temporally filtered input video by the base layer. Using a GUI, an operator can adjust the tone mapping parameters to obtain the desired visual style. The base and detail layers are combined according to the selected parameters, which results in the tone mapped video. This latter part of the tone mapping pipeline is quite efficient, and is not changed in the proposed real-time tone mapping workflow, which we discuss in the next section.

4. REAL-TIME METHOD

The data flow diagram of the proposed real-time tone mapping method is illustrated in Figure 2, where the changes over the previous offline method are highlighted with blue. As we mentioned earlier in Section 2, while the PGF filter is computationally efficient, the sheer volume of the temporal neighborhood as well as the per-pixel warping operations prevents temporal PGF filtering from being suitable to real-time applications. In our real-time framework, the PGF filter is applied only spatially (xy PGF) to each video frame independently. The result of the spatial filtering gives a first approximation of the base layer, which is not yet temporally coherent. Similar to the offline method the detail layer is obtained by dividing the input HDR video by the base layer.

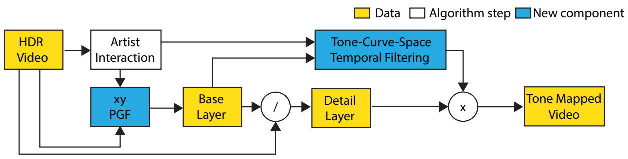

Fig. 2. Real-time HDR video tone mapping workflow with the proposed modifications to the previous state-of-the-art. See text for details.

In our real-time framework, temporal coherence is enforced by a process that we call tone-curve-space temporal filtering. This process is based on the observation that most of the visually disturbing temporal incoherencies in video tone mapping are caused by abrupt changes in the tone curve parameters. Tone curve parameters are often functions of some image statistics such as the mean luminance of the frame. In HDR video sequences, such statistics can change notably between consecutive frames, which in turn causes the tone-curve to behave in a temporally incoherent manner. If we can ensure that the tone-curve parameters change smoothly despite the rapidly changing image statistics, we can eliminate these visual artifacts. This idea has been investigated in previous work in the context of global tone mapping, where the tone curve with temporally smooth parameters is applied to the input HDR video [17]. Temporal smoothing can be achieved by using the convex combination of the tone-curve parameter computed for the current frame ($\hat{p}_t$) and the corresponding tone-curve parameter from the previous frame (p_{t-1}) as the parameter value for the current frame (p_t):

$$p_t = \alpha \, p_{t-1} + (1 - \alpha) \, \hat{p}_t, \quad (1)$$

where α is the weight of the parameter value from the previous frame. In our implementation, we use the same auto-regressive process, which in principle could also be extended to an auto-regressive moving-average process.

On the negative side, Kiser et al.'s [17] approach of applying a temporally smooth tone curve directly to the HDR video severely limits artistic freedom, since using their method one can only perform global tone mapping. Therefore in our workflow we propose decomposing the input HDR video first into base and detail layers, and performing the tone-curve parameter smoothing on the base layer. Doing so has two major advantages. First and foremost, it significantly enhances artistic freedom by enabling local tone mapping operations. Treating the base and detail layers separately allows precise control over the visibility of local details. Second, tone mapping parameters are more stable to begin with when they are computed on the base layer instead of the input HDR image. This is due to the edge-aware PGF filter reducing local variation and removing outliers.

Once we obtain a temporally coherent base layer through tone-curve-space filtering, we simply multiply the detail layer back to obtain the final tone mapped frame. Importantly, unlike the offline method (Figure 1) where artist interaction is performed at the very end, in the proposed real-time framework artistic interaction is possible from the start. In the offline method, due to the computational cost of the spatio-temporal filtering, the filtering parameters have to be set in advance and the base and detail layers have to be pre-computed before any artistic interaction can take place. Thanks to the efficiency of the proposed framework the artist also has the ability to adjust filtering parameters in real-time, and no pre-processing is necessary.

Fig. 3. Comparison between the proposed real-time tone mapping workflow (left column) and the previous offline method (right column). The results suggest that the real-time workflow has the potential to match the offline method in terms of image quality.

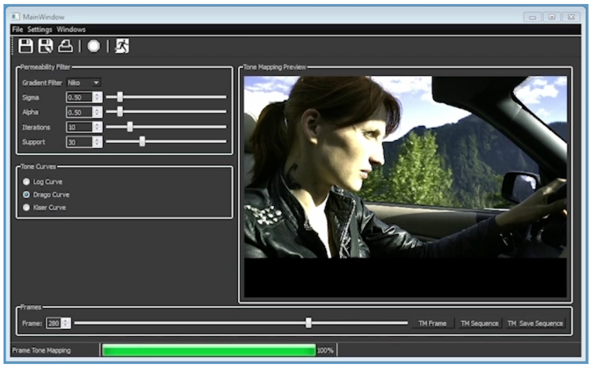

Fig. 4. The graphical user interface of our prototype implementation. Users can change parameters of the permeability guided filtering as well as the tone curve. Parameters can be set on a desired number of key frames, which our system then applies to the entire video sequence by interpolation. All changes are reflected at interactive rates through a rendering of the currently selected frame that is tone mapped using the desired parameters.

5. RESULTS

Figure 3 shows a visual comparison between the offline method [2] (right column) and the proposed real-time method (left column). These results show that differences between the two methods are subtle and only noticeable after careful visual inspection. For most practical purposes, the proposed real-time method achieves comparable image quality with respect to the offline method, with the advantage of being significantly more efficient. Full video sequences of these comparisons, as well as additional results can be found in the accompanying video, which also shows that the proposed method achieves a high level of temporal stability comparable to the offline method. [3].

The results we present in this paper are generated using a prototype implementation of our method. Figure 4 shows a screen-shot during a typical tone mapping session. Our implementation allows the user to change tone mapping parameters on a key frame level while giving feedback at interactive rates. Our currently unoptimized research code (written in C++ and Cuda) can process 720p frames at 24 fps on a desktop PC using the parameter configuration to generate our results. This performance can be further enhanced by limiting the number of PGF iterations and reducing the maximum neighborhood size of the PGF, but in our experience this may compromise image quality. Conversely, for applications where lower frame-rates are acceptable, the PGF parameters can be adjusted for higher image quality.

[3]For video results please refer to our supplemental video at `http://zurich.disneyresearch.com/~taydin/rttmo.mov`

6. LIMITATIONS

The limitations of the proposed real-time tone mapping method are mostly related to the simplified temporal filtering process. Due to the absence of per-pixel filtering as in the offline method, our method does not automatically remove camera noise. Our method also can only prevent temporal incoherencies due to tone mapping. If the input HDR video is temporally unstable because of the limitations in the capture process, these incoherencies carry over to the final tone mapped video.

7. CONCLUSION

We proposed a real-time approximation to the state-of-the-art HDR video tone mapping method by Aydın et al. [2]. We showed that the proposed real-time approximation can retain both the image quality and temporal coherence of the original offline method. Additionally, our method allows significantly higher level of artistic freedom, since it also allow control over filtering parameters and does not rely on pre-processing as the offline method. Our prototype implementation features an intuitive GUI which allows total control over the tone mapping process, including key-frame based parameter adjustment and interpolation. Due to being both high quality and real-time, we believe that the proposed method streamlines the HDR tone mapping process during production significantly, and paves the way for HDR broadcast applications.

8. REFERENCES

[1] Gabriel Eilertsen, Robert Wanat, Rafal K. Mantiuk, and Jonas Unger, “Evaluation of tone mapping operators for hdr-video,” *Computer Graphics Forum*, vol. 32, no. 7, pp. 275–284, 2013.

[2] Tunç O. Aydın, Nikolce Stefanoski, Simone Croci, Markus Seidel, Gross, and Aljoscha Smolic, “Temporally coherent local tone mapping of hdr video,” in *ACM Trans. Graph. (Proc. of SIGGRAPH Asia)*, 2014, vol. 33(6), Article 196.

[3] Erik Reinhard, Greg Ward, Sumanta Pattanaik, Paul Debevec, Wolfgang Heidrich, and Karol Myszkowski, *HDR Imaging - Acquisition, Display, and Image-Based Lighting, Second Edition*, Morgan Kaufmann, 2010.

[4] Michael D. Tocci, Chris Kiser, Nora Tocci, and Pradeep Sen, “A versatile hdr video production system,” *ACM Trans. Graph.*, vol. 30, no. 4, pp. 41:1–41:10, 2011.

[5] James A. Ferwerda, Sumanta N. Pattanaik, Peter Shirley, and Donald P. Greenberg, “A model of visual adaptation for realistic image synthesis,” in *Proceedings of the 23rd Annual Conference on Computer Graphics*

and Interactive Techniques, 1996, SIGGRAPH '96, pp. 249–258.

[6] Sumanta N. Pattanaik, Jack Tumblin, Hector Yee, and Donald P. Greenberg, "Time-dependent visual adaptation for fast realistic image display," in *Proc. of Conf. on Computer Graphics and Interactive Techniques*, 2000, SIGGRAPH '00, pp. 47–54.

[7] Piti Irawan, James A. Ferwerda, and Stephen R. Marschner, "Perceptually based tone mapping of high dynamic range image streams," in *Proceedings of Eurographics Conf. on Rendering Techniques*, 2005, EGSR'05, pp. 231–242.

[8] J. H. Van Hateren, "Encoding of high dynamic range video with a model of human cones," *ACM Trans. Graph.*, vol. 25, no. 4, pp. 1380–1399, 2006.

[9] RafałMantiuk, Scott Daly, and Louis Kerofsky, "Display adaptive tone mapping," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 68:1–68:10, 2008.

[10] Patrick Ledda, Luis Paulo Santos, and Alan Chalmers, "A local model of eye adaptation for high dynamic range images," 2004, AFRIGRAPH, pp. 151–160.

[11] Eric P. Bennett and Leonard McMillan, "Video enhancement using per-pixel virtual exposures," *ACM Trans. Graph.*, vol. 24, no. 3, 2005.

[12] Alexandre Benoit, David Alleysson, Jeanny Herault, and Patrick Le Callet, "Spatiotemporal tone mapping operator based on a retina model.," in *CCIW*. 2009, vol. 5646 of *Lecture Notes in Computer Science*, pp. 12–22, Springer.

[13] Erik Reinhard, Tania Pouli, Timo Kunkel, Ben Long, Anders Ballestad, and Gerwin Damberg, "Calibrated image appearance reproduction," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 201:1–201:11, Nov. 2012.

[14] Ronan Boitard, Kadi Bouatouch, Remi Cozot, Dominique Thoreau, and Adrien Gruson, "Temporal coherency for video tone mapping," 2012, vol. 8499, pp. 84990D–84990D–10.

[15] Sing Bing Kang, Matthew Uyttendaele, Simon Winder, and Richard Szeliski, "High dynamic range video," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 319–325, 2003.

[16] Shaun David Ramsey, J Thomas Johnson III, and Charles Hansen, "Adaptive temporal tone mapping," in *Proceedings of the 7th IASTED International Conference on Computer Graphics and Imaging*, 2004, pp. 124–128.

[17] Chris Kiser, Erik Reinhard, Mike Tocci, and Nora Tocci, "Real time automated tone mapping system for HDR video," in *Proceedings of the IEEE International Conference on Image Processing*, Orlando, Florida, September 2012, IEEE, pp. 2749–2752, IEEE.

[18] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 267–276, 2002.

[19] Chul Lee and Chang-Su Kim, "Gradient domain tone mapping of high dynamic range videos," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, 2007, vol. 3, pp. III – 461–III – 464.

[20] Josselin Petit and Rafal K. Mantiuk, "Assessment of video tone-mapping: Are cameras s-shaped tone-curves good enough?," *Journal of Visual Communication and Image Representation*, vol. 24, no. 7, pp. 1020–1030, 2013.

[21] Ronan Boitard, Thoreau Dominique, and Bouatouch Kadi, "Temporal coherency in video tone mapping , a survey," in *HDRi2013 - First International Conference and SME Workshop on HDR imaging*, 2013, pp. 1–6.

[22] Gabriel Eilertsen, Rafał K. Mantiuk, and Jonas Unger, "Real-time noise-aware tone mapping," *ACM Trans. Graph.*, vol. 34, no. 6, pp. 198:1–198:15, Oct. 2015.

[23] Henning Zimmer, Andrés Bruhn, and Joachim Weickert, "Optic flow in harmony," *Int. J. Comput. Vision*, vol. 93, no. 3, pp. 368–388, July 2011.