A Perceptually Coherent TMO for Visualization of 360° HDR Images on HMD

Ific Goudé, Rémi Cozot, Olivier Le Meur

A perceptually coherent TMO for visualization of 360° HDR images on HMD *

Ific Goudé[1], Rémi Cozot[2], and Olivier Le Meur[1]

[1] Univ Rennes, CNRS, IRISA, France
[2] Littoral Opal Coast University, France

Abstract. We propose a new Tone Mapping Operator dedicated to the visualization of 360° High Dynamic Range images on Head-Mounted Displays. Previous work around this topic has 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 globally the dynamic range of the luminance of the scene. To cope with this problem, we propose a novel Tone Mapping Operator which takes advantage of 1) a view-dependant tone mapping that enhances the contrast, and 2) a Tone Mapping Operator applied to the entire 360° image that preserves the global coherency. Furthermore, the proposed Tone Mapping Operator is adapted to the human eye perception of the luminance on Head-Mounted Displays. We present two subjective studies to model the lightness perception on such Head-Mounted Displays.

Keywords: High Dynamic Range · Tone Mapping Operator · Head-Mounted Display · 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. Moreover, High Dynamic Range (HDR) cameras are now used to capture the whole dynamic of a scene with much more details in brightest and darkest areas, thereby providing realistic panoramas.

Nonetheless, all manufactured Head-Mounted Displays (HMDs) still have Standard Dynamic Range (SDR) screens, which prevent 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 that of SDR displays. Many Tone Mapping Operators (TMOs) exist [1,16] and can be divided into two main groups (global and local) and are often based on how the human perceives lightness and colors. In order

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

to adapt existing TMOs to HMD visualization, we conducted two subjective evaluations to investigate how the Human Visual System (HVS) perceives images on HMDs.

Beyond perception, different approaches to address the problem of 360° image tone mapping on HMD can be considered. One solution is to apply the TMO to the whole 360° image, considering its entire dynamic range. The obtained result is globally coherent but, when considering only a viewport, the contrast can be unpleasantly reduced. 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 problem, 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. A preliminary version of this work has been reported [8]. The main contributions of this extended paper are: (1) a thorough work about lightness and colors perception on HMD that includes two subjective evaluations; (2) an improved TMO for 360° HDR images that ensures a spatial coherency and enhances contrasts while being perceptually coherent with the lightness perception of the human eye on HMD.

The paper in organised as follows. Section 2 introduces related work on perception models and TMOs dedicated to 360° images visualization on HMD. Then, we present the subjective evaluations we conducted to model lightness and colors perception in Section 3. As a result, we show that the perception model on a classic 2D display is slightly different than the one on an HMD. Then, we describe in detail our HMD-TMO in Section 4. 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. All acronyms and definitions are referred in the glossary Table 1.

Table 1: Glossary of definitions and acronyms.

Luminance	Physical quantity of light emitted by an area in $[cd/m^2]$.
Brightness	Attribute of a visual sensation according to which an area appears to emit more or less light [5].
Lightness	The brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting [5].
HDR / SDR	High Dynamic Range / Standard Dynamic Range
HMD	Head-Mounted Display
TMO	Tone Mapping Operator
VR	Virtual Reality
HVS	Human Visual System
JND	Just Noticeable Difference
CAM	Color Appearance Model
FoV	Field of View
CDF	Cumulative Distribution Function

2 Related work

Understanding the HVS and how it reacts to stimuli is essential to ensure an efficient tone mapping. Color Appearance Models (CAMs) seek to describe how a stimulus is perceived depending on viewing conditions, such as the luminance of the background and the surround. Existing TMOs are often based on CAMs [1,16] to ensure the image processing to be coherent with the way we will perceive the tone mapped result. As the viewing conditions change on a HMD, the perception is different and CAMs have to be adapted to those conditions.

2.1 Lightness perception model

The basis of psycho-physical studies about lightness perception comes from the seminal work of Weber. He showed that the human capacity to distinguish a stimulus from a background is linearly proportional to the background luminance. In other words, the lighter the background L, the higher the difference ΔL (between stimulus and background) should be 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 [16]). Fechner integrated Weber's result to obtain the response of the visual system [6]:

$$R(L) = \int_0^L \frac{1}{kL(l)} dl = \frac{1}{k} ln(L), \tag{2}$$

where R is the lightness response for a given luminance L. Accordingly, the subjective perception of luminance, called lightness, is assumed to be the logarithm response to the physical luminance.

More recently, Stevens showed limits of the Fechner's model and proposed to use a power function to model the lightness perception [19]. Stevens psychophysical studies have led to the lightness perception equal to the physical luminance raised to the power of one third. Both of those models are still used and seem to give similar results in comparable conditions (see Figure 1). The debate to know which representation of the lightness perception is the best one, i.e more accurate, is still open and research continue on this topic [2]. Decades later, Whittle conducted a subjective evaluation of lightness perception following a different protocol [21]. He measured the JND between two stimuli (respectively the reference and the test) in front of a unique background. The obtained results are similar to Weber's ones, luminance discrimination is equal to a constant k:

$$\frac{\Delta W}{W} = k, \tag{3}$$

$$W = \frac{\Delta L}{L_{min}}, \tag{4}$$

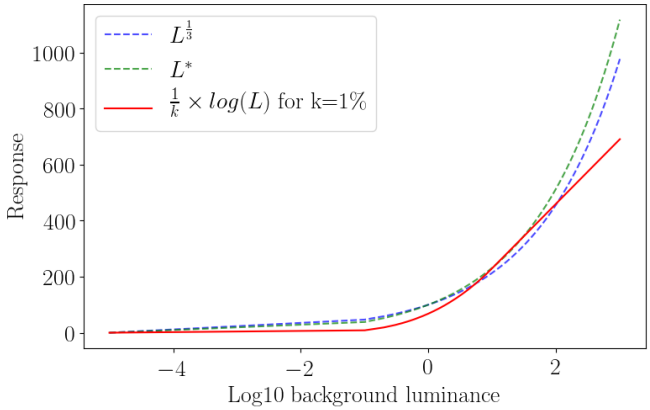


Fig. 1: Lightness functions comparison.

with ΔL the luminance difference between the stimulus and the background, L_{min} the minimum luminance between the reference stimulus and the background. ΔW is the difference between the values of W for the two stimuli (reference and test).

This kind of representation is sufficient as a simple model of lightness perception, but current CAMs take into account more parameters of the HVS, such as the chrominance and the viewing conditions. First proposed by the CIE in 1997 (CIECAM97), this model has been reviewed several times. The most recent version (CAM16) has been proposed by Li *et al.* [12], it is a modified version of CIECAM02 written by Fairchild [5]. In his work, the perceived lightness J of a stimulus depends on the luminance stimulus, the luminance background and the lighting conditions of the surround. It can be expressed as:

$$J = 100 \times \left(\frac{A}{A_w} \right)^{c.z}, \tag{5}$$

$$z = 1.48 \times \sqrt{\frac{Y_b}{Y_w}}, \tag{6}$$

where A and A_w are the achromatic response of stimulus and the achromatic response of white reference respectively. z is defined as the base exponential nonlinearity with Y_b and Y_w the background luminance and the white reference luminance respectively. c is a nonlinearity factor of brightness function depending on the viewing condition (surround enlightening as illustrated in Figure 2a) that can be Dark, Diminish or Average. The enlightening condition of the surround (expressed as c) has a direct effect on the degree of adaptation of the human eye denoted F. On a HMD, the viewing conditions are not well defined. As described later in Section 3, we suppose that the surround component is a function of the background and the background luminance influences much more the lightness perception than on classic 2D display. That why we need a complex perception

model that takes into account the background luminance as a function. The influence of the background size and complex enlightening environment has been studied in [7,11]. Some adjustments have been proposed by Lee and Kyu-Ik [11] to express the adaptation degree F depending on the background luminance L_b instead of the three original viewing conditions (Dark, Dim, Avg):

$$F = \begin{cases} 0.7379 + 0.392\big(1 - exp(0.0221.L_b)\big), & \text{if } L_b < 50cd/m^2 \\ 1, & \text{otherwise} \end{cases} \tag{7}$$

where the adaptation degree F is only a function of the background luminance. Then the new nonlinearity factor, called now c_L, is computed depending on an adaptation luminance L_a:

$$L_a = F.L_b + 0.2(1 - F)L_{d_{max}}, \tag{8}$$

with $L_{d_{max}}$ the maximum luminance of the display and L_a the luminance of adaptation (20% of $L_{d_{max}}$ in traditional viewing condition). Finally, the adapted nonlinearity factor c_L is:

$$c_L = \frac{c.\Delta L_{a|L_a=50}}{\Delta L_a}, \tag{9}$$

$$\Delta L_a = 1.88L_a^{0.23} - 7.24L_a^{0.11} + 8.26, \tag{10}$$

where the nonlinearity factor of brightness c is equal to 0.69. $L_a = 50$ is a limit defined by the authors where $\Delta L_{a|L_a=50}$ is equal to 1.75.

This complex model allows us to express the nonlinearity adaptation factor c_L and the adaptation degree F as a function of the background luminance. Nevertheless, the nonlinearity factor c_L (equation 9) has to be adapted to fit with the particular viewing conditions encounter on HMDs. We ran subjective evaluations to compare perception on HMD with known models and found that the lightness perception is halved on a HMD compared with a 2D display (see Section 3).

2.2 Tone Mapping on HMD

Assuming that perception is the same for HMDs and 2D displays, two user studies performed a subjective comparison of several TMOs applied to many 360° HDR images in order to find the most appropriated one. The first evaluation ran by Perrin *et al.* [15] consists in applying existing TMOs to the entire 360° HDR image and display the obtained result on the HMD. However, none of the evaluated TMOs shows a clear improvement of perceived quality. The year after, Melo *et al.* [13] ran another user study to compare four TMOs on five 360° HDR images and found similar results. These results suggest that existing TMOs should be adapted to meet the requirements of 360° HDR images displayed on HMDs.

A few attempts to propose a TMO dedicated to HMD are presented right after. Yu [22] adapted the Photographic Tone Reproduction operator [17] applied

to the viewport. 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. A second aimed to simulate light and dark adaption of human vision to provide smooth transitions between successive viewports. Indeed, the key value (log-luminance average) can significantly change from a view to another. To prevent flickering artifacts, Yu proposed to smooth the key value between successive views to coarsely reproduce the human eye adaptation behavior.

Cutchin and Li [3] 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. The authors noticed popping effects that happen when two successive views belong to different groups.

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. Especially, these two methods do not tackle the spatial coherency as the TMO is applied to the viewport only. We propose a method that takes advantage of the viewport dependent operation with smooth transitions between successive views to ensure a good contrast while maintaining a global coherency considering the luminance of the entire 360° image. We will now present the subjective evaluations we conducted to model the lightness perception on HMD before detailing our method in Section 4.

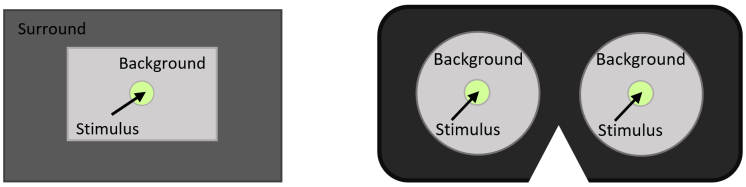
3 Perception on HMD

Before delving into the proposed TMO dedicated to HMD, we describe two subjective experiments we conducted to study the HVS response on HMD. The first one focuses on lightness perception and measures the JND in the dynamic range of the HMD. Moreover, the protocol design follows CIECAM recommendations [5]. This experiment is helpful to validate the use of the logarithm of the luminance as a good representation of the perceived lightness. The second subjective evaluation is intended to be more general, regarding lightness and saturation. The experiment takes Whittle design [21] consisting in presenting two stimuli in front of a uniform background. Its usefulness is twofold. First it confirms results of the previous experiment about lightness. Second, it allows an evaluation of the perception of the chrominance on HMD. Resulting from these two evaluations, we propose to adapt a CAM that is coherent with the perception on HMDs and used in our TMO as describes later in Section 4.

3.1 First experiment: lightness as a function of luminance

The classic Weber's experiment seeks to determine the minimum perceptible difference value of luminance between a stimulus and a background. A fixed background is presented to participants while the stimulus is imperceptibly lighter. Then the stimulus luminance is increased until participants notice a difference between the stimulus and the background. This relative difference is the JND.

More recently, the CIECAM [5] suggests to consider the surround relative luminance as a parameter because the perception acts differently depending on the environment enlightening. Furthermore, the stimulus should have a radius ranging from 2° to 4° in the visual field (corresponding to the foveal vision), a radius of 20° for the background (peripheral vision) and the surround encompasses the rest of the vision field (see Figure 2a). These recommendations are well defined



(a) (b)

Fig. 2: (a) CIECAM02 recommendations for visualization conditions on 2D display. (b) Visualization conditions on HMD.

in case of visualization on a classic 2D display. New constraints are met when considering visualization on HMD. First, the surround field is not considered anymore because the black plastic structure of the HMD encompasses the whole visual field. Second, while the CIECAM model has been designed for distant display visualization that covers vision of about 20°, this angle corresponds to all of the Field of View (FoV) of the HMD (about 100°). To sum up, in our experiment on HMD, we consider a 4° stimulus, the background covers all the Fov of the HMD (100°), and the surround field is ignored (see Figure 2b). The JND has been determined for ten background luminance levels covering all the dynamic range of the HMD. The test lasted about 15 minutes and the 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 using robust estimators and classical regressions, we found the JND is linear and the slope is equal to 2.2% (± 0.3%) as illustrated in Figure 3. The sensitivity is still linear (ΔL as a function of L), resulting in a logarithmic response when using the Fechner's integration (equation 2). This evaluation emphasizes that the logarithmic lightness function is valid to model the human perception on an HMD.

However, this JND approximated to 2% is interpreted as a loss of contrast (two times less) for visualization on HMD compared with visualization on classic 2D display where the JND is usually around 1% (see Figure 1). We suppose this phenomenon is due to the lack of fixed surround luminance. CAM proposed by Fairchild [5] considers three viewing conditions for the surround: Dark, Dim and Average. Indeed, the light emitted from the displays in the headset scatters on the plastic structure. Assuming the structure of the HMD is equivalent to the surround field in the CIECAM recommendations, the surround S is then a

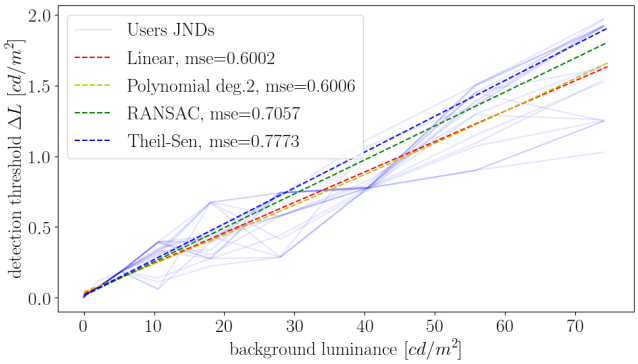


Fig. 3: ΔL as a function of L given the JND on an HMD.

function of the background B ($S = f(B)$). Based on Lee and Kyu-Ik [11] work that takes into account the influence of background luminance in the lightness equations, we also adjusted the viewing-dependent component c_L to fit with our results.

$$c_L = \frac{c.r.\Delta L_{a|L_a=50}}{\Delta L_a}, \tag{11}$$

$$r = \frac{0.01}{k}, \tag{12}$$

where r is the ratio between the classic viewing condition constant at 1% and the found constant $k = 2.2\%$ on HMD. c is still equal to 0.69. We scaled the parameter c_L depending on the perception on HMD, which is almost halved.

We simulate the difference in lightness perception between two stimuli for a solid background in Figure 4. The perceived lightness is strongly attenuated by the background luminance in case of HMD visualization compared with the three viewing conditions proposed in CIECAM02. These two factors (F and c_L equations 7 and 11) are used in the next evaluation and confirm the validity of the lightness model.

3.2 Second experiment: chrominance response function

In order to validate our HMDCAM proposed above, we conducted a second experiment that includes the evaluation of the chrominance perception. This design is inspired by Whittle's experiment on luminance discrimination [21]. Instead of measuring the difference between a stimulus and a background, we compare two stimuli over a uniform grey background. We adapted this evaluation to determine JNDs of luminance and saturation for a set of different stimuli (see Figure 5). The background L_b is an achromatic luminance in the range of the display. V_r is the chromatic or achromatic reference stimulus. Finally, ΔV is the difference between the reference and the test stimulus. ΔV can be either a difference in luminance L, in chroma C or in hue H. In this study, both the

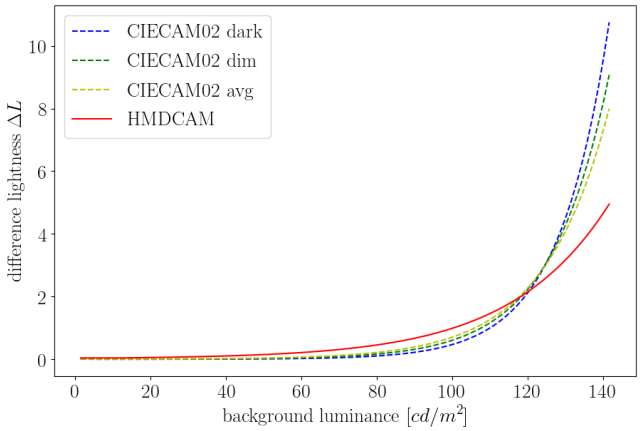


Fig. 4: Proposed HMDCAM compared with the three viewing conditions of CIECAM02.

(a) (b)

Fig. 5: (a) Schematic experiment with the background luminance L_b, the reference stimulus V_r and the difference ΔV between the two stimuli. (b) Schematic values represented on the HMD (binocular).

background and the reference stimulus are fixed while the test stimulus increases in luminance or saturation until participants notice a difference between the two stimuli. The two stimuli have a radius of 2° each and are separated by 8° of vision field. The background still covers the entire FoV of the HMD (100°). The panel consisted of 18 participants, from 20 to 57 years of age, with normal vision and various socio-cultural background. The test lasted about 45 minutes and was split into two parts. First, 8 luminance values are evaluated for 8 different backgrounds, that gives 32 discrimination luminance values. Then saturation has been evaluated for a unique background equal to 20% of the maximum luminance of the HMD (about 30 cd/m^2). The second part of the experiment consists of 8 saturation values for 4 colors: red, yellow, green and blue (32 values in total). The aim of this experiment is to confirm the HMDCAM that has been proposed after the first subjective study. To that end, for the evaluation of the lightness, we compute the CIECAM02 lightness response of both stimuli: J_{ref} the reference, and J_{test} the test average over all values determined by participants.

$$J_{ref} = 100\left(\frac{A_{ref}}{A_w}\right)^{c.z}, \quad (13) \qquad J_{test} = 100\left(\frac{A_{test}}{A_w}\right)^{c.z}, \quad (14)$$

where A_{ref} and A_{test} are respectively the achromatic response of the reference stimulus and the achromatic response of the test stimulus. c is defined according to the Average surround lighting condition. Then, we compute the absolute difference between the two lightness responses:

$$\Delta J = |J_{ref} - J_{test}| \quad (15)$$

As the test value has been determined as the JND between itself and the reference stimulus, ΔJ should be almost constant for any of the 32 conditions. We compute ΔJ for all the tested values (in case of lightness evaluation) using either CIECAM02 or HMDCAM and display resulting curves on Figure 6. We finally

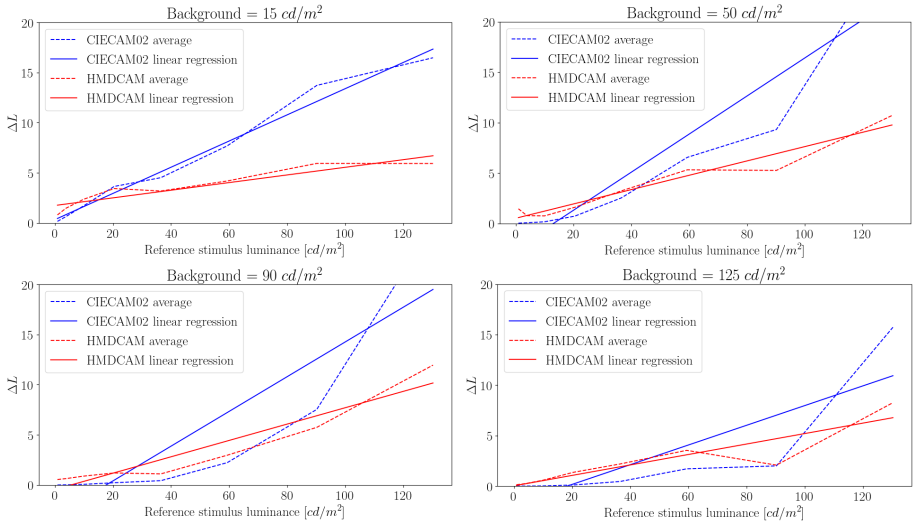


Fig. 6: CIECAM02 compared with our HMDCAM model for the collected subjective data on lightness perception.

compute a linear regression over the two curves in each of the four background luminance condition. We clearly see that the CIECAM02 (blue lines) does not fit with the collected data. The more the stimuli luminance increases, the more the difference ΔJ of perceived lightness is significant. Regarding our proposed CAM modeled for HMD (red lines), the ΔJ is almost constant (about 6% of error in average). We computed the error of estimated lightness for both CIECAM02 and HMDCAM in all the tested conditions and found a clear improvement with our model (see Table 2).

Table 2: Error of estimated perception for lightness and saturation (CIECAM02 [5] compared with our proposed HMDCAM).

Background luminance $[cd/m^2]$	15	50	90	125
CIECAM02 lightness error [%]	13.1	18.8	17.3	9.7
HMDCAM lightness error [%]	3.8	7.1	8.2	5.2
Color	Red	Green	Blue	Yellow
CIECAM02 saturation error [%]	1.3	0.6	3.2	5.3
HMDCAM saturation error [%]	0.7	0.5	1.7	2.8

We reproduced exactly the same protocol for the evaluation of perceived saturation. Δs is the absolute difference between the perceived saturation of the reference stimulus and the test stimulus. The proposed HMDCAM slightly improves the results of the perception of saturation (see Table 2). Nevertheless, the error does not differ so much between classic viewing condition and on a HMD because the error of CIECAM02 stays low. To sum up, we have seen that the perception on HMD differs from classical perception on 2D display. Our two experiments showed that the perception of a difference between two levels of luminance is halved. We then proposed a HMDCAM that better describes the perception on HMD. Further experiments could lead to a more accurate HMDCAM as the current one does not fit perfectly with the data; but we can afford to rely on it for our purpose. In the next section, we present our TMO that uses the HMDCAM to improve the quality of the tone mapped images.

4 A new TMO for HMD

In this section we present the HMD-TMO proposed in [8] and we bring some major revisions that improve significantly the global framework. The framework is presented in Figure 7. Red blocks indicate the proposed improvements. Recall that the input is a 360° HDR image and the output is a tone mapped image of the current viewport visualized on an HMD. The upper branch performs a tone mapping on the entire 360° image and thus preserves the spatial coherency. This operator is based on the log-luminance histogram of the image. We will see that computing a naive histogram of the equirectangular projection of the 360° image leads to an unrepresentative distribution of the luminance. We propose a correction that improves the result of the global TMO. Concurrently, the lower branch performs a tone mapping on the viewport image to enhance the contrast. This operator is based on the Photographic Tone Mapping Operator [17]. It has not been changed compared with the original method [8] as it produces the expected result. In our original version [8], the combination of the resulting luminances of these two TMOs was based on a geometric mean. We relax this combination and propose to use a weighted sum in logarithmic domain. The weight factor α is in the range (0,1) and gives more emphasize on the global or the viewport luminance. The geometric mean is reached when $\alpha = 0.5$. Finally, the resulting tone mapped viewport is colorized using the classic desaturated color method

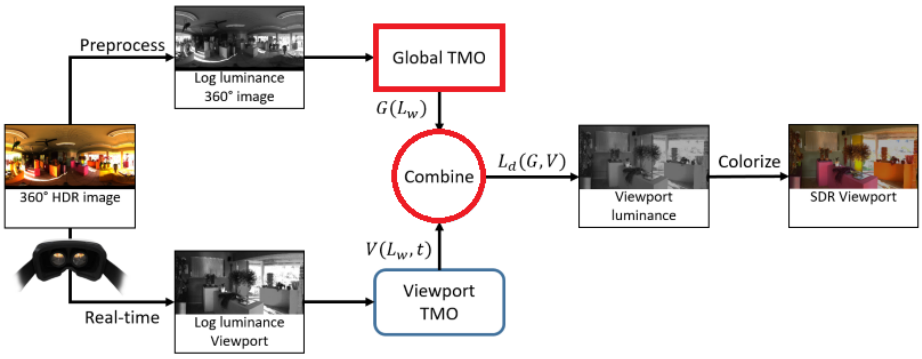


Fig. 7: Our operator combines a Global TMO $G(L_w)$ and a Viewport TMO $V(L_w, t)$. The Global TMO (upper branch) preserves the global coherency of the scene while the Viewport TMO (lower branch) enhances contrast. The combination of both produces our final HMD-TMO $L_d(G, V)$.

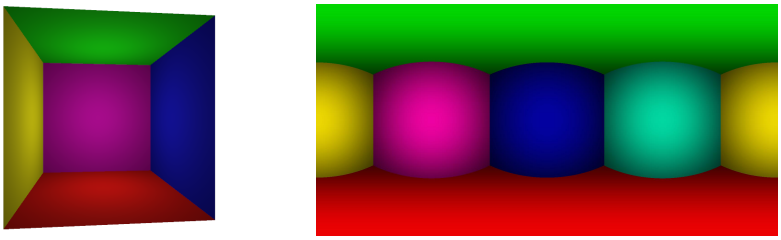
proposed by Schlick [18]. We detail the three operations with the improvements in the following section (global TMO, viewport TMO and combination).

4.1 Global TMO

The global TMO is based on the Visibility Matching Tone Reproduction Operator proposed by Ward *et al.* [9]. It consists of a log-luminance histogram equalization scaled into the display dynamic range. To avoid artifacts due to a too high contrast in the tone mapped image, the authors add a pass of histogram adjustment that matches with the HVS luminance response. We adapted this step using the perception model on HMD. Thus, the log-luminance distribution is needed to compute the TMO, but a naive histogram of the equirectangular projection of the 360° image results in a wrong distribution. Actually, the projection gives more significance to the poles (top and bottom) of the 360° image than to the equatorial area as illustrated by Figure 8. To avoid this over-represented contribution in the histogram, we apply a weight to the pixels depending on the elevation in the equirectangular image to obtain a right distribution [4, 20]:

$$w_{x,y} = cos\left(\pi \times \left(\frac{y}{H} - 0.5\right)\right), \tag{16}$$

where $w_{x,y}$ is the weight of the pixel (x, y) (instead of 1), and H is the image height in number of pixels. The histogram is computed in floating numbers, it is then cumulative and normalized to give the Cumulative Distribution Function (CDF). This correction is especially needed given that, in general cases, the pod of the 360° camera that captures the HDR image lets a black area in the bottom, which produces an offset in black level as illustrated in Figure 9. Finally,



(a) (b)

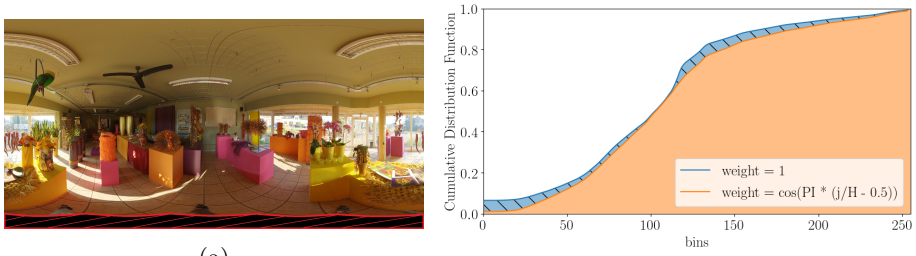
Fig. 8: (a) Inside of a cube, each wall has the same dimension. (b) 360° image (equirectangular projection), the poles (green and red areas) take half of the image.

the weighted log-luminance CDF is given by

$$P(b) = \frac{\sum_{b_i<b} f(b_i)}{\sum_{b_i} f(b_i)}, \tag{17}$$

$$f(b_i) = \sum_{x,y} w_{x,y} \times log\big(L_w(x,y)\big), \tag{18}$$

where $f(b_i)$ is the log-luminance weighted sum of all pixels (x, y) that fall into bin b_i. The number of bins is at least equal to 100 as proposed by the authors to avoid banding artifacts due to quantization. The tone curve G proposed by



(a) (b)

Fig. 9: (a) Equirectangular projection of Florist 360° image. The camera pod (hatched in red) is over-represented. (b) Comparison of CDF without weights (blue curve), and with weights (orange curve). The offset in low luminances (bin 0) is generated by the camera pod.

Ward *et al.* is a scaled version of $P(b)$:

$$G(x,y) = exp\big(ln(L_{dmin}) + \big(ln(L_{dmax}) - ln(L_{dmin})\big) \times P\big(L_w(x,y)\big)\big), \tag{19}$$

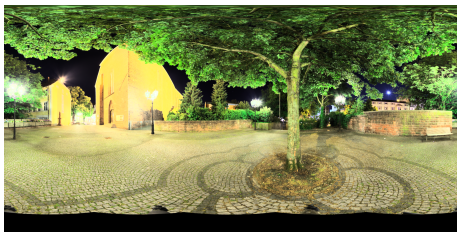
where L_{dmin} and L_{dmax} are respectively the minimum and maximum luminance of the display, $L_w(x, y)$ is the world luminance of the pixel (x, y). $P(L_w(x, y))$ is the CDF defined in Equation 17, and $G(x, y)$ the resulting luminance of the pixel. The last step is a pass of histogram adjustment. If the slope of the CDF is too steep, the contrast produced by the tone mapping is too high. In order to preserve a perceptually coherent contrast in the image, Ward *et al.* [9] proposed to trimming the histogram based on the human perception. If the contrast between two levels of displayed luminance is perceptually higher than it is with the corresponding levels of world luminance, then we have to trim the histogram. It can be written thus:

$$\frac{dL_d}{dL_w} \leq \frac{J_d}{J_w}, \tag{20}$$

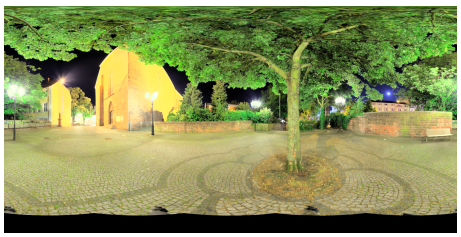
where L_d and L_w are respectively the display and the world luminance while J_d and J_w are the corresponding perceived lightness. We use the iterative process proposed by Ward *et al.* for the ceiling [9].

At the end, we obtain a CDF with a right distribution while the ceiling histogram is perceptually coherent with the HMDCAM. We remind that the human eye is less sensitive on HMD than on classic display (about halved sensitive as seen in section 3). This phenomenon is notable on tone curves in Figure 10, a gentler slope is allowed for luminances that are sparingly present when the ceiling follows the HMDCAM. While the CDF without any ceiling produces a too

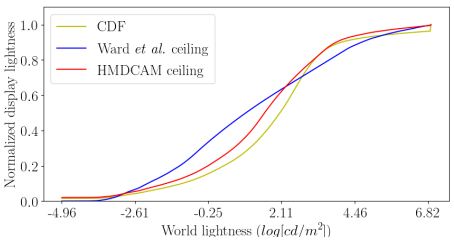
(a) Without ceiling (yellow curve).



(b) HMDCAM ceiling (red curve).



(c) Ward ceiling (blue curve).



(d) Tone curves comparison.

Fig. 10: (a) Histogram adjustment without ceiling, the contrast is too high. (b) With HMDCAM ceiling, the contrast is coherent with the perception on HMD. (c) With Ward ceiling, the contrast is too low. (d) Comparison of corresponding tone curves.

high contrast, trimming the histogram with a classic perception model flatten the image. The ceiling based on our HMDCAM better preserves contrast and stays perceptually coherent when visualized on HMD.

The global TMO preserves the coherency of the scene and is perceptually coherent on an HMD. However, the contrast in the viewport can be improved when not considering the entire 360° image. The viewport TMO enhances the contrast as explained in the following section.

4.2 Viewport TMO

The viewport TMO relies on the Photographic Tone Reproduction operator [17] with temporally smoothed parameters to avoid flickering and simulate eye adaptation as first proposed by Yu [22]. This TMO is based on the log-average luminance of the image:

$$\bar{L}_w\big(V(t)\big) = \frac{1}{N} exp\Big(\sum_{x,y} log\big(\delta + L_w(x,y)\big)\Big), \tag{21}$$

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:

$$\bar{L}_w^{'}(t) = \tau \bar{L}_w\big(V(t)\big) + (1-\tau)\bar{L}_w^{'}(t-1), \tag{22}$$

$$L_{white}^{'}(t) = \tau L_{white}\big(V(t)\big) + (1-\tau)L_{white}^{'}(t-1), \tag{23}$$

where $\bar{L}_w^{'}(t)$ and $L_{white}^{'}(t)$ are respectively the smoothed key and white values between two successive views and τ is a time dependent interpolation variable. The value of τ determines the adaptation time. For $\tau = 1$, there is no adaptation, while for $\tau = 0$, the luminance is never updated. Based on TMOs that use models of eye adaptation [10,14], we decided to fix τ value corresponding to one second of adaptation (for both light and dark): $\tau = \Delta t$ where Δt is the time spent between the previous and the current frame in second. Finally, the luminance is scaled and high values are attenuated to avoid clipping:

$$L(x,y,t) = \frac{a}{\bar{L}_w^{'}(t)} L_w(x,y), \tag{24}$$

$$V(x,y,t) = \frac{L(x,y,t)\left(1 + \frac{L(x,y,t)}{L_{white}^{'2}(t)}\right)}{1 + L(x,y,t)}, \tag{25}$$

where a is a user defined variable which scales the luminance (commonly 0.18) and $L(x,y,t)$ the time dependent scaled luminance. Our Viewport TMO is the

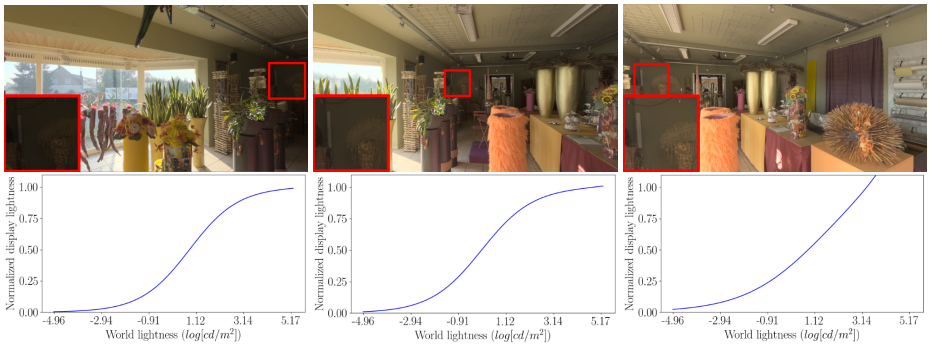


Fig. 11: Photographic Tone Reproduction operator [17] applied to a viewport sequence with smooth transitions. As the key value and the white value evolve from a view to another, the tone curve is modified and a same zone in the scene (red inset) becomes brighter or darker.

displayed luminance $V\big(L_w(x,y),t\big)$ as illustrated in Figure 11. In his operator, Yu actually uses equation (24) 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 ($V(L_w,t)$) that we want to combine to obtain our final tone mapped image.

4.3 TMOs combination

Combination of global and viewport luminances ensures the global coherency to be preserved and the viewport contrast to be enhanced. To obtain a resulting perceptually uniform luminance, the combination has to be done in the perceptual domain (the logarithm is a good representation as seen in section 3):

$$L_d(x,y,t) = e^{\alpha.ln\left(G(x,y)\right)+(1-\alpha).ln\left(V(x,y,t)\right)}, \tag{26}$$

$$L_d(x,y,t) = e^{\alpha.ln\left(G(x,y)\right)} \times e^{(1-\alpha).ln\left(V(x,y,t)\right)}, \tag{27}$$

$$L_d(x,y,t) = \left(e^{ln\left(G(x,y)\right)}\right)^{\alpha} \times \left(e^{ln\left(V(x,y,t)\right)}\right)^{1-\alpha}, \tag{28}$$

$$L_d(x,y,t) = G(x,y)^{\alpha} \times V(x,y,t)^{1-\alpha}, \tag{29}$$

where $G(x,y)$ and $V(x,y,t)$ are our global and viewport TMOs respectively, α is the weight in range $(0,1)$ that gives more emphasize on the global or the viewport result, and finally L_d is the display luminance. The effect of the α value is showed by Figure 12. As expected, the viewport TMO (left) enhances the contrast by exploiting all its dynamic range. Contrarily, the global TMO (right) brightens the image because this area is bright compared with the rest

Fig. 12: α varies linearly from 0 to 1, given more emphasize on the viewport TMO (left), then more and more on the global TMO (right).

of the 360° image. The value of α can change depending on the processed scene. More results are shown in the following section.

4.4 Color saturation

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 [18]:

$$C^{'} = \left(\frac{C}{L_w}\right)^{s} L_d, \tag{30}$$

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.

5 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 a 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 for the binocular vision) 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). The efficiency of our TMO is clearly showed in the example presented in Figure 13. The two images on the left (13a) result from the viewport TMO. A little movement of the camera produces a significant change in displayed luminance, especially for the blue strip behind "Student Service". The middle images (13b) result from the global TMO. The spatial coherency is preserved, the blue color stays the same. Nevertheless, some details are lost due to clipping in the right of top image. The two lasts images (13c) result from the linear combination of both TMOs. The spatial coherency and details are well preserved, the blue strip does not change significantly and light reflects on the wall are visible. Figure 14 presents some additional results of our HMD-TMO. In all of our results, we fixed $\alpha = 0.5$.

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

(a) (b) (c)

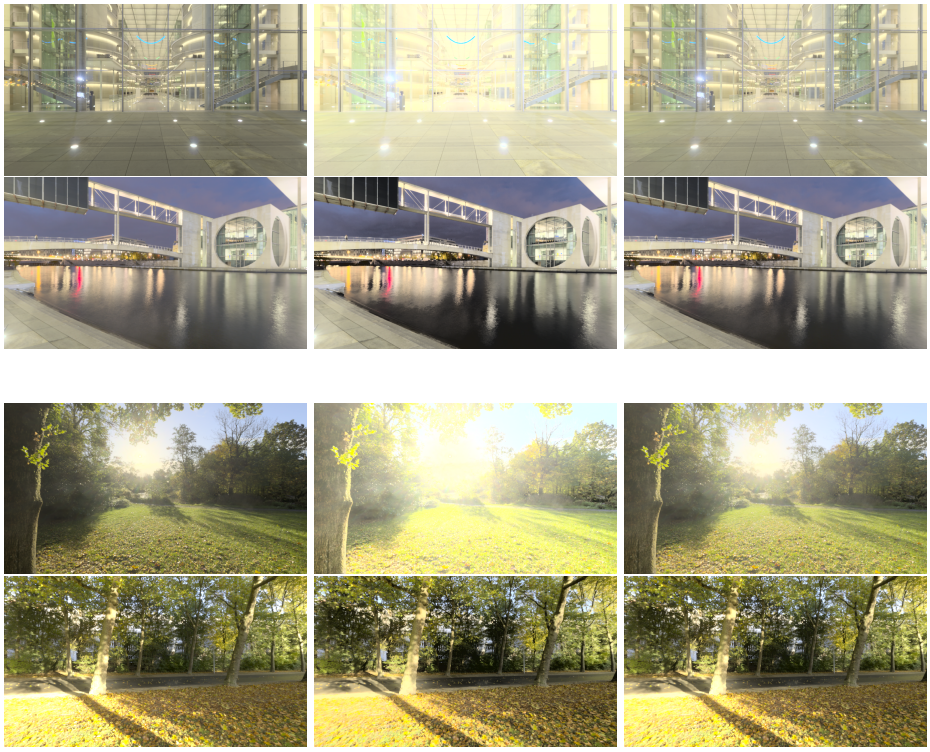
Fig. 13: (a) Viewport TMO: The blue strip change significantly. (b) Global TMO: Some details are lost due to clipping. (c) TMOs combination: the spatial coherency and the details are preserved.

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 CAM that describes well the perception on HMD; (2) a perceptually coherent TMO that combines both global and viewport TMOs. The linear combination allows the TMO to be adaptive depending on the encountered scene. This new TMO does not 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. An automatic tuning of the α parameter used in the combination could tackle those limits.

Acknowledgments

All 360° HDR images come from free SYNS, LizardQ and HDRI Haven 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 reviewers for their constructive commentaries. Thanks to all experiment participants for their contributions.



(a) (b) (c)

Fig. 14: (a) Viewport TMO: the local contrast is enhanced. (b) Global TMO: the coherency is preserved. (c) TMOs combination: the spatial coherency is preserved while the local contrast is enhanced.

References

1. Banterle, F., Artusi, A., Debattista, K., Chalmers, A.: Advanced high dynamic range imaging. AK Peters/CRC Press (2017)
2. Brill, M.H., Carter, R.C.: Does lightness obey a log or a power law? or is that the right question? Color Research & Application **39**(1), 99–101 (2014)
3. Cutchin, S., Li, Y.: View dependent tone mapping of hdr panoramas for head mounted displays. In: Proceedings of the 26th International Conference on Artificial Reality and Telexistence and the 21st Eurographics Symposium on Virtual Environments. pp. 29–36. Eurographics Association (2016)
4. De Abreu, A., Ozcinar, C., Smolic, A.: Look around you: Saliency maps for omnidirectional images in vr applications. In: 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX). pp. 1–6. IEEE (2017)
5. Fairchild, M.D.: Color appearance models. John Wiley & Sons (2013)
6. Fechner, G.T., Howes, D.H., Boring, E.G.: Elements of psychophysics, vol. 1. Holt, Rinehart and Winston New York (1966)

7. Gloriani, A.H., Matesanz, B.M., Barrionuevo, P.A., Arranz, I., Issolio, L., Mar, S., Aparicio, J.A.: Influence of background size, luminance and eccentricity on different adaptation mechanisms. Vision research **125**, 12–22 (2016)
8. Goudé, I., Cozot, R., Banterle, F.: Hmd-tmo: A tone mapping operator for 360° hdr images visualization for head mounted displays. In: Computer Graphics International Conference. pp. 216–227. Springer (2019)
9. Larson, G.W., Rushmeier, H., Piatko, C.: A visibility matching tone reproduction operator for high dynamic range scenes. IEEE Transactions on Visualization and Computer Graphics **3**(4), 291–306 (1997)
10. Ledda, P., Santos, L.P., Chalmers, A.: 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. pp. 151–160. ACM (2004)
11. Lee, S.H., Sohng, K.I.: A model of luminance-adaptation for quantifying brightness in mixed visual adapting conditions. IEICE transactions on electronics **94**(11), 1768–1772 (2011)
12. Li, C., Li, Z., Wang, Z., Xu, Y., Luo, M.R., Cui, G., Melgosa, M., Brill, M.H., Pointer, M.: Comprehensive color solutions: Cam16, cat16, and cam16-ucs. Color Research & Application **42**(6), 703–718 (2017)
13. Melo, M., Bouatouch, K., Bessa, M., Coelho, H., Cozot, R., Chalmers, A.: Tone mapping hdr panoramas for viewing in head mounted displays. In: VISIGRAPP (1: GRAPP). pp. 232–239 (2018)
14. Pattanaik, S.N., Tumblin, J., Yee, H., Greenberg, D.P.: Time-dependent visual adaptation for fast realistic image display. In: Proceedings of the 27th annual conference on Computer graphics and interactive techniques. pp. 47–54. ACM Press/Addison-Wesley Publishing Co. (2000)
15. Perrin, A.F., Bist, C., Cozot, R., Ebrahimi, T.: Measuring quality of omnidirectional high dynamic range content. In: Applications of Digital Image Processing XL. vol. 10396, p. 1039613. International Society for Optics and Photonics (2017)
16. Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., Myszkowski, K.: High dynamic range imaging: acquisition, display, and image-based lighting. Morgan Kaufmann (2010)
17. Reinhard, E., Stark, M., Shirley, P., Ferwerda, J.: Photographic tone reproduction for digital images. In: ACM transactions on graphics (TOG). vol. 21, pp. 267–276. ACM (2002)
18. Schlick, C.: Quantization techniques for visualization of high dynamic range pictures. In: Photorealistic Rendering Techniques, pp. 7–20. Springer (1995)
19. Stevens, S.S.: To honor fechner and repeal his law. Science **133**(3446), 80–86 (1961)
20. Upenik, E., Ebrahimi, T.: A simple method to obtain visual attention data in head mounted virtual reality. In: 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). pp. 73–78. IEEE (2017)
21. Whittle, P.: Increments and decrements: luminance discrimination. Vision research **26**(10), 1677–1691 (1986)
22. Yu, M.: Dynamic tone mapping with head-mounted displays. Standford University Report **5** (2015)