HAL
archives-ouvertes.fr

High Dynamic Range Visual Quality of Experience Measurement: Challenges and Perspectives

Manish Narwaria, Matthieu Perreira da Silva, Patrick Le Callet

To cite this version:

Manish Narwaria, Matthieu Perreira da Silva, Patrick Le Callet. High Dynamic Range Visual Quality of Experience Measurement: Challenges and Perspectives. Visual Signal Quality Assessment - Quality of Experience (QoE), Springer International Publishing, pp.129-155, 2015, 978-3-319-10368-6. <10.1007/978-3-319-10368-6_5>. <hal-01151591>

HAL Id: hal-01151591
https://hal.archives-ouvertes.fr/hal-01151591

Submitted on 13 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

High Dynamic Range Visual Quality of Experience Measurement: Challenges and Perspectives

Manish Narwaria, Matthieu Perreira Da Silva, Patrick Le Callet

1 Abstract

Traditional capture and display devices can only support a limited dynamic range (contrast) and color gamut given the hardware limitations. As a result, the real physical luminance present in a natural scene cannot be captured by these. However, with the recent advancements in the related software and hardware technologies, it is now possible to capture or reproduce higher contrast and luminance ranges. Such scene-referred visual signals are known as High Dynamic Range (HDR) signals. They are visually more appealing because they can represent the dynamic range of the visual stimuli present in the real world. Not surprisingly, the emergence of HDR is seen as an important step towards improving the visual quality of experience (QoE) of the end users. However, HDR comes with its own set of challenges including storage, processing, display and so on. This chapter focuses on some of the issues related to HDR processing from a quality of experience (QoE) viewpoint.

2 Introduction

Humans perceive the outside visual world through the interaction between light energy (usually measured in candela per square meter cd/m^2) and the eyes. Light energy first passes through the cornea, a transparent membrane. Then it enters the pupil, an aperture that is modified by the iris, a muscular diaphragm. Subsequently, light is refracted by the lend and hits the photoreceptors in the retina. There are two types of photoreceptors: cones and rods. The cones are located mostly in the fovea. They are more sensitive at luminance levels between 10^{-2} cd/m^2 to 10^8 cd/m^2 (referred to as the photopic or daylight vision) [7]. Further, color vision is due to three types of cones: short, middle and long wavelength cones. The rods, on the other hand, are sensitive at luminance levels between 10^{-6} cd/m^2 to 10 cd/m^2 (scotopic or night vision). The rods are more sensitive than cones but do not provide color vision. There is only one type of rod photoreceptors and are located around the fovea. Since there are no rods in the fovea, high frequency patterns cannot be distinguished at low lighting conditions [7].

Pertaining to the luminance levels found in the real world, direct sunlight at noon can be of the order in excess of 10^7 cd/m^2 while a starlit night in the range of 10^{-1} cd/m^2. This corresponds to more than 8 orders of magnitude. It is therefore evident that there is a large range of luminance present in different real world scenes.However, given its limitations, the human eye cannot perceive such large range at the same time (instantaneously). In fact, they can perceive about 13 orders of magnitude given sufficient adaptation time [13]. An intuitive example of adaptation is when we arrive in a low lit room on a sunny day. We cannot immediately perceive the visual data in the room and it takes a few minutes before one becomes accustomed (adapted) to new luminance levels. This a remarkable capability to adapt to vastly different luminance conditions. An example comparing the approximate orders of magnitude of the human eye, conventional display and the HDR display is shown in Figure 1. Observe how the traditional LDR display covers only a small range (upto 3 orders of magnitude). On the other hand, HDR displays can match the instantaneous range of the eye of about 5 orders of magnitude.

As for the instantaneous human vision range, it is about 5 orders of magnitude i.e. human eye is capable of dynamically adjusting so that a person can see 5 orders of magnitude throughout the entire range. The conventional display devices however cover only upto 3 orders of magnitude. Consequently, the scenes viewed on typical low dynamic range (LDR) displays have lower contrast and smaller color gamut than what the eyes can perceive. This leads to loss of visual details and in some cases can even lead to misrepresentation of the scene information. To overcome such limitations, High Dynamic Range (HDR) has recently gained popularity in both academia and industry. By representing the scene in terms of physical luminance information, HDR can achieve very high contrasts and a wider color gamut, in effect matching the human instantaneous vision range. Due to allowing more scene information representation, HDR helps to capture very fine details which are otherwise difficult to be retailed with traditional photography. A visual example to illustrate this is shown in Figure 2. The scene in question has very bright sunlight, shadows and other details. With single exposure photograph, we can either retain the information in darker areas (longer exposure time) or the ones in brighter areas (shorter exposure

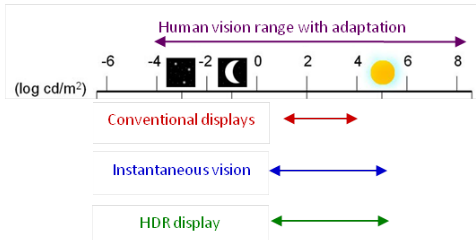


Figure 1: Orders of magnitude of the dynamic range of the eye and displays.

time). In both cases, we tend to lose out information either in dark or bright areas. As shown in this figure, the first two are single exposure images with different exposure values (EV). EV basically controls the amount of light allowed while capturing the scene. The third image is of the same scene but HDR processed (tone mapped to 8-bit precision). The reader will notice that this image preserves more details and has a better overall contrast. In other words, HDR helps to retain visual information in very bright and dark areas. This leads to better visual experience for the viewers.

Quality of Experience (QoE) driven multimedia systems have increasingly come in focus both from research and industry perspectives. The aim to capture the end-users' aesthetic expectations rather than simply delivering content based on a technology-centric approach. HDR is one of the exciting field towards providing the end users a more immersive and realistic viewing experience and thus improving the QoE. The aim of this chapter is therefore to provide the reader an overview of HDR from the view point of visual experience and in the process outline the challenges that exist.

3 The HDR Pipeline

Owing to the characteristics of the HDR signal, its processing right from generation to transmission requires specific tools and different approaches than the traditional LDR processing. A simplified block diagram of a typical HDR processing pipeline is shown in Figure 3. HDR content is first generated by a convenient method. It then needs to be stored appropriately (eg. after compression). The next step involves suitable processing (eg. pre-processing) to transmit it to the end user. The end user can view the processed HDR content directly on an HDR monitor. In case of unavailability of an HDR monitor, a further operation referred to as tone mapping needs to be carried out to make the HDR suitable to be displayed on

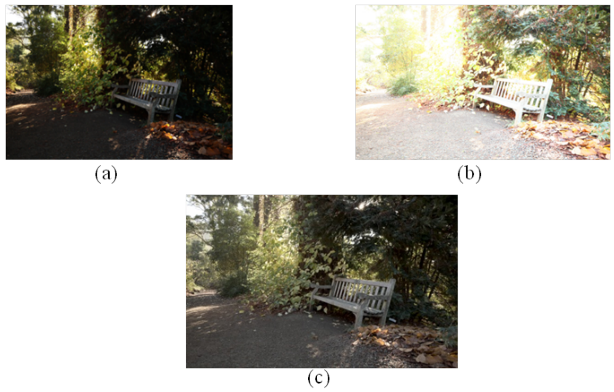


Figure 2: Advantage of HDR over traditional photography. (a) Single exposure (-6.89 EV), (b) Single exposure (-2.89 EV) and (c) HDR processed image (8-bit precision)

more conventional LDR monitors. An important distinction that should be made at this point is regarding the usage of the terms 'content' and 'scene'. Throughout the chapter, we will use them in a generic sense in that they can refer to both still images and videos. While capturing HDR still images and videos follows similar approaches, video involves the additional temporal dimension. This introduces more factors that need careful consideration while capturing HDR video.

3.1 Capture

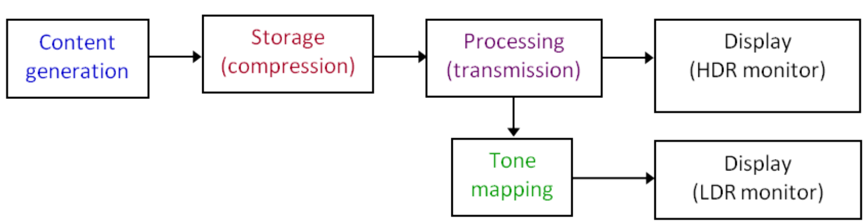


Figure 3: A simplified overview of an HDR processing pipeline

At present, there are several methods to generate HDR content [7]. We will briefly discuss three of them in the following. The first among them employs a weighted fusion of LDR images captured at different exposure levels. Most of the currently available consumer cameras capture 8-bit images (or 14-bit in RAW format). The limited bit-depth is not sufficient to represent the entire dynamic range of a typical real-world scene. Therefore, the authors in [37] proposed the idea of multi-exposure fusion of LDR photographs of the same scene. The underling goal is to incorporate details from each exposure (from brightest to darkest scene areas) and thus obtain a more realistic appearance of the scene. A visual example is illustrated in Figure 4. However, there are two major issues in capturing and fusing multi-exposure LDR images/frames. First, it is cumbersome since one has to manually obtain several photographs from the same scene by varying the shutter speeds. This is further complicated by the fact that different scenes may require different shutter speeds depending on the dynamic range of the scene in question. The second and more serious issue is with regards to pixel alignment and motion. Thus, utmost care is required while capturing (eg. using a tripod stand) the LDR photograph. Since the idea is to capture the same scene with varying exposures, motion or instantaneous change in the scene (eg. a moving person) will severely hinder the effectiveness of this method. This problem is obviously more pronounced in capturing an HDR video and displaying it in real-time. A typical rate to play back a video is 25 frames per second. So creating 25 HDR frames per second should be the goal of HDR video. Assuming that an HDR image can be created from four exposures, a camera would need to capture 100 exposures per second [5]. In addition to this being a challenge for the image sensor and its data interface, there is only one hundredth of a second left for exposing each image. For this to be enough, a lens with a large aperture and possibly an increased gain is required. It is also likely that the camera or the objects in the scene move while acquiring the sequence of exposures. In order to merge them together, the intermediate camera and scene motion must be compensated. Otherwise there would be motion blur or ghosting artifacts. Motion compensation is a computationally costly step. Once the images are aligned, they can be merged together into an HDR frame. Producing 25 HDR frames per second means that there are only 40 ms of processing time available for each frame. Capturing the low dynamic range exposures, aligning and merging them and then tone mapping the result for display thus needs to be performed within 40 ms. These issues can be better handled by the second method for HDR capture through the use of more specialized cameras. There a few commercially available cameras (eg. SpheronCam HDR by SpheronVR [4]) that have an in-built multi-exposure capturing. Given the recent advances in hardware technologies, it is likely that such cameras will become more common. The third method creates HDR content from virtual environments using physically based renderers. This is more commonly employed in entertainment industries (eg. digital cinema).At this point, the reader is referred

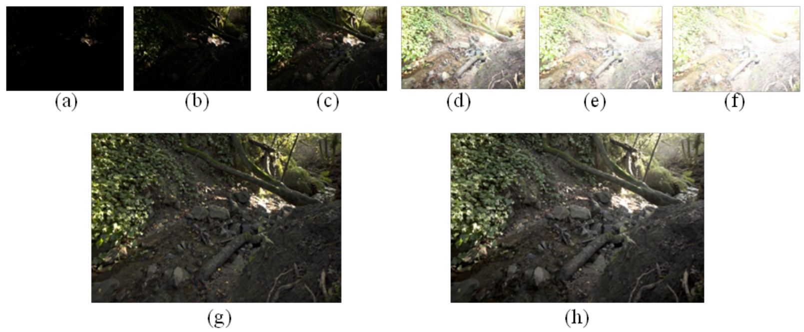


Figure 4: Multi-exposure images and the resulting HDR (one mapped) images, (a) -7.21 EV, (b) -5.47, (c) -3.89 EV, (d) -0.19 EV, (e) 0.76 EV EV, (f) 1.44 EV, HDR image (with luminance range from 10^{-1} cd/m^2 to 10^4 cd/m^2) formed by fusing (a), (b), (c), (d), (e) and (f). (g) and (h) Locally rendered tone mapped images obtained by tone mapping the HDR using two different TMOs

to [7] for further details on the methods for HDR content capture. Given the focus of this chapter, it is assumed that we have a well-captured and realistic HDR scene from either of the mentioned methods and the goal is to process this further keeping in mind the visual appearance to the end user.

3.2 Storage

Once the HDR content is generated, it needs to be stored. As pointed out earlier, an HDR pixel is represented using three single precision floating point numbers. This implies that each pixel requires 12 bytes of memory. A simple computation will reveal that this corresponds to approximately 24 MB of data for high definition (HD) resolution of (1920 by 1080 pixels). In contrast, an equivalent uncompressed LDR representation (24 bits per pixel) of the same scene will require only a fourth (about 6MB) of this memory. It is therefore clear that there is need for efficient compression methods to allow for a more compact HDR storage given the high memory demands.

One of the first solutions towards this was proposed in [14] by the introduction of RGBE. This method stores a shared exponent between the three color channels under the assumption that it does not vary much between them. Consequently, RGBE leads to four bytes per pixel (one byte for each color channel plus one byte for the shared exponent). This immediately reduces the memory requirement to one-third of the uncompressed version (which needs 12 bytes per pixel).Another HDR compression approach proposed in [16] is known as the LogLuv encoding. As the name implies, this format stores the luminance in the logarithmic domain and also assigns more bits to it than to the colors. The underlying principle for LogLuv encoding is that human eyes are more sensitive to luminance information than color (so more bits are devoted to the luminance component). In addition, it has been found that the response of human eyes to the absolute luminance levels is approximately logarithmic. Thus, LogLuv encodes logarithm of luminance (as an additional advantage logarithm operation also help in dynamic range compression). Another common HDR format is the half-floating point format, which is a part of the specification of the OpenExr format [2].

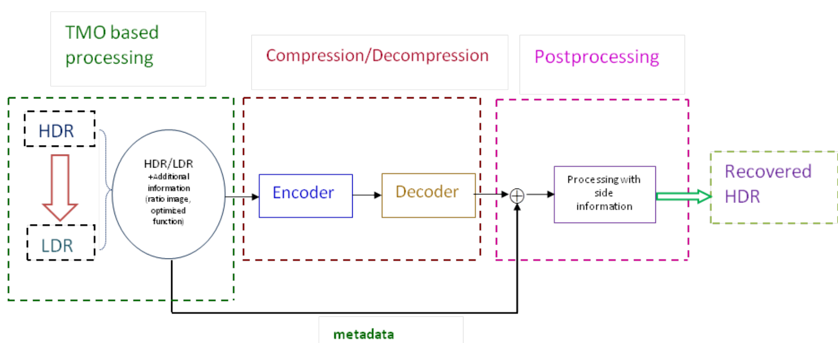


Figure 5: Block diagram of a typical backward compatible HDR compression pipeline.

Despite the existence of efficient formats for HDR storage (like RGBE, OpenExr), there is a need for developing techniques for further HDR compression. This is because even with the HDR formats there is a huge memory requirement. Consequently, HDR content stored in a standard HDR format should be compressed further to enable more practical deployment and real-time processing. So there is need for research into effective HDR compression schemes and this therefore has been an important research area. A crucial and related issue is that the existing coding architectures have become widely adopted standards supported by almost all software and hardware equipment dealing with digital imaging. As a result, it will be of great interest to design HDR compression schemes that are compatible with existing coding architectures. Not surprisingly, substantial research effort has been put into designing HDR compression systems that are backward compatible (for example [17], [30], [33]) with the standard image (e.g. JPEG and JPEG 2000) and video coders (e.g. H.264/AVC). So the aim is to customize the existing codecs so that they can cater to HDR images and videos. Towards that end, dynamic range reduction (or tone mapping) is usually adopted as the first step towards backward compatible HDR compression. In fact, the output of tone mapping is an LDR signal, which requires much smaller memory for storage.

A block diagram showing the steps in a typical backwards compatible HDR compression pipeline is shown in Figure 5. In this figure, we can separate out 3 main blocks. The first one is the Tone Mapping Operator (TMO) based processing module. Here, a TMO is often used to create an LDR version of the HDR image or video frame. Based on the HDR image, the LDR image and/or the TMO configuration, side information (for instance a ratio image as in [17] or a non-linear mapping function as in [30]) that will facilitate the decoder's operation is generated. The second block involves the compression of either the tone mapped LDR content or a modified HDR via an existing compression scheme. The encoded bit streams along with metadata are subsequently transmitted to the decoder. The last block, i.e. the post-processing, performs the inverse tone mapping based upon the side information delivered together with the LDR bit stream, re-converting the decoded LDR image into its HDR format. The reader will notice that the HDR compression pipeline shown in Figure 5 is almost entirely compatible with existing coding architectures. The main difference is the extra metadata that needs to be transmitted for enabling the reconstruction of the HDR image. With such architecture, the focus basically shifts to the first and third blocks namely tone mapping and post processing (or inverse tone mapping). Thus, the problem of HDR compression becomes one of identifying the appropriate tone mapping and inverse tone mapping algorithms. However, this is not straightforward given the fact that it is not easy to convert an HDR image/video to LDR without losing perceivable visual information (related to the difficulty in designing a generic TMO that can retain perceptually useful details). As a result, tone mapping is itself an ongoing research topic while inverse tone mapping is an even less investigated topic with few works addressing it (eg. [8]).

3.3 Visualization

The processed HDR content can be visualized either on the traditional LDR displays or on HDR displays. The former requires some kind of range reduction operation (tone mapping) in order to fit the dynamic range of the LDR display. On the other hand, the latter is a more native and effective approach to HDR visualization. However, this prompts the need for the availability of an HDR display. Display technologies that can natively visualize HDR content without the need for tone mapping are now becoming available. One such display is the SIM2 Solar47 HDR display [34]. The Solar 47 is a 47-inch, 1080p LCD TV with 2202 white LEDs arrayed behind the imaging panel, and unlike other local-dimming sets, each LED is individually addressable. The core technology in this HDR monitor follows the one proposed in [18]. HDR monitors are typically based on two technologies: (a) modulating a Liquid Crystal Display (LCD) panel using a set of powerful Light Emitting Diodes (LEDs) as the backlight, (b) video projector based. Particularly, the SIM2 HDR display [3] uses an LCD panel, replacing its common Back Light Unit (B.L.U.), typically based on a set of Cold Cathode Fluorescent tube Lamps (CCFL), with an array of high-power white LEDs. The idea is to light each small zone of the picture displayed on the LCD, with an LED driven by the specific luminous intensity of that small area of the picture. That means if a scene has black details, the LEDs under those details are turned off to achieve a true black, while where is a high luminous intensity area, LEDs under it are turned on to maximum power. Grey scale areas are then obtained by modulating to intermediate levels the intensity of the LEDs using the HDR picture processing [3].

Even though the HDR display technology is available, it is yet to reach consumer levels. In such scenario, the only alternative is to display HDR contents directly on commonly available devices such as CRT, LCD monitors, printers etc. which have a significantly low dynamic range (LDR). It follows that these cannot provide the necessary luminance range (usually their range lies between 1 to 300 cd/m) for a true HDR experience. Additionally, their contrast ratio is not good enough for displaying HDR contents. 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 106:1. Therefore an important issue in HDRI is to reduce the dynamic range of the HDR content. This problem has been commonly addressed by employing tone mapping operators (TMOs). It is also important to highlight that even the HDR displays cannot reach the luminance levels found in real world and some kind of tone mapping is needed before display the signal.

Thus, tone mapping is an important aspect in HDR processing and display. On the one hand it facilitates the development of backward compatible HDR compression whereby existing coding standards can be exploited for HDR coding, it enables the visualization of HDR on LDR displays, on the other. It is therefore crucial to analyze and understand the impact of tone mapping operators (TMOs) on the overall appearance of the HDR content. This is discussed in the next section.

4 Tone mapping and its impact on visual experience

Tone mapping is the operation that adapts the dynamic range of HDR content to suit the lower dynamic range available on a given display. The idea is to process the HDR content so that the discrepancy between the tone mapped content and the HDR content is minimal from the view point of two observers, one observing the tone mapped content while the other viewing the actual HDR content. Thus, tone mapping attempts to retain important characteristics of the original HDR content such as local and global contrast, details, naturalness etc.

4.1 Tone mapping operators

Several TMOs have been developed over the past years. 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 contents. 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.

TMOs can be broadly classified into 2 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. introduced [22] 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. [32] 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. [11] presented a TMO based on the assumption that an HDR image can be decomposed into a base image and a detail image. The contrast of the base layer is reduced using an edge-preserving filter (known as the bilateral filter). The tone mapped image is obtained as a result of multiplication of the contrast reduced base layer with the detail image. Drago et al. [9] adopted logarithmic compression of the luminance values for dynamic range reduction in HDR images. They use adaptively varying logarithmic bases in order to preserve local details and contrast. The TMO proposed by Ashikimin [23] 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. applied [6] 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 images was introduced by Mantiuk et al. [35]. 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. Another TMO known as iCAM06 [19] 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. 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 average output value). Another global TMO is the one proposed by Ward [15] which focuses on the preservation of perceived contrast. In this method, the scaling factor is derived from a psychophysical contrast sensitivity model. Tumblin et al. [21] have reported a TMO based on the assumption that a real-world observer should be the same as a display observer. These are some of the existing TMOs and the list is by no means exhaustive. The interested reader is also referred to survey papers on the topic (eg. [25]) for a more complete and detailed study of TMOs.

The reader may have noticed that local TMOs seem to have received more attention than the global ones. This is partly due to the fact that as a result of their design local TMOs perform well in preserving the local details (but are less effective in reproducing the overall brightness and contrast). On the other hand, although global TMOs preserve the overall contrast they usually lead to loss of local details. But as an important advantage, global operators are generally computationally more efficient than the local ones. So local and global TMOs have their own advantages and disadvantages. Since tone mapping reduces the dynamic range, it will invariably lead to loss of visual details and as a result affect the perceived appearance of the HDR content. Given that tone mapping is often needed at different stages of HDR pipeline (eg. for compression and visualization), it is therefore necessary to analyze how they affect the visual experience of the processed HDR content. It should be mentioned that evaluating the overall HDR viewing experience is not an easy task since it is a multi-dimensional phenomenon. Nonetheless, we identify three important attributes that play a significant role in the viewing experience: perceptual visual quality, visual attention and naturalness. Therefore, we first analyze how TMOs affect perceptual quality and then discuss their impact on visual attention.

4.2 Tone mapping and visual quality

There have been several studies related to how TMOs affect visual quality of the tone mapped content. We first briefly describe some of the existing studies related to subjective evaluation of TMOs.

The psychophysical experiments carried out by Drago et al. [10] aimed to evaluate six TMOs with regard to similarity and preference. Three perceptual attributes namely apparent image contrast, apparent level of detail (visibility of scene features), and apparent naturalness (the degree to which the image resembled a realistic scene) were investigated. It was found that naturalness and details are important attributes for perceptual evaluation of TMOs. The study by Kuang et al. [20] performed a series of three experiments. The first one aimed to test the performance of TMOs with regard to image preference. For this experiment, 12 HDR images were tone mapped using six different TMOs and evaluation was done using the paired comparison methodology. The second experiment dealt with the criteria (or attributes) observers used to scale image preference. The attributes that were investigated included highlight details, shadow details, overall contrast, sharpness, colorfulness, and the appearance of artifacts. The subsequent regression analysis showed that the rating scale of a single image appearance attribute is often capable of predicting the overall preference. The third experiment was designed to evaluate HDR rendering algorithms for their perceptual accuracy of reproducing the appearance of real-world scenes. To that end, a direct comparison between three HDR real-world scenes and their corresponding rendered images displayed on a low dynamic-range LCD

Figure 6: Different visual qualities of LDR images generated by tone mapping an HDR image by different TMOs, (a) Ashikmin TMO, (b) Drago TMO, (c) Drand TMO, (d) icam06 TMO, (e) linear TMO

monitor was employed. Yoshida et al. conducted [1] psychophysical experiments which involved the comparison between two real-world scenes and their corresponding tone mapped images (obtained by applying 7 different TMOs to the HDR images of those scenes).

Similar to other studies, this one was also aimed at assessing the differences in how tone mapped images are perceived by human observers and was based on four attributes: image naturalness, overall contrast, overall brightness, and detail reproduction in dark and bright image regions. In the experiments conducted [31] by Ledda et al., the subjects were presented three images at a time: the reference HDR image displayed on an HDR display and two tone mapped images viewed on LCD monitors. They had to choose the image closest to the reference. Because an HDR display was used, factors such as controlling screen resolution, dimensions, colorimetry, viewing distance and ambient lighting could be controlled. This is in contrast to using a real-world scene as a reference which might introduce uncontrolled variables. The authors have also reported the statistical analysis of the subjective data with respect to the overall quality and to the reproduction of features and details. Different from the mentioned studies, Cadik et al. adopted [25] both a direct rating (with reference) comparison of the tone mapped images to the real scenes, and a subjective ranking of tone mapped images without a real reference. They further derived an overall image quality estimate by defining a relationship (based on multivariate linear regression) between the attributes: reproduction of brightness, color, contrast, detail and visibility of artifacts. The analysis further revealed that contrast, color and artifacts are the major contributing factors in the overall judgment of the perceptual quality. However, it was also argued that the effect of attributes such as brightness is indirectly incorporated through other attributes. Another conclusion from this study was that there was agreement between the ranking (of two tone mapped images) and rating (with respect to a real scene) experiments. In contrast to this last observation, Ashikimin et al. found [24] that there were significant differences in subjective opinions depending on whether a real scene is used as a reference or not. A recent survey can be found in [12] that evaluated TMOs for HDR video.

It should be emphasized that most of these studies either ranked the TMOs based on the performance in the respective subjective experiments or outlined the factors affecting visual quality of the tone mapped content. However, it might be misleading to generalize the results from these studies since the number of HDR stimuli was limited. Nevertheless, all of them establish beyond doubt that tone mapping (both in still images and videos) tends to not only reduce the visual quality but also affects the naturalness of the processed HDR content (in addition for video stimuli there could be visible temporal artifacts). Because the underlying philosophy of TMOs concerns with reducing they range, they inevitably saturate visual information leading to loss of details. As a consequence, their use for HDR visualization calls for extreme care. An example to show that TMOs can result in very different visual qualities for the same HDR content is given in Figure 6. Observe how some TMO preserve only indoor details while some do so for outdoor information. Ashikmin and icam06 [19] TMOs seem to provide a better trade-off in maintaining visual details in outdoor and indoor simultaneously.

A related aspect in tone mapping is that of naturalness. While it is quite clear that tone mapping reduces visual quality, how it affects naturalness remains an unanswered question. In fact, all the user studies described previously implicitly account for naturalness. This is because when human observers judge the visual quality of tone mapped content, not only the presence or absence of visual details affects their choice but is also affected implicitly by the naturalness of the tone mapped content. Naturalness is a subjective quality which is difficult to be quantified. In the light of this, it is not surprising that most of the TMOs only focus on retaining details and/or maintaining local and global contrast but do not consider naturalness for processing the HDR content. For instance, an over enhanced tone mapped image might have a very large number of details but can still have poor visual appearance due to being unnatural.

Thus, we have provided a brief overview of the impact of one mapping on visual quality of HDR content. We have also provided several useful references for the reader to explore further. In summary, tone mapping in general, degrades visual quality by destroying scene details, ad-hoc saturation of pixels as well as affecting the natural appearance of the content. In

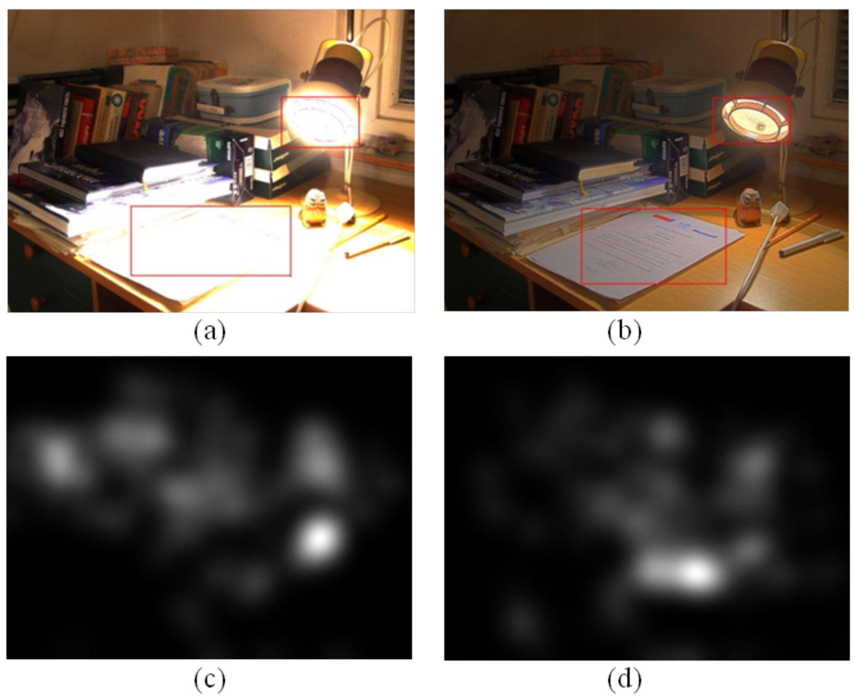


Figure 7: Illustration of the effect of global and local TMOs. (a) Image processed by Tumblin TMO (global), (b) Image processed by Ashikimin TMO (local), (c) VA map for (a) and (d) VA map for (b). The red boxes highlight two areas in the images where details are lost and preserved by global and local TMOs respectively.

the next section, we discuss the effect of tone mapping on visual attention.

4.3 Tone mapping and visual attention

As mentioned in the previous section, the current effort in subjective evaluation has been mainly directed towards assessing the impact of TMOs from quality and aesthetic appeal point of view. From these we may be able to study and analyze peoples preference regarding visual appeal of the tone mapped content. However, visual quality is just one of the several aspects that need to be considered to make conclusions on how TMOs affect the overall quality of experience (QoE). One such issue is that of visual attention (VA) which has been well recognized as a crucial aspect in perceptual visual signal processing. It is well known that human eyes tend to focus more on certain areas in an image/video than others. Stated differently, some regions attract more eye attention and these are termed as salient regions. VA is therefore the ability of the HVS to find and focus on relevant information quickly and efficiently [36]. This has several applications since the more important signal information can be extracted and processed accordingly. For example in image/video coding, the visually salient parts can be assigned more bits in order to achieve higher efficiency and better visual quality. Further from an artistic viewpoint, TMOs could possibly change the way a scene is perceived by human eyes. This may lead to changes in the feelings and emotions conveyed by the image. Thus the intention of the artist /content author may not be represented correctly to the viewer. For instance, intricate details (like very fine texture) in some part of an image which attract viewer attention might be lost due to tone mapping and the photographers intention of producing a compelling picture is jeopardized. It is thus clear that VA plays an important in human perception and therefore the impact of TMOs should also be analyzed from this view point as well.

In the context of VA, eye-tracking is the term commonly used to denote the way of exploring what people look at in any given situation and record their VA strategies with location and duration. The impact of TMOs on visual attention is best explained through VA maps obtained from human observers. To begin the analysis, we recall that TMOs tend to destroy visual details. Moreover, given that some TMOs can preserve details better than others, the VA behavior can change with TMO. To visually exemplify this, we have shown in Figure 7, the tone mapped versions of the an HDR image processed by Ashikimin and Tumblin TMOs. We can immediately make two observations from this Figure. Firstly, the image processed by Ashikimin TMO has more details preserved in the regions highlighted by

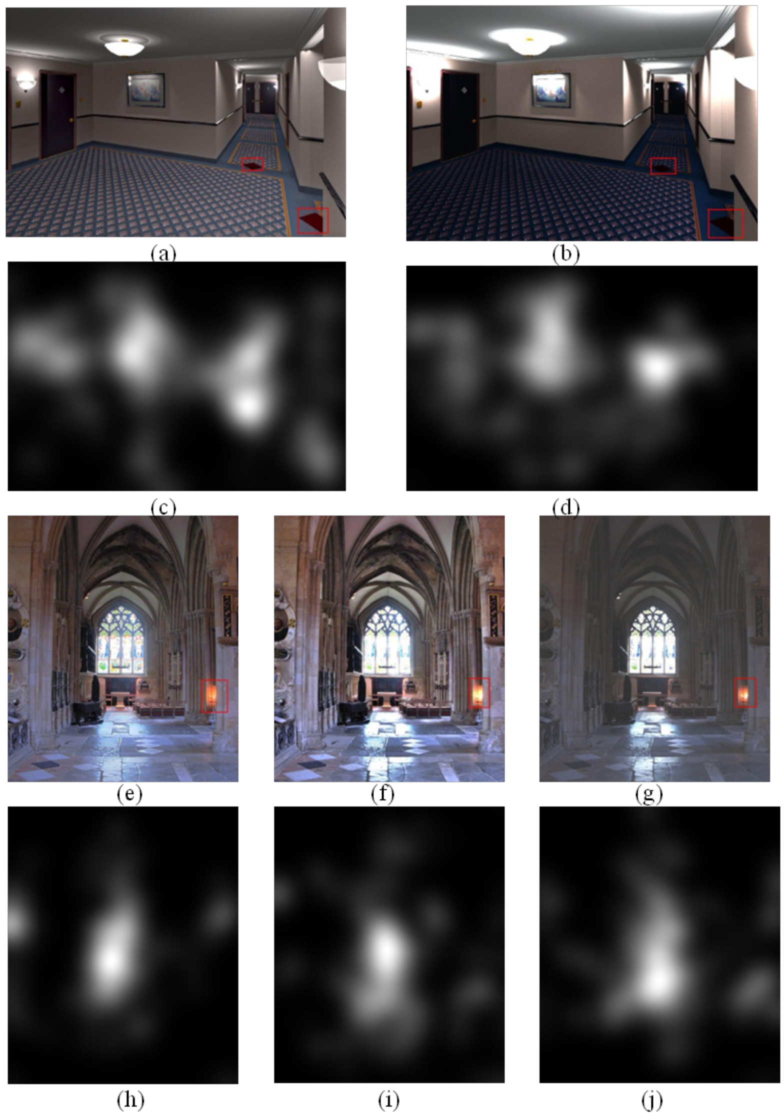
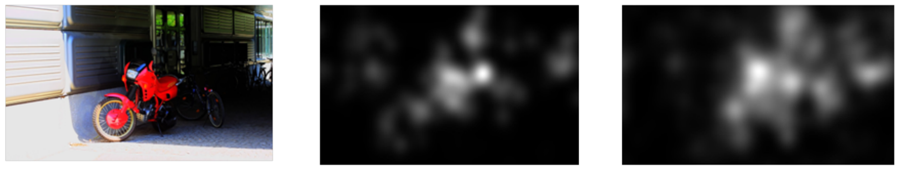


Figure 8: Effect of TMOs on VA. (a) rend02_oC95 image processed by iCAM06 TMO, (b) rend02_oC95 image processed by Drago TMO, (c) VA map for (a) and (d) VA map for (b), (e)-(g) Oxford_Church image processed by Tumblin, iCAM06 and linear TMOs respectively, (i)-(k) VA maps for the images shown in (e)-(g) respectively. The red boxes highlight the area(s) in the images which become salient or non-salient depending on the overall impact of TMO.



(a) (b) (c)

Figure 9: Effect of TMOs on VA. (a) tone mapped version of 'moto' (Reinhard TMO), (b) VA map of image (a) obtained from eye-tracking and (c) VA map of 'moto' HDR image obtained from eye-tracking.

red boxes. As opposed to this, in the same regions of the image processed by Tumblin TMO, the details are clearly missing. Secondly, the overall contrast of the image in Figure 7 (a) is clearly better than the one in Figure 7 (b). To visualize and relate this to the impact of these TMOs on VA, we have shown the corresponding VA maps obtained from eye-tracking. The reader can notice that for the image processed by Tumblin, the attention regions are mainly the books in the background while the letter pad (in the foreground) is nearly unnoticed by the observers. We hypothesize that this happens because with Tumblin being a global TMO the overall image contrast is maintained and the details in the dark areas of the image (like the books in background) are well retained. Further the owl below the lamp is also clearly visible and is a salient region. However, as already mentioned, all this comes with the price of losing finer details mainly in the bright areas (like the lamp and the letter pad) as highlighted. As a result, the letter pad is nearly non-salient since the useful information (the text inside) has been washed away by the TMO. In contrast to this, the VA map of the image processed by Ashikimin TMO shows that the written text on the letter pad is the most salient portion. Also, the darker background (mainly the books) seems to have become less eye-catching since the contrast in that part is reduced. Another example to illustrate that TMOs can modify the attention regions is shown in Figure 8. Here the images in the first and second rows are the tone mapped versions of rend02_oC95 image processed by Drago and iCAM06 TMOs and the corresponding human priority maps (VA maps) respectively. It can be seen that the two red mats (highlighted by red boxes) are more clearly visible in the image processed by iCAM06 since there is high contrast preserved in and around that region.

Consequently, one can see from the corresponding VA map that these indeed are salient regions for the human observers. On the other hand, in the image processed by Drago [9] there is much lower contrast in the said regions. As a result of these attract much lesser eye attention as seen in the corresponding VA map. A second set of examples is shown in the third and fourth rows of Figure 8. Here the third row shows three tone mapped versions of Oxford_Church image (tone mapped by Tumblin, iCAM06 and Linear TMOs) while the corresponding VA maps are shown in the fourth row below each image. Again, one finds that the orange spot (highlighted by red box) is a salient region only in case on linear TMO (see the VA map in Figure 8 (k)) since this TMO destroys contrast on other regions which makes the orange spot stand out and thus eye catching. As opposed to this, Tumblin and iCAM06 TMOs provide much better contrast in other parts of the image as well. So the orange spot is nearly non-salient in these two images as the observers attention is attracted to other parts. Therefore, based on our experiments and analysis of the VA maps obtained from eye-tracking experiments, we can say that contrast of the resultant tone mapped image plays an important role in VA behaviour. As exemplified by visual examples in Figures 7 and 8, the areas that attract eye-attention can vary even within the same image depending on whether contrast is preserved or destroyed by the TMO. This therefore suggests that contrast indeed is a vital dimension in HDR content processing from VA view point.

Tone mapping can also be viewed in terms of reduced signal contrast due to tone mapping. With the reduced contrast, regions that may have attracted observers' attention in the HDR content might be reduced. As a direct consequence, the number of salient regions in tone mapped HDR content tend to decrease. To visually exemplify this, consider Figure 9, where image (a) is a tone mapped version (using Reinhard TMO) of an HDR image. In this image, we can easily identify the foreground (mainly comprising of the headlight and front wheel of the bike) and the background (bicycles and the door). The image (b) shows the VA map of image (a) while the HDR VA map of is shown in (c). Observe how the VA map in (b) indicates very few salient points in the background. As opposed to this, the HDR VA map shows that background also had regions which attracted eye attention. The reason is obvious: tone mapping in this case destroys details mainly in the background but the foreground is fairly well preserved in terms of contrast. As a result, the number of salient points in the background reduce drastically. The last visual example is shown in Figure 10. In this Figure, the second row shows the corresponding VA maps of the images shown in the first row. Since HDR image does not display properly, we have shown image processed by icam06 (instead of the original HDR image) for the sake of explanation. Also note that we have used green

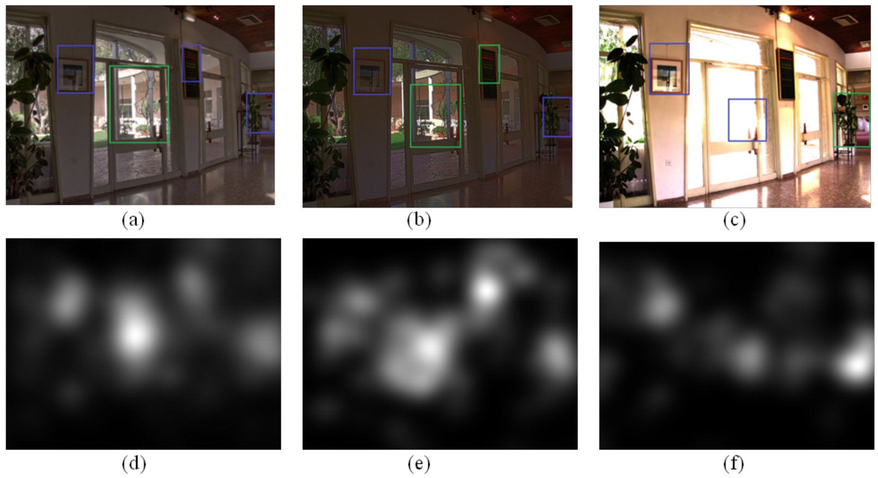


Figure 10: Effect of TMOs on VA. (a) 'dani_belgium' HDR image (for sake of better visualization tone mapped image processed by icam06 TMO has been displayed), (b) tone mapped version of 'dani_belgium' (Ashikimin TMO), (c) tone mapped version of 'dani_belgium' (Drago TMO), 'd) VA map of HDR image, (e)-(f) corresponding VA maps of the images shown above.

box to highlight the area(s) that attracted maximum human attention and blue box for area(s) with relatively lower attention.

Considering the HDR VA map in Figure 10 (d), it shows that there are 4 main regions which are salient according to human observers. These have been highlighted in image (a) shown just above the HDR VA map and include the outside area seen through the door (highlighted through the green box) and the paintings/board (highlighted through blue boxes). Also, notice that the area highlighted in green attracts more attention as compared to the other 3 identified regions. Now we observe the effect of tone mapping on these 4 identified regions. We find that the image processed by Ashikimin TMO (shown in Figure 10 (b)) shows that now there are 2 regions (highlighted by green boxes) which attract the maximum attention (see the corresponding VA map below this image). Thus, tone mapping modified the visual signal in such a manner that a region which was less salient in the original HDR content has become more salient. Likewise, looking at the VA map in Figure 10 (b), we find that there is only one region now that attracts maximum attention (this has again been marked in green in the image shown above this VA map) while the attention for other regions reduced considerably. This once again drives home the point: tone mapping

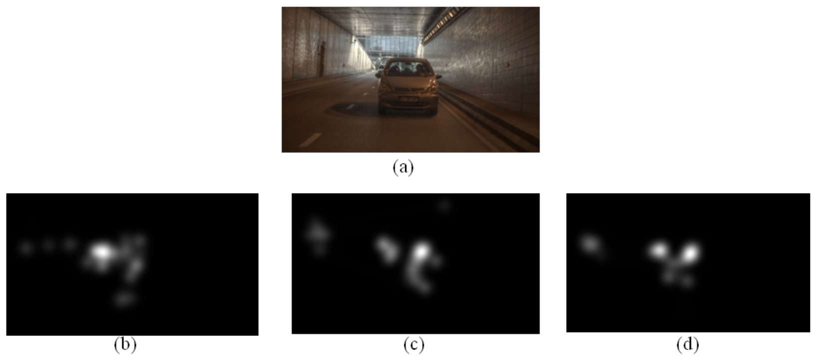


Figure 11: Change in VA behavior in videos , (a) Video frame, (b) HDR VA map, (c) VA map of frame tone mapped by Tumblin TMO and (d) VA map of frame tone mapped by Durand TMO

can change attentional regions in addition to increasing or decreasing the magnitude of attention. Therefore, as analyzed and explained we note large differences for both intra (i.e. for each image content) and inter (i.e. between different image contents) cases. A theoretical explanation for this is the manner in which TMOs operate. Most of them sacrifice one or the other type of visual information in order to reduce the dynamic range. In the process, additional artifacts (such as additional contours) might be introduced.

The eventual result is that a non-attentional region in the HDR image becomes attentional one in the tone mapped version. The opposite case is that in which structural information is destroyed due tone mapping. In such cases, an attentional region in the HDR image becomes less important (or less eye catching) in the tone mapped image. For example, a contrast that was visible in the HDR image becomes invisible in the processed image (loss of visible contrast). We have already provided some visual examples to illustrate these points.

Video signals differ from images due to the addition of a temporal dimension in addition to the spatial one. Given that, it will be interesting to analyze the changes in VA behavior due to tone mapping of HDR video sequences. The analysis of the VA maps from different video stimuli leads to similar conclusions as still images. That is TMOs have a large impact on the VA behavior as compared to the HDR video. We present an example in Figure 11. This is from the video sequence [1] 'Tunnel1' in which a car is

[1]This video sequence was shot as part of the NEVEx project FUI11, related to HDR video chain study.

shown to enter into a tunnel (with normal traffic). Inside the tunnel, there is relatively lower illumination and so as the car enters into it, there is a large change in scene illumination. Figure 11 (a) shows the car inside the tunnel and another car just behind it which also enters the tunnel. Before this time, we found that car was the main region of subject's attention right from the start of the video. However, when the other car enters the frame from behind, it attracts subjects' attention. This is expected since the entry of the new car in the frame is a 'new' or a 'rare' event (upto this point the subjects attention is focused on the first car only). That directs the attention to the second car. This is what was observed when the HDR video was viewed on an HDR screen. The corresponding HDR VA map is shown in Figure 11 (b) where one can see that the 'second' car is the main region of attention. However, when tone mapped video was shown to the subjects we observed different behavior. In this case, it was found that the 'first' car still remains the main focus of attention. This can be clearly observed in the VA map corresponding to Tumblin TMO shown in Figure 11 (c). That is, despite the occurrence of a new event (the entry of the second car), attention behavior did not change. We can attribute this to the fact that Tumblin TMO could not maintain proper contrast at the tunnel entrance where there is a large change in the intensity (dark inside the tunnel and bright outside it). Due to this, the subjects attention was not fully diverted towards the 'second' car. A different observation was however made for in case of Durand TMO. This TMO could maintain relatively better contrast at the tunnel entrance. Due to this, we have a situation where both the 'first' and 'second' cars became the regions of attention. This can be seen from the VA map shown in Figure 11 (d). Thus, depending on the TMO we have different VA behavior for the same scene in the video. This suggests that similar to the case of still images, tone mapping changes VA behavior over time.

Tone mapping is an important HDR processing that enables HDR visualization on traditional display devices. This section was therefore devoted to the study of its impact on the overall QoE. To facilitate discussion, we identified three important dimensions of HDR that tone mapping can affect. These include visual quality, naturalness and visual attention. The discussion was focused on specifics of how TMO affect these dimensions and several visual examples were provided as illustrations. Regarding visual quality, tone mapping generally leads to loss of contrast (both locally and globally) and, result in loss details. This can also have an adverse effect on the naturalness of the tone mapped HDR content. For instance, saturation of color or over-enhancement of details can render the processed content as being unnatural (despite preserving details). In other words, tone mapping can reduce the overall coherency associated with natural signals. We also discussed how tone mapping can impact visual attention behavior. This has the consequence of altering the artistic intention. For example a photographer might capture a scene with the intention that viewers will pay attention to some areas/aspects of the photograph. However, tone mapping can cause changes that can divert viewers' attention to areas/aspects that are not the same as what the photographer intended. Thus, tone mapping can interfere with the artistic intentions and that can ultimately reduce the visual experience [29], [26] of the processed HDR content. Such joint effect of changes in visual quality, naturalness and/or visual attention have the potential of lowering the enjoyment level of the end-users with regards to HDR viewing.

The reader will notice that the discussions pertaining to tone mapping have been in the context of employing TMOs for HDR visualization. In the next section, we discuss some aspects of quality measurement when viewing an HDR scene.

5 HDR quality of experience

HDR quality of experience is a rather wide term in that it can include several dimensions including perceptual quality, naturalness, visual attention, aesthetic appeal and so on. Of course these are not necessarily independent because for instance perceptual quality can implicitly account for naturalness. In the following, we first outline the differences between HDR and LDR from the angle of viewing conditions and then introduce the reader to the topic of subjective and objective HDR quality assessment.

5.1 Viewing conditions in HDR

The most important distinction of HDR from LDR is with respect to luminance range (which in turn leads to high dynamic range). Traditional LDR defines a white point (255 for the 8-bit representation). Thus, any intensity more than the defined white needs to be saturated. Moreover, with LDR the pixel values are typically gamma encoded and perceptually uniform. As a result, change in the pixel values can directly be related to the change in visual perception. However, with HDR there is more flexibility to represent the real-world scene luminance without too much saturation. Consequently, there is no fixed white point in HDR that can correspond to the maximum luminance (as it can vary from scene to scene). There is only brighter (or

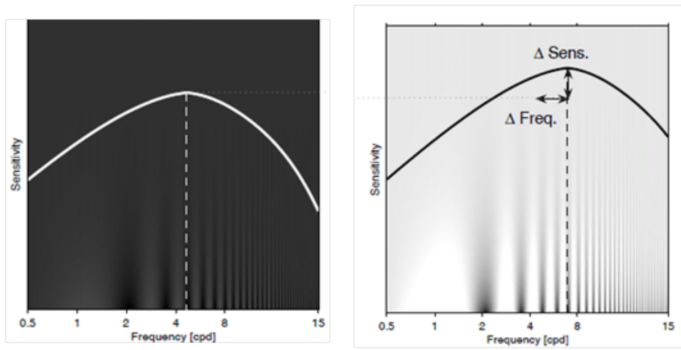


Figure 12: Contrast sensitivity function (CSF) of the human eye in dark (left) and bright (right) viewing conditions.Arrows labeled as $\Delta Sens.$ and $\Delta Freq.$ denote the amount of difference in magnitude and frequency of the peak sensitivity between the dark and bright cases. The figure is reproduced from [38].

darker) scene intensity. Therefore, HDR viewing involves much higher levels of brightness. Since human vision is sensitive to luminance ratio (rather than absolute luminance), changes in the luminance may not necessarily lead to the same change in visual perception of HDR.

The effect of luminance level on the sensitivity of the human visual system is often referred as luminance masking. Figure 12 shows the Campbell-Robson contrast sensitivity chart for two different background luminance levels [38]. For the best viewing, the figure should be viewed on an LCD display of about 200 cd/m2 and the display function close to the sRGB non-linearity. The solid lines denote the contrast sensitivity of the HVS, which is the contrast level at which the sinusoidal contrast patterns become invisible. Even though the same scales were used for both left and right plots, the CSF is shifted upwards (higher sensitivity) and right (towards higher spatial frequencies) for the brighter pattern. This shows that we are more likely to notice contrast changes, if the stimuli is brighter, as is the case of a brighter display. A further validation of this was done in [38] through a subjective experiment. It was found with statistical evidence that distortions of the same type and with the same magnitude are more annoying when the overall brightness of the image is higher. Thus, with HDR, one needs to take into account the display luminance conditions as it can have a significant impact on the perceived quality of the stimuli.

Another crucial factor with regards to HDR viewing condition is the ambient lighting. Given the high levels of luminance, HDR will require a higher level of ambient lighting as compared to LDR. Obviously with low ambient lighting, HDR viewing can be uncomfortable for viewers. With regards to LDR, the International Telecommunication Union Recommendation (ITU-R) BT500-11 recommendation specifies the room (ambient) illumination to be about 15% of the perceived screen brightness. It is not clear if this can be applied to HDR viewing. For example, with SIM2 HDR display [3], the maximum luminance is 4000 cd/m and so the room illumination should be around 600 cd/m according to BT500-11 recommendations. However, it is known that the response of the human eye to luminance is approximately logarithmic and it is likely that a little lower ambient lighting level might be suitable. In any case, it is clear that the current ambient lighting specifications for LDR will not be entirely suitable for HDR viewing.

5.2 Subjective assessment of HDR quality

Human judgment of visual quality remain the gold standard as far the accuracy of quality prediction is concerned. HDR is no exception. However, as outlined in the previous section, subjective measurement of HDR quality calls for more careful considerations of viewing conditions. Otherwise the results may not reflect the actual perceived quality. Another important factor in HDR subjective test design is the use of TMOs. They will not be used for visualization but for HDR processing (eg. compression). A problem with TMOs is that they usually require one or more parameters that are left for the user to tune. The issue of best parameter selection is further complicated since it can be HDR content specific. That is, a set of parameters which is suitable for one content may not be optimal for the other (generally speaking the default parameter values might not yield reasonable quality for every HDR content). Therefore, it requires care to find TMO parameters when preparing HDR content for subjective evaluation. Concerning the sources of distortions in HDR, the first is related to tone mapping. Another common distortion is compression related artifacts. Another category of specific artifacts that occur in HDR are those due to inverse tone mapping. Inverse tone mapping is the final step in a typical backwards compatible HDR compression pipeline and can cause saturation, excess or lack of contrast in the HDR scene. Thus, it can be highlighted that HDR processing includes specific distortion sources (that are not typically present in LDR regime) like tone mapping and inverse tone mapping in addition to common artifacts (due to compression, transmission, post-processing etc.). It is also interesting to note that distortions due to tone mapping and inverse tone mapping are not necessarily additive. That is, inverse tone

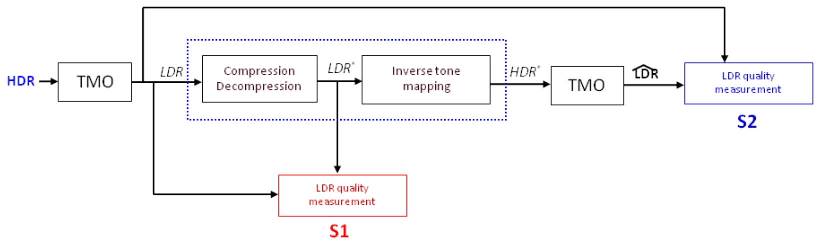


Figure 13: LDR approach to HDR quality assessment.

mapping can offset some artifacts from tone mapping. Further, as explained in previous sections, visual attention can be significantly modified due to tone mapping and this can degrade the overall HDR viewing experience. Thus, HDR quality measurement is challenging in that the processed HDR content can suffer from multiple distortions (which need not be independent of each other). This coupled with the fact that the high luminance in HDR can potentially amplify artifacts suggests that extra care needs to be taken for accurate subjective measurement of HDR quality.

Very few research efforts have been reported for subjective HDR quality assessment. The reasons for this are related to the requirement of specialized HDR displays and the unavailability of real HDR content. Nevertheless, two recent studies have employed an HDR display for QoE evaluation. The first one [27] investigated into codec optimization criterion and perceptual quality issues in HDR. The second study [28] analyzed the impact of TMOs in HDR compression. The conclusions from these studies revealed that, indeed, the perceptual quality of the decompressed HDR signal is dependent on the tone mapping method employed and statistical evidences were also presented to support that. The reader will note that both these works are in contrast to most of the existing studies (some of which have been described in previous sections) which focused only on the quality loss in the resultant LDR signal (obtained via tone mapping) and typically employed LDR displays for subjective viewing experiments.

The final point for subjective HDR quality assessment is related to the need of specialized displays. As we have already pointed out, such HDR displays are still not common on a large scale (although this could change within a reasonable time frame). In the light of such constraint, it is natural to ask if HDR quality measurement can be tackled with existing LDR set up (LDR displays, specifications etc.). Indeed, it is not absurd to think of tone mapping the HDR content and estimate its quality subjectively on an LDR monitor. Specifically with regards to quality measurement in a backwards compatible HDR compression system, there can be two possible scenarios to convert HDR quality assessment to an LDR one. These have been illustrated in Figure 13. The first scenario (indicated as S1) is to judge the subjective quality of the decompressed LDR content (before inverse tone mapping) with respect to reference tone mapped LDR. A second possible scenario (indicated as S2 in Figure 13) is to tone map the decompressed HDR content. Then, this can be compared with the tone mapped reference scene to determine its quality. Unfortunately, both the scenarios have their own limitations. The first one is that by viewing HDR on an LDR display, we ignore the distinct aspects of high luminance associated with HDR viewing (explained in the previous section). The second major issue is that the reference tone mapped HDR is more pristine and as a result can bias the human judgments. Another limitation of assessing HDR quality with scenario 1 (S1) is that it ignores the impact of inverse tone mapping completely and takes into account only the compression artifacts. Recall that inverse tone mapping can introduce visible artifacts such as saturation, excess or lack of contrast etc. Therefore, with this scenario (S1) formulating HDR quality assessment as an LDR one is expected to be less accurate. With scenario 2 (S2) we can note that the tone mapping will operate on different HDR content (pristine HDR and decompressed HDR). Consequently, they will be modified differently and the resulting judgment of visual quality can be erroneous. So it can be concluded that HDR quality can be better judged by simulating proper HDR conditions (use of an HDR display, appropriate ambient lighting etc.). However, it is fair to reiterate that even HDR displays have their own limitations (in particular due to the dual modulation process) and cannot fully represent HDR content. This is however the best available solution at the time of writing.

5.3 Objective assessment of HDR quality

Objective quality measurement is the use of computational model to predict quality. An objective method for quality prediction is a useful tool in cases where subjective assessment is not feasible (such as real-time applications). Being a mathematical model, an objective method is more convenient to be deployed. However, objective methods cannot be as accurate as the subjective ones. Thus, one line of thinking in the research community has been towards developing more accurate objective methods. While this is reasonable, it is important to understand that the human visual system (HVS) represents a complex visual information processing system. Consequently, all the objective methods (for LDR and HDR) are merely approximations and they cannot be relied upon as a generic solution to quality prediction. Nonetheless, it is also important to highlight that objective methods can achieve a reasonable prediction accuracy in the limited context of an application scenario. For example, the mean squared error (MSE) continues to be deployed extensively in visual data compression. Pertaining to LDR, one finds that a lot of research effort has been spent over the past decade. Most of it is devoted to the development of full-reference methods that require both the reference and processed visual signal for quality computations. In contrast, there exists very few methods for objective HDR quality prediction. The reasons for this are already outlined and related to different viewing conditions as compared to LDR. Thus, mathematical models of HVS's functioning (eg. contrast sensitivity) used in LDR methods can no longer be effective for HDR. Another reason for slow progress of objective HDR quality measurement can be attributed to the lack of standard databases.

The HDR-VDP-2 (High Dynamic Range Visual Difference Predictor) [34] is a fairly recent and comprehensive method for objective measurement of HDR quality. It is an extension of the Visible Differences Predictor (VDP) algorithm. The HDR-VDP-2 uses an approximate model of the human visual system (HVS) derived from new contrast sensitivity measurements. Specifically, a customized contrast sensitivity function (CSF) was employed to cover large luminance range as compared to the conventional CSFs. HDR-VDP-2 is essentially a visibility prediction metric. That is, it provides a 2D map with probabilities of detection at each pixel point and this is obviously related to the perceived quality because a higher detection probability implies a higher distortion level at the specific point. Nevertheless, in many cases, it is crucial to know an overall quality score (rather than just the local distortion visibility probability). Pooling is a crucial aspect in converting local error distribution into a single score that denotes the perceptual quality and the human visual system (HVS) can very easily do that accurately. But it is much more difficult to realize that in an objective quality prediction model given the underlying complexities and lack of knowledge of the HVS's pooling mechanisms. It is believed that multiple features jointly affect the HVSs perception of visual quality, and their relationship with the overall quality is possibly nonlinear and difficult to be determined apriori. Therefore, the approach that HDR-VDP- takes is that finding the pooling parameters via optimization of correlation with subjective scores.

In its original implementation, the authors of HDR-VDP-2 tried over 20 different combinations of aggregating (or pooling) functions. These included maximum value, percentiles (50, 75, 95) and a range of power means (normalized Minkowski summation) with the exponent ranging from 0.5 to 16. The aim was to maximize the value of Spearman's correlation coefficient in order to find the best pooling function and its parameters. While HDR-VDP-2 is a fairly comprehensive method for HDR quality assessment, there is an issue with regards to pooling in HDR-VDP-2. This is related to parameter optimization. That is, the parameters of the pooling function in HDR-VDP-2 were found by maximizing (optimizing) correlation using existing LDR image databases. Therefore, its effectiveness in predicting the visual quality of HDR images is questionable given the different characteristics LDR and HDR images especially in terms of distortion visibility and overall visual appeal [34]. The reader will notice that objective HDR quality assessment requires much more efforts both in terms of research as well as implementation. This is more so in the light of the fact that an LDR approach to HDR quality assessment is not as effective and cannot be a substitute to account for the effects that distortions have on HDR viewing.

6 Concluding Remarks and Perspectives

HDR imaging is an emerging area within the realm of visual signal processing. It brings to that table two major advantages over the traditional imaging systems. First, it can provide a more immersive and realistic viewing experience to the users. Second, the higher bit depth required in HDR will allow for more signal manipulation (eg. pre-processing towards efficient encoding) as compared to the traditional content. However, to exploit HDR technology to its fullest potential, several challenges remain and this chapter has focused on a few of them pertaining to their impact the overall HDR QoE. With regards to HDR processing, tone mapping is often required for HDR viewing on LDR displays, compression, and in many other scenarios where backward compatibility is desired. The aim of this chapter was to throw light on the impact of tone mapping on visual experience. Specifically, we discussed its impact on perceptual quality, visual attention, and naturalness. It is worth highlighting that these play an important role in the QoE in HDR viewing. We reiterate that HDR viewing experience is more immersive than traditional content due to that fact that HDR attempts to reproduce real-world scene information without undue saturation of visual information. In other words, with HDR we directly deal with physical luminance related information and this makes HDR experience more wholesome and enjoyable.

From the objective viewpoint, measurement of HDR QoE remains a challenge primarily due to a larger number of factors involved as compared to traditional video quality. In particular, unlike QoE judgment of traditional visual content, the impact of factors such as naturalness and visual attention modification can be more profound in HDR. Therefore, single measures such as signal fidelity alone cannot be expected to be a reasonable substitute for the overall QoE. On the operational front, HDR poses difficulties because the information is stored in luminance-related format, unlike perceptually scaled pixel values in LDR signals. Finally, native HDR visualization is not possible even with the current HDR display technologies and there is saturation of signal contrast (this is of course due to inherent hardware limitations such as the upper limit on power consumption, heating etc). Addressing some of the mentioned issues will ultimately be the key to large scale practical deployment of HDR and further interesting applications.

References

[1] Yoshida A., Blanz V., Myszkowski K., and Seidel H. Perceptual evaluation of tone mapping operators with real-world scenes. In *Proceedings of SPIE Human Vision & Electronic Imaging X*, pages 192–203, San Jose, CA, USA, 2005.

[2] Industrial Light & Magic (2008) OpenEXR. Available at: http://www.openexr.com.

[3] SIM2 MULTIMEDIA Available at: http://www.sim2.com/HDR/.

[4] Spheron HDR VR. Available at: http://www.spheron.com/home.html.

[5] Guthier B. Real-time algorithms for high dynamic range video. *PhD. Thesis*, 2012.

[6] Reinhard E., Stark M., Shirley P., and Ferwerda J. Photographic tone reproduction for digital images. *ACM Transactions on Graphics (TOG)*, 21(3):267–276, July 2002.

[7] Banterle F., Artusi A., Debattista K., and Chalmers A. *Advanced High Dynamic Range Imaging: Theory and Practice*. AK Peters (CRC Press), Natrick, MA, USA, 2011.

[8] Banterle F., Debattista K., Artusi A., Pattanaik S., Myszkowski K., Ledda P., and Chalmers A. High dynamic range imaging and low dynamic range expansion for generating HDR content. *Computer Graphics Forum*, 28(8):3243–2367, December 2009.

[9] Drago F., Myszkowski K., Annen T., and Chiba N. Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum*, 22(3):419–426, September 2003.

[10] Drago F., Martens W., Myszkowski K., and Seidel H. Perceptual evaluation of tone mapping operators. In *Procedings of ACM SIGGRAPH 2003 Sketches & Applications*, pages 1–1. ACM Press, 2003.

[11] Durand F. and Dorsey J. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics (TOG)*, 21(3):257–266, July 2002.

[12] Eilertsen G., Wanat R., Mantiuk R., and Unger J. Evaluation of tone mapping operators for HDR-video. *Computer Graphics Forum*, 32(7):275–284, October 2013.

[13] Mather G. *Foundations of Perception.* Psychology Press, Hove, East Sussex, 2006.

[14] Ward G. Real pixels. In *Graphic Gems II*, pages 80–83. Academic Press, 1991.

[15] Ward G. A contrast-based scalefactor for luminance display. In *Graphic Gems IV*, pages 415–421. Academic Press, 1994.

[16] Ward G. The LogLuv encoding for full gamut, high dynamic range images. *Journal of Graphics Tools*, 3(1):15–31, March 1998.

[17] Ward G. and Simmons M. JPEG-HDR: A backwards-compatible high dynamic range extension to jpeg. In *ACM SIGGRAPH 2006 Courses*, 2006.

[18] Seetzen H., Heidrich W., Stuerzlinger W., Ward G., Whitehead L., Trentacoste M., Ghosh A., and Vorozcovs A. High dynamic range display systems. *ACM Transactions on Graphics (TOG)*, 23(3):760–768, August 2004.

[19] Kuang J., Johnson G., and Fairchild M. iCAM06: A refined image appearance model for hdr image rendering. *J. Visual Communication and Image Representation (JVCI)*, 18(5):406–414, October 2007.

[20] Kuang J., Yamaguchi H., Liu C., Johnson G., and Fairchild M. Evaluating hdr rendering algorithms. *ACM Transactions on Applied Perception (TAP)*, 4(2), Article No. 9,July 2007.

[21] Tumblin J., Hodgins J., and Guenter B. Two methods for display of high contrast images. *ACM Transactions on Graphics (TOG)*, 18(1):56–94, January 1999.

[22] Chiu K., Herf M., Shirley P., Swamy S., Wang C., and Zimmerman K. Spatially nonuniform scaling functions for high contrast images. In *Proceedings of Graphics Interface*, pages 245–253, 1993.

[23] Ashikhmin M. A tone mapping algorithm for high contrast images. In *Proceedings of the 13th Eurographics workshop on Rendering (EGRW)*, pages 145–156, 2002.

[24] Ashikhmin M. and Goyal J. A reality check for tone-mapping operators. *ACM Transactions on Applied Perception (TAP)*, 3(4):399–411, October 2006.

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

[26] Narwaria M., Silva M., Callet P., and Pepion R. Effect of tone mapping operators on visual attention deployment. In *Proceedings of SPIE 8499, Applications of Digital Image Processing XXXV*, San Diego, California, USA, 2012.

[27] Narwaria M., Silva M., Callet P., and Pepion R. Tone mapping-based high-dynamic-range image compression: study of optimization criterion and perceptual quality. *Optical Engineering*, 52(10):102008–102008, 2013.

[28] Narwaria M., Silva M., Callet P., and Pepion R. Impact of tone mapping in high dynamic range image compression. In *Proceedings of Eighth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, Chandler, Arizona, USA, 2014.

[29] Narwaria M., Silva M., Callet P., and Pepion R. Tone mapping based hdr compression: Does it affect visual experience? *Signal Processing: Image Communication*, 29(2):257–273, February 2014.

[30] Sugiyama N., Kaida H., Xue X., Jinno T., Adami N., and Okuda M. HDR compression using optimized tone mapping model. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1001–1004, 2009.

[31] Ledda P., Chalmers A., Troscianko T., and Seetzen H. Evaluation of tone mapping operators using a high dynamic range display. *ACM Transactions on Graphics (TOG)*, 24(3):640–648, July 2005.

[32] Fattal R., Lischinski D., and Werman M. Gradient domain high dynamic range compression. *ACM Transactions on Graphics (TOG)*, 21(3):249–256, July 2002.

[33] Mantiuk R., Efremov A., Myszkowski K., and Seidel H. Backward compatible high dynamic range MPEG video compression. *ACM Transactions on Graphics (TOG)*, 25(3):713–723, July 2006.

[34] Mantiuk R., Jim K., Rempel A., and Heidrich W. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Transactions on Graphics (TOG)*, 30(4), July 2011.

[35] Mantiuk R., Myszkowski K., and Seidel H. A perceptual framework for contrast processing of high dynamic range images. *ACM Transactions on Applied Perception (TAP)*, 3(3):267–276, July 2006.

[36] Rensink R. Visual attention. In *Encyclopedia of Cognitive Science*. Nature Publishing Group, London, 2003.

[37] Mann S. and Picard R. Being 'undigitial' with digital cameras: Extending dynamic range by combining differently exposed pictures. In *Proceedings of IS&T 48th Annual Conference*, pages 422–428. Society for Imaging Science and Technology, 1995.

[38] Aydin T., Mantiuk R., and Seidel H. Extending quality metrics to full luminance range images. In *Proceedings of SPIE Human Vision & Electronic Imaging XIII*, pages 68060B–68060B–10, San Jose, CA, USA, 2008.