TONE MAPPING OPERATOR FOR HIGH DYNAMIC RANGE VIDEO

by

Stylianos Ploumis

B.S.in Informatics, University of Piraeus, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2017

© Stylianos Ploumis, 2017
Abstract

High Dynamic Range (HDR) technology is emerging as the new revolution in digital media and has recently been adopted by industry as the new standard for capturing, transmitting and displaying video content. However, as the majority of the existing commercial displays is still only limited to the Standard Dynamic Range (SDR) technology, backward compatibility of HDR with these legacy displays is a topic of high importance. Over the years, several Tone Mapping Operators (TMOs) have been proposed to adapt HDR content, mainly images, to the SDR format. With the recent development of SDR displays, the need for video TMOs became essential. Direct application of image TMOs to HDR video content is not an efficient solution as they yield visual artifacts such as flickering, ghosting and brightness and color inconsistencies.

In this thesis we propose an automated, low complexity content adaptive video TMO which delivers high quality, natural looking SDR content. The proposed method is based histogram equalization of perceptually quantized light information and smart distribution of HDR values in the limited SDR domain. Flickering introduced by the mapping process is reduced by our proposed flickering reduction method, while scene changes are detected by our approach, thus successfully maintaining the original HDR artistic intent. The low complexity of the proposed method along with the fact that it does not require any user interaction, make it a suitable candidate for real time applications, such as live broadcasting.
Lay Summary

High Dynamic Range (HDR) content and displays are available on the market but the vast majority of the user displays is Standard Dynamic Range (SDR), not able to reproduce the enriched, higher quality HDR content. Thus, there is an immediate need for methods that can efficiently map HDR content to the SDR format, making it compatible backward compatible with the existing infrastructure and SDR displays while offering the best possible visual quality. Such methods are known as Tone Mapping Operators (TMOs). The majority of the existing TMOs were designed to address HDR images. Naïve, straightforward application of these operators on HDR video sequences produces visual artifacts such as flickering, ghosting and temporal texture and color inconsistencies. In this work, we propose a real time, content adaptive, automated video TMO which outperforms all existing TMOs, producing high quality SDR video content while maintain the artistic intent of the original HDR content.
Preface

All of the work presented here in this thesis was conducted in the Digital Multimedia Laboratory at the University of British Columbia, Vancouver campus.

A version of Chapter 3, sub-sections 3.2.3.1 and 3.2.3.2, has been published as S. Ploumis, R. Boitard, M. T. Pourazad and P. Nasiopoulos, “Perception-based Histogram Equalization for tone mapping applications,” in IEEE Digital Media Industry & Academic Forum (DMIAF), Santorini, Greece, 2016. I was the lead investigator responsible for all areas of research, data collection, and the majority of manuscript composition. R. Boitard and M. T. Pourazad were involved in the early stages of research concept formation and aided with manuscript edits. P. Nasiopoulos was the supervisor on this project and was involved with research concept formation, and manuscript edits.

A version of Chapter 3, sub-section 3.2.3, is accepted to be presented at IEEE International Conference on Computing, Networking and Communications (ICNC 2018) as S. Ploumis, M. T. Pourazad and P. Nasiopoulos, “A Flickering Reduction Scheme for Tone Mapped HDR Video.” I was the lead investigator responsible for all areas of research, data collection, as well as the manuscript edit. M. T. Pourazad was involved in the early stages of research concept formation and aided with manuscript edits. P. Nasiopoulos was the supervisor on this project and was involved with research concept formation, and manuscript edits.
An invention disclosure (18-031) that is based on our developed TMO with title “Tone-mapping Scheme for HDR content” has been submitted to the UBC’s University-Industry Liaison Office (UILO).
# Table of Contents

Abstract .......................................................................................................................................... ii  
Lay Summary .................................................................................................................................. iii  
Preface ........................................................................................................................................... iv  
Table of Contents ........................................................................................................................... vi  
List of Figures ............................................................................................................................... viii  
Acknowledgements ........................................................................................................................ x  
Dedication ..................................................................................................................................... xi  

## Chapter 1: Introduction ..............................................................................................................1  
1.1 Overview ................................................................................................................................. 1  
1.2 Motivation ............................................................................................................................... 3  
1.3 Thesis Organization ................................................................................................................ 3  

## Chapter 2: Background ..............................................................................................................4  
2.1 High Dynamic Range (HDR) Technology ............................................................................. 4  
2.1.1 Overview .......................................................................................................................... 4  
2.1.2 Perceptual Quantization ................................................................................................. 5  
2.2 Tone Mapping Operators ...................................................................................................... 7  
2.2.1 Overview ........................................................................................................................ 7  
2.2.2 Challenges of Video Tone Mapping and Scene Changes Preservation ..................... 9  
2.2.3 Offline Video TMOs ....................................................................................................... 10  
2.2.4 Online Video TMOs ...................................................................................................... 12  

## Chapter 3: Efficient Automated Tone Mapping Operator for HDR Video Content ............14
3.1 Introduction ................................................................................................................... 14

3.2 Our Proposed Method ................................................................................................... 14

3.2.1 Overview ................................................................................................................... 14

3.2.2 Display Model ........................................................................................................... 15

3.2.3 Tone Mapping Process .............................................................................................. 16

3.2.3.1 Histogram Equalization .................................................................................... 16

3.2.3.2 Ceiling Function ............................................................................................... 20

3.2.3.3 Our Redistribution Method ............................................................................... 23

3.2.4 Flickering Reduction and Scene Changes Preservations .......................................... 25

3.3 Results and Discussions ............................................................................................... 27

3.3.1 Visual Results on Quality of Tone Mapped Images ................................................. 27

3.3.2 Subjective Evaluation ............................................................................................... 32

3.3.3 Statistical Analysis for Flickering Reduction ........................................................... 41

Chapter 4: Conclusion and Future Work .............................................................................. 45

4.1 Conclusion .................................................................................................................... 45

4.2 Future Work .................................................................................................................. 46

Bibliography ................................................................................................................................. 47
List of Figures

Figure 2-1. Dynamic Range Demonstration ........................................................................... 4
Figure 2-2 Perceptual Linearity ............................................................................................ 5
Figure 2-3. Comparison between (a) an SDR image and (b) an HDR tone-mapped image .... 7
Figure 2-4. Offline TMO (a), Online TMO (b) ..................................................................... 8
Figure 3-1. Overview of the Proposed TMO ......................................................................... 14
Figure 3-2. Display Characteristics ...................................................................................... 15
Figure 3-3. Histogram Equalization ...................................................................................... 18
Figure 3-4. Transfer both of HDR and SDR range to PQ domain ........................................ 22
Figure 3-5. Visual Results of the proposed method ............................................................... 29
Figure 3-6. Visual Results of the proposed method using compressed images ................... 31
Figure 3-7. Results of the first Subjective Test ..................................................................... 34
Figure 3-8. Visual Results of TMOs on sequence “FireEater” ............................................... 35
Figure 3-9. Results of the second Subjective Test ................................................................. 36
Figure 3-10. Visual Results of TMOs on sequence “Starting” ............................................... 38
Figure 3-11. Results of the third Subjective Test ................................................................. 39
Figure 3-12. Visual Results of TMOs on sequence “Hurdles” ............................................... 39
Figure 3-13. Statistical Evaluation in Flickering Reduction ............................................... 42
Figure 3-14. Statistical Evaluation on Scene Change Preservation ...................................... 42
Figure 3-15. Demonstration of Scene Change Preservation .................................................. 44
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIE</td>
<td>Commission Internationale del’Eclairage</td>
</tr>
<tr>
<td>cd/m²</td>
<td>Candela Per Square Meter</td>
</tr>
<tr>
<td>fps</td>
<td>Frames Per Second</td>
</tr>
<tr>
<td>HDR</td>
<td>High Dynamic Range</td>
</tr>
<tr>
<td>HVS</td>
<td>Human Visual System</td>
</tr>
<tr>
<td>ITU-T</td>
<td>International Telegraph Union-Telecommunication Standardization Sector</td>
</tr>
<tr>
<td>MPEG</td>
<td>Motion Picture Experts Group</td>
</tr>
<tr>
<td>JND</td>
<td>Just Noticeable Difference</td>
</tr>
<tr>
<td>PQ</td>
<td>Perceptual Quantizer</td>
</tr>
<tr>
<td>SDR</td>
<td>Standard Dynamic Range</td>
</tr>
<tr>
<td>SMPTE</td>
<td>Society of Motion Picture and Television Engineers</td>
</tr>
<tr>
<td>TMO</td>
<td>Tone Mapping Operators</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to start by expressing my most sincere gratitude to my supervisor and mentor Dr. Panos Nasiopoulos for his support through the past three years and patient guidance and inspiration through the thesis.

I would also like to thank Dr. Mahsa T. Pourazad for her constant help and support during different stages of this thesis.

I am also grateful to my lab mate and friend, Dr. Ronan Boitard, who helped me thorough feedback and support during different stages of my research.

My utmost gratitude to my beloved parents, who always believed in me and encouraged me through my whole life. I am always thankful to their unconditional love, unlimited dedication, and all the sacrifices they made for me. Without them, I would never be who I am today.
Dedication

To my beloved family
Chapter 1: Introduction

1.1 Overview

High Dynamic Range (HDR) technology has gained lot of interest recently, by both academia and industry, as it is regarded the next revolution in digital video technology. As such, HDR is being rapidly becoming industry’s new standard for capturing and transmitting high quality video content. HDR captures and displays a wide range of luminance values (0.005 – 10000cd/m²) close to the one the Human Visual System (HVS) can perceive [1]. This is a significant improvement over the Standard Dynamic Range (SDR) technology, which can capture and reproduce only a small proportion of this range (0.1-100cd/ m²). Hence, HDR offers higher Quality of Experience (QoE) and it is expected to replace SDR in the near future [2]. Production houses and broadcasters have already started delivering HDR content, but the fact is that the vast majority of the available displays still have only SDR capabilities. Considering the difference in range and quality on general that the two technologies offer, it is impossible to reproduce HDR content directly to SDR displays. Furthermore, naïve straightforward linear mapping from HDR to SDR does not preserve image contrast, brightness and details, resulting in unwanted visual artifacts that degrade the original artistic intent. It is, thus, imperative the need for designing efficient conversion schemes for HDR to SDR content that will offer the best possible visual quality. This process is known as tone mapping and its aim is to SDR that contains HDR tonal levels with the most important visual information, thus compressing the dynamic range of the content in the most effective way for the best possible visual output [3].

Given the fact that the HDR video technology started to mature only recently, the majority of the existing Tone Mapping Operators (TMOs) were designed for images. These operators can be
classified into two categories: local and global. Local TMOs, such as the ones described in [4, 5, 6], map each pixel based on the information of its spatial neighborhood. These operators preserve efficiently most of the visual information of the original content. However, sometimes they result in halo artifacts around the edges [7]. Another challenge of local TMOs is the fact that they use spatial filters that need to be manually tuned depending of content characteristics, thus making them unsuitable candidates for real-time applications. On the other hand, global TMOs such as [8, 9, 10] compute a monotonously increasing mapping curve for the entire image. The computation of the mapping curve is based on the image statistical information such as the average, peak and lowest brightness and the brightness histogram. Global operators manage to preserve the natural look of the HDR content, while keeping the computational cost at low levels, making them ideal for real time applications such as broadcasting, streaming and online video gaming. Their disadvantage, however, is that sometimes fail to efficiently preserve the details of the original image [11].

HDR video tone mapping is a relatively new research area due to the lack of widely available HDR video content until recently. Since capturing [12,13,14] and distribution [15,16] of HDR video became possible, entertainment and broadcasting industries along with content providers and display manufacturers became highly interested in efficient video TMOs. Straightforward application of image TMOs to video sequences causes several visual artifacts like visual noise, brightness flickering and ghosting as reported in [17,18,19]. While both local and global TMOs are prone to visual noise and flickering, the ghosting effect is limited to the local operators. A widely used flickering reduction technique involves smoothing of brightness differences between successive frames, such as proposed in [20, 21]. While this method may efficiently reduce
flickering caused by brightness and texture inconsistencies, its application should be avoided during scene changes since it will result in altering the artistic intent of the original content.

1.2 Motivation

Since the time that capturing and distribution of HDR video became possible, efficient video TMOs became a necessity and one of the main priorities for the entertainment and broadcasting industries and inevitably its development became a hot topic of research. In this thesis, we focus on HDR distribution and propose an efficient global, low complexity and content adaptive video TMO, addressing the inefficiencies of existing TMOs by significantly improving visual performance and offering a real-time solution to industry.

1.3 Thesis Organization

The rest of the thesis is structured as follows. Chapter 2 provides background information on Tone Mapping Operators and the challenges of tone mapping of HDR video sequences. Chapter 3 explains in detail our proposed Tone Mapping Operator for HDR video sequences. Furthermore, it presents the subjective tests that we performed to evaluate our proposed TMO and analyzes and discusses the results. Finally, conclusions and future work are drawn in Chapter 4.
Chapter 2: Background

2.1 High Dynamic Range (HDR) Technology

2.1.1 Overview

The popularity of the HDR technology is due to the fact that, unlike the existing SDR technology, it supports the capture, distribution and display of a range of luminance that is close to what the Human Visual System (HVS) can perceive and interpret. Here, we refer to Luminance as the intensity of light that is emitted from a display and it is measured in cd/m². As we observe on Fig. 2.1, HVS is capable of perceiving and interpreting a luminance range of more than 14 orders of magnitude. Note that orders of magnitude for dynamic ranges define the difference in powers of ten between the highest and lowest value of the range. The human eye can adapt simultaneously to only five out of the 14 orders of magnitude at one instance [22].

The luminance range that is supported by the HDR technology is between 0.005 and 10,000 cd/m² [23], which covers almost the dynamic range that the HVS can interpret simultaneously.

Figure 2-1. The dynamic range of real world scenes and capabilities of the human eye, cameras, and displays [1]
On the other hand, the SDR technology is limited to the range of 0.1 and 100 cd/m², with some of the latest high-end displays supporting luminance values up to 500 cd/m². This difference in the supported dynamic range results in better contrast and color representation by the HDR technology. The supported range of HDR allows the capturing of all the light and color information of a scene in contrast with the SDR technology where the supported range is limited to only a small proportion of the visual information present in a real-world scene. As a result, HDR content displayed on HDR display contains more details and more vivid colors than if it was captured using SDR technology and displayed on a conventional SDR display.

2.1.2 Perceptual Quantization

The Perceptual Quantizer (PQ) [24] is an inverse Electro Optical Transfer Function (EOTF) designed to optimize the distribution of light intensities with respect to the Human Visual System (HVS) properties. Although, traditionally cameras store light in a linear way, HVS does not

![Figure 2-2 Barten Ramp CSF (dashed line), PQ curve (black line), BT.1886 curve (blue line), and log curve (red line) [24]](image)
interpret the light the same way. In fact, humans interpret the same amount of light change differently, depending on whether it is daylight or moonlight. In general, HVS is more sensitive to light changes that happen in moonlight than daylight [25]. Based on the above observations, PQ is designed to convert light values to a perceptually linear domain, meaning that any variation of intensity at any brightness level is seen the same way by the human eye. On the other hand, an inverse EOTF aims at transforming physically linear values into perceptually linear values.

In digital imagery, inverse EOTFs transform linear light values into integer codewords for pixel representation. EOTFs mainly aim at encoding the highest amount of visual information using the lowest possible number of codewords. The degree of perceptual linearity that the transform domain of each EOTF offers, has a direct impact on its performance, since higher perceptual linearity leads to fewer codewords. The EOTF specifically designed for SDR technology is standardized as BT.1886 [26], also known as Gamma encoding. BT.1886 efficiently encodes the luminance values that fall into the SDR luminance range. However, the HDR technology offers a significantly wider luminance range with much larger peak values, making the BT.1886 unsuitable to be used as the primary EOTF for the new technology. To cover the need for efficient encoding of the luminance values in the range between 0.005 and 10000 cd/m², in 2012 SMPTE standardized the SMPTE ST 2084 inverse EOTF, also known as Perceptual Quantizer (PQ). The derivation of PQ is based on the peak sensitivity values of the Barten’s Contrast Sensitivity Function (CSF) [27]. The Barten’s CSF models the human vision contrast detection threshold with respect to spatial frequency, background luminance and viewing angle. Fig. 2-2 shows the Barten Ramp (dashed line), which separates the contrast steps that are noticeable by
HVS (pink area) form the contrast steps that the HVS cannot detect (green area). The ideal EOTF should sit exactly below the Barten Ramp on the green area, maximizing its efficiency. We observe that the PQ, using the relatively low bit-depth of 12 bits, manages to maintain almost the same small distance from the Barten Ramp for the entire HDR range. This results in an efficient distribution of the available codewords throughout the luminance range, since two consecutive codewords are assigned to two different luminance levels if their difference is just below the threshold that is just noticeable by the HVS. In contrast to PQ, BT.1886 (gamma encoding) and the logarithmic encoding, even using 15 and 13 bits respectively, fail to remain close to the Barten Ramp (Fig. 2-2). This results in oversampling of the bright areas when BT.1886 is used and oversampling of the dark areas in the case of the log.

2.2 Tone Mapping Operators

2.2.1 Overview

Since the time that capturing and distribution of HDR video became possible, efficient video
TMOs also became the need and priority for the entertainment and broadcasting industries as well as content providers and TV manufacturers. It is worth noting that although tone mapping results in SDR content with much less information and lower visual quality than the original HDR, it still offers much better visual experience than that of the same content originally captured by SDR cameras as shown in Fig. 2-3.

In addition to local and global classification, Tone Mapping Operators can also be classified as “offline” and “online” TMOs. The former are mainly used during the post production phase of video content (Fig. 2-4 (a)) while the latter are used in the transmission pipeline (Fig 2-4 (b)). Since offline TMOs are not applied during the transmission phase, they are not required to be neither real time or adaptive. Therefore, local TMOs are used for post processing since they can

![Figure 2-4. HDR distribution pipelines: (a) Offline TMOs may be used during post production; (b) real-time TMOs are needed for this broadcasting approach.](image)
deliver good quality SDR by using the spatial filters that are manually tuned before each scene to adapt to content characteristics and thus try to preserve all the important visual information. On the other hand, online TMOs are required to be real time and automatically adaptive to the content without requiring any user input. These limitations are derived from industry requirements for live broadcasting applications.

2.2.2 Challenges of Video Tone Mapping and Scene Changes Preservation

Tone mapping of HDR videos is a relatively new research topic as HDR video content has only recently become widely available. Naïve, straightforward application of image TMOs on video sequences results in visual artifacts such as ghosting and brightness flickering [17,18,19]. Local TMOs are more prone to the ghosting effect than global TMOs, since fast changes of the spatial neighborhood among successive frames may affect differently parts of each frame. Tuning of the spatial filters that are used by local TMOs before tone mapping of each scene can effectively reduce ghosting, but that is time consuming.

On the other hand, both local and global TMOs are prone to brightness flickering. Flickering is mainly caused by significant variation of statistical information, including maximum/minimum/average brightness and the brightness histogram between consecutive frames. The most common causes of brightness difference between HDR frames are scene changes, motion, and imperfect HDR content post processing and grading. When these variations occur among successive frames, TMOs tend to enlarge them through the mapping process to the new values, resulting in noticeable brightness differences between consecutive SDR frames.
These abrupt changes are desirable when they are due to scene changes, but they result in flickering if they are generated by any other cause.

A widely used flickering reduction technique involves smoothing of brightness differences between successive tone-mapped SDR frames [20, 21]. While this technique may efficiently reduce flickering caused by brightness and texture inconsistencies, its application should be avoided during scene changes, since it may result in altering the artistic intent of the original content. By continuously smoothing the brightness differences between successive frames, the mapping scheme cannot fully adapt to the new scene, resulting in loss of visual information.

Scene changes in a video sequence can either be abrupt (hard cut) or gradual (soft cut) [28]. Hard cut scene changes result in large variations of statistical brightness information of the frames. Hence, their detection can be considered more trivial than the soft cut changes, which result on similar variation of image statistics generated by other causes, such as motion or imperfect post processing. An ideal flickering reduction method should be able to detect scene changes and separate them from all other causes of brightness changes, thus preserving the artistic intent of the content.

2.2.3 Offline Video TMOs

As we described before, offline TMOs are not required to be neither real time nor adaptive, since they are mostly used during post production process. During the offline process, artists have access to unlimited number of frames to achieve flickering reduction, which is not the case with online applications where the frames used depend on the buffer size and real-time requirements.
One of the state-of-the-art video tone mapping methods is the local operator described in [29], which first decomposes HDR frames to a base layer and detailed (edges) and then maps these layers separately so that the frame details are preserved. Forward and backward optical flow estimation is used for wrapping each frame’s temporal neighborhood so that the temporal consistency is ensured and ghosting is minimized.

Another state-of-the-art video TMO, is the Zonal TMO presented in [30]. This is a local TMO that also decomposes the frames to a base and detailed layers and maps them separately. Along with this decomposition, it also separates each frame into large segments depending on their brightness. Due to the fact that the number of segments and their corresponding boundaries change from frame to frame, they track these segments throughout the entire sequence, defining video zones on each detected scene. The temporal coherency is ensured by eliminating the brightness differences on the defined video zones for each scene on the tone mapped SDR sequence.

In [31] authors propose an image TMO, which is also extended for video applications. This image TMO aims at preserving the contrast of the original HDR images by minimizing the contrast distortions caused by dynamic range compression. This TMO works well when applied on images with middle range luminance, but results in noise when it deals with dark scenes. For video applications, authors suggest to apply a low pass filter on the mapping curve to restrict extreme changes of the curve between successive frames, smoothing the brightness differences on the resulting SDR video sequence in expense of visual efficiency.
In [32] authors extend the TMO introduced in [31] by introducing a spatial filter to decompose the image in base and detailed layers. They propose to map the base layer using the method introduced in [31], which they also improve by lowering the complexity of the method as well as by introducing a noise control function to overcome the problem with the night scenes. Finally the details are restored to the frame, delivering noise free SDR content with rich global and local contrast.

Finally, in [33] authors propose a TMO that mainly aims at retaining the perception consistency of an object through the SDR scene. The proposed method first calculates the average luminance of the entire HDR scene and restricts accordingly the mapping process of each frame of the scene. This method also ensures the overall contrast consistency among consecutive frames of the resulting SDR video sequence.

2.2.4 Online Video TMOs

Online TMOs are designed to be used in the transmission pipeline and as such they are real-time. Thus, they should deliver high quality SDR content while meeting the industry requirements for broadcasting applications. These requirements include low computational complexity, content adaptiveness and no user interaction. Regarding the visual artifacts caused by tone mapping, online real-time TMOs should reduce them using information from a limited number of consecutive frames, in contrast with the offline TMOs that tend to use all the frames of a scene.
One state-of-the-art online TMO is presented by Kiser et al. in [34]. This global TMO extends the photographic operator, presented in [9], to be applied on HDR video sequences. For each HDR frame, this TMO measures the black and white levels of the previous tone mapped frame and restricts accordingly the photographic operator during the calculation of the mapping curve for the current frame. This way, the mapping curve is not allowed to rapidly change between successive frames, resulting in flickering reduction in the SDR sequence.

Another popular online TMO is the CameraTMO presented in [35]. This TMO represents the S-shaped curve that is used by most of the consumer cameras to map the captured light values to the color gamut of the required storage format. The S-curve of this TMO is derived from the experimental measurements using the Canon 500D DSLR. Similar to the previous method, flickering reduction is achieved by restricting the changes of the mapping curve among successive frames.

All of the described TMOs, offline and online, seem to reduce flickering, but it is not clear how these methods handle scene changes. The common way of reducing flickering for all these methods is to smooth the brightness differences between successive frames. However, smoothing the brightness of frames that belong to scene changes results in loss of visual information and alteration of the artistic intent of original HDR sequence. Exception to this is the Zonal TMO [30], which requires as input the entire HDR sequence and thus detects and preserves all the scene changes.
Chapter 3: Our Content Adaptive Tone Mapping Operator

3.1 Introduction

In this thesis, we propose TMO that is designed with three main objectives in mind: low computational cost, content adaptation, and preservation of the original HDR artistic intent. The first two objectives are derived by the demands of real time applications, especially for live broadcasting. The last one is a must for content owners and artists and a requirement enforced upon broadcasters and delivery service providers.

3.2 Our Proposed Method

3.2.1 Overview

The workflow of our TMO is presented on Fig. 3-1. Our method could be divided into three main components, the Display Model, the Tone Mapping process and the Flickering Reduction

![Figure 3-1. Block diagram of the proposed TMO](image-url)
scheme. The display model determines the technical characteristics of the destination display such as the available luminance range and the supported bit-depth. The tone mapping process, using the parameters of the display modes, derives the mapping curve and consequently maps the HDR content to SDR range according to the capabilities of the destination display. Finally, the flickering reduction scheme smoothen the brightness differences caused by inconsistencies among successive frames, while avoiding scene changes and thus preserving the original artistic intent. The following sub-sections describe in detail the three main components of our scheme.

### 3.2.2 Display Model

Nowadays, a large variety of commercial displays with different projection capabilities (Fig. 3-2) is available in the market, making the need for TMOs that prepare the SDR image according to the capabilities of the targeting display(s) a necessity. Our TMO takes as inputs the maximum and minimum luminance level of the targeting display, the supported bit-depth and the used EOTF. Different combinations of these parameters result in different SDR tonal levels, having a direct impact on the way that images/frames are projected by the display.

![Commercial displays with different dynamic range and bit-depth support](image)

<table>
<thead>
<tr>
<th></th>
<th>Standard SDR TV</th>
<th>Samsung J950</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminance Range</td>
<td>0.1-100 cd/m²</td>
<td>0.01-1000 cd/m²</td>
</tr>
<tr>
<td>EOTF</td>
<td>BT.1886</td>
<td>BT.1886</td>
</tr>
<tr>
<td>Bit Depth</td>
<td>8bit</td>
<td>10bit</td>
</tr>
</tbody>
</table>

Figure 3-3. Commercial displays with different dynamic range and bit-depth support
These parameters can be either retrieved through the HDMI connection or entered by the user. However, if they are not available, then our TMO uses the BT.1886 recommended values: maximum luminance = 100 cd/m2, black level = 0.1 cd/m2, bit-depth = 8 bits/pixel and EOTF = Gamma Encoding (BT.1886).

### 3.2.3 Tone Mapping Process

#### 3.2.3.1 Histogram Equalization

Histogram Equalization [36] is a well-known image processing technique, which increases the global contrast of an image. In most images, pixel intensities are usually gathered in a small portion of the available range, typically the center. Histogram equalization redistributes the contrast intensities to cover the entire available range. As it is shown on Fig.3.3 (a) the technique first computes the histogram of an image and the Cumulative Distribution Function (CDF) (red line) as follows:

\[
P(b) = \frac{\sum_{n=0}^{f} f}{f}
\]

CDF, which is a set of slopes, is used as a mapping curve to proportionally project the bins to the available Luma range according to their size (see Fig. 3.3 (b)). In other words, large histogram bins get assigned to higher portions of the available range than the bins that represent contrast intensities with low frequency. The histogram of the resulting image is presented in Fig. 3.3 (c) where it is clear that the tonal levels from the entire range are used to represent the picture, resulting in global contrast enhancement (Fig. 3.3 (d)). One disadvantage of histogram equalization is that it is prone to visual noise when the original image is relatively dark or
contains uniform areas [37]. In these cases, only a limited number of contrast intensities are needed to represent all the visual
Figure 3-3. Histogram Equalization (a) Initial histogram of image brightness and initial CDF; (b) Allocation of the destination range to histogram bins; (c) Histogram of brightness of the resulting image; (d) Testing image before (left) and after the histogram equalization (right).
information. Consequently, a lot of pixels have the same or similar brightness intensities, forming enormous histogram bins. Due to the proportional nature of histogram equalization, these enormous bins are projected to a wide portion of the available range, overspreading the visual information with result the visual noise amplification, especially in dark scenes. Furthermore, due to the fact that these enormous bins get assigned to most of the available range, only a relatively small portion of the range is left for the rest of the bins of the image, resulting in loss of information.

The use of the histogram equalization for tone mapping applications was first introduced by Larson et al. [8]. Since the SDR range is way smaller than the original HDR range, the loss of tonal levels is inevitable. However, the proportional nature of histogram equalization ensures that bins that represent luminance levels with high frequency of appearance are projected to the new range while bins with none or a few pixel members are dismissed. The authors propose to build the histogram of the luminance values in the log domain instead of the linear one since the response of the human vision system to luminance variations is approximately log linear when photoreceptors are fully responsive [38].

However, as we described in Section 2.1.2, the PQ domain is more perceptual linear than the log domain. In other words, it describes better how HVS interpreters light changes. Therefore, we propose to calculate the histogram of the brightness values of a frame in the PQ domain in order to overcome the low perceptual linearity of the log domain for low luminance values. We propose to use a histogram that consists of 1024 bins, whose size is equal to the luminance range of one PQ codeword for 10 bits per pixel HDR content with range 0.005-10000 cd/m². The
reason for assuming 10-bit quantization is that this has been shown to be sufficient for representing natural images [39]. By calculating the histogram in the PQ domain, we ensure that a bin with some population (pixels) indicates visible information while an empty bin indicates that no visible information exists for that luminance level. In other words, our PQ histogram is representative of the way that HVS interprets light intensities. As explained before, the CDF is calculated based on the frequency of each bin and it is used to initially map the HDR image to the significant smaller SDR range. This straightforward mapping may cause two major problems. First, it may project large HDR bins to visible SDR tonal levels that never existed in the original HDR content, thus resulting in visual noise or banding. Second, the rest of the HDR bins with useful visual information may not be represented on new SDR image due to lack of available SDR tonal levels. Therefore, to address the first issue, we use a ceiling method and to deal with the second issue we propose a novel redistribution function of tonal levels, to improve the visual quality of the resulting SDR image.

3.2.3.2 Ceiling Function

When Larson et al. [8] proposed to use the histogram equalization as TMO, they proposed two ceiling functions in to overcome the issue with the uniform areas. These ceiling functions limit the range on the SDR domain that each bin projected, aiming at the reduction of visual noise on the final SDR image. The first ceiling function is linear and sets the same maximum allowed number of SDR tonal levels for all the bins projections, regardless of the luminance level on the original HDR image each bin represents. However, since HVS does not interpret light in a linear way, this function may still result in visual noise.
To further improve the performance, another ceiling is proposed which is derived from the Blackwell CSF [40]. In this case, the thresholds for the projection of each bin on the SDR range are set based on the luminance level that the original bins represent on the HDR range, reducing effectively the visual noise of the resulting SDR images. On this function, the authors only define five luminance levels, the highest of them starting on 1000 cd/m\(^2\) as it was the large end of the range for most of the HDR images back in 1997. In other words, they propose ceiling the same way all the bins that represent luminance values higher than 1000 cd/m\(^2\) on the original HDR domain. However, the dynamic range of nowadays HDR contain is between 0.005 cd/m\(^2\) and 10000 cd/m\(^2\) and the HVS do not interpreter linearly the light ranging between 1000 and 10000 cd/m\(^2\) as the ceiling function does. Therefore, this ceiling function may fail effectively preserves the image details on the case of bright HDR images.

To overcome the issues with the perceptual function proposed by Larson et al. [8] we propose a novel ceiling function that is based on the HVS properties. The proposed ceiling function ensures that the amount of visual information that is mapped from each bin to the destination range is less or equal to the original amount of visual information that each bin represents on the HDR image.

Direct relation of the number of tonal levels in an HDR bin with the number of tonal levels in the projected SDR range is impossible, since HDR values correspond to linear light intensities while SDR values are relevant light intensities and strongly depend on the display characteristics. For this reason, we transfer both representations to the PQ domain in order to compare them. Fig. 3-4 illustrates the process followed to map the CDF of the histogram equalization to the destination
range (LPQ) and then using parameters from the Display Model (EOTF and bit-depth) to calculate the SDR tonal levels. To convert SDR integer values to physical luminance values, one needs to assume the capabilities of the display on which the content will be reproduced. That is why, our ceiling function will take as input parameters the capabilities of the targeted display (black level, white level, EOTF and bit-depth). Using such a display model and the PQ function, we can convert SDR integer values to the expected light intensity emitted by the targeted display (LSDR). Finally, these light intensities are converted to PQ domain using the PQ transfer function and quantized to 10 bits.

Please note that in 10-bit PQ quantization, two consecutive codewords represent two luminance levels that are just noticeable by the HVS. Combining this with the fact that each bin in the HDR domain represents only one PQ codeword, we can safely assume that if the projection of any HDR bin on the SDR range is assigned to more than one PQ codeword, then this generates visual information that does not exist on the original HDR image (denoted as Overloaded projection).

Figure 3-4. Steps involved in transferring the HDR and SDR ranges to PQ domain
On the other hand, if for the projection of a HDR bin on the SDR image is used only one codeword, then the same amount of visual information is reproduced. Finally, if less than a codeword is used, then we have loss of visual information. Such a loss is acceptable since tone mapping is about selecting which information should be preserved and which dismissed. The proposed ceiling limits the overloaded projections to exactly one PQ10 codeword, thus avoiding the generation of visual noise and banding.

Larson et al. [8] avoid redistribution of the ceiled tonal levels, thus not really achieving the full theoretical potential and visual quality. In the following subsection, we describe our method for redistributing the saved tonal levels (codewords) by the ceiling process to further improve the visual quality of SDR.

3.2.3.3 Our Redistribution Method

As described above, the largest HDR bins were assigned more than necessary codewords in SDR, while some of the rest of the bins did not have a fair share of codewords. A logical way of redistributing would be to redistribute the saved tonal levels by ceiling to the largest of the “unfairly” treated HDR bins. While this redistribution may works relatively well increasing the overall global contrast for most of scenes, it method may fails to deliver desirable results in the extreme cases of very dark and very bright content. In these cases, most of the largest “unfairly” treated HDR bins are still located on the same end of the histogram as the bins that resulted in overload projections and got ceiled. Therefore, if the saved codewords are allocated for the projection of these bins which are located on the same end of the histogram, as is the case for
dark and bright scenes, the resulting SDR content becomes darker if the original scene was bright and vice versa.

To overcome this issue, we introduce a novel dynamic redistribution function that detects and handles these two extreme cases. We first propose to categorize scenes into dark, normal and bright by using the geometric mean of the PQ HDR brightness values. Our experiment results show that when the geometric mean of a frame falls below the 8% of the dynamic range, then the frame is categorized as dark. In the case that the frame geometric mean falls between 8% and 90% of the dynamic range, then the frame represents a normal scene. Finally, if the geometric mean is above 90% of the dynamic range, then the frame represents a bright scene.

The proposed redistribution method handles the normal frames as the method described by assigning the saved codewords for the projection of the largest “unfairly” treated HDR bins, regardless of their location in the histogram. In the case of a dark or bright HDR scene, our redistribution method allocates the saved tonal levels (codewords) to the projections of the largest “unfairly” treated HDR bins that are located farther away from the majority of the largest HDR bins, i.e., centre and the other side of the histogram. In order to avoid project HDR bins with little or no information to the SDR domain, we introduce an “eligibility” threshold for the frequency of luminance levels on the HDR histogram equal to 0.005%, below which the corresponding HDR bins are not projected to the SDR domain. We came up with this threshold after empirically testing a large set of representative HDR frames. The proposed redistribution method efficiently handles the two extreme cases where most of the image intensities are gathered on any of the ends of the histogram.
The proposed redistribution method effectively increases the global contrast while the artistic intend and the natural looking of the original HDR image are retained as well.

### 3.2.4 Flickering Reduction and Scene Changes Preservations

A common visual artifact introduced by global TMOs when applied on HDR video content is brightness flickering. In general, flickering reduction methods apply a temporal low pass filter on the mapping curve in order to control severe mapping variations among successive frames, as proposed in [20, 21]. Although these methods effectively reduce flickering, it is not clear how they avoid smoothing scene changes in the process, which in turn will result in undesirable visual artifacts. In this work, we propose a flickering reduction method that aims at detecting and preserving the scene changes while effectively reducing flickering for the rest of the video sequence.

In order to maintain complexity at low levels for real-time implementations, we first chose to use the variation of the geometric mean of the brightness of the HDR frames on PQ space as indicator for scene changes. Our experimental results showed a hard-cut scene change results in a geometric mean variation greater than 28 PQ10 codewords between two successive frames. However, in the case that the variation of the geometric mean between two consecutive frames falls in the range between 8-28 PQ10 codewords, our experiments show that the resulting inconsistencies may be caused by scene changes, motion, imperfect HDR content post processing or grading. Finally, in the case that the detected variation of the geometric mean is less than 8 PQ10 codewords, then the odds of a scene change are very slim. Therefore,
measuring the variation of the geometric mean of the brightness between successive frames, cannot efficiently distinguish a scene change from the rest of the variation causes. Trying to overcome the problem, we conducted experiments using as scene detection indicator the variation of other low complexity statistical metrics on the brightness information such as the arithmetic mean, the median, the histogram skewness and the histogram kurtosis, variance and standard deviation. Unfortunately, none of these metrics managed to give a better estimation than that of the geometric mean.

Therefore, we decided to depend on the use of the geometric mean. For the case of a scene change detection, variation larger than 28 PQ10 codewords, we propose to do not apply any filter on the mapping curve. That way the curve can be adapted rapidly to the new scene, preserving the way that the scene change is represented on the original HDR video sequence. In the case that distinguishing scene changes from other causes is not trivial, geometric mean variations between 8 – 28 PQ10 codewords, we chose to apply a low pass IIR filter on the mapping curve. Experimental results showed that a 3-tap with cut off frequency of 2Hz gave the best visual performance. This is a trade-off between avoiding to significantly smooth scene changes, while reducing the amount of noticeable flickering caused by other sources of brightness variations. Finally, in the case that no scene change is detected, i.e., geometric mean below PQ10 codewords, we propose to use a temporal 7-tap low pass IIR filter. The cut of frequency of the proposed 7-tap filter is 0.5Hz as in [10], since this limit falls below the detection threshold of temporal changes for human observers. We use a 7-tap IIR instead of the 3-tap low pass FIR filter suggested in [21], since sometimes the latter fails to absorb effectively flickering as it is shown in the results section.
3.3 Results and Discussions

Here, we evaluate the performance of our TMO in terms of overall visual quality. On the following sub-sections we show visual results obtained by our method and the TMO proposed by Larson et al [8]. To further evaluate our proposed TMO, we performed subjective evaluations comparing the proposed method with existing state-of-the-art TMOs [17,41,42]. For our tests, we used MPEG HDR video sequences that contain only one scene. Therefore, to evaluate the performance the flickering reduction and scene detection and preservation part of the proposed method, we also conducted a statistical experiment.

It is worth noting that our newly developed method has the potential to be implemented in real-time for the tested FullHD resolution video streams. Without considering I/O operations, our unoptimized, single thread, Matlab implementation requires on average 0.2 seconds to process a color frame at FullHD resolution, on a computer with CPU Intel i7 5820k. Therefore, it can be classified as online TMO.

3.3.1 Visual Results on Quality of Tone Mapped Images

In this section, we compare the visual quality of SDR generated by our TMO and that of Larson et al. [8], the closest TMO to ours. We measure visual quality in terms of contrast and preservation of the original artistic intent.
Fig. 3-5a depicts the resulting SDR image using the linear ceiling presented by Larson et al. [8].

We observe that the image suffers from noise, which is explained by the fact that the same threshold is applied on the size of each bin regardless of its luminance level. The second ceiling
Figure 3-5. Comparison between Larson et al. linear ceiling (a), perceptual ceiling (b) and the proposed method (c).
method proposed by Ward et al. [8], which is based on HVS, efficiently removes noise as illustrated in Fig. 3-5 (b). However, the overall brightness of the image is dimmer and objects like the speaker and the audience are hardly visible. Fig. 3-5 (c) presents the results obtained using the proposed method. It is worth noting that we manage to eliminate the noise while keeping most of the visual information such as the speakers and advertisement on the railings. Finally, details on the bright part of the image, like the flame, the smoke and the person in front of the flame are also preserved by the proposed method.

Furthermore, we compare the two methods (perceptual Ward and the proposed one) using as input a compressed HDR image. Since our method aims at preventing large uniform areas from being emphasized, it should also prevent common compression artifacts such as blocking. Fig. 3-6 (a) shows the resulting SDR image obtained using the perceptual ceiling method of Ward et al. [8]. We observe that this image suffers from blocking artifacts in the cloudy sky (red square), which is magnified in Fig. 3-6 (a1). On contrast, our method handles better the compressed images, as blocking artifacts are less visible as we observe on Fig. 3-6 (b) and the magnified square shown in Fig. 3-6 (b1).
Figure 3-6. Comparison on compressed images: (a) Larson et al. perceptual ceiling; (a1) magnified square of (a); (b) proposed method; (b1) magnified square of (b)
3.3.2 Subjective Evaluations

Since our TMO is a real-time content adaptive real-time method, we first compared it with the state-of-the-art online TMOs, Kiser et al [34] (denoted as “Kiser”) and Camera TMO [35] (denoted as “Camera”). In addition, we also compared it with non-real time methods (offline TMOs) such as the Display-adaptive global TMO [30] (denoted as “Mantiuk”), the Temporal coherence global TMO [32] (denoted as “Boitard12”), and the Zonal temporal coherence local TMO [33] (denoted as “Boitard13”), which are shown to have high visual performance [17,41,42]. Regarding the video sequences, we used the following representative HDR MPEG video streams: FireEater2, Market3, Sunrise, Balloon, EBU_Hurdles and EBU_Starting. All of the sequences are natural outdoor videos, FireEater is a night shot, Market, BallonFestival and SunRise are daylight scenes and Hurdles and Starting are daylight sports content. For our tests we used in total three displays, the HDR display Sim2 HDR47E and two Samsung KS9800 as SDR displays.

We performed three independent subjective tests. For the first test, the evaluation method was Side-by-Side with the resulting tone mapped video of all the compared TMOs on the SDR display and original HDR video on the HDR display as reference [43]. The scale for visual performance ranged from 1 to 10, depending on the visual fidelity (the brightness, colors, contrast, details, artistic intent, etc.) of the SDR to the reference HDR video sequence. The order of videos in each session of the test was randomized and extra care was taken for the results of the same TMO not to be shown consecutively. For detecting the outliers, we used the method proposed in [43].
For the second subjective test, the evaluation was Sibe-by-Side, using two SDR displays. In this test, each pair of videos was the resulting SDR sequence, produced by two out of the six TMOs. We followed the procedure for simultaneous paired comparison as it is described in [44]. The viewers had to choose between A, B or the same in the case that they could not detect any difference between the two stimuli. Finally, we performed a subjective evaluation similar to the previous one, but having also the HDR display as reference. Therefore, in total we used three displays for this test. In this case, the viewers also had to choose again between A, B or the same. For both tests, the video pairs were randomized in each session.

Eighteen subjects (10 males and 8 females) participated in the tests. All the subjects were tested for visual acuity and color vision as it described in [43]. They were non-expert viewers, with negligible experience of HDR video viewing. The average age of the subjects was 28 years old. Each test session was composed of all the three described subjective tests whose order was randomized between the test sessions. Prior to each subjective test on each test session, a training session was given to introduce the test procedure and rating task, using a set of training stimuli. After collecting the subjective results, the outlier subjects were detected according to appropriate method for each test. Two outliers were detected in these tests and their rates were discarded from the results.

Fig. 3-7 depicts the results of the subjective test in terms of visual fidelity of the tone mapped sequences with the original HDR content for all the tested TMOs. The Mean Opinion Score (MOS) for each TMO was calculated by averaging the scores over all the subjects with 95% confidence interval. We observe that subjects ranked the proposed method (denoted as “DML”)
as the closest to the original HDR content for all the testing sequences. Please note, that the other two adaptive, low complexity TMOs, KiserTMO [34] and CameraTMO [35], were ranked by the subjects much lower than the proposed adaptive method. Finally, the results for the sequence “FireEater”, which depicts a dark scene, indicate the efficiency of the ceiling and redistribution methods. In this case, the rest of TMOs, except Boitard13 [30], delivered SDR results with visual noise as shown in Fig. 3-8.

Figure 3-7. Results of the subjective tests where subjects had to rank the visual fidelity of the tone mapped sequences with the original HDR content for all the tested TMOs.
Figure 3-8. Visual results for the sequence “FireEater” using (a) the proposed method, (b) Mantiuk, (c) Boitard12, (d) Boitard13, (e) Kiser, and (f) Camera.
Fig. 3-9 depicts the results of the Side-by-Side subjective evaluation, using two SDR displays, without providing the original HDR video as reference. The results are presented as % of subjects choosing our method (y-axis) over the other 6 methods (x-axis), for each of the video sequences. We notice that for all the testing sequences, 85%-100% of the subjects preferred the SDR version generated by our TMO from the versions of the rest online TMOs, KiserTMO [34] and CameraTMO [35]. Furthermore, we observe that, for all sequences, the proposed method outperforms all the tested TMOs, including and the offline TMOs. The large error bars on the results of Mantiuk [31] for the sequences Market, Hurdles, Starting and SunRise, indicate that a significant number of subjects chose the resulting SDR videos of DML and Mantiuk [31] as

Figure 3-9. Results of the subjective tests where subjects used two SDR displays without the original HDR content as reference. The results are presented as % of subjects choosing our method (y-axis) over the other 6 methods (x-axis)
“equal”. However, the majority of the rest of the subjects (average above 70%) preferred the DML SDR quality.

The visual results of the first frame of the sequence “Starting” are shown on the Fig. 3-10. This sequence represents a medium to bright scene. We observe that the online TMOs result in dim SDR images (Figs. 3-10 (e) and 3-10 (f)), altering drastically the artistic intent of the original HDR content. On the other hand, Boitard12 [16] (Fig. 3-10 (c)) does not preserve enough details especially on the sky, while Boitard13 [30] overdenotes the edges of the image, giving to it a “cartoonish” effect. Finally, we observe that the SDR image produced by our operator (Fig. 3-10 (a)) retains the global and local contrast of the original HDR frame. However, the difference with Mantiuk’s TMO [31] (Fig. 3-10 (b)) is only visible if the viewer focuses on the clouds, where the proposed method preserves more details. It is for this reason that, a significant number of subjects voted the DML and Mantiuk [31] as “equal”.
Figure 3-10. Visual results of the first frame of the sequence “Starting” using (a) the proposed method, (b) Mantiuk, (c) Boitard12, (d) Boitard13, (e) Kiser, and (f) Camera
Finally, Fig. 3-11 presents the results of the Sibe-by-Side evaluation where the original HDR video was provided as reference. We observe that our TMO outperforms the other two online TMOs (KiserTMO [34] and CameraTMO [35]) with a higher preference than the test without a reference. For the rest of the TMOs, the results are similar to the ones from the previous test, with the exception of Boitard12 which seems to offer better visual results for the sequence “Hurdles”.

Fig 3-12 presents the visual results of the first frame of the sequence Hurdles. Even though Boitard12 [33] (Fig. 3-12 (c)) does not preserve enough details compared to our method, subjects seem to prefer the brighter version of Boitard12.

Figure 3-11. Results of the subjective test where subjects used the original HDR content as reference. The results are presented as % of subjects choosing our method (y-axis) over the other 6 methods (x-axis).
Figure 3-12. Visual results of the first frame of the sequence “Hurdles” using (a) the proposed method, (b) Mantiuk, (c) Boitard12, (d) Boitard13, (e) Kiser, and (f) Camera
In summary, our proposed method significantly outperforms the two state-of-the-art automated real-time TMOs (KiserTMO [34] and CameraTMO [35]). In addition, our evaluation tests have shown that our method also outperforms TMOs with higher computational complexity as they require spatial filtering (Mantiuk [31] and Boitard13 [30]), iterations (e.g., Boitard13 [30] and Boitard12 [33]) and user input interaction (e.g., Boitard13[30]).

### 3.3.3 Statistical Analysis for Flickering Reduction

In this sub-section, we evaluate the performance of our scene detection and flickering reduction methods. Since all the HDR sequences that we have access did not have hard cut scene changes, we created a ten seconds HDR video with frame rate of 24 frames per second, by combining 48 frames (2 seconds) from each of the following five MPEG sequences: Balloon Festival, FireEater2, Starting, Hurdles, and Sunrise.

Fig 3-13 shows the results on the flickering reduction effectiveness comparing the resulting SDR and the original HDR sequences. Fig 3-13 (a) depicts the geometric mean of the brightness, calculated on the PQ domain, for the first 48 frames (2 seconds) of the HDR testing video clip where no scene change exists. Fig 3-13 (b) depicts the geometric mean of the resulting tone mapped HDR video sequence without applying any filter (dotted line), applying the filter proposed in [31] (dashed line) and applying the proposed method (solid line). We notice that the 3 tap low pass filter [31] fails to efficiently absorb the severe brightness distortions that the TMO introduces on the resulting SDR video sequence. However, by applying our proposed method, most of the large variations are absorbed and the flickering is efficiently reduced. In this scenario, since the variations of the geometric mean of the brightness of the HDR frames were in
the range between 1 and 6 PQ10 codewords, lower than the threshold of 8 PQ10 codewords, our method applies only the proposed 7-tap low pass IIR filter to reduce flickering. Fig 3-14 depicts the results on the scene changes preservation. Fig 3-14 (a) shows the geometric mean of the brightness, calculated on the PQ domain, for the entire testing video sequence that includes Balloon Festival, FireEater2, Starting, Hurdles, and Sunrise. Fig 3-14 (b) shows the geometric mean of the resulting SDR video sequence by applying the method proposed in [31] (dashed

Figure 3-13. Statistical evaluation of flickering reduction. (a) The Geometric Mean of the 48 first Frames of the HDR video sequence. (b) The Geometric Mean of the tone mapped HDR video sequence without any filter (dotted line), applying proposed in [31] (dashed line) and applying the proposed method (solid line).

Figure 3-14. Statistical evaluation of scene change detection. (a) The Geometric Mean for all the frames of the testing HDR sequence. (b) The Geometric Mean of the tone mapped HDR video sequence applying the method proposed in [31] (dashed line) and applying the proposed method (solid line).
line) and our proposed method (solid line). We observe that the continued application of the low pass filter [31] results in brightness inconsistencies on the tone mapped video sequence, as there is no special provision for the scene changes. A visual representation of the loss of visual information when the low pass filter is applied during the scene changes is shown in Fig. 3-15. At the scene change, the previous frame is depicting a dark scene, and thus the mapping curve is adapted to preserve visual information that comes from HDR pixels with low luminance values. However, the next frame, is depicting a much brighter scene where most of the pixels have middle to high luminance values. By applying a low pass filter during the scene transition as suggested in [31], the mapping curve cannot instantly adapt to the new scene, resulting in loss of visual information on the bright areas of the scene, such as the clouds on Fig 3-15 (a). On the other hand, however, by using our proposed scene detection method, we manage to successfully
avoid filtering the scene change, thus maintaining the original details and artistic intent as shown in Fig. 3-15 (b).

In summary, our tone mapping approach is content adaptive as it is based on histogram equalization and perceptual quantization, which allow us to optimize the mapping curve according to the input. From a complexity point of view, our method has the potential to be implemented in real-time as our tests showed that its Matlab based employment requires on average 0.2 seconds to process a color frame at FullHD resolution, which is well within the real-time requirements of such software implementation. Finally, our objective of delivering the best possible visual performance is achieved as our subjective tests show that we outperform all existing state-of-the-art online and offline TMOs.

Figure 3-15. Visual artifacts caused by filtering scene changes: (a) existing method [31], (b) our proposed method
Chapter 4: Conclusion and Future Work

4.1 Conclusion

In this thesis, we address the backward compatibility challenge in the emerging delivery of HDR content by investigating the possibility of improving the visual quality of the state-of-the-art real-time online TMOs, aiming at delivering the best possible SDR quality for broadcasting applications.

In Chapter 3, we proposed a novel, efficient, global, low complexity and content adaptive video TMO. The proposed TMO is based on histogram equalization and takes advantage of the HVS properties by using the defined domain as well as the properties of the perceptual quantization. We introduce a flickering reduction method that smooths the brightness and content inconsistencies between consecutive frames caused by the mapping operation. Finally, we propose a scene detection method which allows us to avoid smoothing frames corresponding to scene changes, thus further improving the overall SDR visual quality and maintaining the artistic intent of the original HDR content.

Subjective evaluations show that our approach outperforms the state-of-the-art online and offline TMOs. An average of 95% of the subjects used preferred the SDR content produced by our TMO over the state-of-the-art real-time TMOs. When compared with the offline TMOs, an average 75% of the subjects preferred the SDR produced by the proposed method. Finally, statistical evaluation showed that the proposed flickering reduction method in combination with our scene detection approach efficiently reduce flickering introduced by tone-mapping while preserving the original look of scene changes.
4.2 Future Work

Since HVS is more sensitive to luminance changes than color changes, the majority of TMOs mainly focus on compressing the dynamic range of luminance. However, this approach results in SDR colors being somewhat different from the original HDR colors, with the SDR colors looking more saturated. Therefore, color correction after tone mapping is required to restore the original colors. Over the years, methods such as the ones described in [45] and [46] have been proposed to address the color shifting after tone mapping, but they do not meet the requirements for real time applications. Our plan is to investigate the use of different color spaces, such Lab or ICTCp, to address color shifting and explore the possibility of their real-time implantation in compression standards that only support the existing YCbCr color space.
Bibliography


