

UNIVERSITY OF TARTU

FACULTY OF SCIENCE AND TECHNOLOGY

Institute of Technology

Robotics and Computer Engineering

Aygül Salahlı

Comprehensive Study on High Dynamic
Range Tone Mapping with Subjective
Tests

Master's Thesis (30 ECTS)

Supervisors: Assoc. Prof. Gholamreza Anbarjafari,

Dr. Cagri Ozcinar

Tartu 2017

Resümee/Abstract

Põhjalik uuring ülisuure dünaamilise ulatusega piltide toonivastendamise kohta koos subjektiivsete testidega

Ülisuure dünaamilise ulatusega (HDR) pilte iseloomustab väga suur eredusvahemik, mida traditsioonilised madala dünaamilise ulatusega (LDR) monitorid kuvada ei suuda. Seetõttu transformeeritakse HDR pildid enamasti 8-bitilisele kujule, nii, et iga piksli alfa kanalit kasutatakse eksponendi väärtusena. Kirjeldatud notatsiooni nimetatakse eksponentsiaalseks. Toonivastendamise operaatoreid (TMO-sid) kasutatakse selleks, et piksleid kokku pakkides muuta ülikõrge dünaamilise ulatusega pilte madala dünaamilise ulatusega piltideks, et LDR monitorid saaksid neid kuvada. Antud lõputöö eesmärgiks on leida ja analüüsida erinevusi ning sarnasusi erinevate teaduskirjanduses leiduvate toonivastendamise operaatorite vahel. Iga mainitavat TMO-d on analüüsitud subjektiivsetes uuringutes, võttes arvesse erinevaid tingimusi nagu näiteks taust, eredus ja värv. Lisaks analüüsiti erinevaid inversiivseid toonivastendamise operaatoreid, säri fusiooniga HDR vastendusi, histogrammi kohandamist ning Retinex'i filtreerimist. Kasutades erinevaid HDR pilte, uuriti 19 erinevat TMO-d, mille keskmise arvamuste skoori (MOS-i) arvutamiseks küsitleti 25 erinevat inimest, võttes arvesse nende vanust, nägemisvõimet ja värvipimedust.

CERCS: T120 Süsteemitehnoloogia, arvutitehnoloogia; T125 Automatiseerimine, robotika, control engineering; T111 Pilditehnika

Keywords: Ülisuure dünaamilise ulatusega pildid, toonivastendamine, piltide kuvamine, HDR piltide matemaatiline modelleerimine.

Comprehensive Study on High Dynamic Range Tone Mapping with Subjective Tests

A high dynamic range (HDR) image has a very wide range of luminance levels that traditional low dynamic range (LDR) displays cannot visualize. For this reason, HDR images are usually transformed to 8-bit representations, so that the alpha channel for each pixel is used as an exponent value, sometimes referred to as exponential notation [43]. Tone mapping operators (TMOs) are used to transform high dynamic range to low dynamic range domain by compressing pixels so that traditional LDR display can visualize them. The purpose of this thesis is to identify and analyse differences and similarities between the wide range of tone mapping operators that are available in the literature. Each TMO has been analyzed using subjective studies considering different conditions, which include environment, luminance, and colour. Also, several inverse tone mapping operators, HDR mappings with exposure fusion, histogram adjustment, and retinex have been analysed in this study. 19 different TMOs have been examined using a variety of HDR images. Mean opinion score (MOS) is calculated on those selected TMOs by asking the opinion of 25 independent people considering candidates' age, vision, and colour blindness.

CERCS: T120 Systems engineering, computer technology; T125 Automation, robotics, control engineering; T111 Imaging, image processing

Keywords: High dynamic range imaging, tone mapping, display image, mathematical modelling of HDR images.

Contents

List of Figures	8
List of Tables	10
Abbreviations, constants, generic terms	11
1 Introduction	12
1.1 High Dynamic Range Imaging	12
1.2 Tone Mapping Operators	13
1.3 Problem Overview	14
1.4 Goals	15
2 Literature review	16
2.1 Subjective Evaluation of Tone-Mapping Methods on 3D Images	16
2.2 Analysis of Tone Mapping Operators on High Dynamic Range Images	17
2.2.1 Tone Mapping Operators	18
2.3 High Dynamic Range Image Tone Mapping by Maximizing a Structural Fidelity Measure	22

3	Motivation	26
4	TMO Analysis	27
4.1	Global Tone Mapping Operators for Images	27
4.1.1	Adaptive Tone-Mapping Operator for HDR Images Based on Image Statistics	27
4.1.2	Time-Dependent Visual Adaptation for Fast Realistic Image Display	28
4.1.3	Tone-Mapping High Dynamic Range Images by Novel His- togram Adjustment	29
4.1.4	An Efficient Unified-Tone-Mapping Operation for HDR Images with Various Formats	31
4.1.5	Modeling Radiometric Uncertainty for Vision with Tone- Mapped Color Images	32
4.1.6	A Tone Mapping Operator Based on Neural and Psychophysi- cal Models of Visual Perception	33
4.1.7	An Analysis of Visual Adaptation and Contrast Perception for Tone Mapping	35
4.2	Local Tone Mapping Operators for Images	36
4.2.1	GPU-Accelerated Local Tone-Mapping for High Dynamic Range Images	36
4.2.2	Natural HDR Image Tone Mapping Based on Retinex	38
4.2.3	Tone Mapping for HDR Image Using Chromatic Adaptation Method	40

4.2.4	Phase Preserving Tone Mapping of Non-Photographic High Dynamic Range Images	42
4.2.5	New Local Tone Mapping and Two-Layer Coding for HDR Images	44
4.2.6	Local Tone Mapping Using the K-means Algorithm and Automatic Gamma Setting	46
4.2.7	Saliency Modulated High Dynamic Range Image Tone Mapping	48
4.2.8	Region-of-Interest Based Local Tone Mappings	50
4.2.9	A Local Tone Mapping Operator for High Dynamic Range Images	53
4.3	Tone Mapping Algorithms for Videos	55
4.3.1	HDR Video Coding based on a Temporally Constrained Tone Mapping Operator	55
4.3.2	An Optimal Video-Surveillance Approach for HDR Videos Tone Mapping	56
4.4	Algorithms using Exposure Fusion	60
4.4.1	Direct High Dynamic Range Imaging Method Using Experiential Optimal Exposure Criterion	60
4.4.2	Alternating Line High Dynamic Range Imaging	61
4.4.3	Probabilistic Exposure Fusion	62
4.4.4	Evaluation of HDR Tone Mapped Image and Single Exposure Image	68
4.5	Inverse Tone mapping	69
4.5.1	Content-adaptive inverse tone mapping	69

5	Experimental Results	72
5.1	Experimental Setup	72
5.2	Results	73
6	Conclusion and Future Work	77
6.1	Conclusion	77
6.2	Future Work	77
	Acknowledgements	78
	References	79
	Non-exclusive licence to reproduce thesis and make thesis public	87

List of Figures

2.1	Average evaluation scores for tonemapping algorithms: 1 - <i>Log</i> , 2 - <i>Bi</i> , 3 - <i>Grad</i> , 4 - <i>Comp</i> , 5 - <i>Disp</i> , 6 - <i>Photo</i> , 7 - <i>Phy</i> , 8 - normal LDR image.	17
2.2	Steps involved in Tone mapping analysis [74]	21
4.1	Overall structure of Kim et al. method	38
4.2	I_{hdr} : HDR image, I_{ldr} : tone mapped image, $\hat{\cdot}$: decoded version of \cdot , GM: gamma mapping, IGM: inverse GM, rgb2y: calculating luminance function, LPF: lowpass filtering, TMO: tone mapping operator, ITMO: inverse TMO. [29]	44
4.3	examples of tone mapping [29]	44
4.4	Flowchart of proposed tone mapping (TM) in[39]	47
4.5	Block diagram of the proposed HDR video coding scheme. [55]	55
4.6	An intuitive block schema of the proposed algorithm[8]	57
4.7	Truncated pyramid interpolation filter[8]	59
4.8	Bi-linear interpolation in the four zones around the ROI [8]	60
4.9	Zhang and Dai method (Direct Fusion) vs regular HDR process	60
4.10	Flowchart of proposed method in[17]	62

4.11	The flow chart of the system [29]	69
4.12	The over-exposure enhancement flow [29]	71
5.1	Average results of subjects' votes for 1 st image	75
5.2	Average results of subjects' votes for 2 nd image	75
5.3	Average results of subjects' votes for 3 rd image	76
5.4	Average results of subjects' votes for 4 th image	76

List of Tables

4.1	Comparative results of five images.	68
5.1	Images used in our experiment	73
5.2	Average results of subjects' votes for images	74

Abbreviations

HDRI - High Dynamic Range imaging

TMO - Tone mapping operator

LDR - Low dynamic range

MOS - Mean opinion score

JND - Just-noticeable difference

ROI - Region of Interest

HVS - Human Visual System

HEVC - High Efficiency Video Coding

WCG - Wide Color Gamut

GOP - Group Of Pictures

SET - Small Exposure Time

LET - Large Exposure Time

IPC - Interlaced to Progressive Conversion

MLGC - Multi-Level Gain Compensation

1 Introduction

In this thesis, we will provide a comprehensive literature on tone mapping operators (TMOs) that are applied on static images and video sequences. Each method will be described in detail, and their experimental results will be provided. Also, a comparative analysis between different TMOs is provided so that one can decide which TMO is more suitable for their applications. It is important to note that there have been works in which tried to address TMO techniques for HDR videos. In such research works, issues such as flickering have been center of the study [44, 33].

1.1 High Dynamic Range Imaging

High dynamic range imaging (HDRI) is a process that contain multiple stages. In the first phase photographers take multiple pictures of the same scene with different exposure, which is in photography sometimes called stop. Then those images are merged to one another and become an image that luminance level is so high that traditional low dynamic range (LDR) devices can not display them. The luminance of a display is measured with nits which is the amount of the light output within a given area (cd/m^2). An average monitor may have the capability to output 200 to 300 nits, meanwhile HDR-compatible monitors may output 450 to 1,500 nits.

Especially, HDR is becoming more interesting due to its ability to provide more appealing and immersive images and video sequences [5]. HDR imaging has introduced a very attractive way of acquiring images from the scenes [34, 40, 71]. As it preserves a

wide range of luminance values for each pixel [34], it is a challenging process to display the HDR image using the traditional LDR devices that have a limited dynamic range [48].

1.2 Tone Mapping Operators

To solve the issue of transforming HDR images to the traditional LDR display devices, TMOs have been developed and used by researchers and industries. In other words, tone mapping is the process of compressing HDR images into a LDR so the traditional displays [67] can display them. Such traditional displays can visualize only the images and videos that have between 0 and 255 discrete intensity levels for each colour channel of RGB. This range is significantly less than the real world dynamic range, as the light intensity in the real world span a wide range from absolute darkness to very bright [3]. TMO can be presented as time dependent for handling animated sequences where temporal or bright adaptation of human visual system (HVS) are used, and time independent designed specifically for static images [76, 69, 54]. In TMO, two main requirements need to be fulfilled. These requirements are:

- keep details or local contrast;
- maintain the sensation of brightness.

In general, tone mapping tries to solve the problem of strong contrast reduction from scene values to displayable ranges in LDR displays while preserving the image details. This is a possible operation because HVS has limited capacity to detect differences in absolute luminance level. The level of surround luminance which decides upon the HVS adoption state and sufficient contrast on the display screen, is important in choosing of an appropriate TMO algorithm. As a result, the visual quality of display scene can be seriously degraded when the tone mapped images and/or video sequences are transmitted without any prior information about the end-point display device.

A typical tone mapping operator must provide consistent results despite the vast diversity of natural scenes and the possible radiance value inaccuracy found in HDR photographs. Additionally, it should be adaptable to the current capabilities of display devices. Tone mapping must capture the physical appearance of the scene, while avoiding the introduction of artifacts such as contrast reversal or black halos. The overall brightness of the output image must be faithful to the context. It must be fast for interactive and real-time applications while avoiding any trade-off between speed and quality.

Also one can classify TMOs as local and global operators:

- **Global TMO** is scaling each pixel based on its intensity and global image characteristics. It can be very fast, although may also cause loss of important details.
- **Local TMO** is scaling each pixels by its neighbourhood. It might seem good as a result, although it can also create artifacts and can take a lot of calculation time.

1.3 Problem Overview

Photorealistic computer graphics algorithms and high dynamic-range photography [17] are capable of reproducing full range of luminances encountered in the real world. Most modern display devices, on the other hand, are severely limited in their ability to generate light levels necessary to reproduce high dynamic range, very bright, or very dark environments. Therefore, one important problem in computer graphics is that of tone mapping, which deals with optimal ways of displaying such images. The necessity of tone mapping has been implicitly recognized from the inception of realistic computer graphics and, significantly earlier, in the field of photography [52].

1.4 Goals

For the last two decades, there has been a significant attention for tone mapping in computer graphics and photography, several new algorithms have been proposed to solve this problem. This thesis analyzes these TMOs and finds out which one is more accurate, real scene producing and luminance level dependent. To understand this field better and progress further, those tone mapping operators were compared extensively.

2 Literature review

2.1 Subjective Evaluation of Tone-Mapping Methods on 3D Images

Mai et al. conducted a comparison of various tone mapping algorithms based on the subjective evaluations of test subjects [45]. The study was concentrated based on 3D HDR tone mapping problem, and seven algorithms were compared as follows:

Global tone mapping algorithms:

- Adaptive logarithmic tone-mapping (*Log*) [18],
- Tone-mapping for backward-compatible HRD compression (*Comp*) [44],
- Display adaptive tone-mapping (*Disp*) [47],
- Tone-mapping by photoreceptor physiology (*Phy*) [62]

Local tone mapping algorithms:

- Photographic tone-mapping (*Photo*) [64],
- Fast bilateral filtering (*Bi*) [20],
- Gradient domain tone-mapping (*Grad*) [23].

In the experiment, sixteen subjects were participated and eight scenes (four indoors, four outdoors) were evaluated. Each participant scored the result points ranging from 0 to 10 in two categories, overall quality and 3D effect. Average scores for the algorithms are listed in Fig. 2.1.

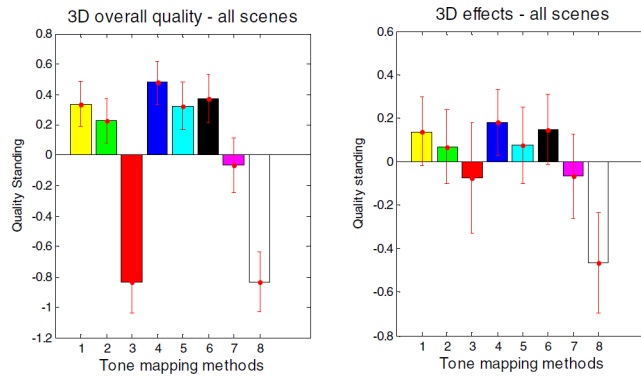


Figure 2.1: Average evaluation scores for tonemapping algorithms: 1 - *Log*, 2 - *Bi*, 3 - *Grad*, 4 - *Comp*, 5 - *Disp*, 6 - *Photo*, 7 - *Phy*, 8 - normal LDR image.

Subjectively the worst one was the Gradient domain tone mapping algorithm, and the best one was Backward compatible HDR compression algorithm. Also, it was observed that global algorithms have better performance regarding 3D effects than local. Additionally, a strong correlation between average brightness and 3D effect score was observed with brighter images having a better 3D effect score.

2.2 Analysis of Tone Mapping Operators on High Dynamic Range Images

In addition, their work in [45] evaluated the performance of various tone mapping operators such as Drago, Reinhard, Ward, Ashikmin, Piecewise, Retinex which are used for converting HDR images to LDR images. Additionally color correction and Hierarchical tone mappings are applied to enhance the image quality. [74]

2.2.1 Tone Mapping Operators

Ward Tone mapping operator is a global operator proposed by [36] in which single mapping function is derived and used for all pixels in a given image. The display luminance is calculated from the world luminance as follows.

$$L_d = \frac{1}{L_{white}} \left(\frac{1.219 + \left(\frac{L_{white}}{2}\right)^{0.4}}{1.219 + \log average^{0.4}} \right)^{2.5}, \quad (2.1)$$

where L_d = Display Luminance,

L_{white} = White pixel Luminance,

$\log average$ = Average of Luminance.

Drago operator is also known as Adaptive Logarithmic Tone mapping which is proposed by [21]. The display luminance is calculated as follows

$$L_d = \frac{\log(L_w + 1)}{\log(L_{max} + 1)}, \quad (2.2)$$

where L_d = Display Luminance,

L_{max} = Maximum Luminance,

L_w = Input Luminance of the pixel.

Reinhard operator in [63] is a spatially varying tone mapping operator also known as Photographic Tone mapping Operator. The log-average luminance is a useful approximation of the scene. L_w is computed by:

$$L_w = \frac{1}{N} \exp \left(\sum_{x,y} \log(\delta + L_w(x, y)) \right), \quad (2.3)$$

where $L_w(x, y)$ is the world luminance of the image and δ is a small value to avoid singularity for black pixels. N is the number of pixels in the image. Based on this

formula, the display luminance $L_d(x, y)$ is calculated as follows.

$$L_d = \frac{L(x, y) \left(1 + \frac{L(x, y)}{L_{white}^2}\right)}{1 + L(x, y)}, \quad (2.4)$$

where L_{white} = maximum luminance

$L(x, y)$ = local average of the neighborhood pixels

The local average of the neighborhood pixels is obtained by applying convolution operation of Gaussian Filter on the input image. The mask filter is calculated as follows.

$$M = \frac{1}{\sigma\sqrt{2\pi}} \frac{\exp(-x^2)}{2\sigma^2} \quad (2.5)$$

Ashikmin Operator is [1] a spatially varying tone mapping operator that uses surround center weighted average method. The local adaptation luminance $L_a(x, y)$ is calculated by contrast limited neighborhood growing procedure. The Gaussian filter is used to find the weighted average of the neighborhood pixels for a pixel. The display luminance $L_d(x, y)$ for a pixel is calculated as:

$$L_d(x, y) = \frac{L(x, y)TM(L_a(x, y))}{L_a(x, y)} \quad (2.6)$$

Piecewise tone mapping proposed by [28] is tone reproduction operator with chromatic adaptation. The tone mapping is done as follows.

Equation for global luminance adaption.

$$f = \frac{(Y_{av} - \log_2 Y_{min}) + (Y_{av} - \log_2 Y_{max})}{\log_2 Y_{max} - \log_2 Y_{min}} a = 2^f \quad (2.7)$$

$$Y'(x, y) = a.Y(x, y) \quad (2.8)$$

Chromatic adaption is performed on the input luminance.

$$\begin{aligned} R_c &= \left[Y_w \left(\frac{D}{R_w} \right) + (1 - D) \right] R \\ B_c &= \left[Y_w \left(\frac{D}{G_w} \right) + (1 - D) \right] B \end{aligned} \quad (2.9)$$

Where D is degree of adaption, R, G, B are the Red, Green and Blue input pixels and Y_w is the white luminance. Local luminance adaption is performed using histogram method.

$$\Delta b = \frac{\log_2(Y_{max}) - \log_2(Y_{min})}{NB} \quad (2.10)$$

Finally the chromatic adapted pixels are combined with luminance to form R, G, and B values.

Retinex adaptive filter operator in [51] is a local tone mapping operator in which global luminance is calculated first and then local luminance processing is done to enhance the image quality. The formula for calculating global display luminance:

$$L_d = (Lum_{input})^v \quad v = \min \left(1, \left(\frac{1}{6} Lum_{avg} \right) \right) \quad (2.11)$$

To evaluate these methods the tone mapping is applied on the HDR image and resulting images are compared with the original image. A luminance preserving color correction formula [46] is applied on the tone mapped images to improve the image quality with color correction. In [60], a computationally simple algorithm is suggested for image enhancement using histogram equalization. This method is combined with tone mapping operators and the effect of the Hierarchical method tone mapping is evaluated.

The evaluation is done, using the interactive data language (IDL) software. The original radiance picture is in standard HDR image format (".hdr") which is supported by most HDR image editors. Each pixel is stored as 4 bytes, one byte mantissa for each r, g, b and a shared one byte exponent. Red, Green, Blue values can be combined with

exponent to form world coordinate Red, Green, Blue channel pixel values as follows.

$$\begin{aligned} R_w &= \frac{R_M+0.5}{256} 2^{E-128} \\ G_w &= \frac{G_M+0.5}{256} 2^{E-128} \\ B_w &= \frac{B_M+0.5}{256} 2^{E-128} \end{aligned} \quad (2.12)$$

where R_w , G_w , B_w represents Red, Green and Blue world coordinates and R_M , G_M , B_M represents the Mantissa and E is exponent for the coordinates.

The steps involved in the analysis process are detailed in the following diagram. A step

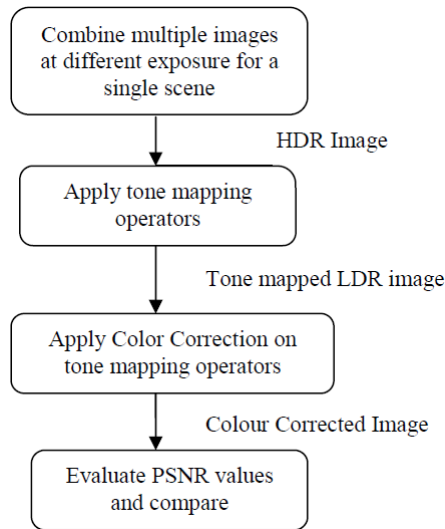


Figure 2.2: Steps involved in Tone mapping analysis [74]

involved in reading an HDR image and perform tone mapping is given below:

1. Read the image in *.hdr* format
2. Get the dimension details, exposure and other format details
3. Read the pixel data from *.raw* format file for the corresponding HDR image
4. Apply tone mapping on the pixel data read in Step 3
5. Display the tone mapped image
6. Compare the tone mapped image with the *.jpg* image for the corresponding HDR image.

According to color correction proposed by R. Mantiuk [46] can be achieved by applying

the following formula (2.13).

$$C_{out} = \left(\left(\frac{C_{in}}{L_{in}} - 1 \right) s + 1 \right) L_{out} \quad (2.13)$$

It was established that Color correction formula performs well on Retinex, Ashikmin and Piecewise tone mapping operators.

Hierarchical Tone mapping operator uses the following steps:

Step 1: Divide the intensity range into N ($N < 56$, $N = 4$ or 8 worked well) intervals.

Step 2: For pixels falling into each interval, use Traditional histogram equalization to re-map the pixels within the interval. Each interval is histogram equalized independently.

It was concluded that Hierarchical tone mapping shows good results on Ashikmin, Piecewise and Reinhard tone mapping operators.

Their work provide a method how to evaluate and use color correction and Hierarchical tone mapping to enhance the final image quality of different tone mapping operators such as Drago, Reinhard, Ward, Ashikmin, Piecewise and Retinex.

2.3 High Dynamic Range Image Tone Mapping by Maximizing a Structural Fidelity Measure

Yeganeh and Wang studied the issue of evaluating the quality of HDR to LDR mappings performed by various TMOs proposed in the literature [78]. They explained that there is no existing metric indicating the goodness of an HDR to LDR mapping that is needed to objectively compare existing methods and overall to formulate the problem in a more systematic and scientifically sound manner.

The authors cite a metric (proposed by themselves in a recent paper) and call it structural fidelity. The metric is somewhat inspired by the well-known structural similarity index (SSIM). The main insight is that both the original HDR image and the created LDR image should have similar contrast: it is allowed to have either low or high con-

trast in both images similarly, but cases where one image has low and another has high contrast, are penalised. This intuitive criterion is formalised using a function called Galton's give that is further discredited for the needs of processing the image patch by patch. That approach enables to create a goodness map for a given HDR to LDR tone mapping.

The described metric enables to formulate everything as an optimisation problem. The authors employ a simple gradient descent method and search for a structural fidelity local maxima in image space. Different starting criteria are used: a constant image, a linearly mapped image, and images provided by state-of-the-art tone-mapping methods. In all cases, they show that gradient descent can be used to improve the structural fidelity scores, and claim that the resulting images also have higher visual quality.

The paper seems promising. It is hard to say how good their metric is without implementing it and carrying out lots of tests, but the approach should be correct. Assuming the metric is in fact as good as they claim, a more sophisticated optimisation method can be used and very likely the results will improve a lot. This method is for image tone-mapping. If used on video in a naive way e.g. one image at a time, it will result in some mean shifting.

Adaptive LogLuv Transform [53] is also used widely in many TMO techniques. In [38], a TMO is proposed which tried to preserve as much information as possible from the HDR data, in order to enable the reconstruction of a high visual quality HDR image. For this purpose, they adopted adaptive LogLuv transform [53] and applied it to luminance values before conduct any coding.

Adaptive LogLuv transform considers the minimum and maximum luminance of the HDR image and maps it into the range of $[0, 2^n - 1]$ where n denotes the bit depth of the target LDR. The forward transformation can be considered of the following steps:

- the HDR image is transformed into the XYZ colour space using:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.497 & 0.339 & 0.164 \\ 0.256 & 0.678 & 0.066 \\ 0.023 & 0.113 & 0.864 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.14)$$

- the transition variables, x and y are defined as:

$$\begin{aligned} x &= \frac{X}{X + Y + Z} \\ y &= \frac{Y}{X + Y + Z} \end{aligned} \quad (2.15)$$

- the adaptive LogLuv representation is calculated by:

$$\begin{aligned} u_e &= \lfloor 410 \frac{4x}{-2x + 12y + 3} \rfloor \\ v_e &= \lfloor 410 \frac{9y}{-2x + 12y + 3} \rfloor \\ L &= \begin{cases} 0 & Y = 0 \\ \lfloor a (\log_2(Y) + b) \rfloor & \text{otherwise} \end{cases} \end{aligned} \quad (2.16)$$

where the parameters a and b can be obtain by Eqn. 2.17

$$\begin{aligned} a &= \frac{2^n - 1}{\log_2 \frac{\max(Y)}{\min(Y)}} \\ b &= \frac{1}{a} - \log_2(\min(Y)) \end{aligned} \quad (2.17)$$

The transferred metadata includes a and b and at the receiver part the following backward transformation is applied:

- firstly the value of luminance is calculated as:

$$L = \begin{cases} 0 & L = 0 \\ 2^{\frac{L+0.5}{a}-b} & \text{otherwise} \end{cases} \quad (2.18)$$

- the X and Z components are obtained by:

$$\begin{aligned} X &= \frac{Yx}{y} \\ Z &= \frac{Y(1-x-y)}{y} \end{aligned} \quad (2.19)$$

where x and y are calculated using:

$$\begin{aligned}
 u &= \frac{u_e + 0.5}{410} \\
 v &= \frac{v_e + 0.5}{410} \\
 x &= \frac{9u}{6u - 16v + 12} \\
 y &= \frac{4v}{6u - 16v + 12}
 \end{aligned} \tag{2.20}$$

- then the reconstructed HDR image is obtained by:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 2.690 & -1.276 & -0.414 \\ -1.022 & 1.978 & 0.044 \\ 0.061 & -0.224 & 1.163 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2.21}$$

Lauga et al. in their method updated the perceptually uniform (PU) values so that they are a closer approximation to the luminance sensitivity of the HVS [38]. The PU encoding was obtained from the contrast sensitivity function [41]. They observed that in most HDR image histograms, the energy of the signal is concentrated in a small range (most of the time in the low values). Nevertheless, a significant amount of information is contained in a much wider range of luminance values. Based on these observations the input HDR image was segmented into two regions, which were processed independently. In order to perform the binary segmentation, a threshold on the luminance is defined. In their work, they set this threshold so that 70% of the luminance values are below the threshold. Then a TMO was applied on the low-range region, rather than the whole HDR image, resulting in preserving more information. In parallel, another TMO was applied to the high-range regions. Hence, saturation was avoided in bright regions, which were conserving details. In order to stretch the histogram and fully exploit the whole range available, they applied a logarithm on the PU values [38].

3 Motivation

The needs of a realtime vision aid require that it functions immediately, not merely for production of a picture or video record to be viewed later. [27] Therefore the vision system must offer a high dynamic range (HDR) (often hundreds of millions to one) that functions in real time. In addition, the processing requires preservation of details captured in the extended dynamic range but not compromising the users in critical situations by having them constantly adjusting many system settings.

There are many real-time and efficient HDR compositing methods developed. However these composited HDR images have a wider range than the conventional media, such as monitors, projectors or in print. The HDR images therefore cannot be displayed without being compressed or using specially designed HDR media [70].

This compressing method is the tone mapping we are studying in this thesis. HDRI has been a huge evolution in the field of image processing. Although there is a great amount of tone mapping operators introduced, they all have their own limitations. To acquire more knowledge and contribute to this field we need to study and analyze all tone mapping algorithms out there.

4 TMO Analysis

4.1 Global Tone Mapping Operators for Images

In this section we analyze global TMOs.

4.1.1 Adaptive Tone-Mapping Operator for HDR Images Based on Image Statistics

Bae et al. proposed a tone mapping method [2] by first dividing input into regions, using K-means clustering, then using a adjusted global tone mapping operator on each region.

First, the number of regions, k , must be decided. The following algorithm is proposed

1. Using luminance values, L , construct a histogram consisting of N bins for $\log_2 L$, where

$$N = \lfloor \sigma \cdot 10 \rfloor \quad (4.1)$$

and σ is variance of $\log_2 L$.

2. Determine the k value

$$k = \left| \left\{ i \mid f(b_i) > \frac{\text{image size}}{8} \right\} \right|, \quad (4.2)$$

where $f(b_i)$ is the pixel count in the histogram bin i .

After the image has been separated into regions, tone mapping algorithm is applied to each region

$$D(L) = (D_{\max} - D_{\min}) \frac{\log(L+\tau) - \log(L_{\min} + \tau)}{\log(L_{\max} + \tau) - \log(L_{\min} + \tau)} + D_{\min}, \quad (4.3)$$

where L_{\min} and L_{\max} are minimum and maximum luminance, D_{\min} and D_{\max} are minimum and maximum display level of visualisation device and τ is a parameter that regulates overall image brightness. For calculation of τ the use of cluster centroids, C_i , is proposed

$$\tau_i = \lfloor 0.3 \cdot C_i \rfloor. \quad (4.4)$$

Compared to most local tone mapping methods, the proposed method has relatively low computational complexity.

4.1.2 Time-Dependent Visual Adaptation for Fast Realistic Image Display

In [56] Sumanta N. Pattanaik et al. includes a new tone mapping operator for appearance changes in animations or interactive real time simulations. The operator accepts a stream of scene intensity frames and creates stream of color display images. The given operator can accept any desired input like, static or dynamic, real or synthetic. It uses global tone mapping method and according to the author it's quite robust. It follows Tumblin and Rushmeier's method and follows the structure below:

Scene intensity > Adaptation model> Appearance model> Inverse Appearance model> Inverse Adaptation Model > Display intensity

Adaptation model: scene intensity to retinal like vectors R Appearance model: R to appearance vectors These two compute display intensity of scene and the inverse pairs compute display intensity.

This model is general accepting time sequences of arbitrary scene intensities. It captures widely varying amounts of adaptation and it is firmly grounded by psychophysics, physiology and color science.

In adaptation model the operator works globally and it is simplification of Hunt's static model of colors by adding time-dependent adaptation components.

The appearance model assumes that human can assign equivalent appearance to dim displays. Again this model follows Hunt's suggestion and determine white as five times the current adaptation level and reference black as $1/32$ the intensity of white reference.

The inverse adaptation model finds display RGB values for a given set of visual response values. The inverse appearance model attempts to map scene appearance values to display values.

4.1.3 Tone-Mapping High Dynamic Range Images by Novel Histogram Adjustment

In [19] a novel histogram adjustment methods for displaying high dynamic range image is presented. First a global histogram adjustment based tone mapping operator reproducing global contrast for high dynamic range images is given, then a novel bilateral weighting scheme to eliminate associated boundary artifacts and halo artifacts caused by the segmentation is introduced.

For global tone mapping method, they have considered the maximum and minimum luminance of the scene, minimum, the display level of the visualisation devices(usually 0-255), and the overall brightness of the mapped image(τ). Usually, a larger τ makes the mapped image darker and vice versa smaller τ causes brighter images. To automatically estimate τ following steps have been developer in the work:

1. Calculate the log- average luminance of the scene.
2. Calculate the key value - Larger key values correspond to brighter images- range

between 0-1

3. Decide the offset such that average luminance in the image is mapped to determined key value The operator is related to the histogram adjustment technique.

The histogram adjustment method has several drawbacks:

1. The method caps only display contrast mapped by histogram equalisation. It means that features visible to human eyes can be lost.
2. If the dynamic range of the scene is already in the device range, the method uses linear scaling.

The method given in the paper overcomes these issues, although it has its own drawbacks like lack of local contrast and details.

After global logarithmic mapping, the next step is to determine which local regions this technique should be applied. To segment an image into local areas, in the paper non-overlapping regular rectangular blocks are used. In order to solve the boundary artifacts, for each pixel value in the image, the final mapped pixel value is the weighted average of the results from tone mapping functions according to a distance weighting function. In order to solve the halo artifact problem, a similarity weighting function is added, which increases the chance that similar pixels in the uniform areas are mapped to similar values even when they belong to different blocks.

operators. To map a $1024 * 768$ pixel image on a T7300 with 2.0 GHz CPU takes about 0.4 s using given ALHA operator, which is about 5 times faster than Fattal et al.'s operator and 6 times faster than Reinhard et al.'s operator, and slightly slower than Durand & Dorsey's bilateral filter technique. Another property is that the derivation of the tone mapping functions in different blocks are independent of each other and the weighting process for each pixel takes only $5*5$ neighboring elements which make it suitable for parallel computing.

To sum up, in the paper 2 algorithms were given: global HALEQ and local ALHA. Global HALEQ causes different artifacts for local contrast, where ALHA is able to

produce high quality with very fast speed.

4.1.4 An Efficient Unified-Tone-Mapping Operation for HDR Images with Various Formats

The article [26] introduces unified TMO for various HDR formats with fixed-point arithmetic. Conventional TMOs are generally expressed in a floating data format so they require a lot of computational and memory cost. The given method replace a floating-point number with 8-bit integer numbers to reduce the resources. However it has its own limitations such as available HDR image formats. The given unified method with various formats such as RGBE and OpenEXR gives almost the same quality compared to other TMOs but it reduces the complexity.

Unlike the conventional TMOs the given method is focused on complexity of memory and computation rather than compression techniques and quality of tone mapped images.

The unified method is first converting various HDR formats like RGBE and OpenEXR to intermediate format and then apply a TMO for it.

Unlike RGBE format, the exponent part of each RGB channel in the intermediate format is independent, which reduces the error of the format conversion.

The integer TMO is implemented with integer input and integer output unlike the conventional methods. This TMO works well with the given intermediate method. The numerical is significantly reduced because exponent part and the mantissa part are separated as two integer numbers. Although it's not working properly with IEEE754 format because it has denormalized numbers as well as OpenEXR.

However the internal arithmetic of the integer TMO is still with floating-point. The proposed method is introducing fixed-point arithmetic to reduce the computational complexity as well.

Errors may occur by these format conversion and fixed-point arithmetic. Experiments and evaluation were carried out to measure peak signal-to-noise ratio(PSNR) of result-LDR images, processing time of the TMO and memory usage. Experiments indicate that the PSNR values get better with increase in the bit length and they can be a little less than 60dB at 8 bit. Afterward the PSNR of the LDR images have been calculated and from the results it was confirmed that the proposed method can execute the TMO with high accuracy even it involves the format conversion and the fixed-point arithmetic. The method cancelled 8208 bits out if the size of the input HDR image is larger than 42 pixels.

Overall, the proposed method is applied with global tone mapping operations.

4.1.5 Modeling Radiometric Uncertainty for Vision with Tone-Mapped Color Images

In the paper [11] Ayan Chakrabarti et al. considers the inherent uncertainty of undoing the effects of tone-mapping. It's observed that this uncertainty varies substantially across color space, making some pixels more reliable than others. A new model for this uncertainty and a method for fitting it to a given camera or imaging pipeline are introduced. Once fit, the model provides for each pixel in a tone-mapped digital photograph a probability distribution over linear scene colors that could have induced it.

In their paper, they argue that vision systems can benefit substantially by incorporating a model of radiometric uncertainty when analyzing tone-mapped, JPEG-color images. [11] A probabilistic approach is introduced for visual inference.

Image Rendering Model

The model is given as below:

$$\tilde{y} = \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \tilde{y}_3 \end{bmatrix} = \begin{bmatrix} f(v_1^T x) \\ f(v_2^T x) \\ f(v_3^T x) \end{bmatrix} \quad (4.5)$$

$$y = Q\left(B(\tilde{y}) + \begin{bmatrix} g_1(\tilde{y}) \\ g_2(\tilde{y}) \\ g_3(\tilde{y}) \end{bmatrix}\right) \quad (4.6)$$

where, v_1, v_2, v_3 define a linear color space transform, $B(\cdot)$ bounds its argument to 8-bit integers, $f(\cdot)$ is assumed to polynomial of model parameters.

Finally, to estimate parameters of standard map, the proposed algorithm can be given as following formula :

$$\{\hat{v}_c\}, \{\hat{\alpha}_i\} = \underset{\{v_c\}, \{\alpha_i\}}{\operatorname{argmin}} \sum_t w_t \sum_c \|f(v_c^T x_t) - y_{t:c}\|^2 \quad (4.7)$$

With high intensities and saturated colors greater compression is experienced and therefore more uncertainty is being faced for JPEG values. They define the inverse probabilistically as a distribution instead of deterministic inverse function.

The framework is based on global tone-mapping processes. This might be considered as a form of signal dependent Gaussian noise. They introduce image fusion, photometric stereo and deblurring.

4.1.6 A Tone Mapping Operator Based on Neural and Psychophysical Models of Visual Perception

In [16] a two stage tone mapping approach is proposed, in which the first stage is a global method for range compression based on a gamma curve equalizing the lightness,

and the second stage performs local contrast enhancement and color induction using neural activity models for the visual cortex.

The displayed image I is presented with a given system gamma γ_{sys} , where I_0 is the original linear radiance map normalized to between 0 and 1

$$I = I_0^{\gamma_{sys}} \quad (4.8)$$

The perceived lightness of the image L^* is modeled as a gamma function of value γ_{psy} on screen luminance, this value is fixed to 0.4 in this study:

$$L^* = I^{\gamma_{psy}} \quad (4.9)$$

For each value of γ_{sys} , the lightness distribution is computed as :

$$F = 1 - \frac{1}{N} \sqrt{\sum_{i=1}^N (H(L^*)_i - i)^2} \quad (4.10)$$

Where, F is the flatness, $H(L^*)$ is the cumulative histogram of image L^* and N is the number of bins. Finally they search for the value of γ_{sys} that optimizes F and it is the 'optimal' system gamma.

As second stage, local operation, one may use the following model to minimize Bertalmio's energy function as follows :

$$I_t(x) = -\alpha(I(x) - \text{frac12}) + \gamma \sum_y w(x,y) \text{sgn}(I(x) - I(y)) dy - \beta(I(x) - I_0(x)) \quad (4.11)$$

The proposed modification is as follows :

$$I_t(x) = -\alpha(I(x) - m(x)) + \gamma(1 + (\sigma(x))^c) \sum_y w(x, y) \text{sgn}(I(x) - I(y)) dy - \beta(I(x) - I_0(x)) \quad (4.12)$$

where the mid-value of the first term is no longer global but the local mean of the original image computed with a Gaussian kernel: $m(x) = (G * I_0)(x)$ based on local standard deviation σ , while c is a constant.

4.1.7 An Analysis of Visual Adaptation and Contrast Perception for Tone Mapping

Another tone mapping operator with two stages has been proposed in the paper [24]. The first stage is a global method that implements visual adaptation, based on experiments on human perception, in particular the importance of cone saturation. The second stage performs local contrast enhancement, based on a variational model inspired by color vision phenomenology.

In the paper Naka-Rushton (NR) [56] and Weber-Fechner (WF) [37] algorithms have been deeply considered and as the first stage, a new TM curve $c(I)$ has been given combining these two, with the following formula:

$$c(I) = \begin{cases} \hat{k} \log(I + m) + s_0 & I \leq I_M \\ \frac{I^n}{I^n + I_s^n} & I > I_M \end{cases} \quad (4.13)$$

Where, s_0 is chosen so as to ensure the continuity of c at I_M , I_M is the semisaturation radiance. The output of c will be normalized so the interval will be $[0, 1]$. This stage is global TMO that uses nondecreasing mapping curve c .

The second stage of the given TMO is finding the image \tilde{I} that minimizes this expression:

Where, they have chosen specific values for the constants α, β, γ , the gaussian kernel w gives the locality to the contrast measure and μ is the average value of the original image I_0 . The image I_0 is the normalized output of the first stage of the given TMO. This function is a modification of the one proposed by Bertalmo.

4.2 Local Tone Mapping Operators for Images

In this section we will analyze local tone mapping operators.

4.2.1 GPU-Accelerated Local Tone-Mapping for High Dynamic Range Images

In [75] Duan Tian and Qui. present a very fast local tone mapping method for displaying high dynamic range (HDR) images. High computational efficiency is obtained by a highly parallel algorithm, which can be implemented on a Graphics Processing Unit (GPU). To produce good results, the proposed method mimics the local adaption mechanism of the human visual system.

In the proposed algorithm the image is divided into non-overlapping rectangle blocks and reproduces contrast and brightness in each of them simultaneously using a highly parallel global tone mapping operator. Individual pixels are rendered in parallel. HALEQ global tone mapping operator is used within the individual blocks. HALEQ works by striking balance between linear compression and histogram equalized compression in the mapping process as Eq. (4.14).

$$d(x, y) = \beta \cdot EC[D(x, y)] + (1 - \beta) \cdot LC[D(x, y)] \quad (4.14)$$

$D(x, y)$ is the input luminance while $d(x, y)$ is the output display intensity level. EC and LC is the histogram equalization and linear mapping function. β controls the contrast enhancement level.

To get rid of boundry artifacts. In the proposed method final mapped pixel value is the weighted average of the results from mapping functions $HALEQ(n)$ of neighboring blocks as Eq.(4.15).

$$d(x, y) = \frac{\sum_{n=1}^{n=N} HALEQ_n[D(x, y)] \cdot w_d(n)w_d(n)}{\sum_{n=1}^{n=N} w_d(n)w_d(n)} \quad (4.15)$$

$$w_d(n) = e^{-(d_n/\sigma_n)}, w_s(n) = e^{-(s_n/\sigma_n)} \quad (4.16)$$

N is the number of used blocks. w_d and w_s is the distance weighting function and pixel value similarity weighting fuction. d_n is the Euclidean distance between the current pixel position and the center of used blocks. s_n is the normalized difference between the current pixel value and the average pixel value of block n . σ_d and σ_s control the smoothness between blocks. To reduce noises in uniform areas a local contrast enhancement adhumstment is performed. To deal with this problem β value is adaptively decreased for uniform ares as Eq. (4.17).

$$\beta = 0.6 \cdot [1 - e^{-(SD_{nmax} - SD_n)}] \quad (SD_n > \eta) \quad (4.17)$$

SD_n is the deviation population for each region. The threshold η is set manually for different images.

The presented tone mapping method uses CUDA (a scalable parallel programming model for GPU) to achieve fast and effective HDR image display and has the potential to render real time HDR videos.

4.2.2 Natural HDR Image Tone Mapping Based on Retinex

Kim et al. propose an improvement to the well known Retinex algorithm by dynamically choosing the k factor and compressing bright areas [31]. The goal is to fix the two main limitation of the Retinex algorithm [35], namely large variation of performance by image and emphasizing only highlight region details.

General Structure

Figure 4.1 shows the overall structure of the method.

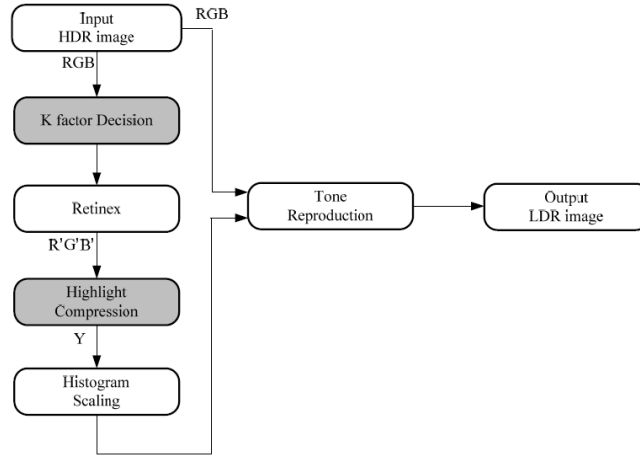


Figure 4.1: Overall structure of Kim et al. method

k Factor Calculation

Based on experimental results, it was observed that the best value for k and image dynamic range have linear relationship, the higher the range the higher the k value should be chosen. Simple formula was derived for calculating the k factor

$$k = \begin{cases} 0.2 & \text{for } D \leq 0.25 \\ 0.1 \times (D - 2.5) + 0.2 & \text{for } 0.25 \leq D \leq 0.55 \\ 0.5 & \text{for } D > 0.55 \end{cases} \quad (4.18)$$

where dynamic range D is calculated

$$D = \log I_{\max} - \log I_{\min} \quad (4.19)$$

and I_{\max} , I_{\min} are minimum and maximum values in the image.

Highlight Compression

Highlight compression is done after applying the Retinex algorithm. Compression is applied on the luminance, L , calculated from the new RGB values using the LHS system

$$L = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B. \quad (4.20)$$

The goal of the compression is to preserve dark area detail and compress bright area detail. It is achieved by following formula

$$L_{out} = \frac{2 \cdot I_{max}}{1 + \exp \frac{-a \cdot I}{I_{max}}} - I_{max}, \quad (4.21)$$

where a controls the curve degree of compression. Experimentally, the a is chosen 4.

After the compression, about 1% (determined by experiments) of the brightest pixels are removed and the rest scaled, in order to give image more natural look.

Tone Reproduction

Finally, the output LDR image is composed by

$$\begin{pmatrix} R_{out} \\ G_{out} \\ B_{out} \end{pmatrix} = \begin{pmatrix} L' \cdot \left(\frac{R_{in}}{L_{in}} \right)^s \\ L' \cdot \left(\frac{G_{in}}{L_{in}} \right)^s \\ L' \cdot \left(\frac{B_{in}}{L_{in}} \right)^s \end{pmatrix} \quad (4.22)$$

where R_{in} , G_{in} and B_{in} represent input HDR data, L' the result of histogram scaling and s the gamma correction coefficient, which is chosen as 0.75.

4.2.3 Tone Mapping for HDR Image Using Chromatic Adaptation Method

In [14] they present a color correction method using chromatic adaptation method.

The mapping in the proposed method uses local adaptation level ($L_a(x, y)$) to apply a locally linear mapping over the input image with the applied coefficients depending on the particular image neighbourhood. Local adaptation level is defined as the average luminance over some local image area, and it can apply a tone-mapping function ($TM(L)$) to $L_a(x, y)$, which would create a tone-mapping adaptation image $TM(L_a(x, y))$. The formula for final mapping

$$L_d = \frac{L(x, y)TM(L_a(x, y))}{L_a(x, y)}. \quad (4.23)$$

where $L(x, y)$ and $L_d(x, y)$ represents world luminance and the displayable luminance. $L_a(x, y)$ is obtained with Laplacian pyramid method [9] based on XYZ tristimulus value.

Chromatic adaptation transformation The proposed method is chromatic adaptation transformation to deal with perceptual mismatches between real world and displayed images. From the input image based on high contrast tone mapping the computation of the transformation is described as follows

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M_{CAT02} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad M_{CAT02} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix} \quad (4.24)$$

$$D = 0.3F \left[1 - \left(\frac{1}{3.6} \right) e^{\frac{-(L_A - 42)}{92}} \right] \quad (4.25)$$

$$R_c = \left[\left(Y_w \frac{D}{R_w} \right) + (1 - D) \right] R \quad (4.26)$$

$$G_c = \left[\left(Y_w \frac{D}{G_w} \right) + (1 - D) \right] G \quad (4.27)$$

$$B_c = \left[\left(Y_w \frac{D}{B_w} \right) + (1 - D) \right] B \quad (4.28)$$

where the R_c , G_c and B_c values are the chromatically adapted cone response by applying D cone response and R_w , G_w and B_w values represent corresponding R , G and B values for the white point. X_w , Y_w and Z_w are the relative tristimulus values of with input XYZ. From eq 7. R' , G' and B' values are the chromatic adaptation values using inverse M_{CAT02}^{-1} transformation and then converted to Hunt-Point Estevez space as follows:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = M_{HPE} M_{CAT02}^{-1} \begin{bmatrix} R_c \\ G_c \\ B_c \end{bmatrix} \quad (4.29)$$

$$M_{HPE} = \begin{bmatrix} 0.38971 & 0.68898 & -0.07868 \\ -0.22981 & 1.18340 & 0.04641 \\ 0.0 & 0.0 & 1.0 \end{bmatrix} \quad (4.30)$$

$$M_{CAT02}^{-1} = \begin{bmatrix} 1.096124 & -0.278869 & 0.182745 \\ 0.454369 & 0.473533 & 0.072098 \\ -0.009628 & -0.005698 & 1.015326 \end{bmatrix} \quad (4.31)$$

Where M_{HPE} is a single transformation matrix from sharpened cone response to the Hunt-Point-Estevéz cone responses.

The proposed method consists of the tone-mapping and chromatic adaptation transformation. As results, the local contrast and image detail are well improved without glare

limit and colour distribution.

4.2.4 Phase Preserving Tone Mapping of Non-Photographic High Dynamic Range Images

The presented method in [32] compresses the dynamic range of an image while preserving local phase information and is intended for non-photographic images. The method is based on extracting local phase and amplitude values using monogenic filters and then reducing the dynamic range by applying a range reducing function to the amplitude values.

In the algorithm the local phase values of the image must be preserved in the tone mapped output. The image is decomposed into its local phase and amplitude values and these amplitude values are attenuated via some function and the image then reconstructed using the original phase values and the attenuated amplitude values.

Monogenic filters are formed by combining a radial band-pass or high-pass filter with its Riesz transform which is made up of two components. If we define two filters in the 2D frequency domain u_1, u_2

$$H_1 = i \frac{u_1}{\sqrt{u_1^2 + u_2^2}}, \quad H_2 = i \frac{u_2}{\sqrt{u_1^2 + u_2^2}} \quad (4.32)$$

then the spatial representation of the vector $H = (H_1, H_2)$ defines the convolution kernel of the Riesz transform. To obtain local phase and amplitude information the image, I , is convolved with the band-pass or high-pass filter f and the two Riesz transform filtered versions of f , $h_1 f$ and $h_2 f$. This provides three outputs, $I * f$, $I * h_1 f$ and $I * h_2 f$, where $*$ denotes convolution.

The output from convolution with the band-pass filter f corresponds to the vertical coordinate, and the convolutions with the Riesz transform filters $h_1 f$ and $h_2 f$ specify the two horizontal coordinates.

The local amplitude at image location (x, y) is given by

$$A(x, y) = \sqrt{f(x, y)^2 + h_1 f(x, y)^2 + h_2 f(x, y)^2}. \quad (4.33)$$

The local phase is given by

$$\phi(x, y) = \text{atan2}(f(x, y), \sqrt{h_1 f(x, y)^2 + h_2 f(x, y)^2}) \quad (4.34)$$

and orientation given by

$$\theta(x, y) = \text{atan2}(h_2 f(x, y), h_1 f(x, y)) \quad (4.35)$$

Dynamic range reduction is achieved by applying a range reducing function to the amplitude and then reconstructing using the original phase, by using the logarithm of the amplitude, $\log(A + 1)$ or a nested logarithm ($\log(\log(A + 1) + 1)$) of the amplitude.

The reconstructed, tone mapped image values, $T(x, y)$ are given by

$$T(x, y) = \log(A(x, y) + 1) \cdot \sin(\phi(x, y)). \quad (4.36)$$

The algorithm ensures that the details are maintained by preserving the local phase values within the image data. The main parameter in the algorithm is the cutoff frequency of the high-pass filter used. This parameter controls the scale of the features that are highlighted by the algorithm.

For better results it is useful to generate a sequence of images with the high-pass filter cutoff frequency increasing geometrically. These images can be inspected separately or blended together.

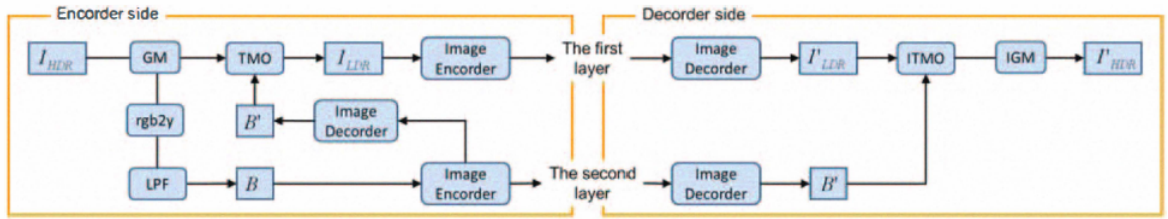


Figure 4.2: I_{hdr} : HDR image, I_{hdr} : tone mapped image, $'$: decoded version of $.$, GM: gamma mapping, IGM: inverse GM, rgb2y: calculating luminance function, LPF: lowpass filtering, TMO: tone mapping operator, ITMO: inverse TMO. [29]

4.2.5 New Local Tone Mapping and Two-Layer Coding for HDR Images

A two-layer High Dynamic Range (HDR) coding scheme using a new tone mapping is proposed where the HDR image is transformed onto a Low Dynamic Range (LDR) image by using a base map that is a smoothed version of the HDR luminance. [29]

The basic architecture of the encoder/decoder are illustrated in Fig.4.2.

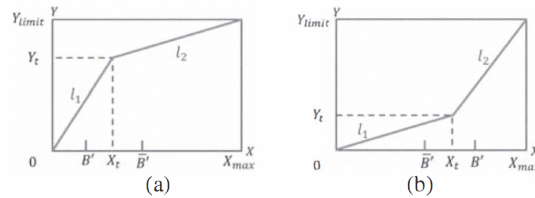


Figure 4.3: examples of tone mapping [29]

The tone mapping is based on the simple two-line curves as illustrated in Fig.4.3. The curve varies pixel by pixel, depending on a base map. The base map B in Fig. 4.2 is calculated by applying a lowpass filter to the HDR luminance. The method encodes a tone mapped image I_{hdr} and the base map B , as the first and second layer. To decode the HDR image, the second layer B' is decode and inversely tone maps the first layer I_{hdr} with it.

Coding scheme

The first step of the encoder is to apply the gamma mapping (GM in Fig.4.2) to the HDR images to mimic the HVS's nonlinearity. In the next step, a lowpass filtering is

applied to the nonlinear HDR luminance that yields the base map B. The method uses this base map for the tone mapping, and encodes it as the second layer. The final quality of the HDR image depends only on the fidelity of the LDR image I_{hdr} .

Finally we encode the tone mapped LDR image by a conventional image encoder and transmit the LDR and the base map to the decoder as the first and second layer. The decoder can obtain the LDR image directly without any tone mapping operations. The inverse tone mapping method is performed by using the LDR and the base map. Finally after the inverse tone mapping, the inverse gamma mapping is applied to obtain the I_{hdr} . The method maps the HDR pixel value to the LDR one by using a tone mapping function composed of two linear transforms as is depicted in Fig.4.3 . To acquire the base map, a linear Gaussian lowpass filtering is performed to the nonlinear HDR luminance.

Tone mapping

The local illumination is approximately given by the base map. The HDR luminance region is separated into two ranges, and maps the luminance in each range by the corresponding line where the mapping functions in low and high luminance ranges are shown as, l_1 and l_2 (see Fig.4.3).

The tone map consisted of those two lines are calculated by following.

$$\begin{cases} l_1 : Y(i) = \frac{Y_t(i)}{X_t(i)} X(i), & \text{if } X(i) \leq X_t(i) \\ l_2 : Y(i) = \frac{Y_{limit} - Y_t(i)}{X_{max} - X_t(i)} X_{l2}(i) + Y_t(i), & \text{else} \end{cases} \quad (4.37)$$

where $X_{l2}(i) = X(i) - X_t(i)$, i is the pixel index, X_{max} is the maximum pixel value of X , and (X_t, Y_t) is the coordinates of intersection between two lines. X_t, Y_t and Y_{limit} are calculated by

$$X_t(i) = \bar{B}' + \alpha_X(\bar{B}'(i) - \bar{B}'), \quad (4.38)$$

$$Y_t(i) = \{ \bar{B}' - \alpha_Y(B'(i) - \bar{B}') \} \frac{Y_{limit}(i)}{X_{max}}, \quad (4.39)$$

$$Y_{limit}(i) = \frac{X_{max}}{1 + B'(i)}, \quad (4.40)$$

$$\bar{B}' = exp \left\{ \frac{1}{N} \sum_{i=1}^N \log(B(i) + \epsilon) \right\}, \quad (4.41)$$

where \bar{B}' , α_X and α_Y are parameters which control the tone mappin effects, N is a number of pixels and ϵ is recomended to be a small value. X_t , Y_t and X_{limit} construct the tone map, and those parameters change according to the base map. The inverse tone mapping is performed by applying the inverse function of (4.37), the tone mapping is invertible in the presence of B' and X_{max} .

The method was compared with the standard two-layer HDR coding method proposed by Ward et al.[77]. The local contrasts for the proposed method can be seen more clearly than the conventional one both in low and high luminance range.

4.2.6 Local Tone Mapping Using the K-means Algorithm and Automatic Gamma Setting

Lee et al. propose a local tone mapping algorithm [39] that first segments the HDR image into regions using K-means algorithm and then adjusts display gamma parameter dynamically for each segmented region based on the mean value of luminance in the region.

First, the luminance of the input HDR image, L_{in} , is filtered by a bilateral filter

$$\tilde{L}(p) = \frac{1}{k(p)} \sum_{q \in \Omega} G_{\sigma_s}(p - q) G_{\sigma_r}(L_{in}(p) - L_{in}(q)) L_{in}(q), \quad (4.42)$$

where G represents Gaussian function, Ω represents pixel p neighbourhood and σ_s , σ_r represent Gaussian weight function standard deviations in the spatial and intensity

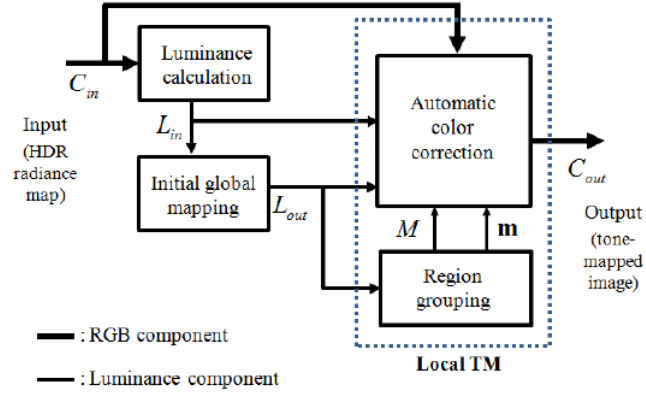


Figure 4.4: Flowchart of proposed tone mapping (TM) in[39]

domain. Normalisation term is defined by

$$k(p) = \sum_{q \in \Omega} G_{\sigma_s}(p-1)G_{\sigma_r}(L_{in}(p) - L_{in}(q)). \quad (4.43)$$

Next, simple global tone mapping is done by logarithmic transformation

$$L_{out}(p) = \log \tilde{L}(q). \quad (4.44)$$

This global mapping is used for the input of K-means algorithm [30]. Let m_k be a mean value of the k^{th} cluster, then mean vector of K clusters is

$$\vec{m} = [m_1, \dots, m_k, \dots, m_K]^T, \quad (4.45)$$

and segmented region is denoted by

$$M(p) = k, \quad 1 \leq k \leq K. \quad (4.46)$$

Finally, local tone mapping is performed based on the mean vector \vec{m} . Gamma value for k^{th} cluster is calculated by

$$\gamma_k = \min \left(\gamma_{\max}, \frac{a}{m_k} \right), \quad (4.47)$$

where a is a scale constant $0 \leq a \leq 1$ and $0 \leq \gamma_k \leq \gamma_{\max}$. Then the calculated gamma vector is defined as

$$\vec{\gamma} = [\gamma_1, \dots, \gamma_k, \dots, \gamma_K]^T. \quad (4.48)$$

The value for a was derived from experimental results. Best value was determined to be 0.3.

The output LDR luminance values are calculated based on the gamma as following

$$C_{out}(p) = \left(\frac{C_{in}(p)}{L_{in}(p)} \right)^\gamma L_{out}(p), \quad (4.49)$$

which is linearised into

$$C_{out}(p) = \left(\gamma \left(\frac{C_{in}(p)}{L_{in}(p)} - 1 \right) + 1 \right) L_{out}(p). \quad (4.50)$$

4.2.7 Saliency Modulated High Dynamic Range Image Tone Mapping

Mei et al. used saliency features of the HDR image to calculate image saliency map and adjust local tone mapping curves according to those saliency values in order to achieve a pleasing LDR image [49].

Saliency Map Calculation

The adapted algorithm (Algorithm 1) from [10] is used for calculating the saliency map from HDR image luminance.

Tone Mapping Based on Saliency

Tone mapping algorithm is based on [19]. It first finds two reference sets of cut points in the HDR value range. First, HDR range is divided into N segments of equal length

Algorithm 1 Saliency map algorithm

Initialisation

$L \leftarrow$ HDR luminance
 $S \leftarrow 5$ (number of scales)
 $\sigma_{\min} \leftarrow 1$ (smallest box filter radius)
for $i = 1$ to S **do**
 $\sigma[i + 1] \leftarrow 2\sigma[i]$
end for

Saliency map calculation

$SaliencyMap \leftarrow \vec{0}$
 $I_m \leftarrow DownSample(L)$
 $BF[1] \leftarrow$ Filter I_m with box filter, $width = 2\sigma[1] + 1$
for $i = 1$ to S **do**
 $BF[i] \leftarrow$ Filter I_m with box filter, $width = 2\sigma[i + 1] + 1$
 $DoB[i] \leftarrow BF[i] - BF[i + 1]$
 $SaliencyMap \leftarrow SaliencyMap + \text{abs}(DoB[i])$
end for**return** $SaliencyMap$

by cut points $\{l_i\}$ and also into N segments containing equal amount of pixels by cut points $\{e_i\}$, $N = 256$. From these, new cut points are constructed by

$$le_i = l_i + \beta(e_n - l_n), \quad (4.51)$$

where $0 \leq \beta \leq 1$ regulates contrast in the LDR image. All HDR pixels between le_i and le_{i+1} are mapped to same LDR value d_i . The values for β are calculated using Algorithm 2. The main idea is that higher saliency areas are applied more contrast than low saliency areas.

Algorithm 2 Tone mapping algorithm

$\beta_{\min} \leftarrow 0.2$
 $p \leftarrow 1.5$
 $n \leftarrow 32$
Partitioning the image into blocks of $n \times n$ pixels
for $i = 1$ to $\{\# \text{ of blocks}\}$ **do**
 $S_i \leftarrow$ Average saliency of i^{th} block
 $\beta_i \leftarrow \beta_{\min} + \left(\frac{S_i}{S_{max}}\right)^p (1 - \beta_{\min})$
 Tone map block i with (4.51) and set $\beta = \beta_i$
end for
Post processing to remove block boundaries

4.2.8 Region-of-Interest Based Local Tone Mappings

Li et al. propose a method [42] for using region of interest (ROI) based tone mapping after decomposing luminance of HDR radiance map into a base layer with large gradients and a detail layer with small gradients.

General Description

Initial HDR radiance map $\{R, G, B\}_h(p)$ is used to calculate the luminance [66]

$$Y_h(p) = 0.299 \cdot R_h(p) + 0.587 \cdot G_h(p) + 0.114 \cdot B_h(p), \quad (4.52)$$

where $p = (x, y)$. The luminance value can be decomposed into base layer, $Y_b(p)$, and detail layer, $Y_d(p)$, as

$$Y_h(p) = Y_b(p)Y_d(p). \quad (4.53)$$

Since the main variation is within the base layer, the ROI based global tone mapping algorithm compresses $Y_b(p)$ to obtain $\hat{Y}_b(p) \in [0, 1]$. The detail layer can be amplified to increase the local contrast to obtain $\hat{Y}_d(p)$. Then, the compressed luminance can be calculated as

$$Y_l(p) = \hat{Y}_b(p)\hat{Y}_d(p). \quad (4.54)$$

Finally, the $\{R, G, B\}_h(p)$, $Y_h(p)$ and $Y_l(p)$ are combined to generate the output LDR image $\{R, G, B\}_l(p)$.

Slight variations of the method are described for fully automatic operation or for some manual input. The ROI selection in automatic operation is based on the bright parts of the radiance map [7], in manual operation the ROI can be chosen by user.

Decomposition of $Y_h(p)$

The log function is used as basis for the compression. Applying the logarithmic transformation for Equation (4.53), we get

$$L_h(p) = L_b(p) + L_d(p), \quad (4.55)$$

where $L_h(p)$, $L_b(p)$, $L_d(p)$ represent $\log L_h(p)$, $\log L_b(p)$ and $\log L_d(p)$ respectively. Next, the detail layer is constructed (after which the base layer can be calculated easily) by using guidance field $\vec{v}(p) = (v_1(p), v_2(p))^T$ such that

$$\vec{v}_1(p) = L_h(p_r) - L_h(p), \quad (4.56)$$

$$\vec{v}_2(p) = L_h(p_b) - L_h(p), \quad (4.57)$$

where p_r and p_b are the right and bottom pixels of pixel p .

The optimal detail layer can be obtained by solving the following minimisation problem

$$\min_{L_d} \sum_p \left[L_d^2(p) + \lambda \left(\frac{(v_1(p) - \frac{\partial L_d(p)}{\partial x})^2}{\psi(v_1(p))} + \frac{(v_2(p) - \frac{\partial L_d(p)}{\partial y})^2}{\psi(v_2(p))} \right) \right], \quad (4.58)$$

where first term represents the energy of the details signal and second term for measuring the fidelity of $\nabla L_d(p)$ with respect to $\vec{v}(p)$. The function $\psi(z)$ is used to obtain a trade-off between two terms

$$\psi(z) = |z|^\gamma + \epsilon_1, \quad (4.59)$$

where γ and ϵ_1 are constants with default values 1.2 and 10^{-4} respectively. If L_d , v_1 and v_2 are vectors containing the pixels in $L_d(p)$, $v_1(p)$ and $v_2(p)$, then the optimal solution

can be obtained by solving

$$\left(\frac{1}{\lambda} I + D_x^T A(\mathbf{v}_1) D_x + D_y^T A(\mathbf{v}_2) D_y \right) \mathbf{L}_d = D_x^T A(\mathbf{v}_1) \mathbf{v}_1 + D_y^T A(\mathbf{v}_2) \mathbf{v}_2, \quad (4.60)$$

where I is the identity matrix, D_x and D_y are discrete differentiation operators and $A(\mathbf{v}_1) = \text{diag}(\frac{1}{\psi(v_1(p))})$.

ROI Selection

In the manual mode, the ROI is a rectangle defined by user

$$ROI = \left\{ (p) \mid \hat{Y}_{b,ROI}^{\max} \geq \hat{Y}_b(p) \geq \hat{Y}_{b,ROI}^{\min} \right\}, \quad (4.61)$$

where $\hat{Y}_{b,ROI}^{\max}$ and $\hat{Y}_{b,ROI}^{\min}$ are the maximal and minimal values of $\hat{Y}_b(p)$ in the ROI.

In the automatic mode, the ROI is defined as range of intensities

$$ROI = \left\{ (p) \mid c_2 \geq \hat{Y}_b(p) \geq c_1 \right\}, \quad (4.62)$$

where c_1 and c_2 are positive constants

$$c_1 = 0.15, \quad c_2 = \frac{Y_{white}}{1 + Y_{white}} \quad (4.63)$$

and Y_{white} is the smallest luminance that is mapped to pure white in [61]. The value for Y_{white} is determined by dynamic range of $Y(p)$ represented by R_t

$$Y_{white} = \frac{3}{64} R_t. \quad (4.64)$$

ROI Based Dynamic Range Allocation

Finally, more of the dynamic range is allocated to the ROI with following monotonic non-decreasing functions

$$\hat{Y}_b(p) = \begin{cases} \hat{Y}_b(p), & \text{if } \hat{Y}_b(p) > \hat{Y}_{b,ROI}^{\max} \\ \hat{Y}_b(p) \frac{\hat{Y}_{b,ROI}^{\min}}{\hat{Y}_{b,ROI}^{\max}}, & \text{if } \hat{Y}_b(p) < \hat{Y}_{b,ROI}^{\max} \\ \hat{Y}_b(p) \frac{\hat{Y}_b(p)}{\hat{Y}_{b,ROI}^{\max}}, & \text{otherwise.} \end{cases} \quad (4.65)$$

4.2.9 A Local Tone Mapping Operator for High Dynamic Range Images

Qian et al. present a local tonemapping algorithm [59] which is derived from Contrast Limited Adaptive histogram Equalization [80] (CLAHE). First, input image is segmented based on standard deviation of scaled luminance, then CLAHE is performed in each segment with different clip limit. Finally, the modified segments are combined to obtain the final tonemapped image.

Luminance Calculation

Input RGB values are converted to CIE XYZ colour space by the following formula [73]

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (4.66)$$

The luminance is represented by the Y component. The scaled luminance, $L(x, y)$ is obtained by

$$L(x, y) = a \cdot Y(x, y) \quad (4.67)$$

where a indicates whether the high dynamic range image is subjectively light, normal or dark. The a value is calculated by

$$a = 2^f \quad (4.68)$$

$$f = \frac{2Y_{av} - \log_2 Y_{min} - \log_2 Y_{max}}{\log_2 Y_{max} - \log_2 Y_{min}} \quad (4.69)$$

Segmentation

Input image is segmented into regions based on the standard deviation, σ , of the scaled luminance, $L(x, y)$. Image is successively divided into quadrants if $\sigma > 5$ for that region. Afterwards adjacent regions whose combined pixels $\sigma \leq 5$ are merged into one region.

Next, morphological operation (opening and closing) was used to close small holes in the image and to remove some noise. Finally, for each region CLAHE algorithm is applied with different clip limit based on the region mean and variance.

Combining the Regions

Color space is changed back to RGB

$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (4.70)$$

Using the CLAHE mapped result, T , and scaled luminance L , the final tone mapped image is obtained by

$$\begin{pmatrix} R_d \\ G_d \\ B_d \end{pmatrix} = T \cdot \begin{pmatrix} (\frac{R'}{L})^{0.5} \\ (\frac{G'}{L})^{0.5} \\ (\frac{B'}{L})^{0.5} \end{pmatrix} \quad (4.71)$$

4.3 Tone Mapping Algorithms for Videos

4.3.1 HDR Video Coding based on a Temporally Constrained Tone Mapping Operator

Cagri Ozcinar et al. present a temporally constrained content adaptive TMO in order to convert the input HDR video into a reduced bit depth video sequence which is then encoded using High Efficiency Video Coding (HEVC) in [55]

To improve Quality of Experience for video services producing highly realistic video is very important. Nowadays there are some investigated ways like high spatial resolutions, high frame rates, Wide Color Gamut(WCG) and HDR.

An approach based on a content-adaptive TMO as described in 4.5 is considered.

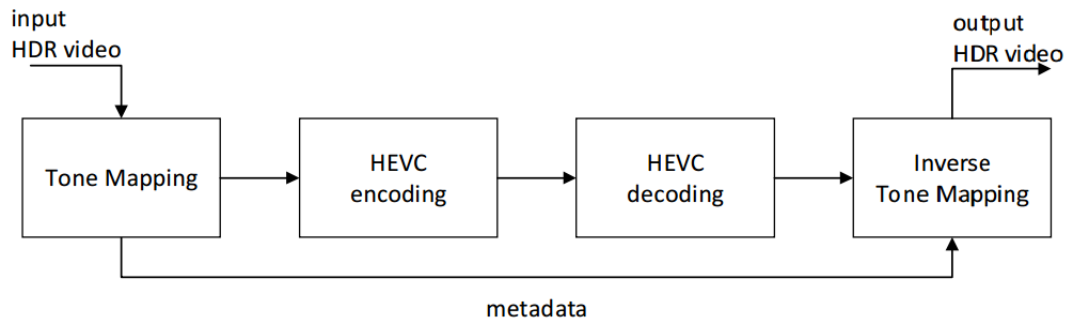


Figure 4.5: Block diagram of the proposed HDR video coding scheme. [55]

Temporal Regularization

A new temporal regularization term is proposed for optimization problem formulation. Optimization problem is defined as in 4.72

$$\min_s \left(\sum_{k=1}^N p_k s_k^{-2} + \lambda C(I_t) \right) \quad s.t. \quad \sum_{k=1}^N s_k = \frac{v_{max}}{\delta} \quad (4.72)$$

where p_k is the summation of the normalized histogram of luminance values for the k^{th} bin, N is the total number of bins in the histogram, v_{max} is the maximum tone mapped value, $C(I_t)$ is a temporal regularization term to favor temporal coherence of the tone mapped LDR video sequence, and λ is a positive constant

The proposed temporal regularization term is defined as

$$C(I_t) = \sum_{i,j} (I_t(i, j) - M(I_{t-1}(i, j)))^2 \quad (4.73)$$

where $I_t(i, j)$ denotes pixels in the tone mapped current LDR frame at time t , and $M(I_{t-1}(i, j))$ corresponds to pixels after motion compensation of the tone mapped LDR frame at time $t - 1$.

Convex Optimization

The problem that is defined in 4.72 can be written as:

$$\min_{s \in \Gamma} (f(s) + \lambda g(s)) \quad (4.74)$$

where the minimization is performed on a convex set.

In this proposed method, they consider the primal-dual M+LFBF algorithm proposed in [15] to address general convex optimization.

4.3.2 An Optimal Video-Surveillance Approach for HDR Videos Tone Mapping

In [8] Leonardi et al. proposes a fast method for tone mapping HDR videos, which combines the benefits of both local and global operators. The method is applied in the context of object detection and tracking.

The goal of the proposed algorithm is to provide an efficient and effective tone map algorithm, by using global and local tone map algorithms. An overview of the approach is shown in Figure 4.6.

The main idea is to apply a local tone map operator only inside the areas of interest for the considered surveillance application, while the remaining part of the picture is tone mapped using a global algorithm. By doing so, the computational time is reduced and the image contrast is enhanced in all the relevant areas (see Fig. 4.6).

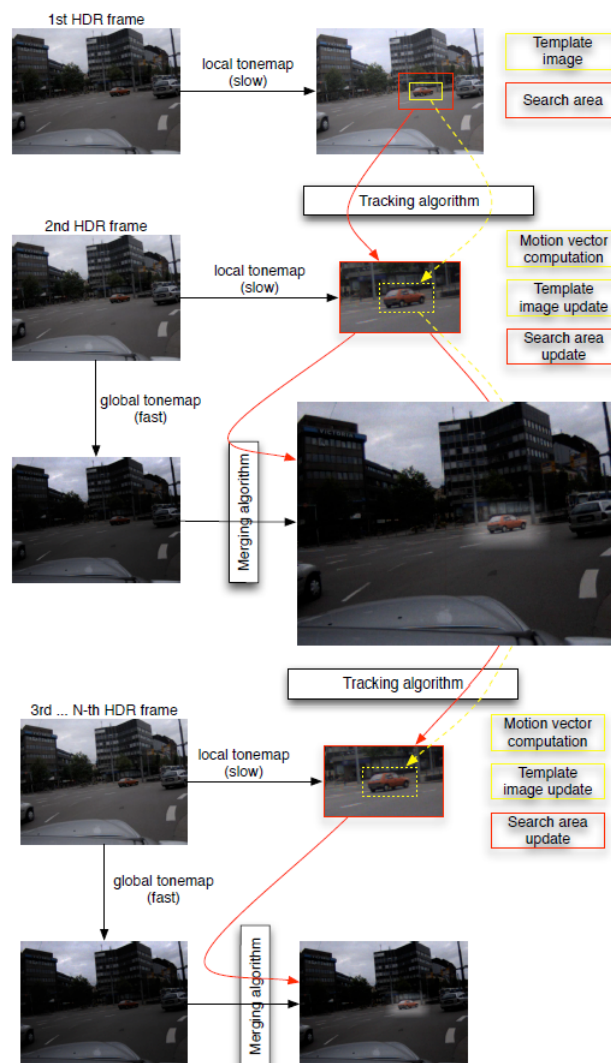


Figure 4.6: An intuitive block schema of the proposed algorithm[8]

In the proposed method the considered HDR video is initially partitioned into GOPs, (Group Of Pictures), where the first frame represents a "Key Picture". All of the Key

Pictures are tone mapped by applying a local algorithm, while all the other frames in a GOP are adaptively tone mapped.

The process is initialised by selecting a number of Regions of Interest (ROI) from the key picture. A local tone mapping is then performed on the previously selected ROIs and enlarged according to the search window of the second frame. Then, the surveillance algorithm is applied to each of these areas (e.g. the red car in Fig. 4.6). Afterwards, ROI parameters (displacement and size) are updated, according to the output of the object tracking algorithm. The updated parameters are then applied to the successive frame. Finally the process is repeated for all the other frames inside the GOP.

The length GL determines the number of frames that compose a GOP. The time required for tone mapping an entire video, composed of N frames, is given by the following equation:

$$T_{TM}[s] = \frac{N}{GL} \overline{T_{TMglobal}} + \left(N + \frac{N}{GL} \right) \overline{T_{TMupdate}} \quad (4.75)$$

where $\overline{T_{TMglobal}}$ is the average time needed for the global tone mapping algorithm and $\overline{T_{TMupdate}}$ is the average time needed for the local tone mapping of ROIs and the bi-linear interpolation. An optimal GL needs to be computed and set during the initialisation phase.

Local tone mapping of ROIs and bi linear interpolation

Let assume that the video-surveillance algorithm, applied on the key picture determines k interesting ROIs, and each of them is a different rectangle with a size (w, h) and centred in the pixel coordinate (x, y) . According to this assumption, each search area is tone mapped with a local TMO and the tracking algorithm is applied to the area. In order to correctly visualise this image, a bi linear interpolation is then used.

The goal of the blending is to merge the local-tone mapped ROIs with the global-tone mapped image. The interpolation filter has the shape of a truncated pyramid, as shown in Figure 4.7.

The bi-linear interpolation process can be split into nine parts, according to Figure 4.8.

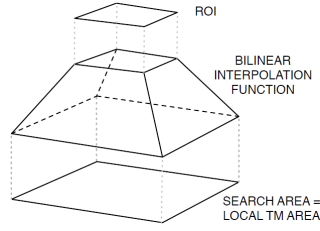


Figure 4.7: Truncated pyramid interpolation filter[8]

In the following formula's $I_G(x, y, c)$ is the value of the component $c(c \in \{R, G, B\})$ of the pixel (x, y) in the global tone mapped image, and $I_L(x, y, c)$ is the same component of the same pixel in the local tone mapped search area. dx and dy are the normalised distance between the pixel (x, y) and the center of the relative zone.

1. Corner zones (A, C, I, K). The interpolated value of (x, y, c) is:

$$I_C(x, y, c) = \min(dx, dy) \cdot I_G(x, y, c) + (1 - \min(dx, dy)) \cdot I_L(x, y, c) \quad (4.76)$$

2. ROI adjacent zones (B, D, F, J). Interpolated value is:

$$I_C(x, y, c) = (1 - dx) \cdot I_G(x, y, c) + dx \cdot I_L(x, y, c) \quad \text{or} \quad (4.77)$$

$$I_C(x, y, c) = (1 - dy) \cdot I_G(x, y, c) + dy \cdot I_L(x, y, c)$$

3. ROI zone (E). Interpolated value is:

$$I_C(x, y, c) = I_L(x, y, c) \quad (4.78)$$

The proposed tone mapping strategy can be effectively used for HDR video-surveillance purposes. The method combines the benefits provided by both local and global TMOs, thus providing a fast and good quality LDR image.

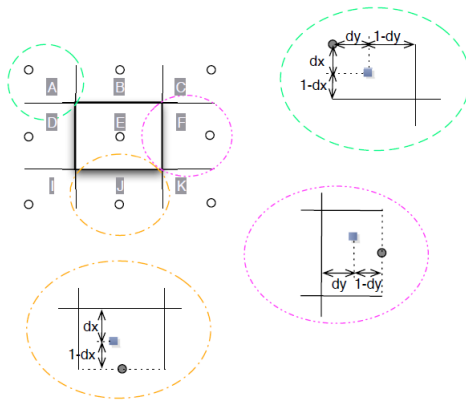


Figure 4.8: Bi-linear interpolation in the four zones around the ROI [8]

4.4 Algorithms using Exposure Fusion

4.4.1 Direct High Dynamic Range Imaging Method Using Experimental Optimal Exposure Criterion

Zhang and Dai propose a method for fusing LDR images with different exposure time directly to achieve improved LDR image directly displayable by regular display device[79], instead of a regular process of constructing HDR radiance map and then using tone mapping algorithm. The idea is based on experimental results that there exists an

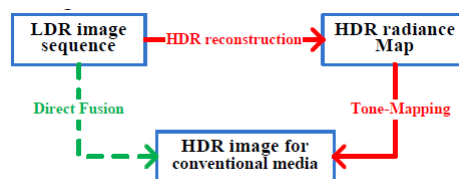


Figure 4.9: Zhang and Dai method (Direct Fusion) vs regular HDR process

optimal exposure rate for each pixel which is then used in the final image. That optimal exposure is observed to be positioned at the sparsest position of the pixel value curve.

First, for each pixel the optimal exposure rate is found by taking the maximum of first derivate of cord length parameterization [22] and then he appropriate pixel value is used. In order to reduce the noise, the fusion map generated from optimal exposure rates, is

smoothed by guided image filtering [25].

Since no HDR map is generated and the LDR images are fused directly, the method is computationally much more efficient than other common HDR methods while providing comparable results.

4.4.2 Alternating Line High Dynamic Range Imaging

Cho et al. propose a method for improving dynamic range of an imaging device by changing the exposure time line by line [13]. This allows to obtain a multi-exposure image without taking several captures. Afterwards, the lost vertical resolution is recovered by image processing algorithms. Finally, two acquired images (dark and bright) are combined using a standard HDR algorithm [17]. The main benefit of the method is that input images are perfectly aligned and not affected by motion artifacts.

The inputs for the algorithm are two images, SET (small exposure time) and LET (large exposure time). They have been obtained by alternating exposure time line by line vertically, thus both have half the vertical resolution of the entire image. The question is, how to fill the missing lines for both images in the best way. When to consider both images separately, the problem of filling in the missing lines is known as *Interlaced to Progressive Conversion* (IPC) problem [6]. However, in the current case, there is some info about the missing lines in the other image. In order to use information from the other image, Multi-Level Gain Compensation (MLGC) method is combined with IPC to obtain the result for the missing lines. MLGC for LET image can be calculated by

$$g_k = \frac{\sum_{x \in B_k, x' \in B'_k} S(x) \cdot L(x')}{\sum_{x \in B_k} S^2(x)} \quad (4.79)$$

where $S(x)$ is pixel value from SET image, $L(x')$ is best estimate for LET image value at x , B_k is image region in SET containing grey value of k and B'_k is neighbourhood of B_k where $L(x') \approx L(x)$. Best estimate of $L(x)$ can be obtained by analysing area B'_k . For flat regions, the average of spatially closest pixels is a good estimation, for edge regions the edge direction might be needed to consider. However, if data in region

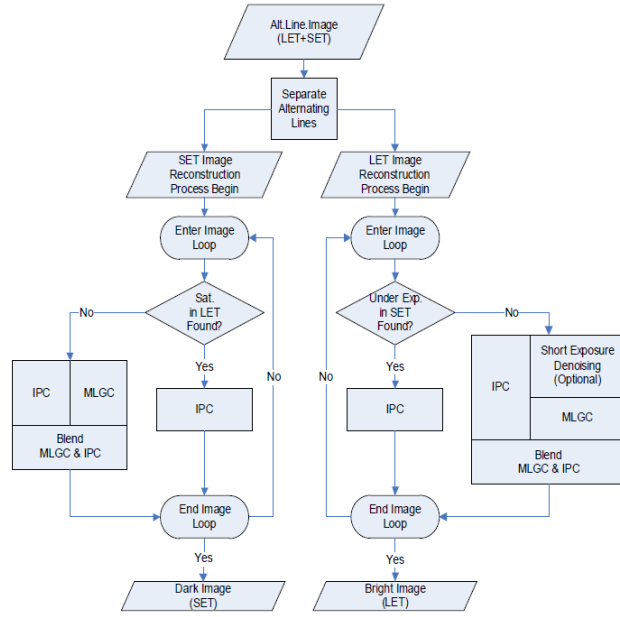


Figure 4.10: Flowchart of proposed method in[17]

B'_k in SET image is saturated or noisy, the true $L(x)$ value might not be possible to estimate. In such a case, data from SET image is not used and IPC is used directly. Similar analysis can be done for obtaining the missing lines in SET image. Finally the two obtained images are combined into single HRD image.

4.4.3 Probabilistic Exposure Fusion

In [72] Mingli Song et al. proposes a new scheme to handle HDR scenes by integrating locally adaptive scene detail capture and suppressing gradient reversals introduced by the local adaptation. The proposed scheme can capture an HDR scene by using a standard dynamic range (SDR) device and synthesize an image suitable for SDR displays. Moreover, the proposed scheme also functions as the tone mapping of an HDR image to the SDR image.

The proposed method contains two steps, which are the scene detail extraction and the SDR image synthesis.

1) Scene detail extraction: By using an SDR capture device, we can record details of a scene in a series of SDR images at different exposure levels.

2) Aprobabilistic model for the SDR image synthesis: Aprobabilistic model is proposed to synthesize an SDR image based on the calculated image luminance and scene gradient at each point in the scene.

By using these captured SDR images, we can extract scene details at each point, i.e., the luminance and the scene gradient. For each point (x, y) , we calculate the image luminance that maximizes the visible contrasts over different captured SDR images. The visible contrast $\nu(x, y)$ is contrast $\Delta I_H(x, y)$ above a predefined threshold, wherein $\Delta I_H(x, y)$ is the luminance difference between point (x, y) and the mean of its surrounding points,

$$\nu(x, y) = T(\Delta I_H(x, y))\Delta I_H(x, y) \quad (4.80)$$

where $T(\gamma) = 1$ when γ is larger than or equal to a predefined threshold; otherwise, it is 0. $\Delta I_H(x, y) = \|\Delta I_H(x, y) - \Delta \bar{I}_H(x, y)\|$, and $\Delta \bar{I}_H(x, y) = (1/M \times M) \sum_{i=x-M/2}^{x+M/2} \sum_{j=y-M/2}^{y+M/2} I_H(i, j)$ is the mean of the surrounding points.

Taking the Weber law into account, a coefficient $c(\bar{I}_H(x, y))$ is defined by (4.81) to make the visible contrast extraction more humanlike, where $c(\bar{I}_H(x, y))$ defines a piecewise function to reflect the slope variation according to different means of the surrounding luminance $I_H(x, y)$.

$$c(\bar{I}_H(x, y)) = \begin{cases} 1/(0.035 - 0.009\bar{I}_H(x, y))(\bar{I}_H(x, y) + 1) & \text{if } 0 \leq \bar{I}_H(x, y) < 60 \\ 1/(0.035(\bar{I}_H(x, y) + 1)) & \text{if } 60 \leq \bar{I}_H(x, y) \leq 200 \\ 1/(0.035 + 0.001(\bar{I}_H(x, y) - 200))(\bar{I}_H(x, y) + 1) & \text{if } 200 < \bar{I}_H(x, y) \leq 255 \end{cases} \quad (4.81)$$

The coefficients are determined according to [26]. By multiplying the directly computed contrast by the adaptive coefficient (4.81), the visible contrast defined in (4.82) is then updated as $I_H(x, y)$.

$$\nu(x, y) = T(c(\bar{I}_H(x, y)))\Delta \bar{I}_H(x, y)c(\bar{I}_H(x, y))\Delta I_H(x, y) \quad (4.82)$$

The predefined threshold values are set for all the tests empirically. To observe stronger

visible contrast defined by (4.82) for each point, the exposure level is adjusted to incident-light quantity adaptively, and the visibility of the contrast is consequently enhanced.

For point (x, y) constructed by multiple SDR images with different exposure the local adaptation can be seen as a process to find a luminance value that maximizes, i.e.,

$$\mu_I(x, y) = \underset{I_H(x, y)}{\arg \max} \nu(x, y) \quad (4.83)$$

μ_I is directly obtained by enumerating N captured images, i.e., $\mu_I(x, y)$ is the luminance corresponding to the maximal visible contrast $\nu(x, y)$ among all N captured SDR images.

The scene gradient of a point is reflected by the gradient that is perceivable by human eyes, called visible gradient, and that can be measured by counting the number of visible differences of luminances between neighboring pixels in the window. The visible gradient can be computed by

$$\varphi(x, y) = \sum_{i=x-M/2}^{x+M/2} \sum_{j=y-M/2}^{y+M/2} T(c(\bar{I}_H(x, y))) \cdot VMAX(|\nabla I_H(x, y)|) \quad (4.84)$$

where $\nabla I_H(x, y) = (\partial I_H(x, y)/\partial x, \partial I_H(x, y)/\partial y)$ and $VMAX \left(\begin{bmatrix} u \\ v \end{bmatrix} \right) = \begin{cases} u & \text{if } u \leq |v| \\ v & \text{if } u < |v|. \end{cases}$

The scene gradient extraction is viewed as a process of finding gradient that maximizes the quantity of the visible gradient $G(x, y)$,

$$G(x, y) = \underset{\nabla I_H(x, y)}{\arg \max} \psi(x, y) . \quad (4.85)$$

$G(x, y)$, is directly obtained by enumerating N captured images. If $\psi(x, y)$ is equal over all $\nabla I_{H, e_i(x, y)}$ for point (x, y) , then $G(x, y)$ can be set as any one of them.

Based on the maximization of the quantity of the visible gradient, we can extract the scene gradient at each point which are used to suppress the gradient reversals in the extracted image luminance gradient.

The Probabilistic Model for a given scene I_H , the corresponding SDR image I_S should preserve the visible contrasts presented in I_H and suppress the gradient reversal between I_H and I_S . This is formulated in the following MAP framework:

$$\begin{aligned}
I_S &= \max_{I_s} p(I_S|I_H) \\
&= \max_{I_s} p(I_H|I_S) p(I_S/p(I_H)) \\
&\propto \max_{I_s} p(I_H|I_S) p(I_S) \\
&= \max_{I_s} \prod_{x,y} p(I_H(x,y)|I_S(x,y)) p(I_S(x,y))
\end{aligned} \tag{4.86}$$

where $I_H(x,y)$ and $I_S(x,y)$ are the intensity of a pixel located at (x,y) in I_H and I_S . $p(I_S(x,y))$ is the probability function of the visible contrasts captured from I_H and preserved in I_S , and $p(I_H(x,y)|I_S(x,y))$ is the probability function of the gradient consistency between I_H and I_S .

To eliminate the gradient reversals, $p(I_H|I_S)$ works as a penalty function by considering the gradient consistency between the scene gradient and that of the synthesized SDR image I_S .

Based on the calculated $\mu_I(x,y)$ and $G(x,y)$ we define $p(I_S(x,y))$ and $p(I_H(x,y)|I_S(x,y))$, as

$$p(I_S(x,y)) \propto \exp(-\sigma_S^{-2}(I_S(x,y) - \mu_I(x,y))^2) \tag{4.87}$$

and

$$p(I_H(x,y)|I_S(x,y)) \propto \exp(-\sigma_C^{-2} \|\nabla I_S(x,y) - G(x,y)\|^2) \tag{4.88}$$

where item $I_S(x,y) - \mu_I(x,y)$ in (4.87) reflects the luminance difference between $I_S(x,y)$ and $\mu_I(x,y)$ i.e., the value that locally maximized the visible contrasts according to (4.83); item $\|\nabla I_S(x,y) - G(x,y)\|$ show the distance between the gradient

of the synthesized SDR image $\nabla I_S(x, y)$ and the gradient of the scene $G(x, y)$; and σ_S^2 and σ_C^2 are predefined covariances.

Based on (4.87) and (4.88), (4.86) can be furthered as a MAP problem in following form

$$\max_{I_S} \sum_{(x,y) \in I} \left[\begin{array}{l} -\sigma_S^{-2} (I_S(x, y) - \mu_I(x, y))^2 \\ -\sigma_C^{-2} \|\nabla I_S(x, y) - G(x, y)\|^2 \end{array} \right] \quad (4.89)$$

For scene modeling by using SDR devices, (4.89) can be simplified as an integral-minimization-based optimization procedure. In the continuous form, I_S minimizes the following integral:

$$\begin{aligned} & \iint F(\nabla I_S(x, y), I_S(x, y)) dx dy \\ &= \iint \left[\sigma_S^{-2} (I_S(x, y) - \mu_I(x, y))^2 \right. \\ & \left. + \sigma_C^{-2} \|\nabla I_S(x, y) - G(x, y)\|^2 \right] dx dy \\ &= \iint \left(\sigma_C^{-2} \left(\frac{\partial I_S(x, y)}{\partial x} - G_x(x, y) \right)^2 \right. \\ & \quad \left. + \sigma_C^{-2} \left(\frac{\partial I_S(x, y)}{\partial y} - G_y(x, y) \right)^2 \right. \\ & \quad \left. + \sigma_S^{-2} (I_S(x, y) - \mu_I(x, y))^2 \right) dx dy . \end{aligned} \quad (4.90)$$

According to the variational principle, function $I_x(x, y)$ that minimizes the integral in (4.90) must satisfy the Euler–Lagrange equation,

$$\frac{\partial F}{\partial I_S} - \frac{d}{dx} \frac{\partial F}{\partial (I_S)_x} - \frac{d}{dy} \frac{\partial F}{\partial (I_S)_y} = 0 \quad (4.91)$$

which is a partial differential equation with respect to I_S . Substituting F in (4.90) to (4.91), we obtain the following equation

$$\begin{aligned} & 2\sigma_S^{-2} (I_S - \mu_I) - 2\sigma_C^{-2} \left(\frac{\partial^2 I_S}{\partial x^2} - \frac{\partial G_x}{\partial x} \right) \\ & \quad - 2\sigma_C^{-2} \left(\frac{\partial^2 I_S}{\partial y^2} - \frac{\partial G_y}{\partial y} \right) = 0 \end{aligned} \quad (4.92)$$

Dividing previous equation by 2 and rearranging the items we obtain

$$\nabla^2 I_S - \frac{\sigma_C^2}{\sigma_S^2} I_S = \text{div} G - \frac{\sigma_C^2}{\sigma_S^2} \mu_I \quad (4.93)$$

where $\nabla^2 I_S$ is the Laplacian operator. $\nabla^2 I_S = (\partial^2 I_S / \partial x^2) + (\partial^2 I_S / \partial y^2)$ and $divG = (\partial G_x / \partial x) + (\partial G_y / \partial y)$ is the divergence of the scene gradient. By integrating all points in the scene, we use (4.93) to form a linear equation system.

The implementation procedure of the proposed algorithm for scene capturing and display by using SDR devices is shown as follows.

- 1) Convert a color image from the red-green-blue (RGB) color space to the YUV color space.
- 2) Initialize the luminance value with $\mu_I(x, y)$ according to (4.83).
- 3) Compute the scene gradient according to (4.85).
- 4) Update the luminance value of each pixel according to (4.93).
- 5) Convert the resulted image from the YUV color space to the RGB color space.

The proposed approach operates on luminances of an image pixel by pixel. In our approach, we use the Y component in the YUV color space as the luminance to be processed. For step 4, we compute Laplacian and divergence values for pixel (x, y) , based on the following two numerical approximations:

$$\begin{aligned} \nabla^2 I_S(x, y) &= I_S(x + 1, y) + I_S(x - 1, y) \\ &+ I_S(x, y + 1) + I_S(x, y - 1) - 4I_S(x, y) \end{aligned} \quad (4.94)$$

$$divG(x, y) = G_x(x, y) - G_x(x - 1, y) + G_y(x, y) - G_y(x, y - 1). \quad (4.95)$$

The derivatives around boundaries are set as 0. Based on (4.94) and (4.95), (4.93) can be transformed to a linear equation system $Ax = b$, where A is a sparse coefficient matrix, x consists of unknown luminance values, and b comes from $divG(x, y) - (\sigma_C^2 / \sigma_S^2) \mu_I(x, y)$ where σ_C^2 / σ_S^2 was set as 1.

This article proposed a new probabilistic exposure fusion scheme by taking into account local adaptation and gradient reversal suppression. A probabilistic model has been then proposed to carry out the scene reproduction (SDR image synthesis) by preserving both the visible contrasts and the gradient consistency. This approach is claimed to be ad-

vantageous compared with representative existing tone mapping operators.

4.4.4 Evaluation of HDR Tone Mapped Image and Single Exposure Image

In [57] YongQing uses subjective evaluation to compare tone mapped image (HDR image after tone mapping) and LDR image captured by LDR camera. The tone mapping is done by Fattal et al's [23] algorithm (eq. 4.96). In their experiment, 40 participants, between 10 to 40 years old, were used to compare the tone mapped images and their corresponding single exposure images, and vote the most preferred one.

$$L_w = \frac{1}{N} \exp \left(\sum_{x,y} \log(\delta + L_w(x, y)) \right) \quad (4.96)$$

Where $L_w(x, y)$ is the "world" luminance for pixel (x, y) , N is the total number of pixels in the image and δ is a small value to avoid singularity if the pixel is black.

Table 4.1: Comparative results of five images.

image	indoor	outdoor	church	tree	ball
LDR	21	24	18	17	21
TMO	19	16	22	23	19

The experimental results show that the preference of images depends on the scenes, but in general the tone mapped image and single exposure image show no large differences.

4.5 Inverse Tone mapping

4.5.1 Content-adaptive inverse tone mapping

The work in [58] proposed a histogram-based method for inverse tone mapping which enhances the over-exposed regions and also the remaining well-exposed regions. The proposed algorithm contains a content-adaptive inverse tone mapping operator, which has different responses with different scene characteristics.

Fig. 4.11 shows the flow chart of the system.

The proposed algorithm mainly consists of two components: scene-adaptive inverse tone mapping and over-exposure enhancement. The scene-adaptive inverse tone mapping algorithm includes a scene classifier and the inverse tone mapping operator. The proposed inverse tone mapping operator is based on the work of Schlick [68] which contain manually set environment-dependant parameters. To counter this problem a SVM scene classifier was trained to recognize scenes so that these parameters can be decided automatically. For the over-exposure enhancement step, the method is based on the works of Rempel et al. [65] and Meylan et al. [50] to make up for the deficiency in each work.

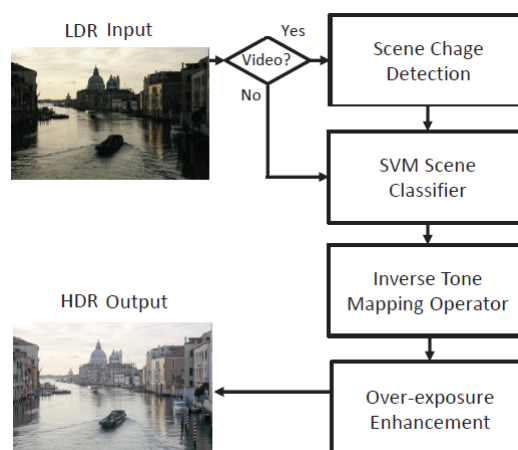


Figure 4.11: The flow chart of the system [29]

Inverse tone mapping operator based on the work of Schlick [68] is

$$L_d = \frac{pL_w(x, y)}{(p - 1)L_w(x, y) + L_{max}}, \text{ where } p \in (1, \infty) \quad (4.97)$$

L_d and L_w refers to the display luminance and world luminance. The parameter p can be approximated by

$$p = \frac{\delta L_0 L_{max}}{N L_{min}} \quad (4.98)$$

The L_0 term in (4.98) is the just-noticeable difference (JND). The minimum luminance that is observed different from black is L_0 . In the experiment, the L_0 is about 7-10. The L_{max} and L_{min} represent to the maximum luminance and minimum luminance in world scenes. L_{max}/L_{min} can be seemed as the contrast ratio of one scene and is set to 103. With $N = 256$, the value of p is 27.34375 (for simplicity, p is set as 30). The proposed inverse tone mapping operator can be directly derived from (4.97) as follows:

$$L_w = \frac{L_d L_{max}}{p(1 - L_d) + L_d} \quad (4.99)$$

With the experiment-decided p , the only parameter to be decided in (4.99) is L_{max} .

We use the libsvm tool provided by Chih-Chung Chang and Chih-Jen Lin [12] to train our scene classifier. One of the features extracted from the data is a histogram in 512 dimensions, where each RGB channel is quantised from 256 bins to 8 bins. This classifier is used for determining the parameters of the operator, taking scene information into consideration. There are three classes in the SVM classifier: "bright", "midtone," and "dim". Which represent scenes of high, middle, and low luminance levels. The class of an image depends on the luminance level and histogram. With three luminance levels and dynamic ranges, the content can be mapped to the parameters matching to the scenes. The L'_{max} s has three conditions: $106cd/m^2$, $103cd/m^2$ and $1cd/m^2$ and the most propitiate one is chosen according to the results of the classification.

The lost information in over-exposed regions is synthesised. Fig.4.12 shows the over-exposure enhancement flow. At first, the input image is passed through one low-pass

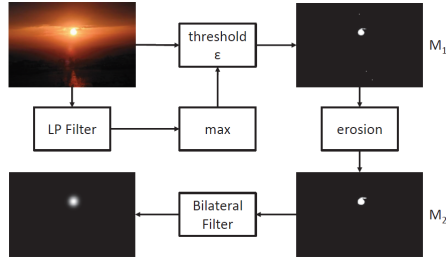


Figure 4.12: The over-exposure enhancement flow [29]

filter. After passing through this filter, the biggest values in the images are chosen to be the threshold ϵ . Then, we have the first binary map M_1 with threshold ϵ

$$M_1(p) = \begin{cases} 1 & \text{if } I(p) > \epsilon \\ 0 & \text{otherwise} \end{cases} \quad (4.100)$$

With morphological erosion to eliminate single 1's resulted from noise, we get binary map M_2 :

$$M_2(p) = \begin{cases} 0 & \text{if } M_1(p) = 0 \text{ or } dil(M_1(p)) = 0 \\ 1 & \text{otherwise} \end{cases} \quad (4.101)$$

1's in M_2 can be regarded as the bright points or the specular points. To determine the over-exposed region, we pass M_2 and the input image through a joint bilateral filter with a large kernel size, leading to the "fade out" effect. The binary map M_2 is computed according to the characteristics of the image, rather than a threshold for every image.

The proposed content-adaptive inverse tone mapping algorithm enhances the over-exposure regions and takes the different scene responses into account. This work is a "histogram-based" method which uses SVM classifier for automatic parameter decision. This paper focuses on inverse tone mapping for videos, which take temporal information into consideration.

5 Experimental Results

5.1 Experimental Setup

We used 19 tone mapping operators which were defined as best by researches. Those are listed below :

Ashikhmin, Banterle, BruceExpoBlend, Chiu, Drago, Durand, Exponential, KimKautz-Consistent , Krawczyk, Kuang, Lischinski, Mertens, Raman, ReinhardDevlin, Reinhard, Schlick, TumblinRushmeier, WardGlobal, WardHistAdj

HDR Toolbox [4] was used to process this tone mapping operators.

We have chosen 4 images which we believe are completely different for their luminance level: inside, outside, very high range, comparingly low dynamic range. The images we have used are as following. In these specific images we have used Drago TMO to display them as in 5.1

There were more than 25 subjects for the study. Each subject was asked to be in the same light environment in this experiment - not so bright, not so dark, but a normal indoor. We asked about their age, color blindness, and corrective glass usages. Subject were given quite enough time for this experiment so that they wouldn't get bored and make their effort more accurate.

Subjects were given TMOs anonymously, so they didn't know what they are voting for,

to make it more accurate. We used different kind of tmos, including global and local operators.

First we did parameter adjustments for each tone mapping operator ranging $-0.5 + 0.5$ mostly. For each image the result was absolutely different from one another. For example, the same tone mapping operator gave better result with $+0.5$ adjustments meanwhile the other image was totally dark. This means luminance level affects image conversion a lot.

Table 5.1: Images used in our experiment



5.2 Results

Subjects' feedback had to be between 0 and 100. The votes were extremely different for each image. We calculated the average of subjects' rating for the tone mapping operators per image. Surprisingly some tmos gave us very different results from one image to the other. For example, WardGlobalTMO got 53.9 out of 100 for one image and 0.75 / 100 for another. 5.2 shows this average results. 5.1, 5.2, 5.3 and 5.4 shows these results for each image separately and more clearly.

Table 5.2: Average results of subjects' votes for images

TMO	Image 1	Image 2	Image 3	Image 4
Ashikmin	52.15	31.25	51.4	46
Banterle	61.3	33.9	55.4	62.2
BruceExpoBlend	21.9	31.9	41.8	57.75
Chiu	20.55	19.5	36.2	35.4
Drago	62.75	29.45	53.15	53.2
Durand	37.75	18.55	12.05	59.55
Exponential	36.85	15.05	44.1	53.5
KimKautzConsistent	48.95	26.05	41.4	46.3
Krawczyk	42.4	12.55	30.4	54.3
Kuang	37.3	26.85	53.85	46.45
Lischinski	48.45	34.3	58.95	31.5
Mertens	54.65	35.9	33.35	44.1
Raman	30.45	24.95	17.2	25.9
ReinhardDevlin	51.4	33.5	23.65	15.55
Reinhard	50.55	31.7	52.55	28.45
Schlick	23	18.5	21	30.4
TumblinRushmeier	6.05	16.85	39.5	30.65
WardGlobal	0.75	10.3	49.45	53.9
WardHistAdj	36.25	33.25	46.4	24

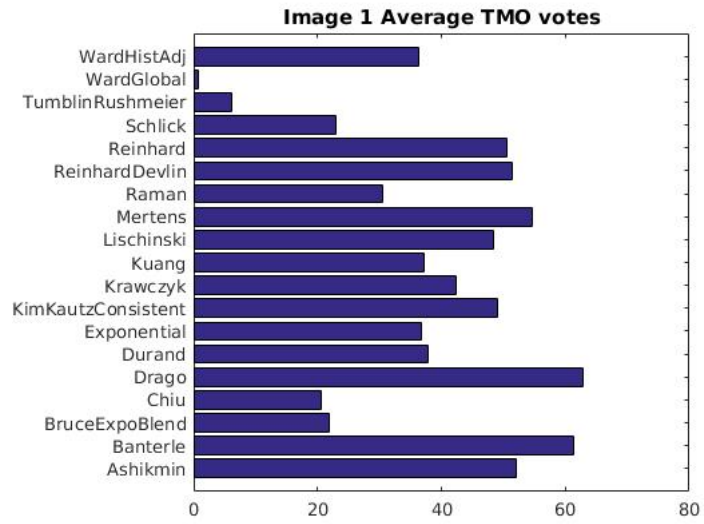


Figure 5.1: Average results of subjects' votes for 1st image

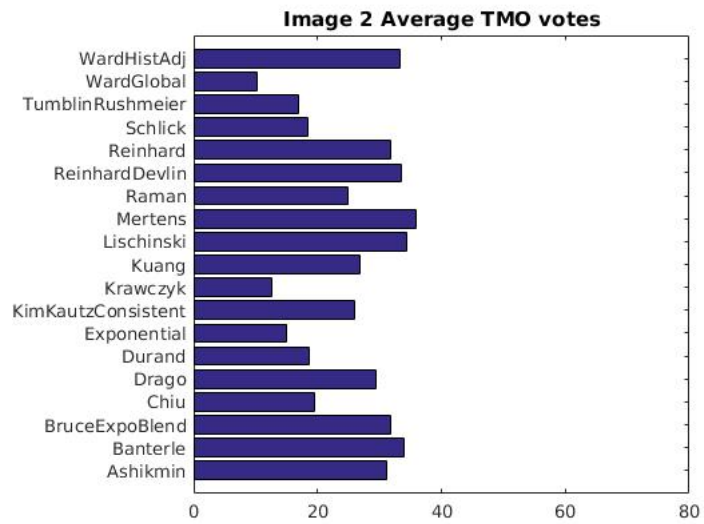


Figure 5.2: Average results of subjects' votes for 2nd image

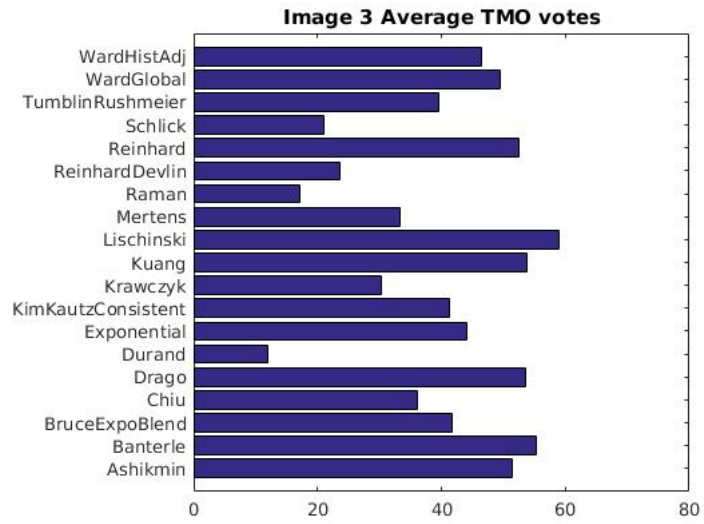


Figure 5.3: Average results of subjects' votes for 3rd image

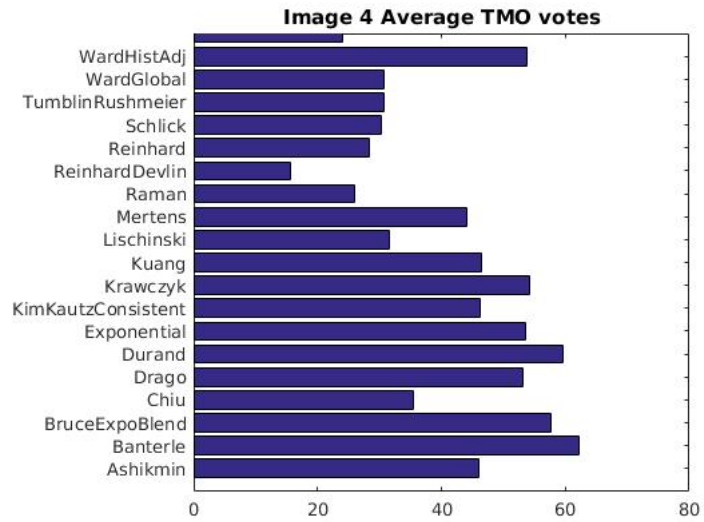


Figure 5.4: Average results of subjects' votes for 4th image

6 Conclusion and Future Work

6.1 Conclusion

Our experiments show that quality rankings of tone mapping operators are directly determined by the evolution criteria. Principally, a global TMOs is better than a local TMO on global criteria experiments.

We did some parameter adjustments ranging -0.5 - $+0.5$ for each tone mapping operator, changing their default values and asked subjects. Some TMOs made difference but, mostly it made the images worse than with the ones with the default values given in each study accordingly.

Evaluated average of the overall tests show that there is no TMO that gives better results among the others for all the images with different luminance levels. So when choosing a TMO, one should compare the image scene he/she is reproducing. And overall, results are not very perfect, limiting itself around 60/100.

6.2 Future Work

Since we have analyzed most of known tone mapping operators, we now are able to define which future of which TMO is reasonable to use. In the future we will combine all these ideas we have got so far from different approaches on high dynamic range imaging and create our own tone mapping operator.

Acknowledgements

I would like to thank my supervisors Gholamreza Anbarjafari and Dr. Cagri Ozcinar for all the advice and guidance on writing this thesis. This thesis has been partially supported by Estonian Research Council Grant PUT638.

References

- [1] Michael Ashikhmin. “A Tone Mapping Algorithm for High Contrast Images”. In: *Proceedings of the 13th Eurographics Workshop on Rendering*. EGRW '02. Eurographics Association, 2002, pp. 145–156.
- [2] Jonghyun Bae et al. “Adaptive tone-mapping operator for HDR images based on image statistics”. In: *TENCON 2011-2011 IEEE Region 10 Conference*. IEEE, 2011, pp. 1435–1438.
- [3] Francesco Banterle et al. “A framework for inverse tone mapping”. In: *The Visual Computer* 23.7 (2007), pp. 467–478.
- [4] Francesco Banterle et al. *Advanced High Dynamic Range Imaging: Theory and Practice*. Natick, MA, USA: AK Peters (CRC Press), 2011. ISBN: 9781568817194.
- [5] Francesco Banterle et al. *Advanced high dynamic range imaging: theory and practice*. CRC Press, 2011.
- [6] Erwin B Bellers and Gerard de Haan. *De-interlacing: A Key Technology for Scan Rate Conversion: A Key Technology for Scan Rate Conversion*. Vol. 9. Elsevier, 2000.
- [7] H Richard Blackwell. “Contrast thresholds of the human eye”. In: *JOSA* 36.11 (1946), pp. 624–643.
- [8] Leonardi R. & Okuda M Boschetti A. Adami N. “An optimal video-surveillance approach for HDR videos tone Mapping”. In: *19th European Signal Processing Conference (EUSIPCO 2011)*. IEEE, 2011, pp. 274–277.

- [9] Burt and Adelson. “The Laplacian Pyramid as a Compact Image Code”. In: vol. 31. 4. IEEE. 1983, pp. 532–540.
- [10] Nicholas J Butko et al. “Visual saliency model for robot cameras”. In: *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. IEEE. 2008, pp. 2398–2403.
- [11] Ayan Chakrabarti et al. “Modeling Radiometric Uncertainty for Vision with Tone-mapped Color Images”. In: IEEE, 2014, pp. 2185–2198.
- [12] Chih-Chung Chang and Chih-Jen Lin. “LIBSVM: A Library for Support Vector Machines”. In: vol. 2. 3. ACM, 2011, 27:1–27:27.
- [13] Seungki Cho et al. “Alternating line high dynamic range imaging”. In: *Digital Signal Processing (DSP), 2011 17th International Conference on*. IEEE. 2011, pp. 1–6.
- [14] Kim Choi and Yun. “Tone mapping for HDR image using chromatic adaptation method”. In: *2012 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE. 2012, pp. 604–605.
- [15] P. L. Combettes and J.-C. Pesquet. “Primal-dual splitting algorithm for solving inclusions with mixtures of composite, lipschitzian, and parallel-sum type monotone operators, Set-Valued and Variational Analysis”. In: Jun. 2012, pp. 307–330.
- [16] Praveen Cyriac et al. “A Tone Mapping Operator Based on Neural and Psychophysical Models of Visual Perception”. In: *Human Vision and Electronic Imaging XX 9394* (2015).
- [17] Paul E Debevec and Jitendra Malik. “Recovering high dynamic range radiance maps from photographs”. In: *ACM SIGGRAPH 2008 classes*. ACM. 2008, p. 31.
- [18] Frédéric Drago et al. “Adaptive logarithmic mapping for displaying high contrast scenes”. In: *Computer Graphics Forum*. Vol. 22. 3. Wiley Online Library. 2003, pp. 419–426.
- [19] Jiang Duan et al. “Tone-mapping high dynamic range images by novel histogram adjustment”. In: *Pattern Recognition* 43.5 (2010), pp. 1847–1862.

- [20] Frédo Durand and Julie Dorsey. “Fast bilateral filtering for the display of high-dynamic-range images”. In: *ACM transactions on graphics (TOG)*. Vol. 21. 3. ACM. 2002, pp. 257–266.
- [21] T. Annen F. Drago K. Myszkowski and N. Chiba. “Adaptive Logarithmic Mapping For Displaying High Contrast Scenes”. In: vol. 22. *Computer Graphics forum*, 2003, pp. 419–426.
- [22] Gerald Farin. *Curves and surfaces for computer-aided geometric design: a practical guide*. Elsevier, 1997.
- [23] Raanan Fattal, Dani Lischinski, and Michael Werman. “Gradient domain high dynamic range compression”. In: *ACM Transactions on Graphics (TOG)*. Vol. 21. 3. ACM. 2002, pp. 249–256.
- [24] Sira Ferradans et al. “An Analysis of Visual Adaptation and Contrast Perception for Tone Mapping”. In: *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE* 33.10 (October 2011), pp. 120–140.
- [25] Kaiming He, Jian Sun, and Xiaoou Tang. “Guided image filtering”. In: *Computer Vision—ECCV 2010*. Springer, 2010, pp. 1–14.
- [26] Toshiyuki Dobashi Hitoshi Kiya. “An Efficient Unified-Tone-Mapping Operation for HDR Images with Various Formats”. In: *ITE Transactions on Media Technology and Applications (MTA)* 4.1 (2016), pp. 2–9.
- [27] Jason Huang and Valmiki Rampersad and Steve Mann. “High Dynamic Range Tone Mapping Based On Per-Pixel Exposure Mapping”. In: (2013).
- [28] G. Jeon J. Lee and J. Jeong. “Piecewise tone reproduction for high dynamic range imaging”. In: *IEEE Transactions on Consumer Electronics*. Vol. 55. IEEE, 2009, pp. 911–918.
- [29] Okuda Jinno and Adami. “New local tone mapping and two-layer coding for hdr images”. In: *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2012, pp. 765–768.

- [30] Tapas Kanungo et al. “An efficient k-means clustering algorithm: Analysis and implementation”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 24.7 (2002), pp. 881–892.
- [31] Kyungman Kim, Jonghyun Bae, and Jaeseok Kim. “Natural HDR image tone mapping based on retinex”. In: *Consumer Electronics, IEEE Transactions on* 57.4 (2011), pp. 1807–1814.
- [32] Peter Kovesi. “Phase Preserving Tone Mapping of Non-Photographic High Dynamic Range Images”. In: *Digital Image Computing Techniques and Applications (DICTA), 2012 International Conference on*. IEEE. 2012, pp. 1–8.
- [33] Alper Koz and Frederic Dufaux. “Optimized tone mapping with flickering constraint for backward-compatible high dynamic range video coding”. In: *Image Analysis for Multimedia Interactive Services (WIAMIS), 2013 14th International Workshop on*. IEEE. 2013, pp. 1–4.
- [34] Dipl-Ing Grzegorz Krawczyk, Dipl-Ing FH Daniel Brosch, et al. “HDR Tone Mapping”. In: *High-Dynamic-Range (HDR) Vision*. Springer, 2007, pp. 147–178.
- [35] Edwin H Land and John J McCann. “Lightness and retinex theory”. In: *JOSA* 61.1 (1971), pp. 1–11.
- [36] Rushmeier Holly Larson Gregory Ward and PiatkoChristine. “A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes”. In: vol. 3. 4. IEEE Educational Activities Department, 1997, pp. 291–306.
- [37] Gregory Ward Larson, Holly Rushmeier, and Christine Piatko. “A visibility matching tone reproduction operator for high dynamic range scenes”. In: *Visualization and Computer Graphics, IEEE Transactions on* 3.4 (1997), pp. 291–306.
- [38] Paul Lauga et al. “Region-based tone mapping for efficient High Dynamic Range video coding”. In: *4th European Workshop on Visual Information Processing (EUVIP)*. 2013, p. 208.

- [39] Ji Won Lee, Rae-Hong Park, and Soonkeun Chang. “Local tone mapping using the K-means algorithm and automatic gamma setting”. In: *Consumer Electronics, IEEE Transactions* 57.1 (2011), pp. 209–217.
- [40] Jing Li et al. “Novel Real-Time Tone-Mapping Operator for Noisy Logarithmic CMOS Image Sensors”. In: *Journal of Imaging Science and Technology* 60.2 (2016), pp. 20404–1.
- [41] Renjie Li et al. “Enhancing the contrast sensitivity function through action video game training”. In: *Nature neuroscience* 12.5 (2009), p. 549.
- [42] ZG Li, JH Zheng, and S Rahardja. “Region-of-interest based local tone mappings”. In: *Industrial Electronics and Applications (ICIEA), 2011 6th IEEE Conference on*. IEEE. 2011, pp. 1553–1559.
- [43] Boaz Livny. *mental ray for Maya, 3ds Max, and XSI: A 3D Artist’s Guide to Rendering*. Indianapolis, Indiana: Wiley Publishing, 2008. ISBN: 9780470008546.
- [44] Zicong Mai et al. “Optimizing a tone curve for backward-compatible high dynamic range image and video compression”. In: *Image Processing, IEEE Transactions on* 20.6 (2011), pp. 1558–1571.
- [45] Zicong Mai et al. “Subjective evaluation of tone-mapping methods on 3D images”. In: *Digital Signal Processing (DSP), 2011 17th International Conference on*. IEEE. 2011, pp. 1–6.
- [46] R. Mantiuk, A. Tomaszewska, and W. Heidrich. “Color correction for tone mapping”. In: vol. 28. 2. *Computer Graphics Forum*, 2009, pp. 193–202.
- [47] Rafał Mantiuk, Scott Daly, and Louis Kerofsky. “Display adaptive tone mapping”. In: *ACM Transactions on Graphics (TOG)* 27.3 (2008), p. 68.
- [48] Martin et al. “Evaluation of HDR tone mapping methods using essential perceptual attributes”. In: *Computers & Graphics* 32.3 (2008), pp. 330–349.
- [49] Yujie Mei, Guoping Qiu, and Kin-Man Lam. “Saliency modulated high dynamic range image tone mapping”. In: *Image and Graphics (ICIG), 2011 Sixth International Conference on*. IEEE. 2011, pp. 22–27.

- [50] Daly S. Meylan L. and Süsstrunk S. “Tone mapping for high dynamic range displays”. In: *Human Vision and Electronic Imaging XII (February 12, 2007)*. Vol. 6492. SPIE, 2007, p. 12.
- [51] L. Meylan and S. Susstrunk. “High dynamic range image rendering with a retinex-based adaptive filter”. In: *IEEE Transactions on Image Processing*. Vol. 15. IEEE, 2006, pp. 2820–2830.
- [52] JAY GOYAL MICHAEL ASHIKHMIN. “A Reality Check for Tone-Mapping Operators”. In: *ACM Transactions on Applied Perception* 3.4 (2016), pp. 399–411.
- [53] Ajit Motra and Herbert Thoma. “An adaptive LogLuv transform for high dynamic range video compression”. In: *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE. 2010, pp. 2061–2064.
- [54] Tae-Hyun Oh et al. “Robust high dynamic range imaging by rank minimization”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 37.6 (2015), pp. 1219–1232.
- [55] Cagri Ozcinar et al. “HDR Video Coding based on a Temporally Constrained Tone Mapping Operator”. In: Télécom ParisTech, Université Paris-Saclay.
- [56] Sumanta N. Pattanaik et al. “Time-Dependent Visual Adaptation for Fast Realistic Image Display”. In: Program of Computer Graphics, Cornell University, 2000.
- [57] YongQing HUO QiCong PENG. “Tone mapping for single-shot HDR imaging”. In: *Schoole of Communication and Information Engineering, 140 Lab, University of Electronic Science and Technology of China, Chengdu Sichuan 611731*. IEEE. 2011.
- [58] Chi-Sun Tang Pin-Hung Kuo and Shao-Yi Chien. “Content-adaptive inverse tone mapping”. In: IEEE, 2012, pp. 1–6.

- [59] Tung Li Qian, Suhaidi Shafie, and M Iqbal Sariipan. “A local tone mapping operator for high dynamic range images”. In: *Modeling, Simulation and Applied Optimization (ICMSAO), 2011 4th International Conference on*. IEEE. 2011, pp. 1–6.
- [60] Guoping Qiu and Jiang Duan. “Fast Tone Mapping for High Dynamic Range Images”. In: *ICPR2004, 17th International Conference on Pattern Recognition*. 2004, p. 26.
- [61] Erik Reinhard. “Parameter estimation for photographic tone reproduction”. In: *Journal of graphics tools* 7.1 (2002), pp. 45–51.
- [62] Erik Reinhard and Kate Devlin. “Dynamic range reduction inspired by photoreceptor physiology”. In: *Visualization and Computer Graphics, IEEE Transactions on* 11.1 (2005), pp. 13–24.
- [63] Erik Reinhard et al. “Photographic Tone Reproduction for Digital Images”. In: vol. 21. 3. ACM, 2002, pp. 267–276.
- [64] Erik Reinhard et al. “Photographic tone reproduction for digital images”. In: *ACM Transactions on Graphics (TOG)*. Vol. 21. 3. ACM. 2002, pp. 267–276.
- [65] Allan G. Rempel et al. “Ldr2Hdr: on-the-fly reverse tone mapping of legacy video and photographs”. In: *In SIGGRAPH '07: ACM SIGGRAPH 2007 papers*. ACM Press, 2007.
- [66] John C Russ. *The image processing handbook*. CRC press, 1998.
- [67] Yasir Salih et al. “Tone mapping of HDR images: A review”. In: *Intelligent and Advanced Systems (ICIAS), 2012 4th International Conference on*. Vol. 1. IEEE. 2012, pp. 368–373.
- [68] Christophe Schlick. “Quantization Techniques for Visualization of High Dynamic Range Pictures”. In: Springer-Verlag, 1994, pp. 7–20.
- [69] Helge Seetzen et al. “High dynamic range display systems”. In: *ACM Transactions on Graphics (TOG)*. Vol. 23. 3. ACM. 2004, pp. 760–768.

- [70] H. Seetzen et al. “High dynamic range display systems”. In: vol. 23. 3. ACM Transactions on Graphics (TOG), 2004, pp. 760–768.
- [71] Kaleigh Smith et al. “Beyond tone mapping: Enhanced depiction of tone mapped HDR images”. In: *Computer Graphics Forum*. Vol. 25. 3. Wiley Online Library. 2006, pp. 427–438.
- [72] Mingli Song et al. “Probabilistic Exposure Fusion”. In: vol. 21. 1. IEEE Press, 2012, pp. 341–357.
- [73] JC Stevens and Stanley S Stevens. “Brightness Function: Effects of Adaptation*”. In: *JOSA* 53.3 (1963), pp. 375–385.
- [74] J. Janet S.Venkata Lakshmi Sujatha and T. Bellarmine. “Analysis of tone mapping operators on high dynamic range images”. In: IEEE, 2012, pp. 1–6.
- [75] Duan Tian and Qui. “GPU-accelerated local tone-mapping for high dynamic range images”. In: *2012 19th IEEE International Conference on Image Processing*. IEEE. 2012, pp. 377–380.
- [76] Brian A Wandell. *Foundations of vision*. Sinauer Associates, 1995.
- [77] Ward and Simmons. “A backwards-compatible, high dynamic range extension to jpeg”. In: *Proceedings of the Thirteenth Color Imaging Conference*. 2005.
- [78] Hojatollah Yeganeh and Zhou Wang. “High dynamic range image tone mapping by maximizing a structural fidelity measure”. In: *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE. 2013, pp. 1879–1883.
- [79] Jun Zhang, Shiqiang Hu, and Xia Dai. “Direct high dynamic range imaging method using experiential optimal exposure criterion”. In: *Image and Signal Processing (CISP), 2011 4th International Congress on*. Vol. 2. IEEE. 2011, pp. 745–749.
- [80] Karel Zuiderveld. “Contrast limited adaptive histogram equalization”. In: *Graphics gems IV*. Academic Press Professional, Inc. 1994, pp. 474–485.

Non-exclusive licence to reproduce thesis and make thesis public

I, Aygül Salahlı(date of birth: 15th of April 1992),

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

1.1 reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

1.2 make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

"Comprehensive Study on High Dynamic Range Tone Mapping with Subjective Tests"

supervised by Assoc. Prof. Gholamreza Anbarjafari and Dr. Cagri Ozcinar

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu 17.05.2017