Comments on Human Visual Attention in High Dynamic Range Images

Ana Gavrovska, *Member, IEEE*, Irini Reljin, *Senior Member, IEEE,* Andreja Samčović, Valery Starovoitov, *Member, IEEE,* and Milan Milivojević, *Student Member, IEEE*

***Abstract*— High dynamic range images tend to provide more details in both dark and light areas. By introducing a high range of luminance, high quality detailed images may bring more realistic scenes. On the other hand, the change in the experience of the images may be found in comparison to traditional images. It is still not clear how high dynamic range may affect human visual attention, and that information can be relevant in different applications including compression. Fixation Density Map images can be helpful in understanding the human visual attention. In this paper, a comparison experiment with the map images is performed based on large deviation analysis for high dynamic range images.**

***Index Terms*— High dynamic range, image statistics, human visual attention, fixation density map, deviation.**

I. INTRODUCTION

HIGH Dynamic Range (HDR) imaging is one of the leading technologies with the potential of bringing high quality of experience to consumers. With the goal of capturing scenes with a high(er) range of luminance, high dynamic range brings new enhancements in details and color in comparison to traditional reproduction devices (cameras, displays) that correspond to low or standard dynamic range [1]. This should bring new settings for the Quality of Experience (QoE).

Lately, besides the compression standards and extensions dealing with high dynamic range (like in high efficiency video coding or JPEG-XT [2, 3]), there have been some additions to ITU-R (*International Telecommunication Union*) Recommendation BT.2020. The BT.2100 extension (BT.2020/BT.2100) includes some of the reference values for high dynamic range content and defines transfer functions for making the relation between the scene light and the signal [4].

Besides the objective evaluation experiments followed by subjective analysis, there is an increasing number of experiments combining the two abovementioned. Judging the way how the new imaging technologies affect the way our visual system reacts is considered very important in order to provide adequate extensions to already available standards related to distribution, manipulation and similar. This is especially important for the compression tasks [2, 5].

Besides Quality of Service (QoS), one of the main issues today is obtaining the high level of QoE. It is not easy to predict how an end-user will react on the new delivered content [5, 6, 7]. There may be similarities among the new imagery and the traditional ones based on the selection of different methods such as tone mapping [8]. Even though, based on perceptual experiments it is believed that new upgrading methods can bring improved results (e.g. new tone mapping operators [9]). Image or frame content may also affect the selection of different techniques, as in SAR (Synthetic Aperture Radar) images, where high intensities can be the result of perceiving specific objects [10].

In this paper a comparison experiment is performed based on singularity spectrum analysis [11, 12, 13, 14]. In particular, large deviation is considered to analyze the details related to high dynamic range images. Introducing high dynamic imagery may have different influence on human visual attention based on the content analyzed. Eye tracking methods are often used to monitor such influence Some of the publicly available benchmarking datasets are presented in [15, 16, 17]. Here, spectral and visual attention analysis is performed using one of them [16].

This paper is organized as follows. After the Introduction, in Section II we briefly describe high dynamic range images and eye-tracking used for human visual attention analysis. Material and methods that are used in simulation are described in Section III. In Section IV experimental setup is explained. Obtained simulation results followed by adequate discussion are given in Section V. Finally, Section VI brings conclusions.

II. HIGH DYNAMIC RANGE AND EYE-TRACKING

There exist numerous approaches how to obtain high dynamic range, where mostly the quality is enhanced by using a larger range of brightness or by enlarging the color gamut. Multi-exposure techniques are quite often applied to enhance details in both dark and light areas [1, 2]. One should have in mind avoiding artifacts and intensifying unwanted high frequency components in compression of these high quality still images or videos.

Tone-mapping operators are applied in order to adapt high dynamic range imagery to already available systems and techniques. They are used for luminance compression

Ana Gavrovska is with the Department of Telecommunications, School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (e-mail: anaga777@gmail.com; anaga777@etf.rs).

Irini Reljin is with the Department of Telecommunications, School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (e-mail: irini@etf.bg.ac.rs).

Andreja Samčović is with the Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11000 Belgrade, Serbia (e-mail: andrej@sf.bg.ac.rs).

Valery Starovoitov is with the United Institute of Informatics Problems, Minsk, Belarus (e-mail: valerystar@mail.ru).

Milan Milivojević is with the Department of Telecommunications, School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11020 Belgrade, Serbia (e-mail: msmilance@etf.rs).

resulting in improved contrast ratio. Inverse operators are also considered to adapt the old media to new high dynamic range channels.

There are numerous tone-mapping operators available. One of the most often applied is Reinhard operator. Its popularity is due to better subjective evaluation for a large number of applications. Such operators are necessary for reproduction of high dynamic range image on a standard display, as it is performed in Fig. 1. The indoor example shows that new details in overexposed area (window). Nevertheless, the selection of proper tone-mapping operator for a specific purpose is a challenge. As it can be seen in Fig. 2 the details in light area are rather compressed meaning the smoothing out of a high intensity level range is noticeable (data labeled as case I6).

Figure 1. Low dynamic (upper) and high dynamic range grayscale image tone mapped by the Reinhard approach (lower).

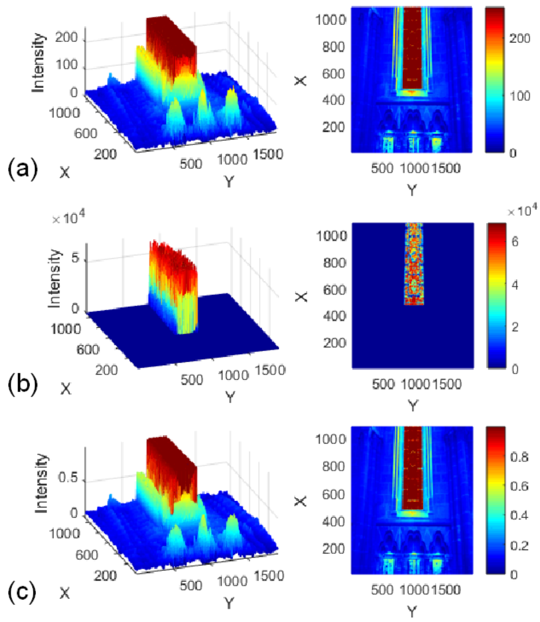

Figure 2. Illustration of (a) low dynamic range image, and high dynamic range image (b) before and (c) after tone-mapping.

High dynamic range images capture richer information closer to the way how human visual system perceives details and adapts to the light.

Information from a captured scene may be reproduced differently from the aspect of visual attention. Namely, eye tracking approaches may reveal different pattern of how high dynamic images change the traditional experience. Generally, there is a tendency to bright objects or areas in an image, as well as to human faces. It is believed that the degree of consumer satisfaction can be modeled according to objective metrics. By recording the gaze points, fixation density maps (FDMs) can be generated and used for further investigating the possible patterns [15, 17].

III. MATERIAL AND METHODS

The human visual attention data are taken from the HDR-Eye dataset [16]. In the dataset a wide variety of content is used covering both indoor and outdoor scenes. It consists of forty six high dynamic range images accompanied with low dynamic images and FDMs for both cases. Each FDM represents an image obtained from recorded gaze points collected from twenty participants (age range 18-56). The system included the Smart Eye Pro 5.8 tracker model and 47" HDR SIM2. The gathered gaze points are used in mapping the fixation density with Gaussian filter taking into account visual angle. Some of the examples of FDMs are presented in Fig. 3 (noted as case I10). FDMs can easily be differentiated in the presented example since a new component can be observed (door area).

Figure 3. Fixation density maps (right) for the low and the high dynamic range image (left) [16] - example of two patterns.

A large deviation technique from [14] is applied for evaluation of the singularities found in images/maps, and calculating multifractal spectrum in, so called, alpha domain. For evaluation of visual attention data, three metrics are calculated. The distribution metric used for evaluation of FDMs is the similarity score from [17]:

$$S = \sum_{i,j} \min\left(L(i,j), H(i,j)\right), \sum_{i,j} L(i,j) = \sum_{i,j} H(i,j) = 1 \quad (1)$$

between low and high dynamic range map images, respectively denoted as L and H. The similarity is one, when the maps are the same, and zero if they do not overlap. The second calculated metric is a Pearson correlation coefficient C between the FDMs for low and high dynamic range image, FDM_L and FDM_H. The third metric is the ratio between the corresponding entropy values:

$$E = entropy(FDM_L) / entropy(FDM_H). \quad (2)$$

IV. Simulation

The simulation is performed using three sets of data. In the first one, six subsets (containing low and high dynamic range image, FDM_L and FDM_H images) related to capturing different scenes are gathered. These are mostly indoor acquired images (except I10 case). The second set is consisted of three subsets/images of similar scene (church), while the third set contains subsets/three images containing characteristic texture areas (walls). All of the numerations are from the publicly available dataset.

Figure 4. Representatives from three sets per column (presented data per column: [I7, I16, I19]; [I37, I38, I39]; [I22, I23, I25]).

The large deviation approach is applied for calculating the spectrum values which are compared under the same circumstances. Moreover, the values are aligned for the comparison reasons (maximum at the same point). In the first step, the comparison is performed for the low and high dynamic range images. In the second step, the experiment is repeated for the FDMs. Finally, the metrics *S*, *C* and *E* are calculated.

V. Results and Discussion

For the purpose of differentiation of the calculated spectrum values red and blue lines are used for low and high dynamic range images, respectively. Similarly, red and blue lines are used for the corresponding FDM images.

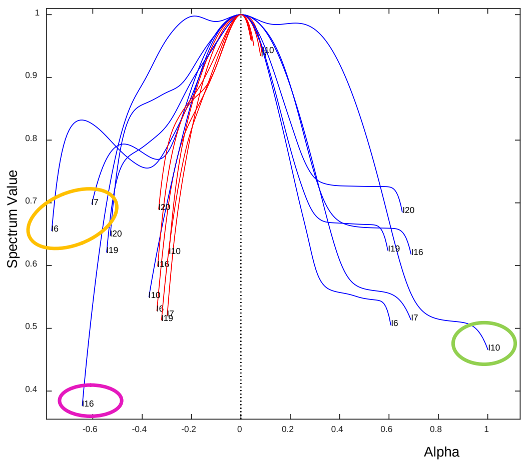

Figure 5. Calculated spectra for the first set.

The results of the first simulation step are presented in Fig. 5 and Fig. 6. It is obvious that all of the high dynamic range images show multifractal behavior, where a much wider spectra is obtained in comparison to traditional images. The left part of diagrams may show some elevations presenting additional processes or pixel-based clusters (cases circled orange, Fig. 5). Particularly interesting is I16 case, where new high value components are found near maximum. The only outdoor image (I10) has more details that can easily detected on the right side of the spectrum (noted green, Fig. 5). The results of the second simulation step are presented in Fig. 7 and Fig. 8. Similar atypical behavior with I10 is noticed for FDM in Fig. 7, where higher values can be found for I10, I16, I19. For I10 both red and blue diagrams are closely detected.

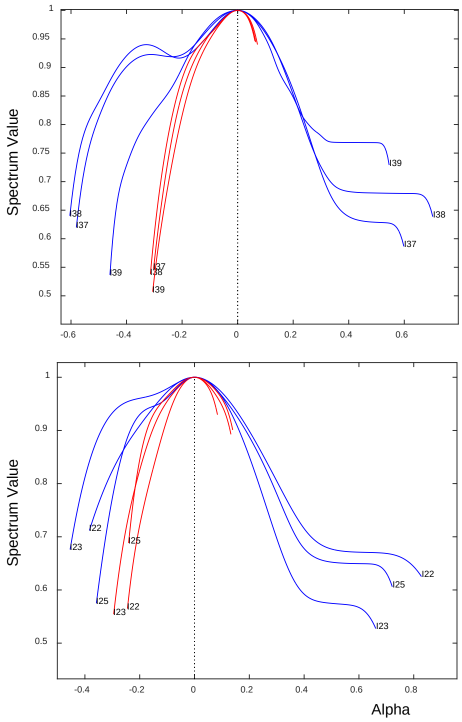

Figure 6. Spectrum values comparison for the second (upper) and the third set (lower).

TABLE I

FDM EVALUATION FOR THE THREE SETS

Set	***Case***	*Similarity S*	*Corr. C*	*Entr. ratio E*
1	I6	0.7976	0.9485	1.0207
	I7	0.7995	0.9026	0.8970
	I10	0.7017	0.8294	**1.4564**
	I16	**0.5448**	**0.5279**	0.8709
	I19	0.7808	0.8641	1.2007
	I20	0.6817	0.6848	0.6501
2	I37	0.7969	0.8549	**1.0136**
	I38	0.7431	0.8169	0.8020
	I39	0.7268	0.7465	0.8059
3	I22	0.7627	0.8414	0.7026
	I23	0.7584	0.8357	**0.8731**
	I25	0.7730	0.8668	0.8008

In Table I FDM evaluation results show that entropy ratio seems as a good metric to describe the behavior found in spectrum (Fig.7 and Fig.8). Namely, in each set the most noticeable difference between FDMs (or corresponding spectra) are found for highest *E* (usually for *E* higher than 1). Case I16 shows additional attention due to structures next to window (Fig. 4), where *S* and *C* values are particularly low.

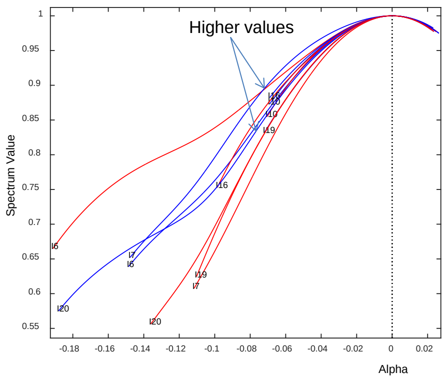

Figure 7. FDM spectrum values for the first set of images.

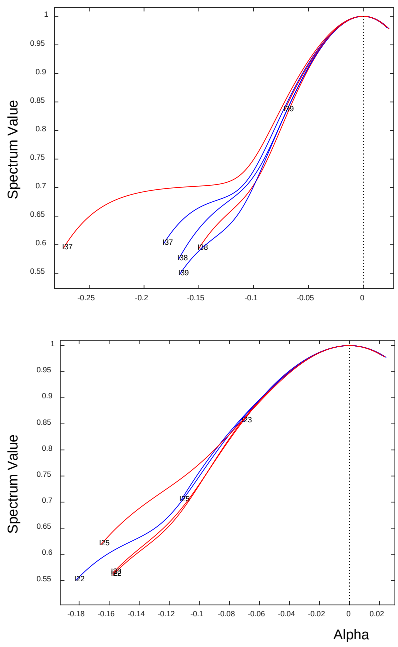

Figure 8. FDM spectrum values comparison for the second (upper) and the third set (lower).

VI. Conclusion

An experiment using FDMs is performed to make some evaluations of high dynamic range content. Such analysis results can be valuable while attempting to preserve visual results and compress data. It is shown that the distribution of human visual attention is not the same for low and high dynamic range images. Namely, we confirm that there are differences between the spectral content for low and high dynamic range images, where for high dynamic range images wider spectrum is obtained in comparison to traditional images. The results show the possibilities of further evaluating the new content, which will be continued in our future work.

Acknowledgment

The research presented in this paper is partially funded by Ministry of Education, Science and Technological Development of the Republic of Serbia as a project TR32048, and Bilateral Project with Republic of Belarus on HDR image compression.

References

[1] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced high dynamic range imaging: theory and practice*, CRC press, 2011.

[2] A. Gavrovska, and A. Samčović, “Mogućnosti unapređenja kvaliteta slike i videa visokog dinamičkog opsega”, PosTel, Beograd, str. 265-274, 29-30. novembar, 2016.

[3] P. Korshunov, P. Hanhart, T. Richter, A. Artusi, R. Mantiuk, and T. Ebrahimi, “Subjective quality assessment database of HDR images compressed with JPEG XT,” In *Quality of Multimedia Experience (QoMEX)*, pp. 1-6, May 2015.

[4] BT.2100: Image parameter values for high dynamic range television for use in production and international programme exchange, 2016.

[5] M. Rerabek, P. Hanhart, P. Korshunov, and T. Ebrahimi, “Subjective and objective evaluation of HDR video compression,” In *9th International Workshop on Video Processing and Quality Metrics for Consumer Electronics* (VPQM), Chandler, Arizona, USA, 5-6 February, 2015.

[6] R. Mukherjee, K. Debattista, T. Bashford-Rogers, B. Waterfield, and A. Chalmers, “A study on user preference of high dynamic range over low dynamic range video,” *The Visual Computer*, *32*(6-8), pp.825-834, 2016.

[7] R. Mukherjee, K. Debattista, T. Bashford-Rogers, P. Vangorp, R. Mantiuk, M. Bessa, B. Waterfield, and A. Chalmers, “Objective and subjective evaluation of high dynamic range video compression,” *Signal Processing: Image Communication*, *47*, pp.426-437, 2016.

[8] M. Barkowsky, and P. Le Callet, “On the perceptual similarity of realistic looking tone mapped high dynamic range images,” *17th IEEE International Conference on Image Processing (ICIP)*, pp. 3245-3248, IEEE, Hong Kong, 26-29 September 2010.

[9] J. Ok, and C. Lee, “HDR tone mapping algorithm based on difference compression with adaptive reference values,” *Journal of Visual Communication and Image Representation*, *43*, pp. 61-76, 2017.

[10] V. Starovoitov, “Comparative Analysis of Some Dynamic Range Reduction Methods for SAR Image Visualization,” In *International Conference on Pattern Recognition and Information Processing*, pp. 63-76, Springer, Cham, 2016.

[11] Paskaš, M., Gavrovska, A., Jevtić, D., Slavković, M. and Reljin, B., “Edge examination using Hölder exponent and image statistics,” In *10th Int. Conf. on Telecommunication in Modern Satellite Cable and Broadcasting Services (TELSIKS)*, Vol. 1, pp. 329-332, IEEE, Nis, Serbia, 5-8 October 2011.

[12] O.V. Spirintseva, “The Multifractal Analysis Approach for Photogrammetric Image Edge Detection,” *Int. Jour. of Image, Graphics and Signal Processing (IJIGSP)*, *8*(12), p.1. 2016.

[13] J.L. Véhel, C. Tricot, C., “On various multifractal spectra,” In *Fractal Geometry and Stochastics III*, Birkhäuser, pp. 23-42, Basel, 2004.

[14] C. Christophe, J.L. Véhel, and C. Tricot, “Continuous large deviation multifractal spectrum: definition and estimation," In *Fractals 98*,World Scientific, Malta, October 1998.

[15] Y. Dong, E. Nasiopoulos, M.T. Pourazad, and P. Nasiopoulos, “High dynamic range video eye tracking dataset,” In *Proc. 2nd Int. Conf. Electron., Signal Process. Commun.*, pp. 56-59, 2014.

[16] HDR-Eye: dataset of high dynamic range images with eye tracking data, http://mmspg.epfl.ch/hdr-eye (*last accessed* 14.03.2017.)

[17] H. Nemoto, P. Korshunov, P. Hanhart, and T. Ebrahimi, “Visual attention in LDR and HDR images,” In *9th Int. Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, Chandler, Arizona, USA, 5-6 February, 2015.