Video logarithmic tone mapping operator with automatic calculation of the logarithmic function

Cédric Walbrecq, Isabelle Marc, Dominique Lafon-Pham, Baptiste Magnier

Video logarithmic tone mapping operator with automatic calculation of the logarithmic function base

Cédric Walbrecq
EUROMOV DHM,
Univ Montpellier, IMT Ales
Alès, France
cedric.walbrecq@mines-ales.fr

Isabelle Marc
EUROMOV DHM,
Univ Montpellier, IMT Ales
Alès, France
isabelle.marc@mines-ales.fr

Baptiste Magnier
EUROMOV DHM
Univ Montpellier, IMT Ales
Alès, France
baptiste.magnier@mines-ales.fr

Dominique Lafon-Pham
EUROMOV DHM
Univ Montpellier, IMT Ales
Alès, France
dominique.lafon@mines-ales.fr

***Abstract*—High Dynamic Range (HDR) videos are intended to capture the full range of luminance that can be found in a real world environment. But, as display devices with restricted luminance range are still the most common ones, the need for efficient tone mapping tools remains. Applying a variable base logarithmic compression curve on HDR values is a method that gives good results on static images. Video mapping requires further processing, in order to maintain correct contrast over the different frames and to minimize undesirable artifacts such as flicker and temporal incoherencies when the overall lighting conditions vary. This paper describes a new method for a real time automatic calculation of the logarithmic function base, depending on the characteristics of successive frames in a sequence. Compared to some well-known operators, our proposal better performs in terms of temporal coherency and overall contrast.**

Keywords— HDR, Video Tone Mapping, Temporal artifact reduction

I. INTRODUCTION

Since the late 1990s, there is a growing interest in developing High Dynamic Range (HDR) imaging methods in order to capture, encode and render images or videos aiming to mimic real world scenes. However, most of the currently commercially available display devices offer a limited dynamic range of two orders of magnitude, well below the range of luminance between the lightest and darkest areas in natural viewing, so works still have to be done to develop and enhance tone mapping operators (TMO), i.e. methods to reduce luminance range while preserving details in dark and light areas as well as overall brightness perception. It is particularly true concerning TMO for videos, as the need for them has appeared later, with the more recent development of HDR video contents. HDR video processing requires a new set of tools, because the simple application of any image operator on each of the successive frames might led to perceptually unacceptable results, in case of failure in temporal correlation between two frames. We propose a method to extend an image operator for video processing, which adapts the tone mapping function to the temporal variations in lighting conditions and scene composition, and prevents the appearance of frames with incoherent luminance content. A comparison is made with others TMOs on a set of sequences with different luminance ranges and variations, and our proposal gives promising results.

II. RELATED WORKS

Tone mapping operators are essentially nonlinear functions: because of the nonlinear behavior of the Human Visual System (HVS), a simple linear scaling of HDR values into a reduced range does not allow for the matching between perception of the tone mapped image and perception of the real-world scene. Hundreds of image TMOs have been proposed in the last decades to meet different objective: fidelity to the HDR scene, with detail preservation in both dim and bright areas, simulation of human visual processing, for instance taking into account light and dark adaptation processes, or artistic rendering [1]. In terms of processing, they fall in two wide categories: global operators apply the same mapping to all pixels in the image, while local ones adapt the mapping function to each pixel depending on its neighborhood. Global and local operators both have advantages and limitations, and at this time, there is no clear evidence of one category being superior to the other. Rather, it appears that their relative performances depend on the way they are evaluated. Numerous psychophysical comparison studies (for a review, see for instance [2], have demonstrated that images processed by global methods are generally better ranked in terms of overall brightness, global contrast, and naturalness. On the contrary, local methods overperform in terms of detail preservation and local contrast.

TMOs designed for static images might obviously be directly applied on each frame of video sequences, but doing that might lead to unacceptable results, with temporal artifacts generation: image TMOs are designed to scale HDR luminance values on the whole range of a LDR display, and are often based on image statistics (i.e. mean, max, histogram, etc). Whenever the range of luminance or image statistics vary between two successive frames, because of a sudden difference in lighting conditions, or motion of the camera during walkthrough, it may break the temporal coherency between these frames (for instance: the second LDR frame appearing brighter though the second HDR one is darker). A detailed definition of these artifacts may be found in [3] and will be detailed in the following. Several subjective quality assessment works have been performed on tone mapped videos. Melo *et al* focused their study on small mobile displays (tablet) and asked sixty participants to compare the subjective performances of six operators for images and videos tone mapping, in reference to an HDR display [4]. In [5], five expert observers rated videos in term of perceptual attributes (overall brightness, overall contrast, overall color saturation, temporal color consistency, temporal flickering, ghosting and noise), for eleven TMOs. , a set of quantitative measurements was used to rank these TMOs: noise visibility, percentage of under and over exposed pixels, loss of contrast, and temporal coherence, which was derived from cross correlation between HDR and LDR values over a few successive frames. It can be drawn from these studies that subjective quality relies on different characteristics for images and videos: for the later, the preservation of contrast in fine details becomes much less important than the general impression of temporal coherence, and global operators are better ranked in general. Brightness flickering and ghosting effect are found to be critical factors.

For the TMOS that are intended to simulate the SVH processing, temporal effects are explicitly taking into consideration by modelling dark and light adaptation mechanisms: Fewerda *et al* scaled HDR values so as just noticeable contrast differences (JND's) in the real scene correspond to contrast JNDs in the mapped image, for the current adaptation state on each frame [6]. Pattanaik *et al* proposed a time dependent TMO, with adaptation process and image appearance modelling [7]. On the other part, TMOs with no explicit reference to adaptation process may also be used for video processing when they are completed by temporal filtering steps. Filtering may be applied to the mapping function parameters. For example, Reinhard *et al* were inspired by dodging & burning photographic practices, and scaled HDR values to LDR ones according to the geometric averaged luminance (named « key value ») [8]. Temporal coherency is increased by filtering this key value over a number of frames which is fixed [9], or variable [10]. Alternately, filtering may be applied to the mapping function itself: in [11] Ploumis *et al* proposed a method to differentiate between flickering effect and actual changes in real scene illumination. Whenever a flicker is detected, the Ward's mapping function is modified. When HDR luminance geometric mean variation appears to be due to a real change in scene characteristics, no filter is perform, and the frame is mapped according to the current conditions. Another approach to reach temporal coherency is post processing of tone mapped videos, as in [12]or [3]. This particular domain is out of the scope of our work and will not be discussed any longer.

III. Proposed algorithm

Logarithmic compression of luminance values is deemed to give a reasonably good fit of the human response to light. Our work is based on Drago's TMO [13], a global operator which offers a good compromise between performances and computational costs. It was one of the best graded techniques in the image subjective quality assessment conducted by [14]. It also has the advantage to be easily adapted to suit the characteristics of the display. Many works already exist that are based on it, and logarithmic mapping is a research axe which is still being investigated. Aydin *et al* proposed to decompose HDR frames into base and details layers, et chose Drago's algorithm to scale HDR values of the base layer [15]. Popovic *et al* also used this algorithm for a real time implementation on a field programmable gate array [16]. In [17] the maximum luminance value of each frame was temporally filtered before being used in the tone mapping function. In a recent work, Lisani presented a local contrast enhancement based on logarithmic tone mapping curve [18]. Pixels belonging to dark regions of the images are modified using a concave tone-mapping function and therefore their values are increased. Conversely, a convex tone-mapping is applied to pixels in bright regions, which produces a decrease of their values, and this results in an increase in visibility in both dark and bright regions.

A. Drago's TMO

This TMO relies on a logarithmic type tone curve applied on the scaled luminance Y channel in XYZ color space:

$$Lw(i,j) = \frac{L(i,j)}{Lw_{avg}} \quad (1)$$

$$Lw_{max} = \frac{L_{max}}{Lw_{avg}} \quad (2)$$

where L(i,j) is the HDR luminance value of pixel (i,j), which is equal to Y value in XYZ space, and L_{max} is the maximum value in the HDR image. Lw_{avg} is the image luminance geometric average, calculated from

$$Lw_{avg} = e^{\frac{1}{M \cdot N}\sum_{i=0,j=0}^{M,N} log(L+\varepsilon)} \quad (3)$$

where ε is a small constant used to avoid computational issue when L = 0. Lw_{avg} is chosen as the scalefactor, as it commonly stands for visual adaptation state. In order to enhance the contrast, the base for the logarithmic function varies from 2 for the darkest pixels to 10 for the brightest ones, and smooth interpolation is performed between these values, based on a bias power function. The mapped values are further scaled to match the maximum luminance of the display $Ldmax$ before a conversion back to RGB space is performed:

$$Ld = \frac{Ldmax}{log10(Lw_{max})} \cdot \frac{log(Lw)}{\left(log\left(2+\left(\left(\frac{Lw}{Lw_{max}}\right)^{bias}\right)\cdot 8\right)\right)} \quad (4)$$

with

$$bias = \frac{log(b)}{log(0.5)} \quad (5)$$

The b parameter is responsible for the trade-off between compression in dark and bright areas. A default value of 0.85 is suggested in [13].

B. Enhanced Drago TMO for Temporal Coherency (EDTC)

We propose to modify Drago's TMO with the objectives to avoid temporal artifact generation while ensuring a good overall contrast in all frames. Temporal coherency is obtained in a classical way, by filtering one of the different statistic parameters that describe each HDR frame. It was found that a leaky integrator applied to the maximum HDR value gives the best results:

$$L_{max}(t) = L_{max}(t-1) + \left(L_{max}(t) - L_{max}(t-1)\right)e^{-\tau} \quad (6)$$

The constant τ is the ratio of the desired number of transition frames over the frame rate.

As already mentioned, contrast in the mapped images relies on the b value, it is then important that this value might be automatically adjusted to the overall luminance of the current frame: contrast in darker (resp. brighter) images is better for b near 0 (resp. 1). In consequence, b was set equal to the Ostu threshold for the histogram of Lw values. Ostu threshold is the value that separate histogram bins into two classes, with minimization of intra-class variance [19]. In our case, the histogram is computed between the minimum and geometric mean values only. All values above this mean are added to its bin. The b value variations are further smoothed, according to the equation (6).

IV. Experimental assessment

We evaluate the temporal performances of our algorithm in comparison with four state-of-the-art TMOs, in terms of temporal artifact generation.

A. Method

1) Definition of temporal artifacts

According to [3], different categories of temporal artifacts may be defined. Temporal noise due to sensor defects is mainly noticeable in dark areas, and more disturbing than in static images, due to its high frequency characteristics. Because most of TMO increase the noise level rather than reducing it, a preliminary step of denoising is often necessary, in particular when videos are captured by HDR logarithmic sensor. Flicker occurs when the mapping function employs parameters deduced from image statistics which may undergo abrupt modifications from frame to frame, leading to transient changes in overall LDR frame brightness or incoherent pixel values. Temporal Brightness Incoherency (TBI) artifacts is observed when the overall relative brightness between two frames is different between HDR and LDR sequences. Temporal Object Incoherency (TOI) corresponds to the same phenomenon restricted to a limited zone or object in the scene. Color perception may also be perturbed by the mapping function, if one color channel gets clipped. Ghosting is another temporal artifact, which may arise on pixels located close to edges, especially for local TMOs, when temporal filtering is performed on a pixelwise basis. In our work, we mainly focus on flicker, TBI and TOI: as we only deal with global TMOs, ghosting is not likely to appear.

2) Others TMOs

We choose to compare our ***EDTC*** algorithm with four state-of-the-art global TMOs that are frequently assessed in comparison and review works. All these operators run in real time. The first one is the Photographic Tone Reproduction operator (***Reinhard02***) described in [8]. The second one, ***Kiser12***, is another version of this TMO, with smooth filtering to reduce variation of the TMO parameters between adjacent frames [20]. Finally, ***Ferwerda96*** [6] and ***Pattanaik00*** [7] are considered, as they perform rather well in [1] and [4].

TMO implementations were obtained from https://github.com/banterle/HDR_Toolbox. They were run with the same set of parameters for all videos, with the values obtained from a preliminary parameter adjustment experiment in [5].

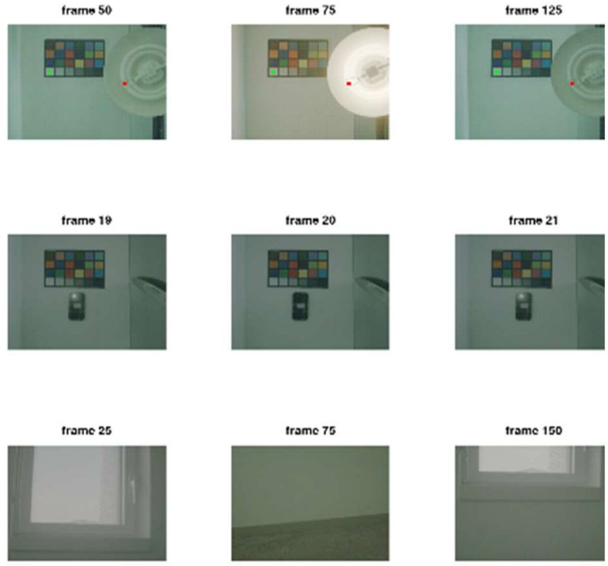

Fig. 1. Three frames of the three video sequences captured for this work. Top row: ***lamp***, with two measurement pixels indicated by green and red points; Middle row: ***blinksource***; Bottom row: ***window***.

3) Video content

Depending on the scene content, performances may be different. We thus decided to use video sequences with a wide range of luminance variations. Some of these videos exhibit more or less rapid change in intensity, due to the camera moving in natural scenes (***hallway***, ***exhibition*** from [5], ***underbridgehigh*** from [3] and ***window***).

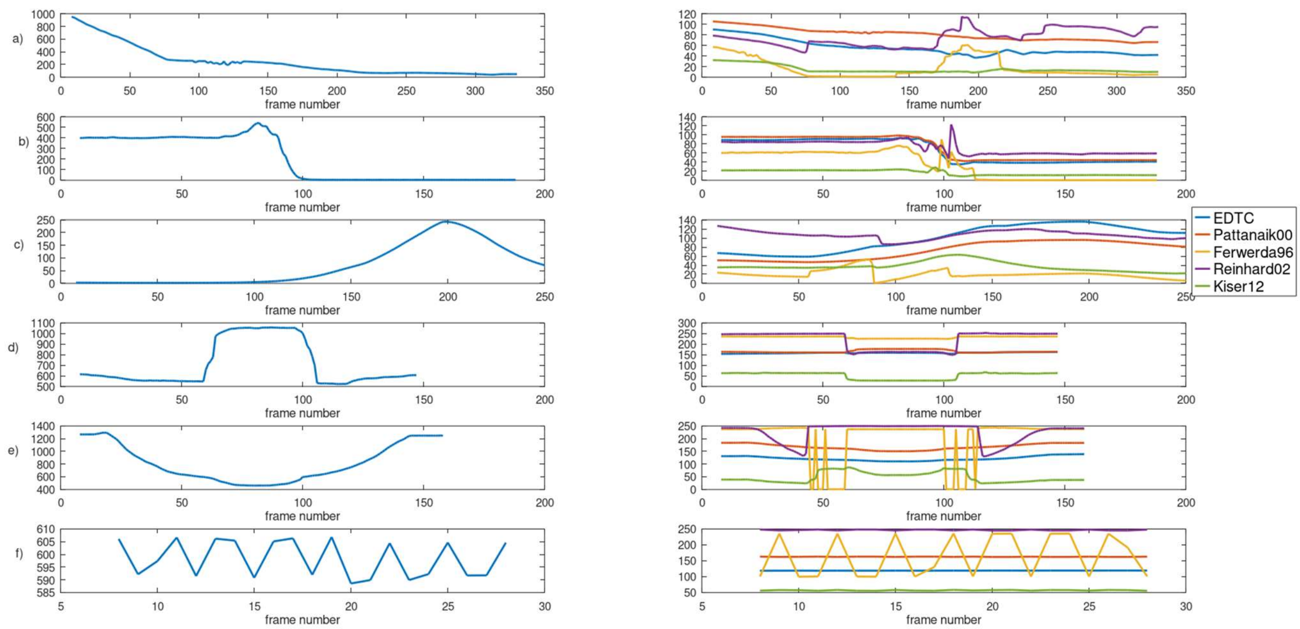

Fig. 2. Temporal variations of the mean HDR luminance (left) and the mean LDR values (right) for the six sequences. a: ***hallway***; b:***exhibition***; c:***underbridgehigh***; d:***lamp***; e:***window***; f:***blinkingsource***.

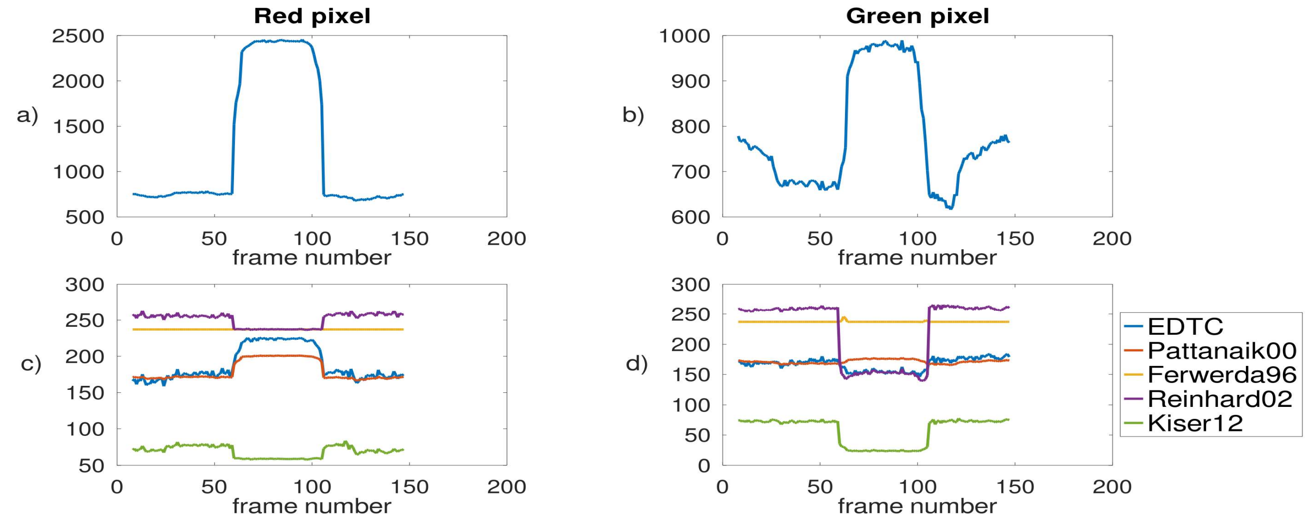

Fig. 3. HDR luminance (top row) and LDR values (bottom row) for two measurement points. Red pixel: in the lamp area. Green pixel: in the non-illuminated area (see Fig. 1 for details)

Two other sequences contain abrupt changes of illumination in static scenes, due to variations in lighting conditions: ***blinking source*** (static scene with small blinking LED source) and ***lamp*** (static scene with halogen lamp turned off, on, and off) figure. The last three videos were captured by a logarithmic sensor camera (NIT MC1003-1VC; 1280×1024 pixels). Two frames from each of these videos are showed on Fig. 1.

B. Results

The panels of Fig. 2 show the HDR mean luminance and LDR mean value over the different sequences.

Flickering effect is demonstrated whenever the LDR mean value undergoes a sharp variation, while HDR mean value remaining merely constant. ***EDTC*** create videos totally free of such artifact, unlike ***Reinhard02***, and ***Kiser12*** (see in particular ***Underbridgehigh***, ***Window***, and ***Blinksource,*** in Fig. 2.)

Temporal coherency may be disturbed, depending on the video under test. For relatively smooth luminance variations due to camera motion, as in ***hallway*** or ***window***, only ***EDTC*** and ***Pattanaik00*** perform well, the others showing strong temporal incoherencies, with LDR values undergoing variations decorrelated from the HDR ones. This behavior is even more pronounced for sharp variations, as in ***blinksource*** and ***lamp***: in these cases, all but ***EDTC*** and ***Pattanaik00*** lead to LDR mean variations opposite to HDR ones.

Temporal object coherency is also better preserved by these two algorithms. The temporal evolution of two pixels in ***lamp*** sequence (their location are indicated on Fig.1) are shown on Fig.3. For the pixel which becomes bright when the lamp is turned on (red point), brightness variations are coherent. For the point in the non-illuminated area (green point), the result is not so good. When the lamp is turned on, pixel value decreases. However, this incoherency is much less pronounced than for the others TMO.

Another characteristic of TMOs to be taken in account in their comparison is the way they manage luminance contrasts. The overall contrast relies on luma range in each frame, which is shown in Fig.4. For the ***lamp*** sequence, when

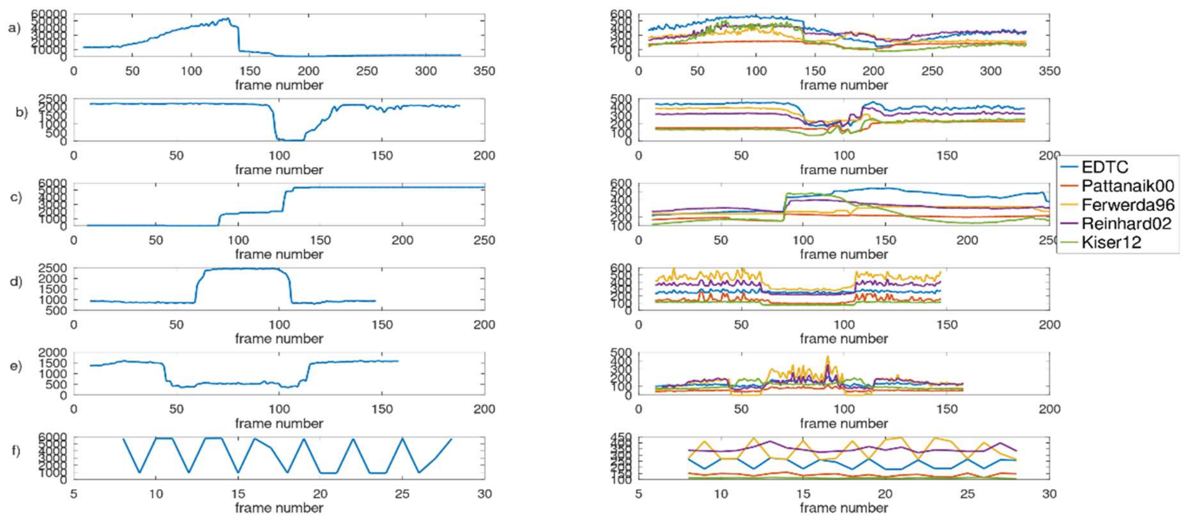

Fig 4 HDR luminance range*: Lwmax-Lwmin* (left) and LDR value range: *Ldmax-Ldmin* (right) for the six sequences. a: ***hallway***; b:***exhibition***; c:***underbridgehigh***; d:***lamp***; e:***window***; f:***blinkingsource***.

the lamp is turned on, the difference between the maximum and minimum values strongly decreases for all but ***EDTC*** TMOs, which reflects the fact that video appears less contrasted, with a kind of glare effect. In ***EDTC***, dynamic range does not change, the feeling of contrast is maintained. This behavior is further highlighted in Fig.5: when the lamp is turned on, some TMOs fail to render contrast in the bright area.

As expected, the ***Reinhard02*** implementation we used, which was designed for image processing, shows the worst results. Its temporal adaptation in ***Kiser12***, more surprisingly, performs only slightly better. ***Ferwerda96*** is also prone to flicker and incoherencies. Only ***Pattanaik00*** performs as well as ***EDTC*** concerning these artefacts. However, it produces quite saturated videos, as shown on Fig.5.

CONCLUSION

A novel method for video tone mapping has been presented, which allows for real time processing. It has been tested on HDR sequences containing various luminance conditions, from smooth variations to sudden changes. In all cases, it produces artifact free videos, and maintains a good overall contrast.

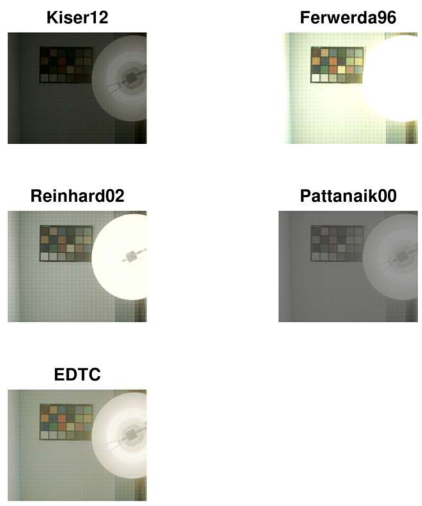

Fig. 5 Comparison of the 50th frame from ***lamp*** sequence, tone mapped by the 5 TMOs

REFERENCES

[1] G. Eilertsen, J. Unger, and R. K. Mantiuk,, «A comparative review of tone-mapping algorithms for high dynamic range video,» *Comput. Graph. Forum,* vol. 36, n° 12, pp. 565–592, 2017.

[2] X. Cerda-Company, C. A. Parraga, and X. Otazu,, «Which tone-mapping operator is the best? A comparative study of perceptual quality,» *J. Opt. Soc. Am. A,* vol. 35, n° 14, pp. 626, 2018.

[3] R. Boitard, R. Cozot, D. Thoreau, and K. Bouatouch,, «Zonal brightness coherency for video tone mapping,» *Signal Process. Image Commun,* vol. 29, n° 12, pp. 229-246, 2014.

[4] M. Melo, M. Bessa, K. Debattista, and A. Chalmers,, «Evaluation of HDR video tone mapping for mobile devices,» *Signal Process. Image Commun,* vol. 29, n° 12, pp. 247–256, 2014.

[5] G. Eilertsen, R. K. Mantiuk, and J. Unger,, «Evaluation of Tone Mapping Operators for HDR Video,» *High dynamic range video*, vol. 36, Academic Press, 2016, pp. 185-207.

[6] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg,, «A model of visual adaptation for realistic image synthesis,» *23rd Annu. Conf. Comput. Graph. Interact. Tec, SIGGRAPH*, pp. 249–258, 1996.

[7] S. N. Pattanaik, J. Tumblin, H. Yee, and D. P. Greenberg, «Time-dependent visual adaptation for fast realistic image display,» *Proc. ACM SIGGRAPH Conf. Comput. Graph SIGGRAPH*, pp. 47–54 2000.

[8] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, «Photo-graphic tone reproduction for digital images,» *ACM Trans. Graph,* pp. 267–276, 2002.

[9] S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski,, «High dynamic range video,» *ACM Trans. Graph,* vol. 22, n° 13, pp. 319–325, 2003.

[10] D. Ramsey, J. T. Johnson, and C. Hansen, «Adaptive temporal tone mapping,» chez *Seventh IASTED Int. Conf. Comput. Graph. Imaging*, pp. 124–128,2004.

[11] S. Ploumis, M. T. Pourazad, and P. Nasiopoulos, «A Flickering Reduction Scheme for Tone Mapped HDR Video,» chez *Int. Conf. Comput. Netw. Commun (ICNC)*, pp. 73–76, 2018.

[12] B. Guthier, S. Kopf, M. Eble, and W. Effelsberg,., «Flicker reduction in tone mapped high dynamic range video,» *Color Imaging XVI Displaying, Process. Hardcopy, Appl.,* vol. 7866, p. 78660C, 2011.

[13] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, «Adaptive Logarithmic Mapping for Displaying High Contrast Scenes,» *Comput. Graph. Forum,,* vol. 22, n° 13, p. 419–426, 2003.

[14] A. Yoshida, V. Blanz, K. Myszkowski and H. P. Seidel, «Testing tone mapping operators with human-perceived reality,» *J. Electron. Imaging,* vol. 16, n° 11, p. 013004, 2007.

[15] T. O. Aydin, N. Stefanoski, S. Croci, M. Gross, and A. Smolic, «Temporally coherent local tone mapping of HDR video,» *ACM Trans. Graph,* vol. 33, n° 16, 2014.

[16] V. Popovic, E. Pignat, and Y. Leblebici,, «Performance optimization and FPGA implementation of real-time tone mapping,» *IEEE Trans. Circuits Syst. II Express Briefs,* vol. 61, n° 110, p. 803–807, 2014.

[17] C. Walbrecq, B. Magnier, and I. Marc, «A Temporal Tone Mapping Adaptation: an Application using Logarithmic Camera.,» *IEEE International Conference on Imaging Systems and Techniques (IST)*, 2018.

[18] J. Lisani, «Local contrast enhancement based on adaptive logarithmic mappings,» *Image Process. on Line,* vol. 10, pp. 43-61, 2020.

[19] N. Otsu, «A Treshold Selection Method from Gray-Level Histograms,» *IEEE Trans. Syst. Man Cybern.,* vol. 20, n° 11, pp. 62-66, 1979.

[20] N. Kiser, C., Reinhard, E., Tocci, M., & Tocci,, «Real time automated tone mapping system for HDR video,» *IEEE International conference on Image Processing*, Orlando, vol. 134, pp. 2749–2752, 2012.