Inverse Tone Mapping of Standard Dynamic Range Content for High Dynamic Range

Applications

by

Pedram Mohammadi

B.A.Sc., Ferdowsi University of Mashhad, Iran
M.A.Sc., Ferdowsi University of Mashhad, Iran

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
Doctor of Philosophy
in
THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

February 2020

© Pedram Mohammadi, 2020
The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Inverse Tone Mapping of Standard Dynamic Range Content for High Dynamic Range Applications

submitted by Pedram Mohammadi in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering

Examinining Committee:

Panos Nasiopoulos
Supervisor

John D. Madden
Supervisory Committee Member

Edmond Cretu
University Examiner

Ian Mitchell
University Examiner
Abstract

High Dynamic Range (HDR) technology has revolutionized the field of digital media, affecting different aspects such as capturing, compression, transmission and display. By modeling the behavior of Human Visual System (HVS) when perceiving brightness and color, HDR technology offers a life-like viewing experience that is far superior to what Standard Dynamic Range (SDR) technology could achieve. While HDR technology has a disruptive impact in different fields, it also opens new revenue sources for SDR content owners and broadcasters that will continue producing real-time events in SDR format for the near future. For the latter case, SDR content need to be efficiently converted to HDR format, taking advantage of the superior visual quality of HDR displays.

Over the years, several attempts aimed at converting SDR content to HDR format, a process known as inverse Tone Mapping (iTM). The design of inverse Tone Mapping Operators (iTMOs) is considered a difficult task, as it tries to expand the brightness and color information to ranges not originally captured by SDR cameras.

In this thesis, we propose novel iTMOs that can effectively deal with all types of SDR content from dark, to normal and bright scenes, producing high visual quality HDR content. Our proposed methods work in the perceptual domain, which allows us to take advantage of the sensitivity of the human eye to brightness changes in different areas of the scene during the mapping process. To preserve the overall artistic impression, we developed methods that divide the SDR frame into dark, normal (average brightness), and bright regions, allowing us to keep intact dark and bright areas, without darkening or brightening up the frame.

We also address the issue of the color shift in SDR to HDR mapping by proposing a perception-based color adjustment method that preserves the hue of colors with insignificant
changes in brightness, producing HDR colors that are faithful to their SDR counterparts. Subjective and objective evaluations have shown that our proposed iTMOS outperform the state-of-the-art methods in terms of overall visual quality of the generated HDR video and generating HDR colors that closely follow their SDR counterparts.
Lay Summary

In this thesis, we propose novel iTMOs that are capable of dealing with all types of SDR frames from dark, to normal and bright frames generating high visual quality HDR videos. Our proposed methods work in the perceptual domain to model the sensitivity of the human eye to brightness changes in different regions of the scene. To maintain the overall visual impression of the input SDR frame, our methods divide the input SDR frame into three regions in terms of their brightness namely dark, normal, and bright. Furthermore, we achieve a good balance between the overall contrast and brightness of the HDR frame by optimizing our mapping process. We also propose a perception-based color adjustment method that maintains the color accuracy between input SDR and generated HDR frames and preserves the hue of colors as we move from SDR to HDR while resulting in negligible luminance change.
Preface

This thesis presents research conducted by Pedram Mohammadi under the guidance of Dr. Panos Nasiopoulos. A list of publications resulting from the work presented in this thesis is provided below.

The main body of Chapter 2 is taken from the publication in [P1]. The content of Chapter 3 appears in the conference paper [P2] and the submitted journal paper of [P3]. The content of Chapter 4 appears in the conference paper [P4] and the submitted journal paper [P5].

The work presented in [P1] – [P5] was performed by Pedram Mohammadi, including literature review, designing and implementing the proposed methods, performing all evaluations, analyzing the results and writing the manuscripts. Dr. Panos Nasiopoulos, Dr. Mahsa T. Pourazad, and Dr. Maryam Azimi provided guidance and editorial input.

The subjective studies in this work were covered under the UBC Ethics Board (H12-00308).

This thesis was written by Pedram Mohammadi, with editing assistance from Dr. Panos Nasiopoulos.

List of Publications Based on Work Presented in This Thesis


Table of Contents

Abstract .................................................................................................................. iii

Lay Summary ........................................................................................................ vi

Preface ................................................................................................................... vii

Table of Contents .................................................................................................. viii

List of Tables ......................................................................................................... xi

List of Figures ....................................................................................................... xiv

List of Abbreviations ............................................................................................ xxi

Acknowledgements ............................................................................................... xxii

Dedication ............................................................................................................... xxiii

Chapter 1: Introduction and Overview .................................................................... 1

1.1 Overview of state-of-the-art iTMOs ................................................................. 3

1.1.1 Akyuz et al.’s Operator .............................................................................. 4

1.1.2 Meylan et al.’s Operator ........................................................................... 6

1.1.3 Banterle et al.’s Operator .......................................................................... 7

1.1.4 Huo et al.’s Operator ................................................................................ 9

1.1.5 Rempel et al.’s Operator .......................................................................... 11

1.1.6 Kovaleski et al.’s Operator ....................................................................... 12

1.1.7 Bist et al.’s Operator ................................................................................ 13

1.1.8 Luzardo et al.’s Operator ........................................................................ 15

1.1.9 Endo et al.’s Operator .............................................................................. 17

1.1.10 Kim et al.’s Operator .............................................................................. 18

1.2 Thesis Outline .................................................................................................. 19
Chapter 2: An Entropy-based inverse Tone Mapping Operator ................................................. 22

2.1 Proposed iTMO ......................................................................................................................... 22

2.1.1 Converting SDR Light Values to Perceptual Values ......................................................... 22

2.1.2 Segmenting the SDR PQ Values to Three Brightness Regions ........................................ 26

2.1.3 Mapping SDR PQ Values to HDR Domain ......................................................................... 31

2.2 Experimental Evaluations ...................................................................................................... 35

2.3 Conclusions .............................................................................................................................. 37

Chapter 3: A Perception-based inverse Tone Mapping Operator for High Dynamic Range Video Applications .................................................................................................................. 38

3.1 Proposed iTMO ......................................................................................................................... 38

3.1.1 Converting SDR Light Values to Perceptual (PQ) Domain, Computing the Histogram and Segmentation to Three Brightness Regions ......................................................... 39

3.1.2 Mapping SDR PQ Values to HDR Domain ......................................................................... 43

3.2 Experimental Evaluations ...................................................................................................... 53

3.3 Conclusions .............................................................................................................................. 69

Chapter 4: A Fully-Automatic Content Adaptive Inverse Tone Mapping Operator with Improved Color Accuracy .......................................................................................................................... 70

4.1 Proposed iTMO ......................................................................................................................... 71

4.1.1 Converting SDR Light Values to PQ Domain and Computing Histogram ......................... 72

4.1.2 Segmenting the SDR PQ Values to Three Brightness Regions ......................................... 72

4.1.3 Mapping SDR PQ Values to HDR Domain ......................................................................... 78

4.1.4 Proposed Color Adjustment Method .................................................................................. 81

4.2 Experimental Evaluations ...................................................................................................... 87
Chapter 5: Conclusions and Future Work

5.1 Summary of the Contributions

5.2 Significance and Potential Applications of the Research

5.3 Future Work

Bibliography
List of Tables

Table 3.1 Visual quality scores by HDR VDP-2.2 for dark frames in Figure 3.10 for three different values of $w_1$ and $w_2$ .......................................................................................................................... 54

Table 3.2 Visual quality scores by HDR VDP-2.2 for normal frames in Figure 3.10 for three different values of $w_1$ and $w_2$ .......................................................................................................................... 54

Table 3.3 Visual quality scores by HDR VDP-2.2 for bright frames in Figure 3.10 for three different values of $w_1$ and $w_2$ .......................................................................................................................... 54

Table 3.4 Visual quality scores by PU-SSIM for dark frames in Figure 3.10 for three different values of $w_1$ and $w_2$ .......................................................................................................................... 54

Table 3.5 Visual quality scores by PU-SSIM for normal frames in Figure 3.10 for three different values of $w_1$ and $w_2$ .......................................................................................................................... 55

Table 3.6 Visual quality scores by PU-SSIM for bright frames in Figure 3.10 for three different values of $w_1$ and $w_2$ .......................................................................................................................... 55

Table 3.7 Mean and standard deviation of MOS scores for our contrast, brightness and hybrid approaches using the dataset in Figure 3.13 ............................................................................................................. 59

Table 3.8 Visual quality scores by HDR VDP-2.2 for dark frames in Figure 3.10 using hybrid, brightness and contrast approaches .......................................................................................................................... 61

Table 3.9 Visual quality scores by HDR VDP-2.2 for normal frames in Figure 3.10 using hybrid, brightness and contrast approaches .......................................................................................................................... 61

Table 3.10 Visual quality scores by HDR VDP-2.2 for bright frames in Figure 3.10 using hybrid, brightness and contrast approaches .......................................................................................................................... 61

Table 3.11 Visual quality scores by PU-SSIM for dark frames in Figure 3.10 using hybrid, brightness and contrast approaches .......................................................................................................................... 62
Table 3.12 Visual quality scores by PU-SSIM for normal frames in Figure 3.10 using hybrid, brightness and contrast approaches.............................................................. 62
Table 3.13 Visual quality scores by PU-SSIM for bright frames in Figure 3.10 using hybrid, brightness and contrast approaches.............................................................. 62
Table 3.14 Mean and standard deviation of MOS scores for the six iTMOs based on the dataset in Figure 3.16 ................................................................. 64
Table 3.15 Visual quality scores by HDR VDP-2.2 for dark frames in Figure 3.19 for six different iTMOs.............................................................................. 67
Table 3.16 Visual quality scores by HDR VDP-2.2 for normal frames in Figure 3.19 for six different iTMOs .............................................................................. 67
Table 3.17 Visual quality scores by HDR VDP-2.2 for bright frames in Figure 3.19 for six different iTMOs .............................................................................. 67
Table 3.18 Visual quality scores by PU-SSIM for dark frames in Figure 3.19 for six different iTMOs.............................................................................. 68
Table 3.19 Visual quality scores by PU-SSIM for normal frames in Figure 3.19 for six different iTMOs.............................................................................. 68
Table 3.20 Visual quality scores by PU-SSIM for bright frames in Figure 3.19 for six different iTMOs.............................................................................. 68
Table 3.21 Execution times of six iTMOs (in seconds) for a 4K SDR frame ....................... 68
Table 4.1 Mean, SD, and 95% CI of the average $\Delta L$ for different values of the saturation factor in our proposed color adjustment method........................................... 87
Table 4.2 Mean, SD, and 95% CI of the average $\Delta L$ using our proposed iTMO with three different color adjustment methods ............................................... 88
Table 4.3 Mean, SD, and 95% CI of the average $\Delta L$ using Kim et al.’s iTMO with three different color adjustment methods ................................................................. 89

Table 4.4 Mean and SD of MOS scores for the six iTMOs based on the dataset in Figure 4.17. 93

Table 4.5 Visual quality scores by HDR VDP-2.2 for dark frames in Figure 4.20 for six different iTMOs ........................................................................................................................................................................ 96

Table 4.6 Visual quality scores by HDR VDP-2.2 for normal frames in Figure 4.20 for six different iTMOs ........................................................................................................................................................................ 96

Table 4.7 Visual quality scores by HDR VDP-2.2 for bright frames in Figure 4.20 for six different iTMOs ........................................................................................................................................................................ 96

Table 4.8 Visual quality scores by PU-SSIM for dark frames in Figure 4.20 for six different iTMOs ........................................................................................................................................................................ 97

Table 4.9 Visual quality scores by PU-SSIM for normal frames in Figure 4.20 for six different iTMOs ........................................................................................................................................................................ 97

Table 4.10 Visual quality scores by PU-SSIM for bright frames in Figure 4.20 for six different iTMOs ........................................................................................................................................................................ 97

Table 4.11 Mean and SD of MOS scores for the six iTMOs based on the dataset in Figure 4.21 ........................................................................................................................................................................ 100
List of Figures

Figure 1.1 (a) Overall brightness range seen by human eye through adaptation. (b) Brightness range seen by human eye at a certain adaptation level. (c) Brightness range supported by SDR technology. (d) Brightness range supported by HDR technology. ......................................................... 1

Figure 1.2 Mapping curves used by Akyuz et al.’s operator for three different $\gamma$ values and $k = 1000$ nits. .................................................................................................................................................. 5

Figure 1.3 (a) An example of a dark SDR frame with (b) its HDR frame generated by Akyuz et al.’s iTMO with $\gamma = 1$, and $k = 1000$ nits. ........................................................................................................ 5

Figure 1.4 Mapping curve used by Meylan et al.’s operator for $Disp_{\text{max}} = 1000$ nits, $\rho = 660$ nits, and $\omega = 30$ nits. ............................................................................................................................................ 7

Figure 1.5 (a) An example of a dark SDR frame along with (b) its HDR frame generated by Meylan et al.’s iTMO with $Disp_{\text{max}} = 1000$ nits and $\rho = 660$ nits. ........................................................................................................ 7

Figure 1.6 Mapping curves used by Banterle et al.’s operator for $Disp_{\text{max}} = 1000$ nits, and three different values for $L_{\text{white}}$. .............................................................................................................................................. 8

Figure 1.7 (a) An example of an SDR frame along with (b) its HDR frame generated by Banterle et al.’s iTMO with $Disp_{\text{max}} = 1000$ nits and $L_{\text{white}} = 10$. ........................................................................................................... 9

Figure 1.8 Mapping curves used by Huo et al.’s operator for $\sigma = 1000$ nits, $L_s = 0$, and $n = 0.86$. ............................................................................................................................................. 10

Figure 1.9 (a) An example of an SDR frame along with (b) its HDR frame generated by Huo et al.’s iTMO with $\sigma = 1000$ nits. .................................................................................................................. 11

Figure 1.10 Mapping curve used by Rempel et al.’s operator. ................................................. 11

Figure 1.11 (a) An example of a bright SDR frame along with (b) its HDR frame generated by Rempel et al.’s iTMO, and (c) HDR frame generated by Kovaleski et al.’s iTMO. ................. 12
Figure 1.12 Mapping curves used by Bist et al.’s operator for three different $\gamma$ values. .......... 14

Figure 1.13 (a) An example of a dark SDR frame along with (b) its HDR frame generated by Bist et al.’s iTMO. ..................................................................................................................... 14

Figure 1.14 Mapping curve used by Luzardo et al.’s operator for $Disp_{\text{max}} = 1000$ nits, $a = 1.25$, $d = 4$, $m_{\text{SDR}} = 21.47$ nits, $m_{\text{HDR}} = 27.60$ nits, $b = -4.2758$, and $c = 5.2758$. ...................................................... 16

Figure 1.15 (a) An example of a dark SDR frame along with (b) its HDR frame generated by Luzardo et al.’s iTMO with $Disp_{\text{max}} = 1000$ nits, $a = 1.25$, $d = 4$, and $m_{\text{SDR}} = 21.47$ nits. .......... 17

Figure 1.16 (a) An SDR frame along with (b) its HDR frame generated by Endo et al.’s iTMO. 17

Figure 1.17 Mapping curve used by Kim et al.’s operator for $a = 5$, $\alpha = 0.4$, and $\beta = 0.6$. ........... 18

Figure 1.18 (a) An example of a dark SDR frame along with (b) its HDR frame generated by Kim et al.’s iTMO with for $a = 5$, $\alpha = 0.4$, and $\beta = 0.6$. ..................................................................................... 19

Figure 2.1 Difference in how our eyes perceive brightness information and how cameras capture that information........................................................................................................................................ 23

Figure 2.2 Luminance values encoded by the PQ and gamma functions. ......................... 24

Figure 2.3 (a) An SDR frame along with (b) the histogram of its luminance channel in PQ domain. ........................................................................................................................................ 25

Figure 2.4 An example of histogram of SDR PQ values. Our proposed segmentation method aims at finding $X_1$ and $X_2$ to divide the PQ values into three brightness regions. ......................... 27

Figure 2.5 (a) Plots of $E_1$, $E_2$, and $E_1 + E_2$. $X_1$ is where $E_1 + E_2$ is maximum. (b) Plots of $E_3$, $E_4$, and $E_3 + E_4$. $X_2$ is where $E_3 + E_4$ is maximum. (c) Result of segmentation using our proposed segmentation method. Black, gray, and white areas represent dark, normal, and bright regions in the SDR frame, respectively. ........................................................................................................ 29
Figure 2.6 (a) – (c) An example of dark, normal, and bright SDR frames, respectively. (d) – (f) Segmentation results using our proposed method for each frame, respectively. Black, gray, and white portions demonstrate dark, normal, and bright regions in the SDR frame, respectively. ...

Figure 2.7 Changes in $X_1$ and $X_2$ using our proposed segmentation method. (a) and (b) show the changes in $X_1$ and $X_2$, respectively.

Figure 2.8 An example of mapping curve used in our proposed iTMO.

Figure 2.9 An example of the two curves associated with slope limits for dark and bright regions. These two curves demonstrate the limits for which we can choose our mapping curves.

Figure 2.10 Preference scores of the subjects in mean opinion score (MOS) when comparing the visual quality generated by our iTMO over that of Meylan et al.’s, and Banterle et al.’s operators.

Figure 3.1 Workflow of our proposed iTMO. The red dotted line encloses components of our approach.

Figure 3.2 (a) An SDR frame along with (b) the histogram of its luminance channel in PQ domain.

Figure 3.3 (a) PQ curve for mapping SDR light values to PQ domain. (b) Rate of change (derivative) of the PQ curve. (c) and (d) Zoomed areas corresponding to dark and bright regions, respectively.

Figure 3.4 (a) - (c) An example of dark, normal and bright SDR frames, respectively. (d) - (f) represent the result of our segmentation method for each frame, respectively. Black, gray, and white portions demonstrate dark, normal, and bright regions in the input SDR frame, respectively.
Figure 3.5 An example of mapping curve used in our iTMO. $X_1$ is the boundary between dark and normal regions, and $X_2$ is the boundary between normal and bright regions. ........................................... 44

Figure 3.6 An example of an overall dark SDR frame with (b) its corresponding HDR frame with our contrast approach. ................................................................. 47

Figure 3.7 An example of an overall bright SDR frame with (b) its corresponding HDR frame with our contrast approach. ................................................................. 47

Figure 3.8 (a) An example of an overall dark SDR frame with (b) its corresponding HDR frame generated using our contrast approach. (c) Corresponding HDR frame generated using our brightness approach. ................................................................. 51

Figure 3.9 (a) An example of an overall bright SDR frame with (b) its corresponding HDR frame generated using our contrast approach. (c) Corresponding HDR frame generated using our brightness approach. ................................................................. 52

Figure 3.10 Nine original tone mapped HDR frames randomly chosen from the dataset in [58]. 53

Figure 3.11 (a) An example of an overall dark SDR frame with (b) its corresponding HDR frame with our contrast approach. (c) Corresponding HDR frame with our brightness approach, and (d) corresponding HDR frame with our hybrid approach with $w_1 = w_2 = 0.5$. .............................. 56

Figure 3.12 (a) An example of an overall bright SDR frame with (b) its corresponding HDR frame with our contrast approach. (c) Corresponding HDR frame with our brightness approach, and (d) corresponding HDR frame with our hybrid approach with $w_1 = w_2 = 0.5$. .............................. 57

Figure 3.13 One frame of each of the nine 4K SDR videos used in our subjective test. ............. 58

Figure 3.14 Preference scores of the subjects in mean opinion score (MOS) when comparing the visual quality of the generated HDR videos using our hybrid approach with $w_1 = w_2 = 0.5$ over that of brightness and contrast approaches separately. ......................................................... 59

xvii
Figure 3.15 Post-hoc analysis using Tukey’s method for mean values of MOS scores and $\alpha=0.05$. Methods within the same group generate statistically insignificant HDR videos. .......................... 60
Figure 3.16 One frame of each of the nine 4K SDR videos used in our subjective test. ............... 63
Figure 3.17 Mean opinion scores (MOS) indicating the visual quality of the HDR videos generated by our iTMO against Meylan et al.’s, Rempel et al.’s, Kovaleski et al.’s, Bist et al.’s, and Kim et al.’s operator ........................................................................................................................................... 64
Figure 3.18 Post-hoc analysis using Tukey’s method for mean values of MOS scores and $\alpha=0.05$. Methods within the same group generate statistically insignificant HDR videos. .................. 65
Figure 3.19 Nine original tone mapped HDR frames randomly chosen from the dataset in [58]. 66
Figure 4.1 The block diagram of our proposed iTMO. The main components of our method are enclosed in the red dotted line. The generated HDR frame in this diagram has maximum brightness of 1000 nits. However, in general, our method is capable of generating HDR videos with arbitrary maximum brightness. ......................................................................................................................... 71
Figure 4.2 (a) An SDR frame along with (b) the histogram of its luminance channel in PQ domain. ..................................................................................................................................................... 73
Figure 4.3 An example of histogram of SDR PQ values. Our proposed segmentation method aims at finding $X_1$ and $X_2$ to divide the SDR PQ values to three brightness regions. ..................... 75
Figure 4.4 Plots of $E_1$, $E_2$, and $E_1+E_2$. $X_1$ is where $E_1+E_2$ is maximum. (b) Plots of $E_3$, $E_4$, and $E_3+E_4$. $X_2$ is where $E_3+E_4$ is maximum. (c) Original SDR frame in Figure 4.2 (a). (d) Segmentation result using our proposed method. Black, gray, and white areas represent dark, normal, and bright regions in the SDR frame, respectively. ............................................................................................................. 76
Figure 4.5 (a) - (c) An example of dark, normal and bright SDR frames, respectively. (d) - (f) represent the result of our proposed segmentation method for each frame, respectively. .......... 78
Figure 4.6 (a) An SDR frame. (b) Our mapping curve associated with it. \( X_1 \) and \( X_2 \) are the thresholds found using our segmentation method and \( Y_1 \) and \( Y_2 \) are their corresponding PQ values in HDR domain.

Figure 4.7 Performance of our proposed iTMO for (a) dark, (b) normal, and (c) bright SDR frames. (d), (e), and (f) show the tone mapped HDR frames generated using our proposed iTMO, respectively.

Figure 4.8 Changes in hue \((h)\) and Lightness \((L)\) with respect to Chroma \((C)\) for six basic colors as we change the saturation factor \(s\) from zero to two for Schlick et al.’s method. (a) and (b) demonstrate the changes in \(h\) and \(L\) with respect to \(C\) for six basic colors, respectively. The points where \(s = 1\) are marked with black squares.

Figure 4.9 Changes in hue \((h)\) and Lightness \((L)\) with respect to Chroma \((C)\) for six basic colors as we change the saturation factor \(s\) from zero to two for Mantiuk et al.’s method. (a) and (b) demonstrate the changes in \(h\) and \(L\) with respect to \(C\) for six basic colors, respectively. The points where \(s = 1\) are marked with black squares.

Figure 4.10 Changes in hue \((h)\) and Lightness \((L)\) with respect to Chroma \((C)\) for six basic colors as we change the saturation factor \(s\) from zero to two for our proposed color adjustment method. (a) and (b) demonstrate the changes in \(h\) and \(L\) with respect to \(C\) for six basic colors, respectively. The points where \(s = 1\) are marked with black squares.

Figure 4.11 A subset of our 400 SDR frames dataset randomly taken from [49].

Figure 4.12 Average of \(\Delta L\) along with its 95% CI over our dataset of 400 SDR frames for three different values of \(s\).

Figure 4.13 Average of \(\Delta L\) along with its 95% CI over our dataset of 400 SDR frames using our proposed iTMO for three different color adjustment methods.
Figure 4.14 Average of $\Delta L$ along with its 95% CI over our dataset of 400 SDR frames using Kim et al.’s iTMO for three different color adjustment methods. ............................................................... 89
Figure 4.15 An example of our mapping curve. $X_1$, and $X_2$ are the boundaries for dark, and bright regions, respectively while $Y_1$, and $Y_2$ are their corresponding PQ values in HDR domain, respectively. ................................................................................................................. 89
Figure 4.16 (a) and (b) Absolute difference for $Y_1$ and $Y_2$ generated by our proposed iTMO using the segmentation method in Chapter 2, respectively, compared to (c) and (d) the proposed segmentation method in this Chapter................................................................................................. 90
Figure 4.17 One frame of each of the nine 4K SDR videos used in our subjective test................. 91
Figure 4.18 Preference scores of the subjects in mean opinion score (MOS) when comparing the visual quality of the HDR videos generated by our iTMO over that of Endo et al.’s, Luzardo et al.’s, Bist et al.’s, iTMO in Chapter 3, and Kim et al.’s iTMO. ................................................................. 93
Figure 4.19 Post-hoc analysis using Tukey’s method for mean values of MOS scores and $\alpha=0.05$. Methods within the same group generate statistically insignificant HDR videos. ................. 94
Figure 4.20 Nine original tone mapped HDR frames randomly chosen from the dataset in [58].95
Figure 4.21 Ten SDR frames randomly chosen from [48] for our subjective test....................... 98
Figure 4.22 Preference scores of the subjects in mean opinion score (MOS) when comparing the color accuracy of output HDR frame generated by our proposed iTMO over that of Endo et al.’s, Luzardo et al.’s, Bist et al.’s, iTMO in Chapter 3, and Kim et al.’s operator......................... 99
Figure 4.23 Post-hoc analysis using Tukey’s method for mean values of MOS scores and $\alpha=0.05$. Methods within the same group generate statistically insignificant HDR videos. ................. 100
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>CIE</td>
<td>Commission Internationale de l'Eclairage</td>
</tr>
<tr>
<td>CNN</td>
<td>Convolutional Neural Network</td>
</tr>
<tr>
<td>HD</td>
<td>High Definition</td>
</tr>
<tr>
<td>HDR</td>
<td>High Dynamic Range</td>
</tr>
<tr>
<td>HVS</td>
<td>Human Visual System</td>
</tr>
<tr>
<td>iTM</td>
<td>inverse Tone Mapping</td>
</tr>
<tr>
<td>iTMO</td>
<td>inverse Tone Mapping Operator</td>
</tr>
<tr>
<td>ITU</td>
<td>International Telecommunication Union</td>
</tr>
<tr>
<td>JND</td>
<td>Just Noticeable Difference</td>
</tr>
<tr>
<td>LF</td>
<td>Light Field</td>
</tr>
<tr>
<td>MOS</td>
<td>Mean Opinion Score</td>
</tr>
<tr>
<td>PQ</td>
<td>Perceptual Quantizer</td>
</tr>
<tr>
<td>PU</td>
<td>Perceptually Uniform</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SDR</td>
<td>Standard Dynamic Range</td>
</tr>
<tr>
<td>SSIM</td>
<td>Structural Similarity Index</td>
</tr>
<tr>
<td>TMO</td>
<td>Tone Mapping Operator</td>
</tr>
<tr>
<td>VDP</td>
<td>Visual Difference Predictor</td>
</tr>
</tbody>
</table>
Acknowledgements

This thesis would never have been done without the help and support from numerous people. I would like to take this opportunity to express my gratitude to the following individuals.

Primarily, I would like to give my sincerest gratitude to my supervisor Dr. Panos Nasiopoulos for his continuous guidance and support throughout my Ph.D. I thank Panos for his encouragement, patience, enthusiasm, and broad knowledge in multimedia. I also thank Panos for the freedom and independence he gave me during my PhD studies. He has always been a great mentor, a role model, and a friend.

I would like to extend my sincere thanks to Dr. Mahsa T. Pourazad for her support and guidance during my PhD studies. I thank Mahsa for her patience, encouragement, and great inspiring suggestions. I was lucky to have the opportunities to collaborate with her throughout my PhD studies.

I would also like to thank my Ph.D. committee members, Dr. Victor Leung, Dr. Jane Wang, and Dr. John Madden for investing time and energy to provide guidance and help throughout my graduate studies. Their encouragement and insightful comments have been extremely valuable to me.

I also thank my other colleagues at the UBC Digital Multimedia Lab, Maryam Azimi, Stelios Ploumis, Ahmad Khaldieh, Hamid Reza Tohidypour, Fujun Xie, Anahita Shojaei, Ilya Ganelin, Nusrat Mehajabin, Doris Xiang, and Joseph Khoury. I thank you all for creating a vibrant and supportive research environment, for the stimulating discussions, for your friendship, and for the fun times we have had together.
Dedication

To my Mom, Mahboobeh, and my Dad, Bahman, whom without their unconditional love and support none of this would be happening.
Chapter 1: Introduction and Overview

The human eye, through adaptation, is capable of perceiving a wide range of brightness values from starlight to sunlight [1]. At a certain adaptation level, our eyes are capable of perceiving a brightness range of 5 orders of magnitude (i.e., the difference in powers of ten between highest and lowest brightness values) measured in units of candelas per square meter (cd/m$^2$) or nits (See Figure 1.1) [1]. However, due to its limitations, Standard Dynamic Range (SDR) technology can capture and display only a small portion (about two to three orders of brightness magnitude) of what our eyes can perceive. High Dynamic Range (HDR) technology aims at overcoming the limitations of its SDR counterpart by diligently modeling the capabilities

![Figure 1.1](image)

Figure 1.1 (a) Overall brightness range seen by human eye through adaptation. (b) Brightness range seen by human eye at a certain adaptation level. (c) Brightness range supported by SDR technology. (d) Brightness range supported by HDR technology.
of Human Visual System (HVS) in perceiving brightness. Doing so, HDR technology can capture and display a brightness range of up to 5 orders of magnitude.

While this emerging video technology has a disruptive impact on the capturing and display industries, it also opens new opportunities for SDR content owners and broadcasters such as Netflix and Hollywood studios that will continue producing real-time events in SDR format for the near future. For the latter case, SDR content will have to be efficiently converted to HDR format, taking advantage of the advanced visual quality of HDR displays to offer to viewers a much better quality of experience than what the original SDR content would ever do. Thus, it is of high importance to be able to generate HDR content by converting existing legacy SDR content to the new format, taking advantage of the higher visual quality that it offers.

Over the years, several attempts have focused on converting existing SDR content to HDR format, a process better known as inverse Tone Mapping (iTM) [2]–[27]. The design of inverse Tone Mapping Operators (iTMOs) has proven to be a challenging task as one tries to recreate the information that is lost during SDR capturing process [28], [29]. Moreover, the fact that most of the existing iTMOs only focus on specific types of SDR content either considering normal (average brightness) SDR content, or focusing on dark, and/or bright ones does not help in providing a general solution. One of the challenges in designing an efficient iTMO arises from the fact that as we try to expand the SDR brightness levels, we inevitably increase the chance of introducing and/or increasing noise in dark regions and banding in bright regions of the input SDR frame. Thus, keeping these visual artifacts at a non-perceivable level is an important consideration when designing iTMOs. Another important challenge in designing iTMOs arises from the fact that as we try to expand the SDR brightness levels, we introduce changes in colors between input SDR and generated HDR frames [30]–[33]. Consequently, an efficient color adjustment method capable
of maintaining the color accuracy between input SDR and generated HDR frame is another important consideration when designing iTMOS.

In this thesis, we propose novel iTMOS capable of dealing with all types of SDR content, from dark, to normal and bright scenes. Our proposed iTMOS keep the visual artifacts at a non-perceivable level and maintain the color accuracy between input SDR and generated HDR frame. In Chapter 2, we propose an entropy-based iTMO to convert SDR content to HDR format. In Chapter 3, we propose a hybrid iTMO for converting SDR content to HDR format capable of achieving the best possible trade-off between overall brightness and contrast of the generated HDR video. In Chapter 4, we propose a content adaptive iTMO that addresses the shortcomings of the methods in Chapters 2 and 3. In addition, we address the issue of color shifts when converting SDR content to HDR format and propose a novel perception-based color adjustment method that is capable of maintaining the color accuracy between input SDR and generated HDR frame.

The following sections in this chapter provide an in-depth look into state-of-the-art iTMOS (Section 1.1) and provide a summary of the scientific contributions of this thesis (Section 1.2).

1.1 Overview of state-of-the-art iTMOS

Over the years, several efforts focused on converting SDR content to HDR format. In general, an iTMO takes as input the luminance channel of the SDR frame and generates the HDR equivalent. Note that luminance is measured in units of candelas per square meter (cd/m\(^2\)) or else known as nits and can be computed based on International Telecommunication Union (ITU) BT.709 standardization \cite{34} using the following equation:

\[
L_{SDR} = 0.2126R_{SDR} + 0.7152G_{SDR} + 0.0722B_{SDR}
\]  \hspace{1cm} (1.1)

where \(L_{SDR}\) is the luminance channel of the SDR frame, and \(R_{SDR}\), \(G_{SDR}\), and \(B_{SDR}\) are the red, green, and blue channels of the SDR frame, respectively. In general, SDR luminance values are in the
range [0.1 – 100] nits. Note that the corresponding calculation of luminance for BT.2020 standard can be found in [35]. Throughout this thesis, we refer to brightness and luminance interchangeably. In the following sub-sections, we provide a comprehensive overview of some of the most prominent iTMOs.

1.1.1 Akyuz et al.’s Operator

One of the simplest iTMOs to implement is the method proposed by Akyuz et al. [2] which uses the following linear mapping curve:

\[ L_{HDR} = k \left( \frac{L_{SDR} - L_{SDR,min}}{L_{SDR,max} - L_{SDR,min}} \right)^\gamma \]  

(1.2)

where \( L_{HDR} \) and \( L_{SDR} \) are the HDR and SDR luminance channels, respectively, \( L_{SDR,min} \) and \( L_{SDR,max} \) are the minimum and maximum SDR luminance values, \( k \) is the maximum brightness of the target HDR display, and \( \gamma \) is a parameter responsible for the shape of the mapping curve that is chosen by the user. Figure 1.2 shows the mapping curves used by this iTMO for three different \( \gamma \) values, and \( k = 1000 \) nits. These curves map the SDR luminance values from [0.1 – 100] nits to [0.01 – 1000] nits in the HDR domain. After performing a series of subjective tests, the authors concluded that a simple linear mapping curve (\( \gamma = 1 \)) is sufficient to have good quality HDR content. This iTMO does a good job in converting normal and bright SDR frames to HDR format. However, it falls short when it comes to dark SDR frames and generates HDR frames that appear overly bright with visible visual artifacts. Figure 1.3 (a) shows a dark SDR frame randomly chosen from [36] while Figure 1.3 (b) shows the HDR frame generated by Akyuz et al.’s operator with \( \gamma \)
As it can be seen, since mapping is done using a simple line, the generated HDR frame is not of the highest visual quality and appears overly bright with visible visual artifacts in dark regions of the frame. Throughout this thesis, generated HDR frames with different iTMOs are tone mapped using Reinhard et al.’s [37] Tone Mapping Operator (TMO) for displaying purposes. Throughout this thesis, since there are no public databases containing SDR and HDR pairs of the same scene, we refrain from comparing generated HDR frames with ground truth.
1.1.2 Meylan et al.’s Operator

Meylan et al. [3], [14] proposed an iTMO with the goal of recovering the highlights in SDR frames. The main idea is to detect dark, and bright regions of the SDR frame and expand them using a piece-wise linear mapping curve with one breaking point. The mapping curve in this iTMO is presented by the following equation:

\[
L_{\text{HDR}} = \begin{cases} 
  s_1 L_{\text{SDR}} & \text{if } L_{\text{SDR}} \leq \omega \\
  s_1 \omega + s_2 (L_{\text{SDR}} - \omega) & \text{otherwise}
\end{cases}
\]

(1.3)

where \( L_{\text{HDR}} \) and \( L_{\text{SDR}} \) are the HDR and SDR luminance channels respectively, \( \omega \) is the boundary between dark, and bright regions, \( s_1 \), and \( s_2 \) are the slopes of the mapping curve associated with dark and bright regions, respectively. These slopes can be calculated as follows:

\[
s_1 = \frac{\rho}{\omega}
\]

(1.4)

\[
s_2 = \frac{\text{Disp}_{\text{max}} - \rho}{100 - \omega}
\]

(1.5)

where \( \text{Disp}_{\text{max}} \) is the maximum brightness of the target HDR display and \( \rho \) is the portion of the maximum brightness of the HDR display that will be allocated to dark regions. To determine a value for \( \rho \) that works well for all types of SDR frames, the authors conducted a series of comprehensive subjective tests and concluded that \( \rho = 0.66 \times \text{Disp}_{\text{max}} \) is a good estimate for generating high quality HDR content.

Figure 1.4 shows the mapping curve used in this iTMO for an HDR display with maximum brightness of 1000 nits (\( \text{Disp}_{\text{max}} = 1000 \) nits), \( \rho = 660 \) nits, and \( \omega = 30 \) nits. This curve maps the SDR luminance values from \([0.1 - 100]\) nits to \([0.01 - 1000]\) nits in the HDR domain. Figure 1.5 (a) shows a dark SDR frame while Figure 1.5 (b) shows the HDR frame generated by Meylan et al.’s operator with \( \text{Disp}_{\text{max}} = 1000 \) nits and \( \rho = 660 \) nits. We observe that although this mapping
does a good job in recovering the highlights, in the case of dark SDR frames the generated HDR frame is often noisy and appears overly bright compared to SDR input, resulting in reduced overall contrast.

1.1.3 Banterle et al.’s Operator

Banterle et al. proposed an iTMO to address the issue of loss of contrast by Akyuz et al.’s and Meylan et al.’s iTMOs [21]–[24]. The main idea behind this method is to choose a previously
proposed tone mapping curve and calculate its inverse curve. Authors chose Reinhard et al.’s tone
mapping curve [37] due to its simplicity. The mapping curve used in this iTMO is presented by
the following equation:

\[
L_{\text{HDR}} = \frac{1}{2} \text{Disp}_\text{max} L_{\text{white}} \left( L_{\text{SDR}} - 1 + \sqrt{\left(1 - L_{\text{SDR}}\right)^2 + \frac{4}{L_{\text{white}}^2} L_{\text{SDR}}} \right)
\]

(1.6)

where \( L_{\text{HDR}} \) and \( L_{\text{SDR}} \) are the HDR and SDR luminance channels, respectively, \( \text{Disp}_\text{max} \) is the
maximum brightness of the target HDR display, and \( L_{\text{white}} \) is a parameter responsible for the shape of the mapping curve and is chosen by the user. Based on comprehensive subjective testing, the authors concluded that if \( L_{\text{white}} \approx \text{Disp}_\text{max} \), the generated HDR frames would demonstrate good
contrast with no banding in bright regions of the frame. After expanding the SDR luminance channel, the bright regions of the frame are identified using the median cut algorithm [38] and are further expanded to increase the contrast in these regions. Figure 1.6 shows the mapping curves used in this iTMO for an HDR display with maximum brightness of 1000 nits (\( \text{Disp}_\text{max} = 1000 \))

Figure 1.6 Mapping curves used by Banterle et al.’s operator for \( \text{Disp}_\text{max} = 1000 \) nits, and three different
values of \( L_{\text{white}} \).
nits), and three different values of $L_{white}$. These curves map the SDR luminance values from [0.1 – 100] nits to [0.01 – 1000] nits in HDR domain. One major disadvantage of this iTMO is that due to the shape of the mapping curve, the dark areas of the SDR frame appear darker in the generated HDR frame, causing loss of visible details in those areas. Figure 1.7 (a) shows an example of an SDR frame and Figure 1.7 (b) depicts its generated HDR frame using Banterle et al.’s iTMO with $Disp_{max} = 1000$ nits and $L_{white} = 10$. We observe that the details in the dark areas of the frame such as the trees in the background are lost in the generated HDR frame.

1.1.4 Huo et al.’s Operator

To take advantage of the capabilities of HVS in perceiving brightness, Huo et al. [25], [26] proposed an iTMO that aims at modeling the eye retina response when viewing an SDR frame. After computing the retina response based on [39], the pixel-based luminance adaptation level of the SDR frame is calculated. Finally, the SDR frame and pixel-based luminance adaptation level are applied to the inverse of the eye retina response and the HDR frame is generated. The mapping curve used in this iTMO is computed using the following equation:
where $L_{HDR}$ and $L_{SDR}$ are the HDR and SDR luminance channels, respectively, $L_{SDR,max}$ is the maximum SDR luminance value, $L_s$ is the surrounding luminance at each pixel of the SDR frame, $\sigma$ is the adaptation level, and $n$ is a sensitivity control parameter [39]. Based on comprehensive subjective tests, authors concluded that $n = 0.86$ results in good quality HDR frames. Figure 1.8 shows the mapping curve used in this iTMO for an HDR display with maximum brightness of 1000 nits ($\sigma = 1000$ nits), surrounding luminance values of zero ($L_s = 0$), and $n = 0.86$. This curve maps the SDR luminance values from $[0.1–100]$ nits to $[0.01–1000]$ nits in HDR domain. Figure 1.9 (a) shows a dark SDR frame while Figure 1.9 (b) presents the HDR frame generated using Huo et al.’s operator with $\sigma = 1000$ nits. As it can be seen, this iTMO underperforms when it comes to dark SDR frames and results in HDR frames that appear darker with loss of visible details (see the brick walls).
1.1.5 Rempel et al.’s Operator

To address the need for real-time iTMOs, Rempel et al. [27] proposed a real-time, fully-automatic (no human intervention) iTMO capable of generating good quality HDR frames. In this iTMO a simple linear curve is chosen as the mapping curve. Afterwards, the bright regions of the SDR frame are identified using a simple fixed threshold and are further expanded to generate the HDR frame. Figure 1.10 depicts the mapping curve used in this iTMO. This curve maps the SDR luminance values to the HDR luminance values.
luminance values from \([0.1 \text{ – } 100]\) nits to \([0.01 \text{ – } 1000]\) nits in the HDR domain. Although normal regions are well represented in the HDR domain, this iTMO falls short in recovering details in dark and bright regions of the SDR frame. Figure 1.11 (a) shows a bright SDR frame while Figure 1.11 (b) depicts the generated HDR frame using Rempel et al.’s iTMO. As it can be seen, details in the dark regions are lost in the HDR frame. In addition, there are visible visual artifacts in the bright regions of the frame, such as the ice cubes on the right side of the HDR frame.

1.1.6 Kovaleski et al.’s Operator

Kovaleski et al. [4], [5] proposed an iTMO with the specific task of improving the performance of Rempel et al.’s operator in dark and bright regions. Kovaleski et al. hypothesized that the generated artifacts in Rempel et al.’s results were due to the simple fixed thresholding for
detecting bright regions. To improve the detection of this threshold, the authors proposed to implement cross-bilateral filtering [40], [41] to detect dark and bright regions of the frame and avoid visual artifacts in those areas. Figure 1.11 (c) shows the generated HDR frame using Kovaleski et al.’s operator. It can be observed that Kovaleski et al.’s iTMO does a better job in recovering the details in dark areas and removing banding in bright regions. However, the details and contrast in the bright regions are reduced.

1.1.7 Bist et al.’s Operator

Bist et al. [6], [7] proposed an iTMO with the aim of preserving the artistic intent of the original SDR content. Preserving artistic intent in this case means that if the input SDR frame is a dark, normal or bright scene, then the generated HDR frame needs to convey the same overall impression. This iTMO is based on the assumption that the input SDR frame is generated using professional stylized workflow and the target HDR display has maximum brightness of 1000 nits. The mapping curve used in this iTMO is calculated as follows:

$$L_{HDR} = 1000 \times L_{SDR}^{\gamma}$$

(1.8)

where $L_{HDR}$ and $L_{SDR}$ are the HDR and SDR luminance channels respectively, and $\gamma$ is an exponent found through comprehensive subjective tests. Based on subjective tests, authors concluded that $\gamma$ could be modeled as a function of the overall brightness of a frame. To calculate $\gamma$, first the SDR frame is converted from RGB to $L^*a^*b^*$ color space [42]. This is done because the $L^*a^*b^*$ color space does a better job in approximating human vision. Afterwards, the median value of the $L^*$ channel ($L_{median}$) is computed and clipped to the range [5 – 95] nits to avoid outlier values. Finally, $\gamma$ is calculated as:

$$\gamma = 1 + \log_{10}\left(\frac{100}{L_{median}}\right).$$

(1.9)
Figure 1.12 Mapping curves used by Bist et al.’s operator for three different γ values.

Figure 1.13 (a) An example of a dark SDR frame along with (b) its HDR frame generated by Bist et al.’s iTMO.

Figure 1.12 shows the mapping curves used by Bist et al.’s iTMO for three different γ values and a target HDR display with maximum brightness of 1000 nits. This curve maps the SDR luminance values from [0.1 -100] nits to [0.01 – 1000] nits in HDR domain. Although this iTMO does a good job in converting normal SDR frames to HDR format, it underperforms in the case of dark and bright scenes, with the generated HDR frame appearing overly dark and with reduced details in dark regions. Figure 1.13 (a) shows a dark SDR frame (γ=2.2) while Figure 1.13 (b)
presents the generated HDR frame using Bist et al.’s operator. We observe that the HDR frame appears darker than its SDR counterpart with loss of visible details in dark regions of the frame.

1.1.8 Luzardo et al.’s Operator

To further improve upon the performance of Bist et al.’s iTMO, Luzardo et al. proposed an iTMO capable of generating HDR content with maximum brightness above 1000 nits while preserving the artistic intent of the SDR content [8]. In this case, the non-linear mapping curve used is described as follows:

\[ L_{HDR} = \frac{a_{SDR}}{L_{SDR}^{ad}b + c} \]  

(1.10)

where \( L_{HDR} \) and \( L_{SDR} \) are the HDR and SDR luminance channels respectively. Based on extensive subjective tests, the authors concluded that when \( a = 1.25 \) and \( d = 4 \), the generated HDR frames have good contrast and no visible visual artifacts. Parameters \( b \) and \( c \) are calculated based on the following equations:

\[ b = \frac{m_{SDR}^{a} Disp_{max} - m_{HDR}}{m_{HDR} \left( m_{SDR}^{ad} - 1 \right) Disp_{max}} \]  

(1.11)

\[ c = \frac{m_{SDR}^{ad} m_{HDR} - m_{SDR}^{a} Disp_{max}}{m_{HDR} \left( m_{SDR}^{ad} - 1 \right) Disp_{max}} \]  

(1.12)

where \( Disp_{max} \) is the maximum brightness of the target HDR display and \( m_{SDR} \) and \( m_{HDR} \) represent the middle gray values in SDR and HDR luminance channels, respectively. As the SDR luminance channel is already provided, \( m_{SDR} = 21.47 \) nits. To calculate \( m_{HDR} \), the authors used the multi-linear regression approach and proposed the following equation:

\[ m_{HDR} = 0.017 + 0.097 L_{g,SDR} + 0.008 C_{SDR} - 0.028 P_{ov,SDR} \]  

(1.13)
where $L_{g, SDR}$ is the geometric average of the SDR luminance channel [8], $C_{SDR}$ is the contrast of the SDR frame [8], and $P_{ov, SDR}$ is the percentage of over-exposed pixels in the SDR frame. Over-exposed pixels are defined as the ones with luminance values greater than or equal 99 nits.

Figure 1.14 depicts the mapping curve used in this iTMO for $Disp_{max} = 1000$ nits, $a = 1.25$, $d = 4$, $m_{SDR} = 21.47$ nits, $m_{HDR} = 27.60$ nits, $b = -4.2758$, and $c = 5.2758$. Although this iTMO preserves the artistic intent of normal and bright frames, when it comes to dark SDR frames, generated HDR frame appears darker as compared to its SDR input resulting in loss of visible details. Figure 1.15 (a) shows the dark SDR frame in Figure 1.13 (a) while Figure 1.15 (b) shows the generated HDR frame using Luzardo et al.’s operator with $Disp_{max} = 1000$ nits, $a = 1.25$, $d = 4$, and $m_{SDR} = 21.47$ nits. As it can be seen from Figure 1.15 (b), Luzardo et al.’s iTMO does a better job in recovering the details in dark areas as compared to Bist et al.’s iTMO. However, the generated HDR frame still appears somewhat brighter compared to its SDR counterpart and loses visible details in dark regions of the frame.
Endo et al.’s Operator

Recently, deep learning-based iTMOs have attracted a lot of attention with these methods being able to generate good quality HDR content [9]–[13], [15], [16]. One of the first iTMOs in this category is the work proposed by Endo et al. [9]. The authors employ supervised learning based on deep Convolutional Neural Networks (CNNs) to generate different exposures of the input SDR frame. Afterwards, the multiple exposures are combined together using Debevec et al.’s method [43] to generate the final HDR content. Although this method is capable of generating good quality HDR content, it falls short in recovering the details of the bright regions of the frame. Another drawback of this approach is its high computational complexity. Figure 1.16 (a) shows an
SDR frame while Figure 1.16 (b) presents the generated HDR frame by Endo et al.’s operator. As it can be seen from Figure 1.16 (b), although the HDR frame is of high visual quality, details of the bright areas (e.g., moon) are lost in the HDR frame.

1.1.10 Kim et al.’s Operator

Kim et al. [10] proposed a learning-based iTMO where the input SDR frame is first divided into its base and detail layers. The relationship between details layers of the input SDR and the generated HDR frames is estimated through a training phase using a set of SDR and HDR frame pairs [10], while at the same time the base layer of the input SDR frame is expanded using a non-linear mapping curve presented by the following equation:

\[ L_{HDR} = \alpha L_{SDR} + \beta \frac{e^{\alpha L_{SDR}} - 1}{e^\alpha - 1} \]  \hspace{1cm} (1.14)

where \( L_{HDR} \) and \( L_{SDR} \) are the HDR and SDR luminance channels respectively, and \( \alpha, \alpha, \) and \( \beta \) are set to 5, 0.4, and 0.6, respectively. Figure 1.17 depicts the mapping curve used in this iTMO for \( a = 5, \alpha = 0.4, \) and \( \beta = 0.6 \). Although this approach does a good job in recovering details of dark
regions in the frame, when it comes to dark SDR frames, HDR frame will appear brighter compared to its SDR counterpart and contains visible visual artifacts in different regions of the generated HDR frame. Figure 1.18 (a) shows a dark SDR frame while Figure 1.18 (b) presents the generate HDR frame by Kim et al.’s operator for $a = 5$, $\alpha = 0.4$, and $\beta = 0.6$. As it can be seen from Figure 1.18 (b), visible details of the dark areas are well represented in the HDR frame, with some visible visual artifacts in the same areas (see the sky). In addition, the generated HDR frame looks overly bright and loses details in the bright regions.

1.2 Thesis Outline

In this thesis, we propose novel iTMOs capable of generating high visual quality HDR videos while keeping the visual artifacts at non-perceivable levels. Our proposed iTMOs are able to deal with all types of SDR content from dark, to normal and bright scenes. Through imitating the behavior of HVS when perceiving brightness, our iTMOs are capable of dividing the SDR frame into three brightness regions and then mapping those regions in an efficient manner to a higher dynamic range. Our final iTMO, based on our subjective and objective evaluations,
achieves the best possible trade-off between overall contrast and brightness of the HDR frame through maximizing a weighted sum of our proposed brightness and contrast difference functions.

In Chapter 2, we propose an entropy-based iTMO for converting SDR content to HDR format. To divide the SDR frame into three regions in terms of their brightness we propose an entropy-based segmentation. The mapping curve in this work is a piece-wise linear curve with two breaking points and the slopes for each segment of the curve are chosen manually based on comprehensive subjective tests. The segmentation approach used in this chapter, although efficient for images, is too sensitive to small variations in pixel distribution between consecutive frames, which causes flickering in video applications. Another challenge is that the proposed iTMO in this chapter is not content adaptive, as the slopes are chosen manually and are fixed for all types of SDR frames.

In Chapter 3, we propose a hybrid iTMO capable of producing high visual quality HDR videos. The mapping curve used in this work is same as the one in Chapter 2. The slopes for each segment of the curve are chosen in a content adaptive manner through our hybrid approach of maximizing our proposed contrast and brightness difference functions. By doing so, our proposed iTMO in this chapter achieves the best possible trade-off between overall contrast and brightness of the generated HDR frame. To make the segmentation process less sensitive to changes in pixel distribution among frames we proposed fixed boundaries for dark, normal, and bright regions. Although this iTMO creates high quality HDR videos, it is not completely content adaptive, as the boundaries for three brightness regions are fixed for all types of SDR frames.

In Chapter 4, we propose a novel fully content adaptive iTMO that works in perceptual domain and can deal with all types of SDR content from dark, to normal and bright scenes, and is capable of producing high visual quality HDR videos. Our proposed iTMO in this chapter
preserves the overall impression of the original SDR content by dividing it to three brightness regions using our proposed segmentation method that is based on the known theory of maximum entropy thresholding [44]. Unlike our proposed segmentation approach in Chapter 2, which results in flickering due to its high sensitivity to pixel distribution variations, this segmentation approach employs a cumulative approach that is less sensitive to these changes and avoids flickering. Although the slopes of the mapping curves are calculated the same way as in Chapter 3, the segmented regions differ as they are determined by our new segmentation method. Furthermore, we address the issue of color shifts when converting SDR frames to HDR format and propose a perception-based color adjustment method that, unlike common color adjustment approaches, results in negligible luminance change and prevents hue shifts.
Chapter 2: An Entropy-based inverse Tone Mapping Operator

To address the need for a general solution that can deal with all types of SDR content, in this chapter we propose an entropy-based iTMO that is capable of dealing with all types of SDR frames from dark, to normal and bright ones. Our proposed iTMO, unlike others, works in the perceptual domain to consider the sensitivity of the human eye to brightness changes in different regions (i.e., dark, normal, and bright) of the scene. This helps our iTMO to treat each region differently based on how our eyes perceive the different levels of brightness. This not only allows us to preserve the overall brightness of each region but also enables our iTMO to deal with all types of SDR frames (e.g., dark, normal, and bright frames), something that most iTMOs fail to address. Our subjective evaluations show that, on average, our method outperforms other state-of-the-art methods in terms of the overall visual quality of the generated HDR frame.

The rest of this chapter is organized as follows: Section 2.1 provides the details of our proposed iTMO. Section 2.2 discusses the results of our experimental evaluations and Section 2.3 concludes this chapter.

2.1 Proposed iTMO

2.1.1 Converting SDR Light Values to Perceptual Values

Our main goal is to design an iTMO capable of dealing with all types of SDR content, ranging from dark to normal and bright scenes. We base our approach on the fact that our visual system is more sensitive to brightness changes in dark regions rather than bright and normal ones. In order to take advantage of this HVS characteristic, we need to convert the captured luminance values to values that match our visual system. As it can be seen in Figure 2.1, cameras capture luminance information in linear fashion, not following the characteristics of our visual system. In fact, humans interpret the same amount of luminance difference in a different (more like a
logarithmic) way, depending on whether it is daylight or moonlight [1]. In general, HVS is more sensitive to light changes that happen in moonlight than daylight. In SDR technology, gamma correction (encoding) [45], also known as the BT.1886 standard, is used to convert camera captured values to values perceived by the HVS. However, gamma encoding was designed to address SDR values with maximum luminance of 100 nits, and thus cannot be used for HDR, which covers a much larger luminance range. For this reason, a new perceptual conversion function has been designed for HDR, known as the Perceptual Quantizer (PQ) function [46]. The PQ function is presented by the following equation:

\[
L_{PQ} = \left( \frac{c_1 + c_2 L_{\text{light}}^{m_1}}{1 + c_3 L_{\text{light}}^{m_2}} \right)^{m_2}
\]  

(2.1)

where \(L_{\text{light}}\) is the input light values in the range \([0.005 – 10000]\) nits, while \(L_{PQ}\) is the corresponding PQ values in the range \([0 – 1]\). The values for \(c_1, c_2, c_3, m_1,\) and \(m_2\) are equal to 0.8359, 18.8515, 18.6875, 0.1593, and 78.8437 respectively.
Figure 2.2, shows the gamma and PQ curves for light values in the range [0.005 – 10000] nits. The difference in the beginning parts of the two curves is due to the fact that PQ assigns more range for mapping of the dark regions of the frame. We observe that the PQ function assigns half of the output values to luminance values ranging from 0.005 to 100 nits, while keeping the rest for luminance values ranging from 100 to 10,000 nit [47]. On the other hand, gamma function uses all of its output values for luminance values ranging from 0.1 to 100. The PQ function is designed to optimize the distribution of light values with respect to the HVS properties, transforming physically linear values into perceptually linear ones. Perceptually linear means that any variation of intensity at any brightness level is seen the same way by the human eye. The latter forms the fundamental base for our implementation as we want to identify what the human eye can see and map that information to a higher dynamic range. To this end, after computing the SDR luminance channel using equation (2.2), we convert the SDR luminance values to the PQ domain, so the mapping process takes advantage of the human visual system characteristics.

\[ L_{SDR} = 0.2126R_{SDR} + 0.7152G_{SDR} + 0.0722B_{SDR} \]  

(2.2)
where $L_{SDR}$ is the luminance channel of the input SDR frame and $R_{SDR}$, $G_{SDR}$, and $B_{SDR}$ are the red, green, and blue channels of the input SDR frame.

Our next step is to compute the histogram of the resulting SDR PQ values, which groups them into uniform size bins (i.e., with equal number of brightness values in each bin). The height of each bin indicates the number of pixels in the frame corresponding to these PQ values. Figure 2.3 (a) shows an example of a normal SDR frame along with the histogram of its luminance channel in PQ domain.
channel in the PQ domain in Figure 2.3 (b). Note that SDR luminance values in this example are from 0.1 to 100 nits, which in the PQ domain correspond to 0.0623 to 0.5081 range.

### 2.1.2 Segmenting the SDR PQ Values to Three Brightness Regions

As we mentioned before, our eyes’ sensitivity to changes in brightness varies in dark, normal and bright regions. One of the main objectives in converting SDR content to HDR format is maintaining the artistic intent and the overall visual impression. To this end, we need to treat dark, normal, and bright regions of the content differently, as the eyes’ sensitivity to color changes based on brightness [48]. Our proposed segmentation method is based on the theory of maximum entropy thresholding [44]. In general, entropy is known as a measure of the uncertainty of a random variable [49]. The lower the probability of occurrence of an event, the more uncertain (higher entropy) we are about it happening and vice versa. Entropy of an image, $E$, is defined as follows [49]:

$$
E = -\sum_i p_i \log_2 (p_i)
$$

(2.3)

where $p_i$ is the number of pixels in the $i$-th bin of the histogram of the frame over the total number of pixels in the frame. Figure 2.4 depicts the histogram of SDR PQ values for the SDR frame in Figure 2.3 (a). Our goal is to find the thresholds $X_1$, and $X_2$ that divide the SDR PQ values to three brightness regions. To this end, the first step is to identify $X_{\text{min}}$ and $X_{\text{max}}$ that are the first and last non-zero elements of the histogram, respectively. To have equal search regions for $X_1$, and $X_2$, we also find $X_{\text{mid}}$ which is the middle point of the histogram. For the histogram in Figure 2.4 we have $X_{\text{min}} = 0.0623$, $X_{\text{max}} = 0.5081$, and $X_{\text{mid}} = 0.2843$. Afterwards, at each PQ point in the interval $[X_{\text{min}} - X_{\text{mid}}]$ we calculate two values: i) first $E_1$, which denotes how uncertain we are that a PQ point in the interval belongs to dark regions of the SDR frame, as follows:
where $N(i)$ is the number of pixels corresponding to $i$-th PQ point in the histogram and $N_{total}(i)$ is the total number of pixels starting from $X_{min}$ until the $i$-th PQ point in the histogram, and ii) $E_2$, that denotes how uncertain we are that a PQ point in the interval belongs to normal regions of the SDR frame, as follows:

$$E_2(X) = - \sum_{i=X_{mid}}^{X_{max}} \frac{N(i)}{N_{total}(i)} \log_2 \left( \frac{N(i)}{N_{total}(i)} \right) , \quad X \in [X_{min} - X_{mid}] \quad (2.5)$$

To find $X_1$ we compute the sum of $E_1$ and $E_2$. At each PQ point, $E_1 + E_2$ denotes how uncertain we are that a PQ point belongs to dark or normal regions. The boundary between dark and normal regions, $X_1$, is the point where $E_1 + E_2$ is maximum. In other words, $X_1$ is the point where we are most uncertain that it belongs to dark or normal regions.
We go through the same procedure for finding \( X_2 \) only in the \([X_{\text{mid}} - X_{\text{max}}]\) interval. At each PQ point in this interval, we compute two values: i) \( E_3 \), that determines how uncertain we are that a PQ point in the interval belongs to normal regions as follows:

\[
E_3(X) = - \sum_{i = X_{\text{mid}}}^{X_{\text{max}}} \frac{N(i)}{N_{\text{total}}(i)} \log_2 \left( \frac{N(i)}{N_{\text{total}}(i)} \right), \quad X \in [X_{\text{mid}} - X_{\text{max}}] \tag{2.6}
\]

where \( N(i) \) is the number of pixels corresponding to the \( i \)-th PQ point in the histogram and \( N_{\text{total}}(i) \) is the total number of pixels starting from \( X_{\text{mid}} \) until the \( i \)-th PQ point in the histogram, and ii) \( E_4 \), that determines how uncertain we are that a PQ point belongs to bright regions, as follows:

\[
E_4(X) = - \sum_{i = X}^{X_{\text{max}}} \frac{N(i)}{N_{\text{total}}(i)} \log_2 \left( \frac{N(i)}{N_{\text{total}}(i)} \right), \quad X \in [X_{\text{mid}} - X_{\text{max}}] \tag{2.7}
\]

The boundary between normal and bright regions, \( X_2 \), is the point where \( E_3 + E_4 \) is maximum. In other words, \( X_2 \) is the point where we are most uncertain that it belongs to normal or bright regions.

Figure 2.5 (a) shows the plot for \( E_1, E_2, \) and the summation of the two for the histogram in Figure 2.4. As we mentioned before, \( E_1 \) denotes how uncertain we are that a PQ point belongs to dark regions. Therefore, as it can be seen from Figure 2.5 (a) as we go towards \( X_{\text{mid}} \) (normal regions) \( E_1 \) increases. In addition, \( E_2 \) denotes how uncertain we are that a PQ point belongs to normal regions. Therefore, as we go towards \( X_{\text{min}} \) (dark regions), \( E_2 \) increases. As we discussed before and can also be seen in Figure 2.5 (a), \( X_1 \) is the point where \( E_1 + E_2 \) is maximum.

Figure 2.5 (b) shows the plot for \( E_3, E_4, \) and the summation of the two for the histogram in Figure 2.4. As we mentioned before, \( E_3 \) denotes how uncertain we are that a PQ point belongs to normal regions. Therefore, as Figure 2.5 (b) shows, as we approach \( X_{\text{max}} \) (bright regions) \( E_3 \) increases. In addition, \( E_4 \) denotes how uncertain we are that a PQ point belongs to bright regions.
Therefore, as we approach \( X_{\text{mid}} \) (normal regions), \( E_4 \) increases. As we mentioned before and is depicted in Figure 2.5 (b), \( X_2 \) is the point where \( E_3 + E_4 \) is maximum.

Figure 2.5 (c) shows the result of our proposed segmentation method for the frame in Figure 2.3 (a). Black, gray, and white areas in the figure depict dark, normal, and bright regions in the frame respectively. As Figure 2.5 (c) shows, our proposed segmentation method works reasonably well in dividing the SDR frame to three brightness regions. To further examine the performance of our proposed segmentation method Figures 2.6 (a), (b), and (c) present three SDR frames.
representing dark, normal, and bright frames, respectively. In addition, Figures 2.6 (d), (e), and (f) represent the result of our proposed segmentation method for each of those frames, respectively. Black, gray, and white portions demonstrate dark, normal, and bright regions in the input SDR frame, respectively. We can observe that our proposed segmentation method does a reasonably good job in dividing the SDR frame into three brightness regions.

One issue concerning the proposed segmentation method is that it is too sensitive to small variations in pixel distribution that can result in flickering for video applications. This sensitivity arises from the fact that when calculating the entropy, we use the number of pixels in each bin of the histogram. Therefore, since in a video sequence the number of pixels vary from one frame to the next, this results in constant change in the values of $X_1$, and $X_2$ which can ultimately cause flickering in the generated HDR video. Figures 2.7 (a) and (b) show the changes in $X_1$, and $X_2$ using our proposed segmentation method for 25 consecutive frames with no scene change,
respectively. As can be seen, our proposed segmentation method, due to its high sensitivity to changes in pixel distribution, results in large variations in the values of $X_1$ and $X_2$. We will address this issue in Chapter 4 and propose a more robust segmentation method that is less sensitive to changes in pixel distribution and prevents flickering.

2.1.3 Mapping SDR PQ Values to HDR Domain

Many different mapping curves may be used to convert SDR PQ values to HDR domain. To choose the best possible mapping curve, we thoroughly tested different types of curves, compared the generated HDR frames (using SIM2 4000 nits HDR display) using each curve using comprehensive subjective tests and evaluated the overall visual quality of the generated frames, and chose the one that achieved the highest visual quality. These curves included piece-wise linear, exponential, and polynomial curves, to mention a few. The performance evaluations concluded that the piece-wise linear function yielded acceptable results while offering the simplest implementation, good control over range, and low computational cost. Since we divide the SDR

![Figure 2.7 Changes in $X_1$ and $X_2$ using our proposed segmentation method. (a) and (b) show the changes in $X_1$ and $X_2$, respectively.](image)
frame into three brightness regions, our mapping curve is a piece-wise linear curve with two breaking points. Figure 2.8 shows an example of a mapping curve used in our iTMO. This curve maps the SDR PQ values from \([0.0623 – 0.5081]\) (corresponding to \([0.1 – 100]\) nits) to \([0.0151 – 0.9026]\) (corresponding to \([0.005 – 4000]\) nits) in the HDR domain. \(s_1, s_2, \text{ and } s_3\) in the figure are the slopes assigned to map dark, normal, and bright regions to HDR domain, respectively.

An important issue in deciding the shape of our piece-wise linear mapping curve is the slope that we assign to dark and bright regions. The slopes should be assigned in a manner that produce high visual quality HDR frames while at the same time keeping the visual artifacts at non-perceivable levels. In our proposed method, to choose the appropriate set of slopes that work well for all types of SDR content we conducted an empirical analysis. In our test, we examined different slopes for mapping dark regions of the SDR frames in our database of 300 frames that were chosen randomly from \([36]\) to represent a very large variety of content. By inspection of the generated HDR frames on a SIM2 HDR display \([50]\), we concluded that there exist a set of slopes \(1 \leq s_1 \leq 3\) for which the output contains no visual artifacts and offers more visible details in dark regions. In
other words, this range represents a good trade-off between visual quality and visual artifacts in dark regions of the generated HDR frame.

We performed the same empirical analysis for a second time, but this time we examined different slopes for mapping the bright regions of the same SDR frames. By inspection of the generated HDR frame on a SIM2 HDR display [50], we concluded that there exist a set of slopes $1.5 \leq s_3 \leq 3.9$ for which the generated HDR frames contain less banding and more visible details in bright regions. In other words, this range offers a good trade-off between visual quality and visual artifacts in bright regions of the generated HDR frame.

Based on our observations, we conclude that if we choose our slopes outside the ranges found for dark and bright scenes, the generated HDR frames will have some visible visual artifacts. Figure 2.9 demonstrates the slope limits for dark and bright regions of the SDR frame along with their corresponding mapping curves. These curves map the SDR PQ values in the range $[0.0623 – 0.5081]$ ([$0.1 – 100$] nits) to $[0.0151 – 0.9026]$ ($[0.005 – 4000$] nits). In order to have a good trade-
off between visual quality and visual artifacts in the generated HDR frame, we are limited to choosing a mapping curve that falls entirely between the red and blue curves. Any curve above the red or below the blue curve results in low visual quality HDR frames. The range of slope limits for dark and bright regions are (in degrees) $45^\circ \leq \theta_1 \leq 71^\circ$, and $56^\circ \leq \theta_2 \leq 75^\circ$, respectively. To have a single mapping curve that works well with all types of SDR content, we took the average of the maximum and minimum slopes that we could use in dark, and bright regions, ending up with a slope of $s_1 = 2$ for dark areas and $s_3 = 2.7$ for bright regions.

One issue with the proposed segmentation method is that since our search region for $X_1$ and $X_2$ begins at the first and ends at the last non-zero elements of the histogram, we may have boundaries that do not correspond to dark and/or bright regions in the frame. Therefore, to get boundaries that are more accurate for each of the three brightness regions, we need to limit the search region for $X_1$, and $X_2$. It has been proven that our eyes start to perceive colors for luminance values above 3 nits, and that the range $0.1$ (0.0623 in PQ) to $3$ (0.2133 in PQ) nits constitutes low but not quite dark lighting situations [51], [52]. Based on these observations, in our proposed segmentation method we limit the search region for $X_1$ to $[X_{\min} - 0.2133]$. As we stated previously, we only chose our slope limits for dark and bright regions of the SDR frame, and we choose the slope for normal regions based on the other two slopes. However, since in a video sequence the values of $X_1$ and $X_2$ change from frame to frame, there is a chance of either having slopes that are less than one (which result in loss of details) or even negative slopes for normal regions of the SDR frame. Therefore, in order to prevent these two cases, the slope for normal regions must be greater than or equal to one ($s_2 \geq 1$). To this end, $X_1$ and $X_2$ must satisfy the following equation:

$$1.7X_2 - X_1 \geq 0.3589.$$  \hspace{1cm} (2.8)
Therefore, after finding $X_1$ in the range $[X_{\text{min}} - 0.2133]$ and $X_2$ in the range $[X_{\text{mid}} - X_{\text{max}}]$ using our proposed segmentation method, we examine whether the value of $X_2$ satisfies equation (2.8). If it does, we keep it. Otherwise, we set its values so that $s_2 = 1$:

$$X_2 = \frac{0.3589 + X_1}{1.7}$$

(2.9)

2.2 Experimental Evaluations

To evaluate the performance of our proposed iTMO we conducted a subjective test and compared the HDR frames generated by our proposed method against the ones generated using Meylan et al.’s and Banterle et al.’s operators in terms of the overall visual quality of the generated HDR frame. We used 23 SDR frames representing dark, normal, and bright scenes. Eighteen adults including 11 males and 7 females with ages between twenty-four and thirty-five years old participated in our test. All subjects were tested for color blindness and visual acuity using Ishihara and Snellen charts respectively. Prior to the test, a training session with three SDR frames representing dark, normal, and bright scenes was conducted to familiarize the subjects with the test procedure. During the test, the HDR frames generated by our proposed iTMO and the other two operators were shown on each side of the SIM2 HDR display [50], with a 14-pixels width black stipe between each pair. Subjects were asked to choose one of the two HDR frames (“A” or “B”) that appears more visually pleasing to them. In case the two HDR frames appear to have the same level of visual quality, subjects were instructed to report them as “equal”.

The subjective test results were analyzed using the well-established paired comparison methodology proposed in [53]. In addition, an outlier detection based on the method proposed in [53] was employed and two outliers were detected and their results were excluded from our analysis.
To validate the subjects' choice consistency as well as taking any variations of brightness across the display into account, the position of the frames processed by our proposed iTMO was randomized from left to right side of the display and shown repeatedly to the subjects during the test.

Figure 2.10 shows the percentage of how many times, on average, subjects preferred the visual quality generated by our iTMO over that of Meylan et al.'s and Banterle et al.'s operators. As it can be observed, on average, our iTMO outperforms Meylan et al.'s operator by 87% and Banterle et al.'s operator by 68% of the time in terms of generating HDR frames that are more visually pleasing to the viewers. It is worth mentioning that we implemented both Banterle et al.'s, and Meylan et al.'s operators using the open source code available in [28] and set the maximum brightness of the generated HDR frame to 4000 nits. In addition, all the other parameters were chosen in a way that both of these operators yield the highest visual quality HDR frames.
2.3 Conclusions

In this chapter, we proposed an iTMO capable of generating high visual quality HDR frames. By working in the perceptual domain, our proposed iTMO is able to model the sensitivity of HVS to changes in brightness in different regions of the scene. Doing so enables our proposed iTMO to maintain the artistic intent, which among other things means to keep the dark areas dark and the bright areas bright, without darkening or brightening up the frame. To this end, we proposed an entropy-based segmentation method that divides the SDR frame into three brightness regions. To convert the SDR PQ values values to HDR domain we chose a piece-wise linear mapping curve with two breaking points. The slopes for each segment of our mapping curve were chosen based on comprehensive subjective testing and in a way that yield high visual quality HDR frames. Based on our subjective evaluations, our proposed iTMO outperforms other existing state-of-the-art methods in terms of generating high visual quality HDR frames.

As we mentioned before, since our segmentation method is too sensitive to variations in pixel distribution, it results in flickering for video applications. On the other hand, our proposed iTMO is not considered content adaptive since the slopes of our mapping curve are fixed for all types of SDR frames. In addition, since the slopes in our empirical analysis were chosen using a 4000 nits HDR display, the proposed iTMO in this chapter was only suitable for converting SDR frames to HDR format with maximum brightness of 4000 nits. To address these shortcomings, in the following chapter we introduce a general solution to the above issues and propose a content adaptive iTMO that prevents flickering and is capable of generating HDR videos for any target HDR range.
Chapter 3: A Perception-based inverse Tone Mapping Operator for High Dynamic Range Video Applications

In this Chapter, we propose a perception-based iTMO capable of producing high visual quality HDR videos. To overcome the issues of our proposed entropy-based segmentation in the previous chapter, we employ fixed boundaries for dark, normal, and bright regions of the SDR frames. By doing so, we eliminate the possibility of flickering for video applications. To choose the slopes of our piece-wise linear mapping curve we propose a hybrid optimization approach where we maximize a weighted sum of our proposed contrast and brightness difference functions. Finally, by taking into account the maximum brightness of the target HDR display, this iTMO is capable of generating HDR videos with arbitrary maximum brightness.

The rest of this chapter is organized as follows: Section 3.1 provides the details of our proposed iTMO. Section 3.2 presents the results of our experimental evaluations. Finally, Section 3.3 concludes this chapter.

3.1 Proposed iTMO

The components of our proposed iTMO are shown as part of the end-to-end workflow in Figure 3.1. The main components of our iTMO are enclosed in the red dotted line. These components include conversion of the SDR light values to perceptual (PQ) domain, computing the histogram of the resulting PQ values, dividing the frame into three brightness regions based on the eye sensitivity to brightness changes throughout the scene, calculating the slopes of the mapping curve by maximizing a weighted sum of contrast and brightness difference functions between input SDR and generated HDR frames, and finally mapping the SDR PQ values to HDR domain and
converting the HDR PQ values back to the light domain. The following sub-sections describe in
detail the main components of our iTMO.

### 3.1.1 Converting SDR Light Values to Perceptual (PQ) Domain, Computing the
Histogram and Segmentation to Three Brightness Regions

We base our design on the perceptual features of the HVS to exploit the sensitivity of our
eye to different brightness levels. This helps us to efficiently apply our mapping process. To this
end, our first task is to convert the SDR light values to perceptual values. This is an essential step
as our eyes do not see light in a linear way, with sensitivity in perceiving brightness changes
directly depending on the overall ambient brightness. More specifically, our eyes are more
sensitive to brightness variations in dark regions than they are to normal and bright ones. The
introduction of HDR led to the design of a Perceptual Quantizer (PQ) that takes advantage of HVS
properties and optimizes the distribution of light values, converting physically linear values to
perceptually linear ones [46]. Perceptual linearity indicates that our eyes perceive any change in
light intensity at any brightness level equally. This is an essential part of our implementation, as

![Figure 3.1 Workflow of our proposed iTMO. The red dotted line encloses components of our approach.](image-url)
we aim to use this knowledge when mapping visual information from a lower dynamic range to a higher one.

Our next step is to compute the histogram of the resulting SDR PQ values, which groups them into uniform size bins (i.e., with equal number of brightness values in each bin). The height of each bin indicates the number of pixels in the frame with these PQ values. Figure 3.2 (a) shows an example of an SDR frame along with the histogram of its luminance channel in PQ domain in Figure 3.2 (b). Please note that the original SDR light values in this example are from 0.1 to 100 nits, which once transferred to the PQ domain range from 0.0623 to 0.5081.
One of the main objectives in converting SDR content to HDR format is maintaining the artistic intent and the overall visual impression of the original SDR frame. To this end, we need to treat dark, normal, and bright regions of the content differently, as the eyes’ sensitivity to color changes when the brightness level varies. For instance, when the light levels drop to near total darkness, the response of the eye changes significantly according to the Scotopic response curve [48]. Therefore, to address this, our next task is to divide the SDR PQ values to dark, normal, and bright regions. Although PQ function takes into account the sensitivity of the human eye to brightness changes in different regions of a scene, it does not indicate where the boundaries of dark, normal, and bright regions are. Figure 3.3 (a) shows the PQ curve for mapping SDR light values in the range \([0.1 – 100]\) nits to \([0.0623 – 0.5081]\) in the PQ domain. As we mentioned before, our eyes are more sensitive to brightness changes in dark areas rather than normal and bright ones. This effect is clearly taken into account in Figure 3.3 (a) by the shape of the PQ curve. By inspecting Figure 3.3 (a) we can observe that dark areas are where the rate of change, as shown in Figure 3.3 (b), in the PQ curve is at its highest, while the bright areas are the ones where the rate of change is reaching its minimum value. Figures 3.3 (c) and (d) depict the zoomed areas corresponding to dark and bright regions, respectively. In bright regions, we observe that at approximately 0.5 nits the rate of change for dark regions is starting to settle to its minimum value, while at 30 nits it begins to settle to its minimum value for bright regions. Therefore, we conservatively define our dark, normal, and bright regions as the ones in the range \([0.1 – 0.5]\) nits \(([0.0623 – 0.1217] \text{ in PQ})\), \((0.5 – 30]\) nits \(([0.1217 – 0.3962]\) in PQ), and \((30 – 100]\) nits \(([0.3962 – 0.5081]\) in PQ), respectively. Our proposed brightness segmentation follows closely the zone system proposed by Ansel Adams [54]. Figures 3.4 (a), (b), and (c) show three SDR frames, while
Figures 3.4 (d), (e), and (f) depict the results of our segmentation method for each SDR frames respectively where black, gray, and white regions represent dark, normal, and bright regions, respectively. As can be seen from the figures our fixed boundaries do a reasonably good job in dividing a frame into three brightness regions.

Figure 3.3 (a) PQ curve for mapping SDR light values to PQ domain. (b) Rate of change (derivative) of the PQ curve. (c) and (d) Zoomed areas corresponding to dark and bright regions, respectively.
3.1.2 Mapping SDR PQ Values to HDR Domain

Over the years, many different mapping curves have been used, varying from polynomial to exponential. As we discussed in Chapter 2, the piece-wise linear function is a great starting point for our implementation as it is computationally efficient, and is capable of generating satisfactory results. Considering that each SDR frame is segmented into three brightness regions perceived somewhat different by the human eye, our mapping curve is a piece-wise linear curve with two breaking points presented by the following equation:

\[
L_{HDR,PQ} = \begin{cases} 
    s_1L_{SDR,PQ} + a_1 & \text{if } L_{SDR,PQ} \leq 0.1217 \\
    s_2L_{SDR,PQ} + a_2 & \text{if } 0.1217 < L_{SDR,PQ} \leq 0.3962 \\
    s_3L_{SDR,PQ} + a_3 & \text{if } L_{SDR,PQ} > 0.3962
\end{cases}
\] (3.1)

Figure 3.4 (a) - (c) An example of dark, normal and bright SDR frames, respectively. (d) - (f) Represent the result of our segmentation method for each frame, respectively. Black, gray, and white portions demonstrate dark, normal, and bright regions in the input SDR frame, respectively.
where $L_{\text{HDR,PQ}}$ and $L_{\text{SDR,PQ}}$ are the HDR and SDR luminance channels in PQ domain respectively, $s_1$, $s_2$, and $s_3$ are the slopes associated with each of our three brightness regions, $a_1$, $a_2$, and $a_3$ are the y-intercepts associated with each of the three lines in our mapping curve.

Figure 3.5 shows an example of mapping curve used in our iTMO. $X_1$ is the boundary between dark and normal regions, and $X_2$ is the boundary between normal and bright regions. This mapping curve maps the PQ values from the range $[0.0623 – 0.5081]$ (or $[0.1 – 100]$ nits) to $[0.0151 – 0.9026]$ (or $[0.005 – 4000]$ nits). It is worth mentioning that the maximum and minimum mapped values are dependent on the maximum and minimum brightness of the target HDR display. This dependency makes our method compatible with HDR displays supporting different brightness ranges.

A critical point in deciding the shape of our piece-wise linear mapping curve is the slope that we assign to dark, normal, and bright regions. The slopes should be assigned in a way that the final HDR frame looks superior to its SDR counterpart in terms of overall visual quality. One of
the main advantages of the HDR content over SDR is that it conveys a much higher contrast. Our initial main objective, thus, is to maximize the contrast in the resulting HDR frame. The contrast of the $k$-th region in HDR domain, $C_{HDR,k}$, will be:

$$C_{HDR,k} = L_{HDR,PQ,max,k} - L_{HDR,PQ,min,k} = s_k \left( L_{SDR,PQ,max,k} - L_{SDR,PQ,min,k} \right) = s_k C_{SDR,k}, \quad k = 1, 2, 3$$

(3.2)

where $L_{HDR,PQ,min,k}$, $L_{HDR,PQ,max,k}$, $L_{SDR,PQ,min,k}$, and $L_{SDR,PQ,max,k}$ are the minimum and maximum PQ values in the $k$-th region of the HDR and SDR domains respectively, $C_{SDR,k}$ is the contrast of the $k$-th region of the SDR frame, and $s_k$ is the slope assigned to $k$-th region of the SDR frame.

To maximize the contrast in each region, we treat this as a convex optimization problem, which allows us to find a global optimum solution in an efficient manner [55]. Let $C_{SDR,1}$, and $C_{HDR,1}$ be the contrast of the first region in the SDR and HDR domains, respectively. Therefore, our objective function would be:

$$\left\| C_{HDR,1} - C_{SDR,1} \right\|^2 = (s_1 - 1)^2 C_{SDR,1}^2$$

(3.3)

where $\left\| \right\|^2_2$ is the second norm ($\left\| \right\|^2$) to the power of two, and $s_1$ is the slope assigned to first region. In order to make our maximization process content adaptive, we need to take into account the distribution of pixels in each of the three brightness regions of the SDR domain. For the first region, we take the distribution of pixel values into account by employing the expectation operator as follows:

$$E \left\{ \left\| C_{HDR,1} - C_{SDR,1} \right\|^2 \right\} = (s_1 - 1)^2 C_{SDR,1}^2 p_1$$

(3.4)

where $p_1$ is the number of pixels in the first region over the total number pixels in the SDR frame, and $E$ is the expectation operator.
For all of the three brightness regions, our final objective function, $F_C(s_1, s_2, s_3)$, would be:

$$F_C(s_1, s_2, s_3) = (s_1 - 1)^2 C_{SDR,1}^2 p_1 + (s_2 - 1)^2 C_{SDR,2}^2 p_2 + (s_3 - 1)^2 C_{SDR,3}^2 p_3 = \sum_{k=1}^{3} (s_k - 1)^2 C_{SDR,k}^2 p_k$$

(3.5)

where $p_1$, $p_2$, and $p_3$ are the number of pixels in the first, second, and third regions over the total number of pixels in the SDR frame respectively and $s_1$, $s_2$, and $s_3$ are the slopes assigned to the first, second, and third regions, respectively.

Our optimization problem is limited by two constraints. The first one is that the resulting HDR contrast in each region should be greater than or equal to the corresponding one in the SDR frame, meaning that the slope for each region must be greater than or equal to one. The second condition is that the sum of the three HDR regions should be less than or equal to the maximum dynamic range supported by the target HDR display. Given the above, our final maximization problem becomes:

$$\text{Maximize} \left\{ F_C(s_1, s_2, s_3) \right\}$$

(3.6)

subject to

$$s_k \geq 1, \quad k = 1, 2, 3$$

(3.7)

$$\sum_{k=1}^{3} s_k D_{SDR,k} \leq D_{HDR}$$

(3.8)

where $s_k$ is the slope assigned to $k$-th region, $D_{SDR,k}$ is the difference between maximum and minimum PQ values in the $k$-th region of the SDR domain (actual bound), and $D_{HDR}$ is the difference between the maximum and minimum brightness values of the target HDR display in PQ domain. Solving this optimization problem for each frame yields the values of the three regional slopes for that frame, given that $D_{SDR,k}$ and $D_{HDR}$ are already known.
One issue with our approach of maximizing the contrast difference is that the resulting slopes may generate a huge difference between the SDR and HDR contrast, which in turn will yield unwanted visual artifacts such as noise in dark areas and banding in bright areas and in some extreme cases change the overall brightness of the frame.

Figures 3.6 and 3.7 show two examples of extreme cases for an overall dark and bright SDR frame, respectively. Figure 3.6 (a) depicts an overall dark SDR frame while Figure 3.6 (b) shows the corresponding HDR frame (tone mapped here using Reinhard’s tone mapping operator [37] for visualization purposes) generated by our contrast approach. Figure 3.7 (a) depicts an
overall bright SDR frame while Figure 3.7 (b) shows the corresponding HDR frame (tone mapped again for visualization purposes) generated by our contrast approach.

As it can be seen from these two examples, although we achieve high contrast in each area of the input SDR frame, in these extreme cases the dark frame became disproportionally brighter and the bright frame disproportionally darker. This is caused by the fact that in our contrast difference maximization process we only consider minimum and maximum PQ values of each region in the SDR domain and try to increase their difference as much as we can in the HDR domain. As a result, the contrast becomes disproportionally bigger, which on one hand affects the overall brightness and on the other forces the PQ values to be farther apart which generates noise and banding. To overcome this issue, we also consider maximizing the difference of the individual PQ values between input SDR and generated HDR frames, first adding the objective functions for each SDR brightness regions and then maximizing the brightness difference of the entire frame.

To maximize the difference of the individual PQ values in each region, we treat this as a convex optimization problem, which allows us to find a global optimum solution in an efficient manner [55]. Let \( L_{SDR,PQ,1} \) and \( L_{HDR,PQ,1} \) be an SDR PQ value in the first region and its corresponding HDR PQ value, respectively. Therefore, based on equation (3.1), our initial objective function for these two points will be:

\[
\|L_{HDR,PQ,1} - L_{SDR,PQ,1}\|_2^2 = \left((s_1 - 1)L_{SDR,PQ,1} + a_1\right)^2
\]  

(3.9)

where \( \| \|_2^2 \) is the second norm (\( \| \|_2 \)) to the power of two, \( s_1 \) is the slope assigned to first region, and \( a_1 \) is the y-intercept associated with the first line of our mapping curve. As equation (3.9) considers only one set of points in SDR and HDR domains, our objective function for all the PQ values in the first region would be as follows:
\[
\left\| L_{\text{HDR,PQ},1} - L_{\text{SDR,PQ},1} \right\|^2 + \left\| L_{\text{HDR,PQ},2} - L_{\text{SDR,PQ},2} \right\|^2 + \ldots \\
+ \left\| L_{\text{HDR,PQ},N_i} - L_{\text{SDR,PQ},N_i} \right\|^2 = \left( (s_1 - 1) L_{\text{SDR,PQ},1} + a_i \right)^2 \\
+ \left( (s_1 - 1) L_{\text{SDR,PQ},2} + a_i \right)^2 + \ldots + \left( (s_1 - 1) L_{\text{SDR,PQ},N_i} + a_i \right)^2 \\
= (s_1 - 1)^2 \sum_{i=1}^{N_i} L_{\text{SDR,PQ},i}^2 + 2a_i (s_1 - 1) \sum_{i=1}^{N_i} L_{\text{SDR,PQ},i} + N_i a_i^2
\]

(3.10)

where \( L_{\text{SDR,PQ},1}, L_{\text{SDR,PQ},2}, \ldots, L_{\text{SDR,PQ},N} \) and \( L_{\text{HDR,PQ},1}, L_{\text{HDR,PQ},2}, \ldots, L_{\text{HDR,PQ},N} \) are SDR and HDR PQ values in the first region respectively, \( N_i \) is the total number of SDR PQ histogram bins in the first region.

To make our maximization process content adaptive, we need to take into account the distribution of pixel values in each of our three brightness regions in the SDR frame. This is achieved by taking the expectation operator of equation (3.10):

\[
E \left\{ (s_1 - 1)^2 \sum_{i=1}^{N_i} L_{\text{SDR,PQ},i}^2 + 2a_i (s_1 - 1) \sum_{i=1}^{N_i} L_{\text{SDR,PQ},i} + N_i a_i^2 \right\} \\
= (s_1 - 1)^2 p_i \sum_{i=1}^{N_i} L_{\text{SDR,PQ},i}^2 + 2a_i (s_1 - 1) p_i \sum_{i=1}^{N_i} L_{\text{SDR,PQ},i} + N_i a_i^2
\]

(3.11)

where \( E \) is the expectation operator, \( L_{\text{SDR,PQ},i} \) is the \( i \)-th PQ value in the first region of SDR domain, and \( p_i \) is the number of pixels in the first region divided by total number of pixels in the frame. It is important to point out that since \( N_i a_i \) is a constant, the expectation operator has no effect on it.

Expanding equation (3.11) over all three brightness regions, the final objective function, \( F_B(s_1,s_2,s_3) \), becomes:

\[
F_B(s_1,s_2,s_3) = (s_1 - 1)^2 p_1 \sum_{i=1}^{N_1} L_{\text{SDR,PQ},i}^2 + 2a_1 (s_1 - 1) p_1 \sum_{i=1}^{N_1} L_{\text{SDR,PQ},i} + N_1 a_1^2 \\
+ (s_2 - 1)^2 p_2 \sum_{i=1}^{N_2} L_{\text{SDR,PQ},i}^2 + 2a_2 (s_2 - 1) p_2 \sum_{i=1}^{N_2} L_{\text{SDR,PQ},i} + N_2 a_2^2 \\
+ (s_3 - 1)^2 p_3 \sum_{i=1}^{N_3} L_{\text{SDR,PQ},i}^2 + 2a_3 (s_3 - 1) p_3 \sum_{i=1}^{N_3} L_{\text{SDR,PQ},i} + N_3 a_3^2
\]

(3.12)
where \( p_1, p_2, p_3 \) are the number of pixels in the first, second, and third brightness regions over the total number pixels in the SDR frame, respectively. Since we are dealing with 8-bit SDR frames that have 256 bins in their histogram, we have \( N_1 + N_2 + N_3 = 2^8 = 256 \).

Our optimization problem is limited by two constraints. The first one is that the resulting HDR PQ values in each region should be greater than or equal to the corresponding ones in the SDR frame, meaning that the slope for each region must be greater than or equal to one. The second condition is that the sum of three HDR regions should be less than or equal to the maximum dynamic range supported by the target HDR display. Given the above, our final maximization problem becomes:

\[
\text{Maximize} \left\{ F_B \left( s_1, s_2, s_3 \right) \right\} \quad (3.13)
\]

subject to:

\[
s_k \geq 1 \quad , \quad k = 1, 2, 3 \quad (3.14)
\]

\[
\sum_{k=1}^{3} s_k D_{SDR,k} \leq D_{HDR} \quad (3.15)
\]

where \( s_k \) is the slope assigned to \( k \)-th region, \( D_{SDR,k} \) is the difference between maximum and minimum PQ values in the \( k \)-th region of the SDR domain, and \( D_{HDR} \) is the difference between the maximum and minimum brightness of the target HDR display in PQ domain. Solving this optimization problem yields the values of the three regional slopes, given that \( D_{SDR,k} \) and \( D_{HDR} \) are already known.

Figures 3.8 and 3.9 show two examples of extreme cases of a dark and bright SDR frames, respectively. Figure 3.8 (a) depicts an overall dark SDR frame along with the HDR frame (tone mapped) generated by our contrast approach in Figure 3.8 (b) and the HDR frame generated by our brightness approach in Figure 3.8 (c). Figure 3.9 (a) depicts an overall bright SDR frame.
along with the HDR frame (tone mapped) generated by our contrast approach in Figure 3. (b) and the HDR frame generated by our brightness approach in Figure 3. (c). As it can be seen from Figures 3.8 and 3.9, the brightness approach, that works more “locally” than the contrast approach, is able to maintain the overall brightness of each region in the generated HDR frame. Therefore, to take advantage of the high contrast offered by contrast approach and good overall brightness offered by brightness approach, we decided to combine them into a hybrid approach, maximizing a weighted sum of the contrast and brightness difference functions. In other words, our maximization problem becomes:

\[
\text{Maximize} \left\{ w_1 F_{B,n} \left( s_1, s_2, s_3 \right) + w_2 F_{C,n} \left( s_1, s_2, s_3 \right) \right\}
\]  

(3.16)
subject to:

\[ s_k \geq 1 \quad , \quad k = 1, 2, 3 \]  \hspace{1cm} (3.17)

\[ \sum_{k=1}^{3} s_k D_{SDR,k} \leq D_{HDR} \]  \hspace{1cm} (3.18)

where \( F_{B,n}(s_1,s_2,s_3) \), and \( F_{C,n}(s_1,s_2,s_3) \) are the normalized brightness and contrast difference functions respectively and \( w_1 \) and \( w_2 \) are two weights with \( w_1 + w_2 = 1 \). After forming our convex hybrid maximization problem, we solve it using first order Karush-Kuhn-Tucker (KKT) optimality condition of the corresponding Lagrangian [55] and obtain our slopes for each of the three brightness regions. The process for choosing the weight values \( w_1 \) and \( w_2 \) is explained in the following section.
3.2 Experimental Evaluations

To choose the best possible combination for $w_1$ and $w_2$ that yields highest visual quality HDR frames we performed objective quality assessment using HDR VDP-2.2 [56] and PU-SSIM [57] quality metrics. These two metrics are designed for evaluating the quality of a test HDR frame with respect to an original (ground truth) HDR frame. For the original HDR frame, we used the dataset in [58] and randomly chose nine HDR frames representing dark, normal, and bright scenes for performance evaluation under different brightness conditions. Figure 3.10 shows the tone mapped version of the nine HDR frames which include Hills, Bridge, Vegas, Store, Kitchen, Park, Bench, Peak, and Church. We tone mapped the nine HDR frames using Reinhard’s tone mapping operator and used the generated SDR frames as input for our hybrid iTMO. We generated different HDR frames using three different sets of values for $w_1$ and $w_2$: 1) $w_1=w_2=0.5$, 2) $w_1=0.3$, $w_2=0.7$ and 3) $w_1=0.7$, $w_2=0.3$. Finally, we evaluated the quality of each generated HDR frame with respect to the original HDR frame for each set of values using HDR VDP-2.2 and PU-SSIM. Tables 3.1, 3.2, and 3.3 show the results of HDR VDP-2.2 and Tables 3.4, 3.5, and 3.6 show the results of PU-

![Nine original tone mapped HDR frames randomly chosen from the dataset in [58].](image-url)
Table 3.1 Visual quality scores by HDR VDP-2.2 for dark frames in Figure 3.10 for three different values of $w_1$ and $w_2$

<table>
<thead>
<tr>
<th>Frame</th>
<th>$w_1=0.5$, $w_2=0.5$</th>
<th>$w_1=0.3$, $w_2=0.7$</th>
<th>$w_1=0.7$, $w_2=0.3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegas</td>
<td>45.70955</td>
<td>45.59459</td>
<td>45.68054</td>
</tr>
<tr>
<td>Store</td>
<td>44.42627</td>
<td>44.4238</td>
<td>44.42818</td>
</tr>
<tr>
<td>Church</td>
<td>46.82211</td>
<td>46.72068</td>
<td>46.8166</td>
</tr>
<tr>
<td>Average</td>
<td><strong>45.65264</strong></td>
<td>45.57969</td>
<td>45.64177</td>
</tr>
</tbody>
</table>

Table 3.2 Visual quality scores by HDR VDP-2.2 for normal frames in Figure 3.10 for three different values of $w_1$ and $w_2$

<table>
<thead>
<tr>
<th>Frame</th>
<th>$w_1=0.5$, $w_2=0.5$</th>
<th>$w_1=0.3$, $w_2=0.7$</th>
<th>$w_1=0.7$, $w_2=0.3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kitchen</td>
<td>47.25607</td>
<td>47.20854</td>
<td>47.23436</td>
</tr>
<tr>
<td>Park</td>
<td>48.91358</td>
<td>48.90155</td>
<td>48.91529</td>
</tr>
<tr>
<td>Peak</td>
<td>46.67786</td>
<td>46.69233</td>
<td>46.67106</td>
</tr>
<tr>
<td>Average</td>
<td><strong>47.61584</strong></td>
<td>47.60081</td>
<td>47.6069</td>
</tr>
</tbody>
</table>

Table 3.3 Visual quality scores by HDR VDP-2.2 for bright frames in Figure 3.10 for three different values of $w_1$ and $w_2$

<table>
<thead>
<tr>
<th>Frame</th>
<th>$w_1=0.5$, $w_2=0.5$</th>
<th>$w_1=0.3$, $w_2=0.7$</th>
<th>$w_1=0.7$, $w_2=0.3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hills</td>
<td>48.62002</td>
<td>48.62003</td>
<td>48.60073</td>
</tr>
<tr>
<td>Bridge</td>
<td>46.78144</td>
<td>46.77396</td>
<td>46.76266</td>
</tr>
<tr>
<td>Bench</td>
<td>50.23429</td>
<td>50.21976</td>
<td>50.1995</td>
</tr>
<tr>
<td>Average</td>
<td><strong>48.54525</strong></td>
<td>48.53792</td>
<td>48.52097</td>
</tr>
</tbody>
</table>

SSIM for different sets of values for $w_1$ and $w_2$ for dark (Vegas, Store, Church), normal (Kitchen, Park, Peak), and bright (Hills, Bridge, Bench) frames, respectively. It is important to point out that the higher the score of HDR VDR-2.2 and PU-SSIM, the higher the visual quality of the generated
HDR frame. Therefore, as it can be seen from the results of both metrics, on average, when \( w_1 = w_2 = 0.5 \), our hybrid approach achieves the highest visual quality for all types of SDR frames. As a result, in our iTMO, we set the values of both \( w_1 \) and \( w_2 \) to 0.5.

### Table 3.4 Visual quality scores by PU-SSIM for dark frames in Figure 3.10 for three different values of \( w_1 \) and \( w_2 \)

<table>
<thead>
<tr>
<th>Frame</th>
<th>( w_1 = w_2 = 0.5 )</th>
<th>( w_1 = 0.3, w_2 = 0.7 )</th>
<th>( w_1 = 0.7, w_2 = 0.3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegas</td>
<td>0.959037</td>
<td>0.957373</td>
<td>0.956808</td>
</tr>
<tr>
<td>Store</td>
<td>0.966426</td>
<td>0.966354</td>
<td>0.966475</td>
</tr>
<tr>
<td>Church</td>
<td>0.970388</td>
<td>0.970428</td>
<td>0.969719</td>
</tr>
<tr>
<td>Average</td>
<td><strong>0.965284</strong></td>
<td>0.964718</td>
<td>0.964334</td>
</tr>
</tbody>
</table>

### Table 3.5 Visual quality scores by PU-SSIM for normal frames in Figure 3.10 for three different values of \( w_1 \) and \( w_2 \)

<table>
<thead>
<tr>
<th>Frame</th>
<th>( w_1 = w_2 = 0.5 )</th>
<th>( w_1 = 0.3, w_2 = 0.7 )</th>
<th>( w_1 = 0.7, w_2 = 0.3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kitchen</td>
<td>0.972699</td>
<td>0.972298</td>
<td>0.971302</td>
</tr>
<tr>
<td>Park</td>
<td>0.969191</td>
<td>0.968345</td>
<td>0.967109</td>
</tr>
<tr>
<td>Peak</td>
<td>0.972083</td>
<td>0.971996</td>
<td>0.97188</td>
</tr>
<tr>
<td>Average</td>
<td><strong>0.971324</strong></td>
<td>0.97088</td>
<td>0.970097</td>
</tr>
</tbody>
</table>

### Table 3.6 Visual quality scores by PU-SSIM for bright frames in Figure 3.10 for three different values of \( w_1 \) and \( w_2 \)

<table>
<thead>
<tr>
<th>Frame</th>
<th>( w_1 = w_2 = 0.5 )</th>
<th>( w_1 = 0.3, w_2 = 0.7 )</th>
<th>( w_1 = 0.7, w_2 = 0.3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hills</td>
<td>0.968888</td>
<td>0.968114</td>
<td>0.968038</td>
</tr>
<tr>
<td>Bridge</td>
<td>0.969359</td>
<td>0.965217</td>
<td>0.96527</td>
</tr>
<tr>
<td>Bench</td>
<td>0.988664</td>
<td>0.98168</td>
<td>0.981628</td>
</tr>
<tr>
<td>Average</td>
<td><strong>0.975637</strong></td>
<td>0.97167</td>
<td>0.971645</td>
</tr>
</tbody>
</table>
Figures 3.11 and 3.12 depict two examples of extreme cases (dark and bright SDR frames).

Figures 3.11 (a), (b), (c), and (d) depict an overall dark frame and the corresponding HDR frames (tone mapped) generated by our contrast approach, our brightness approach, and hybrid approach with $w_1 = w_2 = 0.5$, respectively. Figures 3.12 (a), (b), (c), and (d) depict an overall bright frame and the corresponding HDR frames (tone mapped) generated by our contrast approach, our brightness approach, and hybrid approach with $w_1 = w_2 = 0.5$, respectively. We observe that our hybrid approach produces results that have a good balance between overall brightness and contrast of the generated HDR frame.
To evaluate the performance of our hybrid iTMO against our brightness and contrast methods applied separately we also performed a subjective test. Nine 4K SDR video streams (each 10 seconds long) including normal, dark, and bright content were converted to HDR format using our three approaches and displayed on two similar Sony BVM-X300 professional monitors [60]. Figure 3.13 depicts one frame of each of the nine 4K SDR videos, representative of different content and brightness conditions, used in our subjective test, which include Mountain, Santorini, Niagara, Night, Sunrise, Workplace, Building, Leaves, and Room. Thirteen male, and five female subjects with ages ranging from twenty-four to thirty-two took part in our subjective evaluations. Prior to the test, all subjects were screened for color blindness and visual acuity using the Ishihara
and Snellen charts, respectively. A training session allowed all subjects to become familiar with the test procedure. The subjective test results were analyzed using the methodology in [53].

Subjects were asked to choose if one of the two HDR videos (“A” or “B”) was visually more appealing than the other or if they looked the same (“equal”). To take into account any slight variations between the two displays, the position of the videos was randomized from left to right. Two outliers were detected using the method described in [53] and the data collected from them was omitted from our evaluation. Figure 3.14 shows the average of times the subjects preferred the visual quality produced by our hybrid iTMO against brightness and contrast approaches separately. As it can be seen, our hybrid method outperforms the brightness and contrast methods by 80% and 86% of the times respectively, in terms of being more visually pleasing to the observers. We further analyzed the MOS scores of our contrast, brightness, and hybrid approaches based on $F$-value and $p$-value of one-way analysis of variance (ANOVA) to determine whether there any statistical significance between any of the three separate approaches. To do this, we performed the Levene’s test [61] to guarantee the homogeneity of variances of the gathered data, and we concluded that one-way ANOVA can be applied. Table 3.7 presents the results of the one-way ANOVA test including the mean and standard deviation of the MOS scores for each approach.
The ANOVA test concluded that the null hypothesis (i.e., there is no statistically significant difference between the three approaches) probability ($p$-value) for the dataset in our subjective test is 0.02. Since the $p$-value is less than the significance level $\alpha=0.05$ we reject the null hypothesis and conclude that the difference in MOS values obtained by our contrast, brightness, and hybrid approaches is statistically significant.
To evaluate the MOS-based performance difference between any two of our contrast, brightness and hybrid approaches, we performed Tukey’s post-hoc analysis for their MOS scores. Figure 3.15 shows the results of this test when $\alpha=0.05$. Methods in the same group are the ones with statistically insignificant difference. As it can be seen, our brightness and contrast approaches fall within the same group while our hybrid approach falls in a separate one, meaning that our hybrid approach yields statistically distinct HDR videos as compared to our brightness and contrast methods in terms of subjective quality assessment.

We also compared the performance of our hybrid iTMO against our contrast and brightness approaches using HDR VDP-2.2 and PU-SSIM objective quality metrics. To this end, we used the same nine original HDR frames in Figure 3.10 chosen randomly from the dataset in [58]. We tone mapped these HDR frames using Reinhard’s tone mapping operator and used the generated SDR frames as input for our contrast, brightness and hybrid approaches. Finally, we evaluated the quality of generated HDR frames with respect to the original HDR frames using HDR VDP-2.2 and PU-SSIM. Tables 3.8, 3.9, and 3.10 show the results of HDR VDP-2.2 for our three separate approaches for dark (*Vegas, Store, Church*), normal (*Kitchen, Park, Peak*), and bright (*Hill, Bridge, Bench*) frames respectively. In addition, Tables 3.11, 3.12, and 3.13 show the results of PU-SSIM for our three separate approaches for dark, normal, and bright frames, respectively. As it can be seen from the results of both metrics, on average, our hybrid approach achieves the highest visual quality for all types of SDR frames, reaffirming the result of our subjective test.
We evaluated the performance of our proposed hybrid iTMO by carrying out another subjective test, comparing its performance against those of Meylan et al.’s [14], Rempel et al.’s [27], Kovaleski et al.’s [5], Bist et al.’s [7] and Kim et al.’s [10] iTMOs in terms of the overall visual quality scores. The results are shown in Table 3.8 for dark frames, Table 3.9 for normal frames, and Table 3.10 for bright frames.

**Table 3.8 Visual quality scores by HDR VDP-2.2 for dark frames in Figure 3.10 using hybrid, brightness and contrast approaches.**

<table>
<thead>
<tr>
<th>Frame</th>
<th>Hybrid approach</th>
<th>Contrast approach</th>
<th>Brightness approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegas</td>
<td>45.70955</td>
<td>43.67382</td>
<td>44.13756</td>
</tr>
<tr>
<td>Store</td>
<td>44.42627</td>
<td>41.28935</td>
<td>41.27666</td>
</tr>
<tr>
<td>Church</td>
<td>46.82211</td>
<td>45.11983</td>
<td>45.59701</td>
</tr>
<tr>
<td>Average</td>
<td><strong>45.65264</strong></td>
<td>43.361</td>
<td>43.67041</td>
</tr>
</tbody>
</table>

**Table 3.9 Visual quality scores by HDR VDP-2.2 for normal frames in Figure 3.10 using hybrid, brightness and contrast approaches.**

<table>
<thead>
<tr>
<th>Frame</th>
<th>Hybrid approach</th>
<th>Contrast approach</th>
<th>Brightness approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kitchen</td>
<td>47.25607</td>
<td>46.74428</td>
<td>47.11382</td>
</tr>
<tr>
<td>Park</td>
<td>48.91358</td>
<td>48.97533</td>
<td>48.91015</td>
</tr>
<tr>
<td>Peak</td>
<td>46.67786</td>
<td>46.81748</td>
<td>46.67316</td>
</tr>
<tr>
<td>Average</td>
<td><strong>47.61584</strong></td>
<td>47.51236</td>
<td>47.56571</td>
</tr>
</tbody>
</table>

**Table 3.10 Visual quality scores by HDR VDP-2.2 for bright frames in Figure 3.10 using hybrid, brightness and contrast approaches.**

<table>
<thead>
<tr>
<th>Frame</th>
<th>Hybrid approach</th>
<th>Contrast approach</th>
<th>Brightness approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hills</td>
<td>48.62002</td>
<td>48.57038</td>
<td>48.54986</td>
</tr>
<tr>
<td>Bridge</td>
<td>46.78144</td>
<td>46.74898</td>
<td>46.71023</td>
</tr>
<tr>
<td>Bench</td>
<td>50.23429</td>
<td>49.86225</td>
<td>49.79646</td>
</tr>
<tr>
<td>Average</td>
<td><strong>48.54525</strong></td>
<td>48.39387</td>
<td>48.35218</td>
</tr>
</tbody>
</table>

We evaluated the performance of our proposed hybrid iTMO by carrying out another subjective test, comparing its performance against those of Meylan et al.’s [14], Rempel et al.’s [27], Kovaleski et al.’s [5], Bist et al.’s [7] and Kim et al.’s [10] iTMOs in terms of the overall visual quality scores.
visual quality of the generated HDR video. The parameters of each of the iTMOs were set to their
default values as mentioned in each of their respective papers. For Meylan et al.’s, Rempel et al.’s,
and Kovaleski et al.’s iTMOs we used the publicly available source code in [28]. We implemented
Bist et al.’s iTMO based on the description in its respective paper. Authors of [10] generously provided us with the implementation of their method. We used a new set of nine 4K SDR video streams (each 10 seconds long) [59] involving normal, dark, and bright content for this test. Figure 3.16 depicts one frame of each of the nine 4K SDR videos used in our subjective test, which include *Bellagio, Cave, Fireworks, Flower, Garden, Lake, Desert, Reflection,* and *Surfing.* As can be seen, our dataset includes dark, normal, and bright scenes to evaluate the performance of our iTMO against others under different brightness conditions. Participants in the test were twelve male and six female subjects with ages ranging from twenty-three to thirty-one. All subjects were screened for color blindness and visual acuity by Ishihara and Snellen charts respectively. Subjects who participated in this test were different from the ones who participated in our previous test. A training session allowed subjects to get familiar with the test procedure. Once more, the two HDR videos were shown on two similar Sony BVM-X300 professional monitors, and the subjective test results were analyzed according to the method proposed in [53]. The order of all possible combinations of the video pairs was randomized, yielding a total of 45 video streams. Subjects were required to choose either one of the two HDR videos as more visually appealing or decide if they looked the same. The videos were randomly changed between the two displays to remove any
effect caused by possible differences between the two professional monitors and their physical location. Two outliers were found using the methodology in [53] and their results were omitted from our analysis.

Figure 3.17 shows the frequency of times subjects favored the visual quality produced by our iTMO against Meylan et al.’s, Rempel et al.’s, Kovaleski et al.’s, Bist et al.’s, and Kim et al.’s iTMOs. We observe that, on average, our iTMO outperforms Meylan et al.’s by 78%, Rempel et al.’s by 94%, Kovaleski et al.’s by 81%, Bist et al.’s by 91%, and Kim et al.’s by 69% of the times in terms of being more visually pleasing to the observers. We further analyzed the MOS scores of the six iTMOs in this subjective test in Table 3.14 based on F-value and p-value of one-way analysis of variance (ANOVA). To do this, we performed the Levene’s test to guarantee the homogeneity of variances of the gathered data, and we concluded that one-way ANOVA can be applied. The ANOVA test concluded that the null hypothesis probability (p-value) for the dataset in our subjective test is $3.87 \times 10^{-11}$, meaning that difference in MOS values by the six iTMOs is statistically significant at significance level $\alpha=0.05$. 

![Figure 3.17 Mean opinion scores (MOS) indicating the visual quality of the output HDR video generated by our iTMO against Meylan et al.’s, Rempel et al.’s, Kovaleski et al.’s, Bist et al.’s, and Kim et al.’s](image-url)
To evaluate the MOS-based performance difference between any two of the six iTMOs, we performed Tukey’s post-hoc analysis for their MOS scores. Figure 3.18 shows the results of this test when $\alpha=0.05$. Methods in the same group are the ones with statistically insignificant difference. As it can be seen, Meylan et al.’s, Rempel et al.’s, Kovaleski et al.’s, Bist et al.’s, and Kim et al.’s iTMOs fall within the same group while our proposed hybrid iTMO falls in a separate one, meaning our proposed iTMO yields statistically distinct HDR videos as compared to other five iTMOs in terms of subjective quality assessment.

We also compared the performance of our proposed hybrid iTMO against Meylan et al.’s, Rempel et al.’s, Kovaleski et al.’s, Bist et al.’s, and Kim et al.’s iTMOs using HDR VDP-2.2 and PU-SSIM objective quality metrics. To this end, we used nine original HDR frames randomly

### Table 3.14 Mean and standard deviation of MOS scores for the six iTMOs based on the dataset in Figure 3.16

<table>
<thead>
<tr>
<th>iTMO</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meylan et al.’s</td>
<td>1.75</td>
<td>1.4832</td>
</tr>
<tr>
<td>Rempel et al.’s</td>
<td>0.5</td>
<td>0.7302</td>
</tr>
<tr>
<td>Kovaleski et al.’s</td>
<td>1.5</td>
<td>1.2110</td>
</tr>
<tr>
<td>Bist et al.’s</td>
<td>0.75</td>
<td>0.8563</td>
</tr>
<tr>
<td>Kim et al.’s</td>
<td>2.625</td>
<td>1.1474</td>
</tr>
<tr>
<td>Proposed</td>
<td>35.375</td>
<td>3.7925</td>
</tr>
</tbody>
</table>

$F$-value ($p$-value) $= 15.7863 \times (3.87483 \times 10^{-11})$
chosen from the dataset in [58] representing dark, normal, and bright scenes to evaluate the performance of our iTMO against others under different brightness conditions. Figure 3.19 shows the tone mapped version of the nine HDR frames which include *Waterfall, Snow, Arch, Lamp, Typewriter, Peppermill, Pier, Gorse,* and *Rock*. We tone mapped these HDR frames using Reinhard’s tone mapping operator and used the generated SDR frames as input for the six iTMOs. Finally, we evaluated the quality of the generated HDR frames with respect to the original HDR frame using HDR VDP-2.2 and PU-SSIM. Tables 3.15, 3.16, and 3.17 show the results of HDR VDP-2.2 for the six iTMOs for dark (*Lamp, Typewriter, Peppermill*), normal (*Pier, Gorse, Rock*), and bright (*Waterfall, Snow, Arch*) frames respectively. In addition, Tables 3.18, 3.19, and 3.20 show the results of PU-SSIM for the six iTMOs for dark, normal, and bright frames respectively. As it can be seen from the results of both metrics, on average, our proposed hybrid iTMO achieves the highest visual quality for all types of SDR frames, reaffirming the result of our subjective test.

We also compared the execution time of our proposed iTMO against other five iTMOs. We calculated the execution times based on unoptimized source codes for each iTMO. In addition, the execution times is calculated for a 4K SDR frame using MATLAB R2018a software on a computer with Intel(R) Core(TM) i7-5820K CPU @ 3.3GHz. Table 3.21 shows the results of the
Table 3.15 Visual quality scores by HDR VDP-2.2 for dark frames in Figure 3.19 for six different iTMOs.

<table>
<thead>
<tr>
<th>Frame</th>
<th>Proposed</th>
<th>Kim</th>
<th>Kovaleski</th>
<th>Meylan</th>
<th>Bist</th>
<th>Rempel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lamp</td>
<td>53.0473</td>
<td>46.74879</td>
<td>46.963256</td>
<td>47.519966</td>
<td>49.58417</td>
<td>52.69622</td>
</tr>
<tr>
<td>Typewriter</td>
<td>48.10044</td>
<td>43.2599</td>
<td>44.020972</td>
<td>44.469251</td>
<td>45.87392</td>
<td>46.75414</td>
</tr>
<tr>
<td>Peppermill</td>
<td>46.37453</td>
<td>41.48034</td>
<td>42.870002</td>
<td>42.481676</td>
<td>43.26849</td>
<td>43.60474</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>49.17409</strong></td>
<td>43.82968</td>
<td>44.618077</td>
<td>44.823631</td>
<td>46.24219</td>
<td>47.68503</td>
</tr>
</tbody>
</table>

Table 3.16 Visual quality scores by HDR VDP-2.2 for normal frames in Figure 3.19 for six different iTMOs.

<table>
<thead>
<tr>
<th>Frame</th>
<th>Proposed</th>
<th>Kim</th>
<th>Kovaleski</th>
<th>Meylan</th>
<th>Bist</th>
<th>Rempel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pier</td>
<td>50.72516</td>
<td>43.02445</td>
<td>43.178084</td>
<td>42.929074</td>
<td>45.11719</td>
<td>46.03405</td>
</tr>
<tr>
<td>Gorse</td>
<td>46.3769</td>
<td>43.82617</td>
<td>45.403682</td>
<td>45.463806</td>
<td>44.55143</td>
<td>46.03896</td>
</tr>
<tr>
<td>Rock</td>
<td>46.54312</td>
<td>44.60946</td>
<td>46.654773</td>
<td>46.733931</td>
<td>44.51708</td>
<td>46.93629</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>47.88173</strong></td>
<td>43.82003</td>
<td>45.078846</td>
<td>45.04227</td>
<td>44.72857</td>
<td>46.33643</td>
</tr>
</tbody>
</table>

Table 3.17 Visual quality scores by HDR VDP-2.2 for bright frames in Figure 3.19 for six different iTMOs.

<table>
<thead>
<tr>
<th>Frame</th>
<th>Proposed</th>
<th>Kim</th>
<th>Kovaleski</th>
<th>Meylan</th>
<th>Bist</th>
<th>Rempel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterfall</td>
<td>47.95183</td>
<td>45.67576</td>
<td>47.920803</td>
<td>47.880207</td>
<td>45.50425</td>
<td>47.99985</td>
</tr>
<tr>
<td>Snow</td>
<td>51.72745</td>
<td>46.24283</td>
<td>47.253952</td>
<td>47.31836</td>
<td>48.25494</td>
<td>49.8245</td>
</tr>
<tr>
<td>Arch</td>
<td>47.32708</td>
<td>44.84433</td>
<td>47.432354</td>
<td>47.294075</td>
<td>44.06702</td>
<td>47.08561</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>49.00212</strong></td>
<td>45.58764</td>
<td>47.535703</td>
<td>47.497548</td>
<td>45.94207</td>
<td>48.30332</td>
</tr>
</tbody>
</table>

execution times for each iTMO in seconds. As it can be seen our proposed iTMO ranks second in terms of execution time which is reasonable considering the higher visual quality it provides.
### Table 3.18 Visual quality scores by PU-SSIM for dark frames in Figure 3.19 for six different iTMOs.

<table>
<thead>
<tr>
<th>Frame</th>
<th>Proposed</th>
<th>Kim</th>
<th>Kovaleski</th>
<th>Meylan</th>
<th>Bist</th>
<th>Rempel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lamp</td>
<td>0.995766</td>
<td>0.982229</td>
<td>0.9740865</td>
<td>0.9764124</td>
<td>0.984618</td>
<td>0.996626</td>
</tr>
<tr>
<td>Typewriter</td>
<td>0.990925</td>
<td>0.970268</td>
<td>0.9592009</td>
<td>0.9623684</td>
<td>0.97299</td>
<td>0.992212</td>
</tr>
<tr>
<td>Peppermill</td>
<td>0.982823</td>
<td>0.962577</td>
<td>0.9482766</td>
<td>0.9419287</td>
<td>0.961448</td>
<td>0.975384</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.989838</strong></td>
<td><strong>0.971692</strong></td>
<td><strong>0.9605213</strong></td>
<td><strong>0.9602365</strong></td>
<td><strong>0.973019</strong></td>
<td><strong>0.988074</strong></td>
</tr>
</tbody>
</table>

### Table 3.19 Visual quality scores by PU-SSIM for normal frames in Figure 3.19 for six different iTMOs.

<table>
<thead>
<tr>
<th>Frame</th>
<th>Proposed</th>
<th>Kim</th>
<th>Kovaleski</th>
<th>Meylan</th>
<th>Bist</th>
<th>Rempel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pier</td>
<td>0.97675</td>
<td>0.965017</td>
<td>0.960169</td>
<td>0.9601589</td>
<td>0.965688</td>
<td>0.969436</td>
</tr>
<tr>
<td>Gorse</td>
<td>0.972222</td>
<td>0.963704</td>
<td>0.9598112</td>
<td>0.9609173</td>
<td>0.967228</td>
<td>0.965017</td>
</tr>
<tr>
<td>Rock</td>
<td>0.977566</td>
<td>0.966256</td>
<td>0.9611561</td>
<td>0.961659</td>
<td>0.968312</td>
<td>0.967247</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.975513</strong></td>
<td><strong>0.964992</strong></td>
<td><strong>0.9603788</strong></td>
<td><strong>0.9609118</strong></td>
<td><strong>0.967076</strong></td>
<td><strong>0.967233</strong></td>
</tr>
</tbody>
</table>

### Table 3.20 Visual quality scores by PU-SSIM for bright frames in Figure 3.19 for six different iTMOs.

<table>
<thead>
<tr>
<th>Frame</th>
<th>Proposed</th>
<th>Kim</th>
<th>Kovaleski</th>
<th>Meylan</th>
<th>Bist</th>
<th>Rempel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterfall</td>
<td>0.979287</td>
<td>0.97418</td>
<td>0.9712239</td>
<td>0.9719816</td>
<td>0.973783</td>
<td>0.97463</td>
</tr>
<tr>
<td>Snow</td>
<td>0.976956</td>
<td>0.967769</td>
<td>0.962969</td>
<td>0.9631881</td>
<td>0.966271</td>
<td>0.969094</td>
</tr>
<tr>
<td>Arch</td>
<td>0.983075</td>
<td>0.975587</td>
<td>0.9684621</td>
<td>0.9704864</td>
<td>0.973628</td>
<td>0.974917</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.979772</strong></td>
<td><strong>0.972512</strong></td>
<td><strong>0.9675517</strong></td>
<td><strong>0.9685521</strong></td>
<td><strong>0.971227</strong></td>
<td><strong>0.97288</strong></td>
</tr>
</tbody>
</table>

### Table 3.21 Execution times of six iTMOs (in seconds) for a 4K SDR frame

<table>
<thead>
<tr>
<th>Execution Time (sec.)</th>
<th>Proposed</th>
<th>Kim</th>
<th>Kovaleski</th>
<th>Meylan</th>
<th>Bist</th>
<th>Rempel</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Execution Time (sec.)</strong></td>
<td><strong>2.3973</strong></td>
<td><strong>49.6554</strong></td>
<td><strong>1.0892</strong></td>
<td><strong>2.7858</strong></td>
<td><strong>2.9084</strong></td>
<td><strong>48.0186</strong></td>
</tr>
</tbody>
</table>
3.3 Conclusions

In this chapter, we proposed a hybrid iTMO capable of generating high visual quality HDR videos. Based on our subjective and objective evaluations, our iTMO achieves the best possible trade-off between the overall brightness and contrast of the generated HDR video and is capable of generating HDR videos for any dynamic range supported by the target HDR display. To address the issue of flickering caused by our proposed iTMO in Chapter 2, we proposed fixed boundaries for dark, normal, and bright regions of the SDR frame. To achieve high contrast HDR videos, we aimed at maximizing our proposed contrast difference function between input SDR and generated HDR frames. Although this approach yields high contrast HDR videos, in extreme cases it affects the overall brightness of the frame with dark SDR frames appearing overly bright in HDR domain and bright SDR frames overly dark. To overcome this issue, we proposed a more local optimization approach and aimed at maximizing the difference of individual PQ values in SDR and HDR domain, a process that we called maximizing the brightness difference function between input SDR and generated HDR frames. Although our brightness approach preserves the overall brightness of the HDR frame, the generated results are not of high contrast. As a result, to take advantage of the capabilities of both contrast and brightness approaches, we proposed our hybrid approach and aimed at maximizing a weighted sum of the two functions. Our subjective and objective evaluations demonstrate that our proposed hybrid iTMO outperforms other state-of-the-art methods in terms of overall visual quality of the generated HDR video.

One disadvantage of this iTMO is that the segmentation method is not content adaptive, as the boundaries for our three brightness regions are fixed for all types of SDR frames. In the following chapter, we will address this issue and propose a fully content adaptive segmentation method capable of producing flicker-free high visual quality HDR videos.
Chapter 4: A Fully-Automatic Content Adaptive Inverse Tone Mapping Operator with Improved Color Accuracy

The proposed iTMO in Chapter 3 is not content adaptive as it employs fixed boundaries for determining dark, normal, and bright regions of the SDR frame. Furthermore, the proposed segmentation method in Chapter 2, although content adaptive, is too sensitive to small variations in pixel distribution and results in flickering for video applications. Therefore, a content adaptive segmentation method that is less sensitive to small variations in pixel distribution not only avoids flickering in the generated HDR video, but can also improve the visual quality of it.

In this Chapter, we propose a fully content adaptive iTMO that divides the input SDR frame to three brightness regions based on a novel entropy-based segmentation approach. Our proposed segmentation is based on the theory of maximum entropy thresholding [44] and accurately divides the input SDR frame to three brightness regions (i.e., dark, normal, and bright) to maintain the overall visual impression of the scene, while avoiding flickering. Furthermore, we propose a perception-based color adjustment method that is capable of preserving the color accuracy between the input SDR and the generated HDR frame. By performing the color adjustment in the perceptual domain, the proposed method preserves the hue of the SDR colors and generates HDR colors that closely follow their SDR counterparts, irrespective of the iTMO used. Based on our subjective and objective evaluations, our proposed iTMO outperforms other state-of-the-art operators in terms of overall visual quality of the generated HDR video and generates HDR colors that closely follow their SDR counterpart.
The remainder of this chapter is organized as follows: Section 4.1 discusses the details of our proposed iTMO. Section 4.2 presents the results of our experimental evaluations, and Section 4.3 concludes this chapter.

### 4.1 Proposed iTMO

The block diagram of our proposed iTMO is presented in Figure 4.1. The main components of our proposed iTMO are presented in the red-dotted line. First, the input SDR frame is linearized using inverse gamma encoding [45]. Then, the SDR luminance channel is computed and these normalized [0-1] values are converted to light values in the range [0.1 – 100] nits, which are subsequently converted to perceptually linear values. Afterwards we compute the distribution (histogram) of these values. This step is followed by segmenting these values into three brightness regions (i.e., dark, normal, and bright) using our proposed entropy-based segmentation. Next, we compute the slopes of our mapping curve using our previously proposed hybrid approach of maximizing a weighted sum of brightness and contrast difference between input SDR and output HDR frame in Chapter 3. These slopes help us perform the inverse tone mapping, converting the

![Figure 4.1 The block diagram of our proposed iTMO. The main components of our method are enclosed in the red dotted line. The generated HDR frame in this diagram has maximum brightness of 1000 nits. However, in general, our method is capable of producing HDR videos with arbitrary maximum brightness.](image-url)
SDR PQ values to HDR domain. To have HDR colors that closely match their SDR counterparts we employ our proposed perception-based color adjustment method. Finally, the HDR perceptual values are converted back to light values, generating the HDR frame. The following sub-sections offer an in-depth description of each of these steps.

4.1.1 Converting SDR Light Values to PQ Domain and Computing Histogram

As we mentioned before, our visual system does not perceive the brightness information of a scene in a linear way like cameras do [1]. In fact, they are more sensitive to changes in brightness in dark regions rather than normal and bright ones. To take advantage of this capability, we decided to implement our iTMO in perceptual domain. Doing so enables us to detect regions of the SDR frame that are visually important to our eyes and efficiently map those regions to higher brightness levels. To convert the SDR light values to perceptual domain we employ Perceptual Quantizer (PQ) transform [46], which is designed to optimize the distribution of light values with respect to HVS characteristics.

Therefore, as a first step in our iTMO, we convert the input SDR light values to PQ domain. Afterwards, we compute the histogram of these SDR PQ values. Figure 4.2 (a) shows an example of an SDR frame along with the histogram of its luminance channel in PQ domain in Figure 4.2 (b). The SDR light values in this example range from [0.1 – 100] nits which translates to [0.0623 – 0.5081] in PQ domain.

4.1.2 Segmenting the SDR PQ Values to Three Brightness Regions

As we discussed earlier in this chapter, our eyes are more sensitive to brightness changes in dark regions than they are to normal and bright ones. Therefore, it is important for our proposed iTMO to be able to detect each of these three brightness regions. To do so, we segment the SDR
PQ values to three brightness regions namely dark, normal, and bright. Our proposed segmentation method is based on the theory of maximum entropy thresholding [44].

As we mentioned in Chapter 2, entropy is known as a measure of the uncertainty of a random variable [49]. The lower the probability of occurrence of random variable, the more uncertain (higher entropy) we are about it happening and vice versa. Entropy of an image, $E$, is defined as follows [49]:

![Figure 4.2](image_url)

Figure 4.2 (a) An SDR frame along with (b) the histogram of its luminance channel in PQ domain.

PQ values to three brightness regions namely dark, normal, and bright. Our proposed segmentation method is based on the theory of maximum entropy thresholding [44].

As we mentioned in Chapter 2, entropy is known as a measure of the uncertainty of a random variable [49]. The lower the probability of occurrence of random variable, the more uncertain (higher entropy) we are about it happening and vice versa. Entropy of an image, $E$, is defined as follows [49]:

\[ E = -\sum_{x} p(x) \log p(x) \]
where \( p_i \) is the number of pixels in the \( i \)-th bin of the histogram of the frame over total number of pixels in the frame. In addition, \( H \), and \( W \) denote height and width of the SDR frame respectively.

Figure 4.3 depicts the histogram of SDR luminance values in the PQ domain for the SDR frame shown in Figure 4.2 (a). As a first step in our segmentation method, we define \( X_{\text{min}} \), and \( X_{\text{max}} \) as the minimum and maximum PQ values in the SDR domain respectively. In addition, we define \( X_{\text{gray}} \) as the middle gray point (defined as middle code-word 128) in PQ domain, according to Ansel Adam’s famous zone system [54]. The values for \( X_{\text{min}}, X_{\text{gray}}, \) and \( X_{\text{max}} \) are dependent on the maximum brightness of the SDR content and are fixed for all the frames throughout the SDR video. For instance, for the case of \([0.1 – 100] \text{ nits} \) SDR content, we have \( X_{\text{min}} = 0.0623, X_{\text{gray}} = 0.3651 \) (PQ equivalent of middle code-word 128), and \( X_{\text{max}} = 0.5081 \). The position of \( X_{\text{min}}, X_{\text{gray}}, \) and \( X_{\text{max}} \) are marked with vertical red-dotted lines in Figure 4.3. Our goal is to find the thresholds \( X_1 \), and \( X_2 \) (shown by vertical blue-dotted lines in Figure 4.3) that divide the SDR PQ values to three brightness regions. To make our proposed segmentation method less sensitive to changes in pixel distribution among frames and to avoid flickering, we adopt a cumulative approach and work with the summation of pixels until each bin when calculating the entropy. To do so, at each PQ point in the interval \([X_{\text{min}} – X_{\text{gray}}]\) we calculate two entropy values. First, we compute \( E_1 \), which denotes how uncertain we are that a PQ point in this interval belongs to the dark region of the SDR frame, as follows:

\[
E_1(X) = - \sum_{i=X_{\text{min}}}^{X} \frac{S(i)}{N_{\text{total}}(X_{\text{gray}})} \log_2 \left( \frac{S(i)}{N_{\text{total}}(X_{\text{gray}})} \right), \ X \in [X_{\text{min}} – X_{\text{gray}}]
\]  

\[(4.2)\]
where $S(i)$ is the sum of pixels until the $i$-th PQ point in the histogram. Also, $N_{total}(X_{gray})$ is the total number of pixels starting from $X_{min}$ until $X_{gray}$.

The entropy $E_1$ for the histogram in Figure 4.3 is shown in Figure 4.4 (a) in red. We observe that, since $E_1$ denotes how uncertain we are that a PQ point belongs to dark region, its value increases as we go towards $X_{gray}$ (normal region). We also calculate $E_2$, which denotes how uncertain we are that a PQ point in the interval belongs to normal region of the SDR frame, as follows:

$$E_2(X) = - \sum_{i=X_{min}}^{X_{max}} \frac{S(i)}{N_{total}(X_{gray})} \log_2 \left( \frac{S(i)}{N_{total}(X_{gray})} \right), \ X \in [X_{min} - X_{gray}] \quad (4.3)$$

The entropy $E_2$ for the histogram in Figure 4.3 is shown in Figure 4.4 (a) in blue. In this case, as we move towards $X_{min}$ (dark region), $E_2$ increases as it denotes the uncertainty of belonging to the normal region. To find $X_1$ we compute the sum of $E_1$ and $E_2$. At each PQ point, $E_1 + E_2$ denotes...
how uncertain we are that a PQ value belongs to the dark or normal region. Threshold $X_1$ that defines the boundary between dark and normal region (the point where we are most uncertain that it belongs to dark or normal region) is the point where $E_1 + E_2$ is maximum (see the green curve in Figure 4.4 (a)).

We go through the same process for finding $X_2$ in the $[X_{\text{gray}} - X_{\text{max}}]$ interval, which defines the boundary between the normal and bright region. Once more, at each PQ point in this interval,
we compute two entropy values. First, we compute $E_3$ that determines how uncertain we are that a PQ value in the interval belongs to the normal region as follows:

$$E_3(X) = -\sum_{i=\text{gray}}^{X} \frac{S(i)}{N_{\text{total}}(X_{\text{max}})} \log_2 \left( \frac{S(i)}{N_{\text{total}}(X_{\text{max}})} \right), \ X \in [X_{\text{gray}} - X_{\text{max}}]$$

(4.4)

where $S(i)$ is the sum of pixels until the $i$-th PQ point in the histogram. Also, $N_{\text{total}}(X_{\text{max}})$ is the total number of pixels starting from $X_{\text{gray}}$ until $X_{\text{max}}$.

The entropy $E_3$ for the histogram in Figure 4.3 is shown in Figure 4.4 (b) in red. We observe that since $E_3$ denotes how uncertain we are that a PQ point belongs to normal region, its value increases as we approach $X_{\text{max}}$ (bright region). In addition to $E_3$, we calculate $E_4$, which determines how uncertain we are that a PQ value in the interval belongs to the bright region as follows:

$$E_4(X) = -\sum_{i=X}^{X_{\text{max}}} \frac{S(i)}{N_{\text{total}}(X_{\text{max}})} \log_2 \left( \frac{S(i)}{N_{\text{total}}(X_{\text{max}})} \right), \ X \in [X_{\text{gray}} - X_{\text{max}}].$$

(4.5)

The entropy $E_4$ for the histogram in Figure 4.3 is presented in Figure 4.4 (b) in blue. In this case, as we approach $X_{\text{gray}}$ (normal region), $E_4$ increases as it denotes the uncertainty of belonging to bright region. To find $X_2$ we calculate $E_3+E_4$ which denotes how uncertain we are that a PQ point belongs to normal or bright region. The boundary between the normal and bright region, $X_2$, is the point where $E_3+E_4$ is maximum (the point where we are most uncertain that it belongs to normal or bright region). Difference in amplitude of entropy between Figures 4.4 (a) and (b) is due to difference in pixels distribution between $[X_{\text{min}} - X_{\text{gray}}]$ and $[X_{\text{gray}} - X_{\text{max}}]$ intervals.

It is important to mention that in case the pixel distribution in the $[X_{\text{min}} - X_{\text{gray}}]$ interval is zero, we have $X_1 = X_{\text{gray}}$. The same conclusion is true for the $[X_{\text{gray}} - X_{\text{max}}]$ interval meaning that in case the pixel distribution is zero in this interval, we have $X_2 = X_{\text{gray}}$. 

77
Figure 4.4 (c) shows the original SDR frame in Figure 4.2(a) while Figure 4.4 (d) depicts the result of our proposed segmentation method. Black, gray, and white areas in Figure 4.4 (d) denote the dark, normal, and bright regions in the frame, respectively. Figure 4.5 further demonstrates the performance of our proposed segmentation method. Figures 4.5 (a), (b), and (c) show a dark, normal, and bright SDR frames, respectively, while Figures 4.5 (d), (e), and (f) show the segmentation results, where black, gray, and white areas represent dark, normal, and bright regions, respectively. As it can be observed, our proposed segmentation accurately divides the SDR frame to three brightness regions.

4.1.3 Mapping SDR PQ Values to HDR Domain

As we mentioned in Chapter 2 and 3, to choose the best possible mapping curve for our iTMO, we ran a series of extensive subjective tests and examined different types of mapping curves. We concluded that piece-wise linear curve offers the simplest implementation, and low
computational cost, while at the same time producing high visual quality HDR videos. Since we divide the SDR PQ values into three brightness regions, our mapping curve is a piece-wise linear curve with two breaking points, which can be written as:

\[ L_{HDR,PQ} = \begin{cases} 
    s_1 L_{SDR,PQ} + a_1 & \text{if } X_{min} \leq L_{SDR,PQ} \leq X_1 \\
    s_2 L_{SDR,PQ} + a_2 & \text{if } X_1 < L_{SDR,PQ} \leq X_2 \\
    s_3 L_{SDR,PQ} + a_3 & \text{if } X_2 < L_{SDR,PQ} \leq X_{max} 
\end{cases} \quad (4.6) \]

where \( L_{SDR,PQ} \) and \( L_{HDR,PQ} \) are SDR and HDR PQ values respectively. \( s_1, s_2, \) and \( s_3 \) are the slopes assigned to each of our three brightness regions. \( a_1, a_2, \) and \( a_3 \) are the \( y \)-intercepts associated with each segment of the mapping curve. \( X_{min} \) and \( X_{max} \) are the minimum and maximum PQ values in the SDR domain respectively. \( X_1 \) is the boundary between dark and normal regions and \( X_2 \) is the boundary between normal and bright regions. Figure 4.6(a) shows the SDR frame in Figure 4.2(a) while Figure 4.6(b) shows our mapping curve associated with this frame. This curve maps the SDR PQ values in the range \([0.0623 \text{ – } 0.5081]\) (or \([0.1 \text{ – } 100]\) nits) to \([0.0215 \text{ – } 0.7518]\) (or \([0.01 \text{ – } 1000]\) nits) in HDR domain. It is important to point out that since the brightness range of HDR displays may vary, our proposed iTMO is capable of dealing with HDR displays with arbitrary brightness ranges.

To find the slope of each segment of our mapping curve, we use the hybrid approach introduced in Chapter 3, where we maximized a weighted sum of the brightness and contrast difference functions between input SDR and generated HDR frame. Figures 4.7(a), (b), and (c) represent a dark, normal, and bright SDR frames, respectively, while Figures 4.7(d), (e), and (f) depict the generated HDR frames using our proposed iTMO. All HDR frames are tone mapped.
Figure 4.6 (a) An SDR frame. (b) Our mapping curve associated with it. $X_1$ and $X_2$ are the thresholds found using our segmentation method and $Y_1$ and $Y_2$ are their corresponding PQ values in HDR domain. Using Reinhard et al.’s tone mapping operator [37] for visualization purposes. As it can be
observed, our proposed iTMO achieves a good balance between overall brightness and contrast of
the frame and preserves the overall impression of the input SDR content.

4.1.4 Proposed Color Adjustment Method

An important consideration in designing an efficient iTMO is maintaining the color accuracy between the input SDR and the generated HDR frames. This is a challenge, because as we try to expand the brightness levels from SDR to HDR, we may cause color shifts [30], [31]. Thus, an efficient color adjustment scheme is of high importance when designing iTMOs.

One common approach for color adjustment in SDR to HDR mapping is proposed by Schlick et al. [62]. This method performs the color adjustment in the light domain and the RGB color space, aiming at preserving the hue of the input SDR colors, as follows:

\[ C_{HDR} = \left( \frac{C_{SDR}}{L_{SDR}} \right)^{\gamma} L_{HDR} \] (4.7)
where $C_{HDR}$ denotes one of the color channels (red, green, or blue) of the generated HDR frame, $C_{SDR}$ represents the same color channel in the input SDR frame, $L_{HDR}$ and $L_{SDR}$ are the HDR, and SDR luminance channels, respectively and $s$ is a saturation factor which is greater than zero.

We wish to examine the performance of equation (4.7) in terms of its effect on hue, chroma, and lightness of colors. To this end, we first convert the RGB values to XYZ and then the IPT color space as it has a better hue uniformity than Lab color space [63]. As we need to examine the changes in hue, chroma, and lightness, we convert IPT to LCh [64] color space as it has all the three variables we aim to examine. Figure 4.8 (a) shows the changes in hue ($h$) with respect to Chroma ($C$) as we change the saturation factor $s$ in equation (4.7) for six basic colors. As it can be seen, changing $s$ results in changes in the hue of colors, especially for red and blue. Figure 4.8 (b) presents the changes in Lightness ($L$) with respect to Chroma ($C$) as we change saturation factor $s$. 

![Figure 4.8](image-url)
We observe that the overall brightness of colors first decreases and then starts to increase as $s$ increases.

Another common approach for color adjustment in SDR to HDR mapping is proposed by Mantiuk et al. [30] where the color adjustment is performed in the light domain and the $RGB$ color space with the main aim of limiting the resulting luminance changes, as follows:

$$C_{HDR} = \left( \frac{C_{SDR}}{L_{SDR}} - 1 \right) s + 1 \right) L_{HDR} \tag{4.8}$$

where $C_{HDR}$ denotes one of the color channels (red, green, or blue) of the generated HDR frame, $C_{SDR}$ represents the same color channel in the input SDR frame, $L_{HDR}$ and $L_{SDR}$ are the HDR, and SDR luminance channels, respectively and $s$ is a saturation factor which is greater than zero.

Figure 4.9 (a) shows the changes in hue ($h$) for six basic colors with respect to Chroma ($C$) as the saturation factor $s$ in equation (4.8) changes. As it can be seen, changing $s$ results in changes
in the hue of colors, especially for red and blue. Figure 4.9 (b) presents the changes in Lightness ($L$) with respect to Chroma ($C$) as we change saturation factor $s$ in equation (4.8). As it can be observed, the overall brightness of colors has a near-constant behavior with $s$.

To address the issue of hue shift caused by Mantiuk et al.’s and Schlick et al.’s operators and achieve better color accuracy between input SDR and output HDR frames, we propose to perform the color adjustment in the PQ domain as follows:

$$C_{HDR,PQ} = s \frac{L_{HDR,PQ}}{L_{SDR,PQ}} C_{SDR,PQ}$$

(4.9)

where $C_{HDR,PQ}$ denotes one of the color channels (red, green, or blue) of the generated HDR frame in the PQ domain, respectively. $C_{SDR,PQ}$ represents the same color channel in the SDR frame in the PQ domain. $L_{HDR,PQ}$ and $L_{SDR,PQ}$ are the HDR and SDR luminance channels in the PQ domain, respectively and $s$ is a saturation factor which is greater than zero.

Figure 4.10 (a) shows the changes in hue ($h$) with respect to Chroma ($C$) as we change the saturation factor $s$ in equation (4.9) for six basic colors. We can see that changing $s$ does not affect the hue of colors and it stays constant regardless of the changes in $s$. In addition, Figure 4.10 (b) shows linear changes in Lightness ($L$) with respect to Chroma ($C$) for different values of $s$, meaning that the colors will appear brighter as we increase the saturation factor. In Figure 4.10, points where $s = 1$ are marked with black squares.

Comparing Figures 4.8 (a), 4.9 (a), and 4.10 (a) demonstrates that our proposed color adjustment method preserves the hue of colors irrespective of the changes in $s$ as opposed to Schlick et al.’s and Mantiuk et al.’s methods. Inspection of Figures 4.8 (b), 4.9 (b), and 4.10 (b) shows that Mantiuk et al.’s method does a better job in limiting changes in color Lightness ($L$).
compared to our proposed and Schlick et al.’s methods, with our proposed method showing a linear behavior.

Our next step is to identify a suitable value for the saturation factor $s$ that yields the smallest change in the HDR luminance channel for our color adjustment method. To this end, we conducted an experiment where we randomly selected 400 SDR frames from [36] and generated their corresponding HDR frames using our proposed iTMO coupled with the proposed color adjustment method, using three different saturation factors: 1) $s = 0.8$, 2) $s = 1$, and 3) $s = 1.2$, resulting in $400 \times 3 = 1200$ pairs of SDR and HDR frames. Our dataset included a broad range of color hues and saturation levels representing human faces, indoor, and outdoor scenes. Figure 4.11 shows a small subset of our dataset that includes *Fall, Leaves, Fruits, Vase, Umbrella, Holi, Rose, Parrot, Pencils*, and *Dance*. For each pair, we compute the absolute difference between the luminance channel of the generated HDR frame before and after applying the color adjustment method ($\Delta L$).
An efficient color adjustment method should not affect the overall HDR luminance ($\Delta L \approx 0$). Table 4.1 presents the mean, standard deviation (SD) and the 95% confidence interval (CI) for the average values of $\Delta L$ for each pair of SDR and HDR frames and for three different saturation factors. We observe from Table 4.1 that when $s = 1$, our proposed color adjustment yields on average the lowest change in $\Delta L$. Figure 4.12 is a graphical representation of results in Table 4.1.

Given the above findings, we would like to examine the overall performance of our proposed color adjustment method for the extreme luminance changes generated by inverse tone mapping. To this end, we compared the performance of our proposed method against the methods of Schlick et al.’s and Mantiuk et al.’s for Kim et al.’s [10] and our iTMO in this chapter. For this test, we used the same dataset as our previous test (Figure 4.11), which includes 400 SDR frames randomly chosen from [36]. Similar to our previous test, for each pair we compute the absolute difference between the luminance channel of the generated HDR frame before and after applying the color adjustment methods ($\Delta L$). Table 4.2 presents the mean, SD and the 95% CI for the average values of $\Delta L$ for each SDR and HDR pair using our proposed iTMO with three different color adjustment methods. Similarly, Table 4.3 presents the results for Kim et al.’s iTMO. Figures 4.13 and 4.14 are the graphical representation of the results in Tables 4.2 and 4.3, respectively. We
observe that our proposed method and Mantiuk et al.’s color adjustment method yield negligible changes in HDR luminance irrespective of the iTMO used. In summary, our proposed color adjustment method yields negligible changes in luminance while, unlike other methods, preserves hue.

### 4.2 Experimental Evaluations

In sub-section 4.1.2, we claimed that our proposed entropy-based segmentation method prevents flickering due to its cumulative approach in calculating Entropy. Here, we evaluate our iTMO’s performance regarding flickering. We chose 120 consecutive frames of an SDR video (24 fps) sequence known to cause flickering for most iTMOS. This sequence is converted to HDR

<table>
<thead>
<tr>
<th>$\Delta L$</th>
<th>$s = 0.8$</th>
<th>$s = 1$</th>
<th>$s = 1.2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0889</td>
<td>0.0080</td>
<td>0.0761</td>
</tr>
<tr>
<td>SD</td>
<td>0.0363</td>
<td>0.0131</td>
<td>0.0351</td>
</tr>
<tr>
<td>95% CI</td>
<td>[0.0853, 0.0925]</td>
<td>[0.0067, 0.0093]</td>
<td>[0.0727, 0.0795]</td>
</tr>
</tbody>
</table>

**Figure 4.12** Average of $\Delta L$ along with its 95% CI over our dataset of 400 SDR frames for three different values of $s$. 
Table 4.2 Mean, SD, and 95% CI of the average ΔL using our proposed iTMO with three different color adjustment methods

<table>
<thead>
<tr>
<th>ΔL</th>
<th>Proposed</th>
<th>Mantiuk</th>
<th>Schlick</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.00808</td>
<td>1.9×10⁻¹⁴</td>
<td>0.53034</td>
</tr>
<tr>
<td>SD</td>
<td>0.013134</td>
<td>2.1×10⁻¹⁴</td>
<td>0.77058</td>
</tr>
<tr>
<td>95% CI</td>
<td>[0.006,0.009]</td>
<td>[1.7×10⁻¹⁴, 2.1×10⁻¹⁴]</td>
<td>[0.4548,0.6059]</td>
</tr>
</tbody>
</table>

Figure 4.13 Average of ΔL along with its 95% CI over our dataset of 400 SDR frames using our proposed iTMO for three different color adjustment methods.

For both generated HDR sequences we calculated the absolute difference for Y₁ and Y₂ (see Figure 4.15) for all 120 frames in each sequence as follows:

\[ Diff_{Y_1,i} = \left| Y_{1,F_{i+1}} - Y_{1,F_i} \right|, \quad i = 1,2,...,119 \] (4.10)

\[ Diff_{Y_2,i} = \left| Y_{2,F_{i+1}} - Y_{2,F_i} \right|, \quad i = 1,2,...,119 \] (4.11)

where \( Diff_{Y_1} \) and \( Diff_{Y_2} \) are the absolute difference for \( Y_1 \) and \( Y_2 \), respectively, and \( F_i \) denotes the \( i \)-th frame in the sequence.
Table 4.3 Mean, SD, and 95% CI of the average $\Delta L$ using Kim et al.’s iTMO with three different color adjustment methods

<table>
<thead>
<tr>
<th>$\Delta L$</th>
<th>Proposed</th>
<th>Mantiuk</th>
<th>Schlick</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0079</td>
<td>$1.5 \times 10^{-14}$</td>
<td>0.4261</td>
</tr>
<tr>
<td>SD</td>
<td>0.0130</td>
<td>$1.6 \times 10^{-14}$</td>
<td>0.6295</td>
</tr>
<tr>
<td>95% CI</td>
<td>[0.0068, 0.009]</td>
<td>[1.3 $\times 10^{-14}$, 1.6 $\times 10^{-14}$]</td>
<td>[0.364, 0.487]</td>
</tr>
</tbody>
</table>

Figure 4.14 Average of $\Delta L$ along with its 95% CI over our dataset of 400 SDR frames using Kim et al.’s iTMO for three different color adjustment methods.

Figure 4.15 An example of our mapping curve. $X_1$ and $X_2$ are the boundaries for dark, and bright regions, respectively while $Y_1$ and $Y_2$ are their corresponding PQ values in HDR domain, respectively.
Flickering happens when a noticeable brightness variation occurs between two consecutive frames, meaning the brightness difference between the two consecutive frames exceeds the Just Noticeable Difference (JND) threshold [46]:

\[ \text{Diff}_i^Y > \text{JND}, \quad i = 1, 2 \]  

(4.12)

Figures 4.16 (a) and (b) show \( \text{Diff}_1^Y \) and \( \text{Diff}_2^Y \) when we employ the entropy-based segmentation method in Chapter 2, respectively. Figures 4.16 (c) and (d) show \( \text{Diff}_1^Y \) and \( \text{Diff}_2^Y \) when we use the entropy-based segmentation method presented in this chapter, respectively. The horizontal red-dotted line in Figures 4.16 (a) – (d) denotes the JND value [46]. From Figures 4.16 (a) and (c) we observe that the absolute difference for \( Y_1 \) does not exceed the JND value, meaning that our eyes will not notice these changes. However, Figure 4.16 (b) shows that the entropy-based
segmentation in Chapter 2 causes flickering, as the absolute difference for $Y_2$ passes the JND threshold. On the other hand, Figure 4.16 (d) demonstrates that our proposed iTMO coupled with the new entropy-based segmentation prevents flickering, as the changes fall well below the JND threshold.

We evaluated the visual quality of the generated HDR videos using our proposed iTMO against Endo et al.’s, Luzardo et al.’s, Bist et al.’s, iTMO in Chapter 3, and Kim et al.’s iTMO through a subjective test. For all of the iTMOs, we used their default parameters as mentioned in their respective papers, except the maximum brightness that was set to 1000 nits. For the Endo et al.’s iTMO, we used the open source code available in [65]. We implemented Luzardo et al.’s, and Bist et al.’s iTMOs based on the description in their respective papers. Authors of [10] generously provided us with the implementation of their iTMO.

In this test, we asked our subjects to compare the HDR videos generated by our proposed iTMO against other methods and choose the one that they believe looks more visually pleasing to them. For this test, we used nine 4K SDR videos [59] (each ten seconds long) representing dark, normal, and bright scenes to evaluate the performance of our iTMO against others under different brightness conditions. Figure 4.17 depicts one frame of each of the nine 4K SDR videos used in

**Figure 4.17** One frame of each of the nine 4K SDR videos used in our subjective test.
our subjective test, which include Desert, Mall, Boat, Paris, Crowd, Performance, Wall, Bridge, and Brunch. Eighteen subjects including 12 males, and 6 females participated in our test. Subjects were between twenty-two and thirty-three years old. All subjects were screened for color blindness and visual acuity using the Ishihara and Snellen charts, respectively. Prior to our test, we performed a training session to familiarize subjects with the test procedure. Two HDR videos were shown on two identical Sony BVM-X300 professional monitors. All different combinations of the test videos were randomly included in our subjective test resulting in 45 HDR videos that were evaluated by the subjects. During the test, subjects were asked to choose one of the two HDR videos (“A” or “B”) that looked more visually pleasing. We also provided the subjects with the “equal” choice in case the two HDR videos appeared to have the same level of visual quality. The position of HDR videos were randomly changed between the left and right displays to remove any effect caused by differences between the two displays and their physical location. The subjective test results were analyzed based on the methodology proposed in [53]. We used the outlier detection methodology in [53] and found two outliers and excluded their results from our analysis.

Figure 4.18 depicts the average of times subjects preferred the visual quality of the HDR videos generated by our proposed iTMO against Endo et al.’s, Luzardo et al.’s, Bist et al.’s, iTMO in Chapter 3, and Kim et al.’s operator. We observe that, on average, our iTMO outperforms Endo et al.’s by 85% , Luzardo et al.’s by 94%, Bist et al.’s by 85%, iTMO in Chapter 3 by 67%, and Kim et al.’s by 77% of the times in terms of generating HDR videos that are more visually pleasing to the viewers. We further analyzed the MOS scores of the six iTMOs in this subjective based on F-value and p-value of one-way analysis of variance (ANOVA) to determine whether there any
To do this, we performed the Levene’s test \[61\] to guarantee the homogeneity of variances of the gathered data and we concluded that one-way ANOVA can be applied. Table 4.4 presents the results of the one-way ANOVA test including the mean and standard deviation (SD) of the MOS scores for each approach. The ANOVA test

\[
\begin{array}{|l|l|l|}
\hline
\text{iTMO} & \text{Mean} & \text{SD} \\
\hline
\text{Endo et al.’s} & 1.1875 & 1.1672 \\
\text{Luzardo et al.’s} & 0.5 & 0.7302 \\
\text{Bist et al.’s} & 1.25 & 1 \\
\text{iTMO in Chapter 3} & 2.4375 & 1.5478 \\
\text{Kim et al.’s} & 1.875 & 1.2041 \\
\text{Proposed} & 35.0625 & 3.2139 \\
\hline
\end{array}
\]

\[F\text{-value (p-value)} = 6.8270 (1.9290 \times 10^{-5})\]

Figure 4.18 Preference scores of the subjects in mean opinion score (MOS) when comparing the visual quality of the HDR videos generated by our iTMO over that of Endo et al.’s, Luzardo et al.’s, Bist et al.’s, iTMO in Chapter 3, and Kim et al.’s iTMOS.
concluded that the null hypothesis (i.e., there are no statistical significance between the six iTMOs in this subjective test) probability ($p$-value) for the dataset in this subjective test is $1.92 \times 10^{-5}$. Since the $p$-value is less than the significance level $\alpha=0.05$ we reject the null hypothesis and conclude that the difference in MOS values obtained by the six iTMOs in this subjective test is statistically significant.

To evaluate the MOS-based performance difference between any two of the six iTMOs, we performed Tukey’s post-hoc analysis for their MOS scores. Figure 4.19 shows the results of this test when $\alpha=0.05$. Methods in the same group are the ones with statistically insignificant difference. As it can be seen, Endo et al.’s, Luzardo et al.’s, Bist et al.’s, iTMO in Chapter 3, and Kim et al.’s fall within the same group while our proposed iTMO falls in a separate one, meaning our iTMO yield statistically distinct HDR videos in terms of subjective quality assessment.

We also compared the performance of our proposed iTMO against Endo et al.’s, Luzardo et al.’s, Bist et al.’s, iTMO in Chapter 3, and Kim et al.’s iTMOs using HDR VDP-2.2 and PUSSIM objective quality metrics. To this end, we used nine original HDR frames randomly chosen from the dataset in [58] representing dark, normal, and bright scenes. Figure 4.20 shows the tone mapped version of the nine HDR frames which include View, Arch, Sunlight, Desk, Frontier, Building, Office, Hut, and Stadium. As can be seen, our dataset includes dark, normal, and bright scenes to evaluate the performance of our iTMO against others under different brightness conditions. We tone mapped these HDR frames using Reinhard et al.’s tone mapping operator [37]
and used the generated SDR frames as input for the six iTMOs. Finally, we evaluated the quality of the generated HDR frames with respect to the original HDR frame using HDR VDP-2.2 and PU-SSIM. Tables 4.5, 4.6, and 4.7 show the results of HDR VDP-2.2 for the six iTMOs for dark (Desk, Frontier, Building), normal (Office, Hut, Stadium), and bright (View, Arch, Sunlight) frames, respectively. In addition, Tables 4.8, 4.9, and 4.10 show the results of PU-SSIM for the six iTMOs for dark, normal, and bright frames, respectively. As it can be seen from the results of both metrics, on average, our proposed iTMO achieves the highest visual quality for all types of SDR frames, reaffirming the result of our subjective test.

We also evaluated the performance of our proposed iTMO against Endo et al.’s [9], Luzardo et al.’s [8], Bist et al.’s [7], the iTMO in Chapter 3, and Kim et al.’s [10] iTMO in terms of maintaining the color accuracy between the input SDR and generated HDR frames by means of another subjective test. For all these iTMOs, we used their default parameters as mentioned in their respective papers, except the maximum brightness that was set to 1000 nits. For the Endo et al.’s iTMO, we used the open source code available in [65]. We implemented Luzardo et al.’s, and Bist et al.’s iTMOs based on the description in their respective papers. Authors of [10] generously
provided us with the implementation of their iTMO. Our proposed iTMO uses the proposed color adjustment method in this chapter, while Luzardo et al.’s, Bist et al.’s, Endo et al.’s, and the iTMO in Chapter 3 all use Mantiuk et al.’s color adjustment method. In addition, Kim et al.’s iTMO uses Schlick et al.’s color adjustment method.

### Table 4.5 Visual quality scores by HDR VDP-2.2 for dark frames in Figure 4.20 for six different iTMOS

<table>
<thead>
<tr>
<th>Frame</th>
<th>Proposed</th>
<th>iTMO in Chapter 3</th>
<th>Kim</th>
<th>Endo</th>
<th>Bist</th>
<th>Luzardo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desk</td>
<td>52.2767</td>
<td>49.58417</td>
<td>46.748792</td>
<td>45.866851</td>
<td>49.49307</td>
<td>52.69622</td>
</tr>
<tr>
<td>Frontier</td>
<td>47.67912</td>
<td>45.002</td>
<td>42.651174</td>
<td>42.721272</td>
<td>44.87382</td>
<td>45.7905</td>
</tr>
<tr>
<td>Building</td>
<td>44.78351</td>
<td>41.23503</td>
<td>39.487692</td>
<td>37.408875</td>
<td>41.15466</td>
<td>44.5629</td>
</tr>
<tr>
<td>Average</td>
<td><strong>48.24644</strong></td>
<td>45.27373</td>
<td>42.962552</td>
<td>41.999</td>
<td>45.17385</td>
<td>47.68321</td>
</tr>
</tbody>
</table>

### Table 4.6 Visual quality scores by HDR VDP-2.2 for normal frames in Figure 4.20 for six different iTMOS

<table>
<thead>
<tr>
<th>Frame</th>
<th>Proposed</th>
<th>iTMO in Chapter 3</th>
<th>Kim</th>
<th>Endo</th>
<th>Bist</th>
<th>Luzardo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Office</td>
<td>47.03192</td>
<td>43.7328</td>
<td>41.546788</td>
<td>39.761681</td>
<td>43.33167</td>
<td>43.41994</td>
</tr>
<tr>
<td>Hut</td>
<td>47.87282</td>
<td>45.49702</td>
<td>43.632633</td>
<td>44.234949</td>
<td>43.51107</td>
<td>45.96083</td>
</tr>
<tr>
<td>Stadium</td>
<td>52.49312</td>
<td>49.69019</td>
<td>47.158808</td>
<td>47.541001</td>
<td>47.79539</td>
<td>50.22695</td>
</tr>
<tr>
<td>Average</td>
<td><strong>49.13262</strong></td>
<td>46.30667</td>
<td>44.112743</td>
<td>43.845877</td>
<td>44.87938</td>
<td>46.53591</td>
</tr>
</tbody>
</table>

### Table 4.7 Visual quality scores by HDR VDP-2.2 for bright frames in Figure 4.20 for six different iTMOS

<table>
<thead>
<tr>
<th>Frame</th>
<th>Proposed</th>
<th>iTMO in Chapter 3</th>
<th>Kim</th>
<th>Endo</th>
<th>Bist</th>
<th>Luzardo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lake</td>
<td>47.52308</td>
<td>45.5694</td>
<td>43.640389</td>
<td>45.751585</td>
<td>43.74832</td>
<td>43.93695</td>
</tr>
<tr>
<td>Arch</td>
<td>47.88094</td>
<td>47.32708</td>
<td>44.844326</td>
<td>46.533</td>
<td>44.06702</td>
<td>51.2192</td>
</tr>
<tr>
<td>Sunlight</td>
<td>47.88069</td>
<td>44.45322</td>
<td>41.976296</td>
<td>44.6211</td>
<td>43.90446</td>
<td>43.03713</td>
</tr>
<tr>
<td>Average</td>
<td><strong>47.76157</strong></td>
<td>45.7832</td>
<td>43.487004</td>
<td>45.6352</td>
<td>43.9066</td>
<td>46.06443</td>
</tr>
</tbody>
</table>
In this test, we asked our subjects to compare the HDR frames generated by our proposed iTMO against other state-of-the-art methods and choose the one that they think is closer to the original SDR frame in terms of its colors. For this test, we used ten SDR frames randomly chosen from [36] covering a wide range of color hues and saturation levels representing human faces, indoor, and outdoor scenes. Figure 4.21 shows the dataset used in this subjective test which
includes Powder, Autumn, Mix, Pepper, Lake, Legos, Laughter, Bird, Park, and Jenga. Eighteen subjects (4 females and 14 males) between the ages of twenty-three and thirty-four participated in our test. The subjects who participated in this test were different than the ones in our previous subjective test. All subjects were screened for color blindness and visual acuity using the Ishihara and Snellen charts, respectively. Prior to the test, subject were familiarized with the test procedure through a training session. The original SDR frames were shown on a Sony BVM-X300 professional monitor [60], while the HDR frames were shown on another identical Sony BVM-X300 professional monitor in a side-by-side manner, using a 14-pixel wide black stipe between them. All different combinations of the test frame pairs were included randomly in the experiment, resulting in 50 overall frames that were evaluated by the subjects. Each SDR-HDR pair was shown for 10 seconds on each display.

During the test, subjects were asked to choose one of the two HDR frames (“A” or “B”) that they think its colors more closely represent the ones in the original SDR frame. Subjects were also provided with the “equal” choice in case the colors in both HDR frames appeared equally close to the ones in the original SDR frame.
The position of HDR frames were randomly switched between the left and right side of the display to remove any effect caused by any variations of brightness across the display. Analysis of the results followed the process described in [53]. We used the outlier detection methodology in [53] and found two outliers and excluded their results from our analysis.

Figure 4.22 shows the percentage of how many times, on average, subjects decided that the colors generated by our proposed iTMO more closely represent the ones in the original SDR frame over that of Endo et al.’s, Luzardo et al.’s, Bist et al.’s, the iTMO in Chapter 3, and Kim et al.’s operators.

The position of HDR frames were randomly switched between the left and right side of the display to remove any effect caused by any variations of brightness across the display. Analysis of the results followed the process described in [53]. We used the outlier detection methodology in [53] and found two outliers and excluded their results from our analysis.

Figure 4.22 shows the percentage of how many times, on average, subjects decided that the colors generated by our proposed iTMO more closely represent the ones in the original SDR frame over that of Endo et al.’s, Luzardo et al.’s, Bist et al.’s, the iTMO in Chapter 3, and Kim et al.’s iTMOs. As it can be observed, on average, our proposed iTMO outperforms Endo et al.’s by 79%, Luzardo et al.’s by 68%, Bist et al.’s by 96%, the iTMO in Chapter 3 by 72% and Kim et al.’s by 65% of the times in terms of how closely the colors of the output HDR frame represent the ones in the original SDR frame. We further analyzed the MOS scores of the six iTMOs in this subjective based on $F$-value and $p$-value of one-way analysis of variance (ANOVA) to determine
Table 4.11 Mean and SD of MOS scores for the six iTMOs based on the dataset in Figure 4.21

<table>
<thead>
<tr>
<th>iTMO</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endo et al.’s</td>
<td>1.9375</td>
<td>0.9287</td>
</tr>
<tr>
<td>Luzardo et al.’s</td>
<td>3</td>
<td>0.9660</td>
</tr>
<tr>
<td>Bist et al.’s</td>
<td>0.3125</td>
<td>0.6020</td>
</tr>
<tr>
<td>iTMO in Chapter 3</td>
<td>2.5625</td>
<td>1.3149</td>
</tr>
<tr>
<td>Kim et al.’s</td>
<td>3.3125</td>
<td>0.9464</td>
</tr>
<tr>
<td>Proposed</td>
<td>36.625</td>
<td>3.7925</td>
</tr>
</tbody>
</table>

F-value (p-value) 7.7570 (4.1840×10^-6)

Figure 4.23 Post-hoc analysis using Tukey’s method for mean values of MOS scores and α=0.05. Methods within the same group generate statistically insignificant HDR output.

whether there any statistical significance between any of the iTMOs. To do this, we performed the Levene’s test [61] to guarantee the homogeneity of variances of the gathered data and we concluded that one-way ANOVA can be applied. Table 4.11 presents the results of the one-way ANOVA test including the mean and SD of the MOS scores for each approach. The ANOVA test concluded that the null hypothesis (i.e., there is no statistical significance between the six iTMOs in this subjective test) probability (p-value) for the dataset in this subjective test is 4.184×10^-6. Since the p-value is less than the significance level α=0.05 we reject the null hypothesis and conclude that the difference in MOS values obtained by the six iTMOs in this subjective test is statistically significant.

To evaluate the MOS-based performance difference between any two of the six iTMOs, we performed Tukey’s post-hoc analysis for their MOS scores. Figure 4.23 shows the results of
this test when $\alpha=0.05$. Methods in the same group are the ones with statistically insignificant difference. As it can be seen, Endo et al.’s, Luzardo et al.’s, Bist et al.’s, the iTMO in Chapter 3, and Kim et al.’s iTMOs fall within the same group while our proposed iTMO falls in a separate one, meaning that it yields statistically distinct HDR colors in terms of subjective quality assessment.

4.3 Conclusions

In this Chapter, we proposed a new content adaptive iTMO designed for video applications that can effectively deal with all types of SDR content from dark, to normal and bright scenes. Our iTMO works in the perceptual domain, allowing mapping to take advantage of the human eye’s sensitivity to brightness changes in different regions of the scene. A segmentation method that is based on maximum entropy thresholding is used to effectively segment the frame into dark, normal, and bright regions in order to better control the overall impression of each region and maintain the artistic intent, while also preventing flickering. The challenge of color shifting in inverse tone mapping is addressed by the introduction of a perception-based color adjustment method that preserves the hue of colors with insignificant changes in brightness, producing HDR colors that are faithful to their SDR counterparts. Subjective and objective evaluations showed that our proposed iTMO outperforms other state-of-the-art operators in terms of overall visual quality of the generated HDR video, while producing HDR colors that are faithful to their SDR counterparts.
Chapter 5: Conclusions and Future Work

5.1 Summary of the Contributions

In this thesis, we proposed novel iTMOs capable of generating high visual quality HDR content. We also proposed a novel perception-based color adjustment method that prevents color shifts in SDR to HDR mapping and is capable of generating HDR colors that are faithful to the original SDR counterparts. Our subjective and objective evaluations demonstrate that our proposed iTMOs outperform other state-of-the-art methods in terms of overall visual quality and maintaining the color accuracy between input SDR and generated HDR content.

In Chapter 2, we proposed an entropy-based iTMO that can deal with all types of SDR content, generating high visual quality HDR frames. Working in perceptual domain allowed us to model the sensitivity of our eyes to brightness changes in different regions of a scene. To keep the dark areas dark and bright areas bright without darkening and brightening up the frame, we proposed an entropy-based segmentation method capable of dividing the SDR frame into dark, normal, and bright regions. Then, we employed a novel piece-wise linear mapping curve with two breaking points to map the three brightness regions to the HDR domain. The slopes for each segment of the mapping curve were chosen based on comprehensive subjective tests. Our subjective evaluations demonstrated that our proposed iTMO outperformed other state-of-the-art methods in terms of generating high visual quality HDR images. Despite its good performance, this proposed iTMO is only capable of generating 4000 nits HDR frames and is not completely content adaptive as the slopes of the mapping curve were fixed for all types of SDR content. Another limitation is that the application of this iTMO is limited to SDR images, as the proposed segmentation method will cause flickering in video applications due to its high sensitivity to small changes in pixels distribution.
In Chapter 3, we proposed a perceptually optimized iTMO that addressed the shortcomings of the method in Chapter 2 while generating high visual quality HDR videos for any arbitrary dynamic range supported by the target HDR display. To address the high sensitivity of the segmentation method proposed in Chapter 2 to small variations in pixel distribution, we proposed fixed boundaries for the dark, normal, and bright regions of the SDR frame. Similar to Chapter 2, a piece-wise linear mapping curve with two breaking points was used to convert SDR video to HDR format. However, to overcome the issue of having fixed slopes for all types of SDR content, we calculated the slope of each segment of the mapping curve using a hybrid approach that optimizes the trade-off between contrast and brightness in the generated HDR frame. Our subjective and objective evaluations demonstrated that this proposed iTMO outperforms other state-of-the-art iTMOs in terms of generating high visual quality HDR videos. Despite that, the proposed iTMO is not fully content adaptive, as it uses fixed boundaries for detecting the dark, normal, and bright regions of an SDR frame.

In Chapter 4, we proposed a fully content adaptive iTMO that addressed the shortcomings of the methods in Chapters 2, and 3 while generating high visual quality HDR videos for any dynamic range supported by the target HDR display. To overcome the issue of non-content adaptive limitation of the iTMO in Chapter 3 that results from using fixed boundaries for detecting dark, normal, and bright regions, we proposed a novel entropy-based segmentation method. We also demonstrated that unlike the segmentation method in Chapter 2, this cumulative segmentation is less sensitive to small changes in pixels distribution and prevents flickering. Similar to Chapters 2 and 3, a piece-wise linear mapping curve with two breaking points was used to convert SDR video to HDR format. The slope of each segment of the mapping curve is optimized by finding the best trade-off between contrast and brightness throughout the entire dynamic range. We also
addressed the issue of color shifts in SDR to HDR mapping by introducing a perception-based color adjustment method that preserves the color accuracy between input SDR and generated HDR frames that maintain the hue levels while yielding negligible changes in luminance. Our subjective and objective evaluations demonstrated that our proposed iTMO outperforms other state-of-the-art methods in terms of generating high visual quality HDR videos while also producing HDR colors that closely follow their SDR counterparts.

5.2 Significance and Potential Applications of the Research

While HDR video technology has a disruptive impact on the capturing and display industries, it also opens new opportunities for SDR content owners as well as broadcasters that will continue producing real-time events in SDR format for the foreseeable future. For the latter case, SDR content will have to be efficiently converted to HDR format, taking advantage of the advanced visual quality of HDR displays to offer viewers a much better quality of experience than what the original SDR content would ever do. In fact, iTMOs have widespread applications including but not limited to TV manufacturing, broadcasting, video gaming, security, HDR cinema and post-production. The proposed iTMO is an ideal solution, as it is capable of generating high visual quality HDR videos, surpassing in performance every other existing approach. It is worth noting that the work presented in this thesis has been patented and is already being employed by TELUS Communications Inc. to convert their existing SDR content to HDR format.

In the case of video gaming, there has been a growing interest in the market towards HDR streaming. Recently, gaming consoles such as Microsoft’s Xbox One S [66], and Sony’s PS4 Pro [67] support HDR streaming. However, although there exist several HDR monitors such as Samsung KS9800 [68], LG C8 [69], and Sony X800E [70], not all video games are or will be in HDR format. Hence, the proposed iTMOs can be implemented in the game engines or the HDR
monitors to convert existing SDR gaming content to HDR format, offering a higher visual quality and wider range of color and brightness.

In the case of security and surveillance, there has been a growing demand for HDR support as it will drastically improve contrast, improve details in the dark regions and even prevent images from being washed out by harsh sunlight. Therefore, our proposed iTMOs in this thesis can be implemented into the security camera hardware to offer higher visual quality, and visible details in dark and bright regions of the scene.

Another application for the research presented in this thesis is HDR cinema. HDR technology has not yet reached the cinema industry, and movies are still in the SDR format with the maximum brightness of 48 nits [71]. SDR cinema content quality has been widely perceived as better than SDR Home image quality, yet HDR cinema has not yet been established while HDR home content is becoming widely available [71]. Our proposed method can be implemented for converting legacy cinema SDR content to HDR format for theater levels of HDR, providing cinemagoers with life-like viewing experience.

Another important application of the research presented in this thesis is post-production. To enable a more pleasing and immersive viewing experience in the home and cinema applications a range of post-production operations are adopted by content owners. In post-production, color grading will happen manually on HDR display with fixed capabilities. The process of color grading varies depending on the application. For example, color grading for cinema and home viewing requires a separate process since the viewing environment in cinema is dark, while home environment is usually dim. In addition to the viewing environment, capabilities of home displays vary widely. This means that the combination of target display and target viewing environment is unlikely to match the grading environment produced by post-production studios. The manual
process is extremely time consuming. Tests with many studios have shown that our method outperforms manual approaches and TELUS has already replaced such internal process with our iTMO solution.

5.3 Future Work

Future work may include extension of the final proposed iTMO described in Chapter 4 to Light Field (LF) technology and more specifically conversion of SDR LF content to HDR format in order to increase its overall visual quality. Although LF technology allows capturing richer visual information than traditional cameras by capturing light intensity and its direction [72], existing LF capturing technologies are limited to SDR format. Our proposed iTMO can help adjust the overall brightness and contrast along with color information of the light field content for a number of applications such as entertainment, autonomous driving, robotics, human monitoring and endoscopy.

Another possible topic for further research is to develop a fully-automatic content adaptive Tone Mapping Operator (TMO) based on the principals we used in designing our iTMO. Tone Mapping is the process of converting the HDR content to SDR format. To preserve the artistic intent of the original HDR content we can divide it into three brightness regions using our proposed entropy-based segmentation in Chapter 4. Furthermore, to map each of the three HDR region we can find the slopes based on minimizing our proposed contrast and brightness difference functions in Chapter 3. We expect that the values of optimization weights $w_1$ and $w_2$ need to be adjusted through objective assessment for Tone Mapping applications.
Bibliography


Mapping Operator with Improved Color Accuracy for High Dynamic Range Applications,”


[46] S. Miller, M. Nezamabadi, and S. Daly, “Perceptual signal coding for more efficient usage


