DOI: 10.1111/cgf.12606

COMPUTER GRAPHICS *forum*
Volume 00 (2015), number 0 pp. 1–13

Evaluation of Tone-Mapping Operators for HDR Video Under Different Ambient Luminance Levels

M. Melo[1,2], M. Bessa[1,2], K. Debattista[3] and A. Chalmers[3,4]

[1]Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal
emekapa@sapo.pt, maxbessa@utad.pt
[2]INESC TEC, Porto, Portugal
[3]WMG, University of Warwick, Coventry, United Kingdom
{k.debattista, alan.chalmers}@warwick.ac.uk
[4]goHDR Ltd., Coventry, United Kingdom

Abstract
Since high dynamic range (HDR) displays are not yet widely available, there is still a need to perform a dynamic range reduction of HDR content to reproduce it properly on standard dynamic range (SDR) displays. The most common techniques for performing this reduction are termed tone-mapping operators (TMOs). Although mobile devices are becoming widespread, methods for displaying HDR content on these SDR screens are still very much in their infancy. While several studies have been conducted to evaluate TMOs, few have been done with a goal of testing small screen displays (SSDs), common on mobile devices. This paper presents an evaluation of six state-of-the-art HDR video TMOs. The experiments considered three different levels of ambient luminance under which 180 participants were asked to rank the TMOs for seven tone-mapped HDR video sequences. A comparison was conducted between tone-mapped HDR video footage shown on an SSD and on a large screen SDR display using an HDR display as reference. The results show that there are differences between the performance of the TMOs under different ambient lighting levels and the TMOs that perform well on traditional large screen displays also perform well on SSDs at the same given luminance level.

Keywords: image and video processing, high dynamic range/tone mapping

ACM CCS: I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithms, I.4.0 [Computer Graphics]: General—Image displays

1. Introduction

Conventional imagery, known as standard dynamic range (SDR) [or as low dynamic range (LDR)], cannot properly represent the colour gamut and contrast present in real scenes in a way that matches the human visual system (HVS)'s dynamic range. High dynamic range (HDR) imaging was developed to overcome the limitations of SDR with the aim to capture, store, transmit and deliver real-world lighting. By capturing a dynamic range equivalent, or indeed higher, than that of the HVS, HDR brings significant benefits to many applications such as security, entertainment, art, scientific research and health.

While it is possible to deliver HDR content in a relatively straightforward manner on HDR displays [SHS*04], the majority of displays currently available are SDR and thus it is necessary to reduce the content's dynamic range to match that of the targeted display. This is achieved using methods known as tone-mapping operators (TMOs). These can take into account the scene characteristics or the HVS properties in order to provide the best viewing experience from the available HDR data. Although many TMOs have been developed, only a few are designed for HDR video. Since many TMOs have been proposed, the need to evaluate and identify which ones perform best under given conditions in order to successfully show HDR content on SDR displays has been frequently explored [BADC11]. However, most of the evaluations addressed still images and conventional sized displays.

The popularity of mobile devices is growing daily and as a result mobile-cellular penetration is reaching nearly 100%. This means that, on average, there will soon be one mobile device per person [ITU11]. Mobile devices are already being widely used to access multi-media content and it is estimated that around 51% of the traffic on mobile devices is video [CIS13]. Mobile devices are typically more limited than conventional devices and thus other variables need to be considered when thinking about content distribution on them: they can be exposed suddenly to different ambient lighting conditions very easily, the viewing angle and distance are more limited and, obviously, the displays are smaller than traditional desktop displays.

Some previous work has evaluated HDR video on SSDs. The first of these was conducted by Urbano *et al.* [UMM*10] that investigated still HDR images only. Melo *et al.* [MBDC13b, MBDC13a] evaluated HDR video tone mapping and the results and data used for that research form part of this work as will be shortly discussed. More recently, Akyüz *et al.* [AEA13] investigated how TMOs and exposure fusion perform on SSDs.

This paper addresses a number of the concerns with mobile devices in the context of tone mapping for displaying HDR video. In particular, consideration is given to whether the TMOs that are successful for traditional displays also work for SSDs and if different lighting conditions affect the performance of the TMOs. The ability to use the right mobile tone mapper, under diverse viewing conditions, is important in, for example, the content production field, where captured HDR content may need to be immediately displayed on a mobile device under widely varying luminance scenarios from harsh outdoor sunlight to dark indoor settings. The results from the related work are divided and not yet conclusive. Urbano *et al.* [UMM*10] and Akyüz *et al.* [AEA13] found differences across displays while Melo *et al.* [MBDC13a] did not; however these experiments were run under different conditions. Urbano *et al.* used still images, Akyüz *et al.* videos but with a relatively small camera display, while and Melo *et al.* used a larger SSD, identical to the one used in this study. One of the goals of this work serves to further investigate the possibilities of differences in HDR video tone mapping across displays and, to clarify, whether different levels of lighting have an impact on TMO accuracy and to what extent. Knowing if there are certain TMOs that perform better under specific circumstances can be also important to further study their battery usage and optimize the impact of visualization of HDR video on mobile devices similarly to the method proposed by Wanat and Mantiuk [WM14] since battery-dependency is a clear limitation of mobile devices when compared to conventional displays. Also, SSDs are increasing in size; in this work one of the larger size SSDs was targeted as a representative as many variables are being investigated. Future work will consider the impact of various SSD sizes.

The work presented here add significantly to Melo *et al.* [MBDC13b, MBDC13a] while reusing some of the captured data. The previous results are included as just one of the luminance conditions for this work. The experiments in this paper provide novel contributions as they are concerned with different viewing conditions, namely, with how the environmental lighting may affect results of tone mapping on mobile devices, which are frequently exposed to differing lighting conditions due to their portable nature. More specifically, the study is conducted across luminance levels and provides new insights into the previous data through the cross luminance analysis, particularly that the TMO rankings change according to the luminance levels of the surrounding environment for both SSDs and SDR displays.

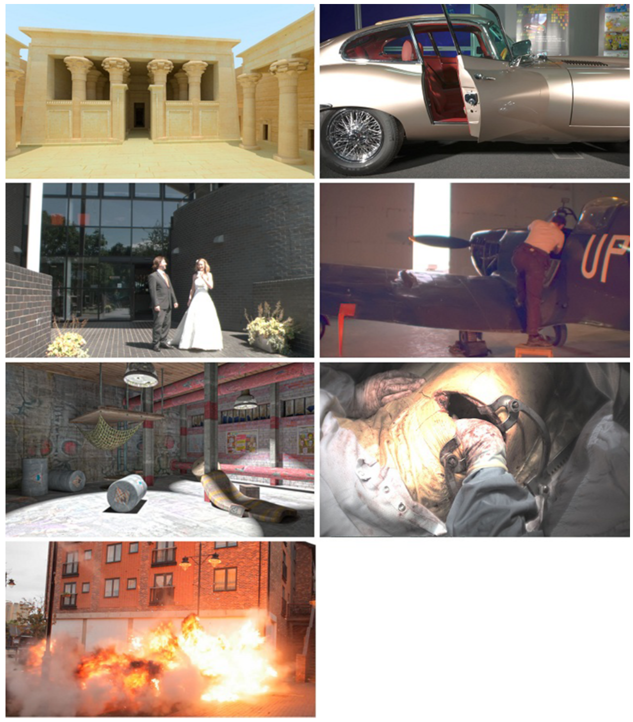

Figure 1: *Frames of the seven HDR videos which were used for the evaluations.*

HDR video tone mappers for SDR and SSD are evaluated under three different scenarios that differ in the ambient luminance: one dark, one medium and one bright scenario. Six HDR TMOs ([BALCH09, PTYG00, VH06, MDK08, FPSG96 and BBC*12]) applied to seven different videos (Figure 1) are compared. This evaluation intends to determine: if the participants rank equally the accuracy of the TMOs for the large screen displays and SSDs; if the ambient luminance has an influence on the accuracy rankings of the TMO; and, whether there is one TMO that performs best for SSDs.

2. Related Work

Various TMOs have been proposed which each take into consideration a number of factors, for example, aspects of the HVS, or the features of the displays in question, but none of the TMOs proposed so far has been specifically developed for SSDs.

TMOs can be broadly divided into two categories: global and local. The main difference is that the global TMOs are spatially invariant and, as such, they map all the pixels of an image evenly. Local TMOs, on the other hand, typically apply a different operation for each pixel, taking into consideration the surrounding pixels of the pixel being processed.

In addition to these categories, TMOs can be time-dependent or time-independent. Time-dependent TMOs consider a determined sequence of frames (adjacent frames or overall frames) when tone mapping a frame. A number of TMOs have been developed with HDR video in mind. TMOs such as [BALCH09, MDK08, BBC*12, ASC*14] take into consideration temporal coherency since significant changes of luminance can occur in HDR videos and failure to take this into account could lead to noticeable flickering artefacts. The TMOs chosen for this study are discussed in Section 2.2.

2.1. TMO evaluation

There have been a number of evaluations of the large number of TMOs that have been proposed over the years. Two main methodologies for evaluations have been proposed: error metrics and psychophysical experiments.

Error metrics are objective methods that computationally compare images. The comparison can be made based on differences in the physical quantities of the images or by attempting to simulate the HVS in order to identify which aspects of the image would be perceived by the HVS as being different [BADC11]. The major drawback of this approach is that even the most complex metrics cannot yet fully simulate the complexity of the HVS, although some studies [BTDH10] are making progress in understanding this complexity. Examples of error metrics used to evaluate HDR images are HDR-VDP (visual difference predictor) that determines the visible differences in HDR images [MMS04, MKRH11], and the DRIVQM (dynamic range independent video quality metric) assessment that predicts the perceived quality of video sequences compared to a reference [AvMS10].

Psychophysical experiments are subjective and based on user studies conducted with human participants. These experiments are conducted in controlled environments and can make use of a number of evaluation methods for comparing images and videos such as rating, pairwise comparison or ranking. The experiments can include a reference to guide the participants' choices. For evaluating TMOs real-world scenes or HDR displays are frequently used as a reference. Several psychophysical experiments have previously been conducted, including [DMMS02, LCTS05, KYL*07, ČWNA08, UMM*10, EUWM13, AEA13, NDSLCP14]. We adopt a similar approach to these in this paper.

Drago *et al.* [DMMS02] was one of the first to subjectively evaluate TMOs. In this evaluation, four different scenes were used and seven TMOs were compared by 11 participants by pairwise comparison on a CRT (Cathode Ray Tube) display without an HDR reference.

Ledda *et al.* [LCTS05] conducted the first experiment that compared TMOs using an HDR display as a reference. In these experiments the authors evaluated six TMOs applied to 23 images that were evaluated by 18 participants. The HDR display was placed in the middle of two SDR displays. On each of the SDR displays a tone-mapped image was shown and the participants were asked to perform a pairwise comparison with the HDR image shown on the HDR display in the middle as the reference.

Kuang *et al.* [KYL*07] studied the preference and accuracy of TMOs. Their work involved 33 participants and considered six TMOs that were divided into three experiments: paired comparison, ranking scale and a ranking scale to compare the real-world scene with the tone-mapped image. Another example of experiments using real-world scenes as reference are the ones conducted by Čadik et al. [ČWNA08] that evaluated 14 TMOs in three different scenes.

More recently, Eilertsen *et al.* [EUWM13] undertook an evaluation of TMOs for HDR video where 11 TMOs were analysed and evaluated using both camera-captured and computer-generated videos. This work analysed the response of TMOs to temporal variations as well as dividing TMOs in two categories (visual system simulators and best subjective quality). The evaluation was run with 36 participants who were asked to conduct pairwise comparisons between the TMOs and the HDR content. The results showed that many TMOs introduce artefacts such as flickering, ghosting or oversaturated colours and that less sophisticated global operators could outperform more recent and complex TMOs.

Narwaria *et al.* [NDSLCP14] conducted a subjective study to evaluate if extra details in tone-mapped contents affect user preferences over single exposure content. Seventeen images were evaluated by 38 participants using a pairwise comparison method. The study revealed that overall there was no statistical evidence for users' preference between tone mapped and single exposure images.

A few studies have considered mobile devices but none of them have considered ambient lighting levels as an independent variable. Urbano *et al.* [UMM*10] was one of the first that was aimed specifically at SSDs. Their experiments evaluated several TMOs on different size displays using a pairwise comparison of tone-mapped images with a real scene reference. Three different displays were used: two 17” displays, and one 2.8” display with a resolution of 1024×682 and 240×320, respectively. The mobile devices were placed in a fixed position on a table in order to maintain the same viewing distance and viewing angle among participants. The authors concluded that TMO preference order between displays was different and that for mobile devices content that offered stronger detail reproduction, more saturated colours and overall brighter image appearance were preferred.

Another recent study aimed at SSDs was conducted by Akyüz *et al.* [AEA13]. This evaluation took into account two groups of algorithms that display HDR content on SDR displays: TMOs and EFAs (exposure fusion algorithms). EFAs differ from TMOs since they take as input a set of SDR images with different exposures instead of having to generate HDR content as a first step. The algorithms merge the best exposed pixels from these SDR exposures into an SDR image that is rich in detail. The evaluation consisted of two pairwise comparison experiments. For the first experiment, three HDR images of real scenes (one with reference and two without reference) were evaluated on a 24” display by 15 participants judging on the criteria of colour, contrast and detail. The second experiment involved a 3” display and one real-world scene and one computer-generated scene. Pairwise comparisons based on the similarity criteria were conducted using the real scene as reference. The authors concluded that the TMOs outperformed the EFAs although the differences between the algorithms on SSDs were barely perceptible and thus more straightforward methods can be used to reproduce HDR images on SSDs. However, since the display size was relatively small for the chosen SSD, the results cannot be generalized to all mobile devices.

Earlier work on the evaluation of HDR video tone mapping for mobile devices for a fixed luminance condition was conducted by Melo *et al.* [MBDC13b]. This study used an HDR display as reference to compare TMOs between an SSD and a 37” SDR display. The results demonstrated that there was a statistically significant difference between the choice of TMOs between the SSD and the large screen display; however, this was subtle and the ordering accuracy of the TMOs remained constant across the two displays. The data used for Melo *et al.* [MBDC13a] forms the medium luminance group for this current paper, which extends the previous work by adding and analysing the data with two different sets of luminance and increasing the number of participants from 60 to 180.

2.2. TMOs considered

All the TMOs used for the experiments are global TMOs. Since the evaluation addresses HDR video, all the considered TMOs are time-dependent. The following are the TMOs that are chosen for the study in this paper:

Benoit (Ben): The spatiotemporal TMO [BALCH09] is based on a model of the retina local adaptation properties developed by Meylan *et al.* [MAS07] and is complemented by spatiotemporal filters of the retina. Some of the foveal retina functionalities are simulated and temporal coherency included to avoid flicker.

Boitard (Boi): Boitard *et al.* [BBC*12] worked on temporal coherency for video tone mapping and developed mechanisms that allow preserving the overall contrast of the video. The key characteristics are perception consistency of an object throughout the video and the preservation of the overall temporal contrast consistency. This operator has two main steps: it first processes each frame of the video individually (that can utilize arbitrary static operators) and then considers the luminance of each frame taking into account the lighting within the whole HDR video.

Ferwerda (Fer): This TMO is based on a model of visual adaptation from psychophysical experiments and includes some aspects of the HVS such as visibility, visual acuity and colour appearance. This operator uses TVI functions for modelling photopic and scotopic vision. The mesotopic range is achieved by a linear combination of both the photopic and scotopic [FPSG96].

Hateren (Hat): This TMO [VH06] uses a model of human cones and takes advantage of their dynamic response characteristics. To achieve the dynamic range reduction, the dynamic non-linearities are first combined and then noise reduced through a low-pass filtering that adapts to the scene luminance.

Mantiuk (Man): For this TMO [MDK08], the visible contrast distortions are minimized for a wide range of devices based on a model of the HVS. The TMO considers the environment luminance levels and display characteristics, such as the peak luminance of the display or the reflectivity of a screen. These characteristics are used to calibrate the tone-mapping process in order to optimize the results for situations with different variables. This TMO can be beneficial for mobile devices since they are used in changing environmental conditions.

Pattanaik (Pat): The HVS does not adapt instantly to big changes of luminance intensities. This is taken into account by this time-dependent visual adaptation for realistic image display TMO proposed by Pattannaik *et al.* [PTYG00]. In this method, the appearance changes to match the user’s visual responses so he/she can experience the view of a scene as it would appear in reality.

3. Experimental Framework

This section presents a user study to evaluate the performance of HDR video tone-mapped by the six TMOs discussed in the previous section.

3.1. Method

The six TMOs were evaluated for seven video sequences on a traditional SDR display and on an SDR SSD. An HDR display was used as reference. All the evaluations took place in a room in which all the environmental variables could be controlled. The evaluations took place under three different luminance levels: dark, medium and bright environment.

The experiments conducted used a ranking method. Ranking was preferred over rating as rating could result in a narrow distribution and these could vary considerably with the participant response style. Pairwise comparison was also not chosen since it would require significantly more time due to the large number of comparisons required to compare all methods across all scenes. Ranking was thus chosen since it provided clear and timely results for the accuracy rankings of the TMOs. Ranking also guarantees that each ranked item has a unique value through all tone-mapped HDR video footage for each scene.

There were four independent variables: the six considered TMOs, the displays, the luminance levels of the room and the scene groups. In total, 12 groups of participants were used. Display, luminance level of the experiments’ room and scene groups were in-between-participant independent variables; the choice of TMOs was a within-participant variable. The display independent variable was composed of the participants that viewed the sequences on a 37” SDR display, henceforth referred to as SDR, and those that ran the experiment on an iPad representing an SDR SSD, henceforth referred to as SSD.

The luminance level–independent variable consisted of three conditions: dark, medium and bright. Dark represented an indoor environment with little lighting, medium a traditional office environment ambient lighting and bright an outdoor environment on a cloudy day. All the scenarios were conducted in an experimental room where the environment variables could be controlled. There were two groups evaluating the same display, one viewed four scenes and the other group viewed a different set composed of three scenes. This was used to control for participant fatigue and also to analyse whether the results were consistent across scenes. All the TMOs were viewed and ranked by all the participants.

Since the video sequences shown during the experiments were played at the same time on the SDR/SSD display and on the HDR display, this could lead to a bias. To avoid the possibility of having

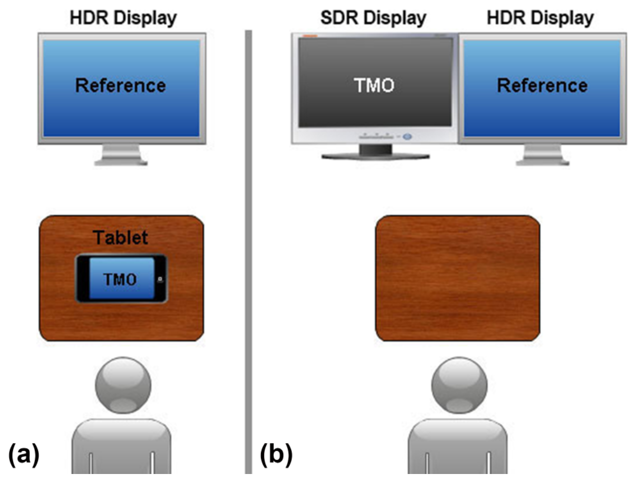

Figure 2: *Experimental setup.*

that bias towards giving preference to TMOs that generate brighter footage since more luminous content is known to be preferred [AFR*07] the question asked to the participants was ‘which video sequence is closest to the reference?’ rather than ‘which video sequence do you prefer?’

A key concern during this study was to minimize possible issues related with bias and luminance adaptation of the human eye especially as we were working with displays that emit different luminance levels. In particular, the SDR display and the HDR display have luminance values of 500 and 4000 cd/m^2, respectively. The possibility of having the reference HDR video played by itself and followed by the tone-mapped sequences was considered and may have helped in reducing possible adaptation issues. However, this would have required the participants to rank based on memory, introducing a recall bias that direct comparison controls for. In order to mitigate for potential maladaptation, a grey screen was introduced between each scene so that the participant could adapt to the ambient lighting conditions of the experimental room.

3.2. Materials

Figure 2 illustrates the configuration of the room where the experiments were undertaken. In the SSD groups, a tablet was placed on a table (a), and for the SDR groups, the SDR display was placed at the left side of the HDR display (b).

The reference HDR display used was a 37” DR37-P from Brightside. The SDR display was a Westinghouse 37” LCD which has the same exact front panel as the HDR display, and the tablet an Apple iPad4. The technical specifications of the displays used are given in Table 1. By mobile devices, we mean mobile phones, phablets and tablets. As discussed earlier, the choice of SSD was on the higher end of the SSD size scale. A 9.7” display device was chosen due to its wide availability, and due to the resolution disparities between the SSD and the SDR and HDR displays a retargeting of the content was necessary. As a consequence, in order to maintain the aspect ratio, the content was not displayed in full screen on the SSD, but centred within a black frame.

Table 1: *Technical specifications of the displays used in the experiments.*

	HDR display	SDR display	Tablet
Brand	Brightside	Westinghouse	Apple
Model	DRP37-P	LVM-37w3	iPad 4
Size	37”	37”	9.7”
Resolution	1920 × 1080	1920 × 1080	2048 × 1536
Contrast ratio	200 000:1	1000:1	877:1
Max luminance (cd/m^2)	4 000	550	476
Min luminance (cd/m^2)	0	0.55	0.48
View angle (horizontal)	40°	176°	175°
View angle (vertical)	15°	176°	175°

The SDR and the HDR displays were placed side by side and the participants stood at approximately 2.0 m from the 37” displays at the table where the tablet was. The two 37” displays were placed at the same distance and height forming a slight angle between them in order to optimize the visualization. The default position of the SSD was on a stand forming an approximately 45° angle at approximately 45 cm from the participants but the participants were able to move the tablet freely to adjust it to a preferred position for viewing the tone-mapped videos.

A total of seven videos were evaluated [labelled ‘CGRoom’, ‘Jaguar’, ‘Kalabsha’, ‘Morgan Lovers’, ‘Explosion’, ‘IDL Wedding’ and ‘Medical’ (Figure 3 shows the example of three of the videos used)]. The length in seconds, the Average Dynamic Range (Avg. DR) in log units, the capture method and the maximum F-Stops for each videos is shown in Table 2.

The experiments were conducted under three ambient luminance scenarios. In the dark scenarios, the experimental room had no light sources besides the light emitted from the displays, corresponding to a dark indoor environment; the luminance in the room was 15 cd/m^2. The second, medium luminance, scenario had only the regular lighting of the experimental room that was from the ceiling illumination at 55 cd/m^2, which corresponds to an environment such as regular family living room or office. The final, bright luminance, scenario was conducted by setting the lighting to 1450 cd/m^2 by using four photographic lights. This corresponds approximately to the equivalent outdoor luminance levels that were measured outside on a series of cloudy days in Portugal in February at noon; the same time period at which the experiments took place. For the illumination of the room, indirect light was used so there were no reflections on the displays. All the values were measured with the Sekonic L-758D DigitalMaster exposure meter.

No retargeting was done between the footage of the HDR and the LCD displays since the panel of the SDR is identical to that of the HDR display. To ensure a similar viewing experience of the HDR video content between the HDR display and the SSD, additional adjustments were required regarding the disparity of resolutions: the resolution of the images was maintained and the videos were centred in a black background. Since the tablet used had a retina display that has a resolution of 2048 × 1536, the videos kept the same resolution but they were centred in a black frame.

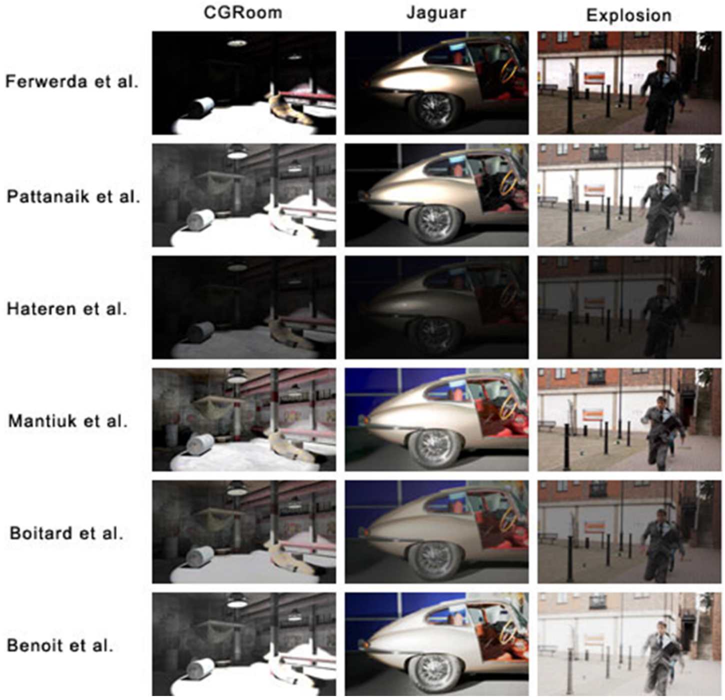

Figure 3: *Thumbnails of three of the videos used for the experiments.*

Table 2: *Features of the HDR videos used.*

Video	Length (s)	Avg. DR (Log units)	Capture	Device Max F-Stops
CGRoom	7	16.6	CG Canon 1D	20
Jaguar	13	13.4	Mark II	14
Kalabsha	11	18.5	CG	20
Morgan Lovers	15	19.5	HDRv Spheron	20
IDL Wedding	10	18.5	HDRv Spheron	20
Explosion	8	12	Canon 5D	12
Medical	4	15.2	HDRv Spheron	20

Software devised for this experiment was used for viewing and ranking the six TMOs for each sequence. Figure 4 shows an image from the software used. This software was only displayed on the HDR display to avoid any bias that a tone-mapped version of the reference thumbnail may have had on the SDR display.

Before applying the TMOs, consensus between three experts (who each had at least 2 years' experience in HDR imaging) was used to determine which were the best settings across all TMOs across all lighting levels. Therefore, the settings used for the TMOs were constant and for Ben the configurable parameters are the histogram clipping value ($h = 0$), the colour saturation factor ($s = 3$), the retina horizontal cells gain ($\beta_{\mathrm{h}} = 40$), the local adaptation photo-receptors ($R_{\mathrm{p},0} = 193$) and the local adaptation ganglion cells ($R_{\mathrm{g},0} = 186$). Boi was used with gamma correction ($\gamma = 2.2$) using the Ramsey *et al.* TMO [RH04] as it was the one advised by the authors to use. Fer was used setting maximum display luminance to 200 cd/m^2 and the minimum display luminance to 50 cd/m^2. Hat was used with the fixed pupil area ($p_{\mathrm{a}} = 10$ mm^2), the calcium feedback constant ($a_{\mathrm{c}} = 9 \times 10^{-2}$), the gain ($k_\beta = 1.6 \times 10^{-4}$) and the residual activity ($C_\beta = 2.8 \times 10^{-3}$). Man was applied using the 'LCD profile' that corresponds to a gamma correction ($\gamma = 2.2$), the maximum display luminance ($L_{\max} = 200$), the black levels of the display ($L_{\mathrm{black}} = 0.8$), the reflectivity of the display ($k = 0.01$) and the ambient illumination ($E_{\mathrm{amb}} = 60$). For Pat, the global version of the TMO was used since the local version cancels the time-dependent effects and the adaptation levels for cones and rods were calculated using the average luminance.

3.3. Participants

A total of 180 participants, 106 men and 74 women with ages between 19 and 28 years, were randomly assigned between the SDR (90 participants) and SSD experiment (90 participants), the three luminance conditions (30 for each luminance scenario) and between the seven scenes (15 evaluated the first four scenes and the other 15 the last three scenes). So for each experimental scenario, there were a total of 15 evaluations for each video on each device for each luminance condition. All the participants reported normal or corrected to normal vision.

3.4. Procedure

Each experimental scenario (luminance level, display and scene group) was randomly assigned between the participants. The participants were asked to view the six tone-mapped videos for the scenes they were allocated.

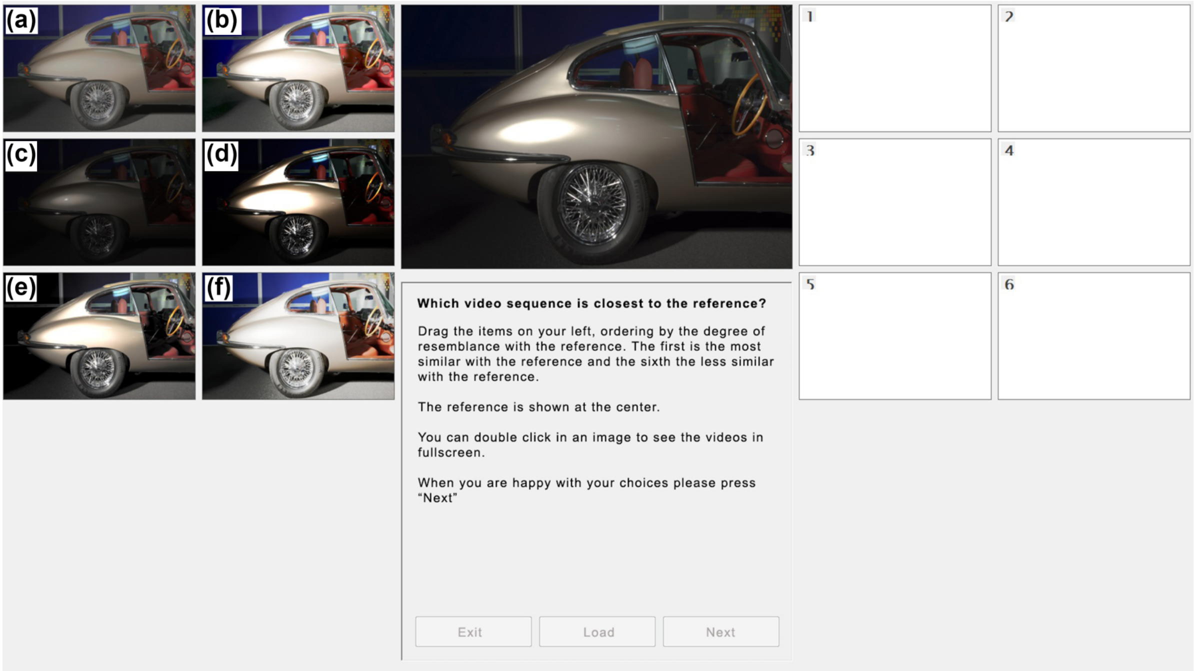

Figure 4: *Experimental software used.*

The software devised for this experiment presented on the left side thumbnails representing all the TMOs was presented (the order of these was automatically randomized). The reference video was presented in the centre of the screen. Clicking on the thumbnails played the appropriate video on the appropriate display. The tone-mapped video played on the SDR or SSD simultaneously as the HDR video played on the HDR display. On the right side, there were six empty slots numbered from 1 to 6. Having seen each video, the participants were asked to drag and drop the thumbnails of the tone-mapped videos according to how similar they thought each was compared to the reference onto what they thought was the appropriate position on the right-hand side. Before starting each experiment, it was told the participants to evaluate the contents as a whole and to have as reference the four main characteristics that are well-known for characterizing an image: colour, contrast, details and naturalness [ČWNA06].

The participants were allowed to look at any of the tone-mapped footage again as often as required before taking a final decision on any given scene. At any point, they could focus on either the SDR/SSD or the HDR directly. Before starting the evaluations, a few minutes was given to each participant to adapt to the ambient luminance of the room. The function of the software was also explained to participants. Furthermore, in order to avoid visual discomfort between scenes due to the variation of the lightning conditions on the displays, a grey image was shown between the scenes to help the eyes adjust.

4. Results

Due to the large amount of data collected, the results are divided into subsections with a discussion following in the next section. Initially, the results across all luminance, displays and groups are presented and subsequently similar results across displays shown for each of the luminance values. Finally, the results for individual video sequences are given.

4.1. Results across all displays across all luminance levels

In order to calculate the main effects, the results were collapsed across the video sequences by averaging the values across the seven video sequences for each TMO. The main effects of luminance levels (dark, medium and bright), display (SSD or SDR), TMO (Ben, Boi, Fer, Hat, Man, Pat) and group (the first group that saw the first four video sequences, and the second group that saw the subsequent three video sequences) were analysed using a $3 \times$ (luminance levels) $\times 2$ (group) $\times 2$ (display) $\times 6$ (TMO) mixed design factorial ANOVA for all the 180 participants. TMO was a within-participant variable and luminance level, group and display were between-participant variables. Kendall's Coefficient of Concordance W was also used to give an estimate of the agreement among participants. A Coefficient of Concordance of $W = 1$ signifies perfect agreement among the participants and $W = 0$ indicates complete disagreement. The results are shown in Table 3.

The group main effect was not significant, $F(1, 168) = 0$, $p > 0.05$ meaning that although different participants viewed different sequences the differences in the results between these two groups were insignificant. This result demonstrates that the different groups broadly performed the same.

The main effect of the display was not significant, $F(1, 168) = 1.289$, $p > 0.05$ signifying that, overall, there was no difference in the results across the displays. This does not mean that for a specific luminance the results were not significant and as shall be shown below this indeed was not the case.

Table 3: *Overall results obtained for each scene. Coloured groupings represent TMOs that were not found to be significantly different using pairwise comparisons to each other, via Bonferroni adjustment, at* $p < 0.05$.

	Kendall's Co-efficient of Concordance	1st	2nd	3rd	4th	5th	6th
Across all methods	0.71*	Man	Boi	Pat	Ben	Fer	Hat
Dark	0.83*	Man	Boi	Pat	Ben	Fer	Hat
Medium	0.86*	Man	Boi	Pat	Ben	Fer	Hat
High	0.73*	Pat	Man	Boi	Ben	Hat	Fer

*for (p < 0.01)

Table 4: *Overall similarity results obtained for each scene under dark luminance.*

	SDR							SSD						
	Kendall's Co-efficient of Concordance	1st	2nd	3rd	4th	5th	6th	Kendall's Co-efficient of Concordance	1st	2nd	3rd	4th	5th	6th
CGRoom	0.70*	Boi	Man	Ben	Pat	Hat	Fer	0.65*	Man	Ben	Boi	Hat	Pat	Fer
Jaguar	0.63*	Pat	Man	Boi	Fer	Ben	Pat	0.57*	Pat	Man	Boi	Ben	Fer	Hat
Kalabsha	0.69*	Man	Pat	Boi	Hat	Fer	Ben	0.81*	Man	Pat	Boi	Hat	Fer	Ben
Morgan Lovers	0.63*	Boi	Pat	Man	Ben	Fer	Hat	0.66*	Boi	Pat	Man	Ben	Fer	Hat
IDL Wedding	0.70*	Ben	Man	Pat	Boi	Hat	Fer	0.74*	Man	Ben	Boi	Pat	Fer	Hat
Explosion	0.71*	Man	Pat	Boi	Ben	Fer	Hat	0.76*	Boi	Man	Pat	Fer	Ben	Hat
Medical	0.78*	Boi	Man	Ben	Pat	Fer	Hat	0.69*	Man	Boi	Ben	Pat	Fer	Hat
Across Scenes	0.81*	Man	Boi	Pat	Ben	Fer	Hat	0.86*	Man	Boi	Pat	Ben	Fer	Hat

*for (p < 0.01)

For the luminance levels, there was a significant main effect, $F(1, 168) = 586.272$, $p < 0.05$ indicating that the luminance levels lead to different results on the rankings made by the participants. Pairwise comparisons with Bonferroni corrections at $p < 0.05$ indicated a significant difference between bright and both dark and medium luminance. There was no significant difference between the dark and medium luminance levels.

Sphericity was violated for the main effect of TMO (Maulchy's test, $p < 0.05$) so Greenhouse–Geisser correction was applied. The main effect of TMO was statistically significant, $F(4.383, 736) = 772.334$, $p < 0.05$ meaning that there was a clear distinction between the TMOs ranked and the accuracy ranking order was Man, Pat, Boi, Ben, Hat and Fer. There was a significant interaction for TMO × luminance, $F(8.765, 736) = 56.37$, $p < 0.05$, indicating the change TMOs exhibited across luminance levels. This is expanded on when discussing the TMOs at each of the luminance levels. None of the other interactions were significant. For all displays across all luminance levels, the Kendall Coefficient of Concordance was significant at $W = 0.71$.

4.2. Results for dark luminance

For the dark environment, the results obtained for the 60 participants who undertook the experiment under the dark luminance level were analysed. As with the results across the entire data set, the results were collapsed across the video sequences and analysed using a 2 (group) ×2 (display) × 6 (TMO) mixed design factorial ANOVA.

The results obtained for the dark environment scenario show that the main effect of the group was again insignificant, $F(1, 56) = 0.011$, $p > 0.05$.

For the displays, the main effect was also insignificant, $F(1, 56) = 0.524$, $p > 0.05$. The main effect of the TMOs was significant, $F(4.050, 227) = 298.381$, $p < 0.05$ (sphericity test was violated; Maulchy's test, $p < 0.05$ therefore Greenhouse–Geisser correction was applied). The accuracy ranking order was Man, Boi, Pat, Ben, Fer, Hat (see Table 3). The pairwise comparisons with Bonferroni corrections show that that Man was considered the most accurate representation of the HDR video and Boi and Pat as well as Fer and Hat showed no significant differences between each other. Kendall's coefficient of Concordance for this set of data was $W = 0.83$ ($p < 0.05$) which indicates a very high level of agreement among the participants when assigning the scores. Table 4 shows that the order of the TMO rankings was the same for both SSD and SDR with a little difference (not considered significant) in the groupings.

4.3. Results for medium luminance

For the medium luminance environment, results were originally published in Melo *et al.* [MBDC13a] but are reported here for completeness. The results were analysed as in the case of the dark environment. The main effect of the group was not significant, $F(1, 56) = 0$.

In this occasion, the main effect of the display was significant, $F(1, 56) = 6.179$ meaning that there is a significant difference on the accuracy ranking order between displays unlike the results for dark luminance and indeed across the displays.

As sphericity was violated for the main effect of TMO Greenhouse–Geisser correction was applied, $F(3.839, 215) = 2.836$, $p < 0.05$; this is again a significant difference. The accuracy ranking order on this case was Man, Boi, Pat, Ben, Fer and Hat with significant differences between them all. Kendall's coefficient of Concordance was $W = 0{,}86$ ($p < 0.05$). While, as Table 5 shows, the results obtained for the order of ranking between SDR and SSD was the same there were significant enough differences to consider this significantly different statistically.

4.4. Results for bright luminance

The results for bright luminance were analysed in the same manner as dark and medium luminance. The main effect of group was again insignificant, $F(1, 56) = 0$, $p > 0.05$. The main effect of display was not significant, $F(1, 56) = 0$, $p > 0.05$ meaning that the display had no influence on the ranking order made by the participants. In fact, as described below, the rankings of the TMOs were identical across the displays. The main effect

Table 5: *Overall similarity results obtained for each scene under medium luminance.*

	SDR							SSD						
	Kendall's Co-efficient of Concordance	1st	2nd	3rd	4th	5th	6th	Kendall's Co-efficient of Concordance	1st	2nd	3rd	4th	5th	6th
CGRoom	0.79*	Man	Boi	Ben	Pat	Fer	Hat	0.84*	Man	Boi	Ben	Pat	Hat	Fer
Jaguar	0.61*	Pat	Boi	Man	Ben	Fer	Hat	0.75*	Pat	Man	Boi	Ben	Fer	Hat
Kalabsha	0.73*	Man	Pat	Boi	Hat	Ben	Fer	0.76*	Man	Pat	Boi	Fer	Hat	Ben
Morgan Lovers	0.71*	Man	Boi	Ben	Pat	Fer	Hat	0.23*	Man	Pat	Ben	Boi	Hat	Fer
IDL Wedding	0.59*	Man	Ben	Boi	Pat	Fer	Hat	0.77*	Man	Boi	Ben	Pat	Fer	Hat
Explosion	0.77*	Pat	Man	Boi	Fer	Ben	Hat	0.76*	Man	Boi	Pat	Fer	Ben	Hat
Medical	0.75*	Man	Boi	Ben	Pat	Fer	Hat	0.88*	Boi	Man	Ben	Pat	Fer	Hat
Across Scenes	0.85*	Man	Boi	Pat	Ben	Fer	Hat	0.89*	Man	Boi	Pat	Ben	Fer	Hat

*for (p < 0.01)

Table 6: *Overall similarity results obtained for each scene under bright luminance environment.*

	SDR							SSD						
	Kendall's Co-efficient of Concordance	1st	2nd	3rd	4th	5th	6th	Kendall's Co-efficient of Concordance	1st	2nd	3rd	4th	5th	6th
CGRoom	0.33*	Boi	Man	Ben	Pat	Hat	Fer	0.57*	Man	Boi	Ben	Pat	Hat	Fer
Jaguar	0.55*	Pat	Man	Boi	Ben	Fer	Hat	0.33*	Pat	Man	Boi	Fer	Ben	Hat
Kalabsha	0.78*	Man	Pat	Boi	Hat	Fer	Ben	0.79*	Man	Pat	Boi	Hat	Fer	Ben
Morgan Lovers	0.67*	Boi	Man	Ben	Pat	Fer	Hat	0.70*	Boi	Pat	Man	Ben	Fer	Hat
IDL Wedding	0.70	Man	Ben	Boi	Pat	Fer	Hat	0.63*	Man	Ben	Pat	Boi	Fer	Hat
Explosion	0.70*	Boi	Man	Pat	Fer	Ben	Hat	0.63*	Boi	Man	Pat	Fer	Hat	Ben
Medical	0.80*	Boi	Man	Ben	Pat	Fer	Hat	0.79*	Boi	Man	Ben	Pat	Fer	Hat
Across Scenes	0.77*	Pat	Man	Boi	Ben	Hat	Fer	0.70*	Pat	Man	Boi	Hat	Ben	Fer

*for (p < 0.01)

of TMO was significant, $F(3.116, 175) = 184, 829$, $p < 0.05$ (sphericity test was violated via Maulchy's test, $p < 0.05$ therefore Greenhouse–Geisser correction was applied) and the accuracy ranking order was Pat, Man, Boi, Ben, Hat and Fer. Table 3 shows how Ben was grouped together with Boi and Hat, Boi was grouped with Ben and Hat with Ben for cases when there were no significant differences between them for pairwise comparisons with Bonferroni corrections at $p < 0.05$. Kendall's Coefficient of Concordance was $W = 0.73 (p < 0.01)$. Table 6 shows that the exact same sequence and grouping was obtained for both the SDR and SSD.

4.5. Overall results for SSD

Since this work is primarily focused on SSD, the results for SSD only across the 90 participants that carried out the experiment for the SSD condition are presented. As above, the results were collapsed across video sequences. The data analysed using a 2 (group) × 3 (luminance) × 6 (TMO) mixed design factorial ANOVA.

The group main effect was again insignificant, $F(1, 84) = 0.02$, $p > 0.05$. The main effect of luminance was significant,

Table 7: *Overall similarity results obtained from the experiments for each display.*

	SDR							SSD						
	Kendall's Co-efficient of Concordance	1st	2nd	3rd	4th	5th	6th	Kendall's Co-efficient of Concordance	1st	2nd	3rd	4th	5th	6th
Across All	0.71*	Man	Pat	Boi	Ben	Hat	Fer	0.71*	Man	Pat	Boi	Ben	Hat	Fer
Dark	0.81*	Man	Boi	Pat	Ben	Fer	Hat	0.86*	Man	Boi	Pat	Ben	Fer	Hat
Medium	0.85*	Man	Boi	Pat	Ben	Fer	Hat	0.89*	Man	Boi	Pat	Ben	Fer	Hat
Bright	0.77*	Pat	Man	Boi	Ben	Hat	Fer	0.70*	Pat	Man	Boi	Ben	Hat	Fer

*for (p < 0.01)

$F(2, 84) = 525.88$, $p < 0.05$. Pairwise comparisons with Bonferroni corrections showed a significant difference between the bright luminance level and the other two levels. There was no significant difference between the dark and medium luminance levels.

The main effect of TMO was also significant. Greenhouse–Geisser was applied as Maulchy's test proved significant, $F(4.148, 348.465) = 369.844$, $p < 0.05$. An order of Man, Pat, Boi, Ben, Fer and Hat was observed and pairwise comparisons of the TMOs with Bonferroni corrections at $p < 0.05$ gave all operators as being significantly different except Fer and Hat which did not have any significant differences between them. These results are presented in Table 7 labelled SSD, across all. Kendall's Coefficient of Concordance with $W = 0.89$ was considered significant, $p < 0.01$.

The above results including the differences between the luminance levels were the same as that seen for the overall results. Although not discussed in detail here, for the SDR display there was also a similar pattern observed for the different luminance levels with dark and medium luminance levels not considered significantly different.

4.6. Overall results for each scene

Tables 4, 5 and 6 show results for all the individual scenes for dark, medium and bright luminance, respectively. As before, the grouping shows the TMOs that have no significant differences for $p < 0.05$. For all scenes, Kendall's Coefficient of Concordance was significant, $p < 0.01$.

A fact that is noticeable is that for dark and medium luminance the accuracy ranking order is the same (Man, Boi, Pat, Ben, Fer, Hat) but when we refer to the high luminance environment the accuracy ranking order changes slightly (Pat, Man, Boi, Ben, Hat, Fer).

5. Discussion

This study investigates whether a TMO's ability to accurately represent an HDR scene would differ for HDR video across screens as reported with HDR images [UMM*10AEA13]. In addition, the study explored if this would occur only under specific lighting conditions or across different lighting conditions. The results demonstrate relatively close accuracy rankings across displays for each of the luminance levels. This is clear for the bright and dark environments as the results are very similar and although the medium environment reports a statistically significant difference across the displays the order is the same and this difference is relatively small ($F(1, 56) = 6.179$, for $p < 0.05$).

The TMO accuracy ranking remains the same for dark and medium luminance environments. For bright luminance environments, the accuracy ranking order changes slightly. This indicates that under conditions such as being outside, different TMOs can deliver a better perceptual match to the HDR content than the ones who provide better perceptual match on lower light level environments. The closeness in results between medium and dark as compared to bright conditions may be attributable to the selected luminance values; bright corresponds to outside values which are significantly higher at 1450 cd/m^2 compared to the 55 cd/m^2 used for the medium and 15 cd/m^2 for the dark which are representative of indoor environments. This is of particular interest to the case of SSD as they are frequently used in both outdoor and indoor environments.

A comparison of the tone-mapped videos against the HDR videos serves to provide a more comprehensive understanding of the results. The analysis performed took into consideration the content's naturalness, colour, detail and contrast as they are known for well characterizing the image [ČWNA06]. It was observed that overall Man seems to better reproduce the details of the image, but produces more saturated colours than Pat which produces more contrast and an image more similar image to the original HDR. Under the high luminance condition, Man's colours seem more unnatural and it is more difficult to perceive detail when compared to lower luminance levels. In fact, since the human eye is attracted to movement, the viewers probably pay more attention to the regions where movement occurs ignoring often the background and paying less attention to details. The study conducted by Narwaria *et al.* [NDSLCP14] also indicates that colour and naturalness are significant characteristics that influence user preference. Taking this into account, the results obtained indicate that under high luminance levels users may give more importance to contrast and naturalness over details and colour saturation. Further research will be required to confirm this observation.

Man TMO had the best performance since it was ranked as the most accurate for dark and medium luminance and the second most accurate for high luminance. This good performance can be explained by the fact that Man compensates for the ambient lightning but on high luminance its accuracy can be less effective since this operator focuses on contrast reproduction; although under high luminance environments the image can look less natural.

Boi performed better for the lower luminance values. The three best performers, Boi, Man and Pat consider the time-adaptation of the HVS to large changes in scene intensities and this could be a further factor that led to them being ranked highly. Boi considers the video sequence as a whole and derives the SDR image based on the video features that preserve the relative luminance levels and the temporal brightness consistency. For scenes with lower dynamic range, Boi had very good rankings suggesting further that below a certain dynamic range the colour appearance characteristics may become less important.

Fer tries to preserve colour appearance and deal with adaptation to luminance but it fails when the dynamic range of the scene is considerably high since it cannot preserve details in dark regions while reproducing details in bright regions. This leads to the introduction of artefacts and that could explain Fer's performance. For scenes of lower dynamic range (Jaguar and Explosion) Fer obtained its highest rankings. Hat aims to preserve the scene luminance and reduce noise but struggles when the scene is brighter since the contrast is smaller in regions with more luminance.

5.1. Comparisons with other findings

Contrary to Urbano *et al.* [UMM*10] and Akyüz *et al.* [AEA13] findings for HDR imagery with HDR video the ranking does not change significantly across displays. This finding could be related to findings by Narwaria *et al.* [NSCP14] that were concerned with the importance of visual attention in HDR content processing. The authors concluded that the mechanisms of visual attention are much more significant for images than for video and associate that to the interframe correlation and shorter viewing time. Our results may be indicative of the same phenomenon due to the motion associated with video. As others have also shown, e.g. [ZS06], while participants are viewing an image they may pay more attention to all details across the whole image whereas when viewing a video it is most likely they devote more attention to regions were motion occurs.

When compared to the work conducted by Akyüz *et al.* [AEA13] that concludes that the differences between TMOs is less noticeable on mobile devices and therefore simpler mapping techniques can be used, our work does not support such a conclusion. In scenarios that could be equivalent to the scenarios investigated by Akyüz *et al.*, i.e. dark and medium, our results show a higher degree of concordance between the SDR and SSD. Furthermore, for video the groupings are better defined in such scenarios for SSDs, indicating that participants potentially found it easier to spot differences and rank the different TMOs. The difference in the results may be explained due to Akyüz *et al.* using a 3" display that is three times smaller than the SSD used on our experiments. Further tests will involve looking into smaller screen sizes in order to further investigate the full spectrum of SSDs as discussed below.

5.2. Study limitations

Throughout this study there were a number of factors that were taken into account in order to minimize possible bias as well as luminance adaptation of the human eye. The displays themselves may have contributed to adaptation issues in the participants. However, as discussed earlier the direct comparison design was chosen to avoid problems with recall bias. The maladaptation issue was mitigated by using a grey screen.

A particular limitation was the choice of the SSD screen size which was closer to upper scale of SSD screen sizes, rather than focussing on individual screen sizes and the difference in these between, say a tablet and a smart phone. The original resolution and aspect ratio of the images were maintained in order to be able to perform a straight comparison between the reference and the tone-mapped sequence on the mobile device to guarantee that no information was lost due to the conversion of the content to a lower resolution. Small screen sizes would have required retargeting the content. Also, participants were not allowed to adjust the brightness of the SDD. Future work will consider the impact of smaller screen sizes, although the trend for smart phones appears to be towards larger screens. The energy source is also a limitation of mobile devices so it is intended to further investigate the impact of different TMOs on battery consumption and optimize their use according to different scenarios. In addition, future work will seek to recreate conditions under which mobile devices are used by considering the impact of different viewing angles and distances.

6. Conclusions

Mobile devices are rapidly becoming the leading platform for the consumption of multi-media content and that raises an urgent need to ensure an optimal experience when viewing HDR content on such devices. One of the goals of this work was to understand if and how the ambient luminance levels can have an impact on how a TMO can accurately reproduce an HDR scene and if this accuracy is the same for the different sized displays under different luminance levels. More than just understanding which TMO would perform better under the different scenarios, we wanted to verify if the different luminance levels would have an impact on the order of the ranking.

The main contribution is to show that the accuracy of the perceptual match of TMOs can change depending on the environment luminance levels. The results indicate that the TMO's ranking remains essentially the same across displays under the same luminance levels and for dark and medium environments the ranking is the same. When dealing with bright environments, e.g. an overcast outdoor environment, such as the one we simulated, the TMO's ranking changes suggesting that for high luminance users tend to prefer more contrast and naturalness in the images than for dark and medium luminance scenarios where they seemed to value more colour and details.

SSDs on mobile devices can be exposed to different luminance levels at the same time due to various factors including being used 'on-the-go'or due to shadows or reflections on the screen. Future work will thus investigate whether dynamically changing the TMO can enhance the viewing experience of HDR video on SSDs as the environmental conditions change as well as it will consider the different viewing angles and viewing distances.

Acknowledgements

We would like to thank to the authors of the TMOs for their help with the code for tone mapping the frames used in the experiments and the participants for taking part in the study. We would like to thank also to Elmedin Selmanovic for providing the Jaguar HDR footage.

This work was partially supported by the Portuguese government, through the National Foundation for Science and Technology - FCT (Fundação para a Ciência e a Tecnologia) through the project SFRH/BD/76384/2011 entitled 'HDR Video for Mobile Devices' and partially supported by the Portuguese government, through the National Foundation for Science and Technology—FCT (Fundação para a Ciência e Tecnologia) and the European Union (COMPETE, QREN and FEDER) through the project REC I/EEI-SII/0360/2012 entitled 'MASSIVE—Multimodal Acknowledgeable multiSenSorial Immersive Virtual Enviroments'.

Debattista is partially supported by a Royal Society Industrial Fellowship. This work was also partially supported by the ICT COST Action IC1005 'HDRi: The digital capture, storage, transmission and display of real-world lighting'.

References

[AEA13] AKYÜZ A. O., EKSERT M. L., AYDIN M. S.: An evaluation of image reproduction algorithms for high contrast scenes on large and small screen display devices. *Computers & Graphics 37*, 7 (2013), 885–895. doi: dx.doi.org/10.1016/j.cag.2013.07.004.

[AFR*07] AKYÜZ A. O., FLEMING R., RIECKE B. E., REINHARD E., BÜLTHOFF H. H.: Do HDR displays support LDR content?: A psychophysical evaluation. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM. doi:10.1145/1275808.1276425.

[ASC*14] AYDIN T. O., STEFANOSKI N., CROCI S., GROSS M., SMOLIC A.: Temporally coherent local tone mapping of HDR video. *ACM Transactions on Graphics 33* 6 (Nov. 2014), 196:1–196:13. doi:10.1145/2661229.2661268.

[AvMS10] AYDIN T. O., ČADÍK M., MYSZKOWSKI K., SEIDEL H.-P.: Video quality assessment for computer graphics applications. In *SIGGRAPH ASIA '10: ACM SIGGRAPH Asia 2010 Papers* (New York, NY, USA, 2010), ACM, pp. 161:1–161:12. doi:10.1145/1866158.1866187.

[BADC11] BANTERLE F., ARTUSI A., DEBATTISTA K., CHALMERS A.: *Advanced High Dynamic Range Imaging: Theory and Practice* (1st edition). AK Peters, Ltd./CRC Press (Natick, MA, USA, Feb. 2011). http://vcg.isti.cnr.it/Publications/2011/BADC11.

[BALCH09] BENOIT A., ALLEYSSON D., LECALLET P., HÉRAULT J.: Spatio-temporal tone mapping operator based on a retina model. In *Proceedings of Computational Color Imaging Workshop* (Saint Etienne, France, Mar. 2009), LNCS, Springer, pp. 12–22. Département Images et Signal Département Images et Signal. http://hal.archives-ouvertes.fr/hal-00366126.

[BBC*12] BOITARD R., BOUATOUCH K., COZOT R., THOREAU D., GRUSON A.: Temporal coherency for video tone mapping. In *Applications of Digital Image Processing XXXV* (San Diego, California, USA, 2012), vol. 8499 of *Proc. SPIE*, pp. 84990D–84990D–10. doi:10.1117/12.929600.

[BTDH10] BRÉMOND R., TAREL J.-P., DUMONT E., HAUTIERE N.: Vision models for image quality assessment: One is not enough. *Journal of Electronic Imaging 19* 4, (2010), 043004–043004–14. doi:10.1117/1.3495989.

[CIS13] CISCO: Cisco visual networking index: Global mobile data traffic forecast update, 2012-2017. http://www.cisco.com/c/en/us/solutions/collateral/serv ice-provider/visual-networking-index-vni/white_paper_c11-520862.html, 2013. Accessed 21 November 2013.

[ČWNA06] Čadík M., Wimmer M., Neumann L., Artusi A.: Image attributes and quality for evaluation of tone mapping operators. In *Proceedings of Pacific Graphics 2006 (14th Pacific Conference on Computer Graphics and Applications)* (Oct. 2006), National Taiwan University Press, pp. 35–44. http://www.cg.tuwien.ac.at/research/publications/2006/CADIK-2006-IAQ/.

[ČWNA08] Čadík M., Wimmer M., Neumann L., Artusi A.: Evaluation of HDR tone mapping methods using essential perceptual attributes. *Computers & Graphics 32*, 3 (June 2008), 330–349.

[DMMS02] Drago F., Martens W., Myszkowski K., Seidel H.-P.: *Perceptual evaluation of tone mapping operators with regard to similarity and preference*. Research Report MPI-I-2002-4-002, Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany, August 2002.

[EUWM13] Eilertsen G., Unger J., Wanat R., Mantiuk R.: Survey and evaluation of tone mapping operators for HDR video, 2013.

[FPSG96] Ferwerda J. A., Pattanaik S. N., Shirley P., Greenberg D. P.: A model of visual adaptation for realistic image synthesis. In *SIGGRAPH '96: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1996), ACM, pp. 249–258. doi:10.1145/237170.237262.

[ITU11] ITU: The world in 2013: ICT facts and figures. http://www.itu.int/ITU-D/ict/facts/material/ICTFactsFigures-2011.pdf, 2011. Accessed 21 November 2013.

[KYL*07] Kuang J., Yamaguchi H., Liu C., Johnson G. M., Fairchild M. D.: Evaluating HDR rendering algorithms. *ACM Transactions on Applied Perception 4*, 2 (July 2007). doi:10.1145/1265957.1265958.

[LCTS05] Ledda P., Chalmers A., Troscianko T., Seetzen H.: Evaluation of tone mapping operators using a high dynamic range display. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers* (New York, NY, USA, 2005), ACM, pp. 640–648. doi:10.1145/1186822.1073242.

[MAS07] Meylan L., Alleysson D., Süsstrunk S.: Model of retinal local adaptation for the tone mapping of color filter array images. *Journal of the Optical Society of America A 24*, (2007), 2807–2816.

[MBDC13a] Melo M., Bessa M., Debattista K., Chalmers A.: Evaluation of HDR video tone mapping for mobile devices. *Signal Processing: Image Communication*, *29*, 2 (February 2014), 247–256. http://dx.doi.org/10.1016/j.image.2013.09.010.

[MBDC13b] Melo M., Bessa M., Debattista K., Chalmers A.: Evaluation of tone mapping operators for HDR video on small form factor displays. In *Proceedings of HDRi2013: First International Conference and SME Workshop on HDR Imaging* (Porto, Portugal, 2013), M. Bessa and R. Mantiuk (Eds.).

[MDK08] Mantiuk R., Daly S., Kerofsky L.: Display adaptive tone mapping. *ACM Transactions on Graphics 27*, 3 (Aug. 2008), 68:1–68:10. doi:10.1145/1360612.1360667.

[MKRH11] Mantiuk R., Kim K. J., Rempel A. G., Heidrich W.: HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Transactions on Graphics 30*, 4 (July 2011), 40:1–40:14. doi:10.1145/2010324.1964935.

[MMS04] Mantiuk R., Myszkowski K., Seidel H.-P.: Visible difference predicator for high dynamic range images. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics* (San Diego, CA, USA, Oct. 2004), pp. 2763–2769.

[NDSLCP14] Narwaria M., DaSilva M., LeCallet P., Pepion R.: Single exposure vs tone mapped high dynamic range images: A study based on quality of experience. In *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)* (Lisbon, Sept. 2014), pp. 2140–2144.

[NSCP14] Narwaria M., Silva M. P. D., Callet P. L., Pepion R.: Tone mapping based HDR compression: Does it affect visual experience? *Signal Processing: Image Communication 29*, 2 (2014), 257–273. Special Issue on Advances in High Dynamic Range Video Research. doi:http://dx.doi.org/10.1016/j.image.2013.09.005.

[PTYG00] Pattanaik S. N., Tumblin J., Yee H., Greenberg D. P.: Time-dependent visual adaptation for fast realistic image display. In *SIGGRAPH '00: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 2000), ACM Press/Addison-Wesley Publishing Co., pp. 47–54. doi:10.1145/344779.344810.

[RH04] Ramsey S. D., Johnson T., Hansen C.: Adaptive temporal tone mapping. In *CGIM 2004: Proceedings of Computer Graphics and Imaging* (2004), pp. 3–7.

[SHS*04] Seetzen H., Heidrich W., Stuerzlinger W., Ward G., Whitehead L., Trentacoste M., Ghosh A., Vorozcovs A.: High dynamic range display systems. *ACM Transactions on Graphics 23*, 3 (Aug. 2004), 760–768. doi:10.1145/1015706.1015797.

[UMM*10] Urbano C., Magalhães L., Moura J., Bessa M., Marcos A., Chalmers A.: Tone mapping operators on small screen devices: An evaluation study. *Computer Graphics Forum 29*, 8 (2010), 2469–2478. doi:10.1111/j.1467-8659.2010.01758.x.

[VH06] Van Hateren J. H.: Encoding of high dynamic range video with a model of human cones. *ACM Transactions on Graphics 25*, 4 (Oct. 2006), 1380–1399. doi:10.1145/1183287.1183293.

[WM14] Wanat R., Mantiuk R. K.: Simulating and compensating changes in appearance between day and night vision. *ACM Transactins on Graphics 33*, 4 (July 2014), 147:1–147:12. doi:10.1145/2601097.2601150.

[ZS06] Zhai Y., Shah M.: Visual attention detection in video sequences using spatiotemporal cues. In *MULTIMEDIA '06: Proceedings of the 14th Annual ACM International Conference on Multimedia* (New York, NY, USA, 2006), ACM, pp. 815–824. http://doi.acm.org/10.1145/1180639.1180824, doi:10.1145/1180639.1180824.