2016-07-12

Impact of tone-mapping operators and viewing devices on visual quality of experience

Ifeachor, Emmanuel

http://hdl.handle.net/10026.1/13252

10.1109/ICC.2016.7510690
2016 IEEE International Conference on Communications, ICC 2016
IEEE

Impact of Tone-mapping Operators and Viewing Devices on Visual Quality of Experience

Shaymaa Al-Juboori, Is-Haka Mkwawa, Lingfen Sun and Emmanuel Ifeachor
Centre for Signal Processing and Multimedia Communication
Plymouth University, UK
E-mail: {shaymaa.al-juboori, is-haka.mkwawa, l.sun, e.ifeachor} @ plymouth.ac.uk

***Abstract*— The development of HDR imaging is seen as an important step towards improving the visual quality of experience (QoE) of the end user in many applications. In practice, Tone-mapping operators (TMOs) provide a useful means for converting a high dynamic range (HDR) image to a low dynamic range image (LDR) in order to achieve better visualization on standard displays. Although mobile devices are becoming popular, the techniques for displaying the content of HDR images on the screens of such devices are still in the early stages. While several studies have been conducted to evaluate TMOs on conventional displays, few studies have been carried out to evaluate TMOs on small screen displays, such as those used in mobile devices. In this paper we evaluate, using subjective and objective methods, the most popular Tone-mapping-operators in different mobile displays and resolutions under normal viewing conditions for the end-user. Preliminary results show that small screen displays (SSDs) have an impact on the performance of TMOs compared to computer displays. In general, the larger the mobile resolution, the better the subjective results. We also found clear differences between SSDs and LDRs performances. The best TMO for mobile displays is iCAM06 and for computer displays it is Photographic Reproduction.**

Keywords—HDR, Tone mapping operators, Subjective tests, Objective test, Small screen devices, mobile devices, Low dynamic range, Standard dynamic range, Quality of Experience.

I. Introduction

In recent years, we have witnessed widespread application of High Dynamic Range (HDR) imaging due to its ability to capture a wide range of luminance values, similar to that of the human visual system (HVS). The application areas include home-entertainment, security, scientific image, video processing, computer graphics and multimedia communications [1]. However, in practice the full HDR content cannot be displayed on standard or low dynamic range (LDR) displays, and this diminishes the benefits of HDR technology to many users. To address this, Tone-Mapping Operators (TMO) are used to convert HDR images so that they can be displayed on low-dynamic-range displays and preserve as far as possible the perception of HDR [2].

A large number of different TMO algorithms have been proposed in recent years, with varying degrees of success in preserving the perceptual quality of HDR images. The need to evaluate the performance of TMO algorithms to inform the choice of algorithms for different displays and application is widely recognised [1]. There has been a number of studies undertaken to address this, but most of these were carried out using large conventional displays such as those of TV sets and PC monitors [1,2] and very few using small screen devices such as those of mobile phones [3,4]. There is also no concrete indication of which TMO performs the best.

With advances in mobile wireless communication, the popularity of mobile devices and mobile applications are growing dramatically. It is predicted that by 2019, there will be 8.2 billion handheld or personal mobile-ready devices and 3.2 billion mobile-to-mobile connections [5]. With the ability and convenience to be used anywhere and at any time, smart mobile devices have become the main means for receiving multimedia content [3]. The need remains to understand how TMO algorithms perform on small screen devices, such as those of the mobile phones. This is exacerbated because of the existence of many different mobile devices and brands with different resolutions, sizes and models.

It is unclear how current TMO algorithms perform in small screen devices (SSD), such as mobile phones and tablets, and whether they can be used directly for SSD or as SSD-friendly, or more specifically mobile-friendly. The importance of this issue has recently began to be addressed [3,4]. However, only two studies have been reported so far. Ubano et al. [3] carried out the first subjective evaluation of seven TMO algorithms on three different displays including LDR and a mobile device for still images. They found that the TMOs perform significantly different for SSDs compared with LDRs. However, only one mobile device (with a screen size of 2.8'') was tested. Melo et. al. [4] carried out a subjective evaluation of six different TMO algorithms for video using three displays (HDR, LDR and Tablet) and did not find major differences between SSDs and LDRs. Their work was limited to video and the testing was only based on one tablet. In both studies, the Quality of Experience (QoE) of the end-user was not taken into consideration in the experiments.

QoE driven multimedia systems have increasingly come into focus in both research and industry. Capturing the end-user's aesthetic expectations is the aim rather than simply delivering content based on a technology-centric approach. HDR is one of the important new developments which provide end-users with enhanced realistic viewing experience and thus improving the QoE [6].

QoE assessments are traditionally performed in laboratories under controlled viewing conditions. However, the Web is now considered as an important platform for uncontrolled QoE assessments with large numbers of participants. It also helps to create a realistic test environment, as the assessment is done directly on the participants' devices. However, it is not clear whether different mobile devices have differential impact on the QoE of HDR images, and if so, to what extent the impact is compared to conventional displays.

In this paper, we investigate the impact of different mobile devices and resolutions in assessing QoE of tone-mapping operators and address a number of major concerns regarding TMOs, e.g.: Are the TMOs which were successful for traditional displays also successful for SSDs? Do different device sizes/ resolutions affect the QoE?

The rest of the paper is organized as follows. Section II reviews briefly the related work in evaluating TMOs and Section III discusses the experimental framework. Section IV presents the experimental results. Section V discusses the objective quality metrics and their result. In Section VI, we evaluate the performance indices between four subjective tests on the one hand and between subjective and objective tests on the other. Conclusions and future work are given in Section VII.

II. Related Work

Error metrics and psychophysical experiments are the two main methodologies for evaluating TMO. Error metrics are objective methods that compute quality indices by comparing images [7]. 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. Psychophysical experiments are subjective and based on human participants. These experiments are conducted in controlled environments and can make use of a number of evaluation methods for comparing images such as rating, pairwise comparison or ranking. Several psychophysical experiments have previously been conducted. Cadick [8] adopted a direct rating Full Reference comparison of the tone mapped images of real scenes, and a subjective ranking of tone mapped images without a real reference. They applied 14 methods, and three typical real world HDR scenes. More recently, Salih [9] compared six tone operators using visual rating by comparing the printings and LDR display devices. The study concluded that photographic reproduction TMO is the best in terms of visual preference. Urbano et al [3] was one of the first studies aimed specifically at SSDs. They evaluated several TMOs on displays with different sizes using a pairwise comparison test of the processed images with a reference of real scenes. Three different displays were used, two 17" and one 2.8" displays with resolutions of 1024×682 and 240×320, respectively. The authors concluded that the order of preference for TMO between the displays was different and that for mobile devices, the content that offered stronger detail reproduction, more saturated colors and overall brighter image appearance were preferred.

Despite a large body of research devoted to the evaluation of TMOs, there is no standard methodology for performing such studies. The choice of method depends on the application and what is relevant to the study. In this study, we employ Non-Reference (NR) and Full Reference (FR) methods since in many end-user viewing applications there is no need for comparing with "perfect" or "reference" image. In the FR image quality evaluation, the task is to determine the quality of reproduction with reference to the original image which has to be available. In NR evaluation, the original is not available and image quality features can be used instead [7].

III. Experimental Framework

Two sets of subjective, visual quality assessments were conducted using the same dataset in generic environments. 60 observers were involved and the viewing conditions included indoor and outdoor environments, with natural and artificial light. Participants were free to look at the images on the Websites in the way they felt comfortable. Typically, subjective quality assessment involves quality rating, and the final result is expressed as a Mean Opinion Score (MOS), that is the average of the individual scores.

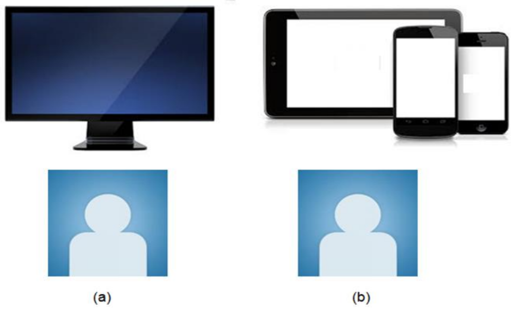


Fig. 1. Experimental setup (a) computer test (b) mobile test

Two experimental setups were designed for this study (c.f., Fig. 1). In the first experiment, a website [11] is designed and accessed from LDR displays of personal computers. The instructions for the test were made available on the website. We chose a discrete, five-level scale rating table for ITU-R quality ratings. This is more suitable for naïve observers (non-experts in image processing) as it is easier for them to quantify the quality from "bad" (1) to "excellent" (5) [15]. Gamma correction of 2.2 was applied to the tone mapped images as a last step of the tone-mapping algorithms in order to compensate for the non-linearity of displaying devices [1]. The experiment has two tests, test 1 and 2 which are FR and NR, respectively. Two websites were created for each test of all TM images and the MOS results were submitted to a database at the end of each test (Continuous test). Participants were asked to read the instructions and then view 30 images (divided into 2 websites 15 images per website).

The web site for the second experiment was designed to be accessed from SSDs, i.e. smart phones and tablets [11] as shown in Fig. 2. The instructions for the second experiment were sent to participants in a recruitment email. The MOS in this case is an eleven-grade quality scale ('10=no further improvement is possible' and' 0=A worse quality cannot be imagined') [15]. There were two tests in this experiment, FR and NR. Each test has three websites for the TM images. For each test, participants submitted their MOS, individually, for each image. Next" and "Previous" buttons allow participants to evaluate next images or to review previous images. Participants can also swipe the screen to move forward and backwards. A progress bar appears below the TM images as an indicator of percentage of progress so far (c.f., Fig. 2).

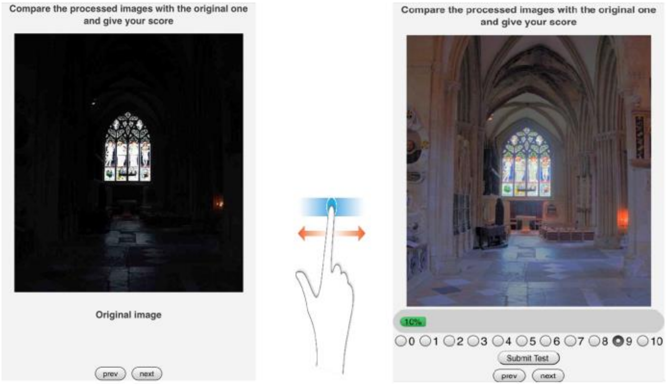


Fig. 2. Mobile website

A. Participants

The total number of participants for the entire study was 60. All of the participants were between 20 and 50 years old and had normal or corrected vision and non-experts in HDR, but have a clear understanding of the test.

B. Devices

In the Mobile experiment, five different devices for a total of 30 participants were used as shown in TABLE I., while for the computer experiment a total of 30 participants were also used; TM images were displayed on Philips Brilliance 221P3LPYES, 21.5-inch LED-backlit, LCD panel display with a native display resolution of 1920×1080.

TABLE I. DEVICES FOR MOBILE EXPERIMENT

Devices	**No of Users /Devise**	**Features**	**Resolution / pixels**
IPhone 6	9	4.7 inch Retina HD display,	1334×750
IPhone 5S	7	4 inch Retina display	1136×640
Samsung Galaxy noteII	5	5.5 inch Super AMOLED display	1280×720
Samsung Galaxy S4	3	5 inch HD Super AMO LED display	1920×1080
IPad mini 3	6	7.9 inches, IPS LCD	2048×1536

C. Considered TMOs

In this study, we used ten local and global well-known tone mapping operators; *Ashikhmin AL1, Ferwerda AL2, Adaptive Logarithmic Mapping AL3, iCAM06 AL4, Fattal AL5, Pattanaik AL6, Photographic Reproduction AL7, Tumblin – Rushmeier AL8, Ward AL9* and *Bilateral Filtering AL10* [8,9,13,14].

D. Dataset

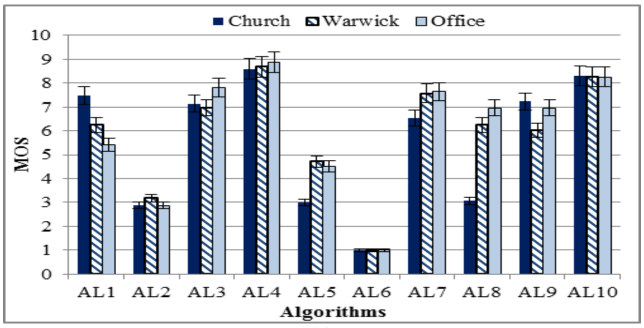
The dataset consists of three HDR images and 30 HDR images obtained from the ten tone mapping algorithms (computed using Banterle's HDR toolbox for MATLAB and iCAM06 source code which are freely available with the default settings of operators' performance as presented in the respective papers)[13,14]. The images were selected for this study, based on their visual content, quality and the dynamic range of the content. We used an existing HDR image database; the indoor scene is *Oxford Church*, Author: Banterle, Resolution: 840×886. The dynamic ranges of images are about 10^0: 10^3 cd/m^2. The outdoor scene is *Warwick*, Author: Banterle, Resolution: 1189×598, the dynamic ranges is about 10^{-1}: 10^1 cd/m^2. Indoor and Outdoor scene *Office* Resolution: 1165×751, the dynamic range of the image is about 10^{-2}: 10^1 cd/m^2.

IV. EXPERIMENTAL RESULTS

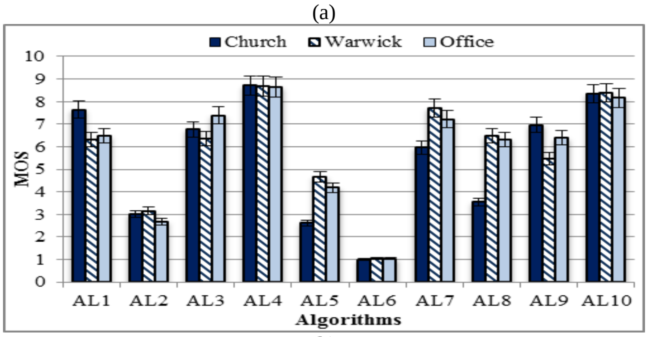
The first step of the analysis of the results is the calculation of the mean opinion score. The raw subjective scores were converted into a corresponding MOS for each sequence with 95% confidence interval. In each test, the quality score values were converted to the range [1:10] by mapping the lowest and highest quality score values to 1 and 10, respectively, Intermediate values were scaled proportionally.

Fig. 3. (a) and (b) shows the results of the mobile experiments. In (a), *iCAM06 Al4* and *Bilateral Filtering AL10* had the best performance from the observers' point of view, with a very good MOS scores between 8.8 and 8.2 for the three images. These two operators preserve good details compared to the reference image. *Adaptive Logarithmic Mapping AL3* obtained MOS less than 8, while The worst TMO was *Pattanaik AL6* with MOS of 1 for all images and *Ferwerda AL2*. Moreover, in (b) for the NR test *iCAM06 Al4* and *Bilateral Filtering AL10* still performs as best TMOs, *Adaptive Logarithmic Mapping AL3* and *Ashikhmin AL1* obtained good results of MOS between 7 and 8. *While Pattanaik AL6* still with the lowest MOS of 1.

The results of the computer experiment are illustrated in Fig. 4. The FR test (a), shows the results of the three images; *Church, Warwick and Office*. *Photographic Reproduction AL7* had the best performance from the observers' point of view, with very good MOS scores around 9 for the three images, while *Adaptive Logarithmic Mapping AL3* and *iCAM06 AL4* performed well as well with MOS between 8 and 9. The global *Drago TMO AL7* is based on logarithmic compression of luminance. While the best performance of local operator of Reinhard came from applying the dodging and burning technique, authors provided an efficient way of compressing the dynamic range while reducing halo artefacts [8]. Less halo results into a very good overall image quality. In the other hand *Fattal AL5* and *Pattanaik AL6* were the worst TMOs. The reason behind the low performance of *Pattanaik* is that it's using a multiscale decomposition of the image according to a comprehensive psychophysically-derived filter banks. However, it may still present halos which affected the overall quality of the image [8]. In the NR test (b) the MOS results were the same on the FR test, but with different MOS results for both the best TMOs and the worst one. While *Fattal et al*. treat HDR images with a gradient attenuation method. Their method is very good at increasing local contrast without creating halo artefacts [17]. By comparing the results of computer and mobile experiments in Fig. 3. and Fig. 4. , in (a) the FR test we can see that in computer came close to each oth-

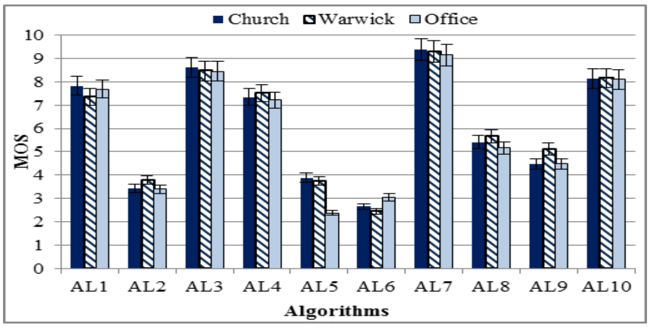


(a)

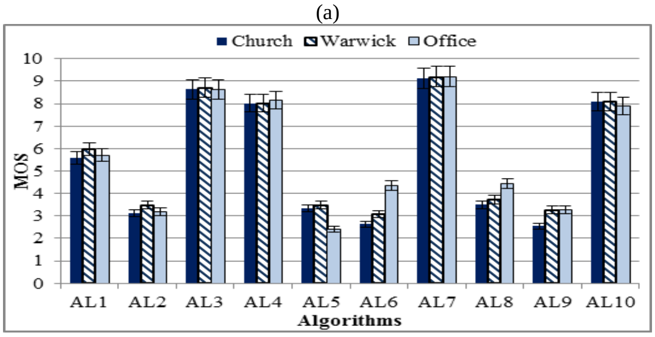


(b)

Fig. 3. MOS for Mobile experiment (a) FR test (b) NR test



(a)



(b)

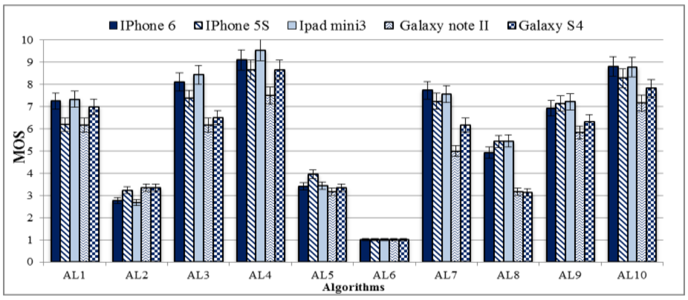
Fig. 4. MOS for Computer experiment (a) FR test (b) NR test.

-er and less variation between the MOS of subjects. The results of *Pattanaik AL6* in the mobile test had the lowest MOS compared with the other TMOs for the three images (MOS=1), but for the computer test it was the lowest MOS as well, with an average of (MOS=2.5) which is significantly higher from the mobile results. While in (b) we can see that less variation in the results appear for both tests. For mobile test *AL4* performed better, while in computer test *AL7* had better MOS results. Moreover, from the results we can see the variance in terms of highest and the lowest performance of TMOs is very clear between SSDs and LDRs.

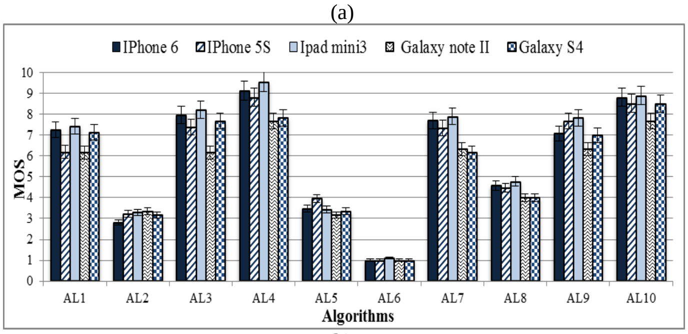
Different mobile and tablet devices have different display features TABLE I. The devices have been used in this study; have screen sizes varying between 4 and 7.9 inch and with different screen resolutions. Fig.5. shows the results of SSDs behavior in uncontrolled viewing conditions (a) FR test (b) NR test. In (a) and (b) the results suggests that the screen resolution and size are particularly important for higher MOS results. iPad mini 3 gave the favorable results compared to other devices; iPhone 6 behaved very well and iPhone 5S comes in third place. Samsung Galaxy note II had the lowest results if compared to overall device types. To analyze the results, we can see that the SSDs resolution effect in the first place, moreover, there is no vast difference in mobile devices performance in uncontrolled viewing conditions for HDR image evaluation whether it was NR or FR tests.

V. Objective Quality Metrics

Subjective rating may be a reliable evaluation method, but it is expensive and time consuming, and more importantly, it is difficult to be embedded into optimization frameworks. The goal of objective image quality assessment research is to provide quality metrics that can predict perceived image quality automatically [10].



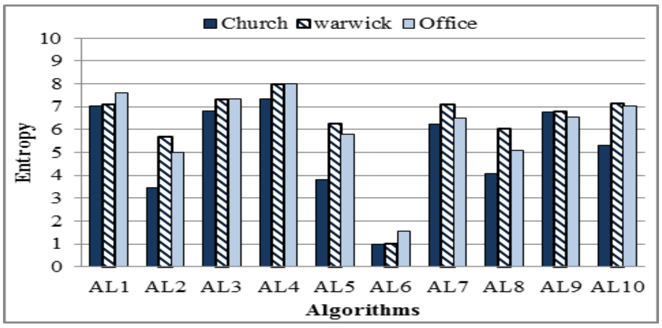
(a)



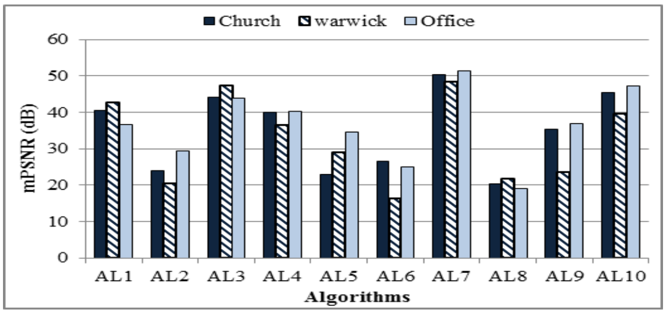
(b)

Fig. 5. SSDs behavior in uncontrolled viewing conditions tests (a) FR (b) NR

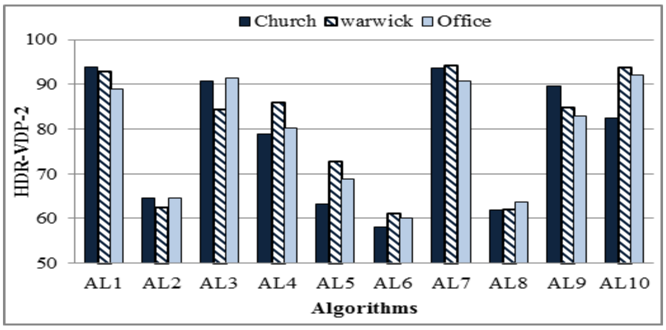
As there is no established standard for evaluating HDR image quality [6][7][10][18], we chose to use three error metrics; Shannon Entropy (E), The Multi-Exposure Peak Signal Noise Ratio (mPSNR) and Visual difference predictor for HDR images HDR-VDP-2. Entropy is used to measure the salient features of the image. Large entropy means that the fused image contains more information and implies a better image fusion [10]. The mPSNR metric is an extension of the peak signal-to-noise ratios (PSNR) metric to HDR domain.



(a)



(b)



(c)

Fig. 6. Objective quality metrics (a) Entropy (b) mPSNR (c) HDR-VDP-2

It takes a series of exposures which are tone mapped using a simple gamma curve and it is expressed in decibel (dB). This means that the larger the value of the metric the better. [1].The HDR-VDP-2 has been designed and calibrated to predict visibility rather than quality [7]. Fig. 6 shows the objective quality metrics results.

The Entropy of original images is: 6.2457, 4.5064, and 6.5698 for *office, church and Warwick* respectively. Fig. 6. (a) shows the results of Entropy for the TM images; *AL4* and *AL3* have the highest scores. The image with the highest entropy value is selected as the most detailed image because *Adaptive Logarithmic Mapping and iCAM06* preserve details while providing high contrast compression in saturated regions. As a result, the algorithm's output has less saturated pixels and wider histogram. That is why they are rated higher than the others [10]. While the worst Entropy result came from *AL6 Pattanaik and AL2 Ferwerda*.

mPSNR results are shown in Fig. 6. (b), *AL7 Reinhard* came with best TMO results. *Office* and *Church* performed the best images with mPSNR more than 50 dB, and then *Warwick* came with a result of 48 dB. The best second performance came from *AL10 the Bilateral filter*, while the lowest results came from *AL8 Tumblin and Rushmeier* with mPSNR less than 21 dB for the three images. The other algorithms performed in the middle with mPSNR between 45 and 30 dB.

HDR-VDP-2 results are shown in Fig. 6. (c), we can see that *AL7, AL1, AL10 and AL3* which are *Reinhard, Ashikhmin, Bilateral Filter and Drago* respectively performed well, while, *AL6 Pattanaik, AL8 Tumblin-Rushmeier and AL2 Ferwerda* had the lowest performance. The operators that performed well, follows the functionality of the human visual system without attempting to construct its sophisticated model [10]. The output of the HDR-VDP-2 is a map of probability for detecting visible differences between HDR image and its corresponding LDR image, i.e., each pixel position has a corresponding probability that any visual difference can be observed [18]. An algorithm may be considered to produce a better visual quality if its HDR-VDP-2 maps contain more pixels with a lower probability of detecting difference [7].

TABLE II. PERFORMACE INDEX BETWEEN SUBJECTIVE TESTS

Test type	SRCC	PLCC	KRCC
FR Mobile/FR Computer	0.8424	0.8817	0.6444
NR Mobile/NR Computer	0.8788	0.9118	0.7889
FR Mobile/NR Mobile	**0.9758**	**0.9934**	**0.9111**
FR Computer /NR Computer	0.8788	0.8918	0.6889

TABLE III. PERFORMACE INDEX BETWEEN SUBJECTIVE TESTS AND METRICS

Test type	Metrics	SRCC	PLCC	KRCC
FR mobile	***HDR-VD-2***	0.8424	0.8923	0.6444
	mPSNR	0.7455	0.8275	0.5556
	Entropy	**0.9394**	**0.9754**	**0.8222**
NR mobile	***HDR-VD-2***	0.8667	0.9038	0.7333
	mPSNR	0.8303	0.8307	0.6444
	Entropy	**0.9879**	**0.9922**	**0.9556**
FR Computer	***HDR-VD-2***	**0.903**	**0.9411**	**0.7333**
	mPSNR	0.8182	0.8397	0.6444
	Entropy	0.8545	0.9006	0.6521
NR Computer	***HDR-VD-2***	**0.9636**	**0.9695**	**0.8667**
	mPSNR	0.8622	0.7333	0.5111
	Entropy	0.8909	0.8902	0.7443

VI. PERFORMANCE INDEXES

Three performance indexes have been used in this paper, the Pearson linear Correlation Coefficient (PLCC) is used to measure the degree of association between two variables. The Spearman Rank Order Correlation Coefficient (SRCC) measures the prediction monotonicity of a metric. Kendall rank correlation coefficient (KRCC) evaluates the degree of similarity between two. A value close to 1 for SRCC, PLCC and KRCC indicate superior correlation [2]. We have calculated the correlation between the MOS for the four tests as shown in TABLE II. The highest correlation is between FR Mobile/NR Mobile, we can recap that observers preferred mobile experiment TM images rather than the computer ones. Moreover, NR Mobile/NR Computer came with the second highest results that's mean that the NR test was preferable than from the FR test. Performance indexes further confirmed that overall there was a significant difference between MOS of the two tests and the objective quality metrics TABLE III. In the mobile test the best correlation came between Shannon Entropy and the MOS, especially with the NR test. While the computer test HDR-VDP-2 came with higher correlation and also the NR test performed better.

VII. CONCLUSION

Nowadays, mobile devices are becoming the main platform for the consumption of multimedia and the rapid increase in the number of such devices in use emphasizes the need to ensure optimal visual quality of experience for the end user when viewing HDR content. The main goal of this work is to understand if the different SDR displays size and resolution have an impact on how the TMO can accurately reproduce an HDR scene in uncontrolled viewing conditions.

From our results we found that SSDs' gave better subjective results than the LDR displays with different TMOs. Moreover, for the mobile test in FR and NR, higher resolution gave more favorable MOS results. There is no vast difference in the performance of mobile devices' in uncontrolled viewing conditions for HDR image evaluation if it was NR or FR tests.While in LDR, there was large differences between the FR and NR results, with a higher MOS for FR test.

The performance indices further confirmed that overall there was a significant difference between the MOS of the four tests, while the highest correlation was between FR Mobile/NR Mobile results. This implies that the uncontrolled methodology is the best way to view HDR content on SSD's. In the mobile test the highest correlation was between Entropy and the MOS for the NR test. Entropy is a measure of randomness and can be used to measure the details in the TM image. It provides information about the number of saturated pixels and the contrast of the HDR image. Since these parameters are the most important for TM algorithms, TM algorithms can be evaluated by means of entropy.

For the future work, we will focus on extensive subjective tests with larger datasets. Moreover, we will study the impact of HDR imaging on human visual system and investigate and quantify how TMOs modify the QoE by using electroencephalography (EEG) responses [20].

ACKNOWLEDGMENT

The first author would like to thank the Ministry of Higher Education and Scientific Research (MoHESR) –Iraq for their financial support. The first author also would like to thank Dr. E. Jammeh and Dr. L. Anegekuh in Centre for Signal Processing and Multimedia Communication, Plymouth University for their valuable comments and support.

REFERENCES

[1] F. Banterle, A. Artusi and K. Debattista, "Tone Mapping", In *Advanced High Dynamic Range Imaging: Theory and Practice*, CRC Press, (AK Peters Ltd), 2011.

[2] R. Mantiuk, K. Myszkowski and H.P. Seidel, "High Dynamic Range Imaging", Wiley, Encyclopaedia of Electrical and Electronics Engineering, 2015.

[3] C. Urbano, L.Magalhães, J.Moura, M.Bessa, A.Marcos, and A.Chalmers, " Tone Mapping Operators on Small Screen Devices: An Evaluation Study ".In *Computer Graphics Forum*, 2010, 29: pp.2469–2478.

[4] M. Melo, M. Bessa, K. Debattista and A.Chalmers, "Evaluation of HDR video tone mapping for mobile devices". In *Signal Processing: Image Communication*, *29*, Vol.2, February 2014, 247–256.

[5] Cisco Visual Networking."Global Mobile Data Traffic Forecast Update", *White Paper* c11-520862, 2014–2019.."

[6] M. Narwaria,M.P. Da Silva and P. Le Callet "High dynamic range visual quality of experience measurement: Challenges and perspectives", *Visual Signal Quality Assessment, Springer International Publishing*, pp. 129-155, 2015

[7] P. Hanhart, M. Bernardo, P. Korshunov, M. Pereira, A. Pinheiro, and T. Ebrahimi, "HDR image compression: a new challenge for objective quality metrics," in *Sixth International Workshop on Quality of Multimedia Experience (QoMEX)*, September 2014.

[8] M. Cadík, M. Wimmer, L. Neumann and A. Artusi,"Evaluation of HDR tone mapping methods using essential perceptual attributes", in *Computers & Graphics*, pp.330-349, 2008

[9] Y. Salih, A. S. Malik, N. Saad, et al., "Tone mapping of hdr images: A review," in *IEEE ICIAS*, vol. 1, 2012, pp. 368–373.

[10] B. Gözde, "Image Dynamic Range Enhancment", PhD thesis, Middle East Technical University, 2011.

[11] Processed HDR images [Online]. Available.
www.tech.plymouth.ac.uk/see/dcee/research/spmc/staff/ssal-juboori/

[12] F. Banterle, HDR Toolbox [Online].Available.
www.github.com/banterle/HDR_Toolbox

[13] J. Kuang, G. Johnson, and M. Fairchild, "iCAM06: A Refined Image Appearance Model for HDR Image Rendering," *J. Visual Comm. and Image Representation*, vol. 18, pp. 406-414, 2007.

[14] S. Corchs, F.Gasparini, R.Schettini, " Noisy images-jpeg compressed: subjective and objective image quality evaluation". *IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics* 2014, pp. 90160V-90160V.

[15] ITU-R BT.500-13, " Methodology for the subjective assessment of the quality of television pictures", *International Telecommunication Union*, January 2012.

[16] G. Valenzise, F. De Simone, P. Lauga, and F. Dufaux, "Performance evaluation of objective quality metrics for HDR image compression," *in Proc. SPIE 9217, Applications of Digital Image Processing* XXXVII, Aug. 2014.

[17] C. Xim, A. Párraga, and X. Otazu. " Which tone-mapping operator is the best? A comparative study of perceptual quality." *Cerdá-Company, arXiv preprint arXiv:*1601.04450, 2016.

[18] H. Yeganeh and Z. Wang. "Objective quality assessment of tone mapped images." *IEEE Transactions on Image Processing*, 22(2):657–67, 2013.

[19] P. Korshunov, T. Ebrahimi." Influence of Context and Content on Tone-mapping" *In HDRi2013-First International Conference and SME Workshop on HDR imaging,* 2013.

[20] S.E. Moon and J.S. Lee. "Perceptual experience analysis for tone-mapped HDR videos based on EEG and peripheral physiological signals". *IEEE Transactions on Autonomous Mental Development*, 7(3), pp.236-247. 2015