Chapter 6 Video Tone Mapping

Ronan Boitard ^{a,b} Rémi Cozot ^a Kadi Bouatouch ^a

^aIRISA, 263 Avenue du Général Leclerc, 35000 Rennes, France

^b Technicolor, 975 Av. des Champs Blancs, 35576 Cesson-Sevigne, France

Each pixel of a LDR image is stored using color components, usually three. The way LDR displays interpret these components to shape color is defined through a display-dependent color space, for example the BT.709 (ITU 1998) or BT.2020 (ITU 2012). On the contrary, HDR pixels represent, in floating point values, the captured physical intensity of light in cd/m^2 . They can also represent relative floating point values. Hence, adapting an HDR image to a LDR display amounts to retargeting physical values, with a virtually unlimited bit-depth, to a constrained space (2^{2n}) chromaticity values over 2^n tonal level, n being the targeted bitdepth). This operation, which ensures backward compatibility between HDR content and LDR displays, is called tone mapping. The bit-depth limitation means that many similar HDR values will be tone mapped to the same LDR one. Consequently, contrast between neighboring pixel as well as between spatially distant areas will be reduced. Furthermore, LDR displays have a low peak luminance value when compared to the luminance that a real scene can achieve. Consequently, captured color information will have to be reproduced at different luminance levels.

In a nutshell, tone mapping an HDR image amounts to finding a balance between the preservation of details, the spatial coherency of the scene and the fidelity of reproduction. This balance is usually achieved by taking advantage of the many weaknesses of the human visual system. Furthermore, the reproduction of a scene can sometimes be constrained by an artist or application intent. That is why, a lot of Tone Mapping Operators (TMOs) have been designed with different intents, from simulating the human vision to achieving the best subjective quality (Reinhard et al. 2010, Myszkowski et al. 2008, Banterle et al. 2011).

In the early nineties, the main goal of tone mapping was to display computer generated HDR images on traditional display. Indeed, using a simple gamma mapping was not enough to reproduce all the information embedded in HDR images. Although, throughout the years TMOs addressed different types of applications, most of them still focused on finding the optimal subjective quality as the many subjective evaluation attest (Drago et al. 2003, Kuang et al. 2007, Cadík et al. 2008). However, due to the lack of high quality HDR video content, the temporal aspect of tone mapping has been dismissed for a long time. Thanks to recent developments in the HDR video acquisition field (Tocci et al. 2011, Kronander et al. 2013, 2014), more and more HDR video content are now publicly available (Unger 2013, Krawczyk 2006, Digital Multimedia Laboratory 2014, Lasserre et al. 2013, IRISA 2015). In the near future, many applications such as real-time TV broadcasting, cinema movies and user-generated videos will require video tone mapping.

In this chapter, we propose to evaluate the status of the video tone mapping field when trying to achieve a defined subjective quality level. Indeed, naively applying a TMO to each frame of an HDR video sequence leads to temporal artifacts. That is why we describe, in Section 1, different types of temporal artifacts found out through experimentation. Then, Section 2 introduces state of the art Video TMOs, that is to say TMOs that rely on information from frames other than the one being currently tone mapped. In Section 3, we present two new types of temporal artifact that are introduced by Video TMOs: temporal contrast adaptation and ghosting artifacts. Finally, Section 4 presents in more details two recently published Video TMOs.

1 Temporal Artifacts

Through experimentation with different HDR video sequences, we encountered several different types of temporal artifact. In this section we focus only on those occurring when applying naively a TMO to each frame of an HDR video sequence. We propose to classify these artifacts into six categories:

- Section 1.1. Global Flickering Artifacts,
- Section 1.2. Local Flickering Artifacts,
- Section 1.3. Temporal Noise,
- Section 1.4. Temporal Brightness Incoherency,
- Section 1.5. Temporal Object Incoherency,
- Section 1.6. Temporal Hue Incoherency.

This section provides a description of those artifacts along with some examples. Note that all the results are provided using TMOs that do not handle time-dependency, namely TMOs that rely only on statistics of the current frame to tone map.

1.1 Global Flickering Artifacts

Global Flickering Artifacts (GFA) are well known in the video tone mapping literature and are characterized by abrupt changes, in successive frames, of the overall brightness of a tone mapped video sequence. These artifacts appear because TMOs adapt their mapping using image statistics that tend to be unstable over time. Analyzing the overall brightness of each frame over time is usually sufficient to detect those artifacts. An overall brightness metric can be, for example, the mean luma value of an image. Note that if computed on HDR images, the luminance channel must first be perceptually encoded, using for example a log transform as proposed in (Reinhard et al. 2002), before averaging.

To illustrate this type of artifact, we plotted in Fig. 1 the overall brightness indication for both the HDR and the tone mapped



Fig. 1. Example of Global Flickering Artifacts. The overall brightness in the HDR sequence is stable over time while many abrupt variations occur in the LDR sequence. As luminance and luma have different dynamic range, they have been scaled to achieve a meaningful comparison.



Fig. 2. Global Flickering Artifacts due to the use of the 99% percentile on two successive frames of the *Desert* sequence using the Multi-Scale tone mapping operator (Farbman et al. 2008).

sequences. Note how the evolution of the overall brightness is stable over time in the HDR sequence while abrupt peaks occur in the LDR one. These artifacts appear because one of the TMO's parameter, that adapts to each frame, varies over time. Fig. 2 illustrates such an artifact occurring in two successive frames of a tone mapped video sequence. The overall brightness has changed because the relative area of the sky in the second frame is smaller, hence reducing the chosen normalization factor (99th percentile).

To summarize, GFA mostly occur when using TMOs that rely on content adaptive parameters that are unstable over time. They are usually considered as the most disturbing artifacts presented in this section, which explains why they have received a lot of attention as will be seen in Section 2.

1.2 Local Flickering Artifacts

Local Flickering Artifacts (LFA) correspond to the same phenomenon as its global counterpart but on a reduced area. They appear mostly when using TMOs that map a pixel based on its neighborhood, namely local TMOs. Small changes of this neighborhood, in consecutive frames, may result in a different mapping. Edge-aware TMOs are particularly prone to such artifacts as they decompose an HDR image into a base layer and one or more detail layers. As each layer is tone mapped independently, a difference in the filtering in successive frames results in LFA.

The top row of Fig. 3 represents a zoom on a portion of the computed base layer of 3 successive frames. Note how the edges are less filtered out in the middle frame compared to the other two. Applying the bilateral filter (Durand & Dorsey 2002) operator results in a local flickering artifact in the tone mapped result (bottom row). Although visible in a video sequence, it is tenuous to represent local flickering artifacts using successive frames. A side effect of LFA is that they modify the saliency of the tone map sequence as the eye is attracted by these changes of brightness on small areas.

1.3 Temporal Noise

Temporal Noise (TN) is a common artifact occurring in digital video sequences. Noise in digital imaging is mostly due to the camera and is particularly noticeable in low light conditions. On images, camera noise has a small impact on the subjective quality, however for video sequences its variation over time makes it more noticeable. It is the reason why denoising algorithms (Brailean



Fig. 3. Example of Local Flickering Artifacts when applying an the Fast Bilateral Filtering TMO (Durand & Dorsey 2002) to 3 consecutive frames. Top row: base layer computed using the bilateral filter. Bottom row: corresponding tone mapped result.

et al. 1995) are commonly applied to video sequences to increase their subjective quality.

As most TMOs aim at reproducing minute details, they struggle to distinguish information from noise. Consequently most of current TMOs increase the noise rather than reducing it. Local TMOs are particularly prone to such artifacts as they aim at preserving details even in dark areas which tend to be quite noisy. Furthermore, noise is usually reproduced at a higher luma level compared to a native LDR image which makes the noise more visible. An example of temporal noise enhanced by the application of a local TMOs is illustrated in Fig. 4.

1.4 Temporal Brightness Incoherency (TBI)

Temporal Brightness Incoherency (TBI) artifacts occur when the relative brightness between two frames of an HDR sequence is not preserved during the tone mapping. As a TMO uses for each frame all its available range, the temporal brightness relationship between frames is not preserved throughout the tone mapping operation. Consequently, a frame perceived as the brightest in the HDR sequence is not necessarily the brightest in the LDR one.



Fig. 4. Example of Temporal Noise amplification due to the application of a local Edge Aware TMO (Gastal & Oliveira 2011) (left) compared to the global Photographic Tone Reproduction TMO (Reinhard et al. 2002) (right).

For example, TBI occur when a change of illumination condition in the HDR sequence is not preserved during the tone mapping. Consequently, temporal information (i.e. the change of condition) is lost, which changes the perception of the scene (along with its artistic intent). Fig. 5 illustrates a TBI artifact by plotting the overall brightness of both HDR and LDR sequences. Note that, although the mean value greatly varies in the HDR sequence, it remains stable in the LDR one. This is due to the fact that a TMO searches for the best exposure for each frame. As it has no information on temporally close frames, the change of illumination is simply dismissed and the best exposure is defined independently (usually in the middle of the available range). Fig. 6 illustrates an example of a Temporal Brightness Incoherency occurring in consecutive frames of a tone mapped video sequence. The top row displays the HDR luminance of these frames in false color. The transition of illumination conditions occurs when the disco ball light source is turned off. When applying a TMO, this change of illumination condition is lost (bottom row).

TBI artifacts can appear even if no change of illumination condition occurs, that is to say when the tone mapping adapts to



Fig. 5. Example of a Temporal Brightness Incoherency. The change of illumination condition (represented by the mean value) in the HDR sequence is not preserved in the tone mapped result.



Fig. 6. Example of a Temporal Brightness Incoherency when a change of illumination occurs. False color luminance (top row) and tone mapped results using Photographic Tone Reproduction operator (Reinhard et al. 2002) (bottom row). Both frames appear at the same level of brightness although the false color representations indicate otherwise. The color bar indicates the value in cd/m^2 .

the content. When this adaptation occurs abruptly on successive frames, it gives rise to Flickering Artifacts (FA) as seen previously. However, when this adaptation is smoother, say over a longer range of time, the brightness relationship between the HDR and LDR sequences will be slowly disrupted. These arti-



Fig. 7. Example of Temporal Brightness Incoherency and Temporal Object Incoherency artifacts. False color luminance (top row) and tone mapped result using the Photographic Tone Reproduction operator (Reinhard et al. 2002) (bottom row). The Temporal Brightness Incoherency is represented by the overall brightness of each frame that is not coherent between the HDR and LDR frames. The Temporal Object Incoherency is represented by the brightness of the downside of the bridge similar in the HDR sequence while greatly varying in the LDR one. From left to right, frame 50, 100, 150, 200.

facts are similar to those occurring when commercial cameras adapt their exposure during a recording (Farbman & Lischinski 2011). Such an artifact is shown in Fig. 7 as the brightest HDR frame (rightmost) is the dimmest one in the LDR sequence. This second cause of TBI artifacts is also a common cause of Temporal Object Incoherency (TOI) presented hereafter.

1.5 Temporal Object Incoherency

Temporal Object Incoherency (TOI) occurs when an object's brightness, stable in the HDR sequence, varies in the LDR one. Fig. 8 plots the HDR and LDR overall brightness along with the value of a single pixel over several frames. Note that the HDR pixel's value is constant over time while the overall brightness changes. As the TMO adapts to each frame, the LDR pixel's value changes, resulting in a TOI artifact. Fig. 7 illustrates visually such an artifact. When looking at the false color representation of the HDR luminance (top row, Fig. 7) the level of brightness of the downside of the bridge appears stable over time. However, after applying a TMO (bottom row), the bridge, that appears relatively bright at the beginning of the sequence, is almost dark at its end. The temporal coherency of the bridge in the HDR se-



Fig. 8. Illustration of Temporal Brightness Incoherency. A pixel's value that is constant in the HDR sequence varies greatly in the LDR one. The pixel and mean value have been computed on the UnderBridgeHigh sequence shown in Fig. 7.

quence has not been preserved in the LDR one. The adaptation of a TMO to a scene is source of TBI and TOI artifacts. However, TBI are of global nature (difference of overall brightness between frames) while TOI are of local nature (difference of brightness between a reduced area over time).

1.6 Temporal Hue Incoherency

Temporal Hue Incoherency (THI) is closely related to TBI as it corresponds to the variation of the color perception of an object rather than its brightness. These artifacts occur when the balance between tristimulus values in successive frames is not temporally preserved by the tone mapping. The main reason of this imbalance is color clipping. Color clipping corresponds to the saturation of one or more of the tone mapped color channels (e.g. red, green or blue). Color clipping is a common artifact inherent in tone mapping of still images when one aims at reproducing to the best the HDR color (Xu et al. 2011, Pouli et al. 2013). When considering color clipping as a temporal artifact, it is not the difference between the HDR and LDR reproduction that is important but rather the LDR coherency from frame to frame.



(a) Tone mapped frame 100 of the ColorSquare sequence, the Tumblin & Rushmeier operator Tumblin & Rushmeier (1993) (left) along with the temporal evolution of the central pixel of the square (right).



(b) Zoom on the area outlined by the rectangle in frames 25, 50, 75 and 100 (from left to right).

Fig. 9. Example of Temporal Hue Incoherency due to color clipping. Each color channel desaturates at different temporal positions.

Indeed, variations in the tone mapping may saturate one color channel of an area which was not in the previous frame.

To illustrate such an artifact, we generated an HDR sequence with the following characteristics:

- a square area of constant luminance (100 cd/m²) with two gradients along the CIE u' and v' chrominances. The chrominance gradient ranges from -0.25 to 0.25 around the D65 white point.
- a neutral gray border area with a temporally varying luminance ranging from 0.005 to 10000 cd/m^2 .

Fig. 9 illustrates a THI due to the clipping of one or more color channels by a TMO. Note the shift in hue illustrated both in Fig. 9a-right and on a zoom on a portion of the tone mapped frames (Fig. 9b).

2 Video TMOs

Applying a TMO naively to each frame of a video sequence leads to temporal artifacts. The aim of Video TMOs is to prevent or reduce those artifacts. Video TMOs rely on information outside the current frame to perform their mapping. Most of current Video TMOs extend or post-process TMOs designed for still images. We chose to sort these techniques in three categories depending on the type of filtering:

- Section 2.1: Global Temporal Filtering,
- Section 2.2: Local Temporal Filtering,
- Section 2.3: Iterative Filtering.

For each category, we provide a description of the general technique along with different state of the art references.

2.1 Global Temporal Filtering

Global temporal filtering aims at reducing Global Flickering Artifacts (GFA) when using global TMOs. Indeed, global operators compute a monotonously increasing tone map curve that usually adapts to the image statistics of the frame to tone map. However, abrupt changes of this curve in successive frames results in GFA. Two main approaches have been formulated so far to reduce those artifacts: filtering temporally either the tone map curve or the image statistics.

By applying a temporal filter to successive tone map curves, GFA can be reduced. This tone map curves are usually filtered during a second pass as a first pass is required to compute a tone map curve per frame. The Display Adaptive (Mantiuk et al. 2008) operator is able to perform such a temporal filtering on the nodes of a computed piece-wise tone map curve. The efficiency of this filtering is illustrated in Fig. 10. The top row provides the independently tone mapped version of three successive frames of an



(a) the Display Adaptative TMO (Mantiuk et al. 2008) without the temporal filtering.



(b) Histogram and piece-wise tone map curves without the temporal filtering.



(c) Histogram and piece-wise tone map curves with the temporal filtering.



(d) the Display Adaptative TMO (Mantiuk et al. 2008) with the temporal filtering.

Fig. 10. Reduction of Global Flickering Artifacts by temporally filtering the tone mapping curves. pfstmo implementation of the Display Adaptative operator (Mantiuk et al. 2008) was used with options -d pd=lcd_office. From left to right, frames 153, 154 and 154 of the Temple sequence.

HDR video sequence. The second row displays the corresponding piece-wise tone map curves on top of their histogram. Note how the tone map curve of the middle frame is different from the other two, resulting in a change of overall brightness (GFA) in the tone mapped result. The third row shows the temporally filtered version of the piece-wise tone map curves. Finally, the bottom row provides the tone mapped frames after reducing the GFA. Image statistics can be unstable over time (e.g. the 99^{th} percentile, mean value, histogram of the luminance (Ward 1994), etc.). For example, the Photographic Tone Reproduction (Reinhard et al. 2002) operator relies on the geometric mean of an HDR image to scale it to the best exposure. One temporal extension of this operator filters this statistic along a set of previous frames (Kang et al. 2003). As a consequence, this method smooths abrupt variations of the frame geometric mean throughout the video sequence. This technique is capable of reducing flickering for sequences with slow illumination variations. However, for high variations it fails because it considers a fixed number of previous frames. That is why, (Ramsey et al. 2004) proposed a method that adapts dynamically this number. The adaptation process depends on the variation of the current frame key value and that of the previous frame. Moreover, the adaptation discards outliers using a min/max threshold. This solution performs better than (Kang et al. 2003) and for a wider range of video sequences. The computed geometric mean for these techniques and the original algorithm are plotted in Fig. 11. The green curve (Kang et al. 2003) smooths every peak but also propagates them to successive computed key values. The red curve however (Ramsey et al. 2004) reduces the abrupt changes of the key value without propagating it.

Another temporal extension of the Photographic Tone Reproduction operator has been proposed in (Kiser et al. 2012). The temporal filtering consists of a leaky integrator applied to three variables (a, A and B) that modify the scaling of the HDR frame:

$$\mathbf{L}_{\mathbf{s}} = \frac{\epsilon \cdot 2^{2(B-A)/(A+B)}}{k} \mathbf{L}_{\mathbf{w}} = \frac{a}{k} \mathbf{L}_{\mathbf{w}}$$
(1)

where $A = L_{max} - k$ and $B = k - L_{min}$ with L_{max} and L_{min} the maximum and minimum value of $\mathbf{L}_{\mathbf{w}}$ which is the HDR luminance. k corresponds to the geometric mean and the leaky integrator is computed as:

$$v_t = (1 - \alpha_v)v_{(t-1)} + \alpha_v v_t \tag{2}$$



Fig. 11. Evolution of the frame geometric mean computed for every frame of a video sequence. An offset is added to avoid an overlap between the curves, the smoothing effect of both the HDR Video operator (Kang et al. 2003) and the Adaptative Temporal operator (Ramsey et al. 2004) are compared to the Photographic Tone Reproduction operator (Reinhard et al. 2002).

where v_t represents any of the three variables a, A and B at time t and α_v is a time constant giving the strength (leakiness) of the temporal filtering.

Many other TMOs filter their parameters temporally, including (Pattanaik et al. 2000, Durand & Dorsey 2000, Irawan et al. 2005, Van Hateren 2006). Most of them either aim at simulating the temporal adaptation of the HVS or at reducing GFA.

2.2 Local Temporal Filtering

Local temporal filtering consists in performing a pixel-wise temporal filtering with or without motion compensation. Indeed, global temporal filtering cannot apply to local TMOs as such operators rely on a spatially varying mapping function. As outlined previously, local changes in a spatial neighborhood cause Local Flickering Artifacts (LFA). To prevent these local variations of the mapping along successive frames, Video TMOs can rely on pixel-wise temporal filtering. For example, the Gradient Domain Compression operator (Fattal et al. 2002) has been extended by (Lee & Kim 2007) to cope with videos. This TMO computes a LDR result by finding the output image whose gradient field is the closest to a modified gradient field. (Lee & Kim 2007) proposed to add a regularization term which includes a temporal coherency relying on a motion estimation:

$$\sum_{x,y} \|\Delta \mathbf{L}_{\mathbf{d}}(x,y,t) - \mathbf{G}(x,y)\|^{2} + \lambda \sum_{x,y} \|\mathbf{L}_{\mathbf{d}}(x,y,t) - \mathbf{L}_{\mathbf{d}}(x+\delta x,y+\delta y,t-1)\|^{2}$$
(3)

where $\mathbf{L}_{\mathbf{d}}$ is the output LDR luma at the precedent or current frame (t-1 or t) and \mathbf{G} is the modified gradient field. The pairs (x, y) and $(\delta x, \delta y)$ represent respectively the pixel location of a considered pixel and its associated motion vectors. The parameter λ balances the distortion to the modified gradient field and to the previous tone mapped frame.

Another operator (Local Model of Eye Adaptation (Ledda et al. 2004)) performs a pixel-wise temporal filtering. However the goal with this operator is to simulate the temporal adaptation of the human eye on a per-pixel basis. Besides, increasing the temporal coherency, pixel-wise temporal filtering also has denoising properties. Indeed, many denoising operators rely on temporal filtering to reduce noise (Brailean et al. 1995). Performing such a filtering during the tone mapping allows to keep the noise level relatively low.

2.3 Iterative Filtering

The techniques presented so far in this section focused on preventing temporal artifacts (mostly flickering) when tone mapping video sequences. These a priori approaches consist in either preprocessing parameters or modifying the TMO to include a temporal filtering step. Another trend analyzes a posteriori the output of a TMO to detect and reduce temporal artifacts, the reduction consisting of an iterative filtering.

One of these techniques (Guthier et al. 2011) aims at reducing



Fig. 12. Results of the Multi-Scale operator (Farbman et al. 2008) without (left) and with (right) the Flicker Reduction post-processing (Guthier et al. 2011). Each image represents two successive frames. Note the flickering artifact on the left image while it has been removed on the right image after applying the Flicker Reduction post-processing (Guthier et al. 2011).

Global Flickering Artifacts (GFA). Such an artifact is detected if the overall brightness difference between two successive frames of a video sequence is greater than a brightness threshold (defined using either Weber's law (Ferwerda 2001) or Steven's power law (Stevens & Stevens 1963)). As soon as an artifact is located, it is reduced using an iterative brightness adjustment until reaching the chosen brightness threshold. Note that this technique performs an iterative brightness adjustment on the unquantized luma to avoid loss of signal due to clipping and quantization. Consequently, the TMO's implementation needs to embed and apply the iterative filter before the quantization step. This technique relies only on the output of a TMO and hence can be applied to any TMO. Fig. 12 illustrates the reduction of a GFA when applying this post-processing.

3 Temporal Artifacts caused by Video TMOs

In the previous section, we have presented solutions to reduce temporal artifacts when performing video tone mapping. These techniques target mostly flickering artifacts as they are considered as one of the most disturbing ones. However, these techniques can generate two new types of temporal artifact: Temporal Contrast Adaptation (TCA) and Ghosting Artifacts (GA) that we detail in this section. To reduce global flickering artifacts, many TMOs rely on global temporal filtering. Depending on the used TMO, the filter is either applied to the computed tone map curve (Mantiuk et al. 2008) or to the parameter that adapts the mapping to the image (Ramsey et al. 2004, Kiser et al. 2012). However, when a change of illumination occurs, as shown in Fig. 6, it also undergoes temporal filtering. Consequently, the resulting mapping does not correspond to any of the conditions but rather to a transition state. We refer to this artifact as Temporal Contrast Adaptation (TCA). Fig. 13 illustrates the behavior of the temporal filtering when a change of illumination occurs. Note how the tone map curve, plotted on top of the histograms, shifts from the first illumination condition (frame 130) toward the second state of illumination (frame 150, see Fig. 6 for the false color luminance). As the tone map curve has anticipated this change of illumination, frames neighboring the transition of illumination are tone mapped incoherently.

These artifacts also occur when performing a post-processing to detect and reduce artifacts as in (Guthier et al. 2011). Indeed, this technique only relies on the LDR results to detect and reduce artifacts. If one has no information related to the HDR video then a change of illumination suppressed by a TMO cannot be anticipated nor predicted.

3.2 Ghosting Artifacts

Similarly to global temporal filtering, local temporal filtering generates undesired temporal artifacts. Indeed, pixel-wise temporal filtering relies on a motion field estimation which is not robust to change of illumination conditions and object occlusions. When the motion model fails, the temporal filtering is computed along invalid motion trajectories which results in Ghosting Artifacts



Fig. 13. Example of temporal filtering of the tone map curves when a change of illumination occurs. Top row: tone mapped result (frame number 130, 140, 149 and 150) using the Display Adaptative operator (Mantiuk et al. 2008) with the temporal filtering active (pfsTMO implementation (Krawczyk & Mantiuk 2007)). Bottom row: histograms of frames 130 (left) and 150 (right) along with the corresponding tone map curves for frames 130, 140 and 150.

(GA).

Fig. 14 illustrates a ghosting artifact in two successive frames resulting from the application of (Lee & Kim 2007) operator. This artifact proves that pixel-wise temporal filtering is only efficient for accurate motion vectors. A GA occurs when a motion vector associates pixels without temporal relationship. Those "incoherent" motion vectors should be accounted for to prevent GA as these latter are the most disturbing artifacts (Eilertsen et al. 2013).

4 Recent Video TMOs

Recently, two novel contributions have been proposed to the field of video tone mapping. The first one, called Zonal Brightness Coherency (ZBC) (Boitard et al. 2014), aims at reducing TBI and TOI artifacts through a post-processing operation which relies on a video analysis performed prior the tone mapping. The second



Fig. 14. Example of Ghosting Artifacts appearing on two successive frames. It is most noticeable around the two forefront columns (red squares). Bottom row is a zoom on the right-most column.

one (Aydin et al. 2014) performs a ghost-free pixel-wise spatiotemporal filtering to achieve high reproduction of contrast while preserving the temporal stability of the video.

4.1 Zonal Brightness Coherency

The Zonal Brightness Coherency (ZBC) algorithm (Boitard et al. 2014) aims at preserving the HDR relative brightness coherency between every object over the whole sequence. Effectively, it should reduce TBI, TOI and TCA artifacts and in some cases THI artifacts. It is an iterative work based upon the Brightness Coherency (Boitard et al. 2012) technique which considered only overall brightness coherency. This method consists of two steps: a video analysis and a post-processing.

The video analysis relies on a histogram-based segmentation as shown in Fig. 15. A first segmentation on a per frame basis provides several segments per frame. The geometric mean (called key value in the article) of each segment of each HDR frame is computed and used to build a second histogram which is in turn segmented to compute zone boundaries. The key value $k_z(\mathbf{L}_{\mathbf{w}})$ is computed for each zone of each frame. An anchor is then chosen either automatically or by the user to provide an intent to the



Fig. 15. Details on the video analysis. The *Frame Segmentation* function segments each frame of the sequence and computes each segment's key value. The *Video Segmentation* determines the video zone's boundaries and their corresponding key values. The *Anchor* function determines the anchor zone in the HDR sequence $k_{vz}(\mathbf{L}_{\mathbf{w}})$ and computes its corresponding LDR key values $k_{vz}(\mathbf{L}_{\mathbf{m}})$.

rendering.

Once the video analysis has been performed, each frame is tone mapped using any TMO. Then a scale ratio s_z is applied to each pixel luminance $\mathbf{L}_{\mathbf{m},\mathbf{z}}$ of each video zone z to ensure that the brightness ratio between the anchor and the current zone, in the HDR sequence, is preserved in the LDR sequence (Equation 4).

$$\mathbf{I}_{\mathbf{m},\mathbf{z}} = s_z \mathbf{L}_{\mathbf{m},\mathbf{z}}$$

$$s_z = \zeta + (1 - \zeta) \, \frac{k_{vz}(\mathbf{L}_{\mathbf{m}})}{k_{vz}(\mathbf{L}_{\mathbf{w}})} \frac{k_z(\mathbf{L}_{\mathbf{w}})}{k_z(\mathbf{L}_{\mathbf{m}})}$$

$$(4)$$

where $\mathbf{I}_{\mathbf{m},\mathbf{z}}$ is the scaled luminance, ζ is a user-defined parameter, $k_{vz}(\mathbf{L}_{\mathbf{w}})$ is the anchor zone HDR key value, $k_{vz}(\mathbf{L}_{\mathbf{m}})$ is the anchor zone LDR key value, $k_z(\mathbf{L}_{\mathbf{w}})$ is the z zone HDR key value and $k_z(\mathbf{L}_{\mathbf{m}})$ is the z zone LDR key value. Note that the subscript z stands for zone.

At the boundaries between two zones, an alpha blending is used to prevent abrupt spatial variations. The whole workflow of this technique is depicted in Fig. 16.



Fig. 16. Complete Zonal Brightness Coherency workflow with details on the scaling phase. The Zones function determines, for each tone mapped frame and each pixel, the corresponding video zone z^j as well as the video blending zone $b_z^{j,j+1}$. Their respective scaling ratios s^j and $\mathbf{s^{j,j+1}}$ are computed. The Zone Scaling function applies the scale ratios to the tone mapped frames. Finally, Q quantize linearly floating point value in the range [0;1] to integer values on a defined bit-depth n (values ranging from $[0;2^n-1]$)

Fig. 17 presents some results when using the ZBC post-processing on tone mapped video sequences where temporal artifacts occurred. The left plot provides results regarding the reduction of the Temporal Object Incoherency (TOI) artifact that was illustrated in Fig. 8. Thanks to the ZBC technique, the value of the pixel, which was constant in the HDR sequence, is much more stable over time. Note also that the LDR mean value is quite low at the beginning of the sequence, which will most likely result in a loss of spatial contrast in the tone mapped frames. That is why the authors have provided a user-defined parameter which effectively trades off temporal coherency for increase of spatial reproduction capabilities (see ζ in (Equation 4)). On the right plot, we show some results regarding the reduction of Temporal Brightness Incoherency (TBI) artifacts. This plot is to be compared with the one in Fig. 5. Using the ZBC post-processing on the Disco sequence allows to preserve the change of illumination present in the HDR sequence.

In (Boitard et al. 2014) and (Boitard 2014), more results are



Fig. 17. Reduction of Temporal Object Incoherency and Temporal Brightness Incoherency artifacts using the Zonal Brightness Coherency post-processing. Left plot is to be compared with Fig. 8 while the right one with Fig. 5.

available especially regarding the preservation of fade effects and the impact of the different parameters of this technique.

4.2 Temporally Coherent Local Tone Mapping

In Section 3.2, we explained why pixel-wise temporal filtering can cause Ghosting Artifacts (GA). However, this type of filtering is the only known solution to prevent local flickering artifacts that can arise when using local TMOs. Consequently, (Aydin et al. 2014) proposed a spatial filtering process to ensure high reproduction of contrast while performing a ghost-free pixel-wise temporal filtering. Figure 18 illustrates the workflow of this technique.

This technique considers a temporal neighborhood composed of a center frame $\mathbf{I_k}$ and temporally close frames $\mathbf{I_{k\pm i}}$. In a first step, each frame is decomposed into a base and detail layer using a permeability map (spatial diffusion weights). Both subbands are then motion compensated (warped) using a previously computed motion flow. Note that no warping is necessary for the subbands associated with the central frame.

The second step consists of two temporal filters which allow a separate filtering of the base and the detail layers. To prevent GA, the filtering relies on confidence weights composed of a photo-



Fig. 18. General worflow of the Temporally Coherent Local operator (Aydin et al. 2014). The spatial decomposition is illustrated on upper right hand corner while the base and detail filtering are illustrated on the bottom left and right corner respectfully.

constancy permeability map and a penalization on pixels associated with high gradients flow vectors. The photo-constancy permeability map is computed between successive frames and corresponds to a temporal transposition of the spatial diffusion weight on which relies the spatial decomposition. The authors observed that this photo-constancy measure can be tuned to stop temporal filtering at most warping errors, hence preventing the appearance of GA. However, it also defeats the purpose of temporal filtering, which is to smooth medium to low temporal variations. That is why, the penalization term has been introduced as it is a good indication of complex motion where the flow estimation tends to be erroneous. This step provides two images, the spatio-temporally



Fig. 19. Illustrative worflow of the Temporally Coherent Local operator (Aydin et al. 2014). PGF stands for Permeability Guided Filtering. Figure courtesy of Tun Ozan Aydin.

filtered base layer \mathbf{B}_k and a temporally filtered detail layer \mathbf{D}_k .

To obtain the tone mapped frame \mathbf{I}_{k}^{TMO} , the base layer \mathbf{B}_{k} can be fed to any TMO and then combined with the detail layer \mathbf{D}_{k} , a process similar to that detailed in (Durand & Dorsey 2002). This method addresses several of the artifacts presented in this chapter. First, the temporal noise is reduced as the details are filtered temporally. Second, local flickering artifacts are minimized thanks to the pixel-wise temporal filtering. Finally, ghosting artifacts are prevented by adapting the temporal filtering to a motion flow confidence metric. A more illustrative worfkow is depicted in Figure 19.

As this technique is fairly recent, extended results with more HDR sequences could help detect new types of artifacts. In particular, it would be interesting to test this method when changes of illumination or cut occur in a sequence. Furthermore, most of the results provided with this technique rely on a user interaction to achieve the best trade-off between temporal and spatial contrast. This is not achievable for many applications such as live broadcast, tone mapping embedded in set top boxes, etc.

Temporal	Possible	Possible
Artifact	Cause	Solutions
GFA	Temporal unstability of	Global temporal
(Global Flicker)	parameters	filtering
LFA	Different spatial filtering	Pixel-wise
(Local Flicker)	in successive frames	temporal filtering
TN	Camera noise	Spatial and/or temporal
(Noise)		filtering (pixel-wise)
TBI	Change of illumination	Brightness analysis
(Brightness)	adaptation of the TMO	of each frame
TOI	Adaptation of the TMO	Brightness analysis
(Object)		per zone of frames
THI	Saturation of color channel	Hue and brightness analysis
(Hue)	(due to clipping)	per zone of frames?
TCA	Due to global	Brightness analysis
(Contrast)	temporal filtering	per zone of frames
GA	Due to pixel-wise	Confidence weighting of
(Ghosting)	temporal filtering	pixel-wise temporal filtering

Table 1

Summary of temporal artifacts along with their main causes and possible solutions.

5 Summary

In this chapter, we have described known types of temporal artifact that occur when tone mapping HDR video sequences. We have categorized Video TMOs with respect to how they handle temporal information, and we have shown that although these solutions can deal with certain types of temporal artifact, they can also be source of new ones. An evaluation of video TMOs (Eilertsen et al. 2013) reported that none of the current solutions could handle a wide range of sequences. However, this study was performed prior the publication of the two Video TMOs presented in Section 4. These two techniques, albeit significantly different, provide solutions to types of artifact not dealt with before. Table 1 gives an overview of the temporal artifacts presented in this chapter along with possible solutions. From this table, we can notice that all of the presented artifacts in this chapter have a solution. However, none of the Video TMOs presented here encompasses all the tools needed to deal with all different types of artifact. Furthermore, the apparition of more HDR video sequences and applications for video tone mapping will more likely result in new types of temporal artifact. Although the two recent contributions significantly advanced the field of video tone mapping, more work still lay ahead.

References

- Aydin, T. O., Stefanoski, N., Croci, S., Gross, M. & Smolic, A. (2014), 'Temporally coherent local tone mapping of HDR video', ACM Transactions on Graphics 33(6), 196:1–196:13.
- Banterle, F., Artusi, A., Debattista, K. & Chalmers, A. (2011), Advanced High Dynamic Range Imaging: Theory and Practice, AK Peters (CRC Press), Natick, MA, USA.
- Boitard, R. (2014), Temporal Coherency in Video Tone Mapping, Ph.d. thesis, University of Rennes 1.
- Boitard, R., Bouatouch, K., Cozot, R., Thoreau, D. & Gruson, A. (2012), Temporal coherency for video tone mapping, *in* 'Proc. SPIE 8499, Applications of Digital Image Processing XXXV', pp. 84990D–1–84990D–10.
- Boitard, R., Cozot, R., Thoreau, D. & Bouatouch, K. (2014),
 'Zonal brightness coherency for video tone mapping', Signal Processing: Image Communication 29(2), 229–246.
- Brailean, J., Kleihorst, R., Efstratiadis, S., Katsaggelos, A. & Lagendijk, R. (1995), 'Noise reduction filters for dynamic image sequences: a review', *Proceedings of the IEEE* 83(9), 1272–1292.
- Cadík, M., Wimmer, M., Neumann, L. & Artusi, A. (2008), 'Evaluation of HDR tone mapping methods using essential perceptual attributes', *Computers & Graphics* **32**(3), 330–349.
- Digital Multimedia Laboratory (2014), 'DML-HDR Public Database', http://dml.ece.ubc.ca/data/DML-HDR/.

- Drago, F., Martens, W. L., Myszkowski, K. & Seidel, H.-P. (2003), Perceptual evaluation of tone mapping operators, *in* 'Proceedings of the SIGGRAPH 2003 Sketches', ACM Press, New York, New York, USA, pp. 1–1.
- Durand, F. & Dorsey, J. (2000), Interactive tone mapping, *in* 'Proceedings of the Eurographics Workshop on Rendering', Springer Verlag, pp. 219–230.
- Durand, F. & Dorsey, J. (2002), 'Fast bilateral filtering for the display of high-dynamic-range images', ACM Transactions on Graphics 21(3), 257–266.
- Eilertsen, G., Wanat, R., Mantiuk, R. K. & Unger, J. (2013), 'Evaluation of Tone Mapping Operators for HDR-Video', Computer Graphics Forum 32(7), 275–284.
- Farbman, Z., Fattal, R., Lischinski, D. & Szeliski, R. (2008), Edge-preserving decompositions for multi-scale tone and detail manipulation, in 'ACM SIGGRAPH 2008 papers on - SIG-GRAPH '08', ACM Press, New York, New York, USA, pp. 67:1– 67:10.
- Farbman, Z. & Lischinski, D. (2011), 'Tonal stabilization of video', ACM Transactions on Graphics 30(4), 89:1–89:10.
- Fattal, R., Lischinski, D. & Werman, M. (2002), 'Gradient domain high dynamic range compression', ACM Transactions on Graphics 21(3), 249–256.
- Ferwerda, J. (2001), 'Elements of early vision for computer graphics', *IEEE Computer Graphics and Applications* **21**(4), 22–33.
- Gastal, E. S. L. & Oliveira, M. M. (2011), 'Domain transform for edge-aware image and video processing', ACM Transactions on Graphics 30(4), 69:1–69:12.
- Guthier, B., Kopf, S., Eble, M. & Effelsberg, W. (2011), 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', Vol. 7866, pp. 78660C-1-78660C-15.
- Irawan, P., Ferwerda, J. A. & Marschner, S. R. (2005), Perceptually based tone mapping of high dynamic range image streams, *in* 'Proceedings of the Sixteenth Eurographics Conference on

Rendering Techniques', EGSR '05, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp. 231–242.

- IRISA (2015), 'IRISA Rennes FRVSense Public HDR Database', http://people.irisa.fr/Ronan.Boitard/.
- ITU (1998), 'Recommendation ITU-R BT.709-3: Parameter values for the HDTV standards for production and international programme exchange', International Telecommunications Union.
- ITU (2012), 'Recommendation ITU-R BT.2020: Parameter values for ultra-high definition television systems for production and international programme exchange', International Telecommunications Union.
- Kang, S. B., Uyttendaele, M., Winder, S. & Szeliski, R. (2003), 'High dynamic range video', ACM Transactions on Graphics 22(3), 319–325.
- Kiser, C., Reinhard, E., Tocci, M. & Tocci, N. (2012), Real-time Automated Tone Mapping System for HDR Video, *in* 'Proceedings of the IEEE International Conference on Image Processing', pp. 2749–2752.
- Krawczyk, G. (2006), 'MPI Public HDR Database', http://www. mpi-inf.mpg.de/resources/hdr/video/.
- Krawczyk, G. & Mantiuk, R. (2007), 'Display adaptive pfstmo documentation', http://pfstools.sourceforge.net/ man1/pfstmo_mantiuk08.1.html.
- Kronander, J., Gustavson, S., Bonnet, G. & Unger, J. (2013), 'Unified HDR reconstruction from raw CFA data', *IEEE In*ternational Conference on Computational Photography (ICCP) pp. 1–9.
- Kronander, J., Gustavson, S., Bonnet, G., Ynnerman, A. & Unger, J. (2014), 'A unified framework for multi-sensor HDR video reconstruction', *Signal Processing: Image Communica*tion 29(2), 203–215.
- Kuang, J., Yamaguchi, H., Liu, C., Johnson, G. M. & Fairchild, M. D. (2007), 'Evaluating HDR rendering algorithms', ACM Transactions on Applied Perception 4(2), 9:1–9:27.
- Lasserre, S., Le Léannec, F. & Francois, E. (2013), Descrip-

tion of HDR sequences proposed by Technicolor, *in* 'ISO/IEC JTC1/SC29/WG11 JCTVC-P0228', IEEE, San Jose, USA.

- Ledda, P., Santos, L. P. & Chalmers, A. (2004), A local model of eye adaptation for high dynamic range images, *in* 'Proceedings of the 3rd international conference on Computer graphics, virtual reality, visualisation and interaction in Africa - AFRI-GRAPH '04', ACM Press, New York, New York, USA, pp. 151– 160.
- Lee, C. & Kim, C.-S. (2007), Gradient domain tone mapping of high dynamic range videos, in '2007 IEEE International Conference on Image Processing', Vol. 3, IEEE, pp. III – 461 – III – 464.
- Mantiuk, R., Daly, S. & Kerofsky, L. (2008), 'Display adaptive tone mapping', ACM Transactions on Graphics 27(3), 68:1– 68:10.
- Myszkowski, K., Mantiuk, R. & Krawczyk, G. (2008), *High Dy*namic Range Video, Vol. 2, Morgan & Claypool.
- Pattanaik, S. N., Tumblin, J., Yee, H. & Greenberg, D. P. (2000), Time-dependent visual adaptation for fast realistic image display, *in* 'Proceedings of the 27th annual conference on Computer graphics and interactive techniques', SIGGRAPH '00, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, pp. 47–54.
- Pouli, T., Artusi, A. & Banterle, F. (2013), Color Correction for Tone Reproduction, in 'Color and Imaging Conference Final Program and Proceedings', pp. 215–220.
- Ramsey, S. D., Johnson III, J. T. & Hansen, C. (2004), Adaptive temporal tone mapping, in 'Computer Graphics and Imaging -2004', pp. 3–7.
- Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G. & Myszkowski, K. (2010), High Dynamic Range Imaging, 2nd Edition: Acquisition, Display, and Image-Based Lighting, Morgan Kaufmann.
- Reinhard, E., Stark, M., Shirley, P. & Ferwerda, J. (2002), 'Photographic tone reproduction for digital images', ACM Transactions on Graphics 21(3), 267–276.

- Stevens, J. C. & Stevens, S. S. (1963), 'Brightness Function : Effects of Adaptation', Journal of the Optical Society of America 53(3), 375–385.
- Tocci, M. D., Kiser, C., Tocci, N. & Sen, P. (2011), 'A versatile HDR video production system', ACM Transactions on Graphics 30(4), 41:1–41:10.
- Tumblin, J. & Rushmeier, H. (1993), 'Tone reproduction for realistic images', Computer Graphics and Applications IEEE 13(6), 42–48.
- Unger, J. (2013), 'HDRv Public Database', http://www.hdrv. org/Resources.php.
- Van Hateren, J. H. (2006), 'Encoding of high dynamic range video with a model of human cones', ACM Transactions on Graphics 25(4), 1380–1399.
- Ward, G. (1994), A contrast-based scale factor for luminance display, in P. S. Heckbert, ed., 'Graphics Gems IV', Academic Press Professional, Inc., San Diego, CA, USA, pp. 415–421.
- Xu, D., Doutre, C. & Nasiopoulos, P. (2011), 'Correction of clipped pixels in color images.', *IEEE Transactions on Visu*alization and Computer Graphics 17(3), 333–344.