令和二年度 博士学位論文

北海道大学
HOKKAIDO UNIVERSITY

デジタル画像における局所ヒストグラム平均化を用いたリアルタイムコントラスト調節手法

Real-time Contrast Management in Digital Images using Local Histogram Equalization Filter

情報科学研究院 情報エレクトロニクス専攻

機能通信センシング研究室

Prasoon Ambalathankandy

指導教員 池辺 将之 教授

2020 年 8 月

デジタル画像における局所ヒストグラム平均化を用いたリアルタイムコントラスト調節手法

要旨

本論文では，新しい適応型トーンマッピングアルゴリズムと，そのリアルタイムFPGA実装について報告する．グローバルおよびローカルトーン制御を組み合わせることができるアルゴリズムはいくつか存在するが，これらのトーン制御機能を統一的に管理することは難しいことが知られている．本研究で提案するSLHE（Smoothed Local Histogram Equalization:）法は，同じトーンマッピング空間内のグローバルおよびローカルトーンカーブを容易に管理することができる．側抑制は，多くの知覚効果をもたらす視覚処理における重要な初期段階のメカニズムであり，あらゆる画像処理アルゴリズム設計において考慮される必要がある．我々の詳細な分析から，提案アルゴリズムは，他の最新のアルゴリズムと比較して，自然さを保ちつつ知覚される処理画像を生成できると結論付けた．また，放射線科医による主観的な検証により，このアルゴリズムが医用画像の処理手法として有用であることも実証した．また，提案手法によるパイプライン・ハードウェア処理のデモ動画を，YouTube 上に公開した．

はじめに

高品質・高画質な画像表示への需要の高まりは，ハイダイナミックレンジ（HDR: High Dynamic Range）画像処理の市場を継続的に牽引してきた．HDR イメージングは，新しい4K/8K 超高精細（UHD）フォーマットの魅力的な要素となっている [1]．視聴体験を向上させるために，HDR イメージング技術は，標準的なデジタル画像処理に代わるものとして，様々な分野で採用されている．本論文は，平滑化局所ヒストグラム均等法（Smoothed Local Histogram Equalization: SLHE）に基づくリアルタイム・トーンマッピング・アルゴリズムの分野に貢献する．本論文では，アルゴリズムの設計と FPGA デバイスへの実装について議論した．また，画像品質の評価とアルゴリズムの妥当性を検証するために，複数の詳細な主観評価を実施した．本論文の主な貢献は以下の通りである．

- SLHE ベースのトーンマップ・パイプライン構成の設計

- SLHE アルゴリズムの FPGA 実装によるリアルタイム処理の実証

- SLHE の視覚特性に関する詳細評価と考察

- SLHE アルゴリズムの医用画像強調への応用

SLHE アルゴリズム

LHE（Local Histogram Equalization）を用いた局所トーンマッピングは，局所累積ヒストグラムから構築されたトーンカーブを用いて対象画素を変換するものであり，画像強調のために広く利用されている．しかし，LHE 処理では，元の画像の外観が損なわれてしまうことが多い．そこで，本論文では，図１に示すように，階級値を削減した輝度積算ヒストグラムとシグモイド曲線を基底関数として合成した平滑化 LHE 法を提案する．リップルを生じない滑らかな局所トーンカーブを直接導出する．すべての局所的な中央値または平均値を127の定数値にマッピングする．これは，空間的な低周波成分を完全に除去することができることを意味する．この関数の性質により，除去された低周波成分を再構成すること

で，グローバルな画像特性の制御を可能とする。この再構成によるグローバル制御と、局所累積ヒストグラムによるローカル制御は、トーンマッピング空間の制限を局所関数の利得制御を利用することができる．また，本論文では，提案したSLHEトーンマッピングアルゴリズムの FPGA による実時間実装を行った．今回の FPGA 実装では，前フレームからのダウンサンプリングされた局所統計量を用いることで，システムレイテンシを 0.81μs @162MHz のシステムクロックで最小化し，フレームメモリを 98%削減することに成功した[2].

SLHE アルゴリズムの可視化と応用

側抑制（Lateral inhibition: LI）は，脳が周囲から受け取る膨大かつ冗長な視覚情報の扱いを助ける，視覚処理の重要な初期段階のメカニズムである．この抑制機構は多くの知覚効果をもたらし，アルゴリズムの仕様にも考慮する必要がある．我々は，目の錯覚を伴う画像に対する多層レイヤー(ML)フィルタとSLHEフィルタの応答を比較することで，LI現象とその視覚知覚への影響を詳細に研究してきた[3]．その結果は以下の通りである．

- ML アルゴリズムでは，カーネルサイズが小さく，Base および Detail Layer 取得用のフィルタを主観的にしようしている，そして Detail Layer の強調利得も主観的であるため，不自然な画像を生成している．

- また，最新のエッジ保存フィルタによる各 Layer の抽出でも，処理された画像には補正が困難なマッハバンド効果が発生する．

- ノコギリ波状のエッジは，錯視を伴う自然な知覚画像を生成するために実用的である．

- ノコギリ波状のエッジ形状の生成には大きなカーネルサイズのフィルタが有用であり，SLHE は大きなカーネルをサポートすることができる．

カーネルサイズを変化させた場合の ML 法と SLHE 法の応答を解析した結果，Guided Filter[1] による ML 法では，半径が大きくなるとエッジをテクスチャのように扱う傾向があり，ユーザ設定による強調利得により、エッジ成分が飽和してしまうことがわかった．この影響は Domain Transform Filter[2]では小さくなるが，同様の特性を示す．これらのフィルタ群は周波数選択的であり，ML 法の Detail 操作においては少ないレイヤを使用して特定の周波数に限定している．一方，大きなカーネルをサポートしたSLHEフ法は，エッジ形状を乱すことなく画像を処理することが可能である．我々は視覚の専門分野である放射線科医の協力のもと，医用画像強調のための我々のアルゴリズムの有用性を主観的に評価した[4]．以下に応用例を示す．図２に示す健常者の胸部Ｘ線（画像提供：３）は，左心境にマッハバンド効果がある（マーカーを参照）．この状態は肺気腫と区別する必要がある．図２に SLHE で操作されたＸ線の異なる画像群を示す．心臓境界付近のマッハバンド効果を減少させながら，正常な肺血管系を維持していることを示す最良の画像として評価されている．評価点（５：良好，１：不良）を表１に示す．肺水腫などの偽陽性診断を疑わせる可能性があるため，マッハバンド効果を誇張しないことが非常に重要である．SLHE アルゴリズムは，こ

---

[1] He, Kaiming, Jian Sun, and Xiaoou Tang. "Guided image filtering." *IEEE transactions on pattern analysis and machine intelligence* 35.6 (2012): 1397-1409.
[2] Gastal, Eduardo SL, and Manuel M. Oliveira. "Domain transform for edge-aware image and video processing." *ACM SIGGRAPH 2011 papers*. 2011. 1-12.

れらのマッハバンドを持つ低コントラストＸ線の詳細を向上させることが本論文により明らかになった.

結論

本論文では，新しい適応型トーンマッピングアルゴリズムとそのリアルタイム FPGA 実装についての研究を行った．本研究で提案した SLHE アルゴリズムは，同じトーンマッピング空間内の大域的なトーンカーブと局所的なトーンカーブを容易に管理することが可能である．本研究の結果，提案した SLHE アルゴリズムは，大きなカーネルサイズで動作することにより，ノコギリ波状のエッジを生成することができ，LI 効果によるアーティファクトを回避することができることがわかった．本研究では，Ｘ線画像を用いて SLHE アルゴリズムの有用性を実証し，マッハバンドを含むＸ線画像の処理例を示した．提案処理は，放射線科医から高い評価を得た.

参考文献

[1] Prasoon Ambalathankandy, Shinya Takamaeda, Motomura Masato, Tetsuya Asai, Masayuki Ikebe, and Hotaka Kusano, “Real-time HDTV to 4K and 8K-UHD conversions using anti-aliasing based super resolution algorithm on fpga,” Microprocessors and Microsystems, vol. 61, pp. 21–31, 2018.

[2] Prasoon Ambalathankandy, Masayuki Ikebe, Takayuki Yoshida, Takeshi Shimada, Shinya Takamaeda, Masato Motomura, and Tetsuya Asai, “An adaptive global and local tone mapping algorithm implemented on FPGA,” IEEE Transactions on Circuits and Systems for Video Technology, 2019.

[3] Prasoon Ambalathankandy, Takeshi Shimada, Shinya Takamaeda, Masato Motomura, Tetsuya Asai, and Masayuki Ikebe, “Analysis of smoothed lhe methods for processing images with optical illusions,” in 2018 IEEE Visual Communications and Image Processing (VCIP). IEEE, 2018, pp. 1–4.

[4] Prasoon Ambalathankandy, Yafei Ou, Jyotsna Kochiyil, Shinya Takamaeda, Masato Motomura, Tetsuya Asai, and Masayuki Ikebe, “Radiography contrast enhancement: Smoothed LHE filter a practical solution for digital X-rays with Mach band,” in 2019 Digital Image Computing: Techniques and Applications (DICTA). IEEE, 2019, pp. 1–8.

図 1. 提案アルゴリズムの概要
輝度積算ヒストグラムから計算された各局所領域について，平滑化された局所トーンカーブをシグモイド関数の和として直接導出している．

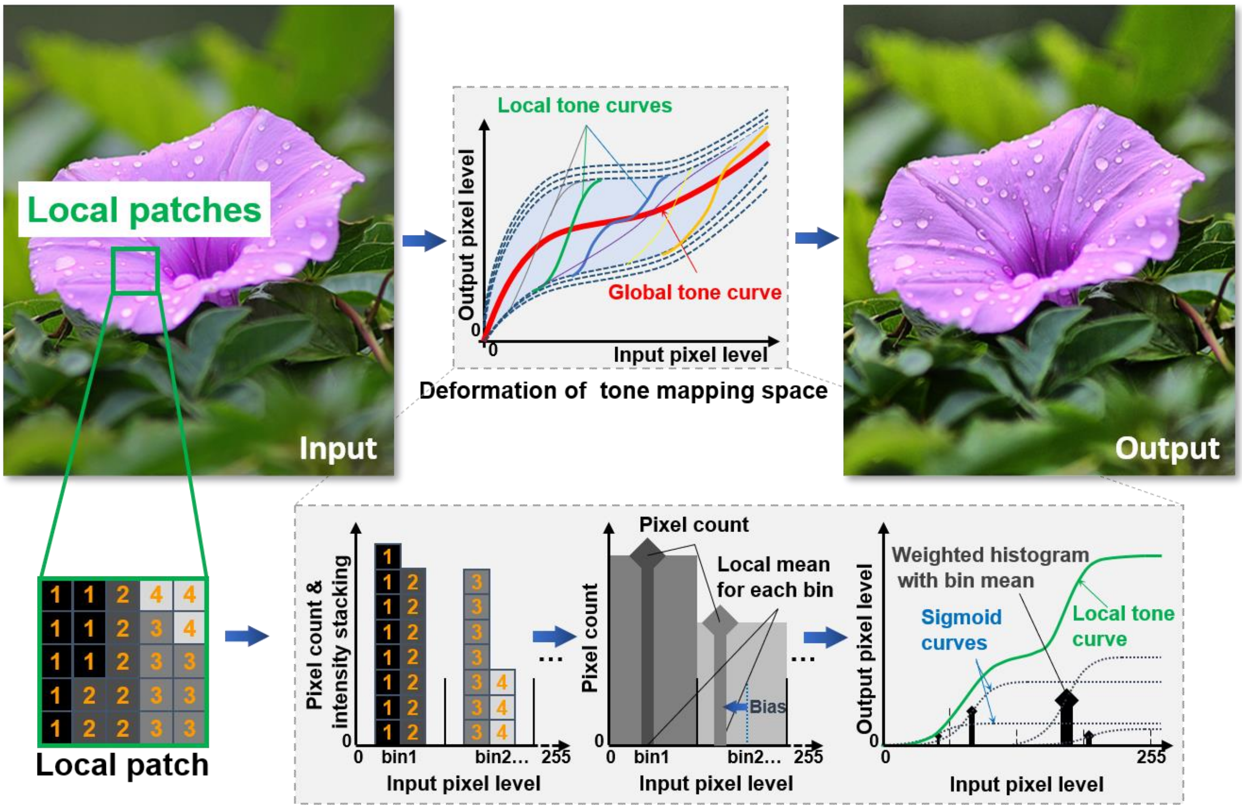
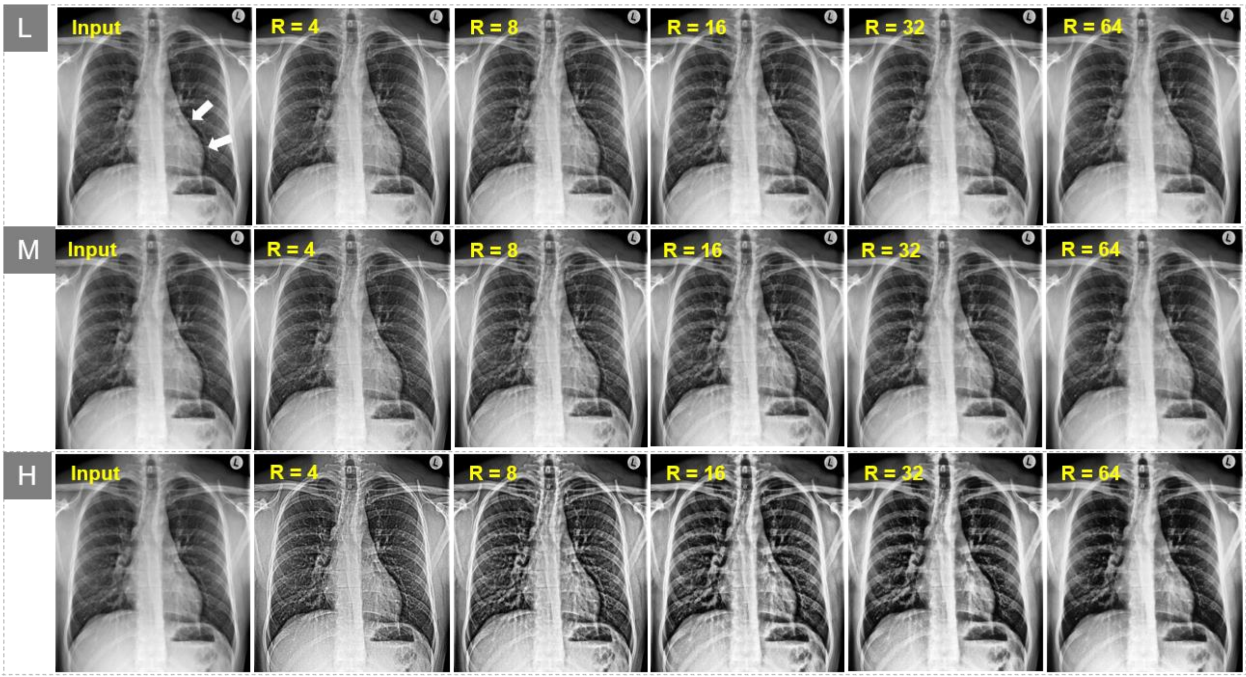


図 2 心臓に隣接したマッハバンド錯視のある胸部 X 線写真

マッハバンド錯視を用いた胸部 X 線画像と，X 線写真コントラストへのカーネルサイズ依存性を示す（半径 4→64）.



Range of the tone mapping space (Max = 100). For SLHE configuration L: Shadow = 10 to 40, Middle = 35 to 65 and Highlight = 60 to 90. Configuration M: Shadow = 0 to 50, Middle = 25 to 75 and Highlight = 50 to 100. Configuration H: Shadow = -25 to 75, Middle = 0 to 100 and Highlight = 25 to 125.

表 1 図 2 への放射線科医による主観評価

Preset	**R=4**	**R=8**	**R=16**	**R=32**	**R=64**
L	1	2	3	4	5
M	5	1	2	3	4
H	5	4	3	2	1

Ph.D. Degree in Engineering

HOKKAIDO UNIVERSITY

Real-time Contrast Management in Digital Images using Local Histogram Equalization Filter

Functional Communication Devices and Circuits Laboratory

Graduate School of Information Science and Technology

Prasoon Ambalathankandy

Supervisor: Prof. Masayuki Ikebe

August 2020

Contents

List of Tables iv

List of Figures v

Acknowledgments xii

Abstract xiv

Abbreviations xvii

Chapter 1 Introduction 1

1.1 High dynamic range . . . 1

1.1.1 Dynamic range of our eyes . . . 3

1.1.2 Dynamic range of a camera . . . 5

1.1.3 Dynamic range of standard display devices . . . 6

1.1.4 Applications . . . 6

1.2 Contributions . . . 6

1.3 Thesis organization . . . 7

Chapter 2 Background 8

2.1 Introduction . . . 8

2.2 High dynamic range imaging . . . 8

2.2.1 HDR merge . . . 9

2.2.2 HDR image sensor . . . 11

2.2.3 Data type . . . 11

2.3 Tone mapping . . . 13

2.3.1 Perceptual match . . . 14

2.3.2 Global tone mapping 17
2.3.3 Local tone mapping 18
2.4 Conclusion 19

Chapter 3 Smoothed local histogram equalization algorithm pipeline 20
3.1 Introduction 20
3.2 Image enhancement algorithms 21
3.2.1 Multilayer methods 22
3.2.2 Edge-preserving filter 23
3.3 Color temperature based perceptual decolorization for tone mapping applications 24
3.3.1 Related work 25
3.3.2 Definition of our proposed decolorization method . . . 27
3.3.3 Luminance mapping using red and blue weighting function 29
3.3.4 Evaluation 32
3.3.5 Perceptual evaluation 33
3.3.6 Limitation of our decolorization method 35
3.4 Smoothed LHE Algorithm 36
3.4.1 Noise characteristics of SLHE filter 38
3.4.2 SLHE halo control 40
3.4.3 $O(1)$ Time exact tone mapping 41
3.5 Image quality assessment 44
3.5.1 Objective IQA 45
3.6 Conclusion 47

Chapter 4 Characteristics of SLHE tonemap visualization 49
4.1 Introduction 49
4.2 Lateral Inhibition in the Human Visual System 50
4.3 Processing images with optical illusion 53
4.4 Analysis: Multilayer filters versus Smoothed LHE filter 54
4.5 Natural tone mapping 60
4.5.1 Background manipulation 60
4.6 Subjective user study 61
4.7 Conclusion 63

Chapter 5 Applications **64**
5.1 Introduction 64
5.2 Contrast enhancement of medical images 65
5.2.1 Contrast enhancement for digital X-rays 65
5.2.2 Medical image processing with smoothed LHE 65
5.2.3 Radiologist Perspective 66
5.3 Demo video 71
5.3.1 Objective video quality assessment 72
5.4 Conclusion 73

Chapter 6 Methodology and design of hardware accelerator for SLHE algorithm **74**
6.1 Introduction 74
6.2 Related work 75
6.3 Design methodologies for SLHE algorithm implementation on hardware platforms 75
6.3.1 Overall view of proposed TMO system 79
6.3.2 Compressed double frame buffers 82
6.3.3 Memory controller 83
6.3.4 Our local histogram estimation 83
6.3.5 Tone mapping module 86
6.3.6 Color converter 88
6.4 Xilinx FPGA synthesis summary 89
6.4.1 Comparison with other related implementations 91
6.5 Conclusion 92

Chapter 7 Conclusions and future work **94**
7.1 Thesis Summary 94
7.2 Future Work 98

Vita **121**

List of Tables

1.1 The table shows the measured illuminance for common scenes. 3
1.2 Table shows the relation between illuminance ratio and dB. . . 4
1.3 The table shows how the precision at which light measurements are translated into digital values also pose limitations on recorded dynamic range. 4

3.1 C2G-SSIM measurement for images in Fig. 3.6 33
3.2 Runtime comparison table with other decolorization algorithms. 34
3.3 Noise suppression performance of our SLHE filter compared to other edge-preserving filters. 39
3.4 Experimental result of SSIM measurement on test images. . . 45
3.5 Experimental result of TMQI measurement on test images. . . 46

5.1 Radiologist Subjective Evaluation. 68
5.2 Noise variance, SSIM, and TMQI experimental measurements of tone mapped video output. 72
5.3 Contrast assessment of TMO video processing 72

6.1 List of tone mapping algorithms implemented on GPU platform 76
6.2 List of tone mapping algorithms implemented on FPGAs . . . 77
6.3 List of full custom designed tone mapping algorithms on ASIC 78
6.4 Memory size of compressed frame/line buffers 83
6.5 Bit depth at each bin 87
6.6 FPGA resource utilization summary 89
6.7 Performance evaluation 89
6.8 Contrast measure for video output 90
6.9 Comparison with related works 90

List of Figures

1.1 This figure illustrates the orders of magnitude of the dynamic range of the eye, camera and display units. Dynamic range of a digital camera is in range of 60 to 72 dB, human eyes have very wide sensitivity through adaptation but instantaneous HVS is about 80 to 110 dB. Standard display monitors offer about 48 dB color and HDR displays with 72 dB dynamic range offers superior quality with 12-bit color. 2

1.2 This figure illustrates how human visual system adapts over a very wide dynamic range. The scotopic, mesopic and photopic regions are defined according to whether rods alone, rods and cones, or cones alone operate. 5

2.1 This figure illustrates how multiple exposure images are required to capture the wide ambient luminance levels that exists in our natural lighting environments. The HDR images produced from these multi-exposure images have higher bit-depths and a TMO is required to faithfully display it on a common display device (images from [1]). 9

2.2 HDR Imaging: Using a camera response curve the full dynamic range of the scene is captured from a set of LDR images with different exposure times. Algorithms like Popadic et al's can directly generate HDR-like images from bracketed images [2]. (images from [1]). 10

2.3 HDR image sensing methodologies (images from [1]). 11

2.4 The histogram here shows the numbers of past tonemap operators input data-type. Most of the papers considered 32-bit HDR data. 12
2.5 Data type based relationship between HDR images, TMO operators, and hardware platforms. 12
2.6 HDR imaging and tone mapping: Global tone mapping functions are good for capturing overall preview of the input image. Local tone mapping function by considering pixel neighborhood information for each input pixel, it can emphasize more local details. Additional filters are used to improve the subjective quality of tonemapped images (original image from [3]). 15
2.7 Objective of a tone mapping function is to achieve perceptual similarity between actual scene and device display image. . . 15
2.8 HDR-LDR Pipeline: HDR imaging, sensing, and tone mapping for display. 16
2.9 Deformation of tone mapping space. (a) Global Tone Mapping. (b) Global and Local Tone Mapping. 17
2.10 General block diagram for global tone mapping system for hardware implementation. 18
2.11 General block diagram for local tone mapping system for hardware implementation. For local calculation full frame buffer/line buffer or compressed frame buffer is required (images from [1]). 19

3.1 Design summary of edge-preserving filter & local histogram equalization: (a) Guided filter (b) Domain transform filter (c) Apical Iridix and (d) Our smoothed LHE filter. 23
3.2 A comparison of luminance components obtained using YCbCr, CIE Lab and our proposed method. 27
3.3 Main concept of our decolorization method: Human perception of warm and cool colors. Warm colors “advance” toward the eye, while cool colors “recede”. In this work we are able to accurately reflect the human perception of warm color, whereas this phenomenon is non-existent in conventional YCbCr color space. 28

3.4 Our red and blue weighted color mapping method. Color chart converted to grayscale using our decolorization method. 30

3.5 Imapct of β_R and β_k on output luminance value. (top) β_k set to 0.8 and for changing β_R we move from Gmax to Rmax. (middle) β_R set to 0.55 and for changing β_k we move from Bmax to Rmax. (bottom) β_R set to 0.55 and for changing β_k we move from Bmax to Gmax. 31

3.6 Comparison with other decolorization methods. Citations of the methods are the same as the table below. 33

3.7 Subjective Visual Test 34

3.8 Examples to demonstrate limitation of our decolorization method when mapping bright green/neon color. 35

3.9 Proposed algorithm's conceptual framework. We directly calculate smoothed local tone curve as sum of sigmoid functions, for each local patch which is computed from intensity stacked histogram [4]. 37

3.10 Our noise suppression method. (a) Input histograms. (b) Our tone blending function. 39

3.11 Image quality assessment: Noise variance estimation using different filters. Test images used for assessment in table 3.3 . . . 40

3.12 Halo control in our method: (a) Check whether P_{IN} belongs to light or dark luminance group. (b) Individually assign weights to light and dark luminance groups. (c) Default edge form of our tonemap function; the dark halo have a positive effect on the contrast enhancement. (d) Image with halos. (e) Halo suppressed using our TMO. 42

3.13 $O(1)$ image processing scheme: (a) Conventional local histogram estimation with box shape and, thus box filtering can be executed with $O(1)$ computation; (b) Updating column histogram with line buffer by one-pixel subtraction and addition; (c) $O(1)$ local histogram estimation using downscaled frame and line buffers and sub-pixel linear interpolation on the estimation scheme. 43

3.14 Image quality assessment. (a) Structural similarity index measurement: Input test images. (b) Tone Mapped image Quality Index: Input test images (linear log-luminance compressed) . . 45

3.15 Comparison with other methods (Local Contrast Enhancement): (a) Column of original images. (b) WLS [5] (using default setting); (c) DMT [6] (each layer is similar to WLS); (d) Apical's Iridix® [7](default setting, base functions =5, $\alpha = 0.5$) ; (e) This work (range of the tone mapping space (Max = 100): Shadow = 10 to 60, Middle = 20 to 80, Highlight = 40 to 90). Relative local patch size is w = 31 at VGA resolution (d,e). . 46

4.1 The brighter bands at 45° and 135° are illusory. This illusion is one of the clearest demonstrations of the difference between physical stimulation and the perception. The two graphs represent similar luminance profile along the diagonal and horizontal path. But we clearly perceive brighter profile along the diagonal. 50

4.2 Illustration of the Mach band illusion. The light stimulus is a light-dark-light bars of equal size, whose actual luminance intensity is flat when plotted (yellow). The perceived luminance distribution is different (green). A hypothetical model of photoreceptors is presented. We assume that each photreceptor inhibits the adjacent neighbor by 1/10 of its own activation. . 51

4.3 Background contrast effect: (a) The Benussi ring, here the ring appears to be of uniform brightness. (b) When a line is drawn along the background boundary, we perceive the ring to have different brightness levels. 52

4.4 (Left) Illustrates the luminance profile of a recent work [8] and (Right) The luminance profile we reconstruct in this work. . . 53

4.5 Processed cross section of Kofkka's ring. 54

4.6 Analysis of filter response for varying kernel sizes. (a) SLHE (b) DTF (c) GF. SLHE filter maintains the integrity of input signal across all kernel radius. 55

4.7 Characteristics of detail layer extraction in tonemap space. Tone mapping functions are expressed as linear function. $L_{Mean,\Omega}$ is always mapped to zero. The high frequency components which are differences between $L_{Mean,\Omega}$ and P_{IN} are output as a detail layer, which get saturated by fixed gain. 56

4.8 Filter output characteristics: (a) Input (b) Multilayer methods (c) Local Histogram Equalization (d) Our smoothed LHE filter 56

4.9 Detail manipulation: (a) Input image (b) using our smoothed LHE [SLHE settings (maximum = 100, 0 $\leq$ shadow $\leq$ 50, 25 $\leq$ middle $\leq$ 75, 50 $\leq$ highlight $\leq$ 100)] (c) DTF [IC Filter, σ_s = 20 and σ_r = 0.08, manipulate the D_0 using sigmoid function described in [5]] [6] (d) GF [r = 16, $\epsilon = 0.1^2$ and the detail layer was boosted five times.][9] (e) and WLS filter [Combined] [5]. . 57

4.10 (Upper left) Input (Upper right) SLHE Output (Bottom left) Horizontal cross section and (Bottom right) Vertical cross section. 59

4.11 Tone mapping for color images. 60

4.12 Application: (a) Ideal image processing pipeline (TMO example) (b) Color boost using local tonemap (c) Individual color control using our algorithm. 61

4.13 Images used for our visual perception test. SLHE settings (maximum = 100, 0 $\leq$ shadow $\leq$ 50, 25 $\leq$ middle $\leq$ 75, 50 $\leq$ highlight $\leq$ 100). DTF settings (IC Filter, σ_s = 20 and σ_r = 0.08, manipulate the D_0 using sigmoid function described in [5]). GF settings (r = 16, $\epsilon = 0.1^2$ and the detail layer was boosted five times.) 62

4.14 Summary of our visual perception test results. 63

5.1 X-ray image enhancement: (a) Cervical spine input (b) Enhanced using GF (c) Enhanced using our SLHE. [SLHE settings (maximum = 100, 0 $\leq$ shadow $\leq$ 50, 25 $\leq$ middle $\leq$ 75, 50 $\leq$ highlight $\leq$ 100)] 67

5.2 (a) Chest X-ray with Mach band effect adjacent to heart marked by pointers. (b) Low AP view of the pelvis with both hip joints. 68

5.4 Effect of Kernel Size on X-Ray contrast enhancement. Range of the tone mapping space (Max = 100). For SLHE configuration L: Shadow = 10 to 40, Middle = 35 to 65 and Highlight = 60 to 90. Configuration M: Shadow = 0 to 50, Middle = 25 to 75 and Highlight = 50 to 100. Configuration H: Shadow = -25 to 75, Middle = 0 to 100 and Highlight = 25 to 125. 69

5.5 Effect of Kernel Size on X-Ray contrast enhancement. Following are the SLHE configuration parameters, range of the tone mapping space (Max = 100). For configuration L: Shadow = 10 to 40, Middle = 35 to 65 and Highlight = 60 to 90. Configuration M: Shadow = 0 to 50, Middle = 25 to 75 and Highlight = 50 to 100. Configuration H: Shadow = -25 to 75, Middle = 0 to 100 and Highlight = 25 to 125. 70

5.3 Ten frames per sequence were grabbed from our test video for quantitative evaluation: (left) Input video. (right) Tone mapped output video. 71

6.1 Choice of hardware platform for developing an algorithm mainly depends upon the application. Other important factors are design time and engineering costs. 78

6.2 Overall view of proposed TMO system. 79

6.3 System pipeline with module latencies. 81

6.4 Compressed image generation routine. 84

6.5 Local histogram update routine. 85

6.6 Tone mapping module: Above is cumulative histogram generation circuit followed by normalization and denoise. Below is blown up view of a mapping function. 86

6.7 Tone mapping module: Upper/lower limit curve circuit. 87

6.8 Output color converter: White to RGB channel 88

6.9 Our hardware experimental setup. 89

7.1 The Chevreul illusion is used to examine the effectiveness of conventional tone mapping methods. The tone mapped image is perceived unnatural due to over enhancement. 95

7.2 Comparing the luminance profile of HVS and perceived luminance we can understand the stacked LI effect in tone mapped images which causes the unnatural feeling. 96
7.3 Trade-off: Visual quality versus hardware specification. 97
7.4 Block diagram of a plausible machine learning-based tonemap operator which can be implemented on hardware. 98

Acknowledgments

Like every PhD thesis, this one too would not have been possible without the help, guidance and support from many people. While it is not possible to name every one, I would like to acknowledge everyone who directly or indirectly helped to make this work a reality. First and foremost I would like to thank Prof. Masayuki Ikebe for giving me this opportunity and then guiding and motivating me through doctoral project. Many thanks to Ikebe sensei for being so patient and helping me throughout this thesis and shaping my research into a good manuscript. I would also like to thank Ms. Momoe Nakajima, Ms. Hidemi Yoshioka of IST, Ms. Yuki Miura and Ms. Atsuko Yokokawa of LALSIE, Ms. Cheiko Akiyama and Ms. Mizuho Tanaka of RCIQE for their timely administrative support.

This work was supported by great collaborators, Dr. Jyotsna Kochiyil (Radiologist) critically assessed the usefulness of the proposed algorithm on medical X-rays. Prof. Goro Nishimura of Research Institute of Electronic Science, helped me understand the mathematical background of biomedical imaging and greatly supported with the infrastructure materials for this research.

I would like to sincerely thank my thesis examination committee of Prof. Tetsuya Asai, Prof. Seiya Kasai, Prof. Sueoka Kazuhisa, Prof. Akihiro Murayama and Prof. Matsumoto Shingo for their detailed review, constructive criticism and excellent advice during the preparation of this thesis.

My initial training year was spent at Laboratories for Advanced LSI Engineering lead by Prof. Masato Motomura and Prof. Tetsuya Asai, whom I would like to thank for providing the great environment filled with many insightful discussions. I would like to thank my lab mates for all the interactions and numerous coffee and lunch breaks we shared during these three and half years. I would like to thank Prof. Shinya Takamaeda, Dr. Kodai Ueyoshi, Kentaro Orimo, Kota Ando, Takeshi Shimada, Sayuri Yokoyama, Takayuki Yoshida, Takahiro Ikegami, Pitchayapatchaya Srikram, and Zhu Keying. Special thanks to Yafei Ou, and Yuri Kanazawa, for being my sounding board, giving valuable suggestions and making my stay at Hokkaido University a memorable one.

I would like to thank my parents and my brothers for their continuous encouragement and support. Lastly, I would like to thank my wife Bincy and son Rohan for their patience as they gave me all the time I needed to complete this thesis. This work was financially supported, in part, by Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (KAKENHI) Grant Number 18H0321309, JSPS Fellows, Grant Number 19J14105, and Toshiba Memory Corporation (now Kioxia Inc) Grant-in-Aid for the Encouragement of Young Scientists (2019) which are gratefully acknowledged.

Abstract

Growing demand for displays with superior quality and higher resolutions has been continuously driving the high dynamic range market. The advances made in high dynamic ranging techniques and availability of suitable hardware platforms to run these algorithms has found applications in AR-VR, gaming, medical, security and many other computer vision related areas. This thesis presents both hardware and software contributions in the field of high dynamic range imaging by designing, implementing and analyzing a smoothed local histogram equalization algorithm.

The proposed global and locally adaptive tone mapping algorithm is inherently fast in execution and can process full HD images and output video at 60 frames per second. Our local histogram equalization (LHE) based algorithm controls global and local characteristics individually. In contrast to other tonemap operators, our algorithm manages light/dark halos separately and by using local tonemap function alone, it is able to effectively suppress noise. Traditionally, tone mapping algorithms operate on luminance channels which can lead to some loss of information. We minimized this data loss by employing a human perception based color to luminance mapping scheme. We proposed a light-weight and high-speed image decolorization method based on human perception of color temperatures. Our grayscale conversion demonstrates that warm colors are lighter than cool ones by using a blending function with R and B channel weighting. Our optimal color conversion method

produces luminance in images that are comparable to other state of the art methods. We validated this by a user study and found that our color to luminance mapping achieves effective luminance distribution and is very suitable as a pre-processing step for tone mapping application, thereby making our TMO operator ideal for many practical applications.

Our smoothed LHE algorithm operates with large kernel size, and forms saw-tooth like edges which is useful for processing images with optical illusions in them. We thoroughly analyzed state of the art edge preserving filters for processing these images with illusion and find they have certain inherent limitations which causes them to produce images with over emphasized edges. Digital X-rays are low contrast images which are also known to have optical illusions like Mach bands and background contrast effects, which are caused by lateral inhibition phenomena. We observe that using multilayer (ML) methods with latest edge preserving filter for contrast enhancement in medical images can be problematic and could lead to faulty diagnosis from detail exaggeration which are caused by uncontrolled texture boosting from user defined gain settings. ML filters are designed with few subjectively selected filter kernel sizes, which can result in unnaturalness in output images. We analyzed, and report that our smoothed LHE filter with adaptive gain control, is more robust and can enhance fine details in digital X-rays while maintaining their intrinsic naturalness. Preserving naturalness in X-ray images are an essential feature for radiographic diagnostics. Our proposed smoothed LHE filter has $O(1)$ complexity and can easily be controlled and operated with a continuously varying kernel size, which functions like an active high pass filter, amplifying all frequencies within the kernel.

The contributions presented in this thesis can be viewed as setup for full HDR pipeline. This thesis work helps in realizing an efficient hardware implemented smoothed LHE tonemap operator for natural appearing display.

We have performed multiple subjective user studies to validate our algorithm's effectiveness in retaining perceptual match.

Keywords: High dynamic range imaging, tone mapping, $O(1)$ local histogram equalization, FPGA, lateral inhibition, optical illusion, visual perception, mach band, background contrast effect, and digital X-ray

Abbreviations

AHE	Adaptive histogram equalization
ALARA	As low as reasonably achievable
AP	Anteroposterior
ASIC	Application-specific integrated circuit
ASIP	Application-specific instruction set processor
C2G-SSIM	Color-to-gray structural similarity index measure
CLAHE	Contrast limited adaptive histogram equalization
CNN	Convolutional neural network
CPU	Central processing unit
DNN	Deep neural nets
FPGA	Field-programmable gate array
FPS	Frames per second
GPGPU	General purpose graphics processing unit
GPU	Graphics processing unit
HDR	High dynamic range

HE	Histogram equalization
HVS	Human visual system
HW	Hardware
LCD	Liquid-crystal display
LDR	Low dynamic range
LHE	Local histogram equalization
LI	Lateral inhibition
ML	Machine learning
ML	Multilayer (edge-preserving filter)
NLM	Non-local mean
SDR	Standard dynamic range
SIMD	Single instruction, multiple data
SLHE	Smoothed local histogram equalization
SSIM	Structural similarity index measure
SW	Software
TMO	Tone mapping operator
TMQI	Tone mapped image quality index

Chapter 1

Introduction

Superior display quality is a dominant feature that has been driving the consumer electronics industry. Unlike the past, the growing demand for definitive viewing experience is not only limited to entertainment, gaming and media industry but has been increasingly sought for applications in security and surveillance, automotive, medical, AR-VR, drones and robotics imaging systems. High Dynamic Range (HDR) imaging has come to be a compelling aspect of the new 4K/8K Ultra-high-definition (UHD) format [10]. Luminance and contrast, which are very important to the human eye, and our modern display systems have problems when dealing with visuals that may have details simultaneously both in sun, and in shadows. HDR imaging looks to solve these problems. This technique accounts for more realistically contrasted visuals by bringing out colors and detail in low-light areas so that visuals in shadow are not compressed, while not saturating the highlights. In other words, HDR can make the dark visuals deeper and the lights brighter, with more color shades with optimized contrast ratio of the display. However, this increase in amount of detail and extended color space comes at the price of higher data-width, thereby requiring more hardware/software resources to create, distribute and display HDR content.

1.1 High dynamic range

Dynamic range of a digital image is defined as the ratio between the darkest and the brightest points captured from a scene [11]. It can be expressed in

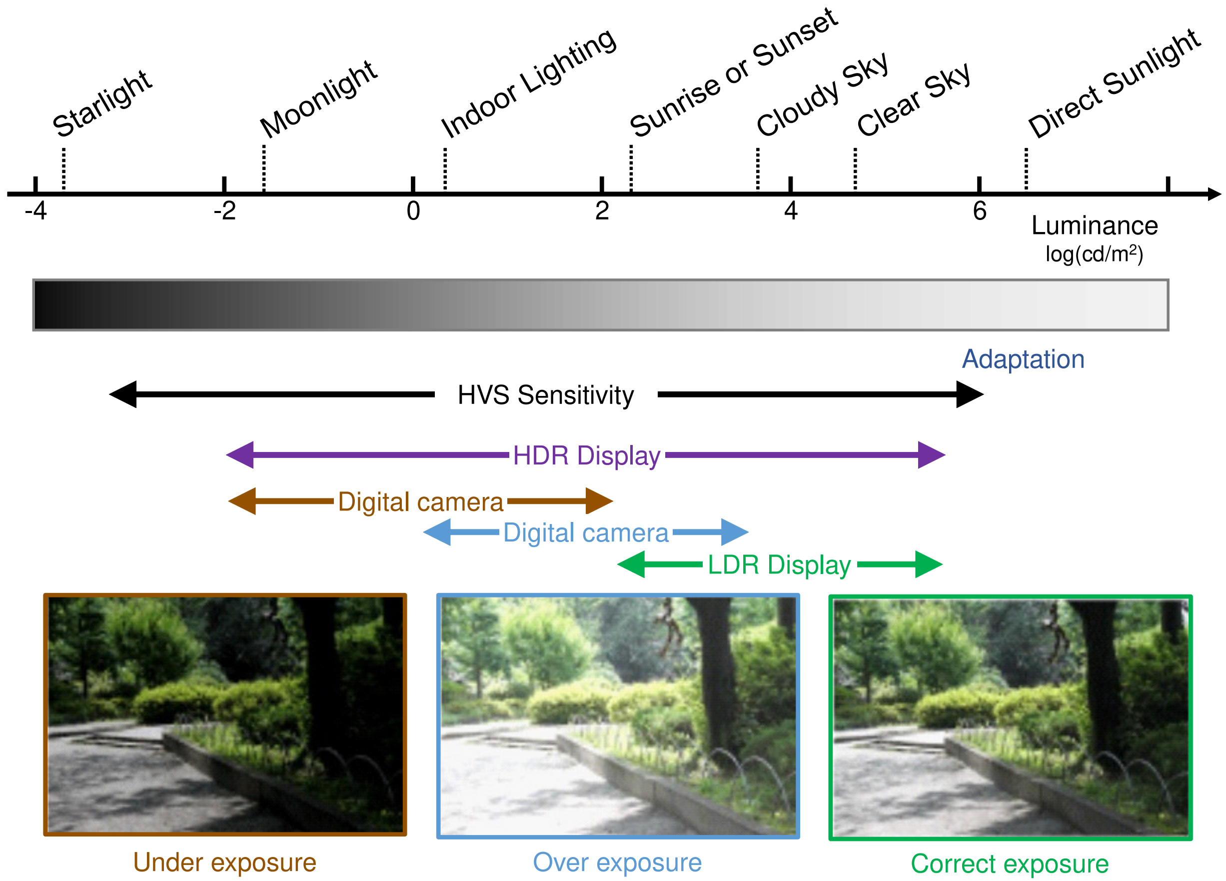


Figure 1.1: This figure illustrates the orders of magnitude of the dynamic range of the eye, camera and display units. Dynamic range of a digital camera is in range of 60 to 72 dB, human eyes have very wide sensitivity through adaptation but instantaneous HVS is about 80 to 110 dB. Standard display monitors offer about 48 dB color and HDR displays with 72 dB dynamic range offers superior quality with 12-bit color.

orders of magnitude (powers of ten), in stops (powers of two) or in decibels (db). From Fig1.1, we can notice that our eyes can see objects both in a dark night and in a sunny day, although the luminance level of a scene in sunlight is about $10^6 cd/m^2$ and one with starlight is about $10^{-3} cd/m^2$ [12]. This means that our human visual system (HVS) is capable of adapting to wide lighting variations within range of nearly 10 orders of magnitude. From [11], it is learnt that HVS can easily adapt up to 5 orders of magnitude within the same scene. HDR imaging aims to increase the dynamic range recorded in a digital image from a given scene. Pixels in an HDR image are proportional to the radiance of the scene, dark and bright areas can be recorded within the same image. But one of the main limitations of our digital cameras is their inability in capturing and storing the HDR of the natural scenes, which visually implies under and over exposure in bright and dark regions. Ordinary digital cameras produce images in a range lower than $1 : 1000 cd/m^2$ [13], thereby most of the digital images are still Low Dynamic Range (LDR), with a data-width of 24 bits per pixel (in RGB format 8-bits per channel). This translates to approximately 2 orders of magnitude while HDR images may have 80 orders dynamic range represented in floating point formats [11]. Tables 1.1, 1.2 and 1.3 shows the relationship between an actual scene and measured illuminance, which is digitized in dBs using analog to digital converters with limited precision.

Table 1.1: The table shows the measured illuminance for common scenes.

Scene	Illuminance (lux)
Night (moonlight)	0.25 - 1
Indoor	100
Office (lighted)	1000
Rainy day	25,000
Sunny day	100,000

1.1.1 Dynamic range of our eyes

The human eye can perceive very wide dynamic range (as shown in Fig 1.2) by adaptation, where our pupil opens and closes for varying light. Thus, our eyes

Table 1.2: Table shows the relation between illuminance ratio and dB.

Ratio	dB
100,000	100
10,000	80
1,000	60
100	40
10	20

Table 1.3: The table shows how the precision at which light measurements are translated into digital values also pose limitations on recorded dynamic range.

Bit Precision of ADC	Contrast ratio	Dynamic range (f-stops)
8	256:1	8
10	1024:1	10
12	4096:1	12
14	16384:1	14
16	65536:1	16

can see over a range of nearly 24 f-stops. For example, in situations of extreme low-light star viewing our eyes can adjust by using rod cells for night vision, by which our eyes approach even higher instantaneous dynamic ranges. This is known as scotopic vision and uses only rods (rods are very effective under low light conditions below $0.001cd/m^2$) to see, meaning that objects are visible, but appear in black and white. Vision above $3cd/m^2$ is based on photopic vision which allows for good color discrimination. Photopic vision typically dominates under normal lighting conditions, for instance during daytime [14]. It is based on three types of cones which are sensitive to short, middle, and long wavelength ranges, which generally appear blue, green and red, respectively to the human eye. Cones are limited in terms of light sensitivity. Mesopic range is the overlap region, where both rods and cones are involved. At any instance, human vision is able to operate over only a fraction of this enormous dynamic range. This subset is called the simultaneous HVS dynamic range as shown in Fig. 1.2. By adaptation HVS shifts to an appropriate light sensitivity by controlling the pupil, and other neuronal adaptive processes [15], so that under

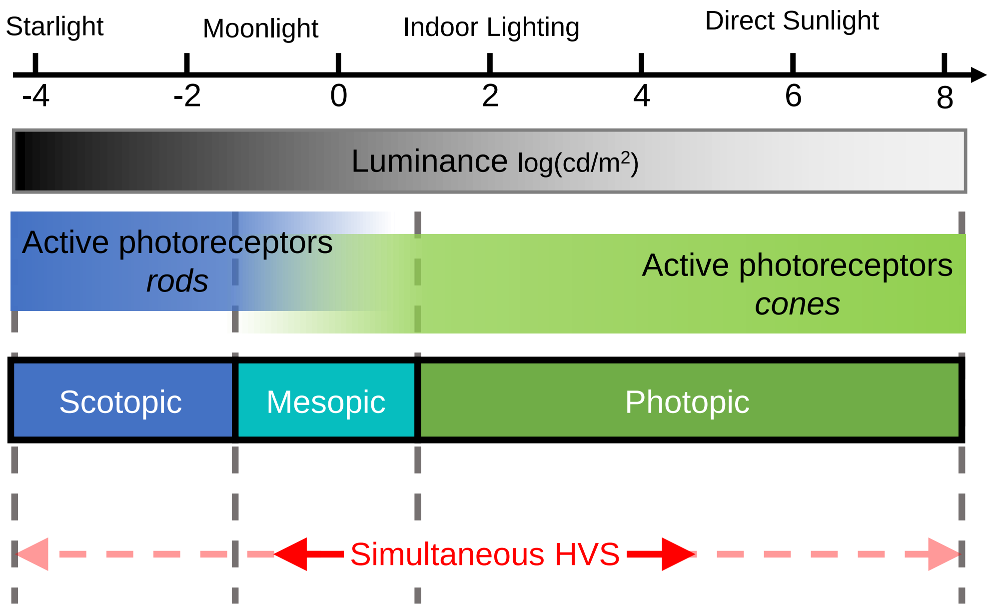


Figure 1.2: This figure illustrates how human visual system adapts over a very wide dynamic range. The scotopic, mesopic and photopic regions are defined according to whether rods alone, rods and cones, or cones alone operate.

any lighting conditions our visual sensitivity is maximized.

1.1.2 Dynamic range of a camera

Camera dynamic range is measured in "stops". An increase or decrease of one stop equals to doubling or halving of the brightness level. Previously we saw that the human eye can perceive about 20 stops of dynamic range in normal situations, which implies that the darkest tones we can perceive at an instant is about 6 orders darker than the brightest tone in the same scene. Cameras have a narrower dynamic range than the human eye, for example on a bright sunny day we can see details in dark shadows, however these details may not be clearly visible in a photo. Modern cameras like the Nikon D810 can achieve just under 15 stops of dynamic range in any one photo, and ordinary digital cameras can capture in the range of 12 to 14 stops. This is why we have to use High Dynamic Range (HDR) photography, in which we combine multiple exposure images (over/under exposure) to create a single final image (correct) as shown in Fig. 1.2.

1.1.3 Dynamic range of standard display devices

The typical dynamic range of conventional display devices is about 1,000:1. When we try to display a high dynamic range image on this ordinary LCD display device, shadow or highlight details are lost. Therefore, some tonemap functions are required to compress the dynamic range so that it can match the dynamic range of the display device. Modern display devices, for example an organic light emitting diode (OLED) screen can support larger dynamic range (for example Sony monitor [16]).

1.1.4 Applications

Computer vision algorithms have greatly benefited from high dynamic range imaging technologies. HDRI has also found applications in astronomical imaging, remote sensing, industrial design, and scientific visualization [11]. Medical imaging systems have also used higher dynamic range images to improve diagnostic quality. Typically, DICOM images have 12 to 16 bits per pixel resolution, which is higher than the conventional display device resolution of 8 bits per pixel [17, 18].

1.2 Contributions

Growing demand for displays with superior quality and higher resolutions has been continuously driving the high dynamic range market. For enhanced viewing experience high dynamic range imaging techniques are being employed in various fields as a replacement for standard digital techniques. This thesis contributes to the field of real-time tone mapping algorithm based on a smoothed local histogram equalization method. There are discussions on algorithm design and implementation on a FPGA device. Multiple detailed subjective user studies have been conducted to assess the image quality and validate the algorithm. Following are the main contributions of this thesis:

- Design and implementation of a smoothed LHE-based (SLHE) algorithm.

- FPGA implementation of the SLHE algorithm for real-time acceleration.
- Detailed study on characteristics of SLHE visualization.
  - Processing images with optical illusion.
  - Analysis of multilayer edge-preserving filters.
- Application of SLHE algorithm for medical image enhancement.

1.3 Thesis organization

This introductory chapter is a brief introduction to the field of HDR imaging. It also lists the main contributions provided in the thesis. The upcoming chapter 2 will provide a more thorough background on HDR imaging and discuss the important tone mapping functions. In Chapter 3, we present the main smoothed local histogram equalization (SLHE) algorithm. This chapter includes a specific color temperature based decolorization scheme, which is used as a pre-processing step for our SLHE algorithm. Chapter 4 discusses the visualization characteristics of our SLHE algorithm, and as an application we process images which have optical illusion in them. This is also validated with a subjective user study, to demonstrate the usefulness of our SLHE scheme for natural tone mapping. We further explore the usefulness of SLHE algorithms for contrast enhancement of medical images and discuss the radiologist's perspective in chapter 5. Chapter 6 is dedicated to hardware implementation of the SLHE algorithm. It also includes a demo video to demonstrate the effectiveness of algorithm acceleration using a Xilinx FPGA board. Finally, in chapter 7, the thesis is wrapped up by discussing an outlook towards the future of hardware accelerated machine learning-based tone mapping system.

Chapter 2

Background

2.1 Introduction

The fundamental objective of an imaging system is to capture and faithfully reproduce a scene as it would have been perceived by a human observer. However, there exists perceptible differences between HVS and existing imaging system's capability to faithfully capture and reproduce the entire range of scene luminance using a standard display device. This difference is highlighted in Fig. 2.1. Physical constraints like limited storage and distribution capacity is what leads to the difference between a scene capture and display. The primary focus of this chapter is to present in detail the basic concepts of HDR pipeline. This pipeline includes scene capture, display and a key mechanism of tone mapping which is required to prepare the HDR images to be displayed on standard display devices. In this chapter we will present a detailed account on HDR imaging pipeline and tone mapping methods.

2.2 High dynamic range imaging

There are many ways to generate or obtain HDR images, for example they can be captured from real world scenes, or rendered on computers by various computer graphics (CG) tools. In this section we will focus on real-time methods of obtaining HDR images using conventional cameras and special HDR sensors. For the CG methods there are well known books describing those

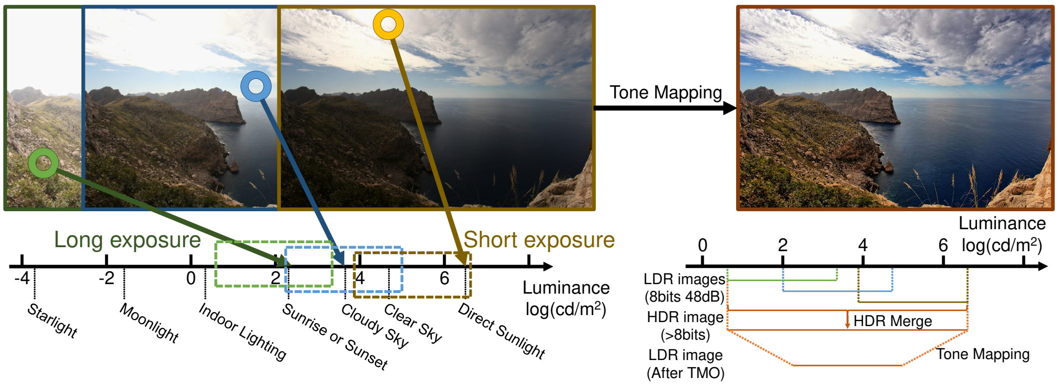


Figure 2.1: This figure illustrates how multiple exposure images are required to capture the wide ambient luminance levels that exists in our natural lighting environments. The HDR images produced from these multi-exposure images have higher bit-depths and a TMO is required to faithfully display it on a common display device (images from [1]).

methods [19, 20]. The gaming industry has been employing HDR rendering for a very long time, they were used for rendering special visual effects like dazzling, slow dark-adaptation there by enhancing the immersive impression [21]. Today HDR imaging is used in many applications to enhance functionality of cinematography and photography [22], biomedical imaging (see DICOM standard [23]) [24], remote sensing [25] and many more computer vision applications [26].

2.2.1 HDR merge

HDR images can be composed by combining multiple LDR with different exposure time in a single HDR. Here, the exposure time which is also known as the shutter speed is the duration of time when the digital sensor inside the camera is permitted to capture light. The amount of light that reaches the film or image sensor is directly proportional to the exposure time. Therefore, a long exposure time image will have an overall greater luminance. It will detect smaller amount of light sources even in darker areas. But the picture might saturate in bright parts of the scene due to too much of light for the sensor. On the other hand, a short exposure image will record bright parts of the scene but would not be able to register darker light sources. Exposure time values are often referred to as "stops". A stop consists in doubling the

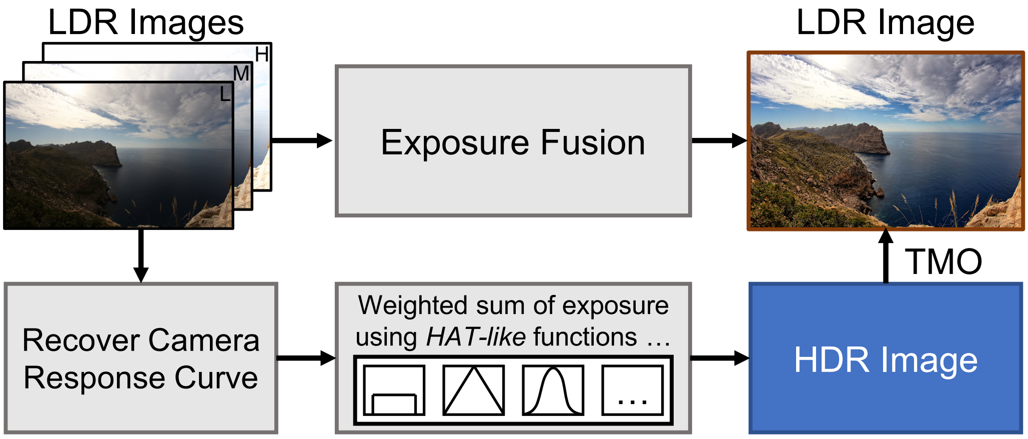


Figure 2.2: HDR Imaging: Using a camera response curve the full dynamic range of the scene is captured from a set of LDR images with different exposure times. Algorithms like Popadic et al's can directly generate HDR-like images from bracketed images [2]. (images from [1]).

exposure time (relative to a reference time). +1 stop is doubling, +2 stops is times 4, and -1 is halving the exposure time [27, 28]. So, there are many well-known techniques to combine multiple images of varying exposures to compose a HDR image [29, 30, 31, 32].

For an LDR image, its dynamic range is bounded by the range of sensor i.e., a camera with 8 bits/pixel can only capture ratios up to 255:1. However, we can achieve a greater dynamic range using the same camera by combining multiple images which have been captured with different exposure time. Each of these LDR images will cover a different range of the luminance in the scene. This allows greater resolution on the luminance captured . Images with a short exposure time will be adapted for capturing very bright parts of the scene but will fail to capture darker parts. Long exposure time images being the opposite, they will saturate in bright parts of the image. All intermediate images captured with various exposure time will help cover the whole range of luminance, this strategy is demonstrated in Fig. 2.2. The final composite image will have a greater dynamic range than it is achievable with a single shot by the camera.

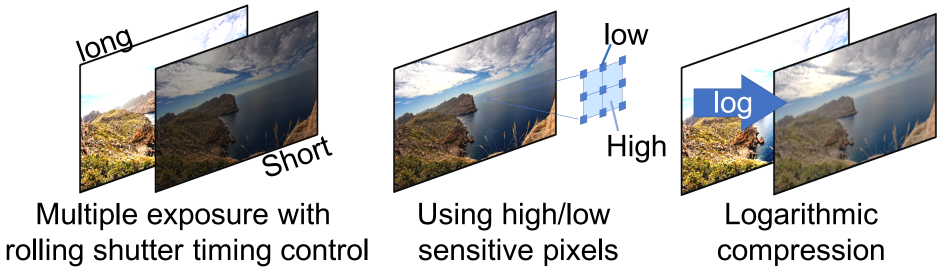


Figure 2.3: HDR image sensing methodologies (images from [1]).

2.2.2 HDR image sensor

There are broadly three different architectures that are used to design HDR image sensors (Fig 2.3). In the first group, by utilizing logarithmic response pixels or by employing likewise circuits to non-linearly extend the dynamic range as in [33, 34]. However, this non-linearity can cause severe problems when reconstructing images. The next group sensors like in [35, 36, 37] extend dynamic range by applying lateral overflow capacitors. But, this group of sensors would need to partially open a transfer gate so that the over-saturated charges can be collected by the overflow capacitor. Again, the threshold voltage of the transfer gates can have a large variation, thereby resulting in variations in saturation level. Also, this group of sensors is known to have higher dark current shot noise [38]. Mase et al., in [39] proposed a sensor design where they would use multiple exposure-time to expand the dynamic range. However, this method also has some issues; different integration time can cause discontinuities in SNR and also cause distortion in moving scenes. A new type of HDR image sensor was designed by Liu et al. [38], which used dual transfer gates. Their HDR image sensor is capable of capturing HDR scenes with low noise. The novelty of this design is it does not use the concept of transfer of threshold voltage and completes charge transfer in one operation.

2.2.3 Data type

HDR image formats and data types are discussed in detail in textbooks [11, 12]. Our focus in this discussion is limited to the data types of images used in hardware tonemap systems. Figure 2.4 shows the frequency of different data types used in hardware tone mapping papers. Most of these past works choose 16-bit

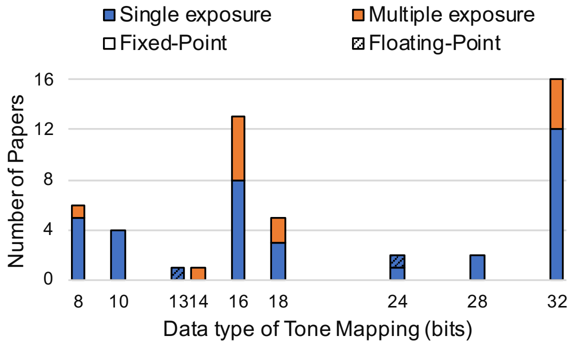


Figure 2.4: The histogram here shows the numbers of past tonemap operators input data-type. Most of the papers considered 32-bit HDR data.

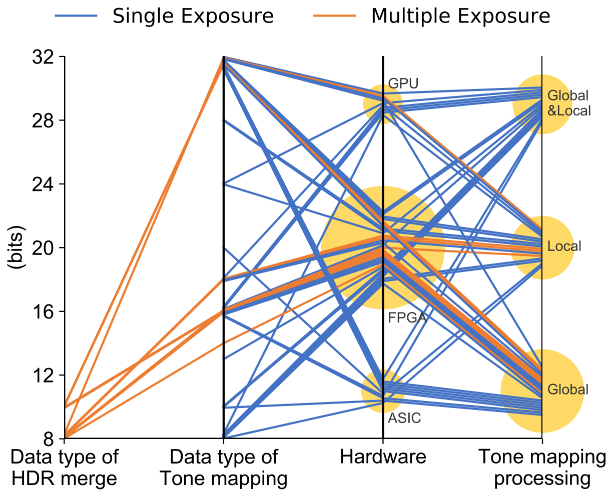


Figure 2.5: Data type based relationship between HDR images, TMO operators, and hardware platforms.

or 32-bit HDR images for tone mapping, and very few works have reported any other data types. Figure 2.5 demonstrates the relationship between data type, hardware and tone mapping processing schemes. As expected, all these proposed TMO operators were implemented on low-power embedded platforms using fixed-point arithmetic. In the same figure, we show if a HDR merge technique has been implemented in the hardware. Compared to floating-point arithmetic, fixed-point arithmetic has several advantages like low-power consumption, smaller footprint and high-speed computing [40, 41, 42]. On the other hand, floating-point arithmetic can realize larger range of luminance with smaller bit depth, Yadid-Pecht et al. in [43, 44] proposed FPGA implementations of tone mapping algorithms for HDR images using floating-point arithmetic.

2.3 Tone mapping

In section 1.1.2 we discussed how there exists a wide gap between the range of light that we can see & perceive and what a common digital camera can capture and display on a monitor respectively. Ordinary digital camera's image sensors have limited capability for recording the entire illumination range in a natural scene. Also, it is a challenge to determine good exposure values, especially in diverse scenes which have good large dark and bright areas. For example, taking a picture on a sunny day we have to chose whether to appropriately expose the bright sky and also account for the details in the shadow of the mountain slopes like in Fig 2.1. Modern digital cameras come with in-built auto-exposure algorithm to automatically set the ISO value, aperture and shutter speed corresponding to the given scene. However, a scene saturation (over-exposure) can occur if the scene is very brightly lit (direct sunlight) and sensor records the maximum allowed value, therefore details in bright areas are clamped to the maximum allowed value which is white. On the other hand, if the scene is poorly lighted and the light energy reaching the sensor is very low it results in under-exposed image. As stated earlier, we have to adapt the image so that we can match the dynamic range of HDR scene with the display device's dynamic range. This process is widely known as tone mapping. Depending upon the dynamic range of the captured image, tone mapping function can expand or compresses it in order to enhance the display quality [45]. The purpose of applying tone mapping on an image can be different and depends on the particular application. In some cases it may be to improve the aesthetics of the image, while for another application it might be to emphasize as many details as possible, or could be to maximize the image contrast [46]. However, the ultimate goal of tone mapping is to match the perception of tone mapped image with the real world perception [11]. A tone mapping operator (TMO) f can be defined as a transformation function $f(I)$:

$$f(I) : R^{w \times h \times c} \rightarrow D^{w \times h \times c} \tag{2.1}$$

Here, the tonemap function f maps real world luminance R to display luminance D [12], and I is an image with dimension $w \times h$ and c is number of color bands which is 3 for RGB image. Tone mapping has been an active area of research for the last two decades, resulting in the design and development of many hundreds of different tone mapping algorithms which can be broadly grouped in to global, local, frequency and segmentation operators [12].

Global TMOs, apply the same function to all pixels in the image. These mapping function treat every pixel of the image independently. These operators are computationally efficient, easy to implement and can be executed in real time. A local tone mapping function compresses a pixel value according to its luminance values and their neighboring pixels luminance values. Hence, for each pixel the computation is adjusted according to an average over a local neighborhood of pixels [11]. Frequency domain-based operators, like the local TMO preserve edges and local contrast by computing in the frequency domain instead of spatial [12, 47]. Segmentation operators divides input image into different uniform regions, and a global mapping operator is applied on each of these regions and finally these are merged to produce output image [48, 49]. HDR image pixels can have large disparity in intensities in small neighborhood thereby exhibiting artifacts in tone mapped images (particularly in local tonemap), like in Fig. 2.6. Therefore, various filters are required to suppress these artifacts and improve aesthetics of output images depending upon the targeted application.

2.3.1 Perceptual match

Fundamental objective of a tone mapping function is to reproduce as realistically as possible, the visual appearance of a scene captured (using a camera) on a conventional display device, by compressing or expanding the dynamic range of the image to fit into the display range while preserving details which perceptually match with our mental image. In this section we briefly review the imaging pipeline from acquisition to display as shown in Fig. 2.7. For any scene we first acquire digital images (for computer graphics, rendering) using a LDR/HDR camera which is then encoded and stored efficiently. Tone mapping is required to display HDR content on LDR display devices, and

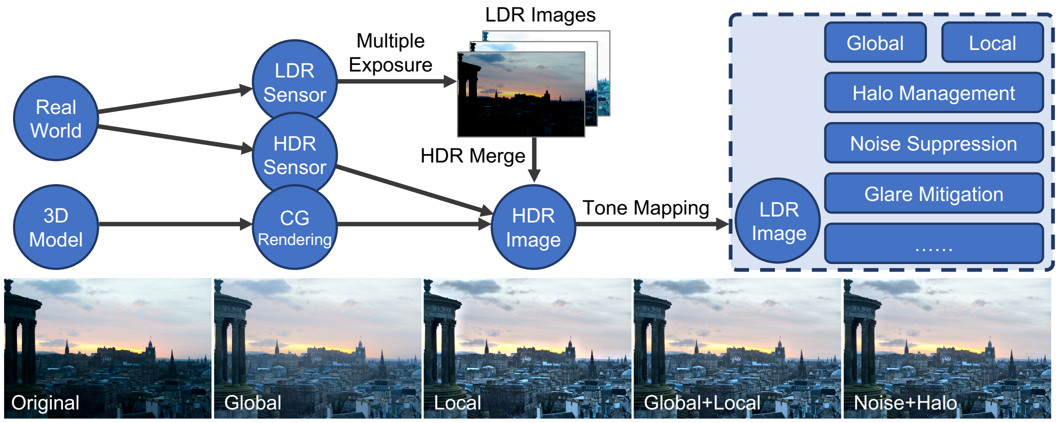


Figure 2.6: HDR imaging and tone mapping: Global tone mapping functions are good for capturing overall preview of the input image. Local tone mapping function by considering pixel neighborhood information for each input pixel, it can emphasize more local details. Additional filters are used to improve the subjective quality of tonemapped images (original image from [3]).

conversely inverse tone mapping to scale-up and match LDR content on HDR displays.

In the previous section 2.2 we discussed how to obtain/produce HDR content, and Fig 2.8 we show the broad HDR to LDR pipeline. This HDR content requires to be stored in a medium with a greater number of bits/pixels than that of a single LDR image. It is necessary in order to achieve a greater dynamic range. HDR display systems do exist [50], and TVs with an extended dynamic range are currently available in the commercial market, but they are

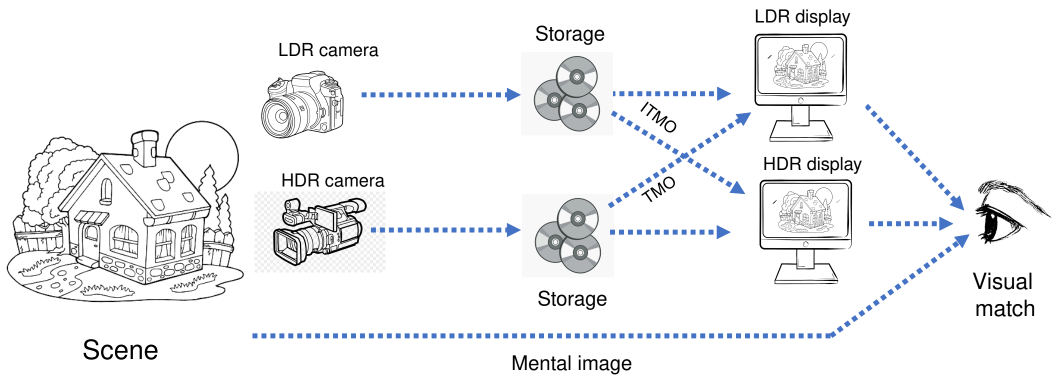


Figure 2.7: Objective of a tone mapping function is to achieve perceptual similarity between actual scene and device display image.

not as widespread due to their limitations in terms of dynamic range and color gamut. The process of tone mapping consists of reducing the dynamic range of the HDR image into an image that can be easily displayed on wide range of display devices which have limited dynamic range and color gamut.

This process can be highly non-linear depending on the result expected. The LDR image after tone mapping will have less information than the original HDR image due to the reduction of information by the tone map function. But it is ready to be displayed, and the image has enough information by revealing details in both dark and bright parts according to our perception. Tone mapping has been an important area of research as is evident from several surveys that have been published over the years [11, 22, 46, 51, 52]. Tone mapping function can be classified in to two broad groups based on the processing function they use.

- Global operators: The same mapping function is applied to all pixels throughout the image.

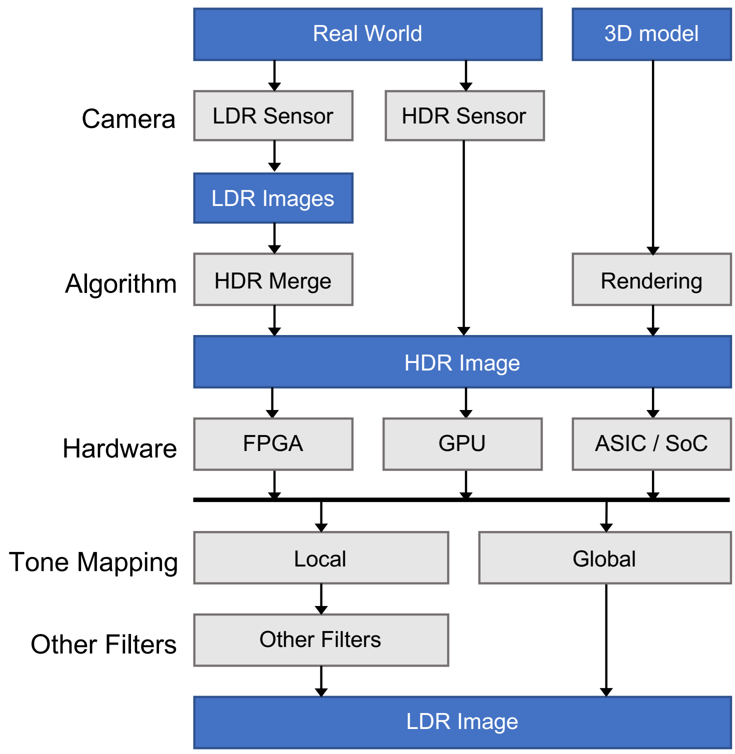


Figure 2.8: HDR-LDR Pipeline: HDR imaging, sensing, and tone mapping for display.

- Local operators: The mapping function applied to a pixel depends on its neighboring pixels.

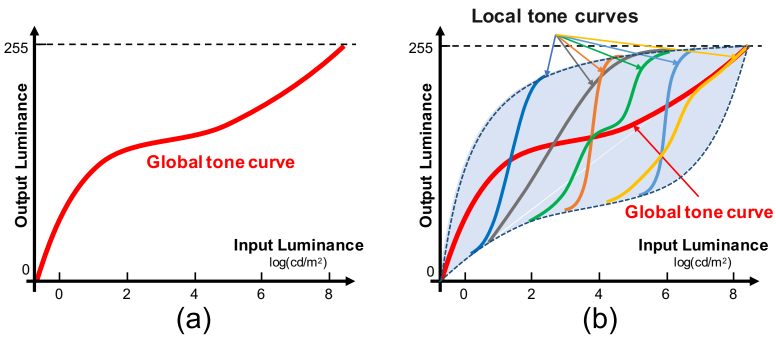


Figure 2.9: Deformation of tone mapping space. (a) Global Tone Mapping. (b) Global and Local Tone Mapping.

2.3.2 Global tone mapping

Global tone mapping algorithms (also known as "tone reproduction curves") are spatially invariant, that is they apply the same function to all pixels in the input image [53, 54, 55, 56, 57]. This results in one and only one output value for a particular input pixel value irrespective of the pixels in its neighborhood. Figure 2.10 shows a general pipeline which is useful for implementing a global tonemap function shown in Fig. 2.9 (a). As we can see, the tonemap pipeline first obtains the luminance image and from that it calculates global statistics (like $L_{max}, L_{min}, L_{average}$). In some algorithms these statistics are also calculated from previous frame based on an assumption that there are very little changes between successive frames when imaging at 30/60 frames per second [58, 59, 60].

In the pipeline next step is to realize a logarithmic or an exponential like function to compute the tone mapped image. The hardware implementation of these functions are a challenge and a common approach is to approximate such functions, by maintaining an acceptable level of error in the realized implementation. As these functions are frequently required to be computed in many engineering and scientific applications many hardware friendly methods have been proposed in literature [61, 62, 63]. The final step in the pipeline after

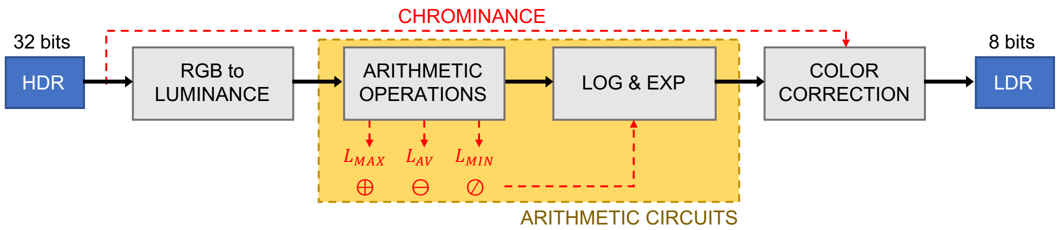


Figure 2.10: General block diagram for global tone mapping system for hardware implementation.

tonemap is to restore the color for displaying the output image. A detailed account for color correction technique is presented in section 3.3.

2.3.3 Local tone mapping

Local tone mapping algorithms (also known as "tone reproduction operators") are spatially variant and apply different functions based on the neighborhood of the input pixel [64, 65, 66, 67, 68, 69, 70]. Therefore, one input pixel could result in different output values based on its position as illustrated in Fig. 2.9 (b). Local tone mapping algorithms are computationally more expensive and time consuming compared to global tone mapping algorithms [71]. We can describe the operation of a local tonemap algorithm using the block diagram shown in Fig 2.11. As was in the case of global tonemap, we initially obtain the luminance values for the input image. The color correction step is described in detail in section 3.3. The high computation cost for local tonemap operator is due to the local information calculation for which a full frame or a few lines of the input image has to be buffered as shown in the Fig 2.11. Some algorithms have also implemented compressed frame buffer (down-sampled images) to reduce the memory cost [72, 73]. To meet the real-time constraints, as a common approach previous frame is used to compute the local information for current frame.

Local tone mapping methods generally produce better quality images as they preserve details, which may not be the case when using global tone mapping methods [11]. However, one of the major drawbacks of local tone mapping algorithms are the creation of halo artifact among the high contrast edges and the graying out of the low-contrast areas [74, 75]. Therefore, local

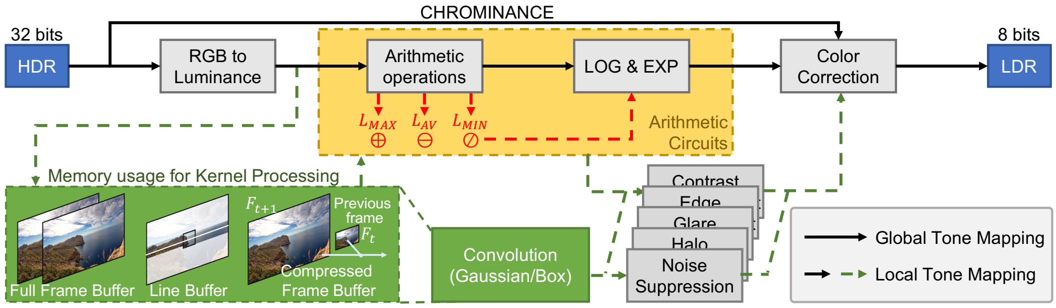


Figure 2.11: General block diagram for local tone mapping system for hardware implementation. For local calculation full frame buffer/line buffer or compressed frame buffer is required (images from [1]).

tonemap methods implement additional filters to suppress these image artifacts like halo and noise. Such filtering may require that the input image (of size $M \times N$) be convolved with a filter (of size $k \times r$). Benedetti et al. demonstrated a simple hardware sliding window convolution block, which can output one pixel every clock [76]. The latency associated with this sliding window method is calculated as given in 2.2.

$$T = Buffer_{depth} \times \left\lfloor \frac{KernelSize}{2} \right\rfloor + \left\lceil \frac{KernelSize}{2} \right\rceil \tag{2.2}$$

2.4 Conclusion

In this chapter we reviewed the HDR imaging pipeline, from capture to display. Given a physical scene, it can be captured using one or more imaging sensors, followed by processing the captured information using techniques for HDR reconstruction. An HDR camera can be used to directly capture the wide dynamic range image, either with a sensor that can cover the entire dynamic range or with a multi-exposure system using a LDR camera. The captured HDR image or video sequence is then stored using HDR format. Many different solutions have been proposed in literature for both static images and video. The next step in the HDR pipeline is to prepare the HDR image for display, using a tone-mapping algorithm. Ideally, the objective of tone mapping algorithms is to compress/expand the dynamic range of the image to the given range of the standard display device.

Chapter 3

Smoothed local histogram equalization algorithm pipeline

3.1 Introduction

Chapter 2 section 2.3 introduced tone mapping concepts and also reported TMO methods implemented in software. This chapter will present in detail the pipeline of our fast global and locally adaptive tone mapping algorithm. A Xilinx FPGA-based real-time implementation of this algorithm is presented in chapter 6. This tone mapping function, is based on smoothed local histogram equalization, and it can control global and local characteristics individually. This is in contrast with other tonemap operators reported earlier. Our smoothed LHE algorithm can manage light and dark halos separately, additionally by controlling the local tonemap function alone, it can effectively suppress noise. We carried out detailed subjective user studies to ensure the effectiveness of our algorithm in terms of important image attributes. To summarize, the main contributions of this chapter are as follows:

- A color temperature based perceptual decolorization method.
  - A luminance mapping using red and blue weighting function.
  - Detailed user study demonstrating the luminance mapping effectiveness.
  - Perceptual evaluation using a tone mapping application.

- A novel smoothed LHE-based tone mapping function which can control global and local characteristics simultaneously.

- Image quality assessment using objective metrics.

3.2 Image enhancement algorithms

With the easy availability of professional digital imaging devices, it is very convenient to capture images. But there are cases when one is not satisfied with the visual quality. Loss of quality may be because of poor lighting, limited camera features, or due to bad camera settings. Some post processing method is required to enhance such kind of images to be closer to human perception. Contrast enhancement is a method for improving the contrast and the visibility of regions of interest in images. Generally, contrast enhancement is conducted by a transform function, which is described as $g(x, y) = T[f(x, y)]$. Here, $f(x, y)$ is the original low contrast image, $g(x, y)$ is the enhanced image. Location of the pixel being processed are given as (x, y) and T is the enhancement operation. There are many contrast enhancement methods proposed in literature of which we will briefly review the following three important ones: Histogram-based Methods, Multi-scale Methods (including retinex) and Machine learning-based Methods.

The most popular histogram-based enhancement method is Histogram Equalization (HE), which is used for improving image quality. HE modifies the original histograms of images to an approximate uniform histogram, by stretching pixel value range of images. This method uses the original histogram of the considered image to obtain the mapping function. HE improves global contrast but fails in amplifying fine details, and it often introduces saturation like visual artifacts. To counter these problems, several other methods were proposed. An adaptive histogram equalization (AHE) within a local window was proposed in [77]. This was effective in enhancing fine details, but still had artifacts and increased noise problems in homogeneous areas. The contrast limited AHE (CLAHE) prevents noise amplification by constraining the contrast enhancement according to user-defined threshold values that clip the local histograms. By constraining the contrast enhancement, and clipping local histograms using user threshold values, noise amplification was suppressed in CLAHE. However, defining content specific threshold values are not easy task for users. A detailed overview on HE-based contrast enhancement methods is presented in [78].

In HE method contrast enhancement was done in spatial domain, Ramponi et. al in [79] divided source image into two frequency layers. They then amplified the high frequency layer and added it back to the low frequency layer. This operation corresponds to boosting fine details. But, in this case [79] the middle frequencies were left out and effect of noise was ignored. However, as the output results were promising, the idea has been extended and studied deeply proposing several multiscale methods. Many new algorithms call for decomposition of an image into a low frequency piece-wise smooth base layer plus one or more high frequency detail layers with different scales. In retinex multilayer (ML) method [80] the base and detail layers are respectively calculated as luminance and reflectance in log domain. By manipulating these layers separately, we can cater to various applications like denoising and detail transfer [81], tone mapping and transfer [66] [82], and image-based editing [83].

Recently, some machine learning-based techniques have also been proposed for image contrast enhancement. For example, in [84] an end-to-end deep convolutional network is proposed to improve photo visual quality, especially for contrast enhancement. This learning-based method can improve both image contrast and color rendition. They presented a composite perceptual error function that combines content, color and texture losses. A large data set was created by capturing images using three smartphones & one DSLR camera, and an image enhancement network model was built using this collected data set.

3.2.1 Multilayer methods

In multilayer methods, image illuminance which have low frequency components are handled as base layer. The remaining high frequency components are image textures which are detail information. Base and detail layers are obtained by filtering and are adequately compensated for over/under exposure while preserving contrast. Multilayer methods suffer from various problems depending on the type of filtering used. For example, if an isotropic filter such as a Gaussian one is used for separating the two layers, halo effects appear near the edges of objects in the image. This is because the edges smoothed by the filtering do not precisely represent the boundaries of objects with a large luminance difference. Therefore, edge-preserving filters are used in ML methods to suppress halo artifacts.

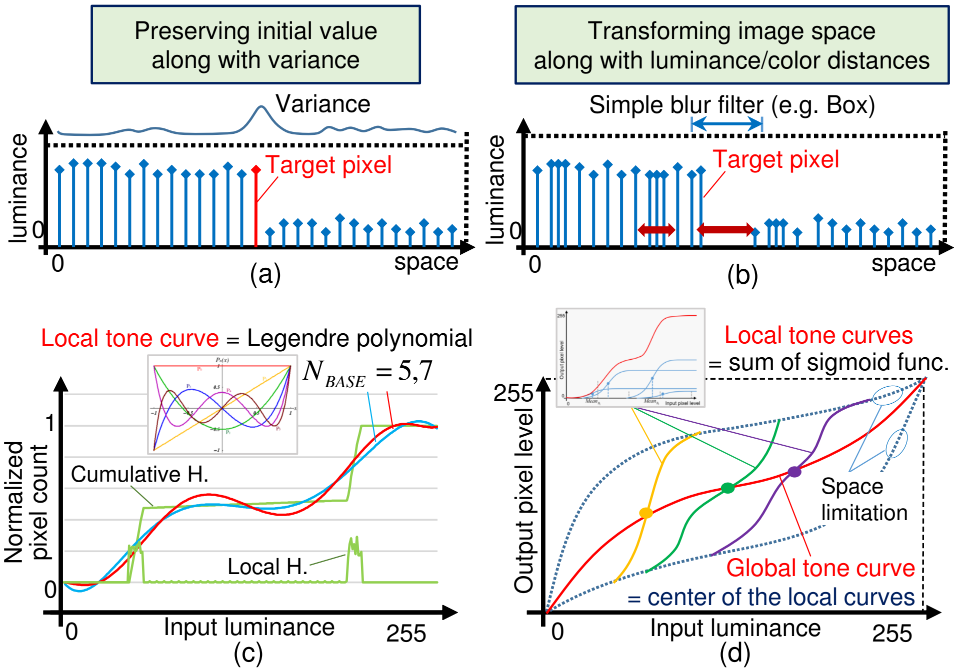


Figure 3.1: Design summary of edge-preserving filter & local histogram equalization: (a) Guided filter (b) Domain transform filter (c) Apical Iridix and (d) Our smoothed LHE filter.

3.2.2 Edge-preserving filter

The most popular edge-preserving filter is a bilateral filter which Tomasi et al. proposed in 1998 [85]. However, this filter excessively enhances the edges of objects, and precise gradients of the edges cannot be preserved. Following the success of bilateral filter, many new edge-preserving filters have been proposed in the last few years. He et al. presented guided filtering in 2010 [9]. This filter works based on the variance of the image; when the variance of local image is large, the filter recognizes that there is an edge part in the local image and preserves the initial pixel value. Gastal et al. proposed domain transform filtering in 2011 [6]. In this filtering technique, distance of the gradients, of both the edge and the color difference are transformed to distance in the image space, simple blur filtering is applied on this transformed image. These filtering techniques preserve the exact gradients of object edges. Brief conceptual overview of the guided filter (GF) and domain transform filter (DTF) is presented in Fig. 3.1 (a) & (b).

3.3 Color temperature based perceptual decolorization for tone mapping applications

Grayscale channels, which reflect image luminance, are used for various applications such as printing, tone mapping, data compression, and feature extraction. Thus, obtaining luminance along with human perception has a key role for decolorization, which converts RGB channels to high-quality gray ones. For example, High Dynamic Range (HDR) compression is performed by tone mapping using the luminance channel. However, applying well-known luminance channels such as Y of YCbCr or V of HSV does not guarantee appropriate tone mapping. This is because these channels do not reflect human perceptions. Therefore, decolorization has gathered considerable attention and various methods to achieve it have recently been proposed. These methods can be classified into global and local methods. Global methods can define only one conversion function for all pixels, and most of these methods use all pixels in the image to determine the function. On the other hand, since local ones process the target from neighboring pixels in the same way as a spatial filter, the function is different for each pixel. However, both types of methods face the issue of calculation cost, which comes from optimization iterations or spatial filter processing.

We have developed a fast decolorization method that reflects the perception of warm and cool colors which is well known in psychophysics studies[86]. Colors are arranged according to their wavelengths on the color wheel. The ones with longest wavelengths are on the right side of the wheel and are known as warm colors, as they evoke warmth. These hues include shades of red, yellow, and orange. On the other hand, green, blue and violet which have shorter wavelengths are placed on the left side of the color wheel, are perceived as cool colors. In our decolorization method, we combine the idea of warm/cool colors with the Helmholtz-Kohlrausch effect [87]. On that account, we make two assumptions, which are: (i) warm colors (mainly including R) are lighter than Y of $YCbCr$ and (ii) mixed colors are darker than the Y or L of CIE with the same luminance. To satisfy these assumptions, we use a weighted blending of RGB channels and remap them to warm/cool colors on the luminance channel. The main contributions of our work can be summarized as follows:

- High-speed color mapping is achieved through exact pixel by pixel processing.
- Luminance comparable to that of optimization-based methods is generated.
- Effective luminance distribution is demonstrated by using global tone mapping.

3.3.1 Related work

There are many well defined methods to convert any color image to a grayscale image. An effortless procedure is to assign different weights to color channels, in order to have the same luminance in the grayscale image as the original color image. For example, in the MATLAB function `rgb2gray`, converts any RGB values to grayscale ($Gray$) values by forming a weighted sum of the R, G, and B components as $Gray = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B$. This function operates under an assumption that human visual system is more sensitive to green color. When operating with $CIELab$ and YUV color spaces, one could directly obtain luminance channel as the grayscale version of the color image, because, they consider the luminance and color channel to be independent. But such crude approaches will fail to preserve image contrast as shown in these examples (see Fig. 3.3 and Fig. 3.2).

In several real-world image/video processing applications like detail enhancement, image matching, and segmentation under different illumination a

1-D grayscale image has to be obtained from its corresponding 3-D color image. However, mapping the 3-D color information onto a 1-D grayscale image while retaining the original contrast and fine details is a challenging problem. Additionally, implementing decolorization algorithms with a reasonable computational efficiency is pivotal for realizing their real-time applications. Many studies have been carried out to develop novel decolorization methods. These mapping methods can be categorized into global [88, 89, 90, 91, 92] and local methods [93, 94, 95, 96]. In local mapping methods, the same color pixel within an image could be mapped into different grayscale values depending on its spatial location. Ideally this undesirable as such output images may be perceived as unnatural. On the other hand, in global mapping methods same color pixels within an image irrespective of its spatial location are mapped to same grayscale values. Thus, global methods are more likely to produce grayscale images that are perceived to appear natural.

In the global methods category, Gooch et al. [88] proposed a global decolorization algorithm that can be implemented by solving the optimization problem for all image pixels. Then, Kim et al. [91] aimed at high-speed processing by simplifying Gooch's method. Smith et al. [90] used unsharp masking and the Helmholtz-Kohlrausch effect (H-K effect) model of Nayatani et al. [87]. Nayatani's model [87] is merely an experimental model for the effect of the $CIELUV$ chrominance component for human perception. On the other hand, the method of Lu et al. [92] is focused on converting a color image into a gray image with high contrast. The main advantage of these methods is transformation consistency, i.e., the same color is converted to the same grayscale. However, speed remains a problem for these methods. Most of the local methods are aimed at speeding up the method of Lu et al [92]. To enhance image contrast, Ancuti et al. [94] adopted the strategy of using a Laplacian filter and Song et al. [96] used a Sobel filter. Although local methods are effective in terms of contrast emphasis, they are disadvantageous in terms of tone mapping because conversion consistency is not maintained, and it differs from human perception.

Recently, some machine learning-based techniques have also been proposed for image decolorization [97, 98, 99, 100, 101]. Cai et al. in [98] proposed a method, that used the perceptual loss function to pretrain VGG-Net [102], however it is difficult to control, and many of their output images are far from human perception. Also, their computational cost are high, for processing an image of 256 × 256 size takes roughly 30 seconds on a single Nvidia GeForceGTX 1080 GPU. Zhang et al., proposed a CNN framework that combines local and global image features [101]. However, their network framework does not account for exposure features [97]. In [100] Lin et al. by using a database of 50 images from the Corel data set produced 50 grayscale images using the Color2Gray algorithm [88]. With these 50 input/output image pairs as training examples for their partial differential equations-based (PDE)learning system, they learn color2gray mapping. The PDE system produced images of comparable quality to [88]. However, for an input image of size $n \times n$ their PDE color mapping algorithm's computational complexity is $O(n^2)$. Under these circumstances, a high-speed method for generating grayscale images that accurately captures human perception has not yet been developed. To make tone maps valid, it is necessary to develop such a method.

3.3.2 Definition of our proposed decolorization method

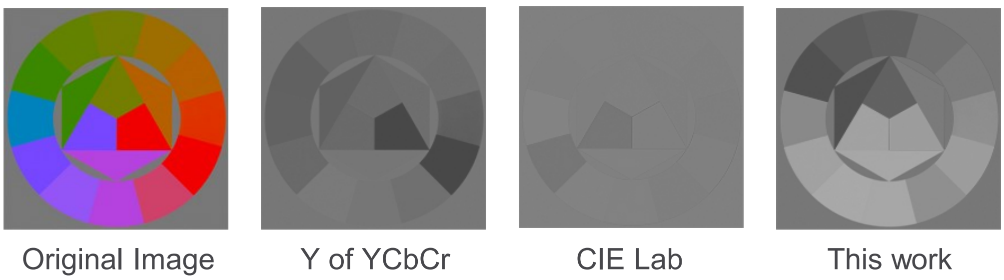


Figure 3.2: A comparison of luminance components obtained using YCbCr, CIE Lab and our proposed method.

Luminance components such as Y of $YCbCr$ and CIE L have been used in various image processing applications; however, they do not accurately reflect human perception (Fig.3.2 and Fig. 3.3). Figure 3.3 shows how the warm colored flower (*Red*) advances towards the eye of the observer, while the background mainly green recedes. Using the luminance channel of the conventional $YCbCr$ color space this phenomenon is absent. But, in our method we are able to capture the warmth R (red) component as perceived in human per-

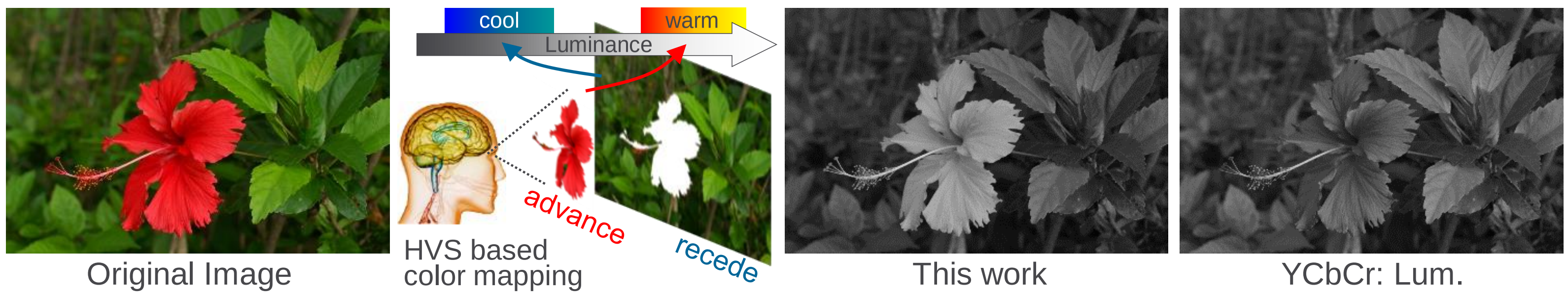


Figure 3.3: Main concept of our decolorization method: Human perception of warm and cool colors. Warm colors "advance" toward the eye, while cool colors "recede". In this work we are able to accurately reflect the human perception of warm color, whereas this phenomenon is non-existent in conventional YCbCr color space.

ception.Figure 3.2 shows that the R (red) and B (blue) components do not come even close to the perception in the luminance component of CIELAB. Moreover, mixed color components such as mud yellow tend to appear dark for people. In this study, we conceived the idea that RGB weight functions for alpha blending can reproduce this phenomenon.

3.3.3 Luminance mapping using red and blue weighting function

We considered the idea of color mapping by combining the warm/cool color and the Helmholtz-Kohlrausch [87]. Psychophysical studies find that, warm and cool colors impact our visual perception of the objects that we see. For example, the red color associated with fire/sun advances toward the eye, creates an illusion of heat and therefore perceived as warmth and comforting. On the other hand, cool colors have reverse effects of warm colors. Receding from the eye of the observer, cool colors reminds of the earthy objects, like meadows and oceans. These hues often are perceived as cool and refreshing [86][103]. In our decolorization method, we developed two weighting functions as shown in Fig 3.4. One for remapping warm colors and the other for remapping cool colors. In our method, actual luminance is defined as the Euclidean distance of weighted warm/cool luminance including the W (white) channel. Essential luminance is given by

$$L_{WHITE} = \sqrt{\frac{R^2 + G^2 + B^2}{3}} \tag{3.1}$$

$$L_B = B \tag{3.2}$$

$$L_{G,R} = \sqrt{\beta_R R^2 + (1 - \beta_R) G^2}, \ (0.5 < \beta_R < 1) \tag{3.3}$$

As we know, $L_{WHITE} : W$ is the Euclidean distance of RGB channels; L_B is the B channel as it is. The warm color function is also defined as the Euclidean distance by β_R weighted R and G(green) vectors. The β_R determines the bias of R components in the replacement; when the β_R becomes higher,

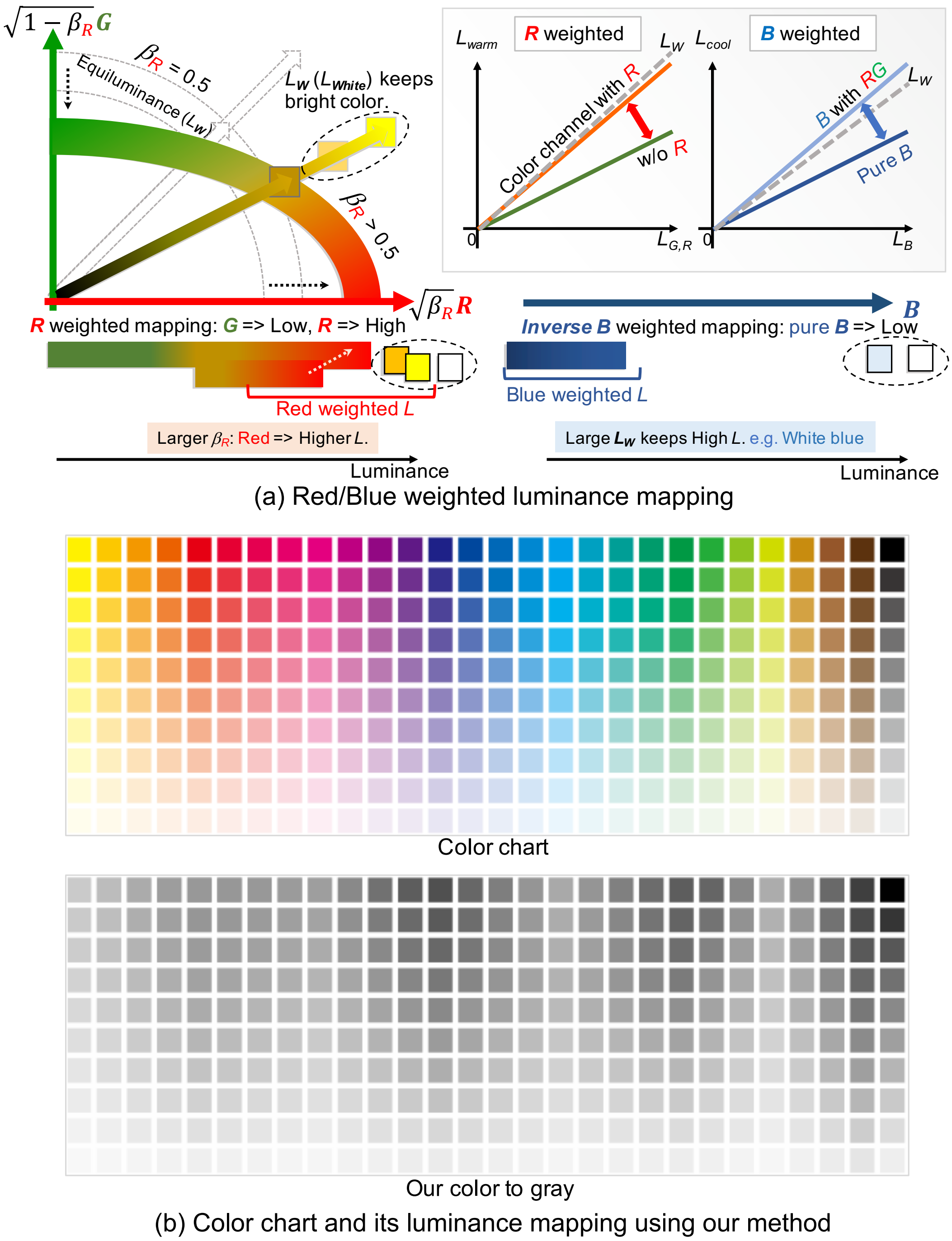


Figure 3.4: Our red and blue weighted color mapping method. Color chart converted to grayscale using our decolorization method.

more R components than G components are rated. The remappings made by the blending function using the R and B ratio of RGB values are given by

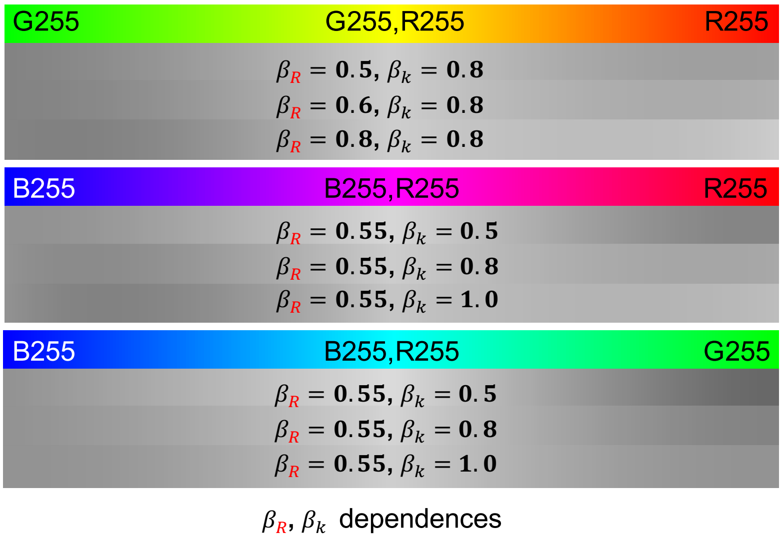


Figure 3.5: Imapct of β_R and β_k on output luminance value. (top) β_k set to 0.8 and for changing β_R we move from Gmax to Rmax. (middle) β_R set to 0.55 and for changing β_k we move from Bmax to Rmax. (bottom) β_R set to 0.55 and for changing β_k we move from Bmax to Gmax.

$$L_{WARM} = \frac{R}{R+G+B} \cdot L_{G,R} + \left(1 - \frac{R}{R+G+B}\right) \cdot L_{WHITE} \tag{3.4}$$

$$L_{COOL} = \left(1 - \frac{B}{R+G+B}\right) \cdot L_B + \frac{B}{R+G+B} \cdot L_{WHITE} \tag{3.5}$$

In the L_{WARM}, using R weighting, we can map $\frac{G}{R}$ to $low/high$ luminance separately. In Inverse B weighting, pure B components are assigned to low luminance in the L_{COOL} . Since both functions are blended with L_{WHITE}, bright orange/yellow and sky blue, which include high white components, are mapped to higher luminance. Finally, we obtain the luminance channel L, which is given by

$$L = \sqrt{\beta_k L_{WARM}^2 + (1-\beta_k) L_{COOL}^2} \; (0.5 < \beta_k < 1) \tag{3.6}$$

In this study, we mainly used warm-color weighting luminance in experiments and set the β_k higher. We set two parameters relating to color component emphasis as follows: ($\beta_R = 0.55; \beta_k = 0.8$). From Fig 3.5 it can be easily understood how β_R and β_k can bias the resulting luminance values.

3.3.4 Evaluation

To evaluate our method, we focused on the following points:

- Comparison with other decolorization methods.
- Processing speed
- Applicability to tone mapping.

Figure 3.6 shows a comparison with other decolorization methods and a quantitative metric listed in table 3.1. For comparison we have used images from Cadik's data set [104] and we chose the recent color-to-gray structural similarity (C2G-SSIM) metric, as it has a good correlation with human perception [105]. The images in column 3 were obtained with the L component of the CIE Lab color space, which is a reversible model reflecting human visual characteristics. L of CIELAB is similar to base luminance Y of YCbCr where the luminance is mapped along with the color order of Y. In the image (col 3 row 4), the luminance of all colors is mapped to same L value. Here, appropriate conversion is required; our method can generate contrast for this image. For global decolorization methods, which means only one conversion function is applied to each pixel. For example, in these methods, optimization techniques referring to whole pixels are applied for luminance conversion along with the brightness of human color perception. In our quantitative assessment Gooch et al. [88] obtained highest average C2G-SSIM score. However, a step artifact occurs in the gradation image (col 4., row 2). Kim et al. [91] proposed an improved version of Gooch's method that achieves high-speed optimization and reflects the H-K effect [87]. In this method, since weighting is applied for expanding luminance distribution of whole pixel colors in an image along with chrominance, its conversion becomes different in each image. Thus, over-enhancement is observed in the images (col 5, row 1) and (col 5, row 4). The output images of our method are similar to those of Smith et al. [90]. This

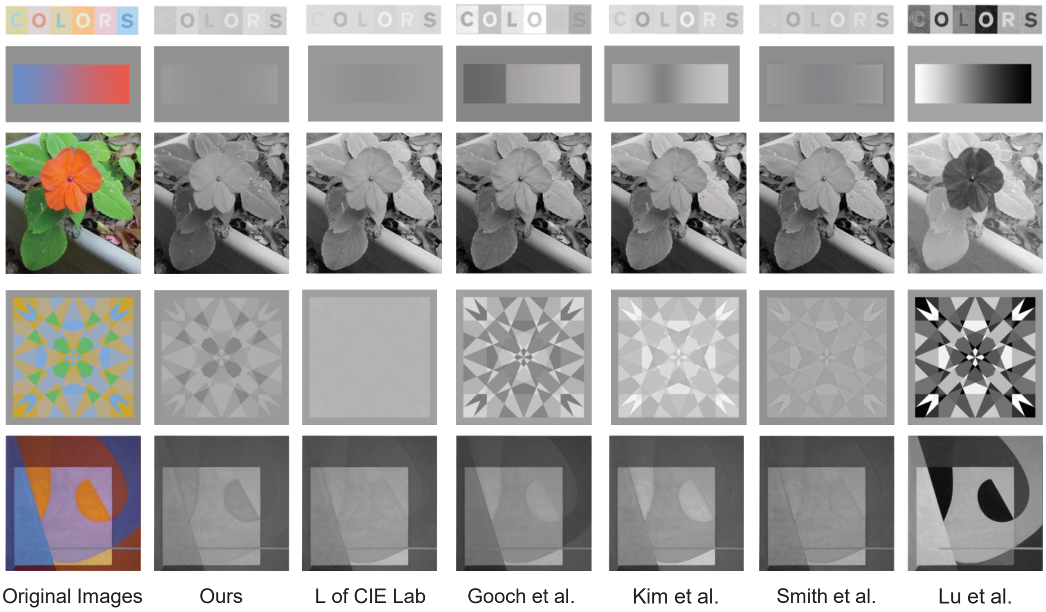


Figure 3.6: Comparison with other decolorization methods. Citations of the methods are the same as the table below.

Table 3.1: C2G-SSIM measurement for images in Fig. 3.6

Image	Our's	CIE Lab	[88]	[90]	[92]	[106]
1	0.8231	0.7309	0.8280	0.7248	0.8437	0.7237
2	0.9380	0.9317	0.9652	0.9395	0.9895	0.9026
3	0.7411	0.7499	0.7397	0.7530	0.6580	0.6907
4	0.9158	0.8071	0.8916	0.8081	0.9711	0.8647
5	0.9274	0.9315	0.9321	0.9292	0.8411	0.9021

method also uses the H-K effect but requires a lot of processing time for post-unsharp-mask filtering. In Lu et al. [92], their method does not reflect human perception; they try generating high contrast images for mask images that are input to an edge preserving filter such as a guided filter.

3.3.5 Perceptual evaluation

For perceptual evaluation we conducted a user study with two objectives:

- Evaluate our decolorization process by measuring accuracy and preference [104, 107].
- To evaluate effectiveness of our decolorization as a pre-processing tool.

Table 3.2: Runtime comparison table with other decolorization algorithms.

Algorithm	Processing time(ms)	Image size	CPU(GHz)	Optimization	Process*	Normalized time* (μs)
[88]	25.7s	200×200	-GPU-	$\surd$	G	N/A
[91]	102	320×240	2.66	$\surd$	G	1.30
[90]	6.7s	570×593	3.0	×	G+L	22.02
[92]	800	600×600	3.80	$\surd$	GC	3.12
[96]	40	320×240	N/A	×	LC	N/A
[94]	100	800×600	2.5	×	LC	0.19
CIE Lab	25.57	800×600	2.7	×	G	0.053
Low res†	**16.71**	800×600	2.7	×	G	**0.034**
High res†	**202.05**	3008×2008	2.7	×	G	**0.033**

† Implemented in this work. * G:Global L:Local C:Contrast.
* Normalized time is the processing time normalized by frequency (2.7 GHz) and divided by the number of pixels. It indicates the effective processing time per pixel.

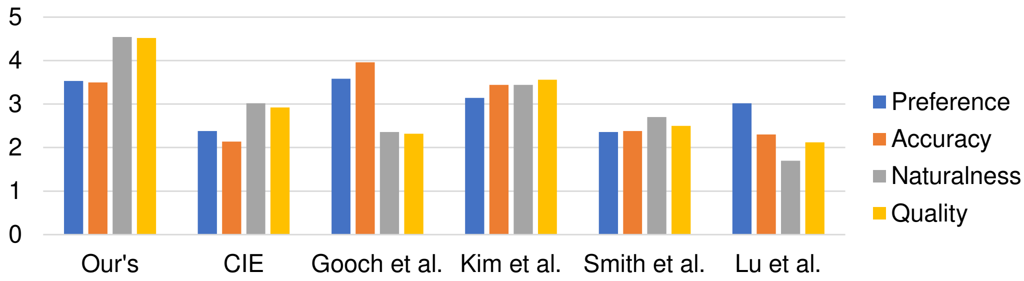


Figure 3.7: Subjective Visual Test

We measured the naturalness and overall quality of the tone mapped images. For tone mapping we chose a MATLAB algorithm of Shan et al. [108] as they can tonemap standard images and by simple parameter settings. Our study group of 15 students were shown six sets of images (Fig. 3.6). They were asked to assign points to these images, on a scale of 1 (low) to 5 (high) for accuracy and preference. Color tone mapped images for the same set were evaluated for naturalness and perceptual quality (see supplementary material). In our test setup we used a Microsoft Surface Pro 7 set to its native resolutions and the lighting of the test room was slightly dim. The compiled response of the volunteers is presented in Fig. 3.7.

Table 3.2 lists the processing speed of various methods, our implementation and CIE Lab were coded in C++; and executed on an Intel Core i5-

5257U (2.70GHz) CPU without any multicore, multithread or SIMD operations. The results confirmed that our method had the fastest run-time among the compared methods. It is notable that it exceeded CIE Lab in run-time; which indicates it also has advantages in total calculation cost including post-processing. Its computational complexity is only $O(1)$ because it performs exact pixel by pixel processing, referring only to the RGB value at each pixel. Among other methods, the one developed by Gooch et al. was reported to have $O(n^4)$ computational complexity. Using $O(1)$ spatial filtering has made it possible to develop $O(1)$ local methods [94, 96], but the filter calculations required degrade their run-time in comparison with our method. Our method has demonstrated faster run-time than local methods by maintaining global coherence, which means the conversions were the same in all pixels.

3.3.6 Limitation of our decolorization method

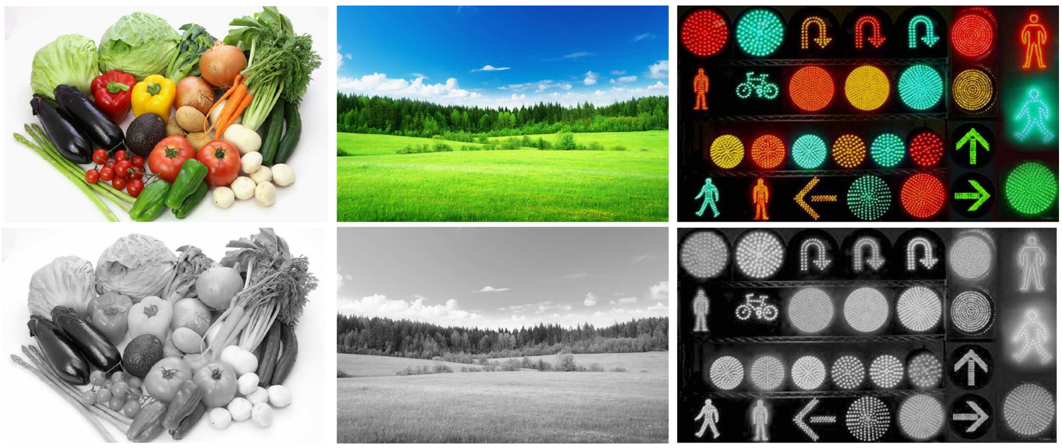


Figure 3.8: Examples to demonstrate limitation of our decolorization method when mapping bright green/neon color.

In our method, pure green is likely to be mapped to a dark luminance value. Light green will also be mapped to a lesser dark luminance part (see Fig. 3.5). Therefore, certain scenes are likely to be perceived as unnatural. For example:

1. Vegetables (e.g. cabbage, etc.)

2. Green meadows under bright sky.

3. Bright green neon lights.

However, in Fig. 3.8 we perceive them as natural. We postulate the following as the possible reasons:

- There is rarely any pure bright green (like G255) scene in nature.
- In our color space, green color with white components follow Eq.(3.1).

Therefore, for the above reasons the color keeps a balance among other color channels.

3.4 Smoothed LHE Algorithm

LHE-based local tone mapping converts target pixels by using tone curves constructed from local cumulative histograms. Apical Iridix ^{TM} [7] now a subsidiary of ARM is a very popular local tone mapping algorithm that has dominated the market. It has been licensed by digital camera manufacturers like Nikon and Sony. The Apical Iridix, uses Legendre basis functions, and its frequency spectrum analysis to shape up the local histogram this is illustrated in Fig. 3.1(c). Following which, the local cumulative histogram is also directly reconstructed by utilizing the product of sum, of the integral form of the Legendre basis functions. By controlling the number of the basis functions, smoothed local cumulated histogram can be obtained with only low frequency components of the histogram shape. These cumulative histograms have some ripples; thus, alpha blending with the initial image is required.

In the proposed smoothed LHE, we use a bin reduced intensity stacked histogram [4] and sum of sigmoid curves as basis functions as illustrated in Fig. 3.9. Smoothed local tone curves without ripples are directly calculated; all local medians are mapped to constant value of 127 by the generated tone curves, which are loci of the function's midpoints. This means that, the spatial low frequency components can be completely converged and removed by our functions. Arranging the loci then becomes reconstruction of the low frequency components which are related to global tone control; limiting of the tone mapping space is utilized as the gain control of the local functions. Summary of our smoothed LHE (SLHE) algorithm is shown in Fig. 3.9. It is worth

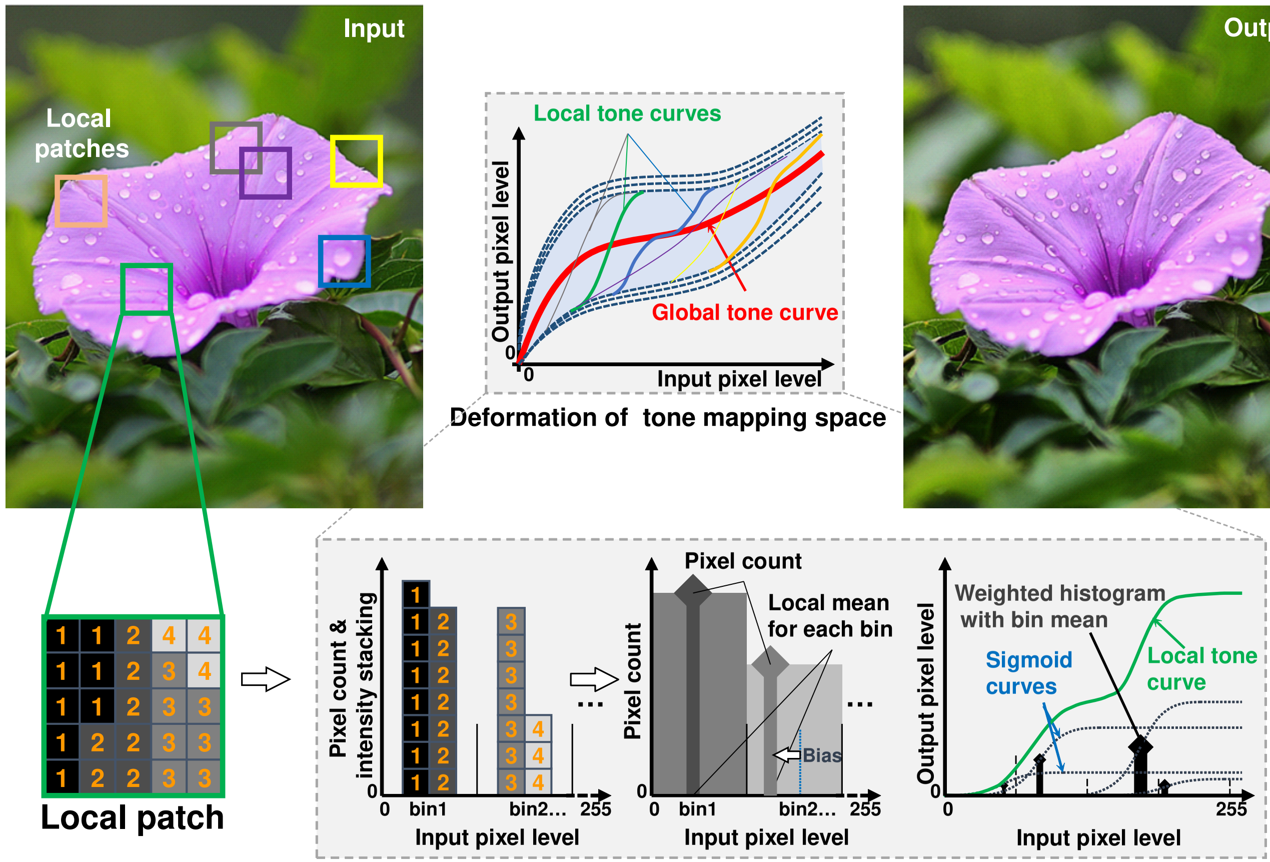


Figure 3.9: Proposed algorithm's conceptual framework. We directly calculate smoothed local tone curve as sum of sigmoid functions, for each local patch which is computed from intensity stacked histogram [4].

noting that, in Apical's algorithm the local tone curve is realized using Legendre's polynomial and in our SLHE filter we make use of the sum of sigmoid functions as shown in the figure 3.9.

Mathematical framework of our algorithm is as follows: For a local patch Ω_j, local tone curve $SLHE_{\Omega_j}$ is computed using Eq. 3.7. Where the pixel count P_{count,B_k} is that of the bin B_k, and whose mean is L_{Mean,Ω_k,B_i}

$$LT_{SLHE,\Omega_j} = \sum_{k=1}^{N_B} \frac{P_{count,B_k}}{1 + exp^{-\left(P_{in} - Mean_{B_k}\right)}}, \tag{3.7}$$

LT_{SLHE,Ω_j} is directly generated from sum of sigmoid functions. In our bin reduced intensity stacked histogram computation we avoid accuracy degradation by calculating mean of each bin which represents the bin bias [4].

$$\begin{aligned} LT_{SLHE,\Omega_j}\left(L_{median,\Omega_j}\right) = CumLH_{SLHE,\Omega_j}\left(L_{median,\Omega_j}\right) \\ = 0.5R_{out} = const, \end{aligned} \tag{3.8}$$

Since the LT_{SLHE,Ω_j} approximates smoothed cumulative histogram $CumLH_{SLHE,\Omega_j}$, local medians L_{median,Ω_j} ($\simeq$ local means), which are center of the local function LT_{SLHE,Ω_j}. They are mapped to constant value which is half of the output range $0.5R_{out}$. Thus, we can obtain global tone curve from the loci of its center points (see Fig. 3.1(d)).

$$\begin{aligned} F_{SLHE,\Omega_j}\left(P_{out}\right) = GT\left(P_{in}\right) + G_{limit} \cdot F_{SLHE,\Omega_j}\left(P_{in}\right) \\ - 0.5R_{out}. \end{aligned} \tag{3.9}$$

Center offset of local tone is denoted as Global tone $GT(x)$ and range of the local tone is denoted as gain G_{limit}. When the gain increases, local contrast also increases. In this way, we realize a smoothed LHE-based on sum of basis functions with global tone control.

3.4.1 Noise characteristics of SLHE filter

Our SLHE tonemap operator uses two curves, one corresponding to the edge region and the other for gradation. Noises in gradation part are much more

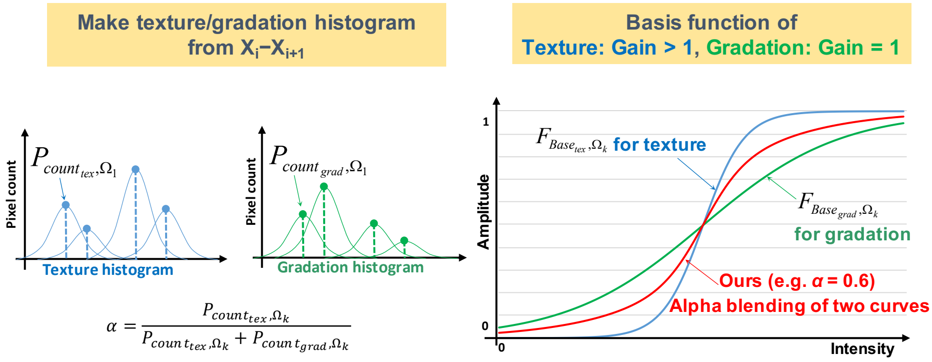


Figure 3.10: Our noise suppression method. (a) Input histograms. (b) Our tone blending function.

noticeable than in the edge and texture part. In the histogram estimation we detect the edge in Y-direction ($P_x - P_{x+1}$) and setting a threshold ≥ 5. Using a simple weighting function described in Eq.3.10 we blend both the tonal curves as shown in Fig. 3.10.

$$Output = \alpha \cdot Y_{edge} + (1-\alpha) \cdot Y_{nonedge}$$
$$where, \alpha = \frac{No.of edgepixels}{Totalpixels} \tag{3.10}$$

In Fig. 3.11 we illustrate the effectiveness of our simple noise suppression technique. We measured the noise variance (NV) using the Laplacian operation described in [109]. Measured NV upon filtering the input test images using our algorithm is presented in table 3.3, which illustrates that the simple weighting function given in Eq. 3.10 can effectively suppress noise.

Table 3.3: Noise suppression performance of our SLHE filter compared to other edge-preserving filters.

Image	SLHE	WLS [5]	DMT [6]	Guided [9]
1	3.3871	11.2545	5.5095	9.2576
2	0.0216	0.2054	0.0855	0.1763
3	2.1140	6.5773	3.7016	4.5315

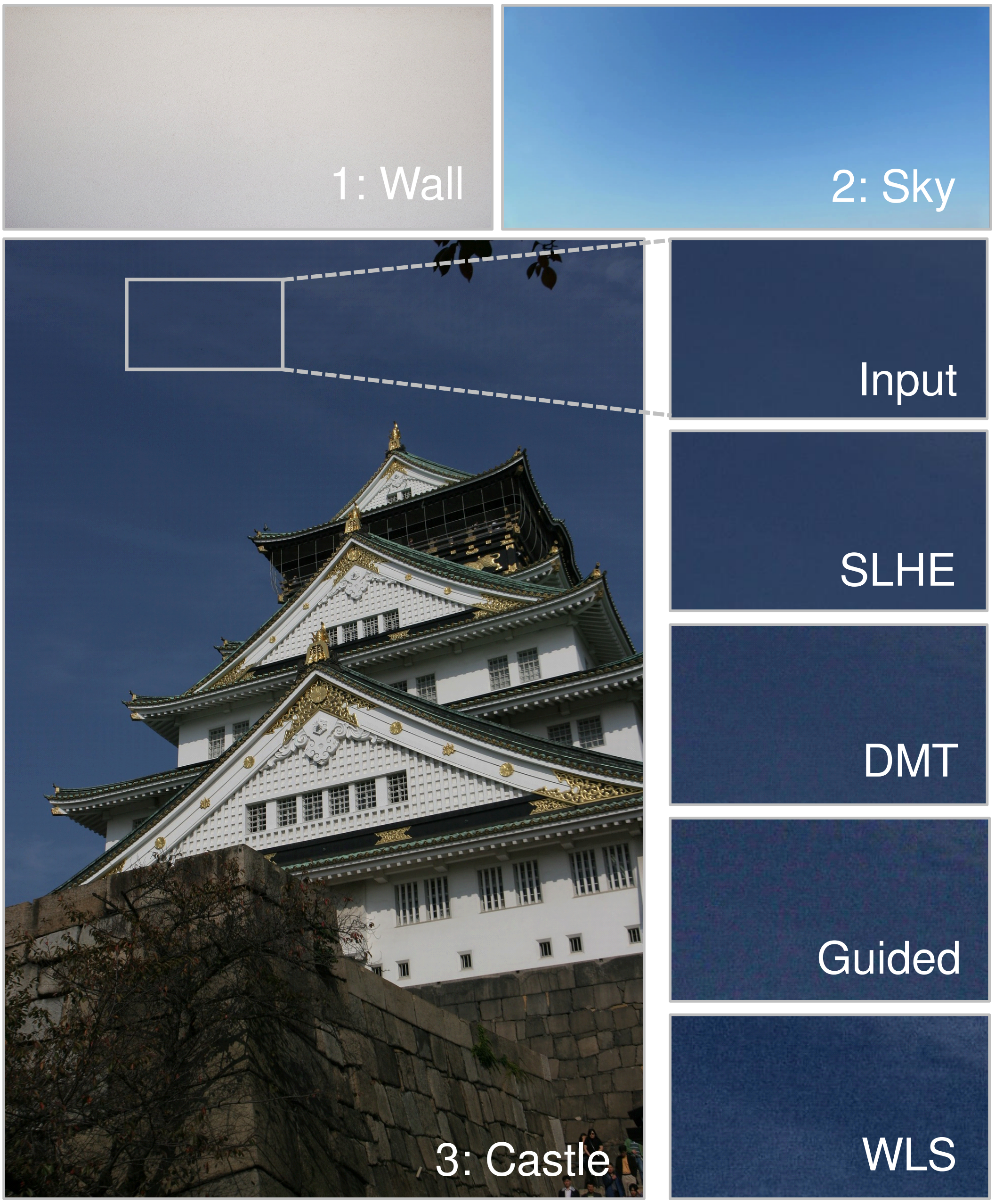


Figure 3.11: Image quality assessment: Noise variance estimation using different filters. Test images used for assessment in table 3.3

3.4.2 SLHE halo control

Halos occur in polarized luminance distributions. In smoothed LHE, as the local window moves across the valley between two adjacent distributions, the input pixel level becomes different from the local mean. Thus, the target pixel level is mapped to a lower or higher level than the input in the low-level or high-level distribution. Halos should generally be removed, and several halo suppression methods have been proposed for use in multilayer methods using some edge-preserving filter. However, in these methods the contrast at the object edges are degraded. Kimura, et al. have proposed individual control of light/dark halos for preserving local contrast [110]. Their study reported that only light halos are required to be suppressed as they appear unnatural to humans. In our TMO, we apply weights to the tone mapping functions in order to control light/dark halo while preserving natural local contrast. The input pixel in polarization luminance distribution is checked to see whether it belongs to the light/dark luminance group by comparing with local mean. If it belongs to the dark group its weight is set to a large value (W_{Dark}), and a smaller value (W_{Light}) for pixels from light group. Figure 3.12 shows the weights for light/dark regions along the edges. We assessed the effectiveness of our halo management method; in Fig. 3.12(e) we show halo suppression in tone mapped image.

3.4.3 $O(1)$ Time exact tone mapping

The conventional tone mapping methods (Fig. 3.13(a)) use overlapping pixels between the current pixel block and the previous block (to the left of the current block). Such methods would require O(N) steps. Here, we focus on pixel access with the "Box-filtering Method" [111]. This method achieves a 2D-moving-average-filter processing in $O(1)$ time with only 2 pixels access for each pixel block. A $O(1)$-median-filter was recently proposed using this pixel access technique [112]. In our implementation, we use a line buffer to store the column local histograms of the previous line and calculate the column histograms of the current line from the histograms of the previous line. Let us consider a case of moving a pixel block from the left of one pixel to the right. The column histogram with 2N+1 pixels is updated by adding/subtracting one new pixel to/from it. The column histogram is calculated using only one addition and subtraction (Fig. 3.13(b)). However, in the HW implementation, this method requires not only the temporary line buffer but also extra buffers,

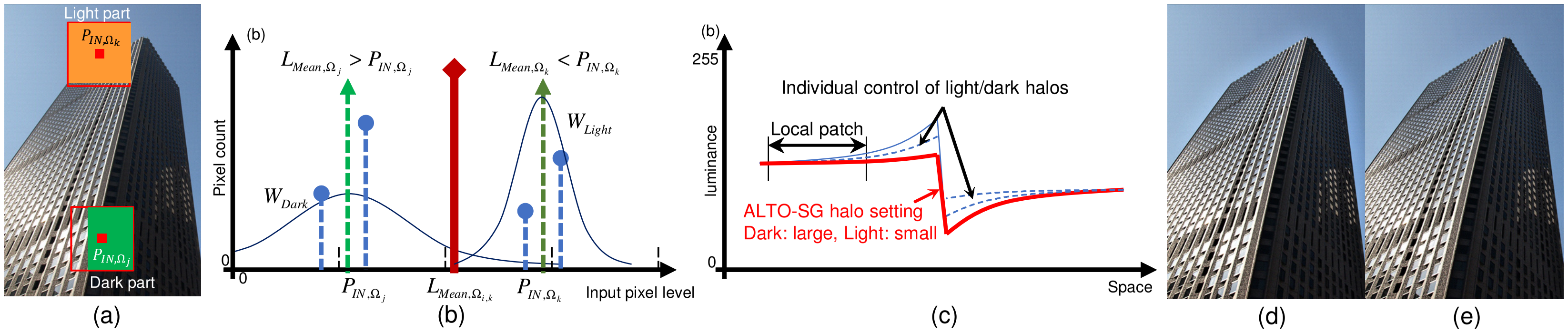


Figure 3.12: Halo control in our method: (a) Check whether P_{IN} belongs to light or dark luminance group. (b) Individually assign weights to light and dark luminance groups. (c) Default edge form of our tonemap function; the dark halo have a positive effect on the contrast enhancement. (d) Image with halos. (e) Halo suppressed using our TMO.

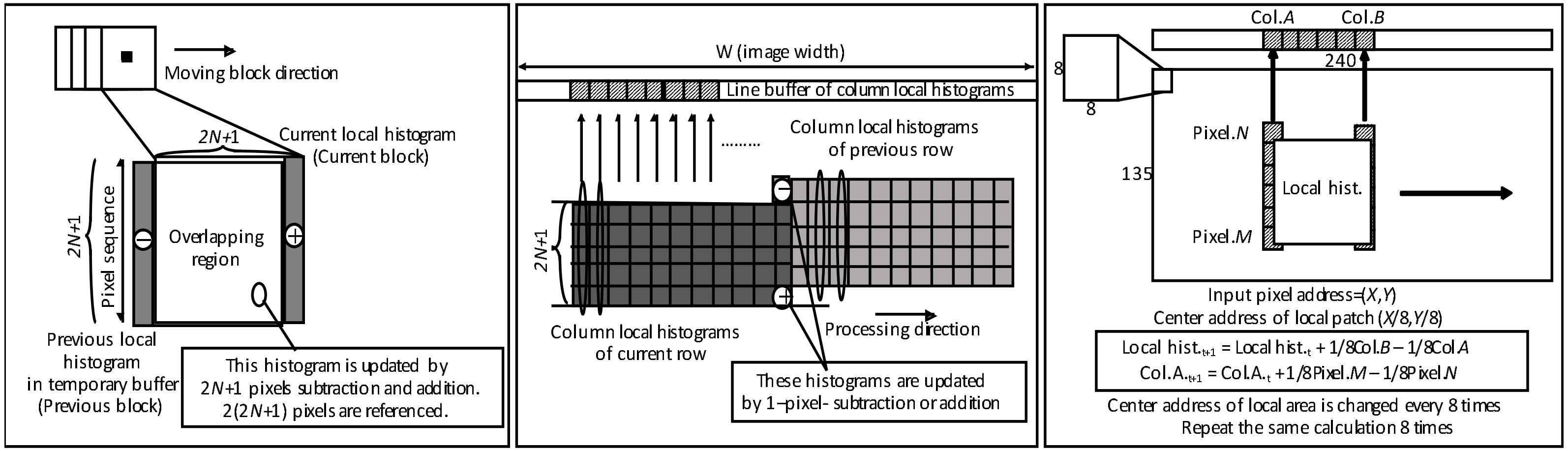


Figure 3.13: $O(1)$ image processing scheme: (a) Conventional local histogram estimation with box shape and, thus box filtering can be executed with $O(1)$ computation; (b) Updating column histogram with line buffer by one-pixel subtraction and addition; (c) $O(1)$ local histogram estimation using downscaled frame and line buffers and sub-pixel linear interpolation on the estimation scheme.

whose size is $2N + 1 \times W$, for a pixel addition and subtraction. For our HW implementation, we improve this method and propose $O(1)$ processing by using a compressed frame buffer. In this one-eighth compressed frame buffer the compressed column histogram is calculated repeatedly as shown in Eq. 3.11 and Fig. 3.13(c).

$$\begin{aligned} Col.A_{t+1} &= Col.A_t + \frac{Pixel.M}{8} - \frac{Pixel.N}{8} \\ Local.H_{t+1} &= Local.H_t + \frac{Col.A_{t+1}}{8} - \frac{Col.B_{t+1}}{8} \end{aligned} \tag{3.11}$$

In this equation, *Col.A, B* are column histograms; Pixels M, N are pixels from the compressed frame buffer; *Local.H* is the temporary local histogram. Repeatedly using the same pixels for the histogram estimation, gradually changes the ratio of the pixels in the column. This leads to a linear interpolation between next column values. Also, the overlapped local histogram estimation shifts to the next position with linear interpolation at each 1/8 pixel step. This calculation is independent of filter-kernel size, therefore, we can achieve $O(1)$ local tone mapping with low memory usage. From Ambalathankandy et al.[113] study it was established that using large kernel size when tone mapping images produces images with a natural feel, and it is also known that LHE's can support large kernels. According to [113] LHE's with large kernels operate like an active high pass filter boosting and balancing all frequencies in the kernel.

3.5 Image quality assessment

Image quality assessment (IQA) helps in determining the actual quality of experience. Thus, playing an important role in ascertaining the worthiness of any imaging system. Subjective user study is the most reliable means to measure image quality. However, it is not always feasible for practical reasons as subjective assessment is a time and resource consuming exercise. On the other hand, objective assessment methods can automatically evaluate the quality of images and can approximate human quality judgments. Over the years, many IQA methods have been proposed which have been covered in detail in the following surveys [114, 115].

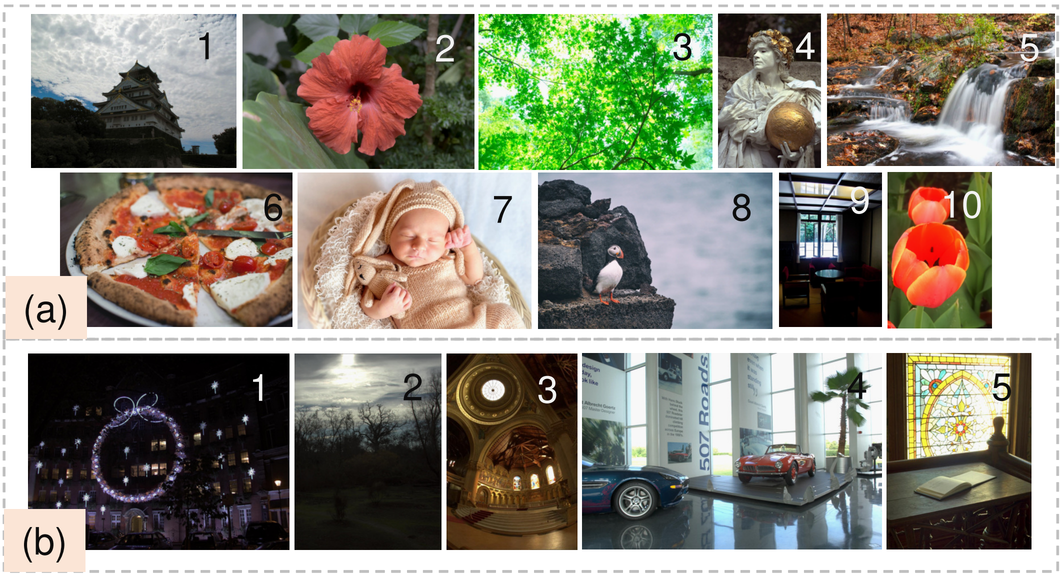


Figure 3.14: Image quality assessment. (a) Structural similarity index measurement: Input test images. (b) Tone Mapped image Quality Index: Input test images (linear log-luminance compressed)

Table 3.4: Experimental result of SSIM measurement on test images.

Image	This work	WLS [5]	DMT [6]	Irdix [7]
1	0.9848	0.6370	0.9583	0.8705
2	0.9507	0.6411	0.9358	0.8986
3	0.9923	0.7165	0.9752	0.9875
4	0.9829	0.6900	0.9609	0.9776
5	0.9503	0.5947	0.8928	0.9332
6	0.9510	0.5816	0.9205	0.9470
7	0.9814	0.6961	0.9493	0.9769
8	0.9752	0.7455	0.9374	0.9449
9	0.9745	0.6239	0.9305	0.9484
10	0.9528	0.6270	0.9015	0.7450
Average	0.9696	0.6510	0.9406	0.9357

3.5.1 Objective IQA

For objective assessment we have performed the well-known structural similarity index measure (SSIM) [116] and measured tone mapped image quality

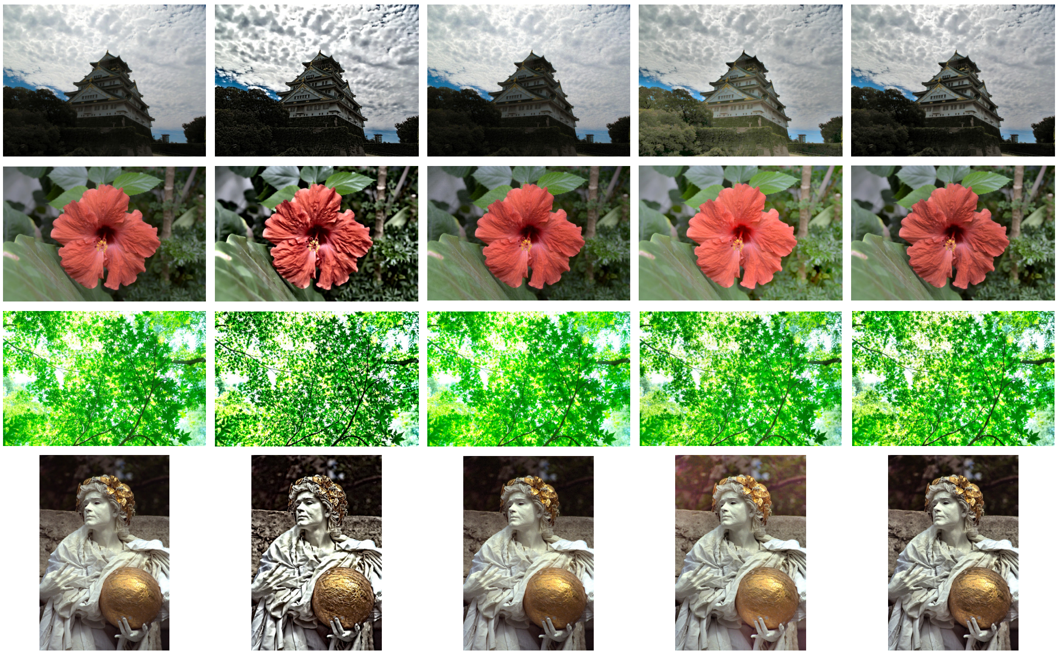


Figure 3.15: Comparison with other methods (Local Contrast Enhancement): (a) Column of original images. (b) WLS [5] (using default setting); (c) DMT [6] (each layer is similar to WLS); (d) Apical's Iridix® [7](default setting, base functions =5, $\alpha = 0.5$) ; (e) This work (range of the tone mapping space (Max = 100): Shadow = 10 to 60, Middle = 20 to 80, Highlight = 40 to 90). Relative local patch size is w = 31 at VGA resolution (d,e).

Table 3.5: Experimental result of TMQI measurement on test images.

Image	This work	WLS [5]	DMT [6]	Irdix [7]
1	0.9547	0.8250	0.8712	0.8219
2	0.9200	0.7596	0.8800	0.8783
3	0.9400	0.8080	0.8805	0.9054
4	0.9152	0.8247	0.8757	0.8800
5	0.9500	0.8203	0.9051	0.9045
Average	0.9360	0.8075	0.8825	0.8780

index (TMQI) scores [117]. In Fig. 3.15 left column images (a) are the original, and next column (b) images have been processed using the detail extraction method of WLS TMO; column (c) images were processed with the detail extraction using the domain transform filter (DMT); column (d) images were modified by using Apical Iridix®and the images in the rightmost column (e) have been modified by our TMO. In WLS method, it uses two detail layers; the parameters given in the references were used while executing their MATLAB code. The images obtained using domain transform filter also has two detail layers and we set these layers with the same visual appearance. WLS method exhibits over-enhancement, while contrast of the DMT processed images appear weaker. Primary limitation of these multilayer methods lies in the fact that they only use few layers. This approach limits it to specific frequency components, and only these frequencies are enhanced to obtain output images. For example, test image 5 has lot of high frequency components and for which DMT filter output has a lower SSIM value. In the Apical Iridix®, which uses the default blending ratio α with the original image, as defined by Chesnokov [7] was used in obtaining the output images. The noticeable degradation in the Iridix processed images are due to its inability in adjusting the DC offsets, and failure in suppressing halos. Our method achieves well balanced tone mapped image, with a natural appearing vivid color with good local contrast enhancement and no visible halos. Distortion in images are measured by SSIM index based on three factors: loss of correlation, contrast and luminance distortion [116]. We used ten images shown in Fig.3.14(a) to compare the performance of our algorithm with other tone mapping algorithms. SSIM values obtained using the test images shown in Fig. 3.14(a) are presented in table 3.4. SSIM value closer to 1 means lesser contrast and luminance distortion, and from the table 3.4 we can notice that our tonemap operator has consistently higher SSIM values. Comparison of TMQI scores (Tone Mapped Image Quality Index [117]) with respect to other filters are given in table 3.5 (higher value is better). TMQI measures the structural fidelity and statistical naturalness, hdr test images used in our experiment is shown in Fig. 3.14(b).

3.6 Conclusion

In this chapter we presented a $O(1)$ global and locally adaptive tone mapping algorithm which is inherently fast in execution and can process full HD images in real-time. Our LHE based algorithm controls global and local characteristics individually. In contrast to other tonemap operators, our algorithm manages light/dark halos separately and by using local tonemap function alone, it is able to effectively suppress noise. Traditionally, tone mapping algorithms operate on luminance channels which can lead to some loss of information. We minimized this data loss by employing a human perception-based color to luminance mapping scheme. We proposed a light-weight and high-speed image decolorization method based on human perception of color temperatures. Our grayscale conversion demonstrates that warm colors are lighter than cool ones by using a blending function with R and B channel weighting. Our optimal color conversion method produces luminance in images that are comparable to other state of the art methods. We validated this by a user study and found that our color to luminance mapping achieves effective luminance distribution and is very suitable as a pre-processing step for tone mapping application, thereby making our TMO operator ideal for many practical applications. We performed detailed objective image quality assessment using two well known metrics SSIM and TMQI to determine the actual quality of experience.

Chapter 4

Characteristics of SLHE tonemap visualization

4.1 Introduction

To replicate human visual perception, we analyze processing images with optical illusion using edge preserving filters and smoothed local histogram equalization (LHE). Images with the optical illusions are good models for gradual/rapid changes in contrast and strong edges, which are good cases for assessing the robustness of image filters. Here, we study and analyze the performance of smoothed LHE filters while processing perceptual illusion. Our studies conclude that, smoothed LHE's are useful in retaining actual edge forms in these images as they can operate using large kernel sizes. These large kernel size filters can construct sawtooth like edge and it corresponds to adequately wide halos. We also demonstrate the usefulness of smoothed LHE like tone mapping techniques in preserving naturalness, and we confirmed it by performing subjective visual test.

Following are the main contributions of this chapter:

- Introduction to lateral inhibition in the human visual system.
- Processing images with optical illusion.
    - Study on filter kernel size influence in preserving wide halo and keeping the good edge form.

- Analyze and compare multilayer filters versus proposed smoothed LHE filter.
- Subjective user study to demonstrate the effectiveness of SLHE filter for processing images with illusion.

4.2 Lateral Inhibition in the Human Visual System

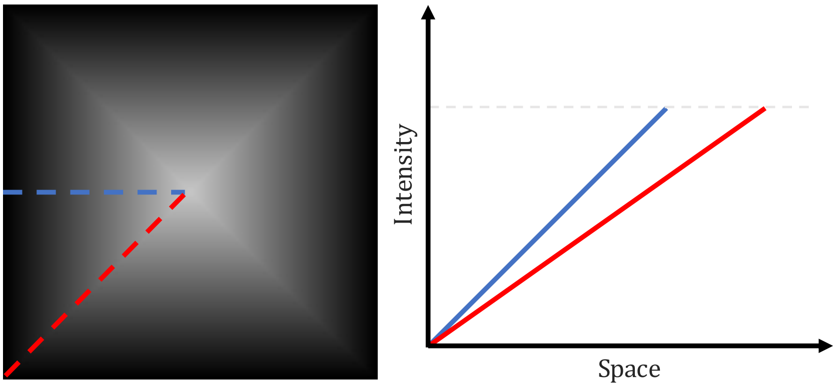


Figure 4.1: The brighter bands at 45° and 135° are illusory. This illusion is one of the clearest demonstrations of the difference between physical stimulation and the perception. The two graphs represent similar luminance profile along the diagonal and horizontal path. But we clearly perceive brighter profile along the diagonal.

Human visual system (HVS) is complex, and it requires numerous components in our eye, optic nerves and brain to work together to enable us to see and perceive the visual information. The moment we focus on a scene, like a camera our eyes capture light through the cornea. From the cornea, the light passes through the pupil which is an opening in the iris. The iris is like a controller and it regulates the amount of light that passes through the lens, which focuses light rays onto the retina. The retina translates an optical image into neural impulses by the patterned excitation of the color-sensitive pigments of its rods and cones (photoreceptor cells). The optic nerves carry these light signals to the visual cortex part of our brain to process these signals for final vision interpretation. However, there is a difference between the image we see and the actual image we perceive [118]. When viewing images, we tend to assume that what we see is a real and un-distorted anatomical reality. In Fig. 4.1 we can see an illusion created by Mach bands, the brighter bands along the 45° and 135° diagonals are illusionary. In the accompanying plot, we can see that the actual intensity values along the horizontal (blue) and diagonal (red) paths are similar.

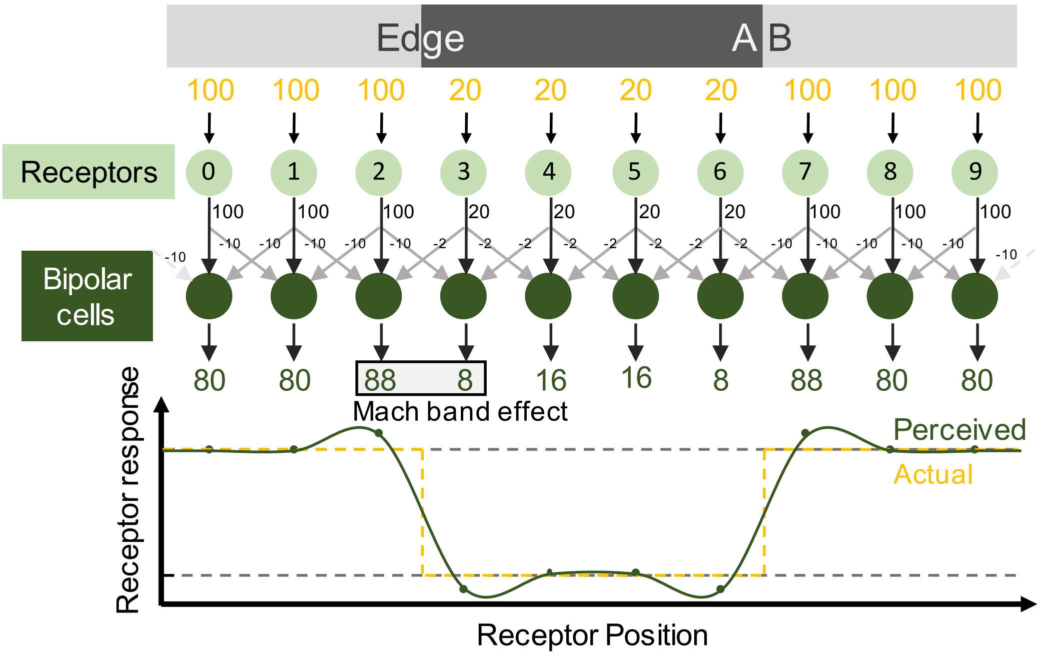


Figure 4.2: Illustration of the Mach band illusion. The light stimulus is a light-dark-light bars of equal size, whose actual luminance intensity is flat when plotted (yellow). The perceived luminance distribution is different (green). A hypothetical model of photoreceptors is presented. We assume that each photreceptor inhibits the adjacent neighbor by 1/10 of its own activation.

The optical illusion in Fig. 4.1 can be explained with the neural processing effects on perception. There is a total of 126 million receptors (6 million cones & 120 million rods) in the retina that converge to just 1 million ganglion cells [119]. On an average, about 120 rods converge to a ganglion cell, as opposed to 6 cones. Neural circuits formed from the convergence of these photoreceptor cells do affect our perception, which we will explain using the illustration in Fig. 4.2. The photoreceptors in retina converge to ganglion

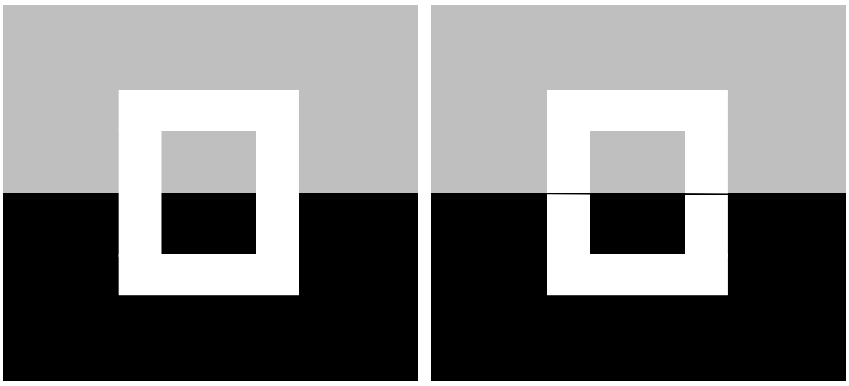


Figure 4.3: Background contrast effect: (a) The Benussi ring, here the ring appears to be of uniform brightness. (b) When a line is drawn along the background boundary, we perceive the ring to have different brightness levels.

cells, creating an excitatory response. These connections pass through the horizontal and amacrine cells which exhibit inhibitory responses. These two features together provide retina with an excitatory-inhibitory property [120]. In Fig. 4.2 the actual intensity of the light stimulus is plotted in dashed yellow line, and it has a flat distribution. However, when we look at the edges, we perceive a small dark band at A and a small light band at B. The perceived signal strength is plotted in green, and it is not flat around the edges. These are Mach bands, and they serve an important function of enhancing edge detection. This phenomenon can be explained using the receptive model shown in Fig. 4.2. Each of the ten receptors send signal to bipolar cells, which inhibit its neighbor on both sides. The strength of inhibition depends upon the strength of the incoming light stimulus. The inhibition caused by the receptors (3/6) stimulated by a lower intensity light is less than those stimulated by the higher intensity (2/8) causing the Mach bands.

Another illusion caused by the lateral inhibition is shown in Fig. 4.3, here the gray ring laid on two different backgrounds appears homogeneous (see Fig. 4.3(a)) until it is split in two halves by a line. Now a contrast effect shows up and the ring appears segregated into two perceptual groups (Fig. 4.3(b)). The Mach band effect illustrated in Fig. 4.2 shows how discontinuity in stimulus signal (edge) gets exaggerated in our perception due to lateral inhibition. In this chapter, we will describe and demonstrate the drawbacks

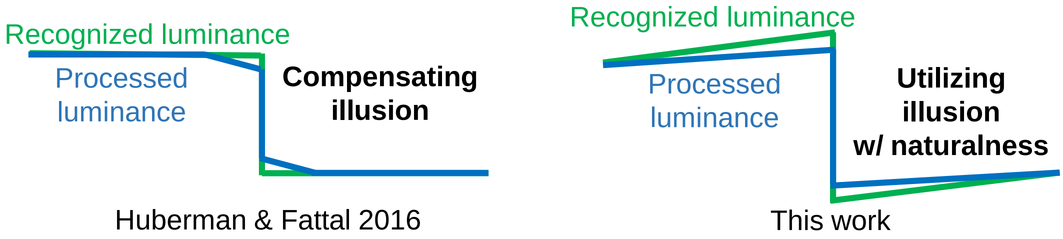


Figure 4.4: (Left) Illustrates the luminance profile of a recent work [8] and (Right) The luminance profile we reconstruct in this work.

of using multilayer methods for processing images with optical illusions.

4.3 Processing images with optical illusion

Lateral inhibition (LI) is a mechanism that assists human in perceiving edges and contrast illusions. LI is an important factor to be considered for local adaptive tone control. Huberman and Fattal in [8] studied the visual biases caused by lateral inhibition and they suggest a compensation technique to counter this visual bias (as illustrated in Fig.4.4). And, they do state that their method suffers from some form of blur, while being able to reduce the bias. From our earlier observations, we noticed that image processing methods which directly reflect LI gives unnatural feeling, when processing illusion images. We propose using LHEs for the following reasons:

- Preserving wide halos can suppress stacked LI effect, which otherwise results in over enhancement.
- Large kernel size plays dominant role in preserving wide halo and keeping the good edge form.
- Latest LHE methods can adequately support large sized kernels.

When filtering Chevreul illusion, guided filter and domain transform filter using their recommended kernel sizes cannot preserve hard-edge profile perfectly and generate slight halos (Fig.4.13). The halos break edge profiles and become barriers against reconstruction of illusion with a natural feeling. On the other hand, LHE group Apical and SLHE with the recommended kernel size can emphasize without stacked LI because they can maintain flatness. In Koffka

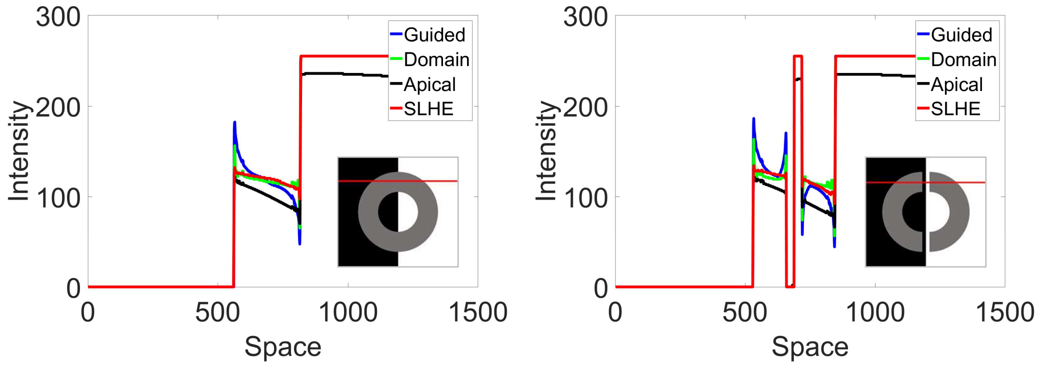


Figure 4.5: Processed cross section of Kofkka's ring.

ring illusion, separation feeling in the ring with line and the feeling of uniformity in the connected ring can be easily evaluated. We understand that the flatness of ring interior causes the uniformity feeling over LI. The LHEs preserve this flatness, and the spatial filters show discontinuity near edges as it can be seen Fig.4.5. Preserving the flatness depends on kernel size; large kernel size is suitable for keeping edge profile with flatness.

4.4 Analysis: Multilayer filters versus Smoothed LHE filter

In a multilayer tone-control method, the base layer, which consists of low frequency components, is considered to contain the gradation and/or global luminance elements of the image, and the detail layer (P_{OUT,Ω_k}), which has only high frequency components, is considered to contain the texture and/or reflectance elements of the objects. These two layers are separated by dividing or subtracting the low-pass filtered image from the input one (P_{IN,Ω_k}). With subtraction, the resulting detail layer is given by

$$P_{OUT,\Omega_k} = P_{IN,\Omega_k} - LPF_{\Omega_k}(P_{IN,\Omega_k}) = HPF_{\Omega_k}(P_{IN,\Omega_k}) \tag{4.1}$$

Here, Ω is the local neighborhood, LPF is the function of the smoothing filter, and HPF is the function of the detail extraction filter. This operation can be

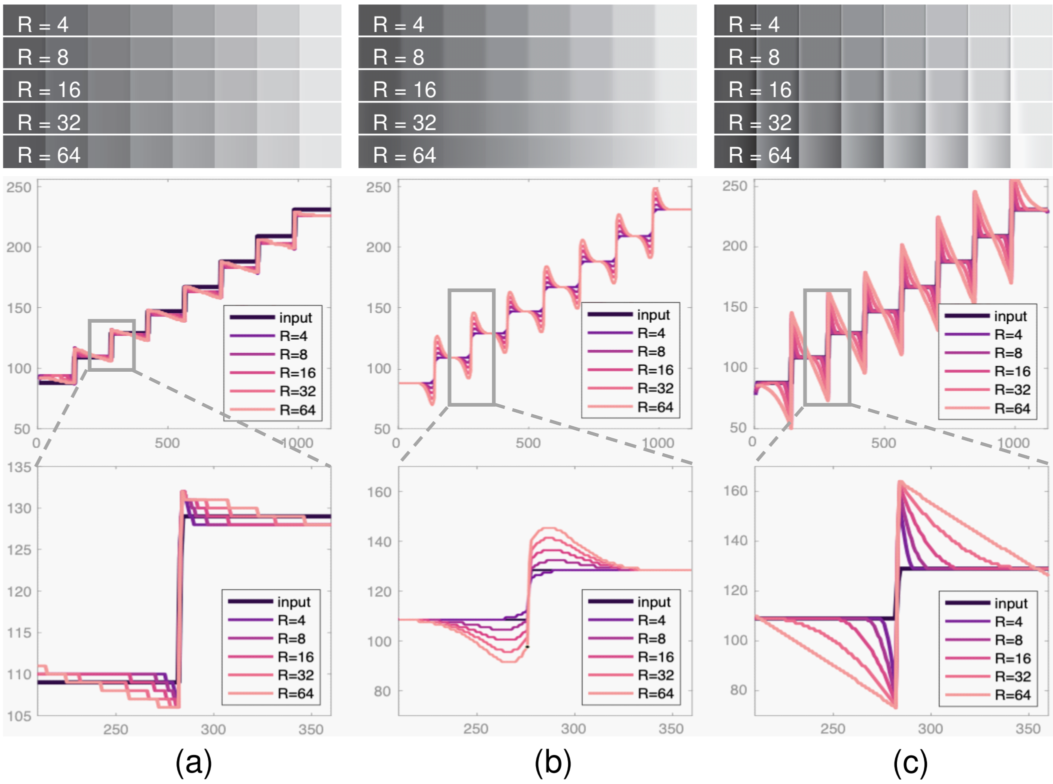


Figure 4.6: Analysis of filter response for varying kernel sizes. (a) SLHE (b) DTF (c) GF. SLHE filter maintains the integrity of input signal across all kernel radius.

expressed as a local tone-mapping function. Low-pass filtered image $LPF_{\Omega(x)}$ is equal to the local mean LPF_{Mean,Ω_k}, so the tone mapping function $F_{Detail}(x)$ is given by

$$P_{OUT,\Omega_k} = F_{Detail,k}(P_{IN,\Omega_k}) = G \cdot (P_{IN,\Omega_k} - L_{Mean,\Omega_k})$$
$$F_{Detail,i}(L_{Mean,\Omega_i}) = F_{Detail,j}(L_{Mean,\Omega_j}) = 0 = const \quad (4.2)$$

Figure 4.7 shows the characteristics of detail layer extraction in the tone-mapping space. The DC low frequency components, which are equal to the local means, are mapped to zero; the high frequency components are plotted linearly (some filters use sigmoid functions) against the output luminance and are handled as a detail layer. The detail layer is enhanced by amplifying its

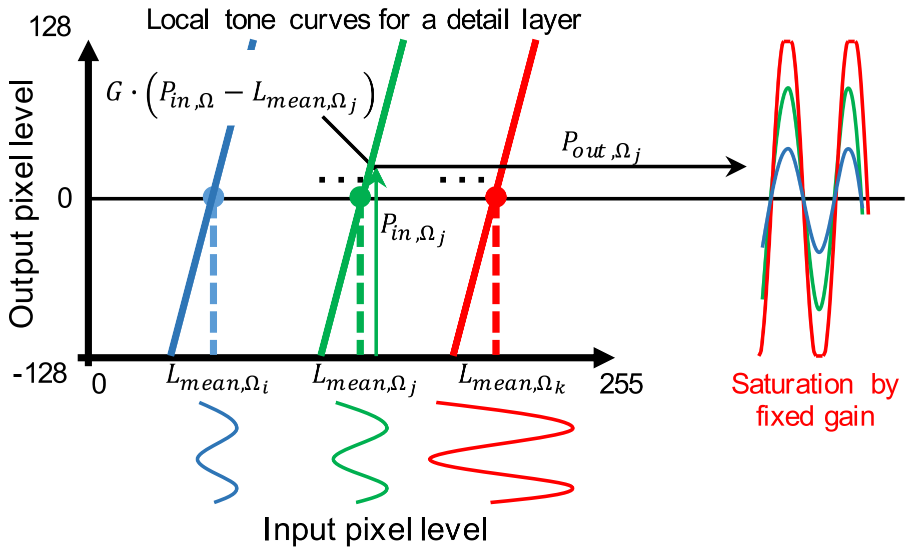


Figure 4.7: Characteristics of detail layer extraction in tonemap space. Tone mapping functions are expressed as linear function. $L_{Mean,\Omega}$ is always mapped to zero. The high frequency components which are differences between $L_{Mean,\Omega}$ and P_{IN} are output as a detail layer, which get saturated by fixed gain.

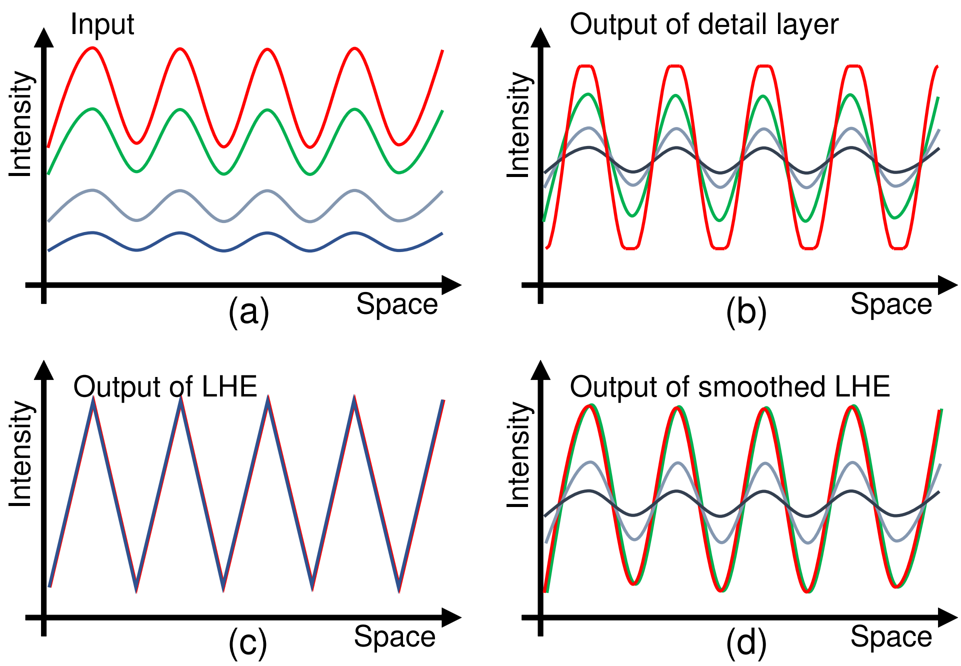


Figure 4.8: Filter output characteristics: (a) Input (b) Multilayer methods (c) Local Histogram Equalization (d) Our smoothed LHE filter

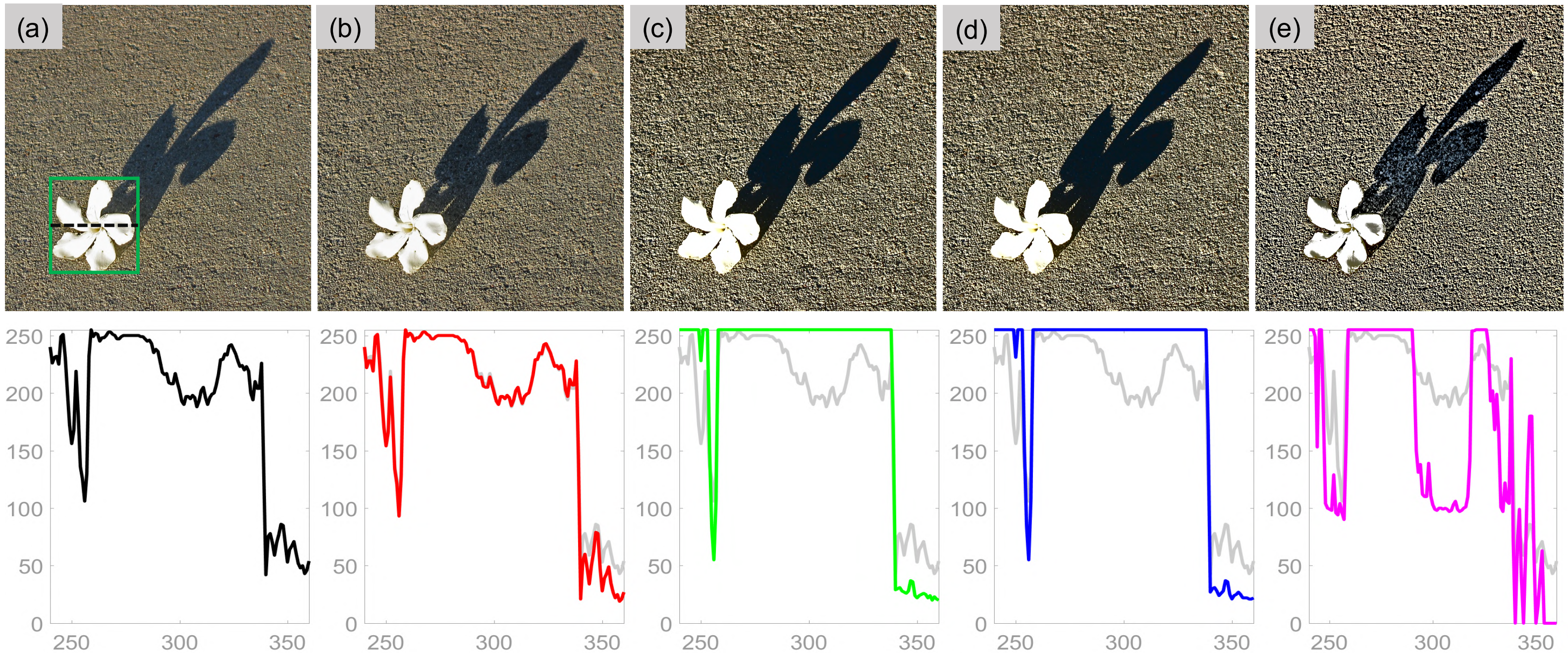


Figure 4.9: Detail manipulation: (a) Input image (b) using our smoothed LHE [SLHE settings (maximum = 100, $0 \leq$ shadow ≤ 50, $25 \leq$ middle ≤ 75, $50 \leq$ highlight ≤ 100)] (c) DTF [IC Filter, $\sigma_s = 20$ and $\sigma_r = 0.08$, manipulate the D_0 using sigmoid function described in [5]] [6] (d) GF [r = 16, $\epsilon = 0.1^2$ and the detail layer was boosted five times.][9] (e) and WLS filter [Combined] [5].

intensity by increasing gain G of the linear functions. The advantage of a ML method is that the linear characteristics of the detail have small amplitudes, however large amplitude causes saturation and uncontrolled over-enhancement as the gain G is set by users. In Fig. 4.8(b) we illustrate the saturation characteristics of ML method. The constant gain used in ML methods lead to disappearance of details as shown in Fig. 4.9. The horizontal scan line drawn from the input image is plotted in gray along with the filtered output by ML methods in green/blue/magenta. In Fig. 4.9(c-e) we can clearly notice how the filter outputs are getting saturated. This results in loss of details due to the uncontrolled over-enhancement.

On the other hand, LHE controls the distribution in the local patch; all textures are converted to full-range triangle waves as illustrated in Fig. 4.8(c). This means that LHE works as active high pass filter. However, since the small amplitude signals are also mapped to full range, uncontrolled over enhancement can occur. SLHE relaxes the characteristics of the LHE which is shown in Fig. 4.8(d) In this work, all local medians are preserved to perform detail enhancement while keeping the low frequency components. From Fig. 4.9(b) we can see that our smoothed LHE method preserves all local details as in original input, while boosting small signals. This is possible, because our method employs an adaptive gain unlike the ML methods which has a constant gain. Furthermore, in our algorithm we set the number of bins to four as in [73]. The bin boundaries of those four bins are defined as shadow, middle and highlight pixels. Ambalathankandy et.al in [121] studied the effect of lateral inhibition on images with illusionary effects. They report that, the edge-preserving filters (like [9], [6]) using their recommended filter size tend to exaggerate details, thus failing to preserve the naturalness in processed images. Therefore, we propose smoothed LHE-like filters for detail enhancement, because of the following disadvantages in multilayer methods:

- They use only a few subjectively selected filters.
- User defined constant gain is applied for emphasizing textures.
- Unmanageable Mach effect even in latest edge-preserving filters.

Here, it is noted that the kernel-size dependency is the same in the filters and LHEs based processing, however the adequate sizes are different in

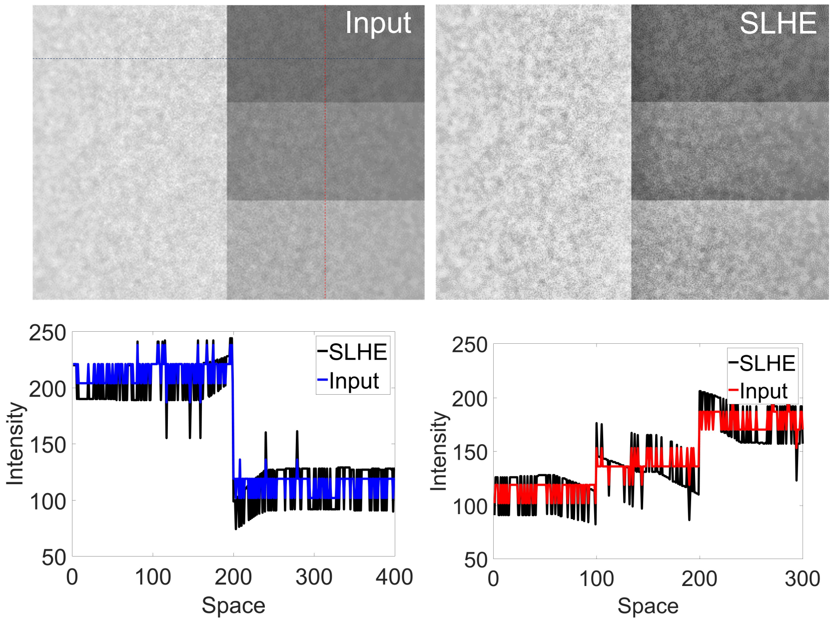
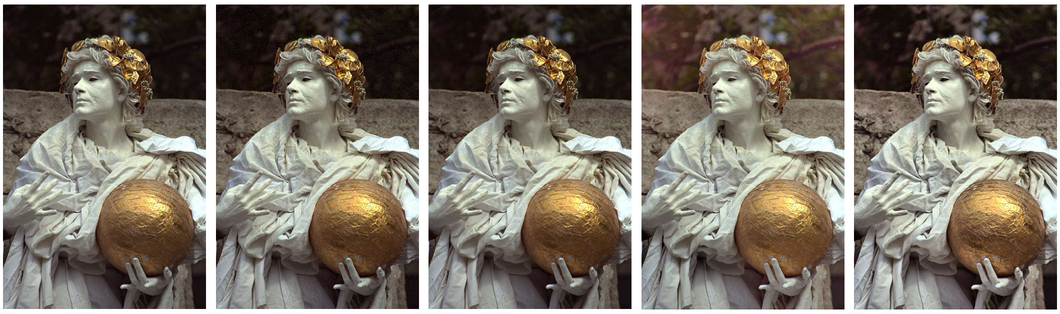


Figure 4.10: (Upper left) Input (Upper right) SLHE Output (Bottom left) Horizontal cross section and (Bottom right) Vertical cross section.

actual processing. In the filter group for multi-layer methods, usually middle or small kernel size is applied; the large size filtering almost averages input images, hence are not useful. In Fig.4.6 we analyzed the filters response for varying kernel sizes. We observe that, with larger radius GF tends to treat edges like textures and leads to saturation, whereas with DTF this response is comparatively lower but does exhibit similar characteristics. These filters also have selectivity for frequency components; detail manipulation makes use of few layers thus limiting to certain frequencies. On the other hand, the LHE group can process the images using large kernel size, because they work like active high pass filtering which amplifies all AC frequency components in the kernel. Thus, we consider that the LHE based processing, which can preserve edge profile including flatness and balance all frequency components, has advantages for natural local tone mapping compared to the multi-layer method that use latest edge preservation filters.



(a) Original (b) DTF (c) GF (d) Apical (e) SLHE

Figure 4.11: Tone mapping for color images.

4.5 Natural tone mapping

In Fig.4.10 we observe wide halo (sawtooth) in a SLHE processed image. However, these enhance the contrast and are not perceived as unnatural. Smoothed LHE methods have been used as tone mapping functions [110], however they are prone to halo artifacts. And a halo reduction method using weighted local histogram is also proposed in [110]. In Fig.4.11 we compare median preserving SLHE tone mapped images with other LHE(Iridix) and multi-layer methods. Multilayer processed images appear weaker. Primary limitation of these techniques lies in the fact that they only use few layers. This approach limits it to specific frequency components, and only those frequencies are enhanced to obtain output images. The noticeable degradation in Iridix processed images are due to its inability in adjusting the DC offsets.

4.5.1 Background manipulation

Our smoothed LHE algorithm operating on red and blue weighted luminance channel (refer section 3.3.3) generates tone mapped images with good quality as shown in Fig . 4.12(b) & (c). Using example shown in Fig .4.12(c) we demonstrate the background manipulation potential of our decolorization algorithm (TM-1: boost and TM-2: suppress). Such manipulations are not possible using CIE Lab or YUV color space.

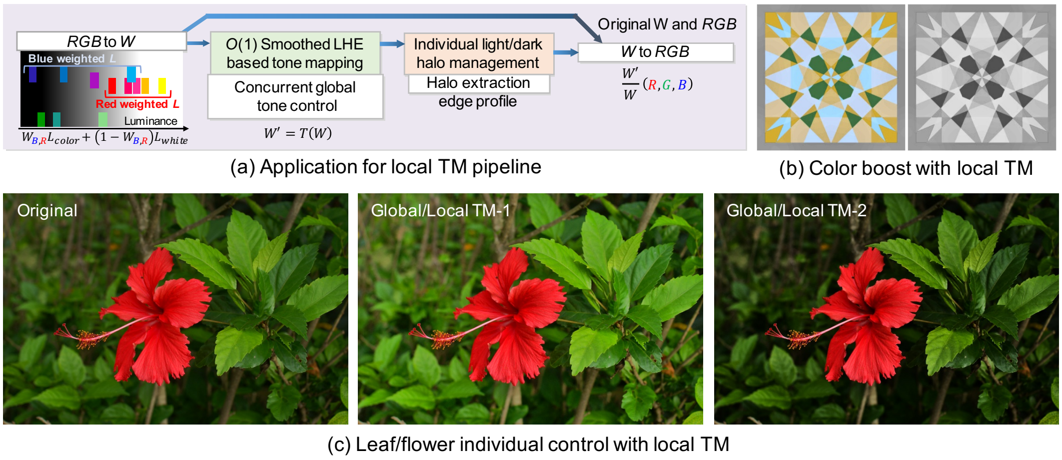


Figure 4.12: Application: (a) Ideal image processing pipeline (TMO example) (b) Color boost using local tonemap (c) Individual color control using our algorithm.

4.6 Subjective user study

We conducted subjective test, to evaluate the naturalness of detail manipulated images. Our objective was to compare the outputs obtained from edge preserving filters versus those processed using our LHE like filter. The edge processed images were obtained using the MATLAB code made available by the authors on their project websites. All the filter settings (kernel size, number of layers) were set to the recommended values (see Fig.4.13). In this evaluation test, we did not consider Apical Iridix as it has DC control issues. The test group consisted of 20 user study volunteers, who were shown five sets of images, which are shown in Fig.4.13. The user study group volunteers were asked to assign points to these images, on a scale of 1 to 5, where 1 corresponds to very unnatural feeling and 5 to good naturalness. They were asked the following questions in this survey:

- Appealing (looks good)
- Appropriate contrast
- Recognize noise
- Is perceived natural

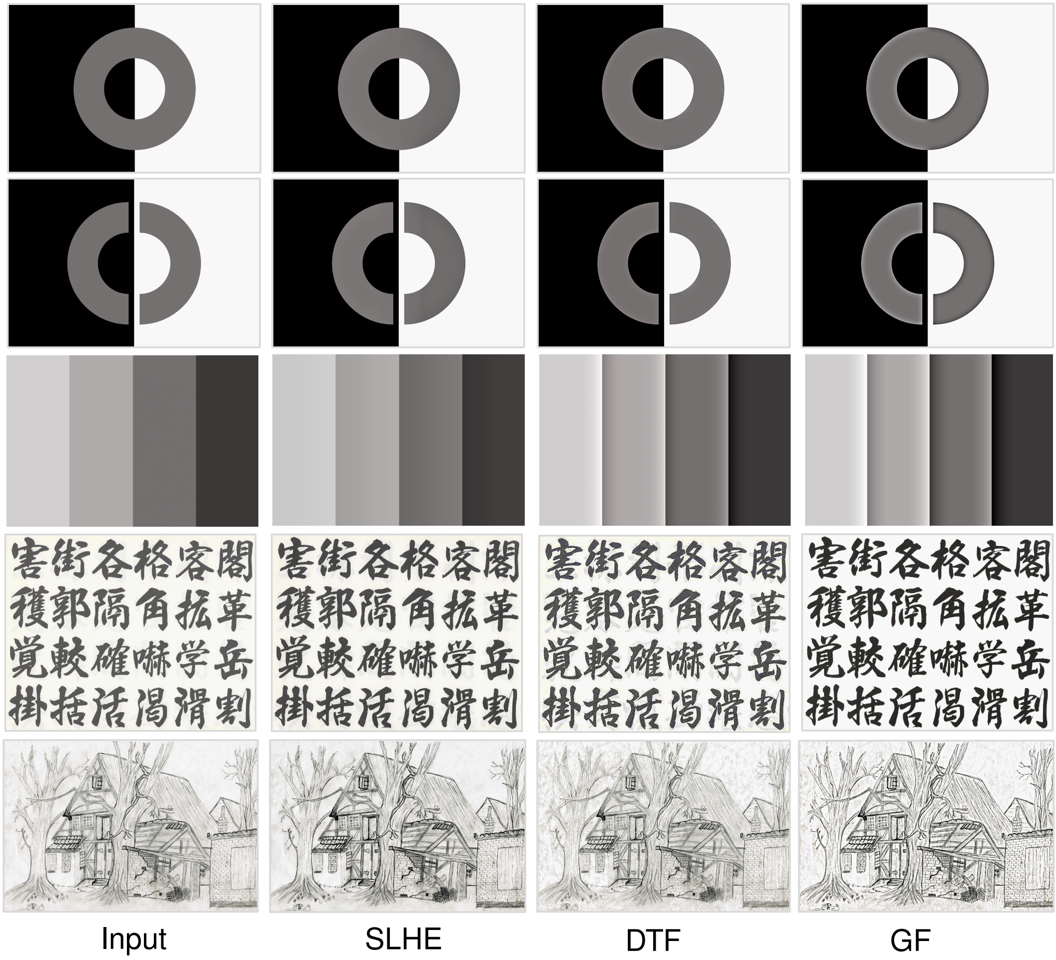


Figure 4.13: Images used for our visual perception test. SLHE settings (maximum = 100, $0 \leq$ shadow ≤ 50, $25 \leq$ middle ≤ 75, $50 \leq$ highlight ≤ 100). DTF settings (IC Filter, $\sigma_s = 20$ and $\sigma_r = 0.08$, manipulate the D_0 using sigmoid function described in [5]). GF settings (r = 16, $\epsilon = 0.1^2$ and the detail layer was boosted five times.)

In our test setup we used Apple MacBooks set to its native resolutions and the lighting of the test room was slightly dim. The compiled response of the volunteers is presented in Fig.4.14. From the subjective test, we observe that, volunteers consistently chose smoothed LHE generated images as natural. Therefore, we state that LHE methods are more suitable for processing images with optical illusion and resulting images have more naturalness to them.

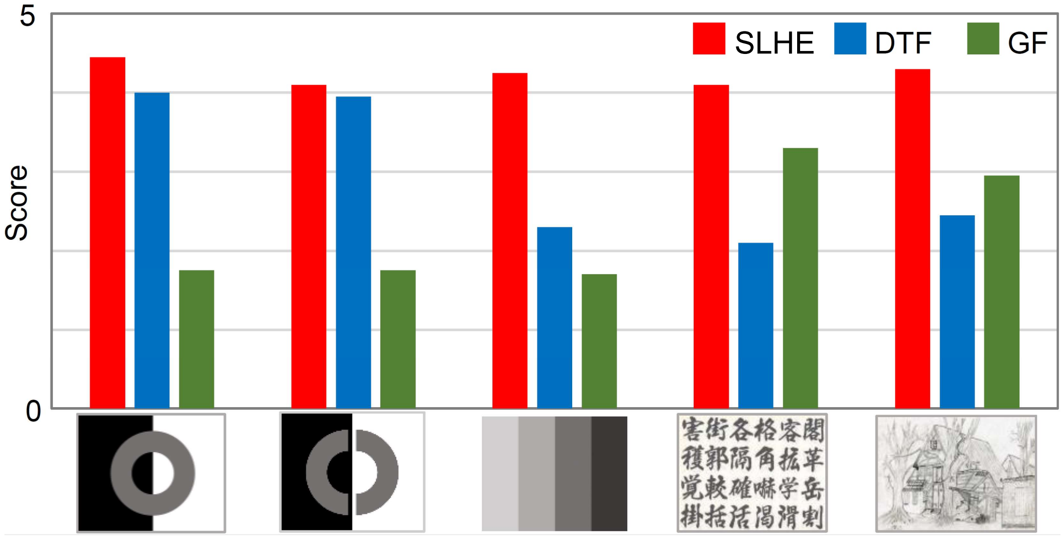


Figure 4.14: Summary of our visual perception test results.

4.7 Conclusion

In this chapter we have analyzed the usefulness of smoothed LHE like filtering methods for image processing. Our observations are as follows. Illusionary images when processed with spatial filters using their recommended kernel size tend to be over enhanced because of the stacking LI effect. This creates unnatural feeling for the observer. On the other hand processing with smoothed LHEs with large kernel sizes, produces illusionary images that retain good edge forms and display a natural-like feeling, as, the LHE techniques work like active high pass filtering which amplifies all AC frequency components in the large kernel.

Chapter 5

Applications

5.1 Introduction

According to Pan American Health Organization estimations about 3.6 billion diagnostic X-ray examinations are performed every year in the world [122]. This shows that radiographic imaging is a decisive diagnostic tool for physicians. X-ray medical imaging is a powerful tool for diagnosis, nevertheless we should remember that the ionizing radiation has long been known to increase the risk of cancer. To give the patients maximum benefit, it is imperative that X-ray images are accurately produced. Correct processing of the images and viewing conditions are essential for accurate diagnosis. Phenomena like Mach band effect can affect correct diagnosis of diseases. The Mach band effect in radiological imaging is an optical illusion, which causes an increased contrast perception between surfaces with different luminance values. Our body is made up of several organ systems, which in turn are made up of complex tissues and they have different radiographic densities. As we know, X-ray imaging is performed by passing an X-ray beam through our body where a part of the beam is absorbed or scattered by internal structures and the rest is transmitted and recorded by the detector. During this process there are chances of capturing various illusions due to overlap shadows and varying background densities. These illusions may result in misinterpretation of important pathology. In medical imaging, contrast enhancement has become a routine task, where it is effectively used to enhance features of regions of interest for characterization and examination [123].

5.2 Contrast enhancement of medical images

Medical image enhancement is one of the most critical application for image processing algorithms because of their great importance and sensitivity in diagnosing various medical conditions of the patients. The quality and accuracy of these images directly affects the quality of life of a patient. To improve the visual aspects of such medical images, several image enhancement algorithms are used, which are also ideally one of the first steps in medical image analysis pipeline. In this section we will focus on contrast enhancement methods for digital X-rays and demonstrate it as a suitable application using our smoothed LHE algorithm presented in section 3.4.

5.2.1 Contrast enhancement for digital X-rays

Post-processing medical images like X-rays for contrast enhancement is a challenging task. Diagnostic imaging centers follow the fundamental radiation protection principle known as *ALARA* (As Low As Reasonably Achievable) [124]. The principle states that the ionizing radiation dose applied to humans or animals should be as low as reasonably achievable while still providing image quality which is adequate to be able to accurately perform diagnosis. Moreover, low dose X-ray imaging produces images with reduced visibility of internal structures, which are affected by noise (poor signal to noise ratio). When x-rays penetrate human body, they are not homogeneously absorbed; some tissues absorb x-rays more than others. This means tissue type also affects the contrast. We know that noise is signal dependent, therefore noise is also affected by tissue density. Accordingly, a good contrast enhancement tool is required to carefully highlight regions of interest in medical X-rays for accurate diagnosis.

5.2.2 Medical image processing with smoothed LHE

Medical imaging is a visual specialty, and radiologist interpret the images presented to them for diagnostic examination. Therefore, it is required, that these images be presented as accurately as possible. As stated earlier, radiographic imaging follow ALARA principle, and the low contrast images undergo post-processing to reveal more details that are not clearly visible in raw input. Irrera et. al [125] use a non-local mean (NLM) patch-based filter to denoise low dose X-ray images. Following which they perform a multiscale contrast enhancement which achieves a good trade-off between noise reduction and improved visibility. However, the choice of NLM may need to be determined experimentally. Yang et. al in [126] presents a multisensor image fusion method using gradient domain guided filtering. A Gaussian smoothing filter is used to decompose source image into base layer and detail layer. Rajendran et.al proposed a guided filter based medical image enhancement method in [127]. However, radiographic images which are full of rich features created by overlaying shadows and tissues of varying radiographic densities, are prone to various optical illusions produced by Mach effect [128]. These are caused by lateral inhibition [129], and in section 4.3 we illustrated the limitation of ML filters when processing such images.

In Fig. 4.6 we had analyzed ML and SLHE filter response for varying kernel sizes. We observe that, with larger radius ML filters tend to treat edges like textures and leads to saturation. This effect is lower in DTF, but it exhibits similar characteristics. These filters have selectivity for frequency components; ML detail manipulation makes use of few layers thus limiting to certain frequencies. On the other hand, our smoothed-LHE with a continuous kernel size control can process the images using varying kernel sizes. This enables SLHE to be perceptive of the neighborhood and boost all AC frequency components within the kernel. This feature enables the radiologist to examine various regions of interest with ease in real-time. Therefore, SLHE-based filters are practical contrast enhancement tool for radiographic images.

5.2.3 Radiologist Perspective

As radiology is a visual specialty, to demonstrate the usefulness of our enhancement algorithm, in this section we present radiologist's subjective assessment. Fig. Fig. 5.1(a) Lateral view of a cervical spine, and its enhancement using the GF and SLHE are presented in Fig. 5.1(b) & (c) respectively. The final processed image using SLHE keeps all the bony detail in the cervical spine X-ray. Additionally, it enhances the visualization of soft tissue planes. For

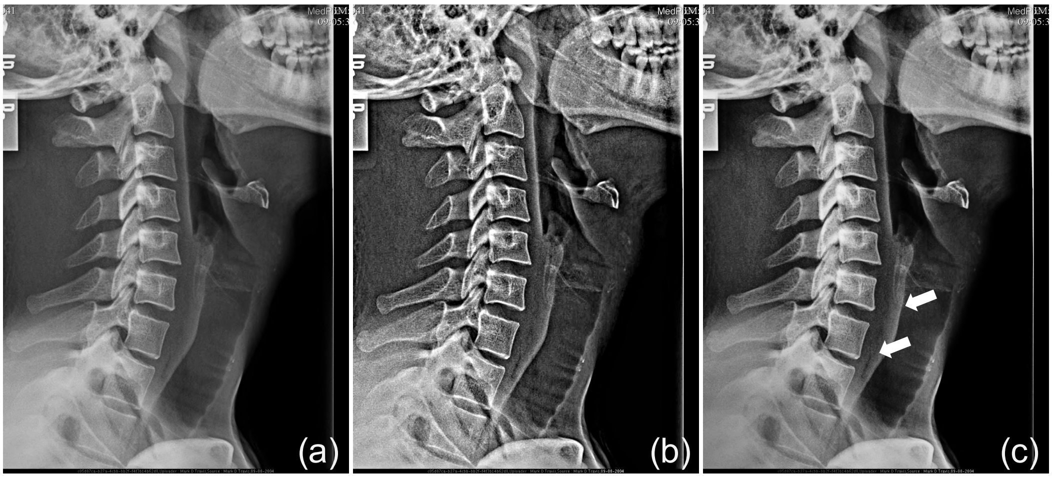


Figure 5.1: X-ray image enhancement: (a) Cervical spine input (b) Enhanced using GF (c) Enhanced using our SLHE. [SLHE settings (maximum = 100, 0 $\leq$ shadow $\leq$ 50, 25 $\leq$ middle $\leq$ 75, 50 $\leq$ highlight $\leq$ 100)]

example: air within the esophageal lumen (see markers in Fig. 5.1(c)) can be clearly seen as compared to the input image. In Fig 5.4, the varying effect of kernel change on Mach band effect and contrast enhancement is demonstrated in a chest X- ray (Fig. 5.2(a)) and an AP view of the pelvis with both hips (Fig. 5.2(b)). These processed images were subjectively graded by a radiologist in diagnostic quality from 1 to 5 (5 being the best interpretive image) as given in table 5.1 case I & II. The qualities taken into consideration includes preservation of normal anatomy while eliminating artifacts thus helping increase diagnostic accuracy.

Chest X-ray with Mach band effect

Chest X-ray of a normal person (image courtesy [1]) demonstrating Mach band effect adjacent to the left heart border as pointed in Fig. 5.2(a), must be differentiated from a pneumomediastinum. The different sets of manipulated images are graded with the best image demonstrating preservation of normal pulmonary vasculature while decreasing the Mach band effect around the heart border. Increased prominence of the pulmonary markings can simulate false positive diagnoses such as pulmonary edema.

[1] RADIOPEDIA (https://bit.ly/2MVXcD9)

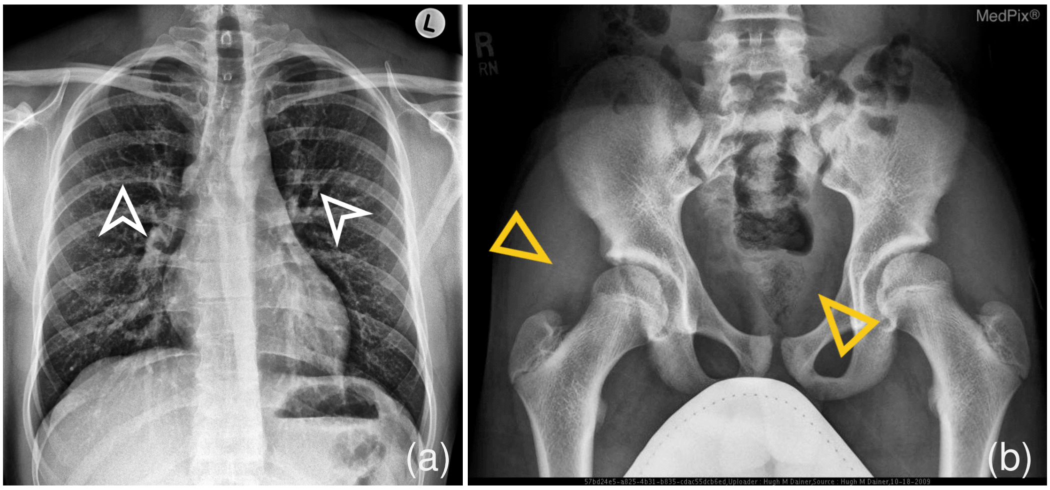


Figure 5.2: (a) Chest X-ray with Mach band effect adjacent to heart marked by pointers. (b) Low AP view of the pelvis with both hip joints.

Table 5.1: Radiologist Subjective Evaluation.

SLHE preset setting		Kernel Radius				
		R = 4	R = 8	R = 16	R = 32	R = 64
I	L	1	2	3	4	5
	M	5	1	2	3	4
	H	5	4	3	2	1
II	L	5	4	3	2	1
	M	4	5	3	2	1
	H	1	2	3	4	5

Range of the tone mapping space (Max = 100). For SLHE configuration L: Shadow = 10 to 40, Middle = 35 to 65 and Highlight = 60 to 90. Configuration M: Shadow = 0 to 50, Middle = 25 to 75 and Highlight = 50 to 100. Configuration H: Shadow = -25 to 75, Middle = 0 to 100 and Highlight = 25 to 125.

Pelvis with bone hip joint X-ray

Low AP (anteroposterior) view of the pelvis with both hip joints of an immature skeleton is shown in Fig. 5.2(b)(image courtesy[2]). Using image manipulation with SLHE filter, the cortical details can be better appreciated while preserving normal anatomy such as soft tissue planes. The subjective grades for this X-ray manipulations are given in Table 5.1 (case II). The preservation of soft tissue planes is key to differentiate normal from pathology on a musculoskeletal film.

[2] MEDPIX (https://bit.ly/2B6buhO)

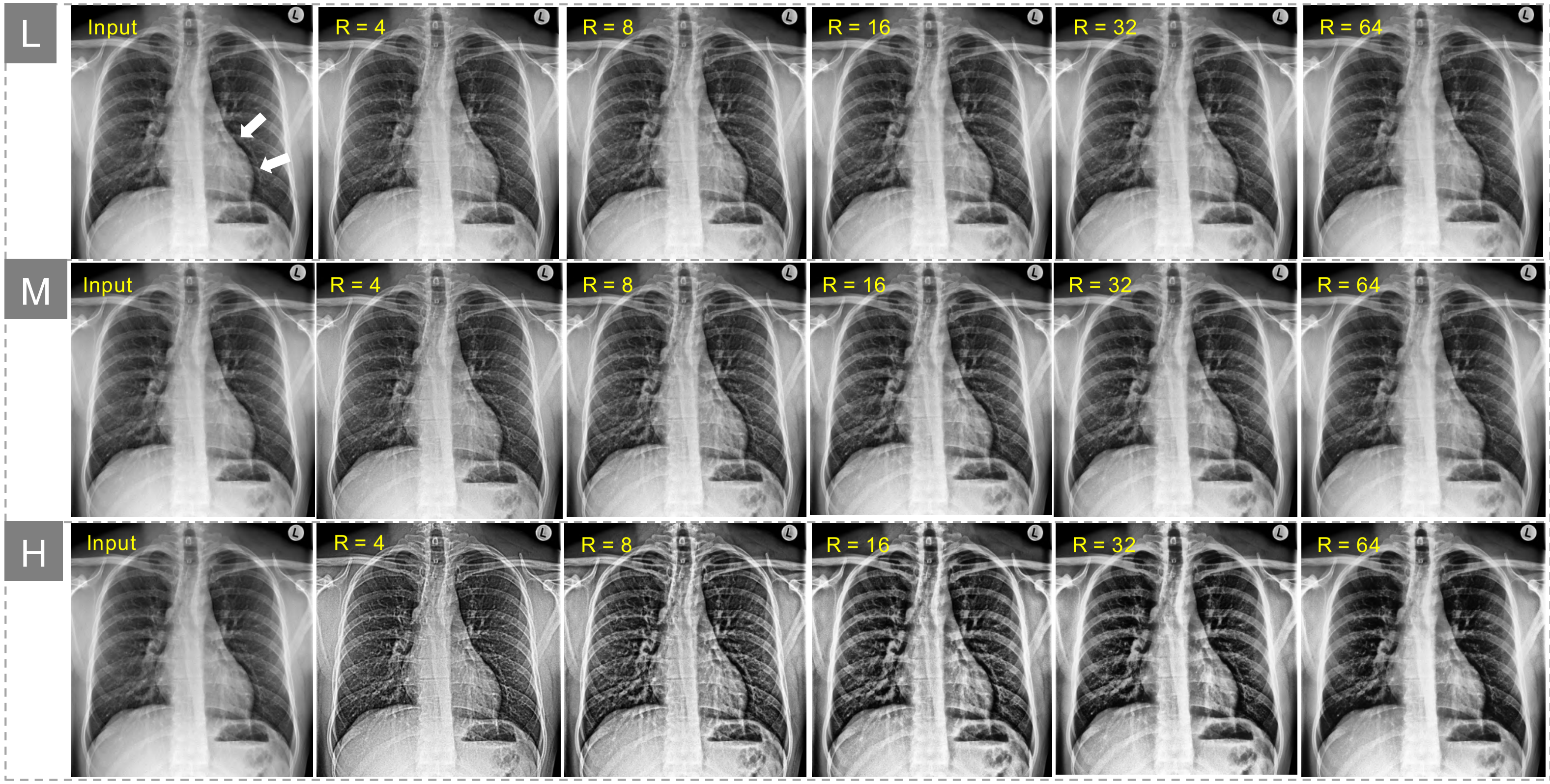


Figure 5.4: Effect of Kernel Size on X-Ray contrast enhancement. Range of the tone mapping space (Max = 100). For SLHE configuration L: Shadow = 10 to 40, Middle = 35 to 65 and Highlight = 60 to 90. Configuration M: Shadow = 0 to 50, Middle = 25 to 75 and Highlight = 50 to 100. Configuration H: Shadow = -25 to 75, Middle = 0 to 100 and Highlight = 25 to 125.

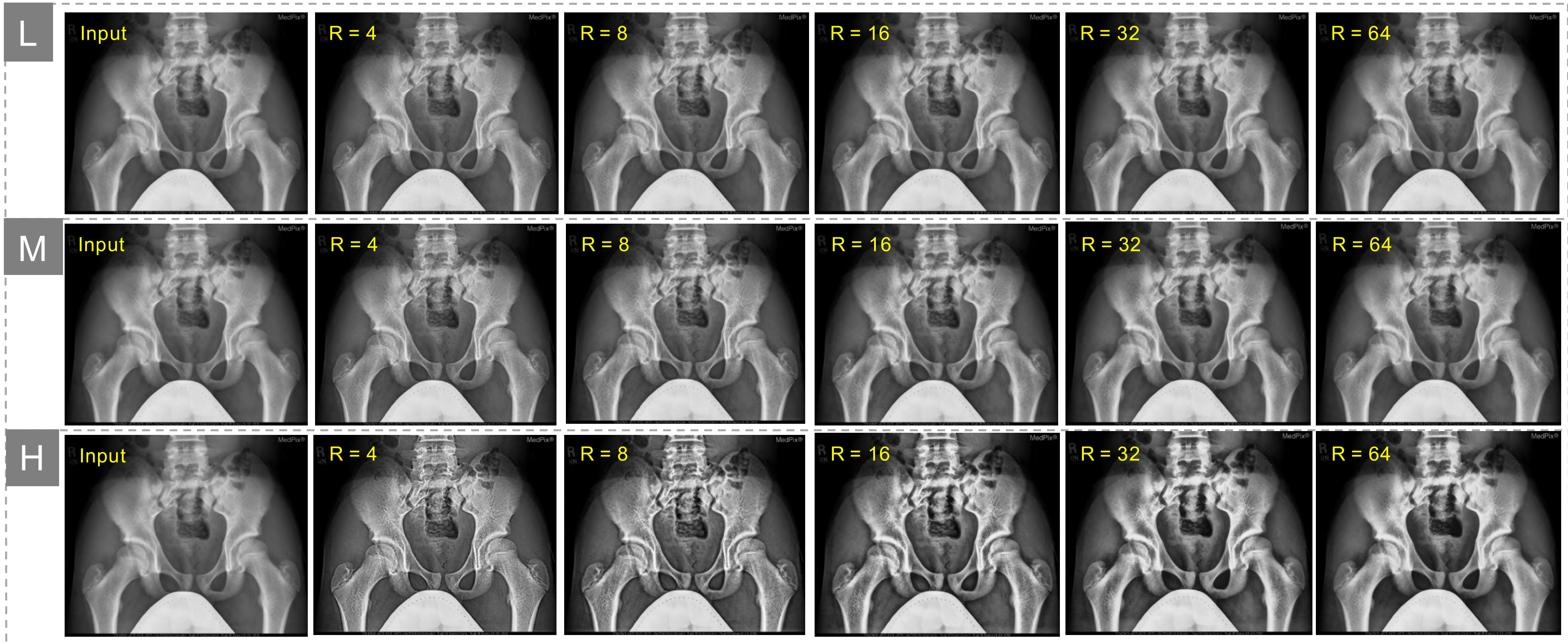


Figure 5.5: Effect of Kernel Size on X-Ray contrast enhancement. Following are the SLHE configuration parameters, range of the tone mapping space (Max = 100). For configuration L: Shadow = 10 to 40, Middle = 35 to 65 and Highlight = 60 to 90. Configuration M: Shadow = 0 to 50, Middle = 25 to 75 and Highlight = 50 to 100. Configuration H: Shadow = -25 to 75, Middle = 0 to 100 and Highlight = 25 to 125.

5.3 Demo video

Our main objective is to realize a real-time tone mapping system for practical applications that can process videos. In this section we present video processing examples from outdoor scenery and indoor scenes. The demo video is published on YouTube for its link YouTube (click here). It is important to establish the level of accuracy of our video processing algorithm and assess the quality of experience measures. For the example videos we have performed objective quality assessment which is presented in the following section.

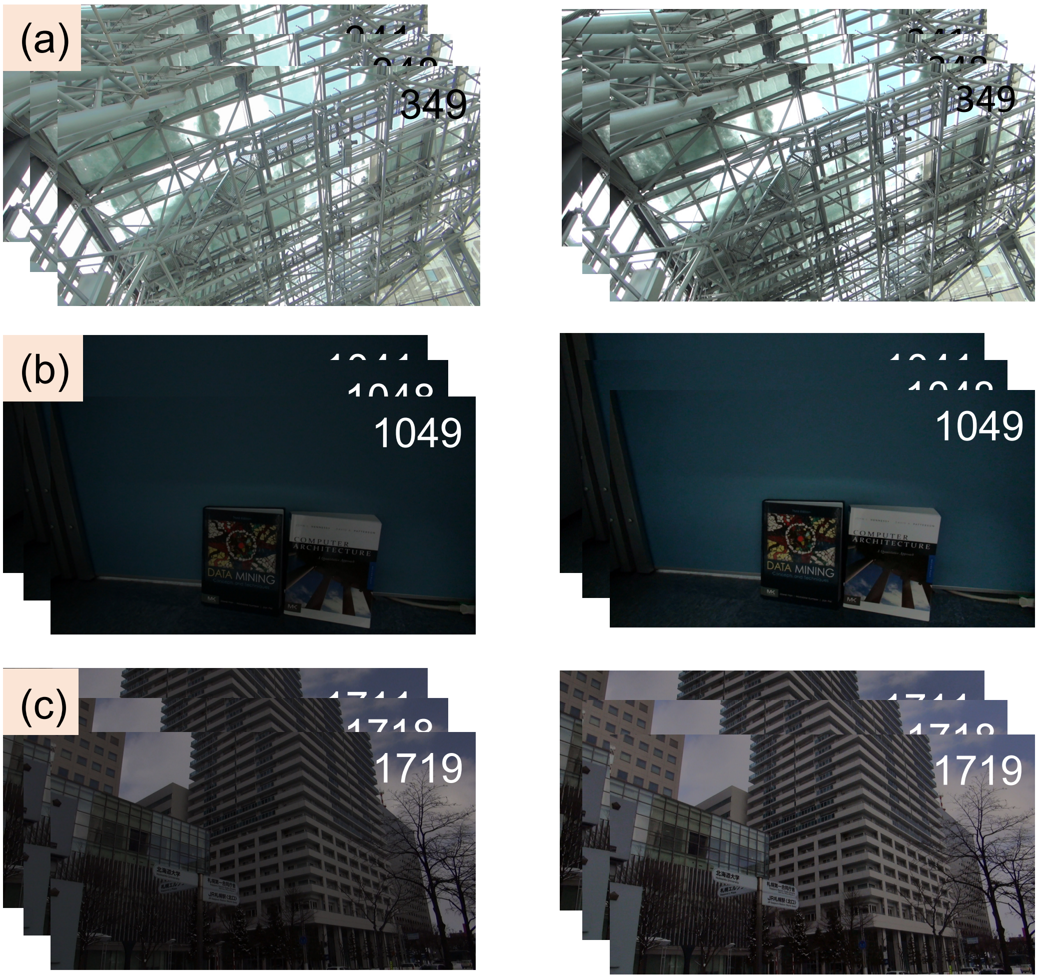


Figure 5.3: Ten frames per sequence were grabbed from our test video for quantitative evaluation: (left) Input video. (right) Tone mapped output video.

5.3.1 Objective video quality assessment

For objective evaluation we chose 10 frames from three different sequences shown in Fig. 5.3. The average SSIM values and TMQI scores for the three video sequences along with its noise variance is presented in table 5.2. These values obtained are similar to our earlier noise variance and SSIM, TMQI results that we had reported in section 3.4.1 and 3.5. This proves the robustness of our hardware implementation while processing real-time full HD videos. In table 5.3 we report contrast enhanced in locally tone mapped output images, when compared to global only images. The results in table 2, were computed as follows: Ten images from a video sequence, and its tone mapped version were captured and, we computed their local edge measure as an absolute value of $\sum Y_{edge} + \sum X_{edge}$ using sobel like edge detector [130]. We chose sobel edge detector as it is less sensitive to noise. From the table 5.3 we can easily notice that for the additional hardware that our implementation requires, it produces images with more contrast.

Table 5.2: Noise variance, SSIM, and TMQI experimental measurements of tone mapped video output.

Sequence	Noise variance	SSIM	TMQI
(a)	1.3206	0.9650	0.9509
(b)	0.7398	0.9872	0.9205
(c)	1.4450	0.9810	0.9688

Table 5.3: Contrast assessment of TMO video processing

	Edge Measure	
Sequence	Global TMO	Global & Local TMO
(a)	2467	70424
(b)	704	5615
(c)	1959	29634

5.4 Conclusion

In this chapter we have analyzed and compared the multilayer methods and SLHE filters for processing low contrast images like digital X-rays. These images are also prone to optical illusions like Mach bands and background contrast effects. These illusions are caused due to lateral inhibition phenomena. In such cases spatial filters used in multilayer methods are observed to cause over-enhancement leading to unnaturalness in output images due to their fixed gains and selective filter kernels. Our proposed $O(1)$ smoothed LHE filter is more robust when processing such images and maintains the naturalness in final output images, which is an essential feature for medical diagnostic application. Our LHE-like filter is easy to control over continuous kernel size and operates like a high pass filter for enhancing detail contrast in real-time. We also presented demo video to demonstrate the real-time performance of our proposed algorithm along with objective video quality metrics. Currently machine learning methods have become an important tool to solve many computer vision tasks. As an extension of this work, we would like to leverage its potential by combining our SLHE algorithm with deep learning networks to develop accurate and fast contrast enhancement tool by learning various low contrast features of digital X-ray images. However, it will be a challenge to obtain good tagged training data and to be able predict the output quality of the network.

Chapter 6

Methodology and design of hardware accelerator for SLHE algorithm

6.1 Introduction

Tone mapping HDR images/videos are a computationally challenging problem, and many hardware-based dedicated accelerators have been proposed for this purpose. In this section we will present a FPGA-based hardware accelerator for our proposed SLHE algorithm. As a preliminary to our hardware design, we will briefly review earlier tonemap operator implementations on GPU, FPGA and ASIC from the literature which are listed in Tables 6.1, 6.2 and 6.3 respectively. Later, in this section we will discuss in detail the hardware design of our tone mapping algorithm and a demonstration video showing its real-time characteristics. To summarize the main contributions of this chapter are as follows:

- Overview and briefly review hardware implemented tone mapping operators.
- Design and development methodologies for SLHE algorithm implementation on a Xilinx Kintex-7 FPGA.
- Comparison with other related implementations.

6.2 Related work

In literature, there have been many state of the art survey reports like [131, 132, 133, 134, 135, 136, 137]. These surveys mainly covered the software algorithms. The choice of a TMO operator is usually application driven. In this report we study algorithms that have been optimized for hardware platforms like Application Specific Integrated Circuits (ASIC), Field Programmable Gate Arrays (FPGA) and dedicate Graphic Processing Unit (GPU). Strong demand for real-time embedded vision-based applications is on the rise in various domains like advanced automotive systems, medical imaging, robotics and Unmanned Aerial Vehicles (UAVs). The main building blocks for such vision-based systems are the image sensors, image processing algorithms and display monitors. For real-time applications with time constraints a hardware acceleration is necessary [138]. Also, embedded applications are energy and resource constrained there by simply porting software algorithms on a hardware platform may result in poor performance or even system failure. Therefore, image processing algorithms have to be optimized for hardware porting [139]. This redesign effort, which can exploit the hardware platform for optimal performance has produced many novel hardware tone mapping algorithms and architectures. In this section, we review such hardware tone mapping algorithms and present the following:

1. A comprehensive introduction to hardware tone mapping operators.
2. A brief survey of tone mapping operators that have been implemented on an ASIC/FPGA and GPU platform.
3. Comparison of tone mapping operators based on their hardware specification and performance.

6.3 Design methodologies for SLHE algorithm implementation on hardware platforms

Real-time image processing applications are bound by timing constraints, i.e., for every pixel clock the processing pipeline should consume one pixel from the camera side and deliver one to the display side. Any missed pixel on either side would lead to loss of information or cause blanking display, which is known as the throughput constraint. When porting software (SW) algorithms to hardware, an inherent design problem is that the SW code is developed on a general-purpose CPU. Therefore, the algorithm is highly sequential, and it is useful to exploit the fast CPU. However, this is not the case on HW platform. For example, FPGAs are clocked at much lower frequencies and designers should exploit this parallelism to implement real-time systems. Another type of constraint that has to be met for real-time tone mapping system, is the pipeline latency. Here, latency implies how many clock cycles are required to process one input pixel to processed/tone-mapped pixel.

Memory bottleneck is crucial for implementing image processing algorithms on hardware. While HW platforms like FPGAs have highly parallel logic blocks and fast re-configurable interconnects to speed up window (kernel) operations, the interface speed between the tonemap accelerator and the rest of the FPGA system is a bottleneck; particularly for image processing applications whose data bandwidth requirements are extremely high volume. The cost of moving data between off-chip memory and the accelerator can be detrimental and outweigh the benefits of implementing the FPGA system. Therefore, well thought out operation sequence that obeys raster order should be chosen, because other computation order would usually require the whole frame to be buffered. Like caching can reduce the memory bottlenecks on

Table 6.1: List of tone mapping algorithms implemented on GPU platform

	Hardware		**Performance**		
	GPU	Technology (nm)	Frame Size (pixel)	Speed (FPS)	Throughput (Mpix/s)
[140]	Radeon 9800 Pro	150	512×512	5	1.3
[141]	GeForce 6800GT	130	1024×768	10	7.9
[142]	GeForce Go 6800	130	2048×2048	7	29.4
[143]	GeForce 8800 GTS	90	-	-	-
[144]	GeForce 7900 GTX	90	640×480	30	9.2
[145]	GeForce GT 550M×2	40	1024×768	2.8	2.2
[146]	GeForce 8800 GTX	90	1002×666	37	24.7
[147]	NVIDIA ION2	40	640×480	27	8.3
[148]	GeForce GTX 980	28	1980×1080	46.5	99.4
[149]	GeForce Titan Black	28	2048×1536	24	75.5
[150]	GeForce GTX 650 Ti	28	4096×4096	7.5	125.8

Table 6.2: List of tone mapping algorithms implemented on FPGAs

	Hardware				**Cost**							**Performance**			
	Camera	Model	Technology (nm)	Clock (MHz)	Latency (clock)	Power (mW)	Memory (bit)	Logic Elements	DSP	Registers	Others	Frame Size (pixel)	Speed (FPS)	Throughput (Mpix/s)	pix/clock
[151]	No	Stratix II	90	77.15	64	-	3,153,408	34,806	54	-	-	1024 × 768	60	47.2	0.61
[72]	No	Virtex II	150	1.3	-	-	20^{γ}	17,280	-	890^{δ}	F/F=1362^{η} Multiplier=16	125 × 86	24	0.258	0.20
[152]	No	Stratix II GX	90	66.66	$2 \times W \times H$ +300	-	2,609,151	$49,763^{\alpha}$	-	43,793	-	2.5M	25	50	0.75
[153]	No	Virtex 5	60	85	3,293,244 (Total)	-	3,932,160	$5,806^{\beta}$	18	-	-	1024 × 768	25	19.7	0.23
[59]	EV76 C560	Virtex 5	60	94.733	65 clock /line	-	40^{γ}	$14,168^{\beta}$	4	8,132	-	1280 × 1024	30	39.3	0.41
[154]	No	Spartan 6	45	75	12288 @163μs	-	3^{γ}	874^{β}	17	1000^{ϵ}	-	1920 × 1080	30	62.2	0.83
[144]	No	Spartan 3	90	40.25	-	900	39^{γ}	$30,086^{\beta}$	26	$16,545^{\epsilon}$	BufgMux=7 DCM=2	640 × 480	60	18.4	0.46
[155]	EyeTap	Spartan 6	45	78.125	-	-	100^{γ}	-	-	-	-	-	30	-	
[156]	EV76 C560	Virtex 6	40	125	136 clocks @1.2us	6	17^{γ}	$16,880^{\beta}$	-	20,192	-	1280 × 1024	60	78.6	0.63
[43]	No	Stratix II	90	114.9	nCols+12+7	-	68,046	$8,546^{\alpha}$	60	10,442	-	1024 × 768	126	99	0.86
		Cyclone III	65	116.5	nCols+12+7	250	86,046	$12,154^{\alpha}$	36	10,518	-				
[157]	No	Stratix II	90	114.18	-	-	307,200	9,019	88	6,586	-	1M	100	100	0.88
[158]	No	Spartan 3	90	40	-	897	720,000	$29,007^{\beta}$	26	11594	Slice=16545	1M	100	100	0.88
[60]	No	Virtex 5	65	214.27	-	-	8^{γ}	$4,536^{\beta}$	30	5,036	-	1920 × 1080	60	124.4	0.58
[159]	No	Cyclone IV	60	55.55	2Rows+18	-	-	1,784	-	-	Linebuff=4 F/F=510^{η}	2560 × 2048	-	-	-
[160]	No	Cyclone II	90	170.24	-	-	98,594	838	10	680	CF=558^{ζ}	1920 × 1080	80	165.9	0.97
		Stratix III	65	288.77	-	-	93,982	80	10	1,153	CF=418^{ζ}	1920 × 1080	140	290.3	1.01
[161]	EV76 C560	Virtex 5	65	114	136 clocks @1.2us	-	30^{γ}	6730^{β}	-	-	F/F=$6,378^{\eta}$	1280 × 1024	60	78.6	0.69
[162]	No	Cyclone III	65	100	83	674.25	87,176	93,989	28	26,004	CF=$67,985^{\zeta}$	1024 × 768	126	99	0.99
[163]	No	Cyclone III	65	100	2088	149.5	31,000	4,20	-	3,031	-	1024 × 768	126	99	0.99
[164]	No	Stratix II	90	114	-	-	614,440	9,019	-	-	-	1280 × 720	123	113.3	0.99
[165]	$-^{\theta}$	Virtex 5×8	65	125	-	31.72W	3^{γ}	$4,764^{\beta}$	54	2,489	-	1024 × 256 × 8	25	52.4	0.42
[166]	Python 2000	Zynq XC7Z020	28	200	-	-	2,160,000	29,850	-	-	-	1920 × 1080	96	199	1.00
[167]	Flare 2KSDI	Zynq XC7Z045	28	-	-	12W	47^{γ}	$24,700^{\beta}$	-	30,405	-	1920 × 1080	30	62	-
[2]	Yes	-	-	400	-	-	-	-	-	-	-	$2M$	33	66	0.17
[168]	Python 2000	Zynq XC7Z020	28	200	-	8W$^{\kappa}$	34^{γ}	$14,706^{\beta}$	38	20,316	-	1920 × 1080	96	199	1.00
[169]	No	Cyclone III	65	100	-	-	77,408	13,216	-	-	-	1024 × 768	126	99	0.99
[44]	Sensor	Cyclone III	65	100	-	-	107,408	15,471	-	-	-	1024 × 768	126	99	0.99
[170]	No	Virtex 7	28	150	-	819	-	-	-	-	-	1024 × 768	189	148.6	0.99
[73]	CX590	Kintex-7	28	162	-	453	9,738,000	$9,799^{\beta}$	21	15,345	-	1920 × 1080	60	124	0.77
		Virtex 5	65	148	-	804	7,488,000	$10,903^{\beta}$	22	12,794	-	1920 × 1080	48	100	0.84
[171]	No	Zynq 7000	28	148.5	@0.241ms	-	2,052,000	$14,888^{\beta}$	-	21,627	-	1920 × 1080	60	124	0.84

$^{\alpha}$Adaptive lookup table(ALUT) $^{\beta}$Lookup table(LUT) $^{\gamma}$Block RAM(BRAM) $^{\delta}$Logic cell $^{\epsilon}$Slice
$^{\zeta}$Combinatorial function(CF) $^{\eta}$Flip-flop(F/F) $^{\theta}$Cell-phone cameras (×16) $^{\kappa}$Include camera

Table 6.3: List of full custom designed tone mapping algorithms on ASIC

	Hardware					Cost			Performance			
	CMOS Sensor	Foundry	Technology (μm)	Monolithic	Clock (MHz)	Area(mm^2)	Power (mW)	Gate Counts	Frame Size (pixel)	Speed (FPS)	Throughput (Mpix/s)	pix/ clock
[172]	No	TSMC	0.13	No	100	2.85×2.85(Core)/ 3.74×3.74(bond)	177.1478	769,620	1024×768	60	47.2	0.47
[173]	Yes	-	-	No	-	1.98×3.88(Chip)	-	-	64×64	-	-	-
[174]	No	AMS	0.35	No	-	-	-	-	256×256	-	-	-
[159]	No	TSMC	0.13	No	200	1.835×1.835 (Core)	-	15,742	2560×2048	37	194	0.97
[175]	Yes (151dB)	AMS	0.35	Yes	-	7.33×6.78(Core)	111.2	-	180×148	1205	32.1	-
[176]	Yes (120dB)	Lfoundry	0.15	in-pixel	-	-	-	-	-	-	-	-
[177]	No	AMS	0.35	in-pixel	-	-	-	-	-	-	-	-
[178]	Yes (102dB)	-	-	No	-	-	-	-	-	-	-	-
[179]	Yes	-	0.15	in-pixel	52	-	54.72	-	1024×768	66	51.9	1
[180]	No	TSMC	0.35	Yes	-	0.039	41	-	-	-	-	-
[181]	Yes (CIS)	-	-	Yes	-	-	-	-	512×512	-	-	-

CPUs, streaming FIFO interfaces can reduce the amount of pixel accesses on FPGA hardware. Also, FPGAs are provided with BRAMs which can be read and written at the same time at every clock cycle, allowing one stream of values to be stored and one stream to be extracted in parallel.

Various technologies are available for implementing image/video processing algorithms. However, the main concerns of design implementation are cost, speed and power. The design methodology adopted for any hardware implementation depends on the application and time to market. The hardware-based implementation can be realized using any one of these hardware platforms: Application-Specific Integrated Circuits (ASIC), Field-Programmable

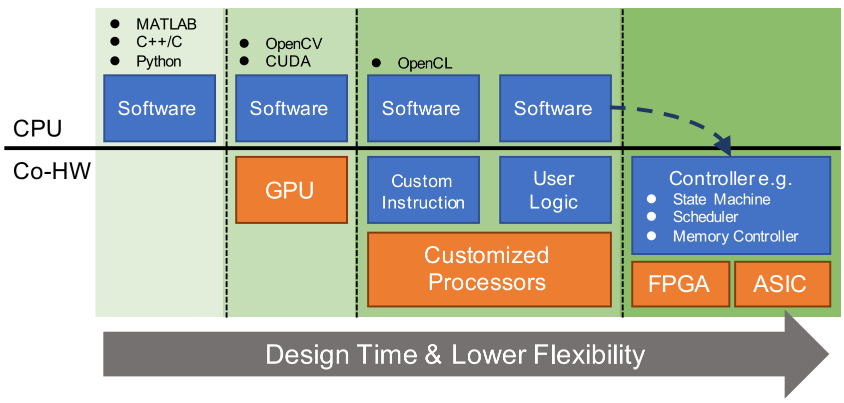


Figure 6.1: Choice of hardware platform for developing an algorithm mainly depends upon the application. Other important factors are design time and engineering costs.

Gate Arrays (FPGA), and Graphics Processing Units (GPU). Each have its own advantages and disadvantages, which we will briefly explain using Fig. 6.1. From this figure, we can notice that the choice of platform depends on various factors like: flexibility, design time and cost. A full custom ASIC design development will be very expensive due to increased manufacturing and design time, and increased non-recurring engineering costs. Even though the ASIC design solution can be very efficient in terms of area and performance, it is only viable for proven designs that can be mass produced. GPUs and FPGA platforms have been preferred for many image processing applications as shown in tables 6.1 and 6.2. OpenCL-based is suitable for implementing algorithms on general-purpose computing on graphics processing units (GPGPU), and it has flavor similar to the proprietary CUDA language from NVIDIA. Recently, FPGA vendors are also supporting openCL development processes.

6.3.1 Overall view of proposed TMO system

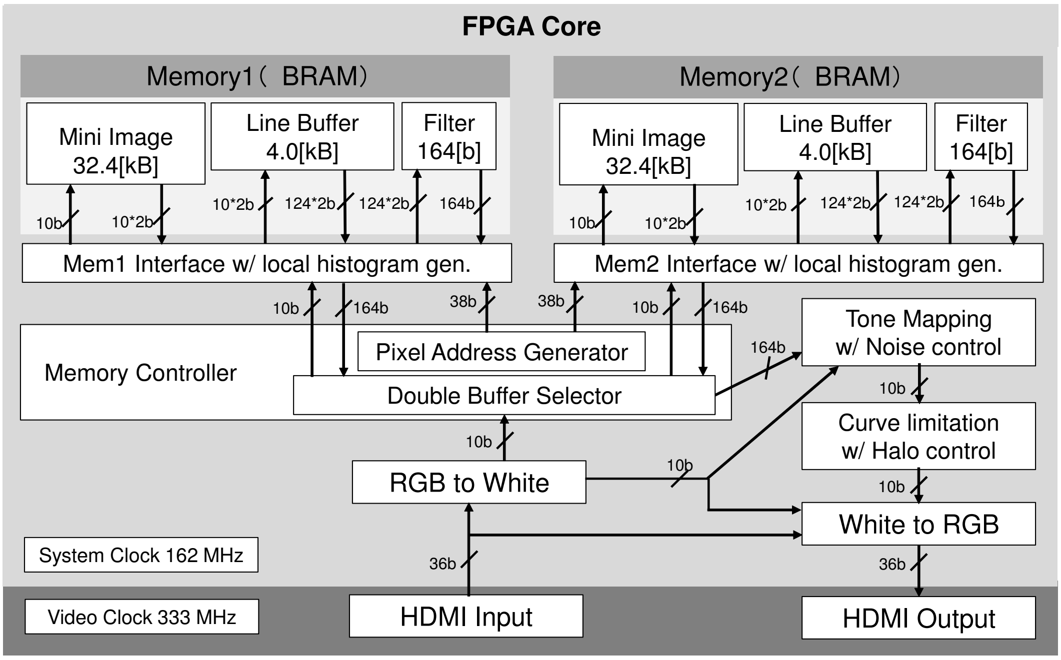


Figure 6.2: Overall view of proposed TMO system.

We designed and implemented a real-time tone mapping system on FPGA. In this section we will explain the constituent modules that we designed to meet the critical timing constraints of processing full HD images, which is of the order of one pixel every $7ns$. The TMO was designed using integer arithmetic and coded in Verilog HDL. Figure 6.2 shows our TMO system implementing the proposed algorithm. Producing one pixel every $7ns$, requires an efficient pipeline of various pixel processing modules which are well-timed. Fig. 6.3 shows our accurately timed pipeline, with each module's latencies marked. Our TMO system consists of six main blocks:

- HDMI I/O circuits
- Color converter $\times$ 2 (RGB to White & White to RGB)
- Memory controller with pixel address generator.
- Compressed double frame buffers
  (Mini image: 32.4 kB, Line buffer: 6.4 kB, Filter: 84 B)
- Tone mapping function generator

Operation of memory controller is critical, as it generates pixel addresses and controls compressed double frame buffer's dataflow. Its control and operation are discussed in 6.3.3. Other modules are controlled using exact delay control and simple wired logic. The input image is obtained from a camera (Sony HDR-CX590) that was connected to the FPGA board through HDMI-compatible FMC daughter cards (Inrevium FMCH-HDMI2-RX). This camera has automatic flicker correction system; therefore our algorithm has no flicker correction module. Since the HDMI and TMO core systems operate at different clock rates (333 and 162 MHz respectively), the input pixels are temporarily buffered and converted to a special white color space, which has the advantage of avoiding color shift. An input RGB color image can be converted to this color space using the below equation:

$$White = \sqrt{\frac{R^2 + G^2 + B^2}{3}} \tag{6.1}$$

This white channel is stored in double frame buffers alternately. The tone mapping function generator reads the down-sampled pixel data to compute the local histogram and the mapping functions. Finally, the tone mapped

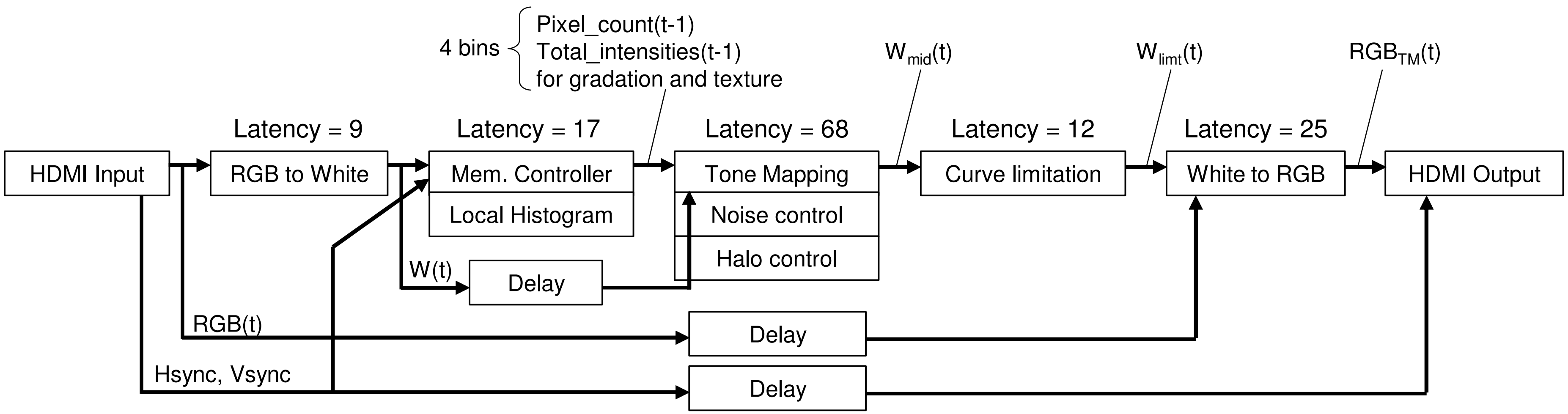


Figure 6.3: System pipeline with module latencies.

pixels are converted to RGB color space buffered and synchronized before they are displayed on the monitor using the FMCH-HDMI2-TX card. The system diagram of the prototype FPGA implementation is given in Fig 6.9.

6.3.2 Compressed double frame buffers

To estimate a local histogram, it is necessary to wait until the processing unit is ready with the local patch data. Therefore, a delay time which is equal to 2N+1 rows (for 2N+1 size local patch) is required in the processing pipeline. For a Full-HD movie processing, this delay is equivalent to 2N+1 $\times$ 1920 pixels. However, in this study, by using the local histograms from the previous frame, we can avoid this latency. Actually, by storing the previous frame data in the compressed frame buffers, we generate the mapping function from the previous frame data and perform tone mapping on the input pixel (current frame). We make use of a one eighth compressed frame/line buffers in our system. Table 6.4 lists memory size of the required buffers; the 10 bit 1920 $\times$ 1080 frame buffer is now down-sampled to a 10 bit 240 $\times$ 135. The reason we use an 8 $\times$ 8 down-sampled image is because the PSNR is at acceptable 40 dB and any further down-sampling leads to lower PSNR, which would mean loss of quality. And by using a downscaled image instead of the actual image saves us 98% memory required by the frame buffer, and 76% for the line buffer. The lesser saving with respect to line buffer is due to the higher bit-width required by the down-sampled version of image. Figure 6.4 shows a routine used for generating a downscaled image. This circuit consists of two blocks:

- Temporary accumulator (Block RAM: 16 bit $\times$ 240 $\times$ 2)
  8 $\times$ 8 $\times$ 240 pixel accumulation
- Compressed double frame buffers
  (Block RAM: 10 bit $\times$ 240 $\times$ 135 $\times$ 2)

Post color conversion, white channel of the line scanned input image is sequentially accumulated in the temporal accumulator. In order to obtain the downscaled image we accumulate 64 pixels which corresponds to an 8$\times$8 patch. Now the average of this accumulated data, which is calculated by 6-bit shift operation, is transferred to the compressed frame buffer. Each downscaled image is alternately stored to one side of double buffers (Fig. 6.4).

Table 6.4: Memory size of compressed frame/line buffers

Frame buffer	Pixel [bit]	Total [kByte]	Line buffer	Unit [bit]	Total [kByte]
Not compressed	10	2,073	Not compressed	92	44.16
Compressed	10	32.4	Compressed	124	7.48

6.3.3 Memory controller

In the memory controller, Hsync and Vsync are used for address generation, for writing W(t) to mini-image buffer. The Vsync alternately selects double buffer and resets the address counter at the time of frame change. The Hsync resets the counter for generating row R/W timing and controlling line buffers. Delayed Hsync and Vsync are reused for HDMI output control.

6.3.4 Our local histogram estimation

Figure 6.5 presents our local histogram estimation method and it is composed of the following modules:

- Compressed double frame buffers.
- Line buffers for column histogram.
- Temporary buffer for local histogram.
- Shifters and adders.

In our local histogram estimation using the down-sampled image data, we must maintain the accuracy of the original image. Since the downscale ratio is 1/64, we apply 64 times expansion to the down-sampled pixel data when estimating the histogram. Upon reading two pixels from the buffer, we perform a 3-bit shift. This operation makes 8 times bit-depth expansion on the down-sampled pixel data, so relative pixel weight is one eighth of original data. Here, two 11-bit pixels are used to update column histograms in the line buffer by performing one pixel addition and subtraction. After updating, using the two 124-bit column data with 3-bit shift, local histogram estimation

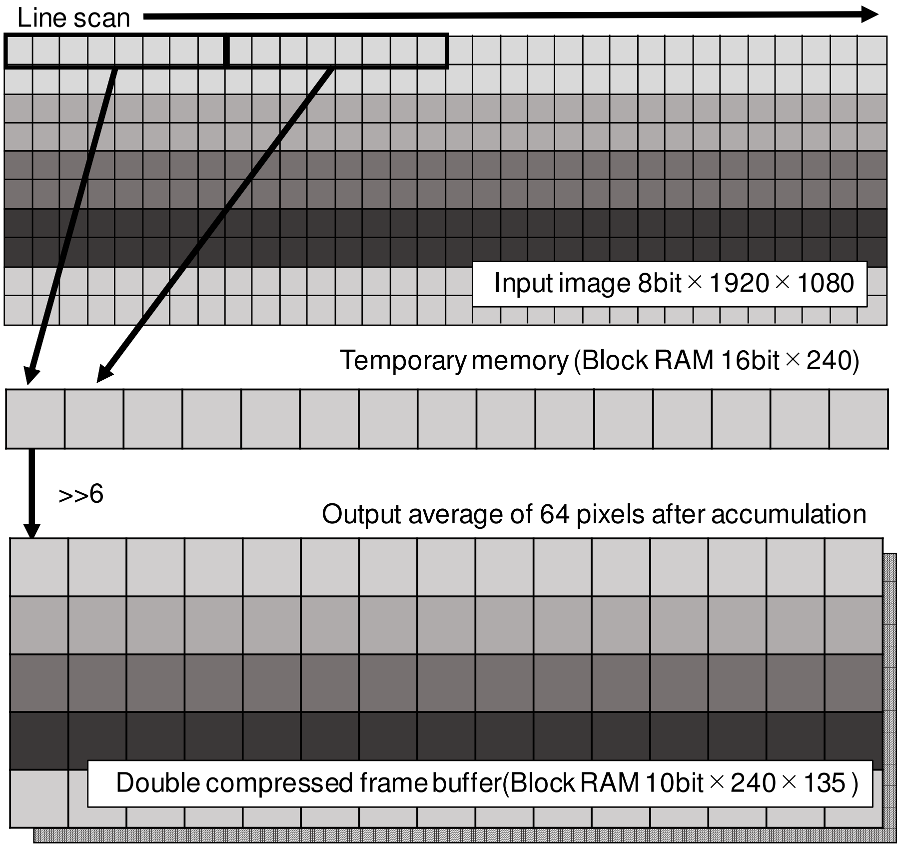


Figure 6.4: Compressed image generation routine.

can be achieved by one column addition and subtraction while maintaining the same accuracy as if working with the original image. Actual bit-width of the downscaled image is 10 bits, which consists of 8-bit data, 1-bit for overflow detection, and 1-bit is used for the following operation. When this bit is set to '1' it corresponds to texture pixel and a '0' corresponds to the gradation pixel. The bit-width of the line/temporary buffer is shown in Table 6.5. The width of the line buffer is determined by max kernel size which is 31 $\times$ 31 box shape, and it corresponds to our relative local area size of 248$\times$248. Also, the width of the temporary buffer is determined by the bit-depth of the line buffer and the kernel size. For smooth operation between frames, it is required to pre-store the line buffer along with the local patch of the first line, of the current frame. And, during the change of line, it is required to pre-store the data corresponding to the local patch of the subsequent line.

Next, using an example we will explain how a local histogram up-

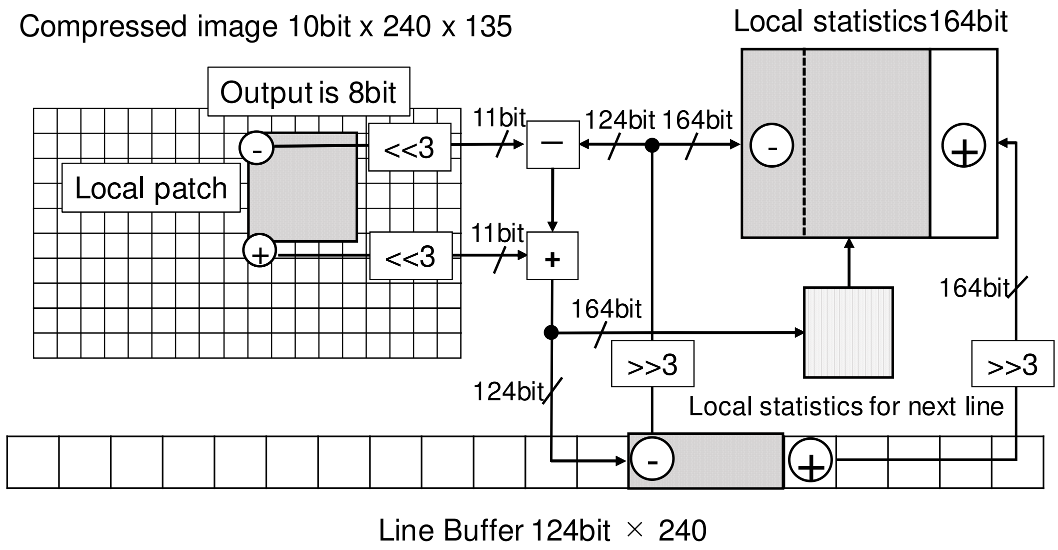


Figure 6.5: Local histogram update routine.

date sequence is performed. Consider that a tone mapped pixel's address is given as (x, y), then the compressed frame buffer address is $(x/8, y/8)$. And, the kernel size is $2N + 1$, then an updated column histogram's address is obtained as shown in Eq. 6.2 and Fig. 3.13(c). Then, the compressed frame buffer address for updating column histogram is given in Eq. 6.3 and Fig. 3.13(b) ($CompppressedImgAddr$ are compressed image's address $0 \leq Xdirection \leq 239$ and $0 \leq Ydirection \leq 134$) and streamlining of column histogram address for local statistics are done as in Eq.6.4 and Fig. 3.13(c). In this implementation, the column histogram is updated by computing line buffer address as in Eq. 6.2, and our kernel size is 31. When the pixel/column accessing is done over the size of line/frame buffers respectively, the address is inverted around the boundary of the image; and a quasi-image mirroring is performed. Output is added/subtracted and input to column histogram but updating cannot work at $x = 239 - ((2N + 1) - 1)/2$ to 239.

$$LineBufAddr = \frac{x}{8} - \frac{(2N+1)-1}{2} \tag{6.2}$$

$$
\begin{aligned}
CompppressedImgAddrX_{Up} &= \frac{x}{8} - \frac{(2N+1)-1}{2} \\
CompppressedImgAddrW_{Up} &= \frac{y}{8} - \frac{(2N+1)-1}{2} \\
CompppressedImgAddrX_{Under} &= \frac{x}{8} - \frac{(2N+1)-1}{2} \\
CompppressedImgAddrW_{Under} &= \frac{y}{8} + \frac{(2N+1)-1}{2} + 1
\end{aligned}
\tag{6.3}
$$

$$
\begin{aligned}
LineBufAddr_{Right} &= \frac{x}{8} - \frac{(2N+1)-1}{2} \\
LineBufAdddr_{Left} &= \frac{x}{8} + \frac{(2N+1)-1}{2} + 1
\end{aligned}
\tag{6.4}
$$

6.3.5 Tone mapping module

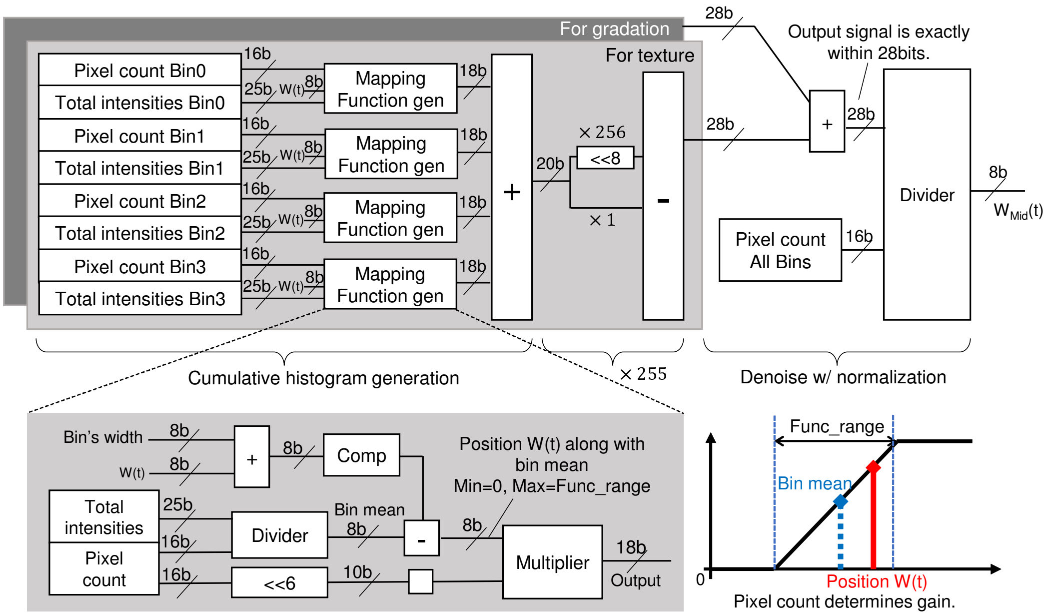


Figure 6.6: Tone mapping module: Above is cumulative histogram generation circuit followed by normalization and denoise. Below is blown up view of a mapping function.

Figures 6.6 and 6.7 presents our tone mapping module which consists of the following blocks:

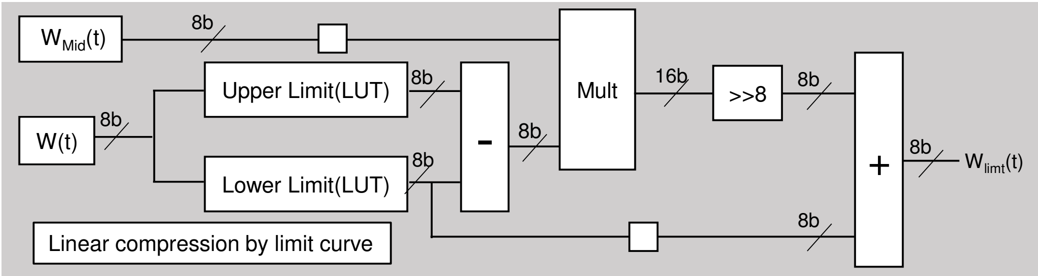


Figure 6.7: Tone mapping module: Upper/lower limit curve circuit.

- Mapping function generator that obtains the average value of each of the four bins and generates a tone curve (each unit has 2 × 8bit, 25bit and 16bit input and 18bit output).

- Tone mapping module that integrates the tone curves for gradation/texture components and merges normalized curve to 8bits (each unit has 4 × 16 bit ,4 × 25bit and 8bit input and 8bit output).

- Linear compression using limiting curve which includes two block RAMs that holds upper and lower limit curve.

- Multipliers and dividers were implemented using Xilinx IP.

Table 6.5: Bit depth at each bin

	Colum histogram	Local Histogram
Pixel	10 bit	16 bit
Luminance	20 bit	25 bit

In our tone mapping function generator, the integral form for each bin is calculated using the following information; number of pixels, accumulated luminance, and $Func_{range}$. This module measures relative $W(t)$ position with bin mean and calculates multiplication between pixel count and minimum(= 0) or maximum(=$Func_{range}$) or $W(t)$ position as integrated local function.The tone mapping module performs summation of integral forms for generating the final local tone mapping function, following which, it calculates tone mapped white channel. This channel including gradation/texture components are generated and merged for noise suppression. Following which, the module performs normalization. The TMO is completed by applying the space limitation curve which restricts the gains for the mapping functions that are related to the local contrast enhancement. The limitation curve is pre-calculated and stored in the lookup table. Equation 6.5 shows the final $W_{TM}(t)$ channel calculation by the limiting curves which linearly compresses it within the upper and lower limiting curves. The pre-determined L_{upper} and L_{lower} (L_{upper} = Upper limit, L_{lower} = Lower limit) are outputted by addressing of the luminance value W_{mid} before conversion (see Fig. 3.9 in section 3.4).

$$W_{TM}(t) = L_{lower} + (L_{upper} - L_{lower})\frac{W_{mid}}{255} \tag{6.5}$$

6.3.6 Color converter

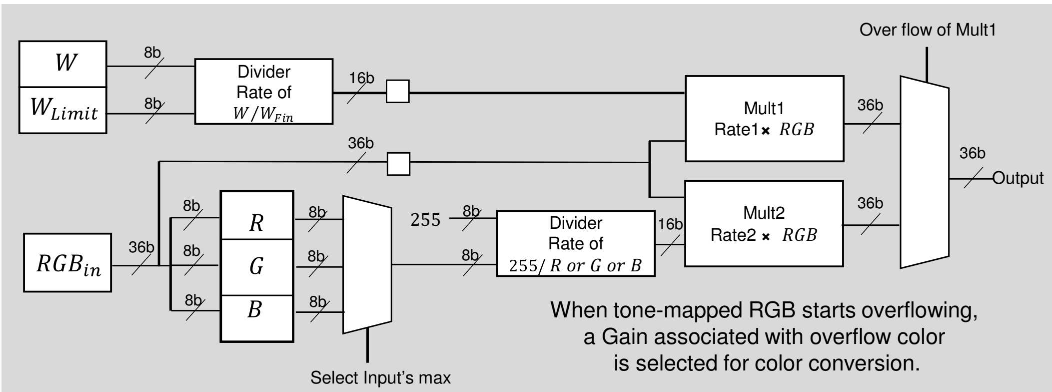


Figure 6.8: Output color converter: White to RGB channel

Fig.6.8 shows our output color converter module, which uses the ratio between W_{in} and W_{fin} to obtain the final RGB_{out}.The output is converted based on the ratio W_{in}/W_{fin} (Eq. (6.6)). However, if one of the RGB channel overflows, another conversion is performed while maintaining color tone using the maximum value of RGB (Eq. (6.7)).

$$RGB_{TM}(t) = \frac{W_{TM}(t)}{W(t)}RGB \tag{6.6}$$

$$RGB_{TM}(t) = \frac{RGB(t)}{M} \quad (M = max(R, G, B)) \tag{6.7}$$

6.4 Xilinx FPGA synthesis summary

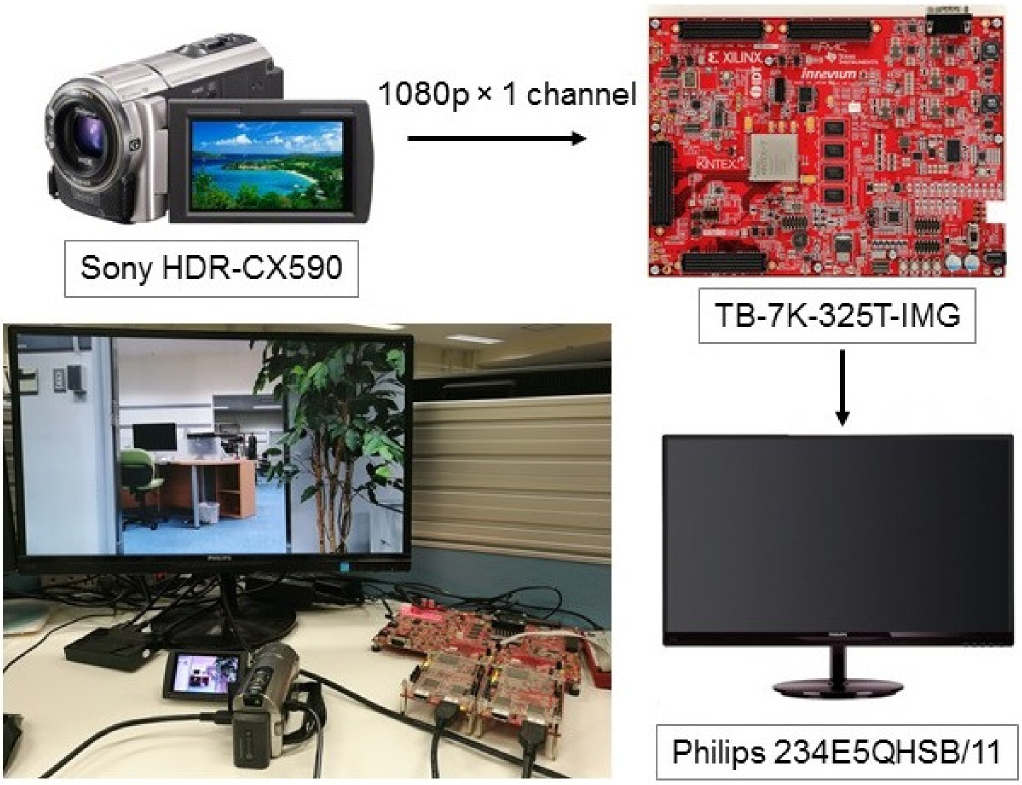


Figure 6.9: Our hardware experimental setup.

Table 6.6: FPGA resource utilization summary

	Slice	Slices Reg	Slice LUTs	BRAM/FIFO	DSP48E1
FPGA Ressources	50,950	203,800	407,600	445+890	840
Used	13,420	43,479	31,685	222 + 97	38
Utilization	26%	10%	15%	32%	4%

Table 6.7: Performance evaluation

Input	1920 × 1080(Full HD)
Output	1920 × 1080(Full HD)
Color space	RGB:10bit †
Framerate	60FPS
FPGA clock	162MHz

†Our tone mapping system converts RGB color to White and calculates only its value for tone mapping operation.

We implemented the proposed tone mapping algorithm on FPGA board (TB - 7K - 325TIMG, Xilinx Kintex - 7). The architectural block diagrams

Table 6.8: Contrast measure for video output

	Edge Measure	
Sequence	Global TMO	Global & Local TMO
(a)	2467	70424
(b)	704	5615
(c)	1959	29634

Table 6.9: Comparison with related works

	[59]	[154]	[60]	[162]	[151]	Our TMO	
Chip	Virtex-5 (60nm)	Spartan-6 (45nm)	Virtex-5 (60nm)	Cyclone-3 (65nm)	Stratix II (90nm)	Virtex-5 (60nm)	Kintex-7 (28nm)
Slice LUTs	14616	874	918	93989	34806	10903	9799
Slice Registers	8132	1000	540	353	-	12794	15345
Block RAM/FIFO(Kb)	144	108	0	-	3153	7488	9738
DSP48Es	4	17	30	28	54	22	21
Image size	1280×1024	1080p*	1024×768	1024×768	1024×768	1920×1080	1920×1080
Kernel size	-	-	-	5	64	31†	31†
TMO	Global only			Global & Local			
Frequency(MHz)	94.733	75	214.7	100	77.15	148	162
Power (mW)	-	-	-	674.25	-	804	453
Throughput(MPixels/s)	39	62	133	116	52.24	100	124
Noise Suppression	×	×	×	×	×	✓	✓
Halo Management	×	×	×	✓	✓‡	✓	✓

*1080p @30Hz[154]. †Relative to our kernel size of (248×248) for full HD. ‡Fixed manually.

of our tone mapping system are shown in Fig. 6.2 to 6.8. This algorithm is coded in Verilog-HDL, synthesized and place-and-routed using Xilinx ISE (14.7). We confirmed our system performance by tone mapping video at 60 FPS. Our experimental setup including video camera (Sony HDR-CX590), HDMI input/output (TB-FMCH-HDMI2 RX/TX), the FPGA board with our tone mapping system, and display monitor (Philips 234E5QHSB) are shown in the Fig. 6.9. The input images (1920×1080) are given through the HDMI FMC board, and the output tone mapped images are also displayed on the monitor via HDMI FMC boards. The system resource usage information is presented in table 6.6. The overall footprint of our system requires only 26% of the total available slices, 10% registers, 15% of LUTs, 4% DSP and 32% of the BRAM. The entire tone mapping pipeline process has a total latency of 131 clocks, which corresponds to $0.81\mu s$ for a 162 MHz system clock. This is very low latency, and goes unnoticed by human eye and then, the pipeline process is able to deliver a full HD video stream at a frame rate of 60 FPS (see table 6.7).

6.4.1 Comparison with other related implementations

Table 6.9 summarizes and compares our SLHE implementation with other FPGA implemented tone mapping systems. Algorithms in [59], [60] and [154] perform only global tone mapping. In comparison they require only fewer computational resources as they apply the same function to all pixels, but may generate lower quality output images, when compared to local tonemap algorithms [182]. Local tone mapping functions are commonly known to generate halos in tone mapped images [183]. Ambalathankandy et al., implemented a global and local tone mapping operator [162]. However, their FPGA implementation resulted in a larger footprint, as their TMO function requires an additional halo reducing filter which is based on a Gaussian-like filter [183]. In their TMO operator, halos were created around small bright features due to strong attenuation of neighboring pixels due to convolution operation with low-pass filtering. The hardware scheme for reducing such halos resulted in a very expensive implementation. The tone mapping operator implemented by Hassan et al., uses global and local image information, but their system can process only gray scale images [151]. The halo control method presented in [151] has been manually fixed, which may require a re-calibration for processing different images. In section 3.4.2 we discuss an effective halo control mechanism which requires no manual control.

Local contrast enhancement algorithms also risk boosting noise. Our proposed SLHE algorithm performs global and local tonemap, and it is likely that local tone mapping function can enhance noise. Therefore, in section 3.4.1 we discussed a method to reduce noise in tone mapped images, and evaluated its effectiveness by performing objective assessment. Our TM circuit does require more FPGA resources than those reported by global only tone mapping circuits. However, for the additional computational resources we are able to obtain more contrast in our output. In table 6.8 we report contrast enhanced in locally tone mapped output images, when compared to global only images. The results in table 6.8, were computed as follows: Ten images from a video sequence, and its tone mapped version were captured. And, we computed their local edge measure as an absolute value of $\sum Y_{edge} + \sum X_{edge}$ using sobel like edge detector [130]. We chose sobel edge detector as it is less sensitive to noise. From this table we can easily notice that for the additional hardware that our implementation requires, it produces images with more contrast. Our hardware implementation, by making use of 1/8th downscaled image and a bin reduced histogram saves large amount of memory and requires fewer pixel access. It is worth noting that, our implementation is self-reliant and requires no external memories for buffering images. The power consumption of our tone mapping function was calculated using Xilinx Power Estimator [184], and it is 453 mW of power for Xilinx Kintex 7 (28nm). The total power consumption of our FPGA system is about 3 Watts, including two HDMI FPGA Mezzanine Cards. We also synthesized our tonemap function targeting a Xilinx Virtex-5 (60nm), and its implementation results are listed in table 6.9. This Virtex-5 implementation consumes 804 mW of power.

6.5 Conclusion

In this chapter we presented the real-time FPGA implementation for the adaptive tone mapping algorithm which we introduced in section 3.4. Even though several algorithms that combine global and local tonal control methods have been proposed, they are known to be difficult to manage in a unified manner. Therefore, they require additional computational circuits or methods to produce quality results. In this chapter, we reported a method by which we could easily manage the global and local tone curves in the same tone mapping space. In our FPGA implementation, we were able to minimize the system latency to $0.81\mu s$ @162 MHz system clock and reduce 98% frame memories by using down-sampled local statistics from the previous frame. Our tone mapping algorithm operates on white luminance statistics (W). As a future direction, we would like to implement the human perception-based color to luminance mapping which we presented in section 3.3. Our immediate next objective is to realize its hardware implementation. This can help us develop a more effective tonemap operator that will be a desirable feature yielding more practical applications.

From our brief survey of real-time implementations that we reported in section 6.2, we saw that from early days FPGAs have been preferred platform for realizing real-time tone mapping applications. As a concluding remark we will list the main reasons for FPGAs popularity as a reliable accelerator platform:

- High Performance: The underlying FPGA fabric supports development of very deep pipeline architectures, with scope for wide parallel computational elements. This flexibility permits the designer to easily develop complex functionalities with strict timing constraints. Current generation of FPGAs can accelerate a whole algorithm while processing full HD images or higher resolution images (for example [73],[166]).

- Re-programmable: This is one of the most attractive feature of FPGA. A designer can iterate his design to make sure that he can tune and optimize his design to meet all specifications of targeted application without incurring additional cost.

- Low development cost: The cost of developing a design on FPGA is comparatively cheaper, as many vendors provide customizable IP's and reference designs which can speed-up the development process.

- Development tools: Good support is available in the form of full development suite. Altera and Xilinx have streamlined development software for design and verification.

- Flexibility for ASIC migration: FPGA proven design can be ported to structured ASICs which are available from many vendors, thereby giving developers a faster route to market their products[185, 186, 187].

Chapter 7

Conclusions and future work

7.1 Thesis Summary

This thesis reports a new adaptive tone mapping algorithm, along with its real-time FPGA implementation. Even though several algorithms exist in literature that can combine global and local tone control, they are known to be difficult to manage in a unified manner. Therefore, they require additional computational circuits or methods to generate quality results. Our smoothed LHE method can easily manage the global and local tone curves in the same tone mapping space. Traditionally, tone mapping algorithms operate on luminance channels which can lead to some loss of information. We minimize this data loss by employing a human perception based color to luminance mapping scheme. Essentially, a decolorization method is expected to have a key role in the pre-processing of tone mapping or edge preserving filters. Therefore, it is ideal to realize luminance mapping methods that require low calculation cost and fast operational time as a pre-processing step. To address this requirement, in this thesis we have developed a fast $O(1)$ decolorization method that reflects human perception. It refers to RGB values in one pixel and performs weighted blending of the Euclidean distances of warm/cool color vectors. This simple conversion outputs a luminance channel that is comparable to the conventional optimization methods using iterations. When our method was applied to tone mapping, it achieved better results than have been obtained with YCbCr/HSV color space.

Lateral inhibition is an important early stage mechanism of visual pro-

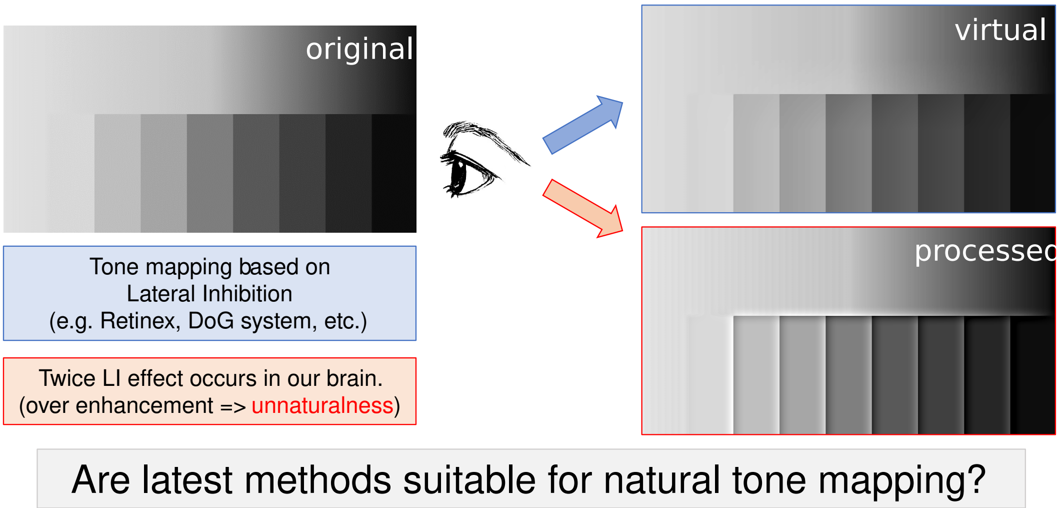


Figure 7.1: The Chevreul illusion is used to examine the effectiveness of conventional tone mapping methods. The tone mapped image is perceived unnatural due to over enhancement.

cessing that helps the brain cope with the enormous redundant visual information it receives from the surrounding. This inhibitory mechanism leads to many perceptual effects and needs to be considered in the algorithm's specification. In this thesis we have studied in detail how LI influences our perceptions when dealing with images which have optical illusionary effects, and an example is presented in Fig . 7.1. From our study we find that, it is the stacked LI effect which causes the unnaturalness. This stacking effect is illustrated using the luminance profile in Fig. 7.2. Our proposed smoothed LHE algorithm which operates with large kernel size, and forms saw-tooth like edges avoids the stacked LI effect by emulating the recognized luminance profile (refer section 4.3 and Fig. 4.4). We thoroughly analyzed state of the art edge preserving filters for processing these illusionary images and find they have certain inherent limitations which causes them to produce images with over emphasized edges due to the stacked LI effect. Medical images like digital X-rays are low contrast images, that are commonly known to have optical illusions like Mach bands and background contrast effects. These are also caused by LI. We observe that using multilayer (ML) methods with latest edge preserving filter for contrast enhancement in medical images can be problematic and could lead to faulty diagnosis from detail exaggeration which are caused by uncontrolled

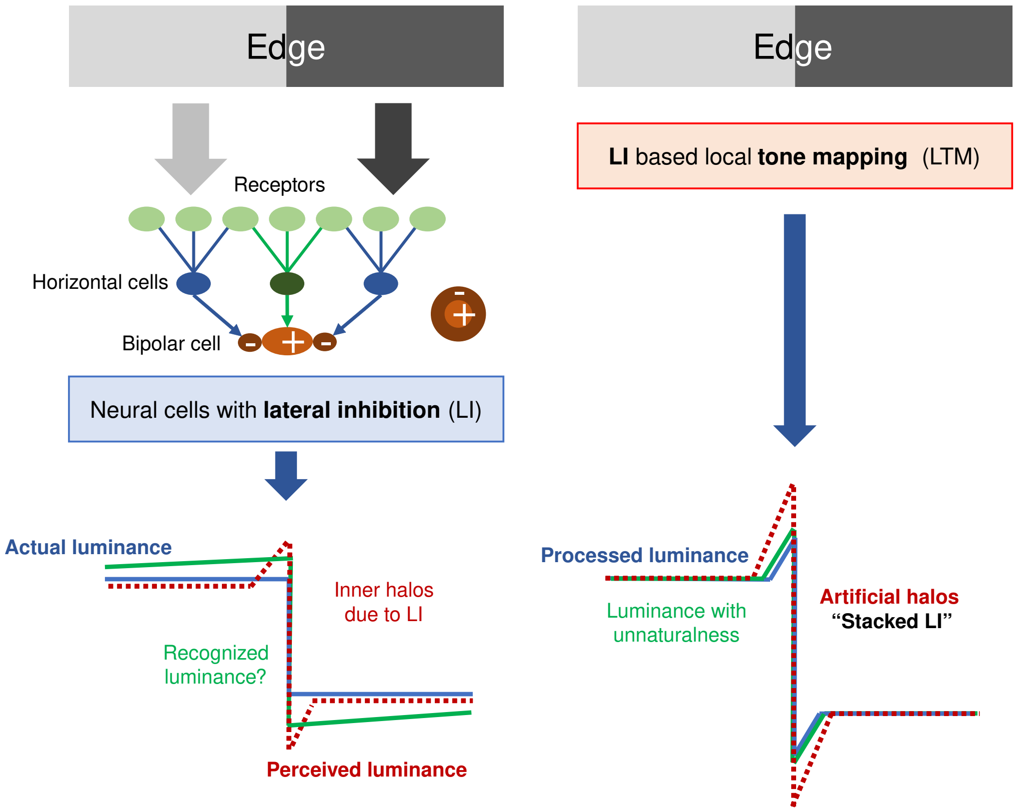


Figure 7.2: Comparing the luminance profile of HVS and perceived luminance we can understand the stacked LI effect in tone mapped images which causes the unnatural feeling.

texture boosting from the stacked LI effect and user defined gain settings. Furthermore, ML filters are designed with few subjectively selected filter kernel sizes, which can result in unnaturalness in output images. We analyze, and report that our smoothed LHE filter with an adaptive gain control, is more robust and can enhance fine details in digital X-rays while maintaining their intrinsic naturalness.

Also, in this thesis we present a real-time FPGA implementation for our proposed smoothed LHE tone mapping algorithm. In our FPGA implementation, we were able to minimize the system latency to $0.81\mu s$ @162 MHz system clock and reduce 98% frame memories by using down-sampled local statistics from a previous frame. Our tone mapping algorithm operates on white luminance statistics (W). As a future direction, we would like to implement the human perception-based color to luminance mapping which we presented in section 3.3, and our immediate next objective is to realize its hardware implementation. This can help us develop a more perceptually appealing tonemap operator, which will be a desirable feature yielding more practical applications.

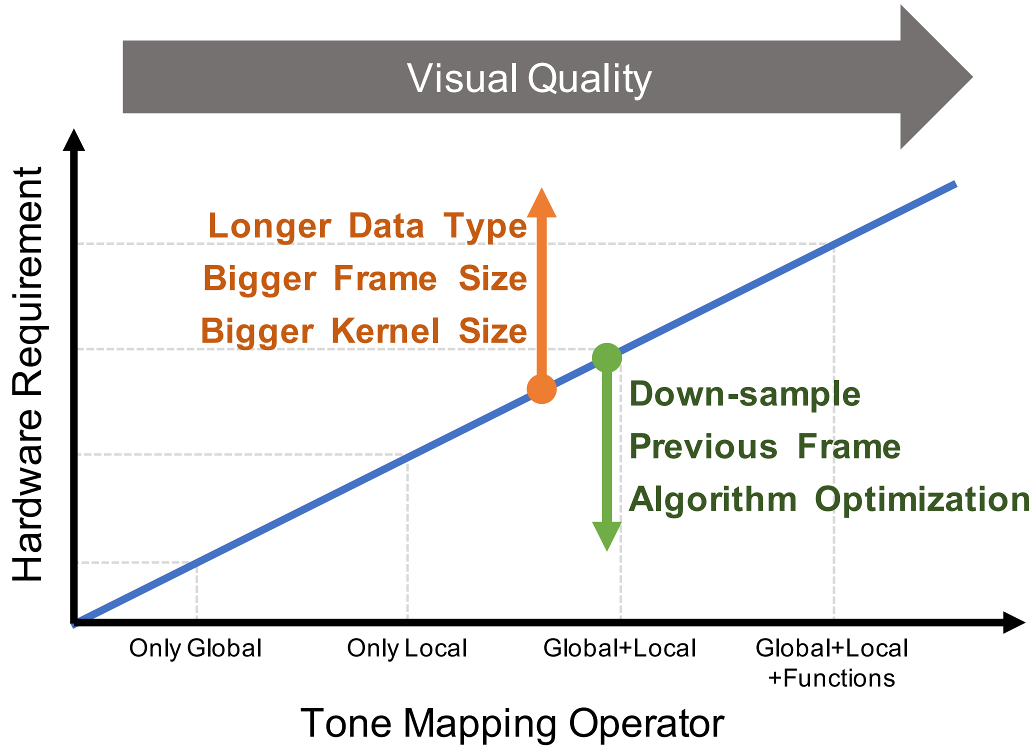


Figure 7.3: Trade-off: Visual quality versus hardware specification.

To summarize, every image processing application seeks to achieve good output image quality, and TMOs are no different. Local TMO's are known to produce better images than global TMOs as they can reproduce both global and local contrast [12]. However, local TMOs are computationally more expensive than global TMOs and may also generate artifacts like halos around edges and amplify noise [66]. Therefore, additional functions are required along with local TMOs to reduce such artifacts. These local operations must be repeated over large amounts of data and usually demand substantial computational effort as shown in Fig .7.3, where the visual quality does improve but at the expense of additional hardware. At the same time, depending on the data-type, image and kernel size will also increase the cost. As shown in Fig .7.3, cost can be reduced by applying techniques like operating the algorithm with previous frame/down-sampled images or employing other optimization techniques that can reduce computation or memory operations.

7.2 Future Work

Currently machine learning-based (ML) methods have become a very important tool to solve many computer vision tasks like image classification, face detection, and video analysis. As a future perspective we would like to leverage its potential by accelerating ML-based TMOs using hardware platforms. In this section we will explore the challenges and opportunities that we will encounter for such systems. Usually image processing tasks would require multiple convolution with fully connected layers, which are exorbitantly computationally intensive (for example, the operations in CNNs are over billion operations [188]). Realizing such systems on resource constrained embedded systems would require very novel architectures and algorithms. FPGAs have been the preferred platform for realizing CNN hardware accelerators for their following well-known features: re-programmability, low-power design features, and quick design time [189, 190].

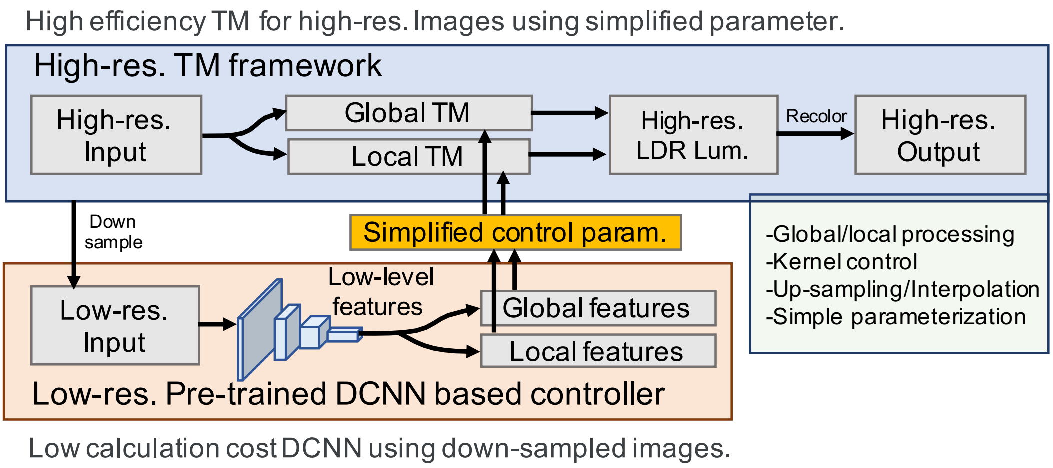


Figure 7.4: Block diagram of a plausible machine learning-based tonemap operator which can be implemented on hardware.

Using Fig. 7.4 we demonstrate how Gharbi et al.'s HDR-Net like architectures are ideal baselines for realizing TMOs with DNN on hardware

[191]. We find such designs are more hardware friendly because of the following features. First, the bilateral grid inspired architecture represents local tone-control as simple parameterized luminance grid in the space. Thus, the HDR-Net like high throughput design can approximate a tonemap system by using a high-resolution guidance map which slices into the grid to produce a unique, interpolated, affine transform to be applied to each input pixel. Second, a lightweight DNN is vital for realizing hardware TMO system. Thus, low-resolution DNN with down-sampling and optimum interpolation are important. Also, simple data transfer between DNN and tone mapping system is required for reducing hardware load. In this scenario, output format of such DNN architectures becomes simple. Finally, good high resolution off-line data set and training method for whole architecture is key for realizing this system.

Bibliography

[1] "Sample hdr photo processed with easyhdr," https://www.easyhdr.com/examples/cap-de-formentor/, accessed on February 1st 2020.

[2] I. Popadić, B. M. Todorović, and I. Reljin, "Method for hdr-like imaging using industrial digital cameras," *Multimedia Tools and Applications*, vol. 76, no. 10, pp. 12 801–12 817, 2017.

[3] "Edinburgh skyline at sunset," https://www.freeimageslive.co.uk/free_stock_image/city-edinburgh-jpg, accessed on March 1st 2020.

[4] M. Igarashi, A. Mizuno, and M. Ikebe, "Accuracy improvement of histogram-based image filtering," in *2013 IEEE International Conference on Image Processing*, pp. 1217–1221. IEEE, 2013.

[5] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," in *ACM TOG*, vol. 27, no. 3, p. 67. ACM, 2008.

[6] E. S. Gastal and M. M. Oliveira, "Domain transform for edge-aware image and video processing," in *ACM ToG*, vol. 30, no. 4, p. 69. ACM, 2011.

[7] V. Chesnokov, "Image enhancement methods and apparatus therefor," Nov. 27 2007, uS Patent 7,302,110.

[8] I. Huberman and R. Fattal, "Reducing lateral visual biases in displays," in *CGF*, vol. 35, no. 8, pp. 19–31. Wiley Online Library, 2016.

[9] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE PAMI*, vol. 35, no. 6, pp. 1397–1409, 2013.

[10] "The best 4k monitors for 2020," https://www.pcmag.com/roundup/370480/the-best-4k-monitors, accessed on December 24, 2019.

[11] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann, 2010.

[12] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced high dynamic range imaging*. AK Peters/CRC Press, 2017.

[13] K. Jacobs, C. Loscos, and G. Ward, "Automatic high-dynamic range image generation for dynamic scenes," *IEEE Computer Graphics and Applications*, vol. 28, no. 2, pp. 84–93, 2008.

[14] T. Kunkel and E. Reinhard, "A reassessment of the simultaneous dynamic range of the human visual system," in *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, pp. 17–24, 2010.

[15] J. A. Ferwerda, "Elements of early vision for computer graphics," *IEEE computer graphics and applications*, vol. 21, no. 5, pp. 22–33, 2001.

[16] "Sony oled broadcast reference monitor bvm-x300 v2," https://pro.sony/ue_US/products/broadcastpromonitors/bvm-x300-v2#ProductSpecificationsBlock-bvm-x300-v2, accessed on May 19 2020.

[17] O. S. Pianykh, *Digital imaging and communications in medicine (DICOM): a practical introduction and survival guide*. Springer Science & Business Media, 2009.

[18] S. Athar, H. Yeganeh, and Z. Wang, "Data rate and dynamic range compression of medical images: Which one goes first?" in *2015 IEEE International Conference on Image Processing (ICIP)*, pp. 4436–4440. IEEE, 2015.

[19] G. Ward and R. Shakespeare, "Rendering with radiance: the art and science of lighting visualization," 1998.

[20] F. X. Sillion and C. Peuch, "Radiosity & global illumination," 1994.

[21] F. Durand and J. Dorsey, "Interactive tone mapping," in *Rendering Techniques 2000*, pp. 219–230. Springer, 2000.

[22] F. Dufaux, P. Le Callet, R. Mantiuk, and M. Mrak, *High dynamic range video: from acquisition, to display and applications.* Academic Press, 2016.

[23] M. Mustra, K. Delac, and M. Grgic, "Overview of the dicom standard," in *2008 50th International Symposium ELMAR*, vol. 1, pp. 39–44. IEEE, 2008.

[24] J. H. Jungmann, L. MacAleese, J. Visser, M. J. Vrakking, and R. M. Heeren, "High dynamic range bio-molecular ion microscopy with the timepix detector," *Analytical chemistry*, vol. 83, no. 20, pp. 7888–7894, 2011.

[25] G. Chander, B. L. Markham, and D. L. Helder, "Summary of current radiometric calibration coefficients for landsat mss, tm, etm+, and eo-1 ali sensors," *Remote sensing of environment*, vol. 113, no. 5, pp. 893–903, 2009.

[26] R. Szeliski, *Computer vision: algorithms and applications.* Springer Science & Business Media, 2010.

[27] "Shutter speed," https://en.wikipedia.org/wiki/Shutter_speed, 5 December 2019.

[28] J. Telleen, A. Sullivan, J. Yee, O. Wang, P. Gunawardane, I. Collins, and J. Davis, "Synthetic shutter speed imaging," in *Computer Graphics Forum*, vol. 26, no. 3, pp. 591–598. Wiley Online Library, 2007.

[29] S. Mann and R. Picard, "On being 'undigital'with digital cameras: Extending dynamic range by combining differently exposed pictures, 7 pages."

[30] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, vol. 1, pp. 374–380. IEEE, 1999.

[31] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *ACM SIGGRAPH 2008 classes*, p. 31. ACM, 2008.

[32] M. D. Tocci, C. Kiser, N. Tocci, and P. Sen, "A versatile hdr video production system," in *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4, p. 41. ACM, 2011.

[33] S. Kavadias, B. Dierickx, D. Scheffer, A. Alaerts, D. Uwaerts, and J. Bogaerts, "A logarithmic response cmos image sensor with on-chip calibration," *IEEE Journal of Solid-state circuits*, vol. 35, no. 8, pp. 1146–1152, 2000.

[34] M. Loose, K. Meier, and J. Schemmel, "A self-calibrating single-chip cmos camera with logarithmic response," *IEEE Journal of Solid-state circuits*, vol. 36, no. 4, pp. 586–596, 2001.

[35] S. Sugawa, N. Akahane, S. Adachi, K. Mori, T. Ishiuchi, and K. Mizobuchi, "A 100 db dynamic range cmos image sensor using a lateral overflow integration capacitor," in *ISSCC. 2005 IEEE International Digest of Technical Papers. Solid-State Circuits Conference, 2005.*, pp. 352–603. IEEE, 2005.

[36] N. Akahane, S. Sugawa, S. Adachi, K. Mori, T. Ishiuchi, and K. Mizobuchi, "A sensitivity and linearity improvement of a 100-db dynamic range cmos image sensor using a lateral overflow integration capacitor," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 4, pp. 851–858, 2006.

[37] N. Ide, W. Lee, N. Akahane, and S. Sugawa, "A wide dr and linear response cmos image sensor with three photocurrent integrations in photodiodes, lateral overflow capacitors, and column capacitors," *IEEE journal of solid-state circuits*, vol. 43, no. 7, pp. 1577–1587, 2008.

[38] Y. Liu, "The design of a high dynamic range cmos image sensor in 110nm technology," *Delft University of Technology*, pp. 1–34, 2012.

[39] M. Mase, S. Kawahito, M. Sasaki, Y. Wakamori, and M. Furuta, "A wide dynamic range cmos image sensor with multiple exposure-time signal

outputs and 12-bit column-parallel cyclic a/d converters," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2787–2795, 2005.

[40] T. Viitanen, P. Jääskeläinen, O. Esko, and J. Takala, "Simplified floating-point division and square root," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2707–2711. IEEE, 2013.

[41] T. Dobashi, A. Tashiro, M. Iwahashi, and H. Kiya, "A fixed-point implementation of tone mapping operation for hdr images expressed in floating-point format," *APSIPA Transactions on Signal and Information Processing*, vol. 3, 2014.

[42] C. H. Lampert and O. Wirjadi, "Anisotropic gaussian filtering using fixed point arithmetic," in *2006 International Conference on Image Processing*, pp. 1565–1568. IEEE, 2006.

[43] C. Ofili, S. Glozman, and O. Yadid-Pecht, "Hardware implementation of an automatic rendering tone mapping algorithm for a wide dynamic range display," *Journal of Low Power Electronics and Applications*, vol. 3, no. 4, pp. 337–367, 2013.

[44] J. Yang, U. Shahnivich, and O. Yadid-Pecht, "Mantissa-exponent based tone mapping for wide dynamic range image sensors," *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2019.

[45] L. Meylan, "Tone mapping for high dynamic range images," EPFL, Tech. Rep., 2006.

[46] B. Hoefflinger, *High-dynamic-range (HDR) vision.* Springer, 2007.

[47] J. Kuang, G. M. Johnson, and M. D. Fairchild, "icam06: A refined image appearance model for hdr image rendering," *Journal of Visual Communication and Image Representation*, vol. 18, no. 5, pp. 406–414, 2007.

[48] D. Lischinski, Z. Farbman, M. Uyttendaele, and R. Szeliski, "Interactive local adjustment of tonal values," in *ACM Transactions on Graphics (TOG)*, vol. 25, no. 3, pp. 646–653. ACM, 2006.

[49] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion," in *15th Pacific Conference on Computer Graphics and Applications (PG'07)*, pp. 382–390. IEEE, 2007.

[50] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, "High dynamic range display systems," *ACM transactions on graphics (TOG)*, vol. 23, no. 3, pp. 760–768, 2004.

[51] K. D. A. C. A. Wilkie and W. Purgathofer, "Tone reproduction and physically based spectral rendering." Eurographics, 2002.

[52] K. Myszkowski, R. Mantiuk, and G. Krawczyk, "High dynamic range video," *Synthesis Lectures on Computer Graphics and Animation*, vol. 1, no. 1, pp. 1–158, 2008.

[53] G. Ward, "A contrast-based scalefactor for luminance display," *Graphics gems IV*, pp. 415–421, 1994.

[54] C. Schlick, "Quantization techniques for visualization of high dynamic range pictures," in *Photorealistic rendering techniques*, pp. 7–20. Springer, 1995.

[55] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg, "A model of visual adaptation for realistic image synthesis," in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pp. 249–258, 1996.

[56] J. Tumblin, J. K. Hodgins, and B. K. Guenter, "Two methods for display of high contrast images," *ACM Transactions on Graphics (TOG)*, vol. 18, no. 1, pp. 56–94, 1999.

[57] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," in *Computer graphics forum*, vol. 22, no. 3, pp. 419–426. Wiley Online Library, 2003.

[58] P.-J. Lapray, B. Heyrman, M. Rossé, and D. Ginhac, "Smart camera design for realtime high dynamic range imaging," in *2011 Fifth ACM/IEEE*

International Conference on Distributed Smart Cameras, pp. 1–7. IEEE, 2011.

[59] P.-J. Lapray, B. Heyrman, M. Rossé, and D. Ginhac, "Hdr-artist: High dynamic range advanced real-time imaging system," in *2012 IEEE International Symposium on Circuits and Systems*, pp. 1428–1431. IEEE, 2012.

[60] V. Popovic, E. Pignat, and Y. Leblebici, "Performance optimization and fpga implementation of real-time tone mapping," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 61, no. 10, pp. 803–807, 2014.

[61] V. Kantabutra, "On hardware for computing exponential and trigonometric functions," *IEEE Transactions on Computers*, vol. 45, no. 3, pp. 328–339, 1996.

[62] I. Koren, *Computer arithmetic algorithms*. AK Peters/CRC Press, 2001.

[63] B. Parhami, *Computer arithmetic*, vol. 20, no. 00. Oxford university press, 2010.

[64] K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang, K. Zimmerman *et al.*, "Spatially nonuniform scaling functions for high contrast images," in *Graphics Interface*, pp. 245–245. Canadian Information Processing Society, 1993.

[65] S. N. Pattanaik, J. Tumblin, H. Yee, and D. P. Greenberg, "Time-dependent visual adaptation for fast realistic image display," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 47–54. ACM Press/Addison-Wesley Publishing Co., 2000.

[66] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *ACM transactions on graphics (TOG)*, vol. 21, no. 3, pp. 257–266. ACM, 2002.

[67] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," in *ACM transactions on graphics (TOG)*, vol. 21, no. 3, pp. 267–276. ACM, 2002.

[68] M. Ashikhmin, "A tone mapping algorithm for high contrast images," in *Proceedings of the 13th Eurographics workshop on Rendering*, pp. 145–156. Eurographics Association, 2002.

[69] S. Pattanaik and H. Yee, "Adaptive gain control for high dynamic range image display," in *SCCG*, vol. 2, pp. 83–87. Citeseer, 2002.

[70] M. Colbert, E. Reinhard, and C. E. Hughes, "Painting in high dynamic range," *Journal of Visual Communication and Image Representation*, vol. 18, no. 5, pp. 387–396, 2007.

[71] T.-H. Wang, W.-S. Wong, F.-C. Chen, and C.-T. Chiu, "Design and implementation of a real-time global tone mapping processor for high dynamic range video," in *2007 IEEE International Conference on Image Processing*, vol. 6, pp. VI–209. IEEE, 2007.

[72] S. Marsi, G. Impoco, A. Ukovich, S. Carrato, and G. Ramponi, "Video enhancement and dynamic range control of hdr sequences for automotive applications," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, p. 080971, 2007.

[73] P. Ambalathankandy, M. Ikebe, T. Yoshida, T. Shimada, S. Takamaeda, M. Motomura, and T. Asai, "An adaptive global and local tone mapping algorithm implemented on fpga," *IEEE Transactions on Circuits and Systems for Video Technology*, 2019.

[74] M. Herscovitz and O. Yadid-Pecht, "A modified multi scale retinex algorithm with an improved global impressionof brightness for wide dynamic range pictures," *Machine Vision and Applications*, vol. 15, no. 4, pp. 220–228, 2004.

[75] L. Meylan and S. Susstrunk, "High dynamic range image rendering with a retinex-based adaptive filter," *IEEE Transactions on image processing*, vol. 15, no. 9, pp. 2820–2830, 2006.

[76] A. Benedetti, A. Prati, and N. Scarabottolo, "Image convolution on fpgas: the implementation of a multi-fpga fifo structure," in *Proceedings. 24th EUROMICRO Conference (Cat. No. 98EX204)*, vol. 1, pp. 123–130. IEEE, 1998.

[77] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355–368, 1987.

[78] D. Menotti, L. Najman, J. Facon, and A. d. A. Araújo, "Multi-histogram equalization methods for contrast enhancement and brightness preserving," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 3, pp. 1186–1194, 2007.

[79] G. Ramponi, N. K. Strobel, S. K. Mitra, and T.-H. Yu, "Nonlinear unsharp masking methods for image contrast enhancement," *Journal of Electronic Imaging*, vol. 5, no. 3, pp. 353–367, 1996.

[80] E. H. Land, "The retinex theory of color vision," *Scientific american*, vol. 237, no. 6, pp. 108–129, 1977.

[81] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM transactions on graphics (TOG)*, vol. 23, no. 3, pp. 664–672, 2004.

[82] S. Bae, S. Paris, and F. Durand, "Two-scale tone management for photographic look," *ACM Transactions on Graphics (TOG)*, vol. 25, no. 3, pp. 637–645, 2006.

[83] E. A. Khan, E. Reinhard, R. W. Fleming, and H. H. Bülthoff, *Image-based material editing*, vol. 25, no. 3. ACM, 2006.

[84] A. Ignatov, N. Kobyshev, R. Timofte, K. Vanhoey, and L. Van Gool, "Dslr-quality photos on mobile devices with deep convolutional networks," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3277–3285, 2017.

[85] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *ICCV 1998*, pp. 839–846. IEEE, 1998.

[86] J. L. Morton, *Color logic.* Colorcom, 1998.

[87] Y. Nayatani, "Simple estimation methods for the helmholtz kohlrausch effect," *Color Research & Application*, vol. 22, no. 6, pp. 385–401, 1997.

[88] A. A. Gooch, S. C. Olsen, J. Tumblin, and B. Gooch, "Color2gray: salience-preserving color removal," *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3, pp. 634–639, 2005.

[89] M. Grundland and N. A. Dodgson, "Decolorize: Fast, contrast enhancing, color to grayscale conversion," *Pattern Recognition*, vol. 40, no. 11, pp. 2891–2896, 2007.

[90] K. Smith, P.-E. Landes, J. Thollot, and K. Myszkowski, "Apparent greyscale: A simple and fast conversion to perceptually accurate images and video," in *Computer Graphics Forum*, vol. 27, no. 2, pp. 193–200. Wiley Online Library, 2008.

[91] Y. Kim, C. Jang, J. Demouth, and S. Lee, "Robust color-to-gray via nonlinear global mapping," *ACM Transactions on Graphics (TOG)*, vol. 28, no. 5, p. 161, 2009.

[92] C. Lu, L. Xu, and J. Jia, "Contrast preserving decolorization," in *Computational Photography (ICCP), 2012 IEEE International Conference on*, pp. 1–7. IEEE, 2012.

[93] C. O. Ancuti, C. Ancuti, C. Hermans, and P. Bekaert, "Image and video decolorization by fusion," in *Asian Conference on Computer Vision*, pp. 79–92. Springer, 2010.

[94] C. Ancuti and C. O. Ancuti, "Laplacian-guided image decolorization," in *Image Processing (ICIP), 2016 IEEE International Conference on*, pp. 4107–4111. IEEE, 2016.

[95] Q. Liu, P. X. Liu, W. Xie, Y. Wang, and D. Liang, "Gcsdecolor: gradient correlation similarity for efficient contrast preserving decolorization,"

IEEE Transactions on Image Processing, vol. 24, no. 9, pp. 2889–2904, 2015.

[96] Y. Song and L. Gong, "O (1) contrast preserving decolorization using linear local mapping," in *Wireless Communications & Signal Processing (WCSP), 2016 8th International Conference on*, pp. 1–6. IEEE, 2016.

[97] S. Liu and X. Zhang, "Image decolorization combining local features and exposure features," *IEEE Transactions on Multimedia*, 2019.

[98] B. Cai, X. Xu, and X. Xing, "Perception preserving decolorization," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 2810–2814. IEEE, 2018.

[99] X. Hou, J. Duan, and G. Qiu, "Deep feature consistent deep image transformations: Downscaling, decolorization and hdr tone mapping," *arXiv preprint arXiv:1707.09482*, 2017.

[100] Z. Lin, W. Zhang, and X. Tang, "Learning partial differential equations for computer vision," *Peking Univ., Chin. Univ. of Hong Kong*, 2008.

[101] X. Zhang and S. Liu, "Contrast preserving image decolorization combining global features and local semantic features," *The Visual Computer*, vol. 34, no. 6-8, pp. 1099–1108, 2018.

[102] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition. 2014," *arXiv preprint arXiv:1409.1556*, 2018.

[103] L.-C. Ou, M. R. Luo, A. Woodcock, and A. Wright, "A study of colour emotion and colour preference. part i: Colour emotions for single colours," *Color Research & Application*, vol. 29, no. 3, pp. 232–240, 2004.

[104] M. Ĉadík, "Perceptual evaluation of color-to-grayscale image conversions," in *Computer Graphics Forum*, vol. 27, no. 7, pp. 1745–1754. Wiley Online Library, 2008.

[105] K. Ma, T. Zhao, K. Zeng, and Z. Wang, "Objective quality assessment for color-to-gray image conversion," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4673–4685, 2015.

[106] L. Neumann, M. Čadik, and A. Nemcsics, "An efficient perception-based adaptive color to gray transformation," in *Proceedings of the Third Eurographics conference on Computational Aesthetics in Graphics, Visualization and Imaging*, pp. 73–80. Eurographics Association, 2007.

[107] C. Lu, L. Xu, and J. Jia, "Contrast preserving decolorization with perception-based quality metrics," *International journal of computer vision*, vol. 110, no. 2, pp. 222–239, 2014.

[108] Q. Shan, J. Jia, and M. S. Brown, "Globally optimized linear windowed tone mapping," *IEEE transactions on visualization and computer graphics*, vol. 16, no. 4, pp. 663–675, 2009.

[109] J. Immerkaer, "Fast noise variance estimation," *Computer vision and image understanding*, vol. 64, no. 2, pp. 300–302, 1996.

[110] Y. Kimura and M. Ikebe, "Halo control for lhe based local adaptive tone mapping," in *2015 IEEE International Conference on Image Processing (ICIP)*, pp. 3911–3915. IEEE, 2015.

[111] M. McDonnell, "Box-filtering techniques," *Computer Graphics and Image Processing*, vol. 17, no. 1, pp. 65–70, 1981.

[112] S. Perreault and P. Hébert, "Median filtering in constant time," *IEEE transactions on image processing*, vol. 16, no. 9, pp. 2389–2394, 2007.

[113] P. Ambalathankandy, T. Shimada, S. Takamaeda, M. Motomura, T. Asai, and M. Ikebe, "Analysis of smoothed lhe methods for processing images with optical illusions," in *2018 IEEE Visual Communications and Image Processing (VCIP)*, pp. 1–4. IEEE, 2019.

[114] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *Journal of visual communication and image representation*, vol. 22, no. 4, pp. 297–312, 2011.

[115] P. Mohammadi, A. Ebrahimi-Moghadam, and S. Shirani, "Subjective and objective quality assessment of image: A survey," *arXiv preprint arXiv:1406.7799*, 2014.

[116] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli *et al.*, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[117] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 657–667, 2012.

[118] K. F. Cater, "Detail to attention: Exploiting limits of the human visual system for selective rendering," Ph.D. dissertation, University of Bristol, 2004.

[119] O. H. Schade, "Optical and photoelectric analog of the eye," *JoSA*, vol. 46, no. 9, pp. 721–739, 1956.

[120] R. D. Glickman, A. Adolph, and J. Dowling, "Inner plexiform circuits in the carp retina: effects of cholinergic agonists, gaba, and substance p on the ganglion cells," *Brain research*, vol. 234, no. 1, pp. 81–99, 1982.

[121] P. Ambalathankandy, T. Shimada, S. Takamaeda, M. Motomura, T. Asai, and M. Ikebe, "Analysis of smoothed lhe methods for processing images with optical illusions," in *2018 IEEE Visual Communications and Image Processing (VCIP)*, pp. 1–4. IEEE, 2018.

[122] "World radiography day: Two-thirds of the world's population has no access to diagnostic imaging," shorturl.at/bqzBU, accessed on March 18, 2020.

[123] H. B. HK, K. Dae-shik *et al.*, *Principles and advanced methods in medical imaging and image analysis*. World Scientific, 2008.

[124] M. Uffmann and C. Schaefer-Prokop, "Digital radiography: the balance between image quality and required radiation dose," *European journal of radiology*, vol. 72, no. 2, pp. 202–208, 2009.

[125] P. Irrera, I. Bloch, and M. Delplanque, "A flexible patch based approach for combined denoising and contrast enhancement of digital x-ray images," *Medical image analysis*, vol. 28, pp. 33–45, 2016.

[126] Y. Yang, Y. Que, S. Huang, and P. Lin, "Multiple visual features measurement with gradient domain guided filtering for multisensor image fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 4, pp. 691–703, 2017.

[127] R. Rajendran, S. P. Rao, S. S. Agaian, and K. Panetta, "A versatile edge preserving image enhancement approach for medical images using guided filter," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 002 341–002 346. IEEE, 2016.

[128] R. Panikkath and D. Panikkath, "Mach band sign: an optical illusion." *Proceedings (Baylor University. Medical Center)*, vol. 27, no. 4, pp. 364–365, 2014.

[129] R. H. Daffner, "Visual illusions in the interpretation of the radiographic image," *Current problems in diagnostic radiology*, vol. 18, no. 2, pp. 62–87, 1989.

[130] A. Beghdadi and A. Le Negrate, "Contrast enhancement technique based on local detection of edges," *Computer Vision, Graphics, and Image Processing*, vol. 46, no. 2, pp. 162–174, 1989.

[131] K. Matkovic, L. Neumann, and W. Purgathofer, "A survey of tone mapping techniques," *esc*, vol. 1, p. 1, 1997.

[132] K. Devlin, "A review of tone reproduction techniques," *Computer Science, University of Bristol, Tech. Rep. CSTR-02-005*, 2002.

[133] F. Drago, W. L. Martens, K. Myszkowski, and H.-P. Seidel, *Perceptual evaluation of tone mapping operators with regard to similarity and preference.* Citeseer, 2002.

[134] M. Čadík, M. Wimmer, L. Neumann, and A. Artusi, "Evaluation of hdr tone mapping methods using essential perceptual attributes," *Computers & Graphics*, vol. 32, no. 3, pp. 330–349, 2008.

[135] Y. Rao and L. Chen, "A survey of video enhancement techniques," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 3, no. 1, pp. 71–99, 2012.

[136] G. Eilertsen, J. Unger, R. Wanat, and R. Mantiuk, "Survey and evaluation of tone mapping operators for hdr video," in *ACM SIGGRAPH 2013 Talks*, p. 11. ACM, 2013.

[137] G. Eilertsen, R. K. Mantiuk, and J. Unger, "A comparative review of tone-mapping algorithms for high dynamic range video," in *Computer Graphics Forum*, vol. 36, no. 2, pp. 565–592. Wiley Online Library, 2017.

[138] T. Kalb, L. Kalms, D. Göhringer, C. Pons, F. Marty, A. Muddukrishna, M. Jahre, P. G. Kjeldsberg, B. Ruf, T. Schuchert *et al.*, "Tulipp: Towards ubiquitous low-power image processing platforms," in *2016 International Conference on Embedded Computer Systems: Architectures, Modeling and Simulation (SAMOS)*, pp. 306–311. IEEE, 2016.

[139] D. G. Bailey, *Design for embedded image processing on FPGAs*. John Wiley & Sons, 2011.

[140] N. Goodnight, R. Wang, C. Woolley, and G. Humphreys, "Interactive time-dependent tone mapping using programmable graphics hardware," in *ACM SIGGRAPH 2005 Courses*, p. 180. ACM, 2005.

[141] G. Krawczyk, K. Myszkowski, and H.-P. Seidel, "Perceptual effects in real-time tone mapping," in *Proceedings of the 21st spring conference on Computer graphics*, pp. 195–202. ACM, 2005.

[142] B. Roch, A. Artusi, D. Michael, Y. Chrysanthou, and A. Chalmers, "Interactive local tone mapping operator with the support of graphics hardware," in *Proceedings of the 23rd Spring Conference on Computer Graphics*, pp. 213–218. ACM, 2007.

[143] H. Zhao, X. Jin, and J. Shen, "Real-time tone mapping for high-resolution hdr images," in *2008 International Conference on Cyberworlds*, pp. 256–262. IEEE, 2008.

[144] R. Ureña, J. M. Gómez-López, C. Morillas, F. Pelayo *et al.*, "Real-time tone mapping on gpu and fpga," *EURASIP Journal on Image and Video Processing*, vol. 2012, no. 1, p. 1, 2012.

[145] Q. Tian, J. Duan, and G. Qiu, "Gpu-accelerated local tone-mapping for high dynamic range images," in *2012 19th IEEE International Conference on Image Processing*, pp. 377–380. IEEE, 2012.

[146] M. Akil, T. Grandpierre, and L. Perroton, "Real-time dynamic tone-mapping operator on gpu," *Journal of Real-Time Image Processing*, vol. 7, no. 3, pp. 165–172, 2012.

[147] R. Ureña, C. Morillas, and F. J. Pelayo, "Real-time bio-inspired contrast enhancement on gpu," *Neurocomputing*, vol. 121, pp. 40–52, 2013.

[148] G. Eilertsen, R. K. Mantiuk, and J. Unger, "Real-time noise-aware tone mapping," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 6, pp. 1–15, 2015.

[149] I. R. Khan, S. Rahardja, M. M. Khan, M. M. Movania, and F. Abed, "A tone-mapping technique based on histogram using a sensitivity model of the human visual system," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 4, pp. 3469–3479, 2017.

[150] C.-Y. Tsai and C.-H. Huang, "Real-time implementation of an adaptive simultaneous dynamic range compression and local contrast enhancement algorithm on a gpu," *Journal of Real-Time Image Processing*, vol. 16, no. 2, pp. 321–337, 2019.

[151] F. Hassan and J. E. Carletta, "An fpga-based architecture for a local tone-mapping operator," *Journal of Real-Time Image Processing*, vol. 2, no. 4, pp. 293–308, 2007.

[152] C. Iakovidou, V. Vonikakis, and I. Andreadis, "Fpga implementation of a real-time biologically inspired image enhancement algorithm," *Journal of Real-Time Image Processing*, vol. 3, no. 4, pp. 269–287, 2008.

[153] S. Vakili, D. C. Gil, J. P. Langlois, Y. Savaria, and G. Bois, "Customized embedded processor design for global photographic tone mapping," in

2011 18th IEEE International Conference on Electronics, Circuits, and Systems, pp. 382–385. IEEE, 2011.

[154] C. Kiser, E. Reinhard, M. Tocci, and N. Tocci, "Real time automated tone mapping system for hdr video," in *IEEE International Conference on Image Processing*, vol. 134. IEEE Orlando, FL, 2012.

[155] S. Mann, R. C. H. Lo, K. Ovtcharov, S. Gu, D. Dai, C. Ngan, and T. Ai, "Realtime hdr (high dynamic range) video for eyetap wearable computers, fpga-based seeing aids, and glasseyes (eyetaps)," in *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1–6. IEEE, 2012.

[156] P.-J. Lapray, B. Heyrman, M. Rossé, and D. Ginhac, "A 1.3 megapixel fpga-based smart camera for high dynamic range real time video," in *2013 Seventh International Conference on Distributed Smart Cameras (ICDSC)*, pp. 1–6. IEEE, 2013.

[157] L. Vytla, F. Hassan, and J. E. Carletta, "A real-time implementation of gradient domain high dynamic range compression using a local poisson solver," *Journal of Real-Time Image Processing*, vol. 8, no. 2, pp. 153–167, 2013.

[158] P. M. Cañada, C. Morillas, R. Ureña, J. G. López, and F. J. Pelayo, "Embedded system for contrast enhancement in low-vision," *Journal of Systems Architecture*, vol. 59, no. 1, pp. 30–38, 2013.

[159] Y.-H. Shiau, P.-Y. Chen, H.-Y. Yang, and S.-Y. Li, "A low-cost hardware architecture for illumination adjustment in real-time applications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 934–946, 2014.

[160] S.-A. Li and C.-Y. Tsai, "Low-cost and high-speed hardware implementation of contrast-preserving image dynamic range compression for full-hd video enhancement," *IET Image Processing*, vol. 9, no. 8, pp. 605–614, 2015.

[161] P.-J. Lapray, B. Heyrman, and D. Ginhac, "Hdr-artist: an adaptive real-time smart camera for high dynamic range imaging," *Journal of Real-Time Image Processing*, vol. 12, no. 4, pp. 747–762, 2016.

[162] P. Ambalathankandy, A. Horé, and O. Yadid-Pecht, "An fpga implementation of a tone mapping algorithm with a halo-reducing filter," *Journal of Real-Time Image Processing*, vol. 16, no. 4, pp. 1317–1333, 2019.

[163] U. Shahnovich, A. Hore, and O. Yadid-Pecht, "Hardware implementation of a real-time tone mapping algorithm based on a mantissa-exponent representation," in *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2210–2213. IEEE, 2016.

[164] J. Liu, F. Hassan, and J. E. Carletta, "A study of hardware-friendly methods for gradient domain tone mapping of high dynamic range images," *Journal of Real-Time Image Processing*, vol. 12, no. 1, pp. 165–181, 2016.

[165] V. Popovic, K. Seyid, E. Pignat, Ö. Çogal, and Y. Leblebici, "Multi-camera platform for panoramic real-time hdr video construction and rendering," *Journal of Real-Time Image Processing*, vol. 12, no. 4, pp. 697–708, 2016.

[166] S. Nosko, M. Musil, P. Musil, and P. Zemcik, "True hdr camera with bilateral filter based tone mapping," in *Proceedings of the 33rd Spring Conference on Computer Graphics*, p. 15. ACM, 2017.

[167] P. Zemčík, P. Musil, and M. Musil, "Real-time hdr video processing and compression using an fpga," in *High Dynamic Range Video*, pp. 145–154. Elsevier, 2017.

[168] S. Nosko, M. Musil, P. Zemcik, and R. Juranek, "Color hdr video processing architecture for smart camera," *Journal of Real-Time Image Processing*, pp. 1–12, 2018.

[169] J. Yang, A. Hore, and O. Yadid-Pecht, "Local tone mapping algorithm and hardware implementation," *Electronics Letters*, vol. 54, no. 9, pp. 560–562, 2018.

[170] L. Liu, X. Xiang, Y. Xie, Y. Li, B. Yan, and J. Zhou, "A high throughput and energy-efficient retina-inspired tone mapping processor," in *2019 IEEE 27th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, pp. 321–321. IEEE, 2019.

[171] J. W. Park, H. Lee, B. Kim, D.-G. Kang, S. O. Jin, H. Kim, and H.-J. Lee, "A low-cost and high-throughput fpga implementation of the retinex algorithm for real-time video enhancement," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2019.

[172] C.-T. Chiu, T.-H. Wang, W.-M. Ke, C.-Y. Chuang, J.-S. Huang, W.-S. Wong, R.-S. Tsay, and C.-J. Wu, "Real-time tone-mapping processor with integrated photographic and gradient compression using 0.13 μm technology on an arm soc platform," *Journal of Signal Processing Systems*, vol. 64, no. 1, pp. 93–107, 2011.

[173] A. Punchihewa, T. Hamamoto, and T. Kojima, "From a review of hdr sensing and tone compression to a novel imaging approach," in *2011 Fifth International Conference on Sensing Technology*, pp. 40–46. IEEE, 2011.

[174] G. Sicard, H. Abbas, H. Amhaz, H. Zimouche, R. Rolland, and D. Alleysson, "A cmos hdr imager with an analog local adaptation," in *Int. Image Sensor Workshop (IISW'13)*, pp. 1–4, 2013.

[175] S. Vargas-Sierra, G. Liñán-Cembrano, and Á. Rodríguez-Vázquez, "A 151 db high dynamic range cmos image sensor chip architecture with tone mapping compression embedded in-pixel," *IEEE Sensors Journal*, vol. 15, no. 1, pp. 180–195, 2014.

[176] L. C. Gouveia, W. Mughal, and B. Choubey, "A reconfigurable cmos pixel for applying tone mapping on high dynamic range images," in *2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, pp. 1098–1101. IEEE, 2014.

[177] W. Mughal and B. Choubey, "Fixed pattern noise correction for wide dynamic range cmos image sensor with reinhard tone mapping operator,"

in *2015 Nordic Circuits and Systems Conference (NORCAS): NORCHIP & International Symposium on System-on-Chip (SoC)*, pp. 1–4. IEEE, 2015.

[178] J. Fernández-Berni, R. Carmona-Galán, and Á. Rodríguez-Vázquez, "Single-exposure hdr technique based on tunable balance between local and global adaptation," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 5, pp. 488–492, 2015.

[179] L. Shi, D. Hadlich, C. Soell, T. Ussmueller, and R. Weigel, "A tone mapping algorithm suited for analog-signal real-time image processing," in *2016 12th Conference on Ph. D. Research in Microelectronics and Electronics (PRIME)*, pp. 1–4. IEEE, 2016.

[180] P. Chen, K. Murari, and O. Yadid-Pecht, "Analog current mode implementation of global and local tone mapping algorithm for wdr image display," *Electronic Imaging*, vol. 2016, no. 12, pp. 1–5, 2016.

[181] W. Guicquero, A. Dupret, and P. Vandergheynst, "An algorithm architecture co-design for cmos compressive high dynamic range imaging," *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 190–203, 2016.

[182] N. Banić and S. Lončarić, "Puma: A high-quality retinex-based tone mapping operator," in *Signal Processing Conference (EUSIPCO), 2016 24th European*, pp. 943–947. IEEE, 2016.

[183] A. Horé and O. Yadid-Pecht, "A new filter for reducing halo artifacts in tone mapped images," in *2014 22nd International Conference on Pattern Recognition*, pp. 889–894. IEEE, 2014.

[184] Xilinx power estimator (xpe). [Online]. Available: https://www.xilinx.com/products/technology/power/xpe.html

[185] "Nec instant silicon solution platforms (issp)," http://www.nec.co.jp/press/en/0203/1801.html, accessed on January 4, 2020.

[186] "Fujitsu accelarray structured asic devices," https://www.fujitsu.com/global/about/resources/news/press-releases/2005/0609-01.html, accessed on January 4, 2020.

[187] "Intel easic devices," https://www.intel.com/content/www/us/en/products/programmable/asic/easic-devices.html, accessed on January 2, 2020.

[188] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[189] C. Farabet, B. Martini, P. Akselrod, S. Talay, Y. LeCun, and E. Culurciello, "Hardware accelerated convolutional neural networks for synthetic vision systems," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, pp. 257–260. IEEE, 2010.

[190] C. Zhang, P. Li, G. Sun, Y. Guan, B. Xiao, and J. Cong, "Optimizing fpga-based accelerator design for deep convolutional neural networks," in *Proceedings of the 2015 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, pp. 161–170, 2015.

[191] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, and F. Durand, "Deep bilateral learning for real-time image enhancement," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, pp. 1–12, 2017.

Vita

Prasoon Ambalathankandy

Master of Science in Electrical Engineering

University of Calgary, Calgary AB T2N 1N4, Canada

Publications: related to this thesis

1. Ambalathankandy, P., Ou, Y., Kochiyil, J., Takamaeda, S., Motomura, M., Asai, T., & Ikebe, M. (2019, December). Radiography Contrast Enhancement: Smoothed LHE Filter a Practical Solution for Digital X-Rays with Mach Band. In 2019 Digital Image Computing: Techniques and Applications (DICTA) (pp. 1-8). IEEE.
2. Ambalathankandy, P., Ikebe, M., Yoshida, T., Shimada, T., Takamaeda, S., Motomura, M., & Asai, T. (2019). An adaptive global and local tone mapping algorithm implemented on fpga. IEEE Transactions on Circuits and Systems for Video Technology.
3. Ambalathankandy, P., Shimada, T., Takamaeda, S., Motomura, M., Asai, T., & Ikebe, M. (2018, December). Analysis of smoothed lhe methods for processing images with optical illusions. In 2018 IEEE Visual Communications and Image Processing (VCIP) (pp. 1-4). IEEE.
4. Ambalathankandy, P., Takamaeda, S., Masato, M., Asai, T., Ikebe, M., & Kusano, H. (2018). Real-time HDTV to 4K and 8K-UHD conversions using anti-aliasing based super resolution algorithm on FPGA. Microprocessors and Microsystems, 61, 21-31.

Other publications:

1. Ou, Y., Ambalathankandy, P., Shimada, T., Kamishima, T., & Ikebe, M. (2019, April). Automatic Radiographic Quantification of Joint Space Narrowing Progression

in Rheumatoid Arthritis Using POC. In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (pp. 1183-1187). IEEE.

2. Kanazawa, Yuri, Sayuri Yokoyama, Shota Hiramatsu, Eiichi Sano, Takahiro Ikegami, Yuma Takida, Prasoon Ambalathankandy, Hiroaki Minamide, and Masayuki Ikebe. "Wideband terahertz imaging pixel with a small on-chip antenna in 180 nm CMOS." Japanese Journal of Applied Physics 58, no. SB (2019): SBBL06.
3. Yokoyama, S., Ikebe, M., Kanazawa, Y., Ikegami, T., Ambalathankandy, P., Hiramatsu, S., & Minamide, H. (2019, February). A 32 $\times$ 32-Pixel 0.9 THz Imager with Pixel-Parallel 12b VCO-Based ADC in 0.18μm CMOS. In 2019 IEEE International Solid-State Circuits Conference-(ISSCC) (pp. 108-110). IEEE.
4. Kanazawa, Y., Sano, E., Yokoyama, S., Ambalathankandy, P., Ikebe, M., Hirmatsu, S., & Minamide, H. (2018, September). CMOS Terahertz Imaging Pixel with a Wideband on-Chip Antenna. In 2018 43rd International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz) (pp. 1-2). IEEE.
5. Shimada, T., Ikebe, M., Ambalathankandy, P., Motomura, M., & Asai, T. (2018, April). Sparse Disparity Estimation Using Global Phase Only Correlation for Stereo Matching Acceleration. In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1842-1846). IEEE.

Awards: This thesis was partially supported by following awards, grants and research assistant employments which are gratefully acknowledged.

1. Japan Society for the Promotion of Science (DC2) Research Fellowships for Young Scientists
2. Toshiba Memory (now Kioxia Inc) Grant-in-Aid for the Encouragement of Young Scientist
3. Foreign Student Scholarship (Hokkaido International Exchange & Cooperation Center)
4. IEEE CAS Travel grant for attending VCIP 2018 Taichung, Taiwan
5. Research Assistantship (2018-2019) from Global Station for Big Data and Cybersecurity, Hokkaido University
6. Japan Student Services Organization The Monbukagakusho Honors Scholarship

Contact Address: Laboratory of Functional Communication Devices and Circuits
Research Center for Integrated Quantum Electronics (RCIQE)
Hokkaido University, Sapporo 060-8628 Japan
https://www.rciqe.hokudai.ac.jp/

This thesis was typeset with LaTeX 2_ε[1] by the author.

[1]LaTeX 2_ε is an extension of LaTeX. LaTeX is a collection of macros for TeX. TeX is a trademark of the American Mathematical Society. The style package *graduatethesis* was used.