HMD-TMO: A Tone Mapping Operator for 360° HDR Images Visualization for Head Mounted Displays

Ific Goudé, Rémi Cozot, Francesco Banterle

To cite this version:

Ific Goudé, Rémi Cozot, Francesco Banterle. HMD-TMO: A Tone Mapping Operator for 360° HDR Images Visualization for Head Mounted Displays. Computer Graphics International, Jun 2019, Calgary, Canada. pp.216-227, 10.1007/978-3-030-22514-8_18 . hal-02301430

HAL Id: hal-02301430

https://hal.archives-ouvertes.fr/hal-02301430

Submitted on 15 Oct 2019

HMD-TMO: A Tone Mapping Operator for 360° HDR images visualization for Head Mounted Displays *

Ific Goudé[1], Rémi Cozot[1], and Francesco Banterle[2]

[1] Univ Rennes, CNRS, IRISA, France
[2] Visual Computing Lab, ISTI-CNR, Pisa, Italy

Abstract. We propose a Tone Mapping Operator, denoted HMD-TMO, dedicated to the visualization of 360° High Dynamic Range images on Head Mounted Displays. The few existing studies about this topic have shown that the existing Tone Mapping Operators for classic 2D images are not adapted to 360° High Dynamic Range images. Consequently, several dedicated operators have been proposed. Instead of operating on the entire 360° image, they only consider the part of the image currently viewed by the user. Tone mapping a part of the 360° image is less challenging as it does not preserve the global luminance dynamic of the scene. To cope with this problem, we propose a novel tone mapping operator which takes advantage of both a view-dependant tone mapping that enhances the contrast, and a Tone Mapping Operator applied to the entire 360° image that preserves global coherency. Furthermore, we present a subjective study to model lightness perception in a Head Mounted Display.

Keywords: Head Mounted Display · High Dynamic Range · Tone Mapping Operator · 360° image.

1 Introduction

Due to the growth of Virtual Reality (VR) technologies over the last years, the visualization of 360° images has become common. 360° images can have a higher dynamic range than classic 2D images. When considering natural outdoor images, the sun can arise in certain zones of the image while dark shadows can appear in other zones. High Dynamic Range (HDR) cameras are now used to capture the whole dynamic of a scene without any loss of information, thereby providing realistic panoramas. The main issue is that all the manufactured Head Mounted Displays (HMDs) still have Standard Dynamic Range (SDR) displays, which prevents them from displaying all the dynamic range of HDR images. To appreciate HDR contents through standard displays, the well known process of Tone Mapping is used to get a limited range corresponding to SDR displays. Many Tone Mapping Operators (TMOs) exist and can be divided into two main

---

* Supported by the ANR project ANR-17-CE23-0020.

groups (global and local) and are often based on how the human perceives lightness. Each TMO can be more appropriated for a particular type of image and some user studies (see Section 2) have been conducted to determine which one is a better candidate for 360° images. These studies show that there is no preferred TMO, and emphasizes on the necessity of developing one dedicated to HMD. For this purpose, two TMOs have been proposed in the literature (see Section 2). There are several different approaches to the problem of tone mapping a 360° HDR image. One of them is to consider the content overall, the 360° image is processed at once, toward its entire dynamic range. The obtained result is globally coherent but, when considering a viewport of the entire 360° image, the contrast is reduced. However, as the user can only watch a limited part of the 360° image at a time, a TMO may be applied to the current viewport. Thus, the viewport contrast is enhanced while the global coherency is lost.

To overcome this problems, we propose a method that takes into account the results of two TMOs: one applied to the entire 360° image, and the other to the current viewport. As will be explained later, the viewport TMO provides a better contrast, while the global TMO preserves the spatial coherency. The main contributions of this paper are: (1) a subjective evaluation to model lightness perception on an HMD; (2) a novel TMO for 360° HDR images that ensures a spatial coherency and enhances contrasts by combining the luminances (physical quantity of light) provided by these two TMOs.

After introducing in Section 2 related works on TMOs dedicated to 360° images visualization on HMD, we present the user study we conducted to model lightness perception (subjective perception of light by the human eye) in Section 3. As a result, we show that the perception model of the lightness on a classic 2D display is still valid on an HMD. Then, we describe in details our HMD-TMO in Section 4. Next, in Section 5, we comment on our results and discuss the efficiency of our approach. Finally, Section 6 concludes the paper and presents some research avenues for future work.

2 Related work

Two approaches have been considered to visualize 360° HDR images on HMDs. The first consists in applying existing TMOs to the entire 360° image and display the result on the HMD. Some studies performed a comparison of many TMOs for many 360° HDR images in order to find the most appropriated TMO. As for the second approach, a new TMO is proposed, it considers the specific visualization conditions on an HMD by applying a TMO to the viewport only. The first comparison of existing TMOs for 360° images using a subjective evaluation has been run by Perrin *et al.* [1]. However, none of the evaluated TMOs show a clear increase of perceived quality. Melo *et al.* [2] have ran another user study to compare four different TMOs on five 360° HDR images and found similar results. So, we cannot rely on existing TMOs in the case of visualization of 360° HDR images on HMD, a specific operator has to be developed. Yu [3] has adapted an existing operator to propose a TMO that takes advantage of the particularity of visualization on HMD. The main contribution was first to take into account the fact that a user only looks at a limited part of the 360° image at a time, and second to simulate the light and dark adaption of human vision to provide smooth transitions between successive views. When compared to previous evaluations, instead of applying the TMO to the entire 360° image, Yu's applies a TMO only to the current viewport. He has adapted the Photographic Tone Reproduction operator [4] accordingly to the specific key value (log-average luminance) of the viewport. Indeed, the key value can significantly change from a view to another. To prevent flickering, Yu proposed to smooth the key value between successive views to coarsely reproduce the human eye adaptation behavior. Cutchin and Li [5] proposed a method that performs a tone mapping on each viewport independently depending on its luminance histogram. The viewport histograms are divided into four groups corresponding to different TMOs. Authors noticed popping effects that happen when two successive views belong to different groups: as the TMO is different, the image shifts dramatically. Both methods benefit from view dependency on an HMD and provide a better perceptible quality, but they still present some limits we want to overcome. These two methods only perform a global operation on the viewport without worrying about the spatial coherency of the entire 360° image. We propose a method that takes advantage of the viewport dependent operation with smooth transitions between successive viewports to ensure a good contrast while maintaining a global coherency considering the luminance of the entire 360° image. Our method implement the logarithm of the luminance to mimic the human perception of the lightness. To ensure the validity of this representation in case of visualization on HMD, we conducted a subjective evaluation that models the human perception.

3 Lightness perception on HMD

For classic 2D displays in a controled visualization environment, the lightness is modelled as the logarithm of the luminance. This result comes from Weber and Fechner [6] studies about lightness perception, and many TMOs are based on this work. We will show, thanks to a subjective evaluation, that this result holds for HMDs.

3.1 The lightness perception model

Weber showed that the human capacity to distinguish a stimulus from the background is linearly proportional to the background luminance. In other words, the lighter the background L, the higher the difference ΔL should be (between stimulus and background) to perceive the stimulus. This ratio is commonly known as the Just Noticeable Difference (JND):

$$JND = \frac{\Delta L}{L} = k, \tag{1}$$

with ΔL the luminance difference between the stimulus and the background (in cd/m^2), L the background luminance (in cd/m^2) and k a constant (around 0.01

for traditional visualization condition on a 2D display [7]). Fechner integrated Weber's result to obtain the response of the visual system based on the luminance transducer:

$$\frac{dR}{dl}(L) = \frac{1}{\Delta L(L)}, \tag{2}$$

$$R(L) = \int_0^L \frac{1}{\Delta L(l)} dl = \frac{1}{k} ln(L), \tag{3}$$

where R is the lightness response for a given luminance L, and ΔL is actually the perceived difference measured by Weber's experiment. Accordingly, the subjective perception of lightness is assumed to be the logarithm response to the physical luminance.

3.2 User study

In our experiment, our objective was to reproduce as closely as possible the Weber's experiment while following the CIECAM recommendations [8]. The CIECAM suggests a circular stimulus with a radius between 2° and 4° to match with the foveal vision. The background has an achromatic color with a radius of 20° to match with the peripheral vision. Finally, the surround field encompasses the rest of the vision field (see Figure 1). Indeed, in case of visualization on HMD, the black plastic structure around the displays is interpreted as the surround. Recall that the CIECAM model has been determined for a background covering a 20° vision for a classic display, while this angle corresponds to all of the Field of View (FoV) of the HMD. We consider a stimulus of 4°, a background covering all the FoV of the used HMD (about 100°), and the surround field is ignored.

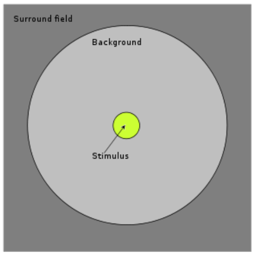

Fig. 1. CIECAM conditions recommendations.

Twenty participants have been presented with ten background luminance levels covering all the dynamic range of the HMD. For each background, a slightly lighter stimulus was displayed and incrementally increased until it gets perceived by the participant. This allows us to determine the JND. The test lasted about 15 minutes and our panel consisted of 20 participants (13 men and 7 women) with normal vision, from 20 to 57 years of age, with various socio-cultural backgrounds. After data fitting (see Figure 2), we found the JND is equal to 2%. Despite of finding linear sensitivity (ΔL as a function of

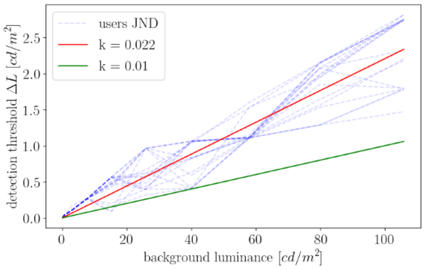

Fig. 2. ΔL as a function of L given the JND on an HMD.

L) resulting in a logarithmic response by the Fechner's integration, in line with 2D visualization, the visual system perceives two times less contrast between the stimulus and the background. Indeed, the JND approaches 2%, while traditional visualization on a 2D display is usually around 1%. This result means that we can lose perceptible fine details when we watch an image on an HMD. This evaluation also emphasizes that the logarithmic lightness function (Equation 3) is valid to model the human perception on an HMD. We designed our HMD-TMO accordingly to this result.

4 HMD-TMO

As seen in previous evaluations [1, 2], applying a TMO to the entire 360° image does not produce a satisfying quality. A TMO applied to the entire 360° image takes into account all the luminance of the scene to produce the tone mapped result. Therefore, it maintains a global coherency but loses contrast in the viewports. Regarding a TMO applied to the viewport, it preserves the contrast but loses the global coherency of the scene. Since none of these methods produces a satisfying result when applied independently, we developed a TMO that combines both methods, global and viewport based, adapted to visualization on HMD in order to preserve global coherency and enhance contrast. Our operator consists in a pipeline with two branches (see Figure 3). The input is a 360° HDR image and the output is a tone mapped image of the current viewport. The upper branch performs a tone mapping on the entire 360° image and thus preserves the spatial coherency, while the lower branch performs a tone mapping on the viewport image to enhance the contrast. Note that the time parameter t means the viewports succession due to the movement of the user

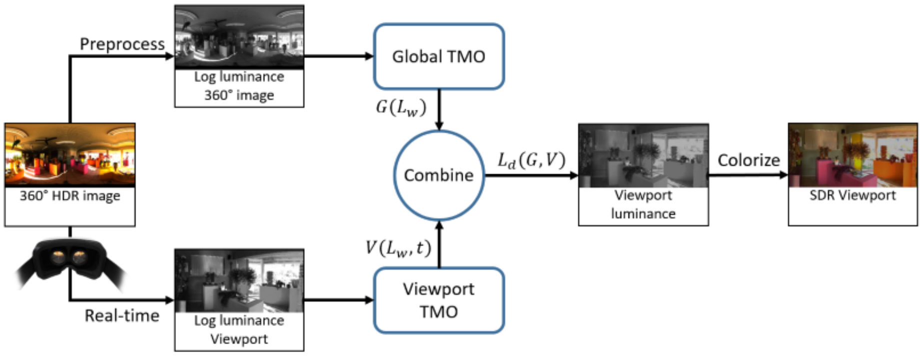

Fig. 3. Our operator combines a Global TMO $G(L_w)$ and a Viewport TMO $V(L_w, t)$. The Global TMO (upper branch) aims to preserve the global coherency of the scene while the Viewport TMO (lower branch) aims to enhances contrast. The combination of both produces our final HMD-TMO $L_d(G, V)$.

who is wearing the HMD. The combination of the resulting luminances of these two TMOs (tone mapping of the entire 360° image $G(L_w)$ and tone mapping of the viewport image $V(L_w, t)$) is calculated by a geometric mean to compute the final tone mapped luminance $L_d(G, V)$. Note that luminance stands for the Y channel in the CIE XYZ color space. Each component of our pipeline (entire image tone mapping, viewport tone mapping, and combination) are detailed in the following subsections.

4.1 Global tone mapping

First, we tone map the entire 360° image in order to preserve the global coherency of the scene. Similarly to Ward *et al.* Visibility Matching Tone Reproduction operator [9], our method is based on the log-luminance histogram of the image represented by a *Cumulative Distribution Function (CDF)*. First, Ward *et al.* compute the normalized cumulative histogram of log-luminance value to find the *CDF*:

$$P(b) = \frac{\sum_{b_i<b} f(b_i)}{\sum_{b_i} f(b_i)}, \tag{4}$$

where $f(b)$ is the number of pixels for a bin b. The number of bins is set to 100 as proposed by the authors to avoid banding artifacts due to quantization. The tone curve G proposed by Ward *et al.* is a scaled version of $P(b)$:

$$G\big(L_w(x,y)\big) = exp\big(log(L_{dmin}) + \big(log(L_{dmax}) - log(L_{dmin})\big) \times P\big(L_w(x,y)\big)\big), \tag{5}$$

where L_{dmin} and L_{dmax} are respectively the minimum and maximum luminance of the display, $L_w(x,y)$ is the world luminance of the pixel (x,y), $P(L_w(x,y))$ the *CDF* defined in Equation 4, and $G\big(L_w(x,y)\big)$ (our Global TMO) the resulting luminance of the pixel. To better match with human perception, Ward *et al.* add a pass of histogram adjustment in case of a too high contrast in the tone mapped image. Given the lightness perception values, the log-luminance histogram is clipped to avoid contrast exaggerations, which results in a flat *CDF*. Figure 4 shows the result of three TMOs applied to the entire 360° image: TMO based on Equation 5 ($G\big(L_w(x,y)\big)$), Ward *et al.*'s TMO [9] (with clipping), and Reinhard *et al.*'s TMO [4]. Reinhard *et al.*'s [4] TMO (first row) preserves the global coherency: the back of the store (red inset) appears dark while the storefront (green inset) is slightly lighted. Ward *et al.*'s [9] TMO (second row) better represents the difference between the dark and the bright zone but it is not enough for the expected result due to the histogram clipping. Finally, the TMO using Equation 5 (third row) seems perceptually too contrasted (the difference between the red and the green inset is significant) but it is the most representative of the global coherency. As will be seen in the next subsection, the contrast is managed by the viewport TMO.

4.2 Viewport tone mapping

The objective of the viewport TMO is to enhance the contrast in the part of the 360° image currently viewed by the user (the viewport). We used an improved

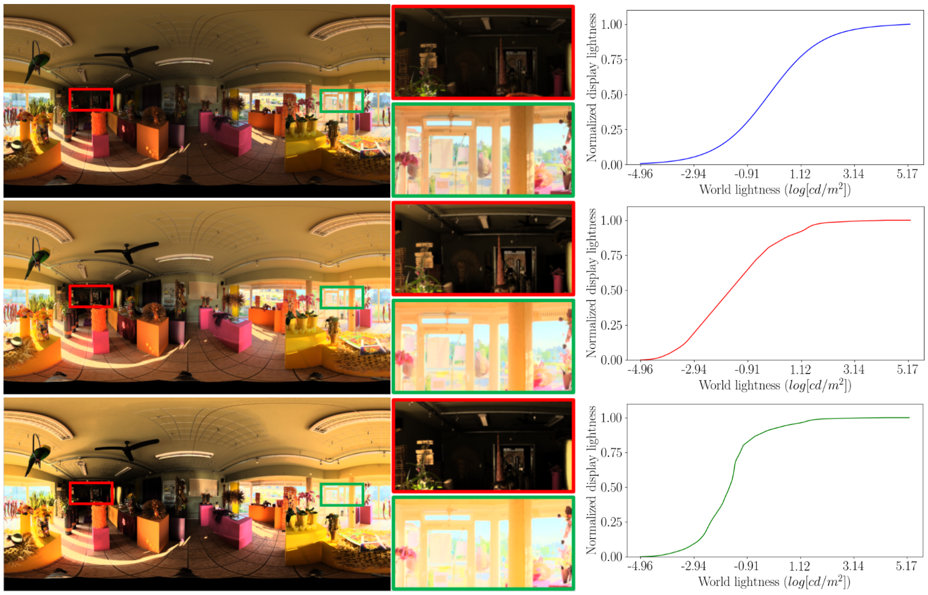

Fig. 4. From top to bottom: the Photographic Tone Reproduction operator [4], the Visibility Tone Reproduction operator [9], and the *CDF* (all are applied to the entire 360° image). To the right of each image, the corresponding tone mapping curves.

version of the viewport TMO proposed by Yu [3]. His method relies on a simplified Photographic Tone Reproduction operator [4]. To avoid any flickering when changing view, Yu proposed to temporally smooth the key values to simulate the adaptation behavior of the human eye (see Figure 5). This TMO is based on the log-average luminance of the image:

$$\bar{L}_w\big(V(t)\big) = \frac{1}{N} exp\left(\sum_{x,y} log\big(\delta + L_w(x,y)\big)\right), \tag{6}$$

where $\bar{L}_w\big(V(t)\big)$ is the viewport key value at a given time, $L_w(x,y)$ the pixel luminance, δ a small value to avoid singularity in case the image contains black pixels, and N the number of pixels in the viewport. Here, time t corresponds to an orientation of the camera due to the head movement. To ensure a smooth transition between two successive viewports, the key and the white values are interpolated as:

$$\bar{L}_w^{'}(t) = \alpha \bar{L}_w\big(V(t)\big) + (1-\alpha)\bar{L}_w^{'}(t-1), \tag{7}$$

$$L_{white}^{'}(t) = \alpha L_{white}\big(V(t)\big) + (1-\alpha)L_{white}^{'}(t-1), \tag{8}$$

where $\bar{L}_w^{'}(t)$ and $L_{white}^{'}(t)$ are respectively the smoothed key and white values between two successive views and α is a time dependent interpolation variable.

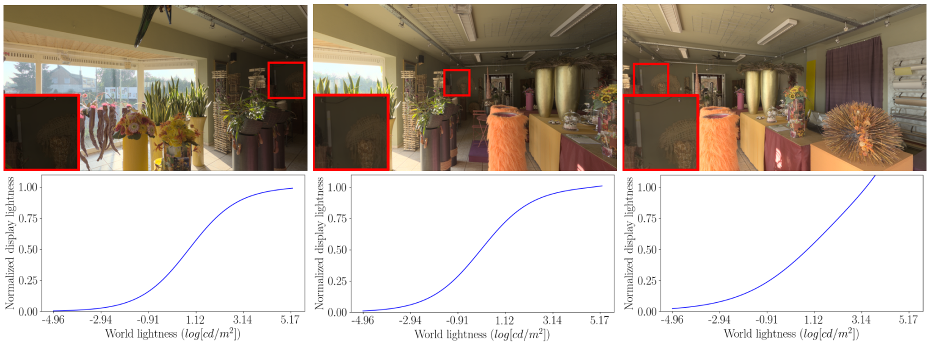

Fig. 5. Photographic Tone Reproduction operator [4] applied to a viewport sequence with smooth transitions. As the key value and the white value evolve from a view to another, the tone curve is modified and a same zone in the scene (red inset) becomes brighter or darker.

Finally, the luminance is scaled and high values are attenuated to avoid clipping:

$$L(x,y,t) = \frac{a}{\bar{L}'_w(t)} L_w(x,y), \tag{9}$$

$$V\big(L_w(x,y),t\big) = \frac{L(x,y,t)\left(1 + \frac{L(x,y,t)}{L'^2_{white}(t)}\right)}{1 + L(x,y,t)}, \tag{10}$$

where a is a user defined variable which scales the luminance (commonly 0.18), $L(x,y,t)$ the time dependent scaled luminance and $V\big(L_w(x,y),t\big)$ (our Viewport TMO) the displayed luminance. In his operator, Yu actually uses Equation 9 that does not avoid clipping in high luminances. We have now the global coherency assured by the 360° image CDF $(G(L_w))$ and the viewport contrast $(V(L_w,t))$ we want to combine to obtain our final tone mapped image.

4.3 TMOs combination

To recap, we want to display a tone mapped viewport of a 360° HDR image. In order to combine both luminances provided by the global and viewport TMOs, ensuring the global coherency to be preserved and the contrast to be enhanced, we propose to use a geometric mean to combine both luminances. After experimenting with different combinations (arithmetic mean, weighted sum and geometric mean), the geometric mean produces the best results (see Figure 6). The resulting images produced by the weighted sum combination are perceptually unpleasant.

As the lightness is the logarithm of the luminance, this combination is interpreted as a perceptual mean:

$$\left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n}} = exp\left(\frac{1}{n}\sum_{i=1}^{n} ln(a_i)\right), \tag{11}$$

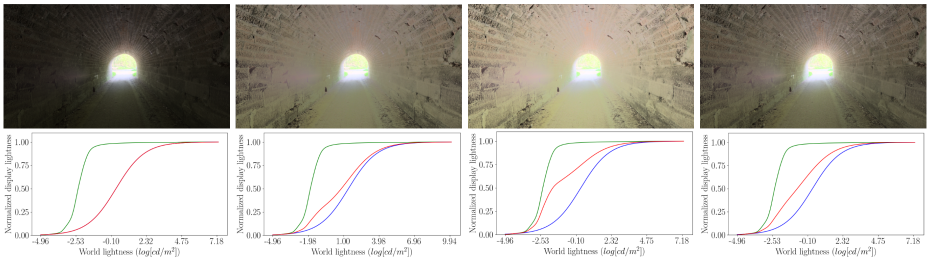

Fig. 6. Comparison between weighted sum and geometric mean combinations. Combinations are respectively: 100% Viewport, 80% Viewport + 20% Global, 50% Viewport + 50% Global, and the geometric mean. The green, blue and red curves below correspond respectively to the Global TMO, the Viewport TMO and the combination.

$$L_d(G,V) = L_d(x,y,t) = \sqrt{V\big(L_w(x,y),t\big) \times G\big(L_w(x,y)\big)}, \tag{12}$$

where $L_w(x,y)$ is the world luminance of the pixel (x,y), t is the time, $V(L_w(x,y),t)$ is the viewport TMO defined in Equation 10 and $G\big(L_w(x,y)\big)$ is the global TMO defined in Equation 5. $L_d(G,V)$ is our final tone mapped luminance we display on the viewport. In Figure 7, we show the behavior of global, viewport and combined TMOs in both dark and bright areas of a 360° image. When only the viewport TMO is applied, the global coherency is lost, both zones seem equally enlightened. Contrarily, regarding the global TMO, the global coherency is preserved but the contrast in the viewport is exaggerated and unnatural. Finally, the combination of these two TMOs preserves global coherency and provides a proper contrast in the viewport.

Fig. 7. Two opposite views in a same scene, a dark side (top) and a bright side (bottom). From left to right: the viewport TMO $V\big(L_w(x,y),t\big)$ (blue curve), the combination of viewport and global TMOs $L_d(G,V)$ (red curve), and the global TMO $G\big(L_w(x,y)\big)$ (green curve).

5 Results

Once our TMO has calculated the tone mapped luminance, we compute the color of all the pixels of the tone mapped image using the Schlick's approach [10]:

$$C^{'} = \left(\frac{C}{L_w}\right)^{s} L_d, \tag{13}$$

where C and $C^{'}$ are respectively the input and output trichromatic values (RGB), L_w the world luminance and L_d the tone mapped luminance. The saturation parameter s is set to 0.7 for our results.

We implemented our HMD-TMO using Unity3D because of its friendly interface for managing VR and its capacity to handle HDR. We used the HTC Vive Pro[3] as HMD. We benefited from GPU programming with shaders to compute 360° image histograms on the 2048×1024 equirectangular projection, and the 1440×1600 viewports (left and right views) key values in real time. Rendering (computation of the colored tone mapped image) is achieved with an image effect shader applied to the HDR viewport. The global TMO is computed once for all and takes less than one second. The navigation (calculation and display of successive viewport images) is performed in real-time: 90 frames are computed per second (Intel Core i7 vPro 7th Gen, NVidia Quadro M2200). We used a dataset of 90 views: 15 views for each of the six different 360° HDR images. We computed the Tone Mapped Image Quality Index (TMQI) [11] score of each view of each 360° image, which amounted to compute 90 scores. We calculated the average of the viewport scores for our method and three other TMOs. Overall, our method had the best mean score (see Table 1).

Table 1. The result of the TMQI quality test: mean value computed on 90 images (Reinhard *et al.*'s [4] and Ward *et al.*'s [9] TMOs are applied to the entire 360° image).

	Reinhard *et al.*	Ward *et al.*	Yu	Ours
TMQI quality	0.798	0.854	0.865	**0.887**

We also compared visually our results to those of Yu's [3] (see Figure 8). In addition to preserving the global coherency, our TMO avoids clipping luminances out of the dynamic range of the HMD. This improvement is shown on the church wall in the tree image, and in the background and at the bottom left corner of the forest image. Furthermore, due to the exaggerated contrast produced by the *CDF*, our HMD-TMO enhances the fine details. Indeed, in the tree image, the contrast between the night sky and the tree leaves is higher with our method, which allows us to distinguish holes through the foliage. The branches lying on the ground are also more detailed in the forest image. The same phenomenon occurs in both examples of the village image and in the folds of the curtain in

---

[3] https://www.vive.com/fr/product/vive-pro/

the florist image. However, an unwelcome effect appears in the florist image. The lighting of the box practically disappears when using our TMO because its luminance is not enough represented by the *CDF* which flattens the zone within the green inset.

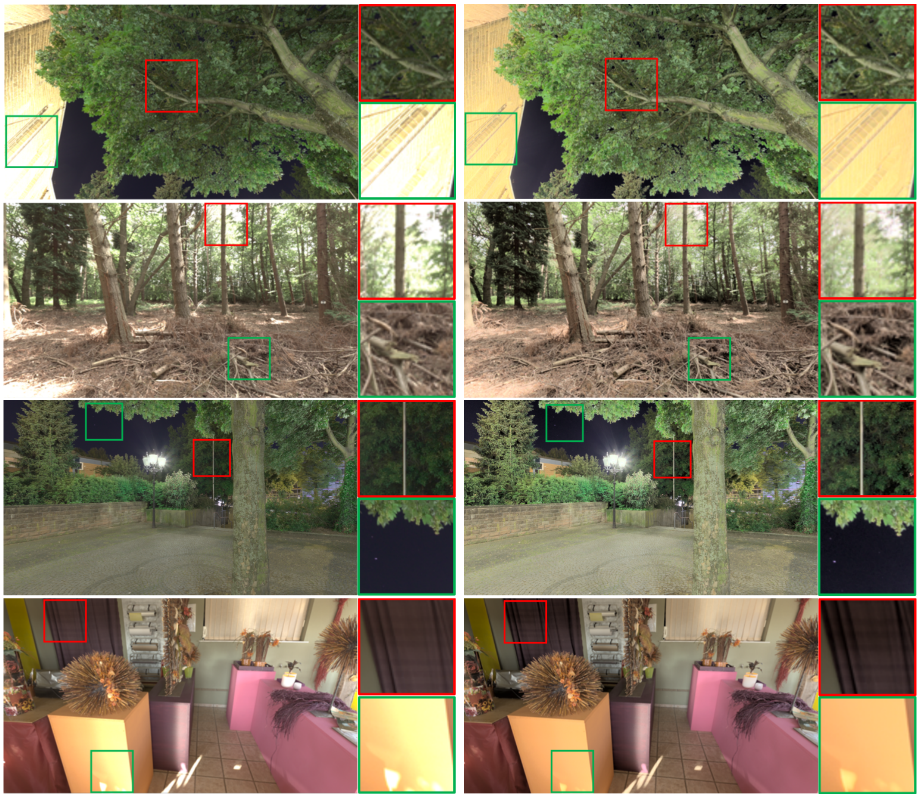

Fig. 8. Yu's TMO [3] (left) compared to our HMD-TMO (right). Our method enhances fine details and removes the clipping in high luminance.

6 Conclusion

HDR imaging enables to capture the whole dynamic of a 360° scene. Previous subjective studies have shown that naive tone mapping of the entire 360° image or tone mapping of a viewport does not provide convincing results. To overcome these limitations, we have proposed a new HMD-TMO. More precisely, our contribution is twofold: (1) a logarithmic model of lightness still valid on an HMD; (2) a novel TMO that combines both global and viewport TMOs. This new TMO doesn't tackle the limits of a viewport tone mapping but ensures a spatial coherency while navigating through the 360° HDR content. Our future work heads toward HDR video tone mapping for visualization on HMD. The main challenge will consist in accounting for: temporal coherency, sudden change in luminance range through time, naturalness of time adaptation, etc.

Acknowledgments

All 360° HDR images come from free SYNS and LizardQ datasets. This work has been supported by the ANR project ANR-17-CE23-0020. We would like to thank Kadi Bouatouch for his help and proofreading. Thanks to all experiment participants for their contributions.

References

1. A.-F. Perrin, C. Bist, R. Cozot, and T. Ebrahimi: Measuring quality of omnidirectional high dynamic range content. Optics+Photonics Optical Engineering+Applications Applications of Digital Image Processing XL. 38 (2017)
2. M. Melo, K. Bouatouch, M. Bessa, H. Coelho, R. Cozot, and A. Chalmers: Tone Mapping HDR Panoramas for Viewing in Head Mounted Displays. International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. 232-–239 (2018)
3. M. Yu: Dynamic Tone Mapping with Head-Mounted Displays. Standford University Report. 5 (2015)
4. E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda: Photographic Tone Reproduction for Digital Images. Conference on Computer Graphics and Interactive Techniques. (29), 267-–276 (2002)
5. S. Cutchin and Y. Li: View Dependent Tone Mapping of HDR Panoramas for Head Mounted Displays. The Eurographics Association. (2016)
6. G. T. Fechner Howes, D. H. Howes, E. G. Boring: Elements of psychophysics. Holt, Rinehart and Winston: New York. **1**, (1966)
7. E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski: High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting. Elsevier Science. (Second Edition), 238 (2010)
8. M. D. Fairchild: Color appearance models. John Wiley & Sons. (2013)
9. G. W. Larson, H. Rushmeier, and C. Piatko: A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes. IEEE Transactions on Visualization and Computer Graphics. **3**(4), 291-–306 (1997)
10. C. Schlick: Quantization techniques for the visualization of high dynamic range pictures. Photorealistic Rendering Techniques, Springer-Verlag Berlin Heidelberg New York, Eurographics, 7–20 (1994)
11. H. Yeganeh and Z. Wang: Objective Quality Assessment of Tone Mapped Images. IEEE Transactions on Image Processing, **22**(2), 657–667, (2013)