Signal Processing: *Image Communication* 29 (2014) 257–273

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Tone mapping based HDR compression: Does it affect visual experience?

Manish Narwaria, Matthieu Perreira Da Silva*, Patrick Le Callet, Romuald Pepion

LUNAM University - IRCCyN CNRS UMR 6597, France

ARTICLE INFO

Available online 11 October 2013

Keywords:
High dynamic range (HDR)
Visual attention
Eye-tracking
Non-parametric analysis

ABSTRACT

High Dynamic Range (HDR) content is visually more appealing since it can represent the real luminance of the scene. However, on the downside, this means that a large amount of data needs to be handled both during storage and processing. The other problem is that HDR content cannot be displayed on the conventional display devices due to their limited dynamic range. To overcome these two problems, dynamic range compression (or range reduction) is often used and this is accomplished by tone mapping operators (TMOs). As result of tone mapping, the HDR content is not only fit to be displayed on a regular display device but also compressed. However from an artistic intention point of view, TMOs are not necessarily transparent and might induce different viewing behavior. It is generally accepted that TMOs reduce visual quality and there have been a number of studies reported in literature which examine the impact of tone mapping from the view point of perceptual quality. In contrast to this, it is largely unclear if tone mapping will induce changes in visual attention (VA) as well and whether these are significant enough to be accounted for in HDR content processing. To our knowledge, no systematic study exists which sheds light on this issue. Given that VA is a crucial visual perception mechanism which affects the way we perceive visual signals, it is important to study the effect of tone mapping on VA deployment. Towards this goal, this paper investigates and quantifies how TMOs modify VA. Comprehensive subjective tests in the form of eye-tracking experiments have been conducted on several HDR content and using a large number of TMOs. Further non-parametric statistical analysis has been carried out to ascertain the statistical significance of the results obtained. Our studies suggest that TMOs can indeed modify human attention and fixation behavior. Based on this we believe that VA needs consideration for evaluating the overall perceptual impact of TMOs on HDR content. As mentioned, since the existing studies so far have only considered the *quality* or *esthetic appeal* angle, this study brings in a new perspective regarding the importance of VA in HDR content processing for visualization on LDR displays.

1. Introduction

High Dynamic Range Imaging (HDRI) has been steadily gaining popularity in both academia and industry [1]. The reason being that HDR content is visually more appealing and realistic as they can represent the dynamic range of the visual stimuli present in the real world. HDR faithfully depicts the dynamic range of the real world luminance (typically varying from 10^{-4} cd/m^2 to 10^6 cd/m²) by storing them as floating point values. The low dynamic range (LDR) images/videos on the other hand store the scene intensities as integer pixel values (typically in the range [0 255]) which represent colors that should appear on a display device and not necessarily proportional to scene

* Correspondence to: Polytech Nantes-Computer Science Department, La Chantrerie-rue Christian Pauc-BP 50609, 44306 Nantes Cedex 3-France. Tel.: +33 2 40 68 30 60; fax: +33 2 40 68 32 32.

E-mail address: matthieu.perreiradasilva@univ-nantes.fr (M. Perreira Da Silva).

0923-5965/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.image.2013.09.005

intensities. This is why LDR is referred to as device-referred or output-referred. Due to this, the areas that are too dark are clipped to black (0) and areas that are too bright are clipped to white (255). Obviously, this leads to loss of visual details and contrast. The floating point representation in HDR overcomes this since the dark and bright areas are assigned values proportional to the actual scene intensity instead of being saturated in an ad-hoc manner as is the case with LDR. So HDR content is scene-referred. As a result, an HDR still image (or HDR video sequence) can capture very high contrasts which in turn enables it to incorporate maximum details that the human eye can discern. Even though HDRI offers obvious advantages over the traditional LDR content in terms of better visual quality of experience (QoE), its large scale deployment at consumer levels is severely hindered due to two major issues. The first one stems from the fact that an HDR file requires large amount of storage space as compared to an LDR file. For instance, an HDR still image may occupy 4 times the space needed by an LDR version of the same image [1]. This, as a consequence, demands for effective compression algorithms for HDR content. Note that compression of LDR images/videos has been an active research area in any case and the advent of HDRI provides a further impetus to this field. The second issue is that HDR content cannot be displayed directly onto conventional display devices as these cannot provide the necessary luminance range (usually their range lies between 1 and 300 cd/m^2) for a true HDR experience. Additionally, their contrast ratio is not good enough for displaying HDR content. For example, even a good In-Plane Switching (IPS) LCD panel can achieve a contrast ratio of only about 1000:1 while the required contrast ratio of typical HDR scenes can be more than 10^6:1.

Fortunately, the two mentioned issues can be tackled at least in the theoretical sense by a single approach namely tone mapping. Tone mapping refers to the reduction of the dynamic range of the HDR content in such a way that maximum perceivable visual information is retained (this is what compresses HDR as elaborated further in Section 3). Not surprisingly, a large amount of research effort have been directed towards the development of tone mapping operators (TMOs) and several of them have been devised in literature [2–12]. Some are simple and based on operations such as linear scaling and clipping while the more sophisticated ones exploit several properties of the Human Visual System (HVS) with the aim of preserving the details. But more often than not, TMOs lead to information loss which can reduce the perceptual quality of the tone mapped content. This is expected since dynamic range compression invariably tends to destroy important details and textures and can introduce additional artifacts related to changes in contrast and brightness.

The remainder of this paper is organized as follows. Section 2 provides a brief overview of some of the existing TMOs. Section 3 discusses HDR content compression via tone mapping. In Section 4, we present the details of the subjective experiments including the test set up, conditions, test scenes, etc. Next, we give detailed analysis of the experimental findings in Section 5. Quantitative comparison is detailed in Section 6 while the non-parametric statistical analysis has been presented in Section 7. In Section 8, we put forth our views on the connection between VA and quality and finally Section 9 gives the concluding remarks.

2. Tone mapping operators: brief overview

TMOs can be broadly classified into two categories namely local operators and global operators. As the name implies, local operators employ a spatially varying mapping which depends on the local image content. As opposed to this, global operators use the same mapping function for the whole image. Chiu et al. [2] introduced one of the first local TMOs by employing a local intensity function based on a low-pass filter to scale the local pixel values. The method proposed by Fattal et al. [3] is based on compressing the magnitudes of large gradients and solves the Poisson equation on the modified gradient field to obtain tone mapped images. Durand et al. [4] presented a TMO based on the assumption that an HDR still image can be decomposed into a base image (mainly low frequency information) and a detail image (mainly high frequency details). This is done using the bilateral filter. Then, the contrast of the base layer is reduced via a scale factor. The tone mapped image is obtained as a result of multiplication of the contrast reduced base layer with the detail image. Drago et al. [5] adopted logarithmic compression of the luminance values for dynamic range reduction in HDR still images. They use adaptively varying logarithmic bases in order to preserve local details and contrast. The TMO proposed by Ashikimin [6] first estimates the local adaptation luminance at each point which is then compressed using a simple mapping function. In the second stage, the details lost in the first stage are re-introduced to obtain the final tone mapped image. Reinhard et al. [7] applied the dodging and burning technique (traditionally used in photography) for dynamic range compression. A TMO based on a perceptual framework for contrast processing in HDR still images was introduced by Mantiuk et al. [8]. This operator involves the transformation of an image from luminance to a pyramid of low-pass contrast images and then to the visual response space. It was claimed that in this framework, dynamic range reduction can be achieved by a simple scaling of the input. Mantiuk et al. [9] also proposed another TMO by formulating tone mapping as an optimization problem: one that minimizes the visible contrast distortions between a human visual system and the display. Another TMO known as iCAM06 [10] has also been developed. It is based on the sophisticated image color appearance model (iCAM) and incorporates the spatial processing models in the HVS for contrast enhancement, photoreceptor light adaptation functions that enhance local details in highlights and shadows. Recently, Boschetti et al. [13] proposed a TMO based on the contrast limited adaptive local histogram equalization (CLAHE) algorithm. This was done by adopting a new strategy for excess pixel distribution: instead of distributing the excess pixels equally to all bins they were redistributed only in the neighborhood of the histogram peaks. Further, the clip limit was determined based on mean and variance of local image blocks. With regards to global TMOs, the simplest one is the linear operation in which the maximum input luminance is mapped to the maximum output value (the maximum luminance mapping) or the average luminance mapping (i.e. mapping average input luminance to the