

DETAIL ENHANCED MULTI-EXPOSURE IMAGE FUSION BASED ON EDGE PRESERVING FILTERS

By

HARBINDER SINGH
(Roll No. 096002)

A THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

IN

ELECTRONICS AND COMMUNICATION ENGINEERING



**JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
WAKNAGHAT**

MARCH, 2015

DECLARATION

I hereby declare that the work reported in the Ph.D. thesis entitled **“Detail Enhanced Multi-Exposure Image Fusion Based on Edge Preserving Filters”** submitted at Jaypee University of Information Technology, Wagnaghat, India, is an authentic record of my work carried out under the supervision of **Dr. Sunil Vidya Bhooshan**. I have not submitted this work elsewhere for any other degree or diploma.

Harbinder Singh

Electronics and Communication Engineering Department

Jaypee University of Information Technology

Wagnaghat, India.

CERTIFICATE

This is to certify that the work reported in the Ph.D. thesis entitled “**Detail Enhanced Multi-Exposure Image Fusion Based on Edge Preserving Filters**”, submitted by **Harbinder Singh** at Jaypee University of Information Technology, Wagnaghat, India, is a bonafide record of his original work carried out under my supervision. This work has not been submitted elsewhere for any other degree or diploma.

Dr. Sunil Vidya Bhooshan

Professor and Head of Department

Electronics and Communication Engineering Department

Jaypee University of Information Technology

Wagnaghat, India.

ACKNOWLEDGMENT

During the course of my PhD, there are many people who deserve a huge amount of thanks and credit for their help. First and foremost, I would like to thank my advisor **Dr. Sunil Vidya Bhooshan** for his insightful guidance and sharing his wealth of experience during the course of this research work. He was very helpful with key suggestions while solving different problems addressed in the thesis. He was highly supportive to provide the solution of problems with meticulously attentive and cheerful approach.

I especially want to express my gratitude to **Dr. Vinay Kumar**, presently at Thapar University Patiala, who has motivated and suggested me to publish my research in top journals. I am extremely fortunate to have the opportunity to work with him and am very grateful for discussions and research advices provided by him.

My sincere thanks also go to **Dr. Shibani Kishen Kak**, our Vice Chancellor for providing constructive criticism, incisive comments and feedback that has enriched this thesis, and **Brig. Balbir Singh** for his encouragement and inspiration.

I am truly indebted to **Dr. T. S. Lamba**, **Dr. Ghanshyam Singh**, and **Dr. S. P. Ghrera** for agreeing to be in my Research Progress Committee and giving useful suggestions in the annual progress seminars. I appreciate the encouragement and valuable suggestions given to me by the committee. I wanted to express my sincere thanks to **Dr. Davinder Singh Saini** who has motivated me to work towards excellence.

I wish to acknowledge the support provided by the librarian **Mr. Shri Ram** for providing me a book written by **Dr. Erik Reinhard et al.** titled “High Dynamic Range Imaging, Acquisition, Display and Image-based Lighting” that helped me to enhance the work and solve the problems addressed in this thesis.

On a more personal note, I devote my thanks to my colleagues **Dr. Vinay Bhatia**, **Dr. Neeru Sharma**, for the encouragement and to the departmental staff of JUIT, for their various helps. Thanks to **Mr. Jacques Joffre**, **Dr. Dani Lischinski**, **Dr. Shree Nayar**, **Dr. Jack Tumblin**, and **Dr. Greg Ward** for the permission to use their image data sets. Also, I want to thank **Dr. Rui Shen** for providing images for analysis purpose.

Lastly and certainly, I must recognize my family. I would like to thank my parents, brother and sister who provided me the inspiration and the reason for my endeavors. No stories of my life would be complete without my lovely wife, **Balwinder Kaur**. To her, I am deeply grateful for love, support, understanding, patience, encouragement, and for making my life truly enjoyable during these years. Thanks to my son, little daughters, nephew and

niece that helped me to relax a bit by playing some funny games and to chill out on occasions in the stressful environment. Finally, I thank God for making all of this possible.

Harbinder Singh

TABLE OF CONTENTS

DECLARATION

CERTIFICATE

ACKNOWLEDGEMENT

LIST OF FIGURES.....	i-xi
LIST OF ABBREVIATIONS.....	xii-xiv
LIST OF NOTATIONS AND VARIABLES.....	xv-xx
ABSTRACT.....	xxi-xxii

1. INTRODUCTION.....	1
1.1 INTRODUCTION.....	1
1.2 PROBLEM STATEMENT.....	6
1.3 MAIN CONTRIBUTION.....	9
1.4 OUTLINE OF THESIS.....	10
2. LITERATURE SURVEY.....	12
2.1 IMAGING TECHNOLOGY.....	12
2.2 HDR IMAGING.....	15
2.3 HDR SOFTWARES.....	18
2.4 HDR SENSOR DESIGN.....	19
2.5 TONE-MAPPING OPERATORS.....	21
2.6 INVERSE TONE-MAPPING.....	24
2.7 IMAGE FUSION.....	25
2.8 EXPOSURE FUSION.....	27
2.9 EDGE-PRESERVING FILTERS AND MULTI-SCALE DECOMPOSITION.....	29
3. ANISOTROPIC DIFFUSION FILTER BASED MULTI-EXPOSURE IMAGE FUSION (ADF).....	34
3.1 INTRODUCTION.....	34
3.2 DATA ACQUISITION AND TWO LAYER DECOMPOSITION.....	35
3.2.1 SCENE DATA ACQUISITION.....	35
3.2.2 EDGE PRESERVING ANISOTROPIC DIFFUSION.....	36

3.3	WEIGHT MAP COMPUTATION: TEXTURE FILTER BASED ON LOCAL RANGE.....	42
3.4	PYRAMID GENERATION AND CONSTRUCTION OF FUSED BASE LAYER ACROSS ALL INPUT BASE LAYERS.....	44
3.4.1	PYRAMID DECOMPOSITION.....	44
3.4.1.1	GAUSSIAN PYRAMID DECOMPOSITION.....	44
3.4.1.2	LAPLACIAN PYRAMID DECOMPOSITION.....	47
3.4.2	BASE LAYERS FUSION.....	50
3.5	CONSTRUCTION OF FUSED DETAIL LAYER, AND DETAIL LAYER ENHANCEMENT.....	52
3.5.1	USER DRIVEN DETAIL LAYER MANIPULATION AND FUSION.....	53
3.5.2	SIGMOID FUNCTION BASED DETAIL LAYER MANIPULATION AND FUSION.....	54
3.6	SUMMARY.....	56
4.	GUIDED FILTER BASED MULTI-EXPOSURE IMAGE FUSION (GFF).....	57
4.1	INTRODUCTION.....	57
4.2	EDGE-PRESERVING GUIDED FILTER.....	58
4.2.1	EDGE-PRESERVING PROPERTY OF GUIDED FILTER.....	61
4.2.2	GUIDED FILTER KERNEL.....	63
4.3	BASE LAYER AND DETAIL LAYER EXTRACTION.....	64
4.4	WEIGH MAP CONSTRUCTION.....	64
4.5	BASE LAYER FUSION BASED ON MULTI-RESOLUTION PYRAMID.....	66
4.6	DETAIL LAYER FUSION AND MANIPULATION.....	67
4.7	SUMMARY.....	68
5.	WEIGHTED LEAST SQUARES FILTER BASED MULTI-EXPOSURE IMAGE FUSION (WLF)	69
5.1	INTRODUCTION.....	69
5.2	EXTRACTION OF COARSER DETAILS AND FINER DETAILS.....	72
5.3	WEIGHT ESTIMATION.....	73
5.4	WLS FILTER BASED WEIGHT REFINEMENT.....	74
5.4.1	REGRESSION ANALYSIS.....	74

5.4.2	EDGE-PRESERVING PROPERTY OF WLS FILTER.....	79
5.4.3	WEIGHT MAP REFINEMENT FOR BASE LAYER.....	80
5.4.4	WEIGHT MAP REFINEMENT FOR DETAIL LAYER.....	81
5.5	WEIGHTED FUSION OF COARSER DETAILS AND FINER DETAILS.....	81
5.6	SUMMARY.....	83
6.	EXPERIMENTAL RESULTS AND ANALYSIS.....	84
6.1	ANISOTROPIC DIFFUSION BASED EXPOSURE FUSION.....	84
6.1.1	COMPARISON WITH OTHER EXPOSURE FUSION AND TONE-MAPPING METHODS.....	84
6.1.2	MULTI-FOCUS IMAGE FUSION.....	87
6.1.3	FLASH AND AMBIENT IMAGE FUSION.....	88
6.1.4	ANALYSIS OF FREE PARAMETERS.....	90
6.2	GUIDED FILTER BASED EXPOSURE FUSION.....	92
6.2.1	COMPARISON WITH OTHER EXPOSURE FUSION.....	92
6.2.2	IMPLEMENTATION AND COMPARISON OF VARIOUS CLASSIC EDGE-PRESERVING FILTERS.....	95
6.2.3	MULTI-FOCUS IMAGE FUSION.....	95
6.2.4	ANALYSIS OF FREE PARAMETERS AND FUSION PERFORMANCE METRICS.....	96
6.3	WLS FILTER BASED EXPOSURE FUSION.....	98
6.3.1	COMPARISON WITH OTHER EXPOSURE FUSION AND TONE-MAPPING METHODS.....	98
6.3.2	MULTI-FOCUS IMAGE FUSION.....	105
6.3.3	ANALYSIS OF FREE PARAMETERS, AND FUSION PERFORMANCE METRICS.....	108
6.3.4	INTERACTIVE TOOL FOR DETAIL ENHANCEMENT.....	112
6.4	OBJECTIVE FUSION QUALITY EVALUATION.....	113
6.5	SUMMARY.....	116
7.	CONCLUSION AND FUTURE SCOPE.....	117
	REFERENCES.....	118
	APPENDIX A.....	133
	APPENDIX B.....	135
	APPENDIX C.....	138

LIST OF FIGURES

Figure Number	Caption	Page Number
1.1:	Shutter efficiency for fully opened shutter.	2
1.2:	Idealized repose of digital camera sensor from highlights, mid-tones and shadows of the scene.	4
1.3:	Multi-exposure shots from unevenly-lit scene depict indoor and outdoor area: images captured at different exposure setting.	5
1.4:	Comparison of high resolution imaging pipeline (i.e. HDR imaging followed by tone-mapping and present exposure fusion approach). The Yellow color depicts HDR and tone-mapping pipeline, and blue color depicts exposure fusion pipeline, which does not require post-processing.	8
2.1:	Tone-mapped HDR images: (a) Global operator and (b) Local operator. Although the input HDR image is same for both cases, non-linear scaling often generates LDR image with complete detail (contrast) than linear scaling.	22
3.1:	Proposed image domain fusion framework. Observation model, illustrating the conceptual framework of the ADF approach. Note that for the concept simplicity; here we have generated BLs and DLs of two input exposures.	36
3.2:	Gradient computation: (a) One dimensional (1-D) grid structure: The signal flow is calculated between two neighboring nodes (solid connection). (b) Two dimensional (2-D) grid structure: The signal flow is calculated between four neighboring nodes (solid connection).	37
3.3:	Two layer decomposition of 1-D signal based on ADI after 5 iterations with $K=30$, $\lambda=1/3$ and $ \eta_s =2$ (left & right). The 1-D input signal (I) is decomposed into two main components: a low frequency BL and a high frequency DL. Notice that the edges are preserved in the diffused image (i.e. BL) and the DL yields	39

	fine details only.	
3.4:	Two layer decomposition of 2-D signal (i.e. “JUIT” image) based on AD (left) after five iterations with $\mathcal{K}=30$, $\lambda=1/7$ and $ \eta_s =4$ (region size of 3×3 pixels). Intensity profiles (right) along a scan lines of 2-D input signal (red), BL (blue) and DL (green). Notice that the strong edges are preserved in the diffused image (i.e. BL) and the DL yields fine details only. Details compressed in the BL are exactly reconstructed in DL.	40
3.5:	Intensity plots (right) along a scan lines of BLs (blue) and DLs (red) obtained with the AD after five iterations with $\mathcal{K}=30$, $\lambda=1/7$ and $ \eta_s =4$ (region size of 3×3 pixels) across all of the input exposures (left). Notice that coarser and finer details are extracted across the visible details adaptively when the scene is captured at different exposure settings.	41
3.6:	Illustration of local range calculation that used as location adaptive weight map. The figure depicts how local range is calculated in the range-filtered image from 3-by-3 neighborhood. Green numbers are the neighbors considered to compute local range for the pixel in interest (i.e. displayed in red number).	43
3.7:	Effect of under exposed, normally exposed, and over exposed regions on local range. Note that local range will be different (i.e. zero for under exposed and overexposed regions) for the same region under different exposure settings. The optimal window size for range calculation is 3-by-3 and the numerical values shown in the box are calculated for small local window.	43
3.8:	Illustration of local range analysis of BLs across the multiple exposures. The local ranges are varying with respect to the different exposure times. ((a), (b)) BLs of underexposed images (top) and their corresponding texture features (bottom); (c) BL of normally exposed image (top) and their corresponding texture features (bottom); (d) BL of overexposed image (top) and their corresponding texture features (bottom). Note that	44

	well-exposed pixels have the brighter texture features (i.e., higher weights) as compared to under-exposed and over-exposed pixels.	
3.9:	First three levels of the Gaussian pyramid and Laplacian pyramid for the "Cameraman" image. The original image, level 0, measures 256 by 256 pixels and each higher level array is roughly half the dimensions of its predecessor. Thus, level 1 measures just 128 by 128 pixels, level 2 measures just 64 by 64 pixels.	45
3.10:	Illustration of possible number of levels in Gaussian pyramid of Cameraman image.	46
3.11:	The equivalent weighting functions $h_l(x)$ for nodes in Gaussian pyramid level l using $\alpha = 0.4$. Note that the resulting equivalent weighting functions closely resemble the Gaussian probability density functions.	47
3.12:	Illustration of Gaussian pyramid expanded to the size of the original image.	48
3.13:	Illustration of Laplacian pyramid expanded to the size of the original image.	49
3.14:	Reconstruction of original image from several spatial frequency bands: first six levels of the Gaussian and Laplacian pyramid. Gaussian images, upper row, were obtained by expanding pyramid arrays through Gaussian interpolation. Each level of the Laplacian pyramid is the difference between the corresponding and next higher levels of the Gaussian pyramid.	49
3.15:	BL manipulation and fusion: illustrating the conceptual framework of BLs fusion in ADF approach. Note that for the concept simplicity, here we have generated the Laplacian pyramid of single BL (LB_1^{adf}) and the Gaussian pyramid of the corresponding GW_1^{adf} . Where $L_2^0, L_3^0, \dots, L_N^0$ are the modified Laplacian pyramids of level zero BLs computed across all of multi-exposures.	53
3.16:	Results of ADF approach; (a) Images representing multiple	54

	exposures; (b) Illustration of ADF result. Note that the fused image contains more details in brightly and poorly illuminated areas with natural contrast. The fine textures on the chair are accurately preserved. (c) The DL enhancement based on sigmoid function across all the inputs reveals more texture details in the fused image and does not depict noticeable artifacts near strong edges.	
3.17:	The affect of weight (i.e. \hat{a}) on sigmoid function and derivative of sigmoid function: (a) The sharpness of the sigmoid in Equation (3.28) varies according to the value of weight. With larger value of \hat{a} , the sigmoid becomes a threshold function. (b) The first derivative of sigmoid function in Equation (3.29) for $\hat{a}=2$ and $\hat{a}=3$.	55
4.1:	GFF based detail-enhanced exposure fusion framework.	58
4.2:	1-D illustration of BL and DL extraction using: (a) BF; and (b) GF.	60
4.3:	Window selection for GF: (a) Illustration of local windows centered at pixel k in the window ω_i ; (b) Example of ideal step edge of a 1-D signal. For a window that exactly centers on the edge, the variables μ and σ are as indicated.	61
4.4:	Edge-preserving property of the GF and the effect of free parameters.	62
4.5:	The DLs computed for different parameters settings used in GF.	63
4.6:	Illustrations of saturation, contrast and exposure measures used for weight map computation in GFF approach.	66
5.1:	Illustration of WLS filter based proposed framework that consists of three principal blocks. (I) BL and DL extraction. The input images are transformed into two scale decomposition. (II) Weight map construction and refinement. (III) Weighted fusion of BLs and DLs. The BLs and DLs across input image series are fused using simple weighted average approach.	70
5.2:	Illustration of uncorrelated variances. (a) homoscedasticity: there is certain correlation between variances across individuals	76

	and (b) heteroscedasticity: there is no correlation between variances across individuals.	
5.3:	Illustration of least squares line of best fit. (a) least squares line of best fit for four data points (1,5), (2,6), (3,7), (4,9); (b) least squares line of best fit for randomly generated data points with normally distributed noise; (c) WLS line of best fit for randomly generated data points with normally distributed noise; and (d) WLS line of best fit for randomly generated data points with normally distributed noise.	78
5.4:	Illustration of weight map refinement and fused image using WLS filter. (a-c) Input exposures; (d-f) weight maps computed for BLs across input exposures; (g-i) weight maps computed for DLs across input exposures; (j) WLF results; and (k) results proposed by Mertens et al. [113]	82
6.1:	Images representing multiple exposures of an outdoor scene depicting highlights from natural sun light and shadows; (b) ADF approach fuses the multiple exposures to obtain high quality image. Note the fused image yields more texture details and natural contrast without the introduction of artifacts. (c) Auto-exposure (d) Mertens et al. [113] (e) Shen et al. [116]. Input images courtesy of Shree Nayar.	85
6.2:	Images representing multiple exposures of an indoor and outdoor scene depicting sunlit details and shadows; (b) ADF approach fuses the multiple exposures to obtain high quality image. Note the fused image yields more texture details and natural contrast without the introduction of artifacts. (c) Auto-exposure (d) Mertens et al. [113] (e) Zhang and Cham [120]. Input images courtesy of Shree Nayar.	85
6.3:	“House” image: Comparison ADF results to other recent exposure fusion technique. (a) Results of ADF method; (b) Mertens et al. [113]; (c) Shen et al. [116]. Note that ADF yields enhanced texture and edge features. Input image sequence courtesy of Tom Mertens.	86

- 6.4:** “Belgium House” image: (a) Series of multiple exposures depicting both indoor and outdoor areas. The exposure value is varying from (1/1000 of a second) to (1/4 of a second). Comparison results to other popular tone-mapping techniques. 86
 (b) Results of ADF method, window size = 3x3; (c) Mertens et al. [113] (d) Shen et al. [116] (e) Results of Ward Larson et al [75]; (f) Results of LCIS method [185]. Note that ADF yields combined features that can only be recorded using different exposures. Input images courtesy of Dani Lischinski.
- 6.5:** “Book” image. (a) Two partially focused images, (focused on different targets), and (b) image generated by the ADF approach, which illustrate that the fused image extracts more color and texture details from the original input images. (Input sequence courtesy of ADu and Wang [110]) 87
- 6.6:** “Table” image. (a,b) two partially focused images, (focused on different targets), (c) image generated by the ADF approach, which illustrate that the fused image extracts more information from the original images and (d) Hod’akov’a et al. [185]. 88
- 6.7:** “Clock” Image. (a,b) two partially focused images, (focused on different targets), (c) image generated by the ADF approach, which illustrate that the fused image extracts more information from the original images and (d) ADu and Wang [110]. (Input sequence courtesy of ADu and Wang [110]). 89
- 6.8:** (a,b) input images photographed with and without flash; (c) Enhanced fused image by ADF algorithm which maintains the warm appearance, and the sharp details after removing strong highlight and (d) Mertens et al. [113]. Images taken from Agrawal et al. [98] 89
- 6.9:** Analysis of number of iterations used for BL computation. MSE is defined as the relative difference from the results generated with $t_a = 1$. Maximum quality score and entropy is only observed when $t_a = 1$. It is observed that MSE and 91

	computational time increases as t_a increases. (a) Effectiveness of t_a on quality score; (b) Entropy; (c) computational time; and (d) Error introduced.	
6.10:	“House” image. The free parameter t_a in Equation (3.11) is used to controls detail enhancement. It is found that $t_a=1$ is sufficient for fine details extraction and gives better results for most cases. Higher value of t_a brings in artifacts near strong edges. (a); $t_a=1$ (b) $t_a=2$; (c) $t_a=3$; (d) $t_a=4$; and (e) $t_a=5$.	91
6.11:	“Igloo” image. The free parameter t_a in Equation (3.11) is used to controls detail enhancement. It is found that $t_a=1$ is sufficient for fine details extraction and gives better results for most cases. Higher value of t_a brings in more details while introducing artifacts near strong edges. (a); $t_a=1$ (b) $t_a=2$; (c) $t_a=3$; (d) $t_a=4$; and (e) $t_a=5$.	92
6.12:	“House” image. The free parameter θ in Equation (3.28) is used to control sharpening. It is found that $\theta=.002$ gives better results for most cases. Higher value of θ brings in more details in highly illuminated areas. (a) $\theta=.002$; (b) $\theta=.003$; (c) $\theta=.004$; and (d) $\theta=.005$.	92
6.13:	Comparison results to other recent exposure fusion techniques. (a) Mertens et al. [113], (b) Burt et al. [89], (c) Shen et al. [116] and (d) Results of GFF method. Note that GFF method yields enhanced texture and edge features. Input image sequence is courtesy of Tom Mertens.	93
6.14:	GFF results for “Hermes” image sequence: (a, b) Different multi-exposure sequences (two input exposures), (c) Fused image by GFF.	93
6.15:	GFF results for “Chairs” image series. (a) Different multi-exposure sequences (five exposures), (b) Fused image by GFF.	93
6.16:	GFF results for Syn image series. (a) Different multi-exposure sequences (seven input exposures), (b) Fused image by GFF.	94
6.17:	Color-coded map comparison (dark blue color indicates over-exposed region and pure white indicates under-exposed region):	95

	Comparison of GFF results to other classic edge preserving and exposure fusion techniques. (a) Source exposure 1; (b) Source exposure 2; (c) BF [167]; (d) WLS [133] (e) Mertens et al. [113]; (f) Results of GFF method. Note that proposed method yields enhanced texture and edge features. Input image sequence courtesy of Tom Mertens.	
6.18:	“Clock” image: (a, b) two partially focused input images (focused on different targets), (c) Image generated by the GFF approach, which illustrates that the fused image extracts more information from input images., and (d) ADu and Wang [110]. Input sequence is courtesy of Adu and Wang.	96
6.19:	Analysis of different free parameters used in the GFF algorithm. Maximum quality score and entropy are only observed when $\varepsilon = 0.01$, $\gamma = 5$, and $r = 2$ (which are set as default parameters). It is observed that VIFF increases as ε , γ , and r increases but the larger values are responsible for over-enhancement. (a) Effectiveness of ε on metrics, (b) Effectiveness of r on metrics, (c) Effectiveness of γ on metrics.	97
6.20:	Visual Inspection: The affect of free parameter r on detail enhancement. It is found that $r=2$ is sufficient for fine details extraction and gives better results for most cases. Higher value of r brings in artifacts near strong edges. (a); $r=1$ (b) $r=3$; and (c) $r=6$	98
6.21:	Results of WLF approach for different input multi-exposure sequences. (a) “Carnival” (top: three input exposures, bottom: fusion results), (b) “House” (top: four input exposures, bottom: fusion results) and (c) “Bellavita” (left: two input exposures, bottom: fusion results). Input image sequences are courtesy of HDRsoft.com, Tom Mertens.	100
6.22:	“Lizard” image: Comparison of WLF results with popular exposure fusion approaches. (a, b) Source images, (c) Results of WLF method. Tree leaves and wall texture appear over-exposed and blurry in (d) Ketan & Chaudhuri [118], (e) Raman &	101

	Chaudhuri [111] and (f) Mertens et al. [113]. In the proposed results, it becomes almost possible to preserve texture and strong edge features simultaneously. Moreover, by applying the WLF, the fine textures are accurately enhanced. Input image sequence is Eric Reinhard, University of Bristol.	
6.23:	“National Cathedral”: Comparison WLF results with popular exposure fusion and tone-mapping methods. (a) Results of WLF method, (b) Mertens et al. [113], (c) iCAM06 [190], (d) WLS. [133] and (e) GRW [116]. Note that the WLF method yields enhanced texture and edge features with better color appearance. Input image sequence is courtesy of Max Lyons.	101
6.24:	Comparison of WLF results with iCAM06, EF and WLS on the “National Cathedral” sequence using DRIVDP [188]. (a), (b) The two source images give good exposures for the paintings on window glass and lamp, and the wall, respectively. In a distortion map, green, blue, and gray pixels indicate visible contrast loss, amplification, and no distortion, respectively. Proposed results are more effective in preserving local details and colors than the others. For the paintings on window glass and lamp, and the wall, proposed results (see c-g) depict the least distortion, followed by iCAM06 (see h-l), Mertens et al. (see m-q), and WLS (see r-v).	102
6.25:	Multi-focus image fusion results: (a, b, d, e) demonstrates the effect of selective focus to capture more details from a particular part of a scene and (c, f) images generated by the WLF approach. Note how in the fused image the all objects appear all-in-focus and enhance the color and texture details present in the foreground and background extracted from the original input images (input sequence is courtesy of Adu and Wang [110]).	106
6.26:	Multi-focus image fusion results: (a, b, d, e) demonstrates the effect of selective focus to capture more details from a particular part of a scene and (c, f) images generated by the	107

- WLF approach. These results demonstrate that WLF approach has helped to handle typical situations of foreground and background present in the scene (input sequence is courtesy Slavica Savić).
- 6.27:** Multi-focus image fusion results: (a, b, d, e) demonstrates the effect of selective focus to capture more details from a particular part of a scene and (c, f) images generated by the WLF approach. These results demonstrate that WLF approach has helped to handle typical situations of foreground and background present in the scene (input sequence is courtesy Slavica Savić). 107
- 6.28:** Comparison of “Clock” image results with recently proposed conventional multi-focus image fusion methods. (a) WLF results, (b) Adu and Wang [110], (c) DWT [192] and (d) Tian et al. [193]. It has been found that the WLF approach helps the viewer to observe enhanced texture and edge features simultaneously without depicting visible artifacts. Input image sequence is courtesy of Adu and Wang. 108
- 6.29:** Sensitivity analysis of free parameter (ψ) and free parameter (\hat{a}) which respectively control the smoothness in weight maps computed for BLs and DLs. (a) Effectiveness of ψ on $Q^{AB/F}$, (b) Effectiveness of ψ on VIFF, (c) Effectiveness of \hat{a} on $Q^{AB/F}$ and (d) Effectiveness of \hat{a} on VIFF. 109
- 6.30:** Visual inspection of fine detail enhancement in WLF approach in typical lighting situations: The free parameter \hat{a} in Equation (5.38) controls detail enhancement and sharpening. It is found that variation of “ \hat{a} ” between 3 and 6 is sufficient for fine details extraction and give better results for various typical situations. By choosing higher value of \hat{a} , the texture details are accurately enhanced. It also shows that selection of much higher value for detail enhancement does not introduce artifacts near object boundaries due to detail-enhancement. (a) $\hat{a} = 1$, (b) $\hat{a} = 6$, (c) $\hat{a} = 12$. 110

6.31:	Sensitivity analysis of errors introduced by the free parameter \hat{a} in WLF approach. The errors increases dramatically as \hat{a} becomes too large, but increases slowly when $\hat{a} \leq 6$. It is observed in Figure (b) and Figure (c) that error increases as \hat{a} increases but still the deviation range is less. (a) Effectiveness of \hat{a} on RMSE, (b) Effectiveness of \hat{a} on LMSE, (c) Effectiveness of \hat{a} on NAE, and (d) Effectiveness of \hat{a} on PSNR.	111
6.32:	Visual inspection of contrast and detail enhancement in WLF approach in multi-focus image fusion. The free parameter ψ in Equation (5.32) and \hat{a} in Equation (5.38) controls detail enhancement and sharpening respectively. WLF is able to achieve aggressive detail enhancement and exaggeration, while avoiding artifacts. “Clock” image: (a) $\psi = 0.5$, $\hat{a} = 7$, (b) $\psi = 0.5$, $\hat{a} = 12$, (c) $\psi = 0.5$, $\hat{a} = 17$. “Leaves” image: (d) $\psi = 0.5$, $\hat{a} = 7$, (e) $\psi = 0.5$, $\hat{a} = 12$, (f) $\psi = 0.5$, $\hat{a} = 17$.	113

LIST OF TABLES

Table Number	Caption	Page Number
I	Quantitative assessment of different multi-exposure image fusion and tone mapping methods.	114
II	Quantitative assessment of different multi-focus image fusion methods.	115

LIST OF ABBREVIATIONS

AD	Anisotropic Diffusion
ADC	Analog to Digital Conversion
ADF	Anisotropic Diffusion Filter Based Multi-Exposure Image Fusion
AEB	Auto Exposure Bracketing
BF	Bilateral Filter
BL	Base Layer
BMP	Bitmap
CCD	Charge Coupled Device
CFA	Colour Filter Array
CIS	Complementary Metal Oxide Semiconductor Image Sensor
CMOS	Complementary Metal Oxide Semiconductor
CRF	Camera Response Function
CRT	Cathode Ray Tube
CT	Computed Tomography
DL	Detail Layer
DLT	Direct Linear Transform
DOF	Depth of Field
DR	Dynamic Range
DRIVDP	Dynamic Range Independent Visible Difference Predictor
DSLR	Digital Single Lens Reflex
DSP	Digital Signal Processor
DWT	Discrete Wavelet Transform
EV	Exposure Value
FPN	Fixed-Pattern-Noise
GF	Guided Filter
GFF	Guided Filter Based Multi-Exposure Image Fusion
GPU	Graphics Processing Units
GRW	Generalized Random Walks
GSM	Gaussian Scale Mixture
HDR	High Dynamic Range
HVS	Human Visual System

IEEE	Institute of Electrical and Electronics Engineering
ILM	Industrial Light and Magic
ISO	International Organization for Standardization
iTMO	Inverse Tone-Mapping Operator
JPEG	Joint Photographic Experts Group
LCD	Liquid Crystal Diode
LCIS	Low Curvature Image Simplifier
LDR	Low Dynamic Range
LED	Light Emitting Diode
LMSE	Laplacian Mean Squared Error
LUT	Look-up Table
MATLAB	Matrix Laboratory
MF	Median Filter
MI	Mutual Information
MLIC	Multi-Light Image Collections
MLS-SVM	Mapped Least Squares Support Vector Machine
MRI	Medical Resonance Image
MSD	Multi-Scale Decomposition
MSE	Mean Square Error
MTB	Median Threshold Bitmaps
NAE	Normalized Absolute Error
ND	Neutral Density
NSCT	Non-Subsampled Contourlet Transform
NTF	Non-linear Translation-variant Filter
OCR	Optical Character Recognition
OSF	Order Statistics Filters
PDE	Partial Differential Equation
PFM	Portable Float Maps
PSNR	Peak Signal to Noise Ratio
QMF	Quadrature Mirror Filters
RANSAC	Random Sample Consensus
RAM	Random Access Memory
RGB	Red Green Blue
RMSE	Root Mean Squared Error

SDR	Standard Dynamic Range
SF	Spatial Frequency
SIFT	Scale-Invariant Feature Transform
SNR	Signal-to-Noise-Ratio
TIFF	Tagged Image File Format
TV	Total Variation
VIFF	Visual Information Fidelity for Fusion
WDR	Wide Dynamic Range
WLF	Weighted Least Squares Filter Based Multi-Exposure Image Fusion
WLS	Weighted Least Squares

LIST OF NOTATIONS AND VARIABLES

E_X	Exposure
I_L	Illuminance incident on the surface of the photographic material (in lux)
t_e	Exposure time as a function of the time that the shutter takes to open (t_{e1}), the time during which the shutter is fully opened (t_{e2}), and the time that the shutter is closing t_{e3}
h	Height of trapezoid used to determine efficiency
η_e	Shutter Efficiency
L_L	Scene luminance as measured at the principle plane of the lens system (in cd/m^2)
θ_L	Angle from a point on the surface of photographic material to the center point of the lens
V_L	Lens barrel vignetting factor
τ_L	Lens transmittance factor
I_{flare}	Illuminance resultant from flare
u_L	Subject to lens, distance (in mm)
F_L	Focal length (in mm)
f_L	Lens aperture (f-number also called focal ratio, f-ratio, f-stop)
D_L	Diameter of the entrance pupil (effective aperture)
EV	Exposure value
DR	Dynamic range
I_{max}	Maximum current handled by the sensor before getting saturated
I_{min}	Minimum current handled by the sensor above the noise floor
Q_{sat}	Charge produced at the saturation level
$f[\cdot]$	Functional of the current $I(t)$ over the integration time $0 \leq t \leq t_{int}$.
W	Weight map
N	Number of input multi-exposure images
k	k represents the k_{th} image from different input multi-exposure images (N)
I_k	Input exposure (subscript k represents k_{th} image from the input data)
I_F	Fused image
(x,y)	Specifies spatial position of pixel
BL_k	BL computed for k_{th} image using edge-preserving filters

DL_k	DL computed for k_{th} image by taking the difference of input exposure and BL
W_k^B	Weight maps computed for k_{th} BL
W_k^D	Weight maps computed for k_{th} DL
$\mathcal{A}(\bullet)$	AD function applied on input image
BL_F	Fused BL
DL_F	Fused DL
BL_k^{adf}	BL computed for k_{th} image using AD
DL_k^{adf}	DL computed for k_{th} image using AD
BL_F^{adf}	Fused BL using ADF approach
DL_F^{adf}	Fused DL using ADF approach
I_F^{adf}	Detail enhanced fused image using ADF approach
∇	Gradient operator
$\ \nabla I\ $	Magnitude of the gradient of image intensity
t_a	Process ordering time parameter in AD
$g(\bullet)$	An edge-stopping function or conduction coefficient in AD
K	Scale parameter (i.e. constant) in AD
\hat{L}, \hat{R}	Left and right neighbor considered in 1-D structure for gradient computation in AD
$\hat{N}, \hat{S}, \hat{E}, \hat{W}$	North, South, East and West neighbor considered in 2-D structure for gradient computation in AD
λ	Scalar that determines the rate of diffusion
$\eta_{x,y}$	Represents the spatial neighborhoods of current sample position x,y
$ \eta_{x,y} $	Number of neighbors
\mathfrak{p}	Determines the position of neighboring pixel for current pixel at x,y location
$R_{x,y,k}^{adf}$	Local range calculated for k_{th} BL at x,y location in 3×3 spatial window
$W_{x,y,k}^{adf}$	Normalized weight map for k_{th} BL in the ADF approach
L_{max}	Maximum pixel values of the neighboring pixels within a 3×3 square window
L_{min}	Minimum pixels values of the neighboring pixels within a 3×3 square window
GB_k^{adf}	Gaussian pyramid of BL in the ADF approach
LB_k^{adf}	Laplacian pyramid of BL in the ADF approach
GW_k^{adf}	Gaussian pyramid of weight map used for BL in the ADF approach

$L_1^{adf,l}$	Modified Laplacian pyramid for l_{th} level for first image in the ADF approach
$L_F^{adf,l}$	Fused Laplacian pyramid for l_{th} level in the ADF approach
$\hat{w}(\bullet)$	Equivalent weighting function
α	A constant used in pyramid decomposition resulting equivalent weighting functions closely resemble the Gaussian probability density functions
L	Modified Laplacian pyramid
l	Represents level of pyramid
α_1	User defined weighting factor (constant) to control the contribution of texture details directly from the input DLs
α_2	User defined weighting factor (constant) in sigmoid function based manipulation to control the contribution of texture details directly from the input DLs
\mathcal{R}	A real number
$t_s \in \mathcal{R}$	Independent variable of the sigmoid function
$\hat{a} \in \mathcal{R}$	Weight parameter of the sigmoid function
θ	A fixed threshold to further control the sharpness of sigmoid function
$\mathcal{S}(\bullet)$	Sigmoid function
q	Filter output from GF
I^G	Guidance image
σ_s, σ_t	Standard deviation parameters used in BF
p	Input image to GF
W	Filter kernel that is a function of guidance image I^G , and independent of input image p
W^{bf}	Weight function used in BF
r	Variable used to define window radius
ω_{ℓ}	Window of a radius $(2r+1) \times (2r+1)$ centered at the pixel ℓ
a_{ℓ}, b_{ℓ}	The linear coefficients in GF
$E(\bullet)$	Cost function in GF
ε	Regularization term on linear coefficient a_{ℓ} for numerical stability
$\mu_{\ell}, \sigma_{\ell}^2$	Mean and variance of I^G in ω_{ℓ} , respectively
SA_k	Saturation measure computed for k_{th} image
CO_k	Contrast measure computed for k_{th} image
EX_k	Exposure measure computed for k_{th} image

BL_k^{gff}	BL computed for k_{th} image using GF
DL_k^{gff}	DL computed for k_{th} image using GF
BL_F^{gff}	Fused BL using GFF approach
DL_F^{gff}	Fused DL using GFF approach
I_F^{gff}	Detail enhanced fused image using GFF approach
GB_k^{gff}	Gaussian pyramid of BL in the GFF approach
LB_k^{gff}	Laplacian pyramid of BL in the GFF approach
GW_k^{gff}	Gaussian pyramid of weight map used for BL fusion in the GFF approach
$L_F^{gff,l}$	Fused Laplacian pyramid for l_{th} level in the GFF approach
γ	User defined parameter to control amplification of texture details in the fused image
$f(\bullet)$	Nonlinear function to achieve detail enhancement while reducing noise and artifacts near strong edges in GFF
\mathbf{r}_i	Residual for the i_{th} data point in regression analysis
y_i and \hat{y}_i	Observed value and the fitted value in regression analysis, respectively
$\beta: \beta_1$ and β_2	Unknowns in regression analysis
S	Cost function used in regression analysis
α_3	User defined parameter to control contrast in the fused DL
U	Gaussian low pass kernel
r_U	Kernel size
σ_U	Standard deviation
FR_k	Combined image features for k_{th} image in WLF approach
SM_k	Saliency maps for k_{th} image in WLF approach
SM_k^n	Saliency map value for pixel n in the k_{th} image in WLF approach
BL_k^{wlf}	BL computed for k_{th} image in the WLF approach
DL_k^{wlf}	DL computed for k_{th} image in the WLF approach
BL_F^{wlf}	Fused BL using WLF approach
DL_F^{wlf}	Fused DL using WLF approach
W_k^{wlf}	Weight map computed in the WLF approach
$W_{BL_k}^{wlf}$	Refined weight map computed for k_{th} BL in the WLF approach

$W_{DL_k}^{wlf}$	Refined weight map computed for k_{th} DL in the WLF approach
I_F^{wlf}	Detail enhanced fused image in the WLF approach
$\#(\bullet)$	WLS filtering operator (function)
\hat{a}_k, \hat{b}_k	The linear coefficients in WLS filter
\hat{q}	Filtered output of WLS filter
v	Input image to WLS filtering operator
w	New image (output image from WLS filtering operator)
$\psi, \alpha_4, \varepsilon'$	Decide the degree of smoothness, sensitivity to the gradients and small constant of the WLS filter, respectively
V_x, V_y	Diagonal matrices containing the smoothness weights $q_x(v)$ and $q_y(v)$, respectively
O_x, O_y	Discrete differentiation operators (forward difference operators)
O_x^T, O_y^T	Backward difference operators
I_{dent}	Identity matrix
L_v	Five-point spatially inhomogeneous Laplacian matrix
ℓ	Log-luminance channel of the input image v

QUALITY METRICS

$Q^{AB/F}$	Fusion quality metric that evaluates the amount of edge information transferred from input images to the fused image from two input images A and B , and a resulting fused image F
$e(m,n)$	Edge strength for each pixel
$\beta(m,n)$	Orientation information for each pixel
$sb_A^x(m,n)$	Horizontal sobel template cantered on pixel $P_A(m,n)$
$sb_A^y(m,n)$	Vertical sobel template cantered on pixel $P_A(m,n)$
L	A constant
w^A, w^B	Perceptual factors which are defined by edge strength $e(m,n)$ for input image A and image B
Q_{MI}	Fusion quality metric that measures how well the original information from source images is preserved in the fused image
$H(A)$	Marginal entropy of A
$H(B)$	Marginal entropy of B

$H(F)$	Marginal entropy of F
$MI(A,F)$	Mutual information between the source image A and the fused image F
$MI(B,F)$	Mutual information between the source image B and the fused image F
Q_{SF}	Quality index that indicates the overall active level (spatial frequency (SF)) in an image by measuring row frequency (RF) and column frequency (CF)
P_i^{res}	Pixel considered in a resulting image for error calculation
P_i^{ref}	Pixel considered in a reference image for error calculation
$E_e(P_i^{res})$	Error calculated using Euclidean distance between a pixel in a resulting image (P_i^{res}) and the corresponding pixel in the reference image (P_i^{ref})
$L(I_F(m,n))$	Laplacian operator applied on the fused image (I_F)
MAX_{I_F}	Maximum possible pixel value of the fused image

ABSTRACT

Recent computational photography techniques play a significant role to overcome the limitation of standard digital cameras for handling wide dynamic range of real-world scenes contain brightly and poorly illuminated areas. In many of such techniques [1,2,3], it is often desirable to fuse details from images captured at different exposure settings, while avoiding visual artifacts. One such technique is High Dynamic Range (HDR) imaging that provides a solution to recover radiance maps from photographs taken with conventional imaging equipment. The process of HDR image composition needs the knowledge of exposure times and Camera Response Function (CRF), which is required to linearize the image data before combining Low Dynamic Range (LDR) exposures into HDR image. One of the long-standing challenges in HDR imaging technology is the limited Dynamic Range (DR) of conventional display devices and printing technology. Due to which these devices are unable to reproduce full DR. Although DR can be reduced by using a tone-mapping, but this comes at an unavoidable trade-off with increased computational cost. Therefore, it is desirable to maximize information content of the synthesized scene from a set of multi-exposure images without computing HDR radiance map and tone-mapping.

This research attempts to develop a novel detail enhanced multi-exposure image fusion approach based on texture features, which exploits the edge preserving and intra-region smoothing property of nonlinear diffusion filters based on Partial Differential Equations (PDE). With the captured multi-exposure image series, we first decompose images into Base Layers (BLs) and Detail Layers (DLs) to extract sharp details and fine details, respectively. The magnitude of the gradient of the image intensity is utilized to encourage smoothness at homogeneous regions in preference to inhomogeneous regions. In the next step texture features of the BL to generate a decision mask (i.e., local range) have been considered that guide the fusion of BLs in multi-resolution fashion. Finally, well-exposed fused image is obtained that combines fused BL and the DL at each scale across all the input exposures. The combination of edge-preserving filters with Laplacian pyramid is shown to lead to texture detail enhancement in the fused image.

Furthermore, Non-linear Translation-Variant filter (NTF) is employed for BL and DL decomposition that has better response near strong edges. The texture details are then added to the fused BL to reconstruct a detail enhanced LDR version of the image. This leads to an increased robustness of the texture details while at the same time avoiding gradient reversal artifacts near strong edges that may appear in fused image after DL enhancement.

Finally, we propose a novel technique for exposure fusion in which Weighted Least Squares (WLS) optimization framework is utilized for weight map refinement of BLs and DLs, which lead to a new simple weighted average fusion framework. Computationally simple texture features (i.e. DL) and color saturation measure are preferred for quickly generating weight maps to control the contribution from an input set of multi-exposure images. Instead of employing intermediate HDR reconstruction and tone-mapping steps, well-exposed fused image is generated for displaying on conventional display devices. Simulation results are compared with a number of existing single resolution and multi-resolution techniques to show the benefits of the proposed scheme for the variety of cases.

Moreover, the approaches proposed in this thesis are effective for blending flash and no-flash image pair, and multi-focus images, that is, input images photographed with and without flash, and images focused on different targets, respectively. A further advantage of the present technique is that it is well suited for detail enhancement in the fused image.

CHAPTER - I

1.1 INTRODUCTION

The research presented in this thesis focuses on solving the problem of preserving and enhancing details present in the poorly illuminated and brightly illuminated regions of the real world scene. It is impossible to capture the entire DR of the real world scene with single exposure. This is due to the limited capabilities of Charge Coupled Device (CCD) or Complementary Metal Oxide Semiconductor (CMOS) sensor chip. Human eye is sensitive to relative rather than absolute luminance values and can observe both indoor and outdoor details simultaneously while digital camera cannot record indoor and outdoor luminance variations in single snapshot. This is because the eye adapts locally as we scan the different regions of the scene and can adapt 10 orders of magnitude of intensity variations in the scene [1], while standard digital cameras are unable to record luminance variation in the entire scene. To circumvent this problem, modern digital photography offers the concept of exposure time variation to capture details in very dark or extremely bright regions, which control the amount of light allowed to fall on the sensor. Currently, there are many applications that involve variable exposure photography to determine the details to be captured optimally in the photographed scene. The intention of exposure setting determination is to control charge capacity of the CCD or CMOS.

Digital cameras have the shutter speed, ISO value and aperture setting to regulate the amount of light captured by the sensors. Shutter speed is directly proportional to ISO setting. Shutter speed should be sufficiently fast to prevent camera shake. By increasing the ISO setting, the shutter speed increases, and the exposure time decreases as shutter speed increases. Higher value of ISO setting yields amplified signal from the sensor, but will also amplify the noise. This will distort the captured image signal and produce noisy image. Most cameras have automatic ISO option, which will increase or decrease ISO to suit the shooting condition. Remember, lower ISO setting and slower shutter speed is used to get noise-free image, if scene is photographed with the help of tripod than holding the camera. The aperture setting controls the Depth of Field (DOF) in the captured image. Large aperture value throws a background out of focus and small aperture increase DOF. In HDR photography [2,3], the aperture setting is kept constant to maintain the same DOF among different exposures.

To provide complete information about specific exposure you need to know the shutter speed, ISO and aperture that control the amount of light falling on the sensor. Therefore, Exposure (E_X) is computed by multiplying the illuminance (I_L) by the time (t_e) of exposure. In the case of a camera, shutter efficiency (η_e) must be considered for determining E_X . Thus, exposure can be expressed as:

$$E_X = I_L \eta_e t_e \quad (1.1)$$

where η_e is the shutter efficiency, which is defined as the ratio of a shutter's actual performance with the performance of an ideal shutter that opens and closes instantaneously. The η_e for fully opened shutter (see Figure 1.1 for graphical representation) can be expressed as [4]:

$$\eta_e = \frac{h(0.5t_{e1} + t_{e2} + 0.5t_{e3})}{ht_e} = \frac{0.5t_{e1} + t_{e2} + 0.5t_{e3}}{t_e} \quad (1.2)$$

where t_{e1} is the time that the shutter takes to open in which the measured light will increase gradually, t_{e2} is the time during which the shutter is fully opened in which the measured light will remain constant, t_{e3} is the time that the shutter is closing in which the measured light will decrease gradually, and $t_e = t_{e1} + t_{e2} + t_{e3}$ is the total exposure time.

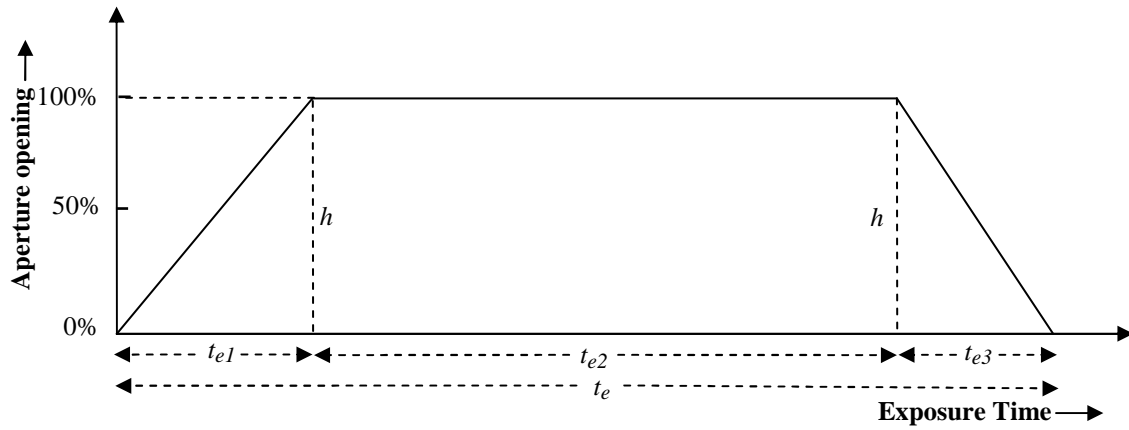


Figure 1.1: Shutter efficiency for fully opened shutter.

Considering numerous factors, the illuminance (I_L) can be mathematically expressed as [4]:

$$I_L = \frac{L_L \cdot \pi \left(\frac{u_L - F_L}{u_L} \right)^2 \cdot \cos^4 \theta_L \cdot V_L(\theta_L) \cdot \tau_L}{4f_L^2} + I_{flare} \quad (1.3)$$

where, I_L is illuminance incident on the surface of the photographic material (in lux), L_L is scene luminance as measured at the principle plane of the lens system (in cd/m^2), θ_L is angle

from a point on the surface of photographic material to the center point of the lens, V_L is lens barrel vignetting factor, τ_L is lens transmittance factor, I_{flare} is illuminance resultant from flare, u_L is subject to lens, distance (in mm), F_L is focal length (in mm) and f_L is the lens aperture (f-number also called focal ratio, f-ratio, f-stop).

f_L is given by the relation:

$$f_L = \frac{F_L}{D_L}, \quad \text{where } D_L \text{ is the diameter of the entrance pupil (effective aperture)} \quad (1.4)$$

For example, if a lens's focal length (F_L) is 20 mm and its entrance pupil diameter (D_L) is 4 mm, the f-number (f_L) is 5, and the aperture size would be expressed as $f_L/5$. f_L is a dimensionless number that is a quantitative measure of lens speed. If we keep increasing the f-number by a factor of 2 then the relative brightness will be decreased by a factor of four. To maintain the same photographic exposure when doubling the f-number, the exposure time would need to be four times as long.

Recently, light metering (spot meter) technique [5] has been developed to measure lighting distribution in the scene. To get right exposure, Digital Single Lens Reflex Camera (DSLR) camera decides shutter speed, aperture and ISO settings automatically. In modern digital photography, stops are also a unit used to quantify ratios of light or exposure, with each added stop meaning a factor of two, and each subtracted stop meaning a factor of one-half. The one-stop unit is also known as the Exposure Value (EV) unit [6]. The combination of exposure time and aperture value determines EV unit, which regulates the amount of light falling on the sensor. The EV unit is defined as follows:

$$EV = \log_2 \left(\frac{f_L^2}{t_e} \right) \quad (1.5)$$

The appropriate value of EV unit for taking a picture in brightly and poorly lit scenes is chosen as per lighting conditions in the scene [7]. In order to handle complex lighting conditions, exposure compensation is performed automatically in the metering mode. A series of exposures of the same subject is collected with the help of Automatic Exposure Bracketing (AEB) whereby exposure settings are determined by the built-in light meter [8].

Moreover, for exposure compensation AEB can be used in aperture priority, shutter priority, program and manual modes. In aperture priority, the aperture you set will remain the same and the shutter speed will be automatically adjusted. In Shutter Priority, the shutter speed you set will remain the same and the aperture will be automatically adjusted. In program and manual modes, the aperture will remain the same and the shutter speed will be automatically adjusted.

If the DR of a scene (i.e. range of intensity variations) is too high to be captured by a conventional digital camera, in single exposure, camera sensor can capture highlights (i.e. brightly illuminated regions) or shadows (i.e. poorly illuminated regions) or mid-tones (i.e. moderate light) but not all in the same exposure. The DR of the sensor determines image sensor quality, which is defined as the ratio of maximum charge that the sensor can collect without saturation and the minimum charge that is just above the sensor's noise level [9]. In decibel (db), the DR of the sensor is given by,

$$DR = 20 \log_{10} \left(\frac{I_{max}}{I_{min}} \right) \quad (1.6)$$

The idealized response of digital camera sensor from highlights, mid-tones and shadows present in the scene is shown in Figure 1.2.

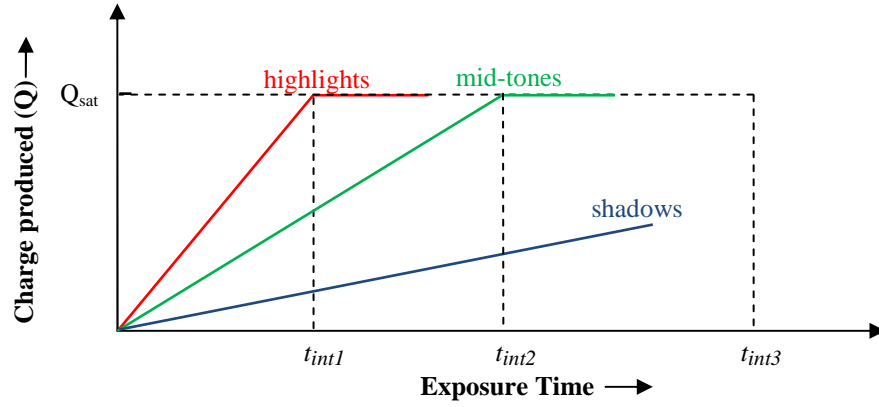


Figure 1.2: Idealized response of digital camera sensor from highlights, mid-tones and shadows present in the scene.

In Figure 1.2, the graph is plotted between collected charges versus exposure times for three illuminations. We assume throughout that both the aperture and ISO setting are constant over integration times (t_{int1} , t_{int2} , and t_{int3}). There is one setting (i.e. exposure time), which determines how long the sensor is exposed to light. The charge (Q) produced at the end of integration is a functional $f[.]$ of the current $I(t)$ over the integration time $0 \leq t \leq t_{int}$. When the sensor is operating in integration mode [9], functional $f[.]$ is given by:

$$f[.] = \int_0^{t_{int}} I(t) dt \quad (1.7)$$

As depicted in the graph, in difficult lighting situations, where highlights, shadows and mid-tones appear simultaneously, the camera sensors under the influence of highlights are saturated at integration time (t_{int1}). As the exposure time increases further, the camera sensors under the influence of mid-tones are saturated at integration time (t_{int2}). Similarly, as we have seen for integration time (t_{int3}), the camera sensors under the influence of shadows

are producing non-saturating signal, while the sensors under the influence of highlights and shadows have been saturated. Therefore, after a short integration time, highlights are captured before the sensor saturates, and adequate integration time is required for capturing shadows present in the scene.

An example is shown in Figure 1.3 (a-d). These four images were captured at the same time with different exposure settings, but they appear very different. Here the camera has exposed for the outdoor details (Figure 1.3 (a)) resulting in an underexposed indoor details. Increasing the EV (Figure 1.3 (d)) resulted in overexposed outdoor details. Images illustrated in Figure 1.3 (b and c) are captured with a slightly different exposure compensation setting. In fact, long exposure yields details in the poorly illuminated areas while short exposure provides detail in the brightly illuminated area. Therefore, each exposure gives us trustworthy information about certain pixels, that is, the optimally exposed pixels for that image. In such type of images, for dark pixels, the relative contribution of noise is high and for bright pixels, the sensor may have been saturated.

In practice, it is desirable to ignore very dark and very bright pixels to achieve supra-threshold viewing conditions [10]. Consequently, the scene contains very dark and very bright areas are partially under- or overexposed in the optimally exposed photograph (please see Figure 1.3 (c)). This is because of limited DR of standard digital cameras (i.e., 10^2). The solution is to photograph the scene several times with variable exposures and reconstruct blended image that contains the whole details, even in brightly and poorly illuminated areas. HDR imaging techniques give the solution to recover radiance maps from photographs taken with conventional imaging equipment.



Figure 1.3: Multi-exposure shots from unevenly-lit scene depict indoor and outdoor area: images captured at different exposure setting.

To make the concept of DR clear, let us redefine some useful terms. Image is said to be LDR when its DR is lower than that of the output medium. A Standard Dynamic Range (SDR) image is the one whose DR corresponds approximately to that of the standard output

medium (i.e., 0–255 OR about 10^2) and is also called display-referred image. A HDR image has DR higher than that of the output medium and it is also called scene-referred image.

Alternatively, the Liquid Crystal Diodes (LCD), Cathode Ray Tube (CRT) and printers have limited contrast ratio. Therefore, these devices are unable to reproduce full DR that leads to tone-mapping problem. Tone-mapping [11,12,13] is the technique to remap the intensities for display HDR images on SDR devices. Although few HDR display devices have been developed and will become generally available in the near future, this technology is very expensive and not accessible by the most users. To display HDR data directly, a number of HDR display prototypes are proposed recently by [14,15,16]. As a result, there will always be a need to prepare HDR imagery for display on LDR devices or directly generate an image that looks like tone-mapped image. The goal of exposure fusion mechanism is to maximize information content of the synthesized scene from a set of multi-exposure images without computing HDR radiance map and tone-mapping.

In exposure fusion, compositing is done on the pixel intensity values rather than irradiance values. Approaches proposed in this thesis do not care about the exposure times and CRF, which is required to linearize the image data before combining LDR exposures into HDR image [17]. Following the consideration of pixel intensity based fusion; the ultimate goal of this thesis is the utilization of computationally simple and robust texture features for the identification of well-exposed regions across input exposures and detail enhancement in the fused image.

Nowadays much research is going on in the area of exposure fusion. The techniques proposed in the present thesis do not limit interest for developing yet another new algorithm. Rather, the aim is to select and apply the aspects of perception of weak textures which should be considered in the context of exposure fusion, to generate weight maps, and to seek further possibilities for detail enhancement by exploiting edge preserving capability of non-linear filters. Although the proposed frameworks do not require human intervention, in practice, the present work provide set of parameters that allow users to interactively control the detail enhancement in the fused image.

1.2 PROBLEM STATEMENT

In recent years several new techniques have been developed that are capable of providing precise representation of complete information of shadows and highlights the real-world natural scenes. The direct 8-bit gray and 24-bit Red Green Blue (RGB) representation

of visual data, with the standard digital cameras in single exposure settings, often causes loss of information in the real-world scenes because the DR of most scenes is beyond what can be captured by the standard digital cameras. Such representation is referred to as LDR image. Digital cameras have the aperture setting, exposure time, and ISO value to regulate the amount of light captured by the sensors. In modern digital cameras, AEB allows us to take all the images without touching the camera between exposures; provided the camera is on a tripod and a cable release is used. Handling the camera between exposures can increase the chance of miss-alignment resulting in an image that is not sharp or has ghosting. However, most scenes can be perfectly captured with nine exposures [14], whereas many more are within reach of a camera that allows 5-7 exposures to be bracketed. When the scene's DR exceeds the DR of camera it is exposure setting that determines that, which part of the scene will be optimally exposed in the photographed image. The DR of a digital camera is typically defined as the charge capacity divided by the noise [14, 17]. At single exposure setting, either detail in the poorly illuminated area (i.e., shadows) is visible with long exposure or brightly illuminated area (i.e., highlights) with short exposure (please see Figure 1.3). Therefore, scene containing highlights and shadows are partially over or under-exposed in the optimally exposed photograph (auto-bracketed shots or auto results). However, the number and value of exposures should be sufficient to capture all luminance variations present in the entire scene.

In practice, for some real-world scenes low bit depth is sufficient to capture the entire details but there are countless situations that are not accommodated by the low bit depth. Although HDR display technology will become generally available in the near future, it will take time before most users have made the transition. Recently Sunnybrook technologies, BrightSide and Dolby prototypes of HDR display devices have been proposed [14,15,16] that can display HDR data directly. Therefore, we must generate 8-bit LDR image directly that looks like a tone-mapped image.

In the present thesis work, an alternate solution of HDR and tone-mapped image is described in which we are combining multi-exposure images directly into detail-enhanced 8-bit single LDR image that does not contain under-exposed and over-exposed regions. Thus, details in both brightly and poorly illuminated areas may be preserved by skipping the construction of HDR image and tone-mapping algorithms. The incorporation of the notion of combining multiple exposures without typical HDR and tone-mapping steps is known as “exposure fusion”, as shown in Figure 1.4.

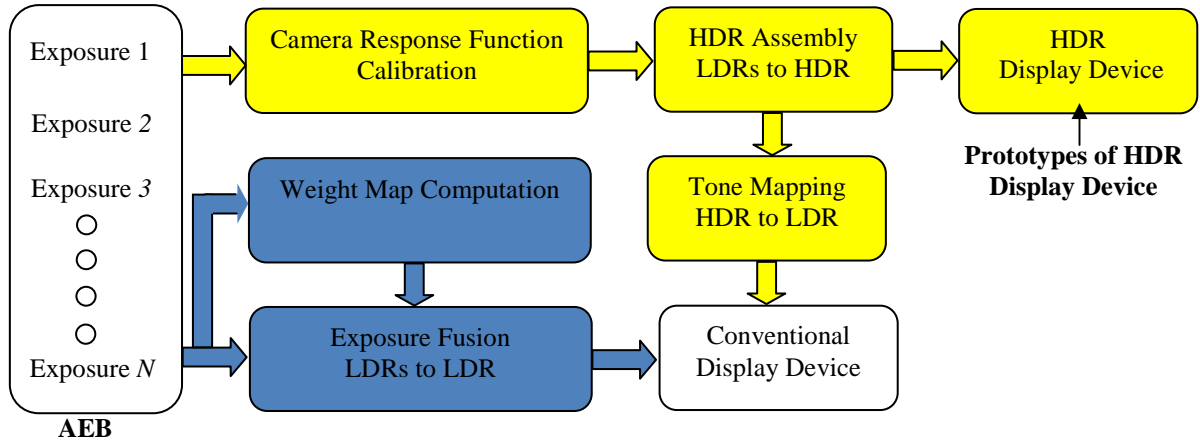


Figure 1.4: Comparison of high resolution imaging pipeline (i.e. HDR imaging followed by tone mapping and present exposure fusion approach). The Yellow color depicts HDR and tone mapping pipeline, and blue color depicts exposure fusion pipeline, which does not require post-processing.

The underlying idea of exposure fusion approaches is based on the utilization of different local measures for generating weight map (W) to preserve details present in the different input exposures (I_k). The fused image (I_F) can be computed as:

$$I_F(x, y) = \sum_k^N W_k(x, y) I_k(x, y) \quad \text{where } k = 1, 2, 3, \dots, N \quad (1.8)$$

where (x, y) specifies spatial position of pixel and N is the number of input images.

In the present thesis work, edge preserving filters are used for two-scale decomposition that separates sharp details called BL and fine details called DL across various input images captured at different exposure levels. The DL is computed as the difference between the original input image and the BL. Therefore, present approach is effective to control fine and coarse details separately during the compositing process and needs no further post-processing to produce detail-enhanced image. Thus, Equation (1.8) can be rewritten as:

$$I_F(x, y) = \sum_k^N W_k^B(x, y) BL_k(x, y) + \sum_k^N W_k^D(x, y) DL_k(x, y) \quad (1.9)$$

where W_k^B and W_k^D are the weight maps computed for k^{th} BL and DL, respectively.

Therefore, sharp details and fine details are manipulated and fused separately:

$$I_F(x, y) = BL_F(x, y) + DL_F(x, y) \quad (1.10)$$

where BL_F is the fused BL and DL_F is the fused DL based on W_k^B and W_k^D weight maps, respectively.

The current state-of-the-art method for automatic exposure fusion exploits the capability of edge preserving filter to generate weight map function that guides the fusion of BLs and DLs. Therefore, to achieve optimal contrast and color details in the fused image, present research develops an appropriate mask based on weak textures and color saturation measure. While previous exposure fusion methods perform the fusion only concern with local contrast, the proposed algorithm takes weak textures, and saturation measure into consideration that can enhance texture details evidently.

1.3 MAIN CONTRIBUTION

The present thesis work approach multi-exposure image fusion by exploiting the edge preserving and intra-region smoothing property of nonlinear diffusion filters. Weak texture details will decide the contribution of corresponding pixel across different exposures to the fused image. A rich texture details means a maximum contribution, which tells that image block has higher weight during the fusion process. Such metric is used to quantify perceived local contrast in an image under different exposure settings, and allows discarding under exposed and overexposed pixels in the fused image. Therefore, instead of employing intermediate HDR reconstruction and tone-mapping steps, well-exposed fused image is generated for displaying on conventional display devices and realistic-looking image that is much closer to what our eyes originally saw. This approach does not care about the exposure times and CRF. A further advantage of present technique is that it is well suited for other image fusion applications. In order to measure the success towards this goal, not only is an exposure fusion approach proposed, but the same algorithm is also employed for the fusion of multi-focus image series and images captured with flash and no-flash. The main contributions in this thesis are highlighted as:

1. Local range based texture analysis is proposed to generate a mask that guides the fusion of BLs computed across input images. Multi-Resolution Pyramid (MRP) approach is described for BLs fusion. The BL is computed by applying nonlinear filter that preserves the locations where the magnitude of the gradient has maximum value. The DLs extracted from input exposures are manipulated and fused separately, which yields enhancement of texture details in the fused image. The present work employs monotonic nonlinear activation function (i.e. sigmoid function) to enhance fine details in the fused image. The algorithm overcomes the major drawbacks of

conventional MRP based fusion, namely, the blurring of edge details and the reduction of fine texture details.

2. A novel detail-enhancing exposure fusion approach using NTF is proposed. With the captured LDR images under different exposure settings, first the fine details are extracted based on Guided Filter (GF). Next, the BLs (i.e. images obtained from NTF) across all input images are fused using MRP. Exposure, contrast and saturation measures are considered to generate a mask that guides the fusion process of BLs. Finally, the fused BL is combined with the extracted fine details to obtain detail enhanced fused image. The algorithm proposed in the thesis shows that GF based BL and DL extraction is more effective for enhancing texture details and reducing gradient reversal artifacts near strong edges in the fused image.

3. Two-scale decomposition based on Anisotropic Diffusion (AD) is proposed for fast exposure fusion, which does not require optimization of number of scales as required in the conventional MRP techniques. Weighted average based blending of coarser details and finer details is performed to form a composite seamless image without blurring or loss of detail near large discontinuities.

4. A novel weight construction approach is proposed to combine texture features and saturation measure for guiding image fusion process. For weight map construction, we seek to utilize the strength of texture details under the change of exposure setting that takes place between an under-exposed and an over-exposed image. WLS filtering is proposed for weight refinement. Furthermore, fast sigmoid function based weight map generation for DLs is proposed that reduce computational complexity of the algorithm.

5. The important contribution of this thesis is the advantages include ease of implementation, quality of compositing, and the provision of DL enhancement without introducing artifacts.

1.3 OUTLINE OF THESIS

In order to solve the multi exposure image fusion problem, this thesis consists of 7 Chapters. This is the first chapter. Chapter 2 discusses the related work, covering methods that are directly related to proposed techniques. This chapter covers the most popular and advanced researches in the imaging technology, HDR imaging, HDR software, sensor design, tone-mapping, inverse tone-mapping, image fusion and exposure fusion.

Chapter 3 explains the elaborate proposed methods for multi exposure image fusion. An edge-preserving AD filter is applied to the input images for BL and DL separation. This chapter will discuss how a MRP technique can be used for preserving and enhancing weak texture details in the fused LDR image. In Chapter 4 explains variational approach for preserving and enhancing details in a final fused LDR image. It discusses the framework of variational method for artifacts reduction near strong edges, which uses a spatial adaptive GF approach for BL and DL separation.

The chapter 5 explains the WLS filter based weight refinement approach for weighted average fusion for generating a LDR image from multi exposure images. The weak textures and saturation measure for weight map generation is discussed that guides the multi-exposure image fusion process. It will also discuss the usefulness of sigmoid function for DL enhancement.

In Chapter 6, the results generated from three approaches are demonstrated. This chapter will examine present algorithms as interactive tools for manipulating the detail and contrast in multi-exposure, multi-focus and multi-flash image fusion problem. Finally, Chapter 7 provides general conclusions and motivates ideas for future directions.

CHAPTER-II

LITERATURE SURVEY

This chapter will give the overview of previous work related to imaging technology, HDR imaging, HDR software, sensor design, tone-mapping, inverse tone-mapping, image fusion and exposure fusion. This chapter will discuss the history, capabilities, and future of existing solutions to handle wide range of illumination conditions present in the natural scenes. This chapter will first survey challenges in recording outdoor and indoor scene that includes a directly visible sun and shadows, respectively. It reviews HDR imaging technology that can encode the color gamut and DR of the original scene into true radiance maps. This chapter provides a brief review of state-of-the-art methods available for sensor design, and the trade-offs for HDR sensor design. Chapter also discusses the difference between the standard digital image representation and the HDR imaging technology. To date, various tone-mapping algorithms have been proposed which prior to display reduces the tonal range of HDR data to match the capabilities of display devices at a minimal side effect on quality. This chapter gives the overview of different global and local tone-mapping methods. It will also briefly review the background work about range expansion in the case of inverse tone-mapping algorithms. Then, this chapter presents brief discussion on exposure fusion approaches available in the literature and indicates their advantages over HDR imaging. At the end of this chapter, chapter provides a brief review of edge-preserving filters, Multi Scale Decomposition (MSD) and their utilization in computer graphics, and image processing applications.

2.1 IMAGING TECHNOLOGY

Digital camera having electronic imaging array and analog camera with film use nonlinear mapping that determines how radiance in the scene becomes pixel values in the image. In the early days of photography, photo-sensitive film was used to capture luminance variations present in the scene. The experimental setup for the formation of the latent image was carried out principally in the Kodak Research Laboratories at Rochester and Harrow in 1947. An overview of the reactions of the photochemistry of silver halide is presented in [18] and [19]. According to them, silver halide crystals are photoconductors which are activated

by the photons. The chemical process by which molecules are broken down into smaller units through the absorption of light is known as photolysis and it involves electronic and ionic migration. In a conventional camera, the photosensitive-film is first exposed to true radiance to form a latent image. The film is then developed to change this latent image into visible image. The film can then be digitized using a film scanner, which projects light through the film onto an electronic photo-sensitive array, converting the image to electrical voltages. These voltages are digitized, and then manipulated before finally being written to the storage medium. If prints of the film are scanned rather than the film itself, then the printing process can also introduce nonlinear mappings. The detailed description of photochemical process involved in analog cameras is provided in [20] and [21]. Although, analog cameras have greater DR than their digital counterparts and development process is used to limit or enhance the information retrieved from the exposed emulsion. Dodge-and-burn technique is commonly used to extract maximum information during the development process which acts as a tone-mapping step in analog photography. Dry developing method was developed by Applied Science Fiction and marketed to extract the full DR from a negative [22].

The primary disadvantages of analog cameras are that they are non real time and have limited features. To circumvent these restrictions digital cameras have been developed to capture, process, store and share image data quickly. In digital photography, camera's sensor performs same function as photosensitive film does in the analog camera. The complex and sophisticated grid arrangement of millions of photodiodes is known as photo-sites. Each photo-site represents one pixel of the captured image. More the number of photo-sites better will be the resolution of captured image.

The CCD or CMOS is a major piece of digital imaging where the luminance variations are initially captured before converted to the digital form. CCD concept was introduced by Boyle et al. in 1970 [23]. The significance behind the name of CCD is that each row of photo-site is connected. The accumulated charge from each photo-site is passed down row-by-row and read at one corner of the array before being deleted. In a CMOS chip, the charge value is read from each individual photo-site and the manufacturing is cheaper than CCD. Moreover, CMOS chip is more energy efficient [24]. The most commonly used sensor type in today's digital cameras is the CCD because this array type has certain features which make it more efficient at gathering light and producing high quality images. Photo-sites record light intensity, but cannot record colour. To generate a colour image, a Colour Filter Array (CFA) is placed over the sensor and the values of the missing color bands at every pixel are often synthesized using some form of interpolation from neighboring pixel values is

known as demosaicing [25]. Numerous demosaicing methods have been proposed through the years.

Gradient was considered first time in [26]. It is reasonable to assume that there is no correlation between the values of neighboring pixels located on the different sides of an edge. They proposed that interpolation can be performed only in smooth regions because smooth hue transition is valid for low frequency regions. The second derivative of the R and B channels is used to estimate the edge direction in the G channel. Later, the G channel is used to compute the missing values in the R and B channels. Other examples of popular demosaicing methods available in literature are [27] and [28]. In [27], the second derivative of the G channel and the first derivative of the R (or B) channels are used to estimate the edge direction in the G channel. Spatial and spectral correlations among neighboring pixels are exploited to define the interpolation step in [27]. Adaptive median filtering was used as the enhancement step. Kodak professionals have invented that in typical digital cameras, Bayer pattern can be preferred in CFA because human eye is much more sensitive to G light [29].

In digital photography, sensor size has impact on the quality of the images being captured. This is quite significant for determining the size of individual photo-sites. Smaller photo-sites are less sensitive to light, less able to capture poorly illuminated area or strongly illuminated area of scene accurately, and can result in increased level of noise [24]. Smaller photo-sites will yield more shadow from dark areas and more highlights from white areas. Smaller CCD typically generates smaller pixels, too. However, compact cameras have drawbacks compared with the cameras having larger sensor size. For a long while now, manufacturers have been employing three CCDs to assign the job of processing one of the three red, green and blue primary colours independently by which observer can enhance or compress individual colour as per the requirement.

The next important category of camera I would like to mention here is DSLR camera in which a mechanical mirror system and pentaprism is used to direct light. DSLR permits the photographer to view through the lens and see exactly what will be captured. One of the big advantages of DSLR is that the lenses are interchangeable. There are numerous types of zoom lens available for DSLR [30]. This 70-200 mm zoom lens from Canon has the ability to zoom 3X.

2.2 HDR IMAGING

There is a tremendous need to record a much wider gamut than standard 24-bit RGB. The direct 8-bit gray and 24-bit RGB representation of visual data, with the standard digital cameras in single exposure settings, often causes loss of information in the real-world scenes. This is because; the DR of most scenes is beyond what can be captured by the conventional digital cameras. Such type of representation is referred to as LDR image. Therefore, the camera is unable to capture what our eyes can see. That is why HDR photography is needed to capture all brightness variations present in the scene.

Conventional digital cameras can only capture either detail present in the poorly illuminated or brightly illuminated regions. To cover the entire DR in such a scene, one can take a series of photographs with different exposures. The contrast range (i.e. difference between the lightest and darkest part in the scene) determines number of exposures you need to capture the full DR of image. To measure contrast range, various metering modes have been proposed in [31]. Spot meter is the most commonly used metering approach that allows the user to measure a very small area of the scene (between 1-5% of the viewfinder area). With this metering mode, EVs are determined accurately because the measured brightness is not influenced by other areas in the frame. The first spot meter was developed by Arthur James Dalladay in 1935. In modern DSLR cameras have built-in spot meters.

HDR image encodings come into play when full range of color and luminance values need to be captured. The HDR photography produces a single photo that is correctly exposed in both the dark and light areas. The practice of assembling HDR image from multiple exposure images [14,32] recover radiance maps from photographs taken with conventional imaging equipment. Debevec and Malik [17] and Mann and Picard [33] proposed a HDR imaging to record the entire range of the scene radiances from different exposures that were acquired with a standard camera. The CRF recovered from differently exposed images is used to create HDR image whose pixel values are equivalent to the true radiance value of a scene. Quadratic objective function based on least square error was used by Debevec and Malik [17] to recover smooth and monotonic response function. Furthermore, to ignore saturated pixel values, the weighting function is used to give higher weight to exposures in which the pixel's value is closer to the middle of the response function. Single-precision floating point values are used to encode recovered radiance map. The approach proposed by Mitsunaga and Nayar [34] utilizes polynomial approximation for their response function. This technique is suitable for low-cost consumer equipments.

In recent years significant progress has been made in the development of algorithms for local and global image registration [35,36]. Three fully automatic methods for eliminating misalignments between a sequence of multi-exposure images have been proposed [37,38,39]. The first method [37] employs Median Threshold Bitmaps (MTB) to accelerate image operations. Then pyramids of MTBs computed for each input image are generated, which are then aligned horizontally and vertically using bitwise shift and differencing operations. The second method [38] can handle both global movement and local movement in a scene, and is based on a variant of the Lucas and Kanade motion estimation technique [40]. In [39], a global registration method was proposed in which Scale Invariant Feature Transform (SIFT) [41] was employed to search for key-points in consecutive multi-exposure images. Random Sample Consensus (RANSAC) [42] algorithm was used to choose the best key-point correspondences and Direct Linear Transform (DLT) [39] technique was employed to calculate transformation matrix.

In some research, disparity maps [2] have been utilized to generate a spherical HDR image for image-based virtual environments from a moving camera. In the first stage of this frame work, rotational alignment of exposures taken by the camera is performed. In the second stage, with a spherical stereo vision toolkit, they described a new method for disparity maps calculation. In the final step, they developed a spherical HDR radiance map, which is obtained from the warped exposure bracket. It showed that a spherical HDR image can be generated for each of the viewpoints of the LDR images captured from indoor and outdoor scenes.

HDR imaging [2] is also called scene-referred representation which represents the original captured scene values as closely as possible. Such representation is sometimes referred to as extrasensory data representation. One of the important applications of HDR capturing techniques for security application is capturing video at entrance of the buildings [14]. Conventional cameras are incapable to faithfully capture the interior and exterior of a building simultaneously while HDR camera, which is based on two-phase workflow, would be able to simultaneously record indoor as well as outdoor activities. Other important applications of HDR representation are satellite, scientific and medical imagery, in which data is analyzed and visualized to record more than what is visible to the naked eye. Therefore, after acquiring HDR data, an efficient encoding technique is needed to avoid taking an excess of disk space and network bandwidth. Various possible formats to store radiance maps are described by Reinhard et al. [14]. Four bytes “hdr” format was developed to encode radiance maps. Run-length encoding is employed to achieve an average of 25%

compression. It is the most commonly used format for radiance lighting simulation, rendering system and image-based lighting [17,43]. It was proposed by Ward-Larson and Shakespeare in 1989 [44].

Another option to encode radiance map is “floating point tiff”, which uses 32 bit/component Institute of Electrical and Electronics Engineering (IEEE) floating-point RGB encoding [45] to encode 79 orders of magnitude approximately. “Floating point tiff” can encode a very high DR with better precision than the radiance format. In 2002, Industrial Light and Magic (ILM) has proposed Extended range format (.exr), which is based on based on a 16-bit half floating-point type with fewer bits [46]. In 24-bit LogLuv encoding, 10 bits are used for log luminance value and remaining 14 bits are used to represent chromaticity [47]. However, the success of HDR image capture has shown that it is possible to produce an image that exhibits details in poorly and brightly illuminated areas. Moreover, HDR formats have since found widespread applications in the computer graphics and HDR photography.

It is not possible to display HDR data on conventional image output devices [3]. CRT and LCD are the leading softcopy LDR display devices over the years. In a CRT, the DR is high but maximum brightness is limited by the amount of energy we can safely deposit on a phosphorescent pixel without damaging it or generating unsafe quantities of X-ray radiation [14]. On the other hand, the LCD can achieve high brightness with a limited DR. In Recently explored HDR display devices, backlight modulation is used to increase DR of LCD. Recently Sunnybrook technologies, BrightSide and Dolby prototypes of HDR display devices have been proposed [14,15,16] that can display HDR data directly. The prototypes of HDR display devices provide direct HDR display capabilities by means of a projector or Light Emitting Diode (LED) array that lights the LCD from behind with a spatially varying light pattern [15,16]. Initially, Sunnybrook technologies employed projector-based backlight modulators and LCD panel as the final view portion of their display. The DR of these devices is 50,000:1 and the maximum luminance can be up to $2,700 \text{ cd/m}^2$ [14]. Later, a new design was proposed in which LED array is used in the backlight with a spatially varying light pattern. The maximum luminance of LED based display can be up to $8,500 \text{ cd/m}^2$. However, these prototype devices can represent high contrast ratio, which are still in the developing stage and not available in the market for the routine customers.

2.3 HDR SOFTWARES

The problem of improving DR by fusing images captured from static scene at different exposure settings is well described in the literature and up until now, various useful software to build HDR image have been developed e.g. Photosphere, HDRshop, and Photomatix.

Photosphere is developed by Greg Ward for building HDR images, which can work on Tiger and later versions of Mac OS X [48]. To date, this software is the best option to explore and enjoy playing with HDR imagery. Recently launched Apple iPhone 4S uses Photosphere to extract CRF for HDR builder application [49].

HDRshop is developed by Paul Debevec at the University of Southern California's Institute for Creative Technologies and designed to view and edit HDR images. Instead of storing a pixel's on-screen, HDRshop stores the amount of light it represents. Since there is no limit to how much light you can have in the real world, HDR Shop stores these pixel values as floating point numbers. That is, instead of storing pixels using the numbers (0, 1, 2, 3, ..., 254, 255) like in an 8-bit image, it uses numbers like 0.01534, 0.9429, 1.0500, and 1.356, 35.0253 [50].

HDR Shop can read and write HDR formats such as Radiance's HDR, 16-bit or floating point TIFF, Portable Float Maps (PFM), and raw binary files. Moreover, HDR Shop can import and export conventional 8-bit image formats, including Joint Photographic Experts Group (JPEG), Windows Bitmap (BMP), and Tagged Image File Format (TIFF) for generating HDR images. Moreover, HDRshop has some nice features that make it flexible and user-friendly: multi-threading option to speeds up work-flow, file format customization to support new file formats, JavaScript scripting interface to create your own custom operations and plug-ins, and image filters to implement Star Blur, Motion Blur, Ward Specular Convolution, Vertical Cosine Falloff and Fast Diffuse Convolution.

Photomatix is developed by HDRsoft to fuse the series of differently exposed photos. Photomatix Pro and Photomatix Essentials are two stand-alone versions of Photomatix that can be run on Windows and Mac OS X. Photomatix Essentials is an excellent simple tool for producing HDR images and is user-friendly. Photomatix Pro offers more options and includes advanced features such as batch processing and selective dehosing. Photomatix Pro is the one of the version that allows the users to create tone-mapped images, contrast optimizer for realistic-looking results, Finishing touch adjustments for contrast and sharpening, and option to enable fusion from a single raw file [51,52].

Extended version of Photoshop uses new HDR Pro technology to produce 32-Bit HDR images for creating widest range and richest detail with automatic ghost removal. It also provides various features including greater control with tone-mapping and contrast adjustments. Adobe Photoshop PCS2, Photoshop CS5 and Photoshop CS5 Extended provide a script that automatically fuses a set of multi-exposure images into a radiance map.

The recently released Image processing toolbox of MATLAB also provides some inbuilt function for reading HDR images, writing HDR images, building HDR image from different exposures and tone-mapping of HDR image for display. `hdrread` function is used for reading HDR data, `hdrwrite` function is used to write a Radiance HDR image, `makehdr` function creates the single-precision HDR image from the set of spatially registered LDR images and `tone-map` function converts the HDR image to a LDR image using a process called tone-mapping, which is suitable for display.

2.4 HDR SENSOR DESIGN

Beside software solution, to increase the DR of real-time digital still imagery and video camera, a lot of computational capability and processing speed is required which is possible with the assistance of graphics processing hardware [34,38,53]. Although this is largely achievable, the hardware design approaches did experience some key issues associated with the development of the real-time HDR reconstruction: These are: computation time, computation cost, registration, sensor design and realization in hardware. Various solutions for designing the HDR image sensors have been proposed throughout the years [54,55,56,57]. An overview of various HDR sensor design techniques is presented in the tutorial by El Gamal and Yang [9].

Pixel-level Analog to Digital Converter (ADC) can be used for widening the DR of the CMOS sensor [58]. Apart from increasing DR of the sensor, such technique can be used to achieve high Signal-to-Noise-Ratio (SNR). Non-uniform bit resolution for each sample can be used to capture all intensity variations present in the scene, which can be represented accurately through binary-floating-point codes. In this approach, a 640×512 image sensor with pixel-level ADC was implemented in a 0.35μm CMOS technology.

Another plenoptic function [59] based camera model, designed for increasing DR, was introduced by Georgiev and Lumsdaine [60]. This model is based on the 3D aspect of light field capture and manipulation, with less attention paid to wavelength and polarization parameters. Their approach employs microlens aperture filters at the main lens aperture,

which allows high spatial resolution as well as increased DR capture. The unique feature of their approach is that different views depicting typical luminance details can be separately filtered by placing filters or darker Neutral Density (ND) filter at the main lens, which avoids blooming [60].

Fixed-Pattern-Noise (FPN) correction procedure is proposed by Dass and Collins in [61]. In their method a 100×98 CMOS Image Sensor (CIS) is fabricated in a standard $0.35\text{-}\mu\text{m}$ CMOS technology. FPN correction procedure is employed to reduce artifacts and estimation of accurate mean photocurrent from a linear response at low photocurrents and a logarithmic response at larger photocurrents of sensor. Look-Up Table (LUT) was used to estimate photocurrent for every sensor. The threshold values used to estimate photocurrent for linear and logarithmic model were chosen 30 and 10^3 , respectively. They have verified that that in the logarithmic region, LUT-based method reduced FPN to approximately 2% of the photocurrent, while in the linear region of operation, the residual noise in the response to uniform scenes increases because of quantization noise. In this manner, this method would determine accurate photocurrent flowing in each sensor irrespective of the high computational cost than the other methods.

One of the best HDR-CMOS sensor design methods recently is a re-configurable pixel circuit [62], which has the capability of capturing images with a Wide Dynamic Range (WDR) of intensities and presenting it in a format fit to be directly displayed on LDR displays. They proposed programmable transduction function which enables it to produce any monotonically increasing response. By introducing this flexibility, the algorithm is capable to adapt any desired operator without requiring any change in the pixel circuit. So they propose an adaptive approach for handling WDR intensity variations present in the scene.

The comparative quantitative analysis of the performance of different WDR sensors has been given by Spivak et al. in [63]. In such analysis SNR, DR extension, noise floor, minimal transistor count, and sensitivity parameters are employed using consistent assumptions and definitions. Advantages and drawbacks of each of the seven sensor categories available in the literature were discussed and summarized. Furthermore, a comprehensive analysis on power consumption of different sensors has been given in [64]. The power-performance tradeoffs of three generations of WDR sensors are also discussed in this survey.

2.5 TONE-MAPPING OPERATORS

Unfortunately, conventional display devices (i.e. CRT and flat panel display) have DRs spanning a few orders of magnitude, much lower than those of the real world scenes, often less than 100:1. In order to display HDR images on monitors or print them on paper [65], it is must to remap the DR of the HDR images to reproduce LDR images suitable for Human visual System (HVS). Therefore, the ultimate goal of tone-mapping operators is to maintain original appearance of the tone-mapped HDR image on existing display devices whose capabilities in terms of DR are insufficient. Therefore, appropriate metrics based on human visual perception need to be accounted to depict their original appearance. Thus, the investigation of response of photoreceptors distributed on the retina with a finite resolution is very worthwhile in the context of tone-mapping operators. The reader is requested to refer [14,66,67,68] for detailed descriptions.

Tone-mapping operators may be classified into two broad categories: Global operator (linear scaling) and Local operator (non-linear scaling). The linear scaling is the one of the fastest and easiest way to achieve the goal of contrast reduction in which radiance values that are above or below certain range are saturated to the upper or lower limit values. This means you need to determine the lower and upper limits to truncate full input data range. The tone-mapped image in Figure 2.1 (a) shows the result using linear scaling [14], which is noticeably producing low contrast image. It illustrates typical ton-mapped HDR image depicts indoor and outdoor details. The area inside the door is illuminated by the interior light and outdoor area is illuminated by the sunlight. In this image, the HDR radiance maps were recovered from the recorded multi-exposure image series using the method of Debevec and Malik [17]. To preserve local and global contrast in the tone-mapped image automatically, non-linear scaling leads to a useful operator that produces plausible results, the results of which are shown in Figure 2.1 (b). However, to reproduce detail-preserving tone-mapped image from HDR data for display on conventional display devices is still an open issue.

In the literature, several tone-mapping methods for converting real-world luminances to display luminances have been developed and fulfilling the fast growing demand of displaying HDR data on conventional display devices. Most tone-reproduction algorithms make use of photoreceptor adaptation [69,70] and simulation of real-world light conditions [71] to achieve visually plausible results. Local light adaption property of HVS is adopted in the local operators to correspond to the visual impression an observer had when watching the original scene, while the global operator are spatially invariant and are less effective than the local



Figure 2.1: Tonemapped HDR images: **(a)** Global operator and **(b)** Local operator. Although the input HDR image is same for both cases, non-linear scaling often generates LDR image with complete detail (contrast) than linear scaling. Images courtesy of the Reinhard et al. [14].

operators.

The algorithm proposed by Reinhard et al. [13] uses Ansel Adams' zone system to handle HDR images. To preserve fine details, sigmoid scaling function and local dodging & burning techniques are employed on log average luminance. Local operators [72-74] involve the spatial manipulation of local neighboring pixel values based on the observation that HVS system is only sensitive to relative local contrast. To model local adaption properties of HVS, input image is convolved with a Gaussian filter kernel. Algorithm proposed in [13] has examined that the performance of algorithm is usually dependent on the size of local adaption area. Indeed if the size of local adaption area is chosen incorrectly, artifacts may occur in the tone-mapped image.

Another popular local operator employs a symmetrical analysis-synthesis filter bank for decomposing and reconstructing images [72]. This model takes into account the sigmoid for controlling the gain at each location and applies local gain control to modify the sub-bands computed for input HDR data. The gain is low for high values and high for low values. This approach is also useful for range expansion by which a LDR image can be expanded into a HDR.

Global operators [75-77] do not involve spatial processing. Such operators apply certain spatially uniform remapping function on every pixel independently. Histogram adjustment based technique has been proposed by Larson et al. [75] in which models of human contrast sensitivity, glare, spatial acuity, and color sensitivity are incorporated to preserve subjective correspondence between the real scene and its displayed image. A simple

adaptive logarithmic mapping technique was proposed for HDR video player by Drago et al. [76], which can be implemented very efficiently in the graphics hardware. Other global operator includes control intensity, contrast, and level of chromatic adaptation parameters for DR compression [77]. In addition this model used separate functions for rods and cones to account for the intensity in scotopic and photopic lighting conditions. However, most of global operators are computationally efficient than local operators.

A simple S-shaped curve (i.e. sigmoid function) has been utilized as tone-mapping function [78]. The middle portion of such sigmoidal function is nearly linear and thus resembles logarithmic behavior. Moreover, sigmoidal functions have two asymptotes: one for very small values and one for large values.

Transform domain tone-mapping approaches [79] became popular compared to intensity domain. The properties of HVS for HDR compression are employed in gradient-domain and the Gaussian pyramid is used to obtain different levels of contrast. Weber's law [1] based attenuation maps are utilized to compress gradient magnitudes in the supra-threshold layer. Pixels whose gradient magnitudes are smaller than 0.0043 are considered as the sub-threshold layer and pixels whose gradient magnitudes are greater than 0.0043 are considered as supra-threshold layer. Then the resulting modified gradient fields are converted to an image by a Poisson solver.

Fattal et al. [73] has introduced gradient based approach to preserve details from HDR image. They attempt to simulate adaptation behavior of HVS for gradient modification at various scales. A reduced, LDR image is then obtained by solving a Poisson equation on the modified gradient field. This local intensity range changes based algorithm reduces the DR in transform domain to preserve local changes of small magnitude. This method is almost free of artifacts and does not require any manual parameter tweaking.

Recently, DR compression based on two-scale decomposition has been proposed [74]. Only the magnitude of the BL is compressed in the log domain, thereby preserving detail. The BL of input HDR image is computed using an edge-preserving filter called the Bilateral Filter (BF) and the DL is the division of the input intensity by the BL. This technique has proven to be a valuable tool to extend beyond the scope of DR compression.

Most of the tone-mapping algorithms suffer from halo artifacts and require human intervention in the parameter adjustment process. Both local and global operators have their own advantages and disadvantages in terms of computational cost, easy implementation, halo effects (artifacts), spatial sharpness, and practical application. The detailed review of various tone-mapping operators is given in Reinhard et al. [14]. Although attempts to reproduce HDR

data using an early stage of HVS have been carried out for better compression in high contrast areas while increasing the visibility of texture details in low contrast areas, it remains an important issue.

2.6 INVERSE TONE-MAPPING

In the last few years, researchers have focused on developing HDR prototype displays [15,16] that can display HDR data without tone-mapping. It is anticipated that the HDR display devices will be available in the consumer market for wide-spread use. The majority of today's cameras are capturing LDR data. Therefore, the expansion of LDR content for displaying on HDR screens is an interesting research topic, which is known as inverse Tone-Mapping Operator (iTMO) [80,81]. [82] has proposed an approach for expanding the DR of legacy video and photographs for viewing on HDR displays. Inverse gamma and image pyramid is used for contrast stretching and brightness enhancement, respectively. For real-time iTMO solutions, this approach is well suited for implementation on Graphics Processing Units (GPUs), and Digital Signal Processors (DSPs).

Meylan et al. [83] has proposed an iTMO with the specific task of representing highlights in LDR images when displayed on HDR monitors. The main idea is to detect the diffuse and specular part of the image and to expand these using different linear functions. Two stage iTMO is proposed by [84]. In their model, firstly, an iTMO is applied to an LDR image or a frame of a video to expand its luminance value. In the second stage, an expand-map from high luminance areas is built using density estimation. These expand maps avoid incorrect expansion from iTMO. For video sequence, the temporal expand-maps are used as weighting for interpolation of the LDR and the inverse tone-mapped image.

In [85], median cut is used to find the areas considered of high luminance and subsequently apply a density estimation to generate an expand-map to extend the range in the high luminance areas using an inverse photographic tone reproduction operator. Rigorous psychophysical investigations to determine how LDR images are best displayed on a state-of-the-art HDR monitor are provided by Ahmet et al. [86]. Moreover, Evaluation of iTMO through varying exposure conditions has been proposed in [87]. In this case, both subjective and objective [88] experiments were employed for performance evaluation. In these validation studies, it has been concluded that instead of considering individual pixel values; global statistics about the image should be taken into account for developing superior iTMO.

2.7 IMAGE FUSION

Image fusion techniques blend information present in different images into a single image. Burt and Adelson [89] first introduced the idea of image fusion based on Laplacian pyramid. Image fusion techniques are generally classified into three categories: pixel level, feature level, and decision level, which are reviewed by Smith and Heather [90].

In recent years, various fusion algorithms have been developed to combine substantial information from multiple input images into a single composite image. The principal motivation for image fusion is to extend the DOF, extend spatial and temporal coverage, to increase reliability, extend DR of the fused image and the compact representation of information. Imaging sensor records the time and space varying light intensity information reflected and emitted from object in a three-dimensional observed physical scene. However, the characteristics recorded from the incident radiations in the source images, such as exposure value, focusing, modality, and environmental conditions, often make fusion extremely challenging. The automated procedure of extracting all the meaningful details from the input images to a final fused image is the main motive of image fusion.

Bhatnagar et al. [91] developed a method for capturing most relevant information from various imaging modalities into a single output, which plays an important role in medical diagnosis. They utilize Non-Subsampled Contourlet Transform (NSCT) as analysis and synthesis tool for image compositing.

To facilitate image fusion, it may be necessary to align input images of the same scene captured at different times, or with different sensors, or with EV settings (called AEB), or from different viewpoint using local and global registration methods [37,92]. Normally it is assumed that the input images are captured with the help of tripod mounting. Hence, in general, image fusion approach expects point-by-point correspondence between different input exposures of a scene. From technical stand point, the fused image reveals all details present in the scene without introducing any artifacts or inconsistencies which would distract the human observer or subsequent image processing stages.

Orgden et al. [93] has proposed pyramid solution for image fusion. The pyramid becomes a multi-resolution sketch pad to fill in the local spatial information at increasingly fine detail (as an artist does when painting). The Laplacian pyramid representation expresses an image as a sum of spatially band-passed images while retaining local spatial information in each band [89]. Another multi-resolution based fusion which employs gradient map of the input images to yield a fused image with true information [94]. This approach takes into

account the horizontal and vertical gradient maps for producing fused gradient map for each orientation and resolution. This gradient fusion approach utilizes Discrete Wavelet Transform (DWT) and Quadrature Mirror Filters (QMFs) in the reconstruction process. This approach was implemented for input data provided by multi-sensory arrays. Another multi-sensor data fusion technique which utilizes Total Variation (TV) [94] is the one by Kumar and Das [96].

In pixel-level approach [97], the TV semi-norm is used to solve the forward model and estimate the pixels of the fused. The input data set used in this approach is taken from Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). Moreover, this algorithm was also applied to visible-band and infrared sensors as well as the aircraft navigation images.

Image gradient orientation coherence model based fusion [98] provides the solution to handle strong highlights and remove self-reflections from flash and ambient images [99]. This model seeks to utilize the properties of image gradients that remain invariant under the change of lighting that takes place between a flash and an ambient image.

The support value computed from Mapped Least Squares Support Vector Machine (MLS-SVM) can be used as an indicator of salient features of image, which could be used for multi-sensory data fusion [100]. SVM is a recently proposed powerful tool for data classification and function estimation, and can be used for a variety of applications in image processing. This approach is based on the fact that, in SVMs [101-104], the data with larger support values have a physical meaning in the sense that they reveal relative more importance of the data points for contributing to the SVM model. In practice, in this approach, the SVM is used as multi-resolution transform, which directly provides the salient features of source image. After decomposing input images into the sequence of support value images and the low-frequency components, choose-max method is used to select salient features in the fused image. This approach was developed for multi-focus and multi-sensor image fusion. Le and yang [105] described region segmentation and spatial frequency based multi-focus image fusion. This idea is inspired by the fact that region-based image fusion methods could be more useful to overcome sensitivity to noise, blurring effects and mis-registration than pixel-based fusion methods. In this approach, normalized cuts algorithm method [106] is used for segmentation of temporary fused image obtained by averaging source images, then the source images are segmented according to the segmenting result of the intermediate fused image, and finally reconstruct the fused image from the segmented regions of source images according to their spatial frequencies.

Shutao et al. developed a multi-focus image fusion approach for dynamic scenes [107]. This approach utilizes morphological filtering for focus measurement and the computed focus information is used for finding mate function [108]. As compared to others approaches, instead of using single map to define in-focus region, the concept of trimap is used to definite focused region, the definite defocused region and the unknown region. Finally, the fused image with increased depth of field is obtained by fusing the focused regions of source images together. In addition, weighted non-negative matrix factorization [108] and focal point analysis based multi-focus fusion method [110] has been proposed to preserve feature information in the fused image in which weight for each pixel is computed from the focal point position and the Euclidean distance. However, reproduction of complete details, avoiding introduction of artifacts and handling of dynamic scene in the image fusion problem are important issues and need to be further investigated.

2.8 EXPOSURE FUSION

The fundamental goal of the exposure fusion is to preserve details in both very dark and extremely bright regions without HDR image representation and tone-mapping step. The underlying idea of various exposure fusion approaches [111-116] is based on the utilization of different local measures to generate weight map for preserving details present in the different exposures. Thus it provides convenient and consistent way for preserving details in both brightly and poorly illuminated by skipping the construction of HDR image and use of tone-mapping operators.

Raman and Chaudhuri [111] have utilized BF for the fusion of multi-exposure images, in which appropriate matte is generated based on local texture details for automatic compositing process. Goshtasby [112] proposed exposure fusion method based on weights determined by blending function. An image block is considered best-exposed within an area if it carries more information (i.e. entropy) about the area than any other image blocks. The optimal block size and width of the blending functions was determined using a gradient-ascent algorithm to maximize information content in the fused image. The optimal block size was varied from image to image. Images representing scenes with highly varying reflectance, highly varying surface orientations, and highly varying environmental factors such as shadows and specularities, produce smaller optimal block size.

Recently Tom Mertens et al. [113] addressed the generation of fused image using multi-resolution approach [89] without extending the DR and tone-mapping of the final

image. The method blends multiple exposures in Laplacian-pyramid code based on quality measures like saturation and contrast. Moreover, flash image can be included in the exposures to enhance the details in the fused image [113]. The performance of this multi-scale technique is dependent on the number of decomposition levels, i.e. the pyramid height. For better performance, larger images would have to be processed in a higher number of pyramid levels than smaller images. An automated exposure fusion approach using optimization of the pyramid height is the one of the recently proposed solution by Kartalov et al. [117].

As compared to others existing exposure fusion approaches, Kotwal and Chaudhuri [118] proposed a new fusion alternative in which optimization technique is attempted to estimate the best possible matte, which act as weight for the fusion purpose. In this approach, multi-objective cost function is developed, which provides an iterative solution using the variational method. The mates are adaptively derived from the data and the corresponding fused image at every iteration. Another matte-less solution was published by Raman and Chaudhuri [119], which used unconstrained optimization problem for the selection of locally high contrast pixels. In another approach, gradient information based visibility and consistency measures are utilized to merge static and dynamic scenes, and concluded that image gradients convey important information about the latent scene [120].

Zhao et al. proposed [121] the QMFs as analysis tool for decomposing the source images into different frequency sub-bands. In order to modify different frequency sub-bands, the weight maps were computed according to the image appearance measurements, such as exposure, contrast and saturation, and to reduce remove the halo artifacts introduced by the sub-band analysis stage, the gain control maps were estimated based on the property of HVS. A pioneer research on combining images of text-documents taken with different exposures is published by Block et al. [122] in which entropy and local edge intensity were used to select well-exposed pixels across source images. In order to evaluate the performance of their approach, Optical Character Recognition System (OCRS) Tesseract Version 2.03 [123] were used as a benchmark. Experimental results showed that the recognition rate was improved from 0.46 to 0.64 by employing exposure blending as preprocessing step to an OCRS. They have further experimented that the proposed approach of blending high-pass filtered images instead of original images further increases the recognition rate to 0.95.

Over the years, various fast and effective weighted average based exposure fusion approaches have been proposed. Among these guided filtering [124] based two-scale decomposition fusion approach [115], global optimization using Generalized Random Walks

(GRW) [124-126] for fusion [116], and median filter and recursive filtering [128] based fusion approach [114], are producing fusion results with better quality. A more modern approach for dynamic environment was proposed by Li and Kang [114], in which histogram equalization [129] and motion estimation using median filter [130] are included.

Recently, Szeliski [131] produces fused image with improved uniformity in exposure and tone based on simple averaging the pixel brightness levels across auto-bracketed shots. Multi-dimensional histogram was used to analyze a set of bracketed images that projects pixels on to a curve that fits the data. Histogram equalization was used as post-processing operator for optimal contrast enhancement in the fused image.

Unfortunately, these methods do not work so well for preserving and enhancing fine details to produce sharper fused image. An importance quadratic optimization-based method [132,133] to extract fine details from a vector field was proposed by Li et al. [100]. Their method adds the extracted fine details to an intermediate LDR image which is obtained by simply using an existing exposure fusion method [113]. However, to date, it is still rare to find an accurate, robust, and automatic detail enhanced exposure fusion method. The most existing exposure fusion methods are unable to preserve fine details and strong edges simultaneously and are specifically designed for one particular goal.

The previous methods were utilizing different image feature for weight calculation and further refined weight were used to control the contribution of pixels from input exposures. Instead, this thesis uses edge-preserving filters which are effective for two-scale decomposition, weight map generation based on image feature such as weak textures, and weight map refinement. The work proposed in this thesis experiments that the information in the resultant fused image can be controlled with the help of proposed free parameters. Moreover, the present thesis is focused on presenting methods that can also be applied to fuse multi-focus images and flash/no-flash image pair.

The next section briefly reviews the edge-preserving filters and their applications in the field of image processing and computer graphics.

2.9 EDGE-PRESERVING FILTERS AND MULTI-SCALE DECOMPOSITION

Digital filters are used in variety of applications, including speech, image processing and video processing. Digital filters can either be linear (also called isotropic) or non-linear (also called adaptive filters). The goal of the adaption process through certain criterion

function is to extract useful components and remove undesired components. Linear filter apply the spatially invariant function to all pixels of the image, i.e. compute the output pixel value by using the weighted sum of the pixel values in a local window. These types of filters are also known as convolution filters. The average (mean) filtering obtain a de-noised pixel by moving through the image pixel by pixel, replacing pixel in interest with the mean of neighboring pixels, including itself. The main feature of these linear filters is that they smooth Gaussian noise perfectly, but they smear image edges [134]. Furthermore, in image processing field, high frequency information (edges) carries very important information for visual perception in the early stages of the HVS visual information processing. Therefore, the requirement of treating each pixel adaptively has led to an extensive set of non-linear filtering techniques, which adapts to the local image content.

Various asymptotic approaches have been published which are helpful in removing impulses of short duration (non-Gaussian) while preserving edges. Order Statistics Filters (OSF) were recognized as an effective alternative to the linear filters [135-137] which perform better where edge preservation is important. OSF is the most fundamental tool used as signal-adaptive filter [138] that has the ability to locally adapt the filtering to the image content. Median Filter (MF) [139, 140] was proposed in 1974 by Tukey [141] in which a de-noised pixel is computed by replacing the median of surrounding neighborhoods. MF is known as non-average filter because non-weighted average approach is employed on the neighborhood to control the smoothing behavior. In practice, a 3×3 square neighborhood is preferred, larger neighborhoods will produce more severe smoothing [142]. Since MF was proposed, various improvements on it have been developed [143-147]. These approaches improve the performance of MF by increasing signal preservation (edges) in the expense of noise reduction. The general idea of these approaches is that they control the smoothing behavior through its weighting scheme and weights are modified by the local information content.

A Review and comprehensive study of spatial domain, transform domain, and dictionary learning based filters was given by Shao et al. [148], in which advantages and disadvantages of each category were discussed and analyzed. This comprehensive study has stimulated new research issues in image de-noising. Another performance investigation of various edge preserving filters, designed for determining edge response of a filter, was introduced by Himayat and Kassam [149]. This evaluation model is based on the mean and variance, in the vicinity of an edge.

Edge-preserving filters have been the traditional method of removing noise from images whilst preserving signal features such as strong edges [149-154]. Moreover, edge-preserving filters have been utilized in several image processing applications such as edge detection [155,156], and image enhancement and noise reduction [157]. Recently, joint BF [98] has been proposed which is effective for detecting and reducing large artifacts such as reflections using gradient projections.

Over the years, a number of non-adaptive MSD techniques have been proposed recently [158-160] and have some limitations. The first one is the introduction of distortions including halos and visible artifacts. Secondly, it fails to preserve edges during the decomposition. The effectiveness of edge preserving image coarsening has been recognized as valuable tool for MSD decomposition. Recently, various MSD [124,133,161-163] based on edge preserving have been widely used by the graphics researchers for the image processing and the computational photography applications. Traditionally, in computational photography techniques, it is desirable to manipulate details at multiple scales that is done using MSD. WLS optimization framework was proposed to decompose an image into a piecewise smooth BL, containing large scale variations in intensity, and a residual DL capturing the smaller scale details in the image [133]. Farbman et al. [133] have employed this tool for HDR tone-mapping, shape and detail enhancement from multi-light image collections, and boosting of the individual scales for detail manipulation. Another MSD approach based on BF was proposed by Fattal et al. [164], in which BF is applied on input images progressively for analysis purpose. To control edge preserving capability of BF, the range parameter was reduced at each iteration for ensuring that edges which were preserved in the previous level would remain preserved in the next one. In this way they were computing a set of difference layer by computing difference between successive levels of these bilateral filtered images. The purpose of this two-layer decomposition is to separate illumination from texture which can be used to enhance texture details and amplify shading detail without introducing halos.

Interactive tools were primarily designed for local adjustments of tonal values, color adjustment, and other visual parameters in an image [165]. These interactive tools offers incredible power for adjusting various parameters of LDR and HDR images, and provides the user with much more creative control than existing tone-mapping algorithms. Another two-scale tone management approach for photographs has been published by Bae et al. [166] in which large scale effects and texture details were controlled independently. This operator utilizes BF [167] to split input image into base and DLs. Histogram matching tool over the

decomposition was used to control tonal distribution. Finally, they have used gradient correction to prevent undesirable modifications which does not generate gradient reversal and halos. In their model, the detail-preserving tone-mapped image from HDR data is obtained by simply scale down the BL [74].

Hussein and Yang [168] proposed an interactive colorization method by introducing a new concept to the BF [167]. The threshold is derived in a Bayesian framework, and the prior used on the wavelet coefficients is the Generalized Gaussian Distribution (GGD) widely used in image processing applications. They proposed that threshold Nikolaou filter [169] is equivalent to the range weights BF, but has better behavior near the edges. In this method, firstly the user selects grayscale image regions manually by directly painting these regions. Then, each selection is automatically expanded from the user's paint brush and aligned with the object boundary.

WLS [133], BF [170], AD [155], and GF [124] are the popular MSD computation techniques. Among these, WLS, GF and AD are the well-posed approaches for preserving edges while the textures are smoothed out. The detailed description of WLS, GF and AD is given in the forthcoming chapters.

BF was first proposed by Tomasi and Manduchi [167] in 1998. The BF is an adaptive smoothing framework that does a weighted sum of the pixels in a local neighborhood; the weights depend on both the spatial domain and the intensity domains which are used to manipulate smooth regions while preserving strong edges. BF based exposure fusion introduced by Raman and Chaudhuri [111] uses the concept of local contrast [167] to preserve edge details.

This chapter has briefly reviewed the background history of imaging technology, HDR imaging, HDR software solutions, sensor design, tone-mapping, inverse tone-mapping, exposure fusion and edge-preserving filters. The work proposed in the present thesis describes an alternate solution of HDR and tone reproduction based on edge-preserving filters. It focuses on preserving and enhancing fine details during the exposure fusion framework. The present thesis work proposes a two-scale decomposition of multi-exposure images that exploits the edge-preserving capability of non-linear filters. It will first discuss three different techniques based on edge-preserving filters for achieving this task while capturing static scenes. Then, it describes some other related applications which can be performed using the methods proposed in the present work.

The present thesis also proposes an extension of the multi-exposure image fusion technique to fuse images captured with varying DOF and flash/no flash photography. To

evaluate the performance of proposed approaches various quality metrics will be used in Chapter 6. Furthermore, the present thesis is focused on analyzing the effect of free parameters used in the variational solutions to the detail-enhanced exposure fusion approaches. The variational solutions to LDR image generation from different exposures captured at variable exposure settings are discussed in the forthcoming chapters.

CHAPTER-III

ANISOTROPIC DIFFUSION FILTER BASED MULTI-EXPOSURE IMAGE FUSION (ADF)

This chapter discusses the utility of edge-preserving capability of AD for the fusion of images captured at different exposure settings. The BLs and DLs are fused separately to preserve texture details. A first step, in ADF algorithm is two-scale decomposition of each input image to extract details at arbitrary scales, based on adaptive and edge preserving filter (i.e. AD) [156].

ADF algorithm takes N identically sized multi-exposure images taken from a fixed viewpoint, and produces output image of the same size, in which well exposed pixel value is computed by combining information from all input images at each scale of the decomposition. Unlike earlier image-based compositing techniques [113], ADF separates coarse scale details (i.e. BL) from fine details (i.e. DL). ADF approach is similar in spirit to the multi-scale shape and detail enhancement from Multi-Light Image Collections (MLIC) approach of Fattal et al. [164]. Therefore, it is effective to manipulate fine and coarse details separately during the compositing process, and needs no further post processing for detail enhancement. After the manipulation of each redundant layer, fused BL, and fused DL are recombined to produce well exposed detail enhanced fused image (please see Figure 3.1). Thus, the magnitude of the BLs is modified based on the decision maps to ensure that resulting fused image contains well-exposed regions, while the magnitude of the DLs is unchanged, thereby preserving fine details. To be able to deal with strong edges separately, ADF use a nonlinear multi-scale edge-preserving image decomposition which permits us to manipulate and combine details at multiple scales without introducing objectionable artifacts.

3.1 INTRODUCTION

The main objective of ADF approach is to preserve and enhance details in both brightly and poorly illuminated areas that significantly improve the quality of the fused image. It must provide optimal contrast within the capabilities of the conventional displaying medium and must not lead to artifacts such as contrast reversal or black halos. Additionally, it should produce realistic and pleasant images. The principal characteristic of ADF is an

adaptive adjustment of local spatial information in the Laplacian pyramid [89] depending on texture features (i.e. local range). To control the contribution of pixels, ADF calculates weight that depends on the maximum and minimum intensities of the neighboring pixels from the pixel under consideration. The weight function and Gaussian-Laplacian pyramid are derived in the following sections. Figure 3.1 shows that the proposed scheme contains three steps, which are *analysis*, *scene detail manipulation based on decision map* and *synthesis*.

More specifically, the goal of ADF algorithms proposed in this thesis is to produce well exposed image by combining the information across all of the input multi-exposure images. In ADF approach, two scale decomposition based on AD [156] is used to separate coarser and finer details from each input image. The BLs and DLs across all N input images are defined as:

$$BL_k^{adf}(x, y) = \mathcal{A}(I_k(x, y)) \text{ and } DL_k^{adf}(x, y) = BL_k(x, y), \quad (3.1)$$

where $k = 1 \dots \dots \dots N$

where $\mathcal{A}(\bullet)$ is the AD (i.e. edge-preserving filter) applied on input exposures.

The detail enhanced fused image generated by ADF is given as:

$$I_F^{adf}(x, y) = BL_F^{adf}(x, y) + DL_F^{adf}(x, y) \quad (3.2)$$

where BL_F^{adf} is the fused BL that maximizes the coarser details across all of the input BLs $BL_1^{adf}, BL_2^{adf}, \dots, BL_N^{adf}$ and DL_F^{adf} is the fused residual (i.e. fused DL) that maximizes finer details across all of the input DLs $DL_1^{adf}, DL_2^{adf}, \dots, DL_N^{adf}$. In the forthcoming section, we briefly describe the AD used in ADF for two-scale decomposition, and the local range measure is proposed to generate weight maps for controlling contribution of pixels from BLs.

3.2 DATA ACQUISITION AND TWO LAYER DECOMPOSITION

3.2.1 SCENE DATA ACQUISITION

Conventional digital photography struggles with the high contrast scenes and can capture brightest part (i.e. highlights) by choosing a low exposure level (i.e. short exposure time) or the darkest part (i.e. shadows) by choosing a high exposure level (i.e. long exposure time). In the present ADF approach, information present in the fused LDR image depends on number of input exposures captured at different exposure settings. It assumes that all input multiple exposure images are photographed from static scene with the help of tripod to avoid

any spatial and global movement. To apply ADF technique successfully, sequence of exposures is captured from a scene with very dark and very bright details. The aperture priority, ISO settings, and the camera's white balance are kept fixed for the entire sequence. Sample of input set of images captured at different exposure settings is illustrated in Figure 1.3.

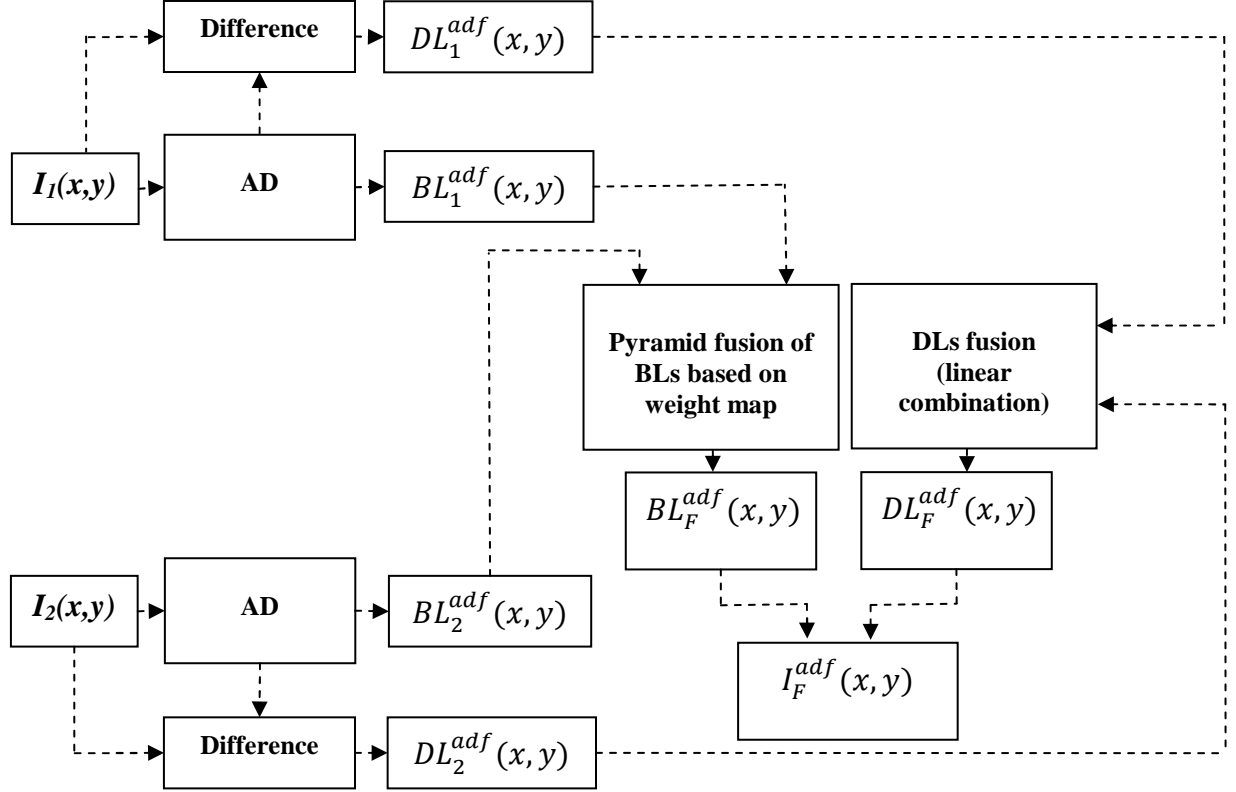


Figure 3.1: Proposed image domain fusion framework. Observation model, illustrates the conceptual framework of ADF approach. Note that for the concept simplicity; here we have generated BLs and DLs of two input exposures.

3.2.2 EDGE PRESERVING PROPERTY OF AD FILTER

AD has led to an efficient new field to remove noise from an image by modifying the image via a PDE. The goal of edge preserving diffusion [156] is to encourage smoothing at homogeneous region in preference to inhomogeneous region (i.e. edge). Mathematically, the isotropic diffusion [156] equation $\partial I(x, y, t_a) = \text{div}(\nabla I)$ is replaced with

$$\frac{\partial I(x, y, t_a)}{\partial t_a} = \text{div}[g(\|\nabla I\|)\nabla I] \quad (3.3)$$

where ∇I is the image gradient, $\|\nabla I\|$ is the magnitude of the gradient of image intensity, $g(\|\nabla I\|)$ is an “edge-stopping function” or “conduction coefficient” that controls the diffusion strength, (x, y) specifies spatial position, and t_a is the process ordering time parameter.

The diffusion strength in the image is influenced by the conduction coefficient which depends on the magnitude of the gradient of the image intensity. The process of gradient computation from the neighbors in 1-D and 2-D structure is illustrated in Figure 3.2(a) and 3.2(b), respectively. If the conduction coefficient is replaced by a constant value (i.e. $g(\bullet) = 1$), the diffusion process will be isotropic diffusion. Since isotropic diffusion does not consider image structure, fine textures as well as edges are smoothed out. Thus in AD, the conduction coefficient is chosen to satisfy $g(x) \rightarrow 0$ when $x \rightarrow \infty$ so that the diffusion process is “stopped” across the region boundaries (i.e. edges) at locations of high gradients.

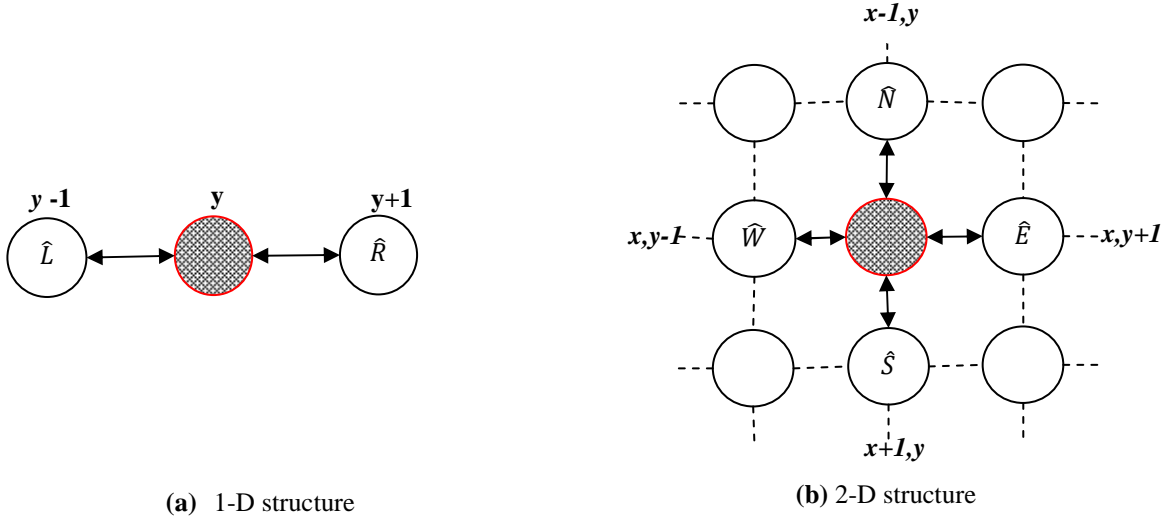


Figure 3.2: Gradient computation: **(a)** One dimensional (1-D) grid structure: The signal flow is calculated between two neighboring nodes (solid connection). **(b)** Two dimensional (2-D) grid structure: The signal flow is calculated between four neighboring nodes (solid connection).

Two different diffusion functions $g(\bullet)$ have been proposed by Parona et al. [156], which result in edge preserving filter defined as:

$$g_1(\nabla I) = e^{\left(-\left(\frac{\|\nabla I\|}{K}\right)^2\right)} \quad (3.4)$$

$$g_2(\nabla I) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{K}\right)^2} \quad (3.5)$$

where \mathcal{K} is a scale parameter (i.e. constant) to be tuned for a particular application. Perona et al. [156] proposed that the value of \mathcal{K} can be fixed manually or using the “noise estimator” described by Canny [171]. Note that the conduction coefficient depends on the magnitude of the gradient of image intensity $\|\nabla I\|$ at (x,y) spatial position. If $\|\nabla I\|$ is small, then $g(\bullet)$ is close to 1; and if $\|\nabla I\|$ is large, then $g(\bullet)$ is small. Hence, $g(\bullet)$ is chosen giving a much higher smoothing at homogeneous regions as compared to edges. The parameter \mathcal{K} controls how quickly the weight of a training example falls off; \mathcal{K} is called the hypothetical scale parameter, and is also something that we’ll get to experiment with in proposed work.

In ADF approach fine details are separated using Equation (3.4), which favors high contrast sharp transitions across multi-exposure input series and the value of $\mathcal{K} = 1/7$ was fixed manually based on experimentation.

The discrete formulation of Perona and Malik [156] AD (i.e. BL in ADF method) is as given by:

$$I_{x,y}^{t_a+1} = I_{x,y}^{t_a} \frac{\lambda}{|\eta_{x,y}|} \sum_{\beta \in \eta_{x,y}} g(\nabla I_{x,y,\beta}) \nabla I_{x,y,\beta} \quad (3.6)$$

where $I_{x,y}^{t_a}$ is a discrete version of input signal, x,y determine the sample position of current pixel in the discrete signal, β determines the neighboring pixel, t_a determines iterations. The constant λ is a scalar that determines the rate of diffusion, $\eta_{x,y}$ represents the spatial neighborhoods of current sample position x,y , and $|\eta_{x,y}|$ is the number of neighbors.

To see the behavior of the Perona et al. [156] filter at edges, we first analyze one dimensional signal into BL and DL. As can be seen in Figure 3.3, at BL (i.e. the coarser level after diffusion), high-frequency textures disappear. The weak texture details lost at the BL are exactly reconstructed at the DL. However, DL is the difference between the input signal and the BL, which is dominated by the large discontinuities characterized by the rapid oscillations (high frequency variations) in the input signal. As a result, AD is able to separate weak texture details from edge transitions that are to be preserved during the fusion process. The continuous diffusive process for 1-D network structure (see Figure 3.2 (a)) is as follows:

$$\frac{\partial I(y, t_a)}{\partial t_a} = \text{div}[g(\|\nabla I\|)\nabla I] \quad (3.7)$$

and the discrete formulation is written as:

$$I_y^{t_a+1} = I_y^{t_a} + \frac{\lambda}{|\eta_y|} [g_{\hat{L}} \cdot \nabla_{\hat{L}} I + g_{\hat{R}} \cdot \nabla_{\hat{R}} I]_y^{t_a} \quad (3.8)$$

where $I_y^{t_a}$ is a discrete version of input signal, y determines the sample position in the discrete signal, t_a determines iterations. ADF finds that one iterations ($t_a=1$) to be sufficient for DL extraction across all of the input images experimented in the thesis.

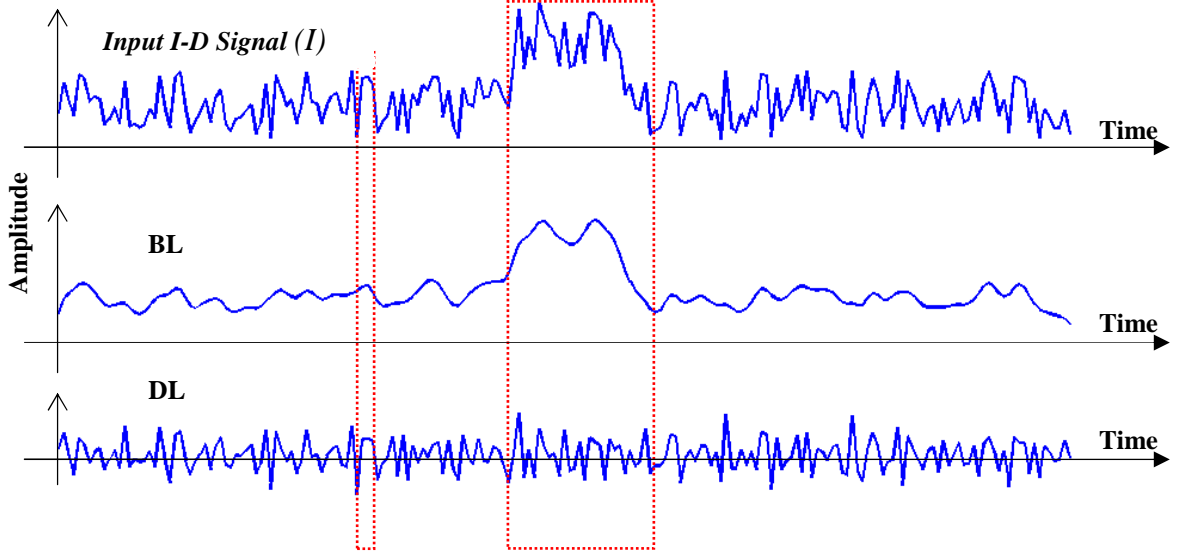


Figure 3.3: Two layer decomposition of 1-D signal based on AD after 5 iterations with $K=30$, $\lambda=1/3$ and $|\eta_y|=2$ (left & right). The 1-D input signal (I) is decomposed into two main components: a low frequency BL and a high frequency DL. Notice that the edges are preserved in the diffused image (i.e. BL) and the DL yields fine details only.

The detailed analysis of affect of number of iteration on computational time and information present in the fused image is given in the results and analysis section. The constant λ is a scalar that determines the rate of diffusion, η_y represents the spatial neighborhood of current sample position y , the subscripts \hat{L} and \hat{R} depicting left and right respectively, and $|\eta_y|$ is the number of neighbors (i.e. two in 1-D case). Where $g_{\hat{L}}$, and $g_{\hat{R}}$ are the conduction coefficients across left and right spatial locations respectively. The symbol $\nabla_{\hat{L}}$ and $\nabla_{\hat{R}}$ indicates the difference of left and right neighbor respectively:

$$\nabla_{\hat{L}} I_y \equiv I_{y-1} - I_y \quad (y \text{ determines sample location in 1-D grid}) \quad (3.9)$$

$$\nabla_{\hat{R}} I_y \equiv I_{y+1} - I_y \quad (3.10)$$

The AD of 2-D grid shown in Figure 3.2 (b) is given by the relation:

$$BL^{adf}(x, y) = I_{x,y}^{t_a+1} = I_{x,y}^{t_a} + \frac{\lambda}{|\eta_{x,y}|} [g_{\hat{N}} \cdot \nabla_{\hat{N}} I + g_{\hat{S}} \cdot \nabla_{\hat{S}} I + g_{\hat{E}} \cdot \nabla_{\hat{E}} I + g_{\hat{W}} \cdot \nabla_{\hat{W}} I]_{x,y}^{t_a} \quad (3.11)$$

where $I_{x,y}^{t_a}$ is a discrete version of input image I , x, y determines the pixel position in the discrete image, t_a determines iterations. The constant λ is a scalar that determines the rate of diffusion, $\eta_{x,y}$ represents the spatial neighborhood of current pixel x, y (North, South, East

and West), and $|\eta_{x,y}|$ is the number of neighbors (usually four). Where $g_{\hat{N}}$, $g_{\hat{S}}$, $g_{\hat{E}}$ and $g_{\hat{W}}$ are the conduction coefficients across North, South, East and West spatial locations. The symbol $\nabla_{\hat{N}}$, $\nabla_{\hat{S}}$, $\nabla_{\hat{E}}$ and $\nabla_{\hat{W}}$ indicates the difference of North, South, East and West neighbor respectively:

$$\nabla_{\hat{N}} I_{x,y} \equiv I_{x-1,y} - I_{x,y} \quad (x,y \text{ determines pixel location in 2-D grid}) \quad (3.12)$$

$$\nabla_{\hat{S}} I_{x,y} \equiv I_{x+1,y} - I_{x,y} \quad (3.13)$$

$$\nabla_{\hat{E}} I_{x,y} \equiv I_{x,y+1} - I_{x,y} \quad (3.14)$$

$$\nabla_{\hat{W}} I_{x,y} \equiv I_{x,y-1} - I_{x,y} \quad (3.15)$$

$$DL^{adf}(x,y) \equiv I(x,y) - BL^{adf}(x,y) \quad (3.16)$$

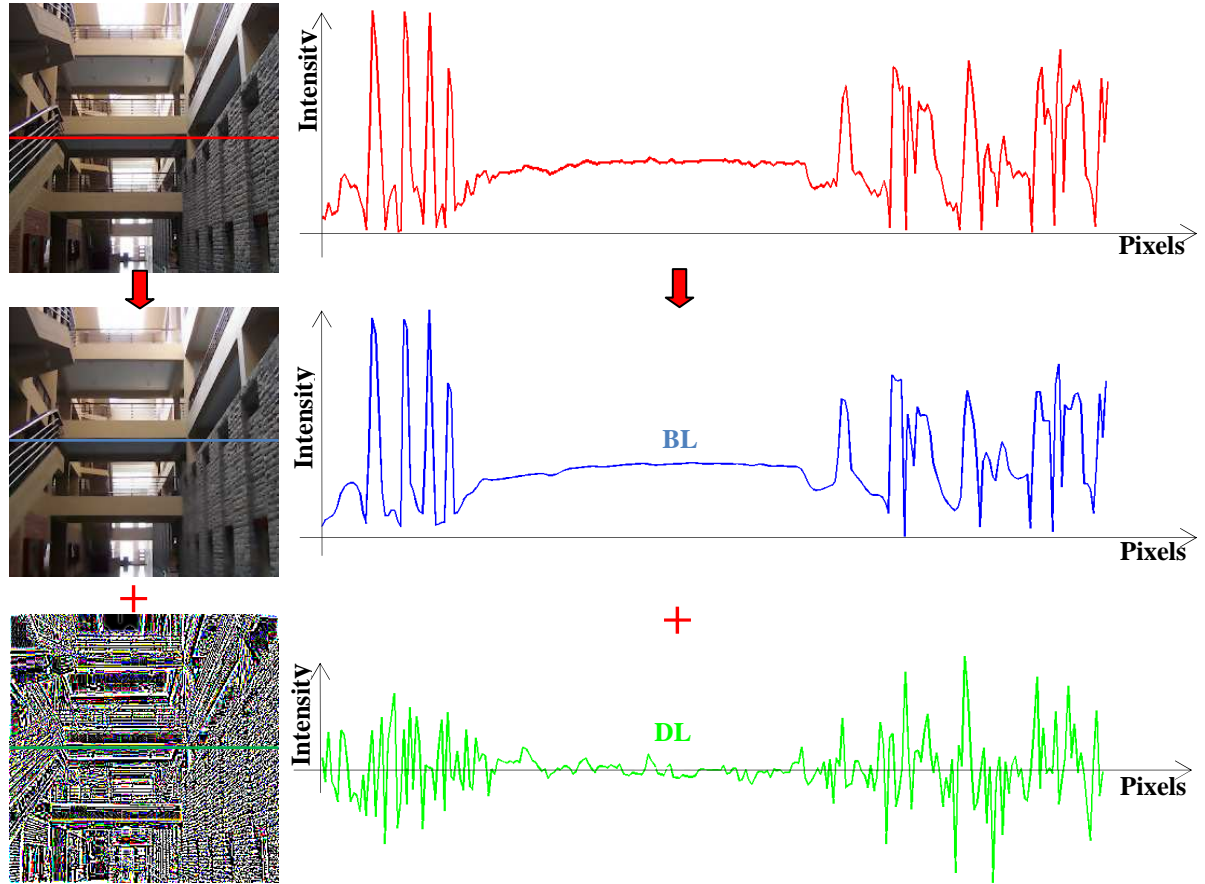


Figure 3.4: Two layer decomposition of 2-D signal (i.e. **JUIT** image) based on AD (left) after five iterations with $K=30$, $\lambda=1/7$ and $|\eta_{x,y}|=4$ (region size of 3×3 pixels). Intensity profiles (right) along a scan lines of 2-D input signal (red), BL (blue) and DL (green). Notice that the strong edges are preserved in the diffused image (i.e. BL) and the DL yields fine details only. Details compressed in the BL are exactly reconstructed in DL.

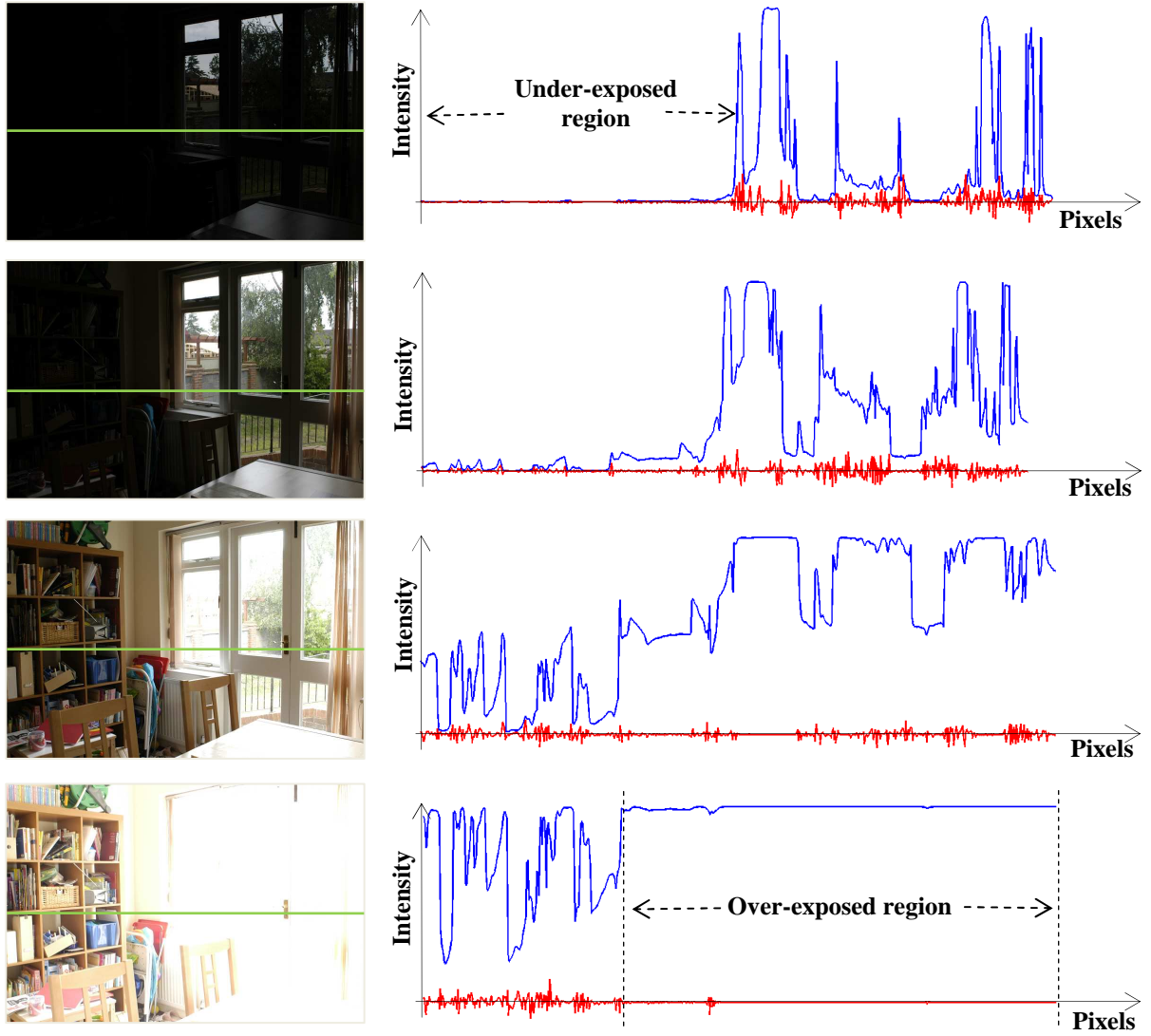


Figure 3.5: Intensity plots (right) along a scan lines of BLs (blue) and DLs (red) obtained with the AD after five iterations with $K=30$, $\lambda=1/7$ and $|\eta_{x,y}|=4$ (region size of 3×3 pixels) across all of the input exposures (left).

Notice that coarser and finer details are extracted across the visible details adaptively when the scene is captured at different exposure settings.

The BL decomposition using Equation (3.11) and DL decomposition using Equation (3.16) of JUIT image is illustrated in Figure 3.4. From Figure 3.4, it can be visually seen that the BL provides coarse details and the textures are almost eliminated. In Figure 3.5, it is illustrated that the intensity profiles of BLs (blue color) and DLs (red color) computed across the images captured at different exposure settings. It is noticed that coarser and finer details are extracted across the visible details adaptively when the scene is captured with variable exposure times.

3.3 WEIGHT MAP COMPUTATION: TEXTURE FILTER BASED ON LOCAL RANGE

In ADF approach, the local range is used to generate weight map function for non-uniform scaling to control contribution of pixels from the BLs computed across all input exposures. In Figure 3.6, it is illustrated that how local range is calculated in the range-filtered image from 3-by-3 neighborhood. This local range is likely to be very different from region to region in a different images captured at variable exposure settings. Well exposed areas will yield higher local range as compared to the over exposed and under exposed regions, which is illustrated in Figures 3.7 and Figure 3.8. Local range is defined as follows:

$$R_{x,y,k}^{adf} = L_{max} - L_{min} \quad (3.17)$$

where L_{max} and $L_{min} \in$ local spatial window at x,y (i.e. 3-by-3) in the k_{th} BL.

The final normalized weight map functions used for the fusion of BLs are computed as follows:

$$W_{x,y,k}^{adf} = \left[\sum_{k=1}^N R_{x,y,k}^{adf} \right]^{-1} R_{x,y,k}^{adf} \quad (3.18)$$

where L_{max} and L_{min} are the maximum and minimum values of the neighboring pixels within a 3-by-3 square window, respectively, and $W_{x,y,k}^{adf}$ is the final weigh map computed in ADF approach at location (x,y) for k_{th} BL.

It is commonly accepted that the higher the luminance variation region is, the stronger the local range of that region to shield a pixel is. However, it is found that the difference between maximum and minimum values of luminance within a local window influences the probability of shielding of well-exposed pixel. To compute such local range, the basic idea is illustrated in Figure 3.6. To illustrate the effect of EV variation on local range computed across multiple exposures, we give four representative images as shown in Figures 3.8 (a-d).

Once the final weight maps are computed by Equation (3.18), the Gaussian pyramid of $W_{x,y,k}^{adf}$ is used to reduce the influence of very high intensities and very low intensities present in the BLs for producing high-resolution fused image, which is described in the following section.

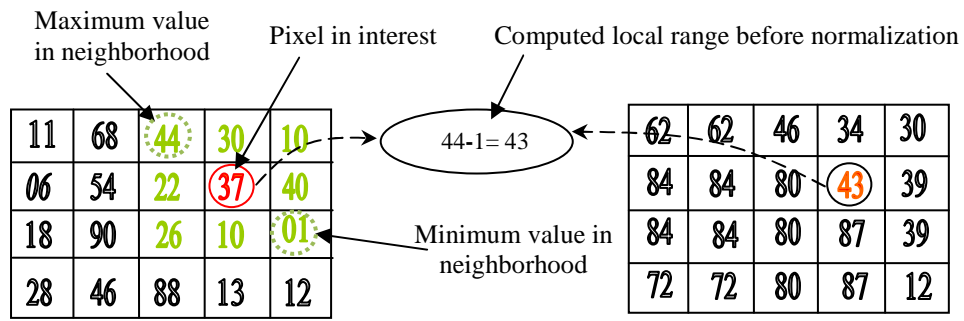


Figure 3.6: Illustration of local range calculation that used as location adaptive weight map. The figure depicts how local range is calculated in the range-filtered image from 3-by-3 neighborhood. Green numbers are the neighbors considered to compute local range for the pixel in interest (i.e. displayed in red number).

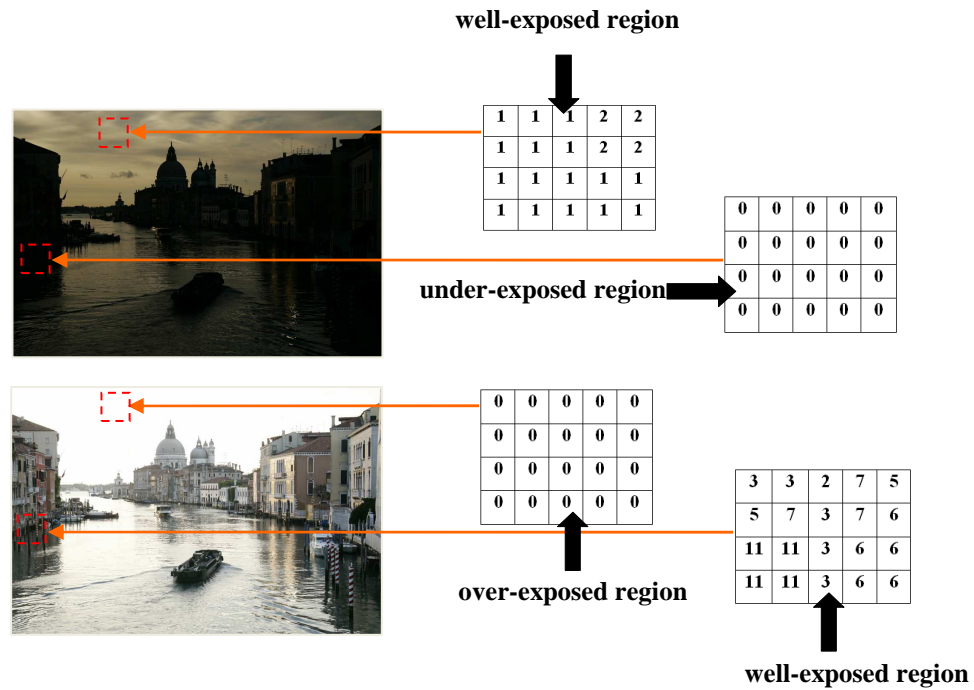


Figure 3.7: Effect of under exposed, normally exposed, and over exposed regions on local range. Note that local range will be different (i.e. zero for under exposed and overexposed regions) for the same region under different exposure settings. The optimal window size for range calculation is 3-by-3 and the numerical values shown in the box are calculated for small local window.

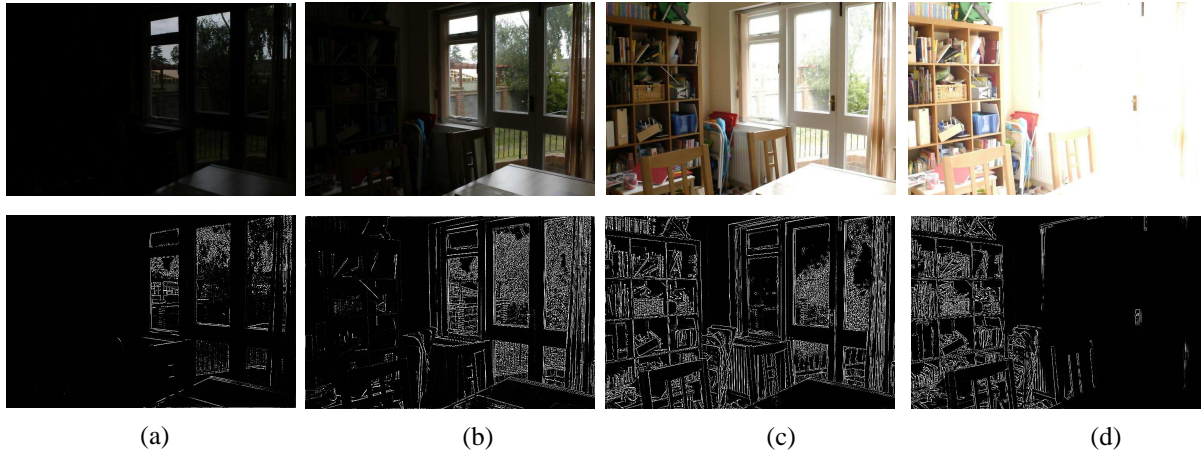


Figure 3.8: Illustration of local range analysis of BLs across the multiple exposures. The local ranges are varying with respect to the different exposure times. ((a), (b)) BLs of underexposed images (top) and their corresponding texture features (bottom); (c) BL of normally exposed image (top) and their corresponding texture features (bottom); (d) BL of overexposed image (top) and their corresponding texture features (bottom). Note that well-exposed pixels have the brighter texture features (i.e., higher weights) as compared to under-exposed and over-exposed pixels. Input images: Jacques Joffre HDR Chief Photographer.

3.4 PYRAMID GENERATION AND CONSTRUCTION OF FUSED BL ACROSS ALL INPUT BLs

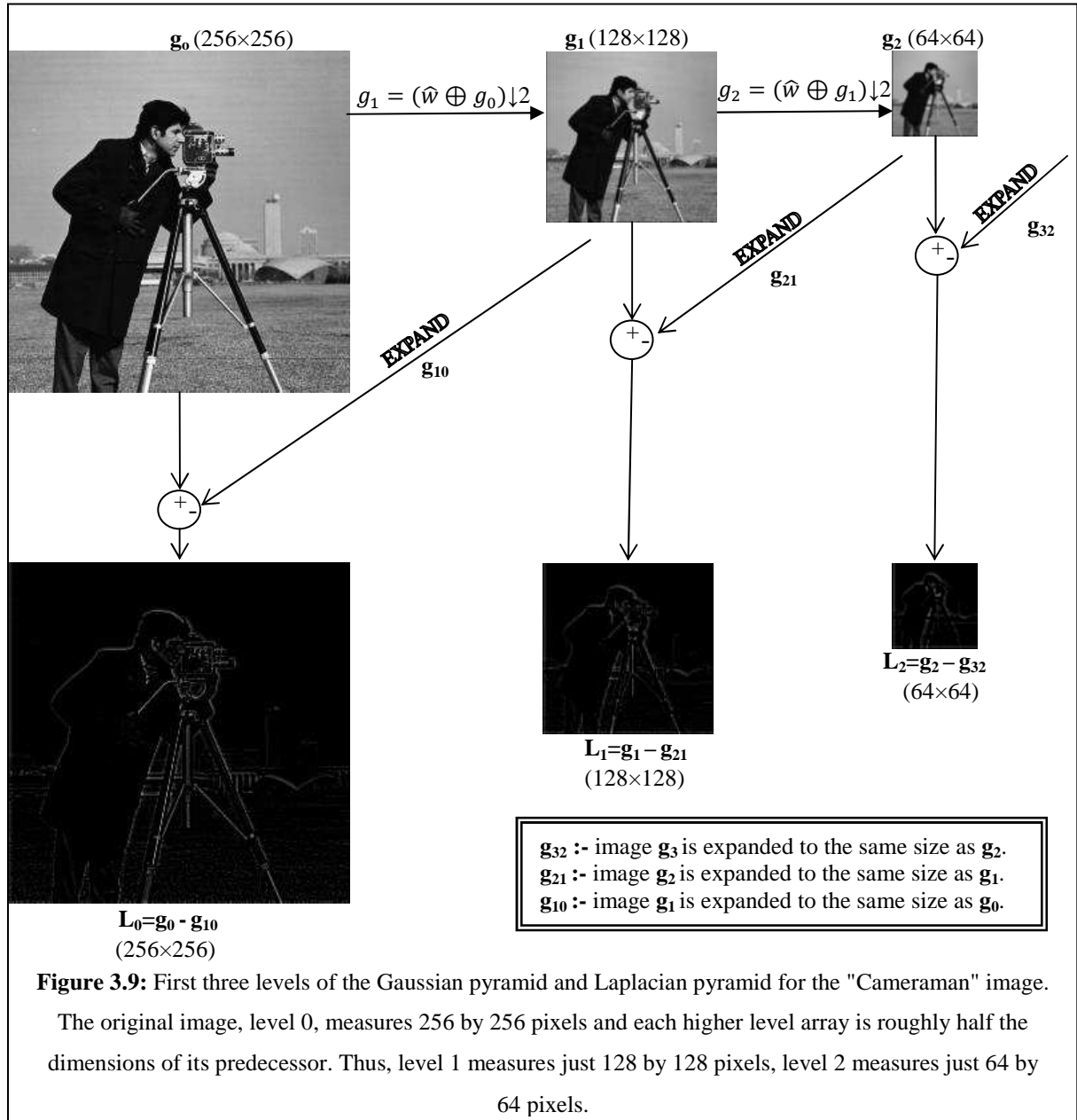
3.4.1 PYRAMID DECOMPOSITION

Pyramid decomposition is a fundamental tool of image encoding. This multi-resolution encoding process decomposes images into several spatial resolutions, to extract features. Each level of the pyramid corresponds to different information such as edge details which is called multi-scale edge estimation. Pyramid decomposition combines the features of predictive and transforms methods [89]. It is being used in variety of image processing and computer graphics applications today. Initially, it was proposed for analysis and progressive transmission of images.

3.4.1.1 GAUSSIAN PYRAMID DECOMPOSITION

The first step in pyramid coding is to convolve original image g_0 with the unimodal Gaussian-like weighting function (generating kernel i.e. a compact two-dimensional filter for image) that yields low-pass filtered image g_1 . Each pixel value within level 1 is computed as a weighted average of pixel values in level 0 within a 5-by-5 window. Then the down-

sampling by a factor of 2 is performed to get the reduced image g_1 at level 1. Therefore, pyramid level 1 contains image g_1 , which is a low-pass filtered and down-sampled version of g_0 . g_1 is then filtered and sub-sampled in the same way to obtain g_2 . By repeating these steps several times we obtain a sequence of stack of successively smaller images $g_0, g_1, g_2, \dots, g_d$, called Gaussian pyramid. The motivation behind such type of decomposition is to produce the sequence of reduced resolution images g_1, g_2, \dots, g_d . A graphical representation of three levels Gaussian pyramid decomposition of Cameraman image is illustrated in Figure 3.9.



The process of pyramid generation can be continued until the image size is 2-by-2 pixel or one pixel. In this way, each successive level of Gaussian pyramid corresponds to

lower frequencies. The possible numbers of Gaussian levels for “Cameraman” image are shown in Figure 3.10.

Let l_k denote a l_{th} level of Gaussian pyramid of k_{th} image and let the averaging process is performed by the function REDUCE [89].

$$g_l = \text{REDUCE}(g_{l-1})$$

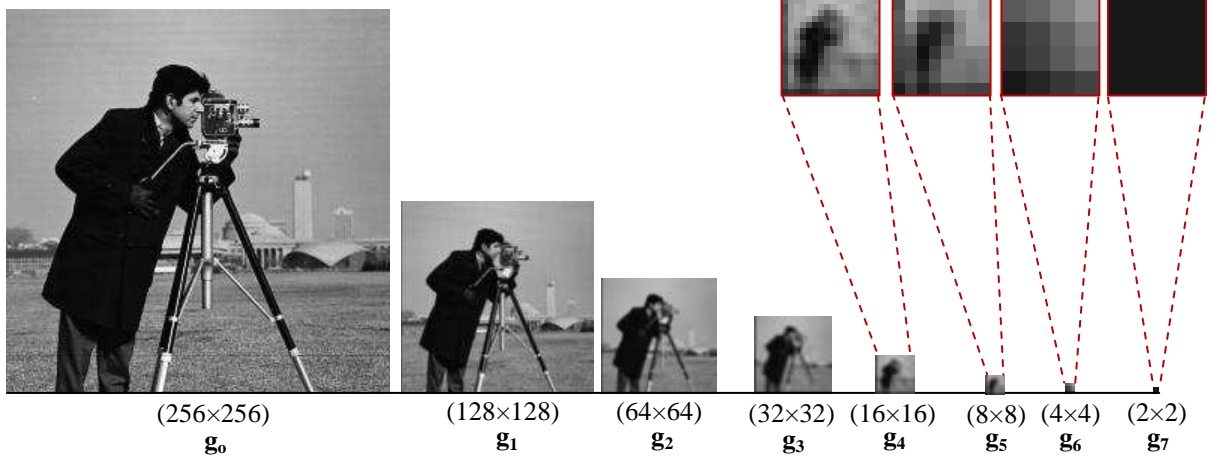


Figure 3.10: Illustration of possible number of levels in Gaussian pyramid of Cameraman image.

For levels $0 < l < d$, the l_{th} Gaussian level is given by

$$g_l(x, y) = \sum_{m=-2}^2 \sum_{n=-2}^2 \hat{w}(m, n) g_{l-1}(2x + m, 2y + n)$$

Here d represents the number of levels and $\hat{w}(m, n)$ refers to the generating kernel or weighting pattern. For images, 5-by-5 generating kernel of weights \hat{w} is utilized to generate each pyramid array from its predecessor. This generating kernel, is chosen subject to certain constraints [172]. For simplicity we make \hat{w} separable:

$$\hat{w}(m, n) = \hat{w}_s(m) \hat{w}_s(n)$$

The one-dimensional, length 5, function \hat{w}_s is normalized:

$$\sum_{m=-2}^2 \hat{w}_s(m) = 1 \text{ and } \sum_{n=-2}^2 \hat{w}_s(n) = 1$$

and symmetric:

$$\hat{w}_s(x) = \hat{w}_s(-x) \text{ for } i = 0, 1, 2.$$

Therefore, $\hat{w}_s(1) = \hat{w}_s(-1)$ and $\hat{w}_s(2) = \hat{w}_s(-2)$.

An additional constraint is called equal contribution. This stipulate that all nodes at a given level must contribute the same total weight ($=1/4$) to nodes at the next higher level. Let

$\hat{w}_s(0) = a$, $\hat{w}_s(1) = \hat{w}_s(-1) = b$, and $\hat{w}_s(2) = \hat{w}_s(-2) = c$. In this case equal distribution requires that $a + 2c = 2b$. These three constraints are satisfied when:

$$\hat{w}_s(0) = a$$

$$\hat{w}_s(-1) = \hat{w}_s(1) = 1/4$$

$$\hat{w}_s(-2) = \hat{w}_s(2) = \frac{1}{4} - a/2$$

Thus, pyramid decomposition can be obtained by convolving the image g_0 with a set of “equivalent weighting functions” h_l .

$$g_l = h_l \oplus g_0$$

$$g_l(i, j) = \sum_{m=-M_l}^{M_l} \sum_{n=-M_l}^{M_l} h_l(m, n) g_0(2^l x + m, 2^l y + n)$$

where M_l is the size of equivalent function which does double from one level to the next level. The equivalent functions double in width with each level.

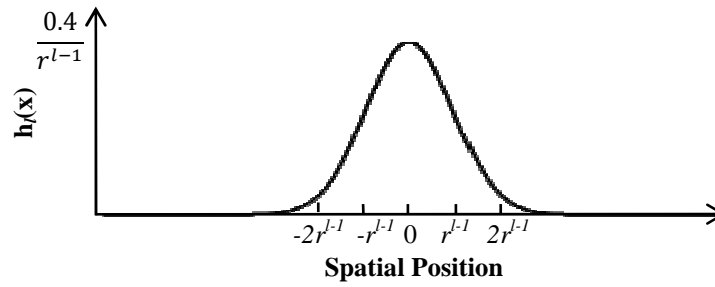


Figure 3.11: The equivalent weighting functions $h_l(x)$ for nodes in Gaussian pyramid level l using $a = 0.4$.

Note that the resulting equivalent weighting functions closely resemble the Gaussian probability density functions.

The equivalent weighting function for Gaussian pyramid level l is shown in Figure 3.11, in which the value of a is chosen 0.4 to produce Gaussian-like shape.

3.4.1.2 LAPLACIAN PYRAMID DECOMPOSITION

The Laplacian pyramid comprises band-pass images known as error images which are obtained directly by taking the difference between the two adjacent levels of Gaussian pyramids. Since these levels differ in their sample density it is necessary to interpolate new sample values between those in a given level before that level is subtracted from the adjacent level. Therefore, Laplacian pyramid can be defined as the difference-of-lowpass filtered

images. Expanding g_l to the same size as g_0 and subtracting yields the band-passed image L_0 .

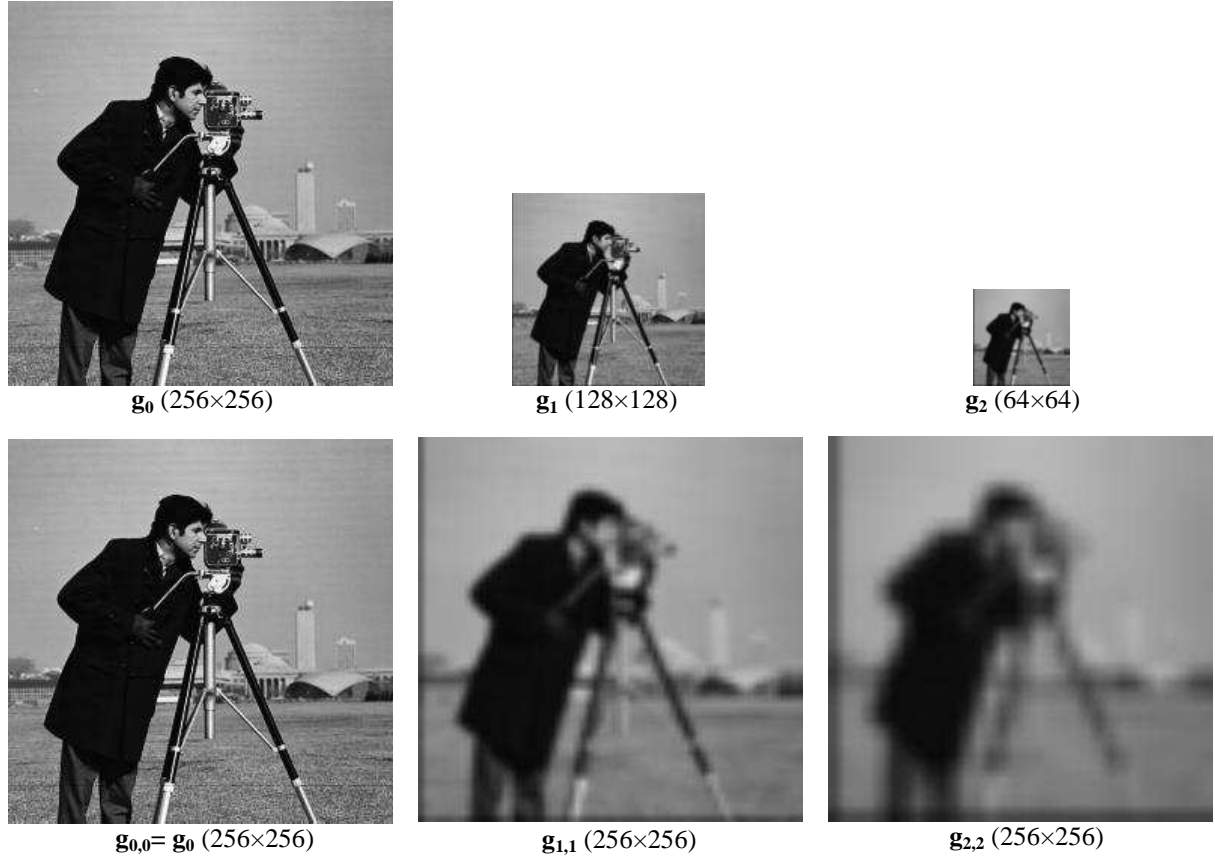


Figure 3.12: Illustration of Gaussian pyramid expanded to the size of the original image.

A Laplacian pyramid $L_0, L_1, L_2, \dots, L_{d-1}$, can be built containing band-passed images of decreasing size and spatial frequency.

Let L_l denote a l_{th} level of Laplacian pyramid and let the interpolating process is performed by the function EXPAND [89].

$$L_l = g_l - \text{EXPAND}(g_{l+1}), 0 < l < d$$

Since there is no image g_{d+1} to serve as the prediction image for g_d , we say $L_d = g_d$.

Where the expanded image g_{l+1} is given by:

$$g_{l+1}(x, y) = \sum_{m=-2}^2 \sum_{n=-2}^2 \hat{w}(m, n) g_l \left(2x + \frac{m}{2}, 2y + \frac{n}{2} \right)$$

The original image can be reconstructed from the expanding band-pass images (each level of Laplacian pyramid expanded to the size of the original image) and summing:

$$\text{Original Image } (g_0) = L_{0,0} + L_{1,1} + L_{2,2} + \dots + L_{d,d}, \quad \text{with } L_{d,d} = g_{d,d}$$

Each Gaussian level contains the lowpass version of the original image of certain spatial frequencies. The zero level contains all spatial frequencies while the highest Gaussian

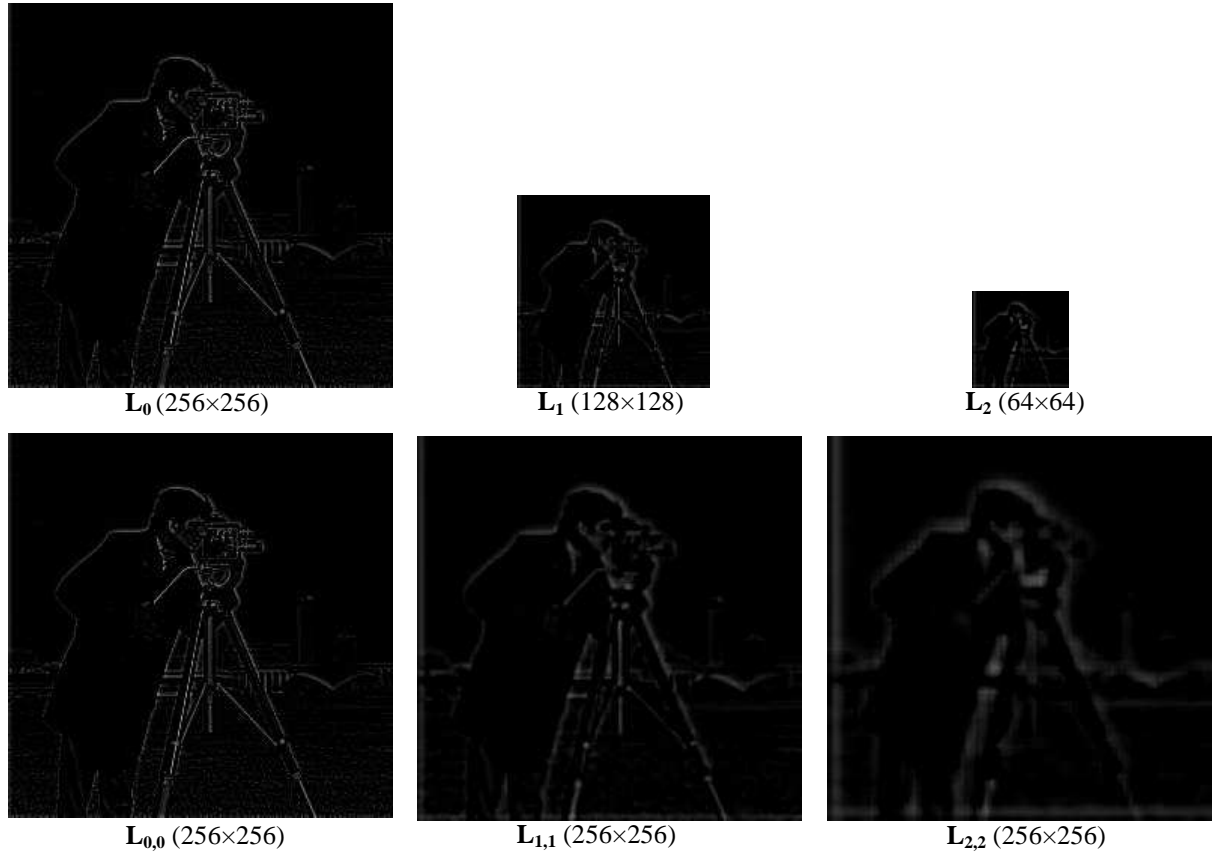


Figure 3.13: Illustration of Laplacian pyramid expanded to the size of the original image.

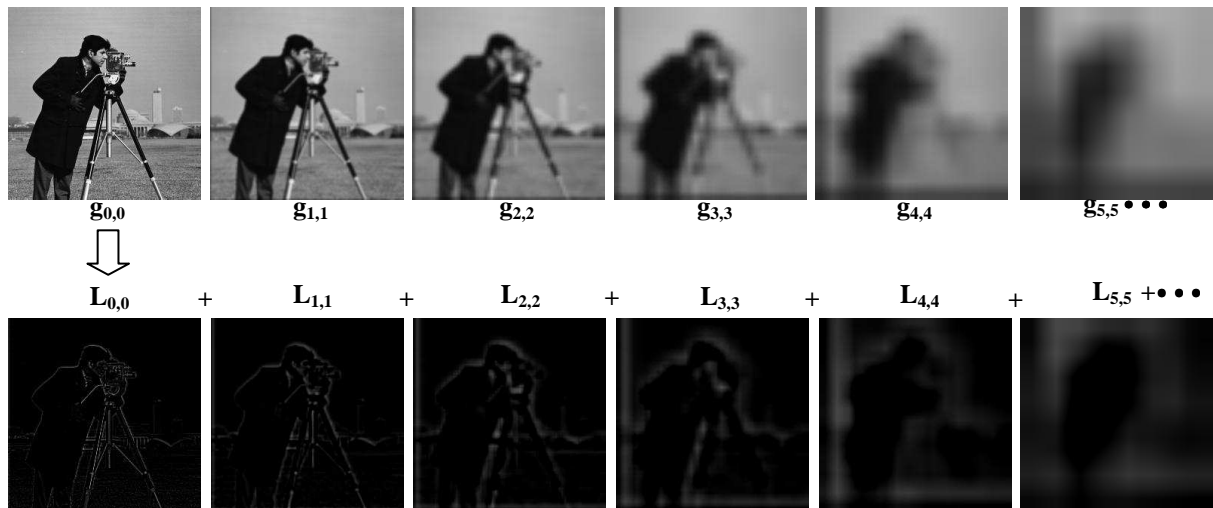


Figure 3.14: Reconstruction of original image from several spatial frequency bands: first six levels of the Gaussian and Laplacian pyramid. Gaussian images, upper row, were obtained by expanding pyramid arrays through Gaussian interpolation. Each level of the Laplacian pyramid is the difference between the corresponding and next higher levels of the Gaussian pyramid.

level contains lowest spatial frequencies. The reduction of spatial frequencies across different Gaussian levels can be clearly seen when the Gaussian pyramid "levels" are expanded to the same size as g_0 (Figure 3.12). Since the Laplacian pyramid is obtained from the adjacent Gaussian levels, it consists of band-passed copies of original image, at progressively weaker edge details. To facilitate comparisons, the expanded Laplacian pyramid levels to the size of the original image for the "Cameraman" image are shown in the bottom row of Figure 3.13. Note that image features such as edges and bars appear enhanced in the Laplacian pyramid. Fine edge details are prominent in $L_{0,0}$, while progressively coarser edge details are prominent in the higher level images. Therefore, each Laplacian level contains the "edges" of a certain size, and spans approximately an octave in spatial frequency.

Since each Laplacian level is a good estimate of band-pass copies (spatial frequency bands) of an original image. Therefore original image can be recovered from these Laplacian levels by expanding and then summing. Figure 3.14 shows an image represented as a sum of several spatial frequency bands.

3.4.2 BASE LAYERS FUSION

Researchers have attempted to synthesize and manipulate the features at several spatial resolutions [113,173] that avoid the introduction of seam and artifacts such as contrast reversal or black halos. In ADF approach, the band-pass components of $BL_{x,y,k}^{adf}$ at different resolutions are manipulated with the help of $W_{x,y,k}^{adf}$ that determines pixel value in the reconstructed BL_F^{adf} . The pyramid representation of BL expresses an image as a sum of spatially band-passed images while retaining local spatial information in each band.

ADF approach begins by constructing a Gaussian pyramid $GB_k^{adf,0}$, $GB_k^{adf,1}$,, $GB_k^{adf,d}$ of k_{th} BL computed across the input exposure, where $GB_k^{adf,0}$ is the full resolution BL obtained from AD, and $GB_k^{adf,d}$ is the coarsest level of the k_{th} BL in the pyramid. Each level of Gaussian pyramid is the low pass filtered and sub-sampled version of the predecessor. Lowpass-filtering of BL is performed by convolving an input BL with an equivalent weighting function, and then sub-sampled for removing every other pixel and every other row that yields a Gaussian pyramid [89]:

$$GB_{xy,k}^{adf,l} = \sum_{m=-2}^2 \sum_{n=-2}^2 \hat{w}(m,n) GB_k^{adf,l-1}(2x+m, 2y+n) \quad (3.19)$$

Here l ($0 < l < d$) refers to the number of levels in the pyramid and k ($1 < k < N$) refers to the number of input BLs and $\hat{w}(\bullet)$ is an equivalent weighting function.

In ADF approach, the Gaussian pyramid is generated with $\alpha = 0.4$, which yields more Gaussian-like equivalent weighting functions.

Expanding $GB_k^{adf,1}$ to the same size as $GB_k^{adf,0}$ and subtracting yields the band-passed image LB_k^0 . A Laplacian pyramid of input BLs $LB_k^{adf,0}, LB_k^{adf,1}, \dots, LB_k^{adf,d}$ is created by taking the difference of two consecutive levels of Gaussian pyramid that contains band-pass images of decreasing size and spatial frequency [89]:

$$LB_k^{adf,l} = GB_k^{adf,l} - \text{EXPAND}(GB_k^{adf,l+1}), k = 1 \dots N \text{ and } l = 0 \dots d \quad (3.20)$$

$$\text{and } LB_k^{adf,l} = \text{EXPAND}(GB_k^{adf,d}) \quad (3.21)$$

where the expanded image $GB_k^{adf,l+1}$ is the same size as the $GB_k^{adf,l-1}$ and $LB_k^{adf,l}$ is the level of Laplacian pyramid of k_{th} BL. Each Laplacian level contains local spatial information at increasing fine details.

The patches extracted from the input BLs are used for texture analysis (i.e. local range). The weight maps are calculated around every pixel within a 3-by-3 window. The value of the weighting function for each pixel depends on the maximum and minimum intensity value of the neighbors within local window. Next, the Gaussian pyramid of W_k^{adf} calculated in Equation (3.18) is generated as follows:

$$GW_{xy,k}^{adf,l} = \sum_{m=-2}^2 \sum_{n=-2}^2 \hat{w}(m,n) GW_k^{adf,l-1}(2x+m, 2y+n) \quad (3.22)$$

where l denotes the level of Gaussian pyramid of k_{th} weight map $GW_k^{adf,0}, GW_k^{adf,1}, \dots, GW_k^{adf,d}$ that is obtained in the range filtered image. Here $GW_k^{adf,0}$ is the full resolution image and $GW_k^{adf,d}$ is the coarsest level in the pyramid.

The Gaussian pyramid of W^{adf} acts as weight map function that determines the contribution of pixels from the BLs across all of the multiple exposures. The Laplacian pyramid of k_{th} BL $LB_k^{adf,l}$ multiplied with the corresponding Gaussian pyramid of weight map function $GW_k^{adf,l}$ and summing over k yield modified and fused Laplacian pyramid $L_F^{adf,l}(x,y)$:

$$L_F^{adf,l}(x,y) = \sum_{k=1}^N LB_k^{adf,l}(x,y) GW_k^{adf,l}(x,y) \quad (3.23)$$

Equation (3.23) is preferred over simple image averaging fusion scheme. Because in the case of image averaging, the output pixels in the fused image are an average of input pixel's luminance values, which reduce noise in the final image, but the contrast of details is compromised. Note, however, it is found that pyramid fusion perform very well on BL fusion when modified with the weight maps, and it produces more pleasing results with optimal contrast enhancement.

Therefore, from Equation (3.23), the fused BL that contains well exposed pixels is reconstructed from modified $L^{adf,l}$ by expanding and summing each level:

$$BL_F^{adf}(x, y) = L_F^{adf,0}(x, y) + L_F^{adf,1}(x, y) + L_F^{adf,2}(x, y) \dots \dots \dots L_F^{adf,d}(x, y) \quad (3.24)$$

It is found that in ADF approach, the modification of Laplacian pyramid in top down fashion eliminate underexposed and overexposed regions in the fused BL that leads to well-exposed image without introducing objectionable artifacts. The proposed idea for BLs fusion in ADF approach is illustrated in Figure 3.15.

3.5 CONSTRUCTION OF FUSED DETAIL LAYER AND DETAIL LAYER ENHANCEMENT

The DLs $DL_1^{adf}, DL_2^{adf}, \dots, DL_3^{adf}$ computed from Equation (3.16) contain the smaller changes in intensity. There are mainly three parameters in AD that control the behavior of BL and DL computation in ADF approach. Referring to Equation (3.11), t_a and constant λ determines the iterations and the rate of diffusion respectively. The constant value K can be chosen manually or by using the “noise estimator” proposed by Perona and Malik [156]. As a consequence, these three parameters can be varied to moderate texture details in the final fused image. When t_a increases, adjacent pixels with large intensity differences are ignored (i.e. more smoothing at edges), which leads to larger details in the residual layers across different exposures. However, if t_a becomes too small, fewer details are preserved in the residual layers across all of the input exposures with smaller computational time. In order to balance the computational time and details in the fused image, ADF has fixed and suggested $t_a=1$, $\lambda = \frac{1}{7}$ and $K=30$ in all experiments, which reveals reasonably good results. More detailed analysis of effects of these free parameters is given in the forthcoming section. In ADF, two alternative options for constructing DL_F^{adf} have been proposed that manipulate

texture details in the final detail enhanced fused image. Both the options can be utilized, depending on the application.

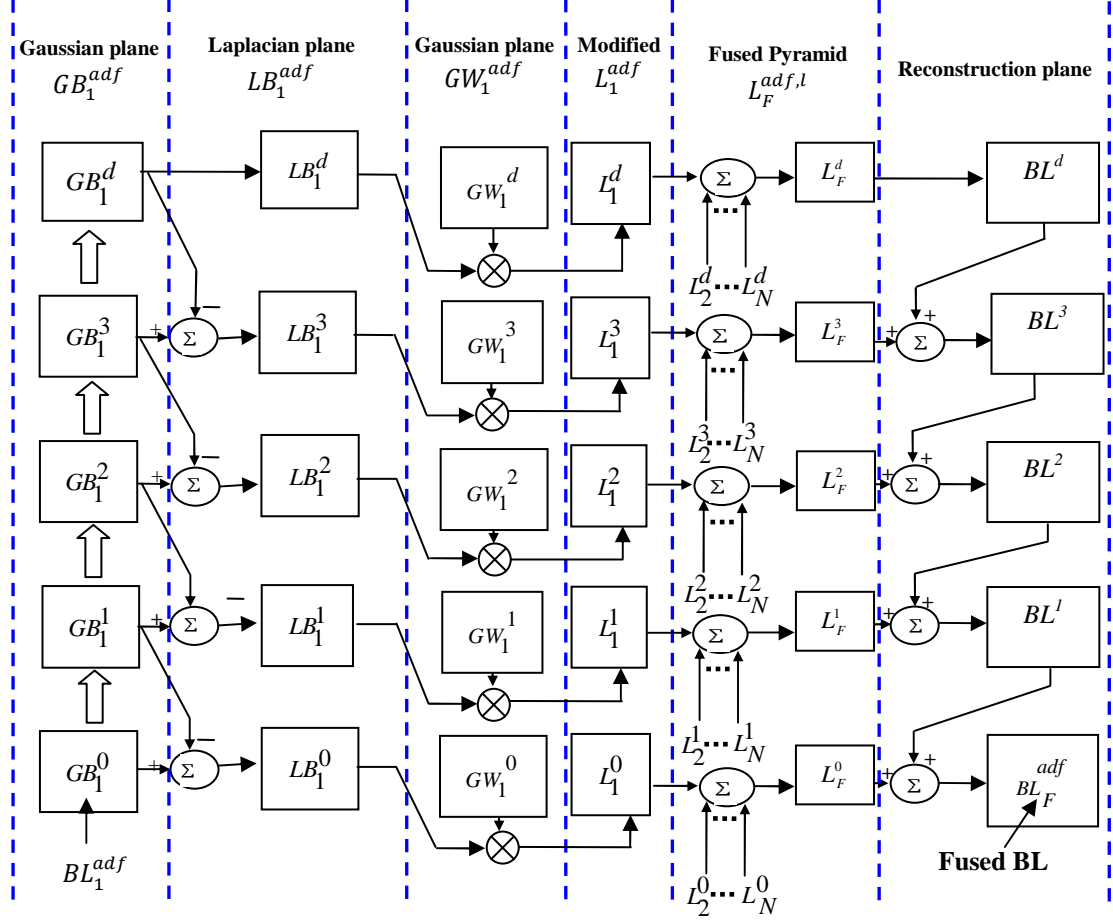


Figure 3.15: BL manipulation and fusion: illustrating the conceptual framework of BLs fusion in ADF approach. Note that for the concept simplicity, here we have generated the Laplacian pyramid of single BL (BL_1^{adf}) and the Gaussian pyramid of the corresponding GW_1^{adf} . Where $L_2^0, L_3^0, \dots, L_N^0$ are the modified Laplacian pyramids of level zero BLs computed across all of multi-exposures.

3.5.1 USER DRIVEN DETAIL LAYER MANIPULATION AND FUSION

In order to compute DL_F^{adf} having rich texture detail, ADF uses a constant weight factor α_1 determined by the user (typical value is 1.2 in ADF approach; please see Figure 3.16(a)). The DL_F^{adf} is obtained as a linear combination of the DLs computed across the input exposure:

$$DL_F^{adf}(x, y) = \frac{\sum_{k=1}^N \alpha_1 DL_k^{adf}(x, y)}{N} \quad (3.25)$$

This straightforward option allows the user to control the contribution of texture details directly from the input DLs across all of the input images. It is found that this simple technique is effective to boost weak texture details in the fused image, but yield over-enhancement at the strong edges. Furthermore, to manipulate DLs across all of the input images precisely, thesis presents a second technique that enhances weak details, while avoids over-enhancement near the edges.

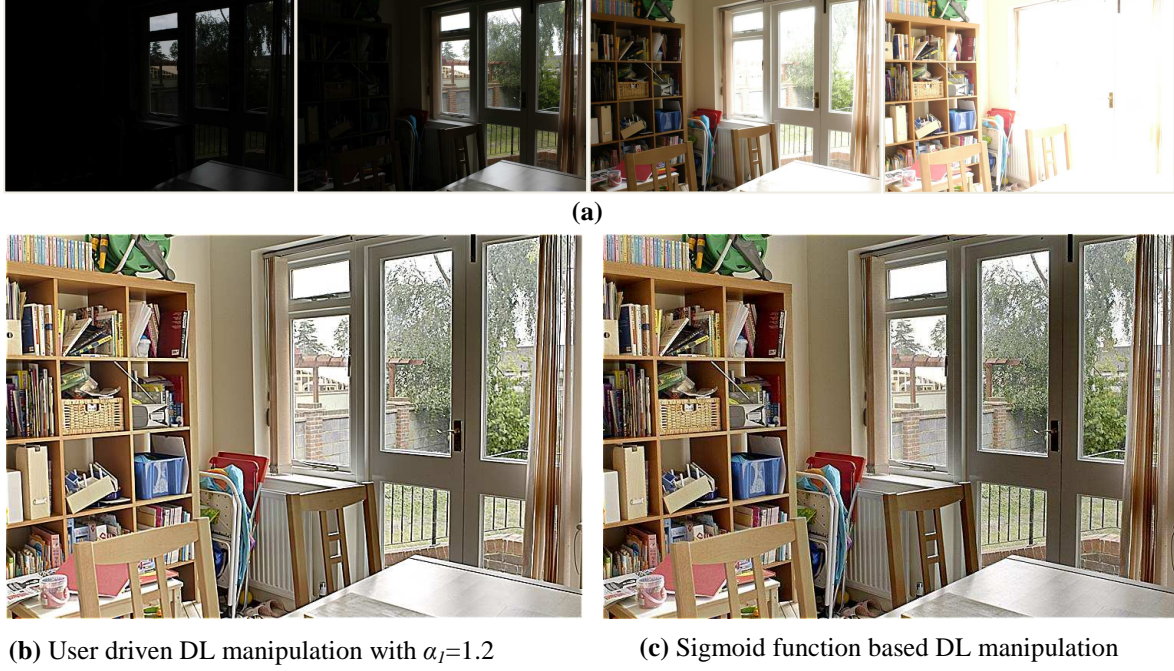


Figure 3.16: Results of ADF approach; (a) Images representing multiple exposures of House; (b) Illustration of ADF result. Note that the fused image contains good details in brightly and poorly illuminated areas with natural contrast. The fine textures on the chair are accurately preserved. (c) The DL enhancement based on sigmoid function across all the inputs reveals more texture details in the fused image and does not depict noticeable artifacts near strong edges.

3.5.2 SIGMOID FUNCTION BASED DETAIL LAYER MANIPULATION AND FUSION

In ADF approach, the second alternative option to enhance fine details in the fused image is based on monotonic non-linear activation function, where the resultant DL_F^{adf} is computed as follows:

$$DL_F^{adf}(x, y) = \sum_x^{n'} \sum_y^{m'} \frac{\alpha_2 \mathcal{S}(DL_k^{adf}(x, y))}{N}, \quad k = 1 \dots N \quad (3.26)$$

where α_2 is a fixed weight ($\alpha_2 = 2$ is found to be suitable in ADF approach in most cases) and $\mathcal{S}(\bullet)$ is the 1-dimensional sigmoid function:

$$\mathcal{S}(t_s) = \frac{1}{1 + e^{-\hat{a}t_s}} \quad (3.27)$$

where $t_s \in \mathcal{R}$ is the independent variable and $\hat{a} \in \mathcal{R}$ is a weight parameter of the sigmoid function. As the Figure 3.16(b) shows, The DL enhancement based on sigmoid function across all the input exposures reveals more texture details in the fused image and does not depict noticeable artifacts near strong edges. Figure 3.17 shows a 1-dimensional sigmoid with different weight values. The weight parameter used in ADF approach was set to 27.

Let θ be a fixed threshold to further control the sharpness of sigmoid function, which is manually chosen by the operator. The 1-dimensional sigmoid function with threshold θ is given by:

$$\mathcal{S}(t_s) = \frac{1}{1 + e^{(-\hat{a}t_s + \theta)}} \quad (3.28)$$

In ADF approach θ is responsible for global contrast management. The detailed analysis of selection of these parameters is given in the forthcoming section. Minai and Williams [174] has presented the sigmoid with threshold as a neuron activation function in artificial neural networks and recurrence relations for calculating derivatives of any order. The first derivative of sigmoid function is computed as:

$$= \hat{a}\mathcal{S}(t_s)(1 - \mathcal{S}(t_s)) \quad (3.29)$$

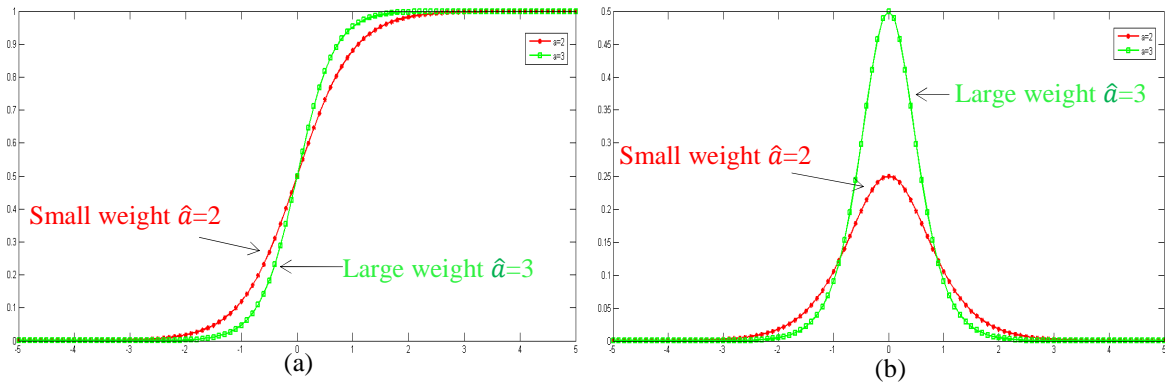


Figure 3.17: The affect of weight (i.e. \hat{a}) on sigmoid function and derivative of sigmoid function: **(a)** The sharpness of the sigmoid function in Equation (3.28) varies according to the value of weight. With larger value of \hat{a} , the sigmoid function becomes a threshold function. **(b)** The first derivative of sigmoid function in Equation (3.29) for $\hat{a}=2$ and $\hat{a}=3$.

In contrast to enhance texture details, DLs can be better handled by the sigmoid function because there is a flexibility to increase or decrease their contribution in the final fused image according to the available texture strength in the input exposures. In ADF approach, a DL manipulation and enhancement approach with free parameters variation flexibility is investigated.

3.6 SUMMARY

This chapter has described texture features based exposure fusion, which has applicability to preserve details in poorly and brightly illuminated regions. In the present method, the Gaussian pyramid of texture features are utilized to modify the Laplacian pyramid of the BLs computed across the input multiple exposures and then constructs a well-exposed LDR image by expanding, then summing all the levels of the fused Laplacian pyramid for the different BLs. For preserving and enhancing fine texture details in the fused image, a nonlinear diffusion filter based on PDE is investigated. In particular, the main contribution of this work is proposal of a technique that fuses details in edge preserving manner from images captured at variable exposure settings without the introduction of artifacts. In this approach, a local range metric for fast calculation of weight map is designed and implemented. Another conclusion is that, sigmoid function has a better response than linear manipulation of DLs for enhancing fine details in the fused image. The work presented in this chapter describes MRP based exposure fusion framework for preserving and enhancing texture details in the fused image. Experimental results demonstrated in Chapter 6 show that present approach improves strong edges and weak textures simultaneously for better visual perception.

CHAPTER-IV

GUIDED FILTER BASED MULTI-EXPOSURE IMAGE FUSION (GFF)

This chapter discuss a detail-enhancing exposure fusion approach using NTF. With the captured LDR images under different exposure settings, first the fine details are extracted based on GF. Next, the BLs (i.e. images obtained from NTF) across all input images are fused using multi-resolution pyramid. Exposure, contrast and saturation measures are considered to generate a weight map function that guides the fusion process of the BLs. Finally, the fused BL is combined with the extracted fine details to obtain detail enhanced fused image. The goal is to preserve details in both very dark and extremely bright regions without HDR image representation and tone-mapping step. Moreover, it will be demonstrated in Chapter 6 that the GFF method is also suitable for multi-focus image fusion without introducing artifacts.

4.1 INTRODUCTION

In GFF approach, the GF is preferred over other existing approaches because the gradients present near the edges are preserved accurately. GFF uses GF [124] for BL and DL extraction which is more effective for enhancing texture details and reducing gradient reversal artifacts near the strong edges in the fused image. Multi-resolution approach is used to fuse computed BLs across all of the input images. The DLs extracted from input exposures are manipulated and fused separately. The final detail enhanced fused image is obtained by integrating the fused BL and the fused DL (please see Figure 4.1). It is worth pointing out that GFF method essentially differs from [113], which aims at enhancing the texture and contrast details in the fused image with a non-linear edge preserving filter (i.e. the GF). Moreover, it is demonstrated in Chapter 6 that the proposed approach fuses multi-focus images effectively and produces the result of rich visual details.

4.2 EDGE-PRESERVING GUIDED FILTER

In this section, the ability of GF [124] derived from local linear model is described to preserve edges, and then show how it avoids gradient reversal artifacts near the strong edges that may appear in fused image after DL enhancement. GFF seeks to maintain the shape of strong edges in the fused image that appears due to exposure time variation across input images.

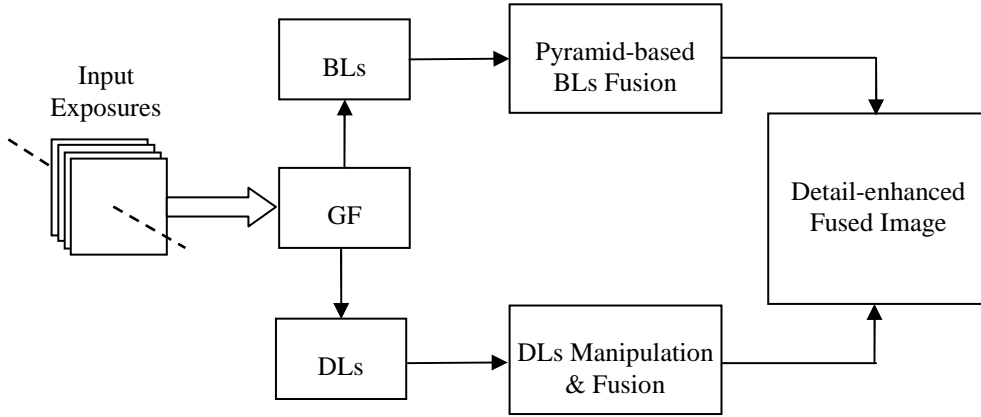


Figure 4.1: GFF based detail-enhanced exposure fusion framework.

Encouraged by the efficient edge preserving capability of GF, GFF will preserve weak texture details across input exposures without introducing gradient reversal artifacts. To examine the effectiveness of GF for separating DL, we show a 1-D illustration of edge-aware smoothing in Figure 4.2. We notice that the BL computed from GF is consistent with the input image near the strong edges which can avoid false texture detail extraction in DL computation compared with the BF filtering method.

GF was developed by Kaiming et al. [124] in 2010 as an alternative to BF [167] and AD [156]. It is an edge-preserving filter where the filtering output is a local linear model between the guidance I^G and the filter output q . The selection of guidance image I^G will depend on the application [124]. In GFF implementation an input image p and guidance image I^G are identical. The output of the GF for a pixel i is computed as a weighted average:

$$q_i = \sum_j W_{ij}(I^G)p_i \quad (4.1)$$

where (ij) are pixel indexes and W_{ij} is the filter kernel that is a function of guidance image I^G , and independent of input image p . Such type of filter is linear with respect to p .

The BF [167] is a well-known explicit weighted-average filter in which Gaussian kernel based on the distance between the pixels and the range based on intensity differences are utilized to compute kernel weights $W_{i,j}^{bf}$:

$$W_{i,j}^{bf} = \frac{1}{\mathcal{K}_i} \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma_s^2}\right) \exp\left(-\frac{\|I_i^G - I_j^G\|^2}{\sigma_t^2}\right) \quad (4.2)$$

where \mathbf{x} is the pixel coordinate and \mathcal{K}_i is a normalizing parameter to ensure that $\sum_j W_{i,j}^{bf} = 1$.

The width of spatial domain Gaussians kernel (i.e. sensitivity of the spatial similarity) and intensity domain Gaussians kernel (i.e. sensitivity of the range (intensity/color) similarity) is controlled by the standard deviation parameters σ_s and σ_t respectively.

Another option is implicit weighted-average filters which optimize a quadratic cost function and solve a linear system:

$$\mathcal{M}q = p \quad (4.3)$$

where q and p are \tilde{N} -by-1 vectors concatenating q_i and p_i , respectively, and \mathcal{M} is an \tilde{N} -by- \tilde{N} matrix only depends on I^G . The solution to Equation (4.3), i.e., $q = \mathcal{M}^{-1}p$, has the same form as Equation (4.1), with $W_{ij} = (\mathcal{M}^{-1})_{ij}$.

Let q be a linear transform of I^G in a window centered at the pixel k :

$$q_i = a_k I_i^G + b_k, \forall i \in \omega_k \quad (4.4)$$

where a_k, b_k are the linear coefficients assumed to be constant in window ω_k centered at the pixel k and calculated in a small square image window of a radius $(2r+1) \times (2r+1)$. The local linear model Equation (4.4) ensures that q has an edge (i.e. discontinuities) only if I^G has an edge, because $\nabla q = a \nabla I^G$. Here a_k and b_k are computed within ω_k to minimize following cost function:

$$E(a_k, b_k) = \sum_{i \in \omega_k} \left((a_k I_i^G + b_k - p_i)^2 + \varepsilon a_k^2 \right) \quad (4.5)$$

where ε is the regularization term on linear coefficient a_k for numerical stability. The relation given in Equation (4.5) is a linear regression [175] in which data are modeled using linear predictor functions, and unknown model parameters are estimated from the data. The detailed description of regression analysis tools is given in the Chapter 5, Sub-section 5.4.1. Therefore a_k and b_k are estimated by minimizing the squared difference between the output image q and the input image p . The significance of selection of window size and relation of ε with the

bilateral kernel [167] is given in [124]. In this thesis, GFF uses $r=2$ (i.e. 5×5 square window) and $\varepsilon = 0.01$.

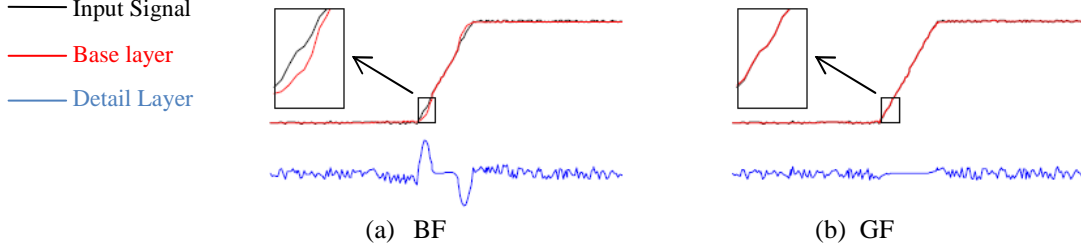


Figure 4.2: 1-D illustration of BL and DL extraction using: (a) BF; and (b) GF.

The linear coefficients used to minimize the cost function in Equation (4.5) are determined by linear regression [175].

$$a_k = \frac{\frac{1}{|\omega|} \sum_{i \in \omega_k} I_i^G p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \varepsilon} \quad (4.6)$$

$$b_k = \bar{p}_k - a_k \mu_k \quad (4.7)$$

$$\bar{p}_k = \frac{1}{|\omega|} \sum_{i \in \omega_k} p_i \quad (4.8)$$

The full derivation of the computation of a_k and b_k is available in Appendix A. In Equations (4.6, 4.7 and 4.8), μ_k and σ_k^2 are the mean and variance of I_i^G in ω_k , $|\omega|$ is the number of pixels in ω_k , and \bar{p}_k is the mean of p in ω_k .

The linear coefficients a_k and b_k are computed for all patches ω_k in the entire image. However, a pixel i is involved in all windows ω_k (please see Figure 4.3) that contains i so the value of q_i in Equation (4.4) will be different for different windows.

$$q_i = \frac{1}{|\omega|} \sum_{k: i \in \omega_k} (a_k I_i^G + b_k) \quad (4.9)$$

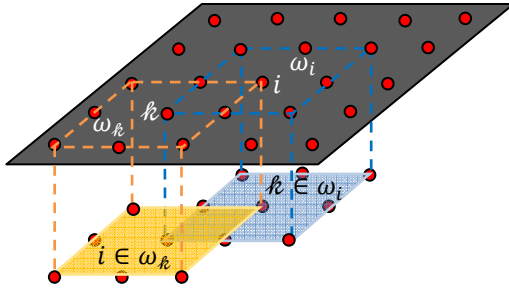
As shown in Figure 4.3 (a), all local windows centered at pixel k in the window i will contain pixel i . So, the value of q_i in Equation (4.4) will change when it is computed in different windows ω_k . To solve this problem, all the possible values of coefficients a_k and b_k are first averaged. Then, the filtered output is determined as follows:

$$q_i = (\bar{a}_i I_i^G + \bar{b}_i) \quad (4.10)$$

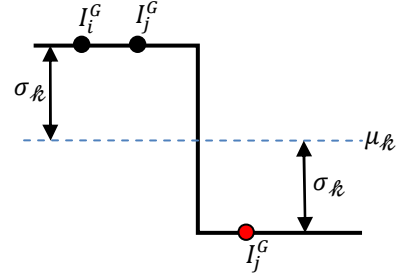
Here \bar{a}_i and \bar{b}_i are computed as:

$$\bar{a}_i = \frac{1}{|\omega|} \sum_{\ell \in \omega_i} a_\ell \text{ and } \bar{b}_i = \frac{1}{|\omega|} \sum_{\ell \in \omega_i} b_\ell \quad (4.11)$$

In practice, it is found that \bar{a}_i and \bar{b}_i in Equation (4.11) are varying spatially to preserve strong edges of I_i^G in q i.e. $\nabla q = \bar{a}_i \nabla I_i^G$. Therefore, q_i computed in Equation (4.10) preserves strongest edges in I_i^G while smoothing small changes in intensity.



(a) Illustration of 2-D window choice.



(b) A 1-D example of an ideal step edge.

Figure 4.3: Window selection for GF: (a) Illustration of local windows centered at pixel ℓ in the window ω_i ; (b) Example of ideal step edge of a 1-D signal. For a window that exactly centers on the edge, the variables μ and σ are as indicated.

4.2.1 EDGE-PRESERVING PROPERTY OF GUIDED FILTER

For instance, Figure 4.4 shows examples of the GF as an edge-preserving smoothing operator with various sets of parameters. It can be observed that the edge-preserving property of guided filtering can be controlled by tuning the free parameters ε and r . Figure 4.5 shows the corresponding DLs produced from GF smoothing of the images in Figure 4.4. The resulting DLs yield fine texture that are extracted by taking the difference between the original images and the guided filtered images shown in Figure 4.4, and the effects of various free parameters used in guided filtering are visible in Figure 4.5. Notice that the coarsening is increased by increasing the values of free parameters ε and r . It is found that the most of the fine textures have been filtered away from the coarsest image, which are perfectly reconstructed in the DL.

The edge-preserving smoothing property of the DF can be explained intuitively as following. If we consider that input image and guidance image are identical (i.e. $I^G \equiv p$). In

this case, $a_k = \sigma_k^2 / (\sigma_k^2 + \varepsilon)$ in Equation (4.6) and $b_k = (1 - a_k)\mu_k$ in Equation (4.7). It is clear that if $\varepsilon = 0$, then $a_k = 1$ and $b_k = 0$. If $\varepsilon > 0$, we can consider two cases.

Case 1: “High variance.” If the image I^G changes a lot within ω_k , we have $\sigma_k^2 \gg \varepsilon$, so $a_k \approx 1$ and $b_k \approx 0$.

Case 2: “Flat patch.” If the image I^G is almost constant in ω_k , we have $\sigma_k^2 \ll \varepsilon$, so $a_k \approx 0$ and $b_k \approx \mu_k$.

When a_k and b_k are averaged to get \bar{a}_l and \bar{b}_l , combined in Equation (4.10) to get the output, we have that if a pixel is in the middle of a “high variance” area, then its value is unchanged ($a \approx 1, b \approx 0, q \approx p$), whereas if it is in the middle of a “flat patch” area, its value becomes the average of the pixels nearby ($a \approx 0, b \approx \mu, q \approx \bar{\mu}$).

More specifically, the criterion of a “flat patch” or a “high variance” one is given by the parameter ε . The patches with variance (σ^2) much smaller than ε are smoothed, whereas those with variance much larger than ε are preserved. The effect of ε in the GF is similar to the range variance σ_r^2 in the BF Equation (4.2): Both determine “what is an edge/a high variance patch that should be preserved.”

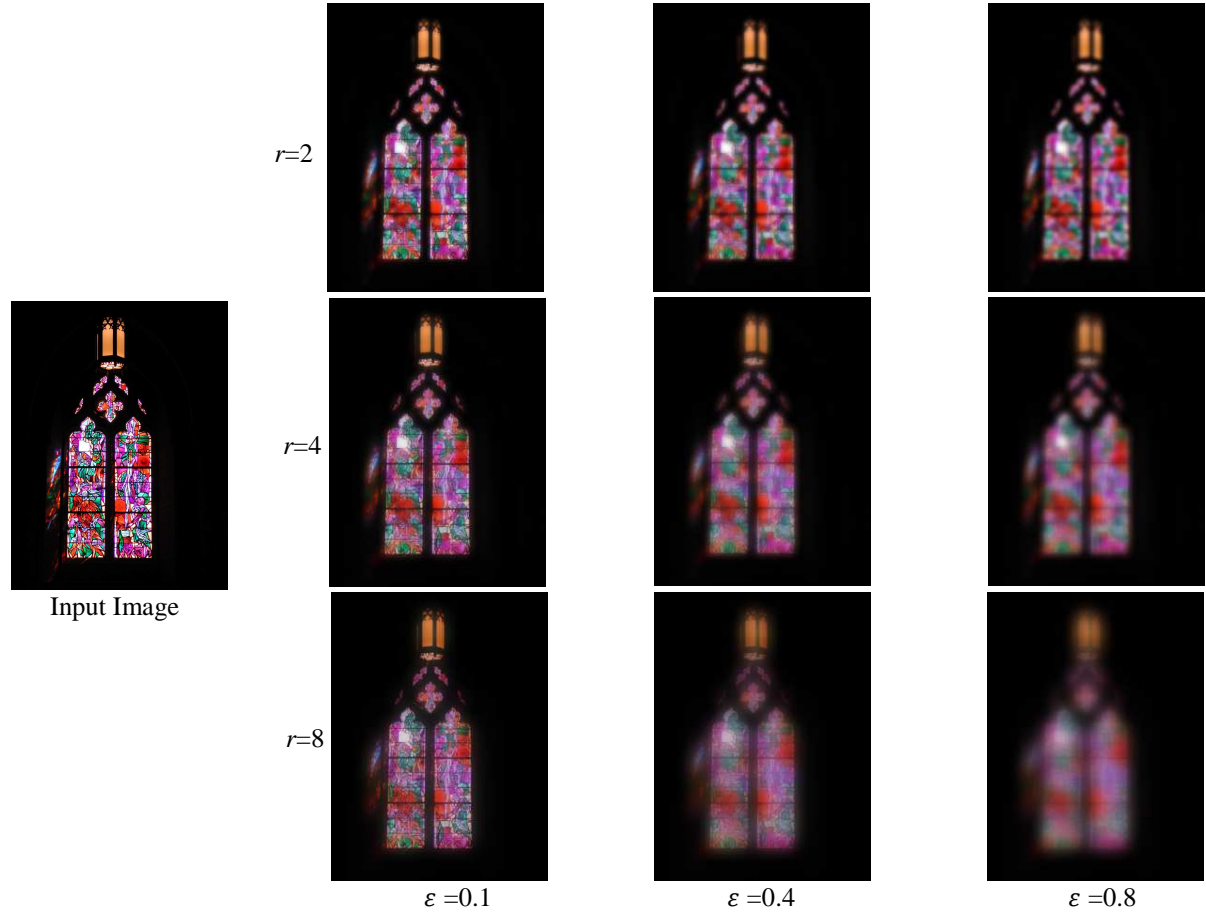


Figure 4.4: Edge-preserving property of the GF and the effect of free parameters.

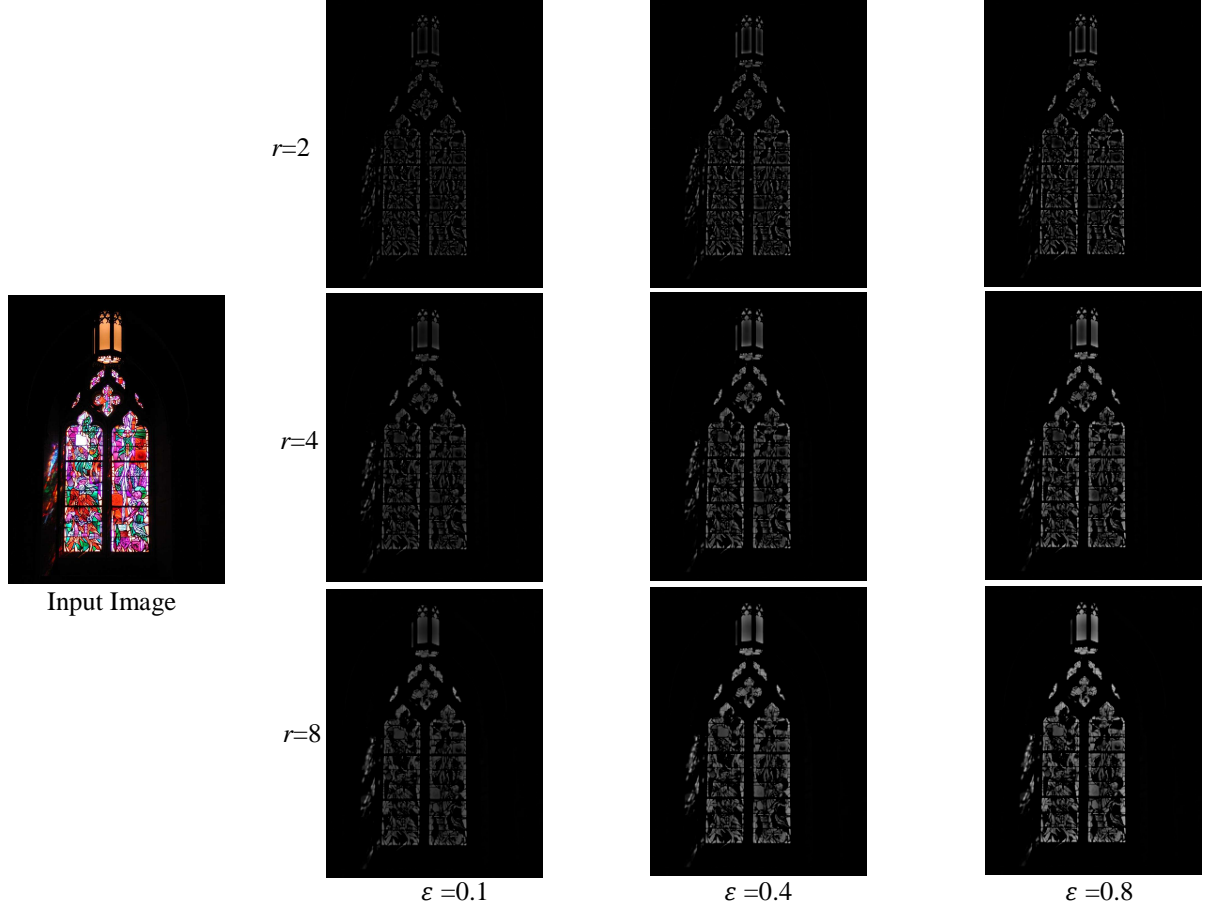


Figure 4.5: The DLs computed for different parameters settings used in GF.

4.2.2 GUIDED FILTER KERNEL

We point out that the relationship among I^G , p , and q given by Equation (4.6), Equation (4.7), and Equation (4.10) are indeed in the form of image filtering Equation (4.1). In fact, a_k in Equation (4.6) can be rewritten as a weighted sum of p : $a_k = \sum_j A_{kj}(I^G) p_j$, where A_{ij} are the weights only dependent on I^G . For the same reason, we also have $b_k = \sum_j B_{kj}(I^G) p_j$ from Equation (4.7) and $q_i = \sum_j W_{ij}(I^G) p_j$ from Equation (4.10). It can be proven that the kernel weights can be explicitly expressed by:

$$W_{ij}(I^G) = \frac{1}{|\omega|^2} \sum_{k:(i,j) \in \omega_k} \left(1 + \frac{(I_i^G - \mu_k)(I_j^G - \mu_k)}{\sigma_k^2 + \varepsilon} \right) \quad (4.12)$$

The edge-preserving smoothing property can also be understood by investigating the filter kernel Equation (4.12). Take an ideal step edge of a 1D signal as an example (see Figure 4.3 (b)). The terms $I_i^G - \mu_k$ and $I_j^G - \mu_k$ have the same sign (+/-) when I_i^G and I_j^G are on the same side of an edge, while they have opposite signs when the two pixels are on

different sides. So in Equation (4.12) the term $1 + \frac{(I_i^G - \mu_k)(I_j^G - \mu_k)}{\sigma_k^2 + \varepsilon}$ is much smaller (and close to zero) for two pixels on different sides than on the same sides. This means that the pixels across an edge are almost not averaged together. We can also understand the smoothing effect of ε from Equation (4.12). When $\sigma_k^2 \ll \varepsilon$ (“flat patch”), the kernel becomes $W_{ij}(I^G) = \frac{1}{|\omega|^2} \sum_{k:(i,j) \in \omega_k} 1$: This is a Linear Translation-Invariant (LTI) low-pass filter (it is a cascade of two mean filters) [115].

4.3 BASE LAYER AND DETAIL LAYER EXTRACTION

Let $BL_k^{gff}(x, y)$ is the BL computed from GF in Equation (4.10) (i.e. $BL_k^{gff}(x, y) = q_i$) for k_{th} input image denoted by $I_k(x, y)$. The DL is defined as the difference between the GF output and the input image, which is defined as:

$$DL_k^{gff}(x, y) = I_k(x, y) - BL_k^{gff}(x, y) \quad (4.13)$$

4.4 WEIGH MAP CONSTRUCTION

In GFF method, saturation (SA_k), contrast (CO_k), and exposure (EX_k) measures are considered to generate a mask that guides the fusion process of the BLs. The saturation [176] of a color described the intensity of color in the photograph which would be useful to produce well exposed pixels in the fused image. When image is captured at long exposure setting the color details present in the brightly illuminated areas become less saturated. A desaturated photograph will look washed out and has overly faded colors. The approach proposed in the thesis would avoid desaturated pixels by producing the weight map with saturation measure. The saturation measure SA_k is computed for k_{th} source image as the standard deviation within the R, G, and B channel, at each pixel. The mathematical description of SA_k is given in Chapter 5, Section (5.3).

Contrast is determined by the difference in luminance or color within the same field of view that makes an object distinguishable. The response of HVS is more sensitive to contrast than absolute luminance. In other words, the maximum contrast of a photograph can be determined by the ratio of the luminance of the brightest color (white) to that of the darkest color (black), which is known as contrast ratio or DR. In order to measure local contrast, GFF

used isotropic derivative operator [177]. The isotropic derivative of an image $I(x,y)$ having two independent variables, denoted by $\nabla^2(x,y)$, is defined as:

$$\nabla^2 I(x,y) = \frac{\partial^2 I(x,y)}{\partial x^2} + \frac{\partial^2 I(x,y)}{\partial y^2} \quad (4.14)$$

For digital images, the second order derivative in the x-direction is estimated as:

$$\frac{\partial^2 I(x,y)}{\partial x^2} = I(x+1,y) + I(x-1,y) - 2I(x,y) \quad (4.15)$$

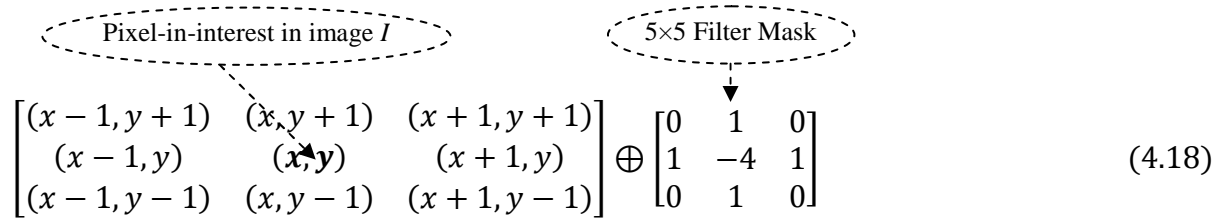
and, similarly in the y-direction is estimated as:

$$\frac{\partial^2 I(x,y)}{\partial y^2} = I(x,y+1) + I(x,y-1) - 2I(x,y) \quad (4.16)$$

From Equation (4.14), two-dimensional Laplacian operator can be obtained by summing Equation (4.15) and Equation (4.16) as:

$$\nabla^2 I(x,y) = I(x+1,y) + I(x-1,y) + I(x,y+1) + I(x,y-1) - 4I(x,y) \quad (4.17)$$

This Equation (4.17) can be implemented in an image by convolving the image with the following 5-by-5 mask:



$$\begin{bmatrix} (x-1,y+1) & (x,y+1) & (x+1,y+1) \\ (x-1,y) & (x,y) & (x+1,y) \\ (x-1,y-1) & (x,y-1) & (x+1,y-1) \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (4.18)$$

If the image is of pixel intensities are between 0 and 1. A correctly exposed picture is one that has intensities not near zero (under-exposed) or one (over-exposed). Therefore, a pixel is said to be well-exposed if intensity value is close to 0.5 [113]. The weight intensity value (\tilde{i}) of each pixel at (x,y) location based on how close it is to 0.5 using a Gauss curve:

$$EX(x,y) = e^{\left(-\frac{(i-0.5)^2}{2\sigma^2}\right)} \quad (4.19)$$

where σ' equals 0.2 in GFF implementation, which can control the quality of the fused image. When computing the weights in GFF, the fusion performance will be worse when the value of σ' is too large or too small. In this paper, the default parameters is set as $\sigma' = 0.2$. In order to balance the brightness in fused image, we have found that $\sigma' = 0.2$ generates reasonably good results for most of the cases. To account for multiple color channels, GFF apply the Gauss

curve to each R, G, and B channel separately, and multiply the results for computing final exposedness measure EX_k .

The weight map function used in GFF approach is computed as the product of three quality metrics:

$$W_k^{gff}(x, y) = SA_k(x, y) \times CO_k(x, y) \times EX_k(x, y) \quad (4.20)$$

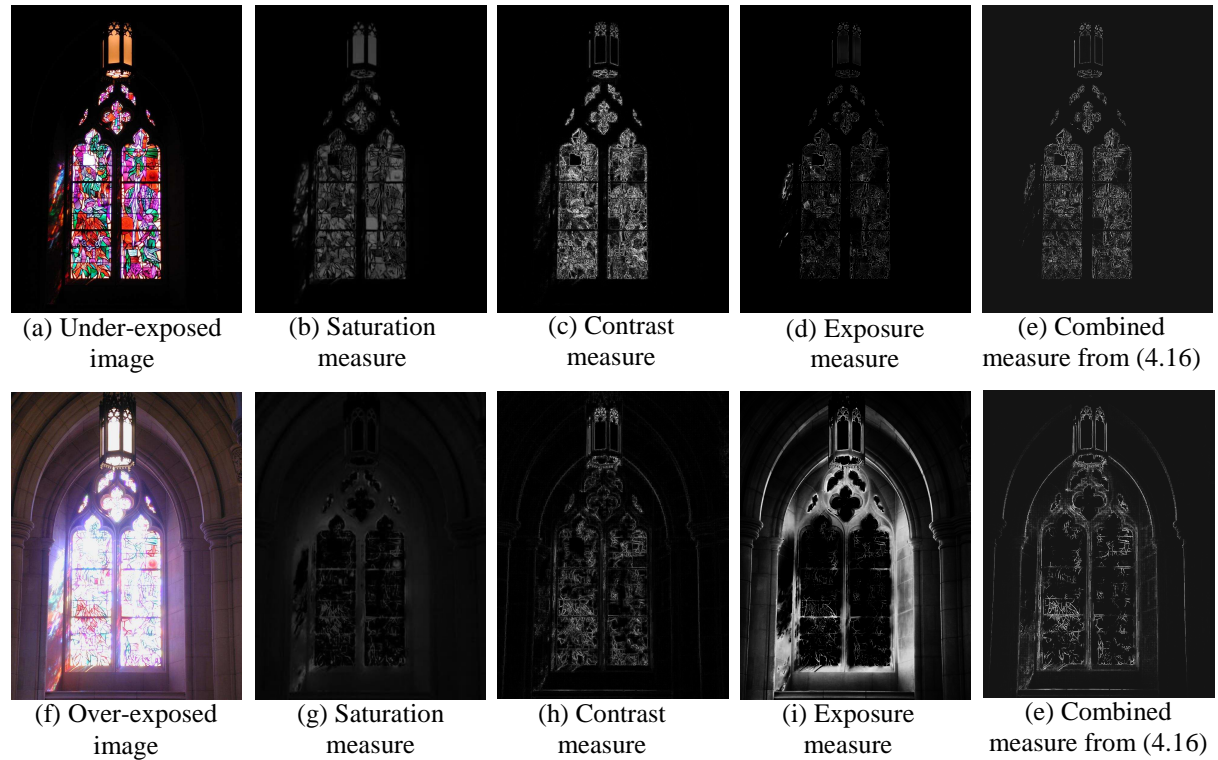


Figure 4.6: Illustrations of saturation, contrast and exposure measures used for weight map computation in GFF approach.

4.5 BASE LAYER FUSION BASED ON MULTI-RESOLUTION PYRAMID

In GFF based exposure fusion framework, the fused BL $BL_F^{gff}(x, y)$ is computed as the weighted sum of the BLs obtained from GF $BL_1^{gff}(x, y)$, $BL_2^{gff}(x, y)$, ..., $BL_N^{gff}(x, y)$ obtained across N input exposures. GFF uses the pyramid approach proposed by Burt et al. [89], which generate Laplacian pyramid of the BL $LB_k^{gff,l}(x, y)$ and Gaussian pyramid of weight map functions $GW_k^{gff,l}(x, y)$ estimated from three quality measures (i.e. $SA_k(i, j)$, $CO_k(i, j)$, and $EX_k(i, j)$). The Laplacian pyramid and Gaussian pyramid of BLs and weight map

function are computed in similar manner as computed in ADF approach, respectively. Here l ($0 < l < d$) refers to the number of levels in the pyramid and k ($1 < k < N$) refers to the number of input images.

The Laplacian pyramid of k_{th} BL ($LB_k^{gff,l}(x,y)$) is multiplied with the corresponding Gaussian pyramid of weight maps ($GW_k^{gff,l}(x,y)$) and summed over k yield modified and fused Laplacian pyramid $L_F^{gff,l}(x,y)$:

$$L_F^{gff,l}(x,y) = \sum_{k=1}^N LB_k^{gff,l}(x,y) GW_k^{gff,l}(x,y) \quad (4.21)$$

The $BL_F^{gff}(x,y)$ that contains well exposed pixels is reconstructed by expanding and summing each level of modified and fused Laplacian pyramid:

$$BL_F^{gff}(x,y) = \sum_{l=0}^d L_F^{gff,l}(x,y) \quad (4.22)$$

4.6 DETAIL LAYER FUSION AND MANIPULATION

The DLs computed in Equation (4.13) across all the input exposures are linearly combined to produce fused detail layer $DL_F^{gff}(x,y)$ that yields combined texture information:

$$DL_F^{gff}(x,y) = \frac{\sum_{k=0}^N \gamma f_k(DL_k^{gff}(x,y))}{N} \quad (4.23)$$

where γ is the user defined parameter to control amplification of texture details (typically set to 5) and $f_k(\bullet)$ is the nonlinear function to achieve detail enhancement while reducing noise and artifacts near strong edges due to over-enhancement. GFF follow the approach of [157] to reduce noise across all DLs. The non-linear function $f_k(\bullet)$ is defined as:

$$f_k(x,y) = \wp(DL_k^{gff}(x,y))^{\alpha_3} + (1 - \wp)DL_k^{gff}(x,y) \quad (4.24)$$

where \wp is a smooth step function equal to 0 if $DL_k^{gff}(x,y)$ is less than 1% of the maximum intensity, 1 if it is more than 2%, with a smooth transition in between and the parameter α_3 is used to control contrast in the fused DL. While it is found that $\alpha_3 = 0.2$ is a good default setting for all experiments.

Finally, the detail enhanced fused image $I_F^{gff}(x, y)$ is easily computed by simply adding up the fused BL $BL_F^{gff}(x, y)$ computed in Equation (4.22) and the manipulated fused DL $DL_F^{gff}(x, y)$ in Equation (4.23):

$$I_F^{gff}(x, y) = BL_F^{gff}(x, y) + DL_k^{gff}(x, y) \quad (4.25)$$

4.7 SUMMARY

In this chapter the edge-preserving property of GF is extended to BL and DL decomposition, which has applicability to avoid false texture detail extraction in DL computation. The two layer decomposition based on GF is used to extract fine textures for detail enhancement. In the present method, the MRP approach is used for the fusion of BLs, and the utilization of contrast, saturation and well-exposedness for the calculation of weight map is investigated.

An alternate approach for DL enhancement using non-linear function has been described. By this approach the over-enhancement of strong edges in the fused image to be avoided.

CHAPTER-V

WLS FILTER BASED MULTI-EXPOSURE IMAGE FUSION (WLF)

This chapter describes a novel technique for exposure fusion in which WLS optimization framework is utilized for weight map refinement. Computationally simple texture features (i.e. DL extracted with the help of edge preserving filter) and color saturation measure are preferred for quickly generating weight maps that control the contribution from an input set of multi-exposure images. Instead of employing intermediate HDR reconstruction and tone-mapping steps, well-exposed fused image is generated for displaying on conventional display devices. A further advantage of the present technique is that it is well suited for multi-focus image fusion. In Chapter 6, simulation results are compared with a number of existing single resolution and multi-resolution techniques to show the benefits of the proposed scheme for variety of cases.

5.1 INTRODUCTION

The block diagrammatic representation of detail enhanced WLF framework is shown in Figure 5.1. WLF uses edge preserving filter based on PDE for two-scale decomposition that separates sharp details and fine details across various input images with different exposure levels. The current state-of-the-art method for automatic exposure fusion exploits the capability of edge preserving filter [156] to generate weight function that guides the fusion of different exposures based on two-scale decomposition. WLF proposes WLS optimization [133] and sigmoid function based weight map refinement for BLs and DLs fusion, respectively. Farbman et al. [133] has utilized WLS filter to construct a multi-scale edge-preserving decomposition multi-scale tone and detail manipulation. To achieve optimal contrast in the fused image present approach develops an appropriate mask based on weak textures and color saturation measure for compositing multi-exposure images. The WFF approach is applicable for the fusion of broad range of textured images.

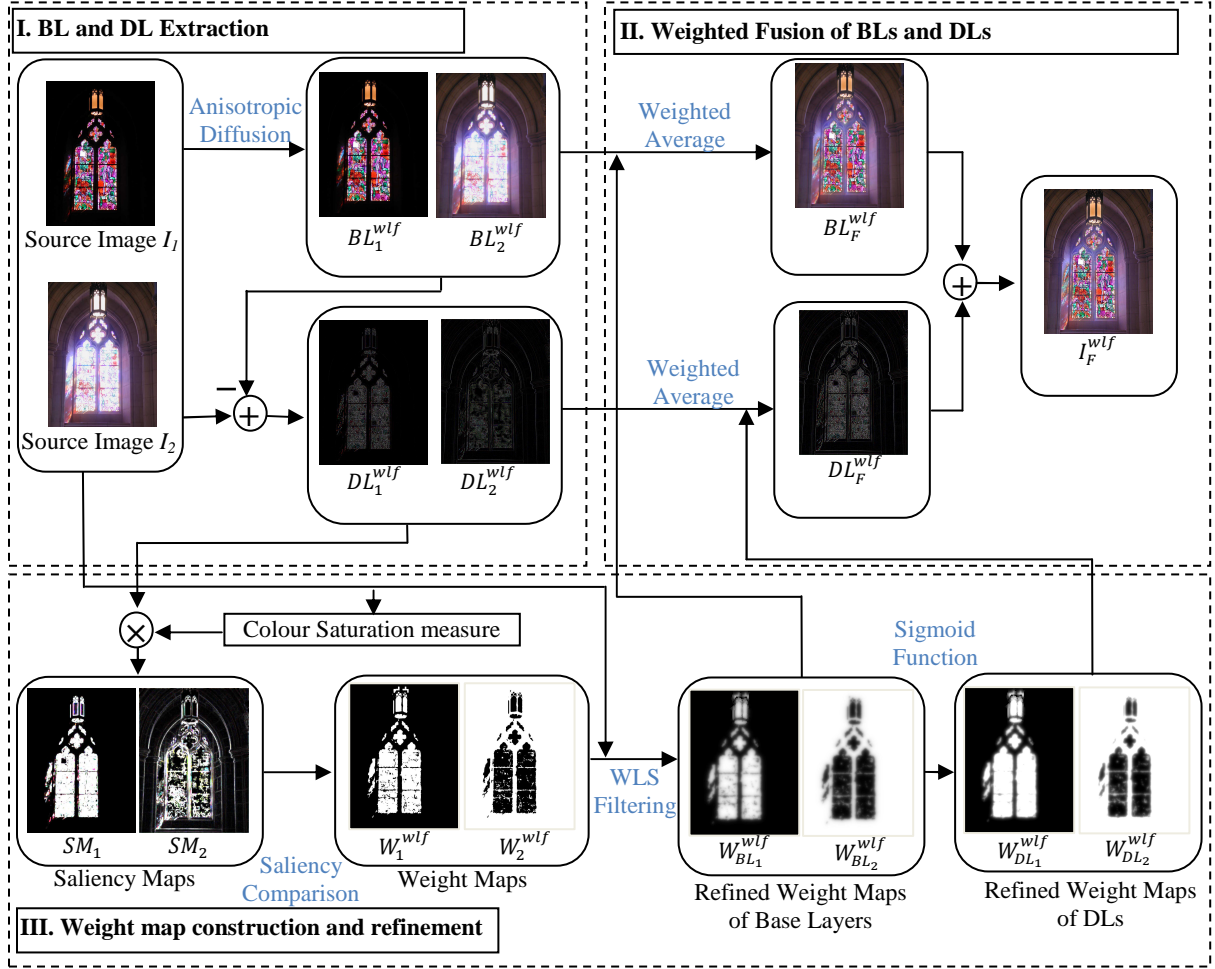


Figure 5.1: Illustration of WLS filter based proposed framework that consists of three principal blocks. (I) BL and DL extraction. The input images are transformed into two scale decomposition. (II) Weight map construction and refinement. (III) Weighted fusion of BLs and DLs. The BLs and DLs across input image series are fused using simple weighted average approach.

Texture features [178] refer to the characterization of regions in an image by their spatial arrangement of color or intensities. Image textures are one way that can be used to help in classification of images [179]. Weak edges or texture information are the ideal indicators to detect over (or under) exposed regions in the image [111]. Raman and Chaudhuri [111] employ a BF for compositing multi-exposure images, in which weak edges were considered to design weight map. Interestingly, thus an analysis of weak textures seems to be the definition of perceived contrast. WLF takes advantage of such possibility and design the appropriate matting function based on AD for exposure fusion. The detailed description of AD filter is given in Chapter 3.

To analyze an image texture, there are primarily two approaches; structural approach and statistical approach. Structural approach uses a set of primitive texture elements in some regular or repeated pattern to characterize spatial relationship. While statistical approach

defines an image texture as a quantitative measure of the arrangement of intensities in a region. In general later approach is easier to compute and is more widely used in computer graphics applications, since natural textures are made of patterns of irregular sub-elements. It has been noticed that simple averaging to fuse details from multi-exposure image data yields low contrast in the fused image: especially in brightly and poorly illuminated areas. In the present approach, texture details will decide the contribution of corresponding pixel from different exposures in the fused image. A rich texture details means a maximum contribution, which tells that image block has higher weight during the fusion process. Such metric is used to quantify the perceived local contrast of an image under different exposure settings and allows discarding under exposed and overexposed pixels. Therefore, to handle under exposed and overexposed regions, WLF propose a texture feature analysis based on AD [156] that has the applicability to design weight map function as shown in Figure 5.1. The goal of present work is to exploit the edge preserving property of AD to produce well exposed image from input images captured under different exposure settings. The detailed description of AD based two-layer decomposition and weight map computation is given in the later sections. The main contributions in the present WLF approach are highlighted as:

1. Two-scale decomposition based on AD is proposed for fast exposure fusion, which does not require optimization of number of scales as required in the traditional multi-scale techniques.
2. A novel weight construction approach is proposed to combine texture features and saturation measure for guiding image fusion process. For weight map construction, WLF seek to utilize the strength of texture details under the change of exposure setting that takes place between an under-exposed and an over-exposed image. WLS filtering is proposed for weight refinement for BLs. Furthermore, fast sigmoid function based weight map generation for DLs is proposed that reduce computational complexity of the algorithm.
3. The important contribution of this paper is the advantages include ease of implementation, quality of compositing, and the provision of DL enhancement without introducing artifacts.

A new type of exposure fusion technique is developed to avoid the limitation of conventional digital camera to handle the luminance variation in the entire scene. The primary focus of this paper is the development of a fast and robust exposure fusion approach based on local texture features computed from edge-preserving filter. Unlike most previous multi-exposure fusion methods, we build on AD, a non-linear filter introduced by Perona et al. [156] in 1990 that

has the ability to preserve large discontinuities (edges). It derives from magnitude of the gradient of the image intensity and controls diffusion strength in the image to prevent blurring across edges. As such, the algorithm implemented (please see Figure 5.1) include four steps:

1. A first step, in WLF algorithm, is two-scale decomposition based on AD which is used to separate coarser details (BLs) and finer details (DLs) across each input exposure.
2. Weak textures (DLs computed in the previous step) and saturation based weight mask generation which provides precise control of contribution of well exposed pixels from input image series.
3. WLS filter and sigmoid function based weight map refinement is performed for coarser details and finer details computed in the first step, respectively.
4. Weighted average based blending of coarser details and finer details is performed to form a composite seamless image without blurring or loss of detail near large discontinuities.

5.2 EXTRACTION OF COARSER DETAILS AND FINER DETAILS

Edge preserving filters have received considerable attention in computational photography over the last decade. BF [167] and AD [156] are the most popular edge-preserving operators. Standard BF uses distances of neighboring pixels in space and range. The space-varying weighting function is computed at a space of higher dimensionality than the signal being filtered. As a result, such filters have high computational costs [180]. AD has led to an excellent tool for smoothing fine details of an image while preserving the coarser details (i.e. edges). It is modeled using PDEs and based on non-linear iterative process. The diffusion Equation (3.11) is used for BL and DL computation. The BL and DL in WLF approach are computed as follows:

$$BL_k^{wlf}(x, y) = \mathcal{A}(I_k(x, y)) \quad \text{and} \quad DL_k^{wlf}(x, y) = I_k(x, y) - BL_k^{wlf}(x, y) \quad (5.1)$$

where $k = 1 \dots \dots \dots N$

where $\mathcal{A}(\bullet)$ is the AD (edge-preserving filter) applied on input image. The detailed description of AD is given in Chapter 3.

The well exposed fused image using WLF approach is given by

$$I_F^{wlf}(x, y) = BL_F^{wlf}(x, y) + DL_F^{wlf}(x, y) \quad (5.2)$$

where BL_F^{wlf} is the fused BL that maximizes the coarser details across all of the input BLs $BL_1^{wlf}, BL_2^{wlf}, \dots, BL_N^{wlf}$, and DL_F^{wlf} is the fused DL that maximizes the finer details across all of the input DLs $DL_1^{wlf}, DL_2^{wlf}, \dots, DL_N^{wlf}$ computed against all input exposures.

5.3 WEIGHT ESTIMATION

The motivation behind weight map computation is to yield non-linear adaptive function for controlling the contribution of pixels from BLs and DLs computed across all input exposures.

Interestingly, the DL computed for k_{th} source image (DL_k^{wlf}) in Equation (5.1) yields analysis of weak textures that seems to be the indicator of contrast variation in the image. WLF adopted such metric to quantify the perceived local contrast of an image under different exposure settings and allow discarding under exposed and overexposed pixels.

Furthermore, in order to accomplish optimal contrast and color details in the fused image, WLF additionally incorporate the color saturation measure (SA_k) to the final weighting function. In practice for k_{th} RGB source image, SA_k is computed for k_{th} source image as the standard deviation within the R, G, and B channel, at each pixel.

$$SA_k(x, y) = \sqrt{\frac{1}{3} \{ [R(x, y) - \mu(x, y)]^2 + [G(x, y) - \mu(x, y)]^2 + [B(x, y) - \mu(x, y)]^2 \}}$$

$$\text{where } \mu(x, y) = \frac{1}{3} [R(x, y) + G(x, y) + B(x, y)]$$

As shown in Figure 5.1, in order to remove the influences of under-exposed and over-exposed pixels for producing well-exposed image, the two image features i.e., DL_k^{wlf} and SA_k are combined together by multiplication to estimate combined features (FR_k).

$$FR_k(x, y) = DL_k^{wlf}(x, y) \times SA_k(x, y) \quad (5.3)$$

Then FR_k is convolved with the symmetric Gaussian low pass kernel (U) having 5×5 kernel size (r_U) with standard deviation (σ_U) 5 to construct the saliency maps SM_k .

$$SM_k = |FR_k| * U_{r_U, \sigma_U}$$

Next, the saliency maps are compared to determine the weight maps as follows:

$$W_k^{wlf}(x, y) = \begin{cases} 1 & \text{if } SM_k(x, y) = \max(SM_1(x, y), SM_2(x, y), \dots, SM_N(x, y)) \\ 0 & \text{otherwise} \end{cases} \quad (5.4)$$

where N is the number of source images and SM_k^n is the saliency value of the pixel at location (x,y) in the k_{th} image.

5.4 WLS FILTER BASED WEIGHT REFINEMENT

This section presents WLS optimization [133] and sigmoid function based weight map refinement approach to obtain noiseless and smooth weight maps. First, WLS filtering is performed on each weight map W_k^{wlf} with the corresponding source image serving as the source image for the affinity matrix [133]. The WLS is the most commonly used statistical regression analysis model [181]. The motivation behind weight maps refinement is as follows. The fusion rules (weight map) computed in Equation (5.4) are hard (the value of weight maps are changing abruptly), noisy and not aligned with the object boundaries. Weight maps need to be as smooth as possible, since rapid changes in the weight maps will introduce seam and artifacts in fused image.

5.4.1 REGRESSION ANALYSIS

Regression analysis is a standard tool to provided relationship between predictor variable (or independent variable) and response variable (or dependent variable). The most important application of regression analysis tool is in data fitting. Least square is an approach that requires a parametric model in fitting process. The resulting fitted model requires an estimate of the model coefficients which are called unknowns. Therefore, the ultimate goal of regression analysis is to find the unknowns for the model which "best" fits the data.

To obtain the coefficient estimates, the least-squares method minimizes the sum of square of residuals. The residual for the i_{th} data point \mathbf{r}_i is defined as the difference between the observed response value y_i and the fitted response value \hat{y}_i , and is identified as the error associated with the data.

$$\text{residual}(\mathbf{r}_i) = \text{data}(y_i) - \text{fit}(\hat{y}_i) \quad (5.5)$$

Equation (5.5) can be defined as the difference between the actual value of the dependent variable and the value predicted by the model:

$$\mathbf{r}_i = y_i - f(\kappa_i, \beta) \quad (5.6)$$

where κ_i is an independent variable and $f(\kappa_i, \beta)$ is a model function with adjustable parameters or unknowns held in the vector β . The Linear squares problem is said to be

overdetermined if the n is greater than the number of unknowns, the sum of square of \mathbf{r}_i is given by

$$\mathbb{S} = \sum_i^n \mathbf{r}_i^2 = \sum_i^n (y_i - \hat{y}_i)^2 \quad (5.7)$$

$$= \sum_i^n (y_i - f(\kappa_i, \beta))^2 \quad (5.8)$$

where n is the number of data points included in the fit and \mathbb{S} is the sum of squares error estimate.

The least-squares problems fall into two categories: linear least squares and non-least squares, depending on whether or not the residuals are linear in all unknowns [181]. A linear model is defined as an equation that is linear in the parameters to be estimated. To illustrate the linear least-squares fitting process, suppose we have n data points that can be modeled by a first-degree polynomial.

$$\hat{y} = \beta_1 \kappa + \beta_2 \quad (5.9)$$

This model represents a straight line in two dimensions. To solve this equation for the unknown coefficients β_1 and β_2 , we write \mathbb{S} as a system of n simultaneous linear equations in two unknowns. If n is greater than the number of unknowns, then the system of equations is overdetermined.

$$\mathbb{S} = \sum_{i=1}^n (y_i - (\beta_1 \kappa_i + \beta_2))^2 \quad (5.10)$$

Because the least-squares fitting process minimizes the summed square of the residuals, the coefficients are determined by differentiating \mathbb{S} with respect to each parameter (i.e. unknown), and setting the result equal to zero.

$$\frac{\partial \mathbb{S}}{\partial \beta_1} = -2 \sum_{i=1}^n \kappa_i (y_i - (\beta_1 \kappa_i + \beta_2)) = 0 \quad (5.11)$$

$$\frac{\partial \mathbb{S}}{\partial \beta_2} = -2 \sum_{i=1}^n (y_i - (\beta_1 \kappa_i + \beta_2)) = 0 \quad (5.12)$$

Substituting b_1 and b_2 for β_1 and β_2 , the previous equations become:

$$\sum_{i=1}^n \kappa_i (y_i - (b_1 \kappa_i + b_2)) = 0 \quad (5.13)$$

$$\sum_{i=1}^n (y_i - (b_1 \kappa_i + b_2)) = 0 \quad (5.14)$$

The normal equations are defined as

$$b_1 \sum_{i=1}^n x_i^2 + b_2 \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i \quad (5.15)$$

$$b_1 \sum_{i=1}^n x_i + n b_2 = \sum_{i=1}^n y_i \quad (5.16)$$

Solving for b_1

$$b_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (5.17)$$

Solving for b_2 using the b_1 value

$$b_2 = \frac{1}{n} \left(\sum_{i=1}^n y_i - b_1 \sum_{i=1}^n x_i \right) \quad (5.18)$$

In various practical applications the assumption of constant error variance which is also known as homoscedasticity of variance is violated. Let us consider the case where the data is of unequal quality and, therefore, has no correlations exist among the observed variances which is known as heteroscedasticity [182]. In such type of cases some observations are important than the other and need to be given more weights to important observations. In Figure 5.2 the examples of homoscedastic and heteroscedastic data set of the least squares estimator for straight lines are provided.

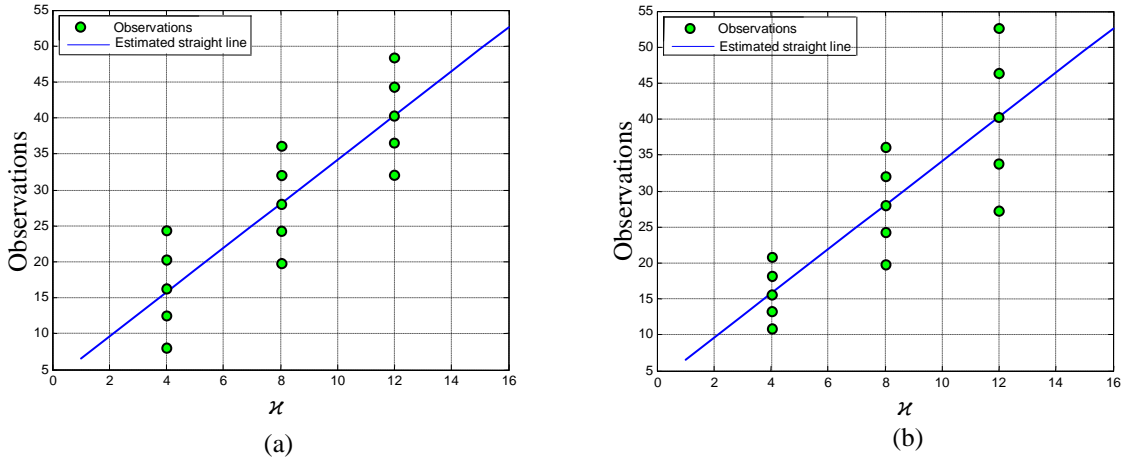


Figure 5.2: Illustration of uncorrelated variances. (a) homoscedasticity: there is certain correlation between variances across individuals and (b) heteroscedasticity: there is no correlation between variances across individuals.

If, however, the measurements are uncorrelated but have different uncertainties, an ordinary least squares approach yields under-fitting or over-fitting. To improve the fit, an

additional scale factor (the weight) is included in the fitting process for minimizing the weighted residual sum of squares:

$$\mathbb{S} = \sum_i^n \mathcal{W}_i r_i^2 = \sum_i^n \mathcal{W}_i (y_i - \hat{y}_i)^2 \quad (5.19)$$

where \mathcal{W}_i is a non-negative valued weight given to the i_{th} observation.

The expression given above is known as WLS estimation in which the weights are inversely proportional to the corresponding variances; points with low variance will be given higher weights and points with higher variance are given lower weights:

$$\mathcal{W}_i = \frac{1}{\sigma^2} \quad (5.20)$$

WLS is a robust regression analysis model in which outliers are given less weight to improve fitting. Suppose that we have model given in Equation (5.9), the weighted least squares estimates of $\hat{\beta}_1$ and $\hat{\beta}_2$ minimize the quantity

$$\mathbb{S} = \sum_{i=1}^n \mathcal{W}_i (y_i - (\hat{\beta}_1 x_i + \hat{\beta}_2))^2 \quad (5.21)$$

The WLS estimates or unknowns are formulated as

$$\hat{b}_1 = \frac{\sum_{i=1}^n \mathcal{W}_i (x_i - \bar{x}_w) (y_i - \bar{y}_w)}{\sum_{i=1}^n \mathcal{W}_i (x_i - \bar{x}_w)^2} \quad (5.22)$$

$$\hat{b}_2 = \bar{y}_w - \hat{b}_1 \bar{x}_w \quad (5.23)$$

where \bar{x}_w and \bar{y}_w are the weighted means

$$\bar{x}_w = \frac{\sum_{i=1}^n \mathcal{W}_i x_i}{\sum_{i=1}^n \mathcal{W}_i} \quad \text{and} \quad \bar{y}_w = \frac{\sum_{i=1}^n \mathcal{W}_i y_i}{\sum_{i=1}^n \mathcal{W}_i} \quad (5.24)$$

For example, the result of fitting a straight line function $y_i = \beta_1 x_i + \beta_2$ through a set of data points (x_i, y_i) : (1,5), (2,6), (3,7) and (4,9) is shown in Figure 5.3 (a). In the plot shown in Figure 5.3 (a), the residuals for data points 3 (i.e. -0.4) and 4 (i.e. 0.3) are shown in red line. We can notice that the least residual is obtained for data point 2 (i.e. -0.1) and the sum of square of the residuals for all data points \mathbb{S} is 0.3. Moreover, for randomly generated data points with normally distributed noise (in green), the least squares line of best fit (in blue), is shown in Figure 5.3 (b).

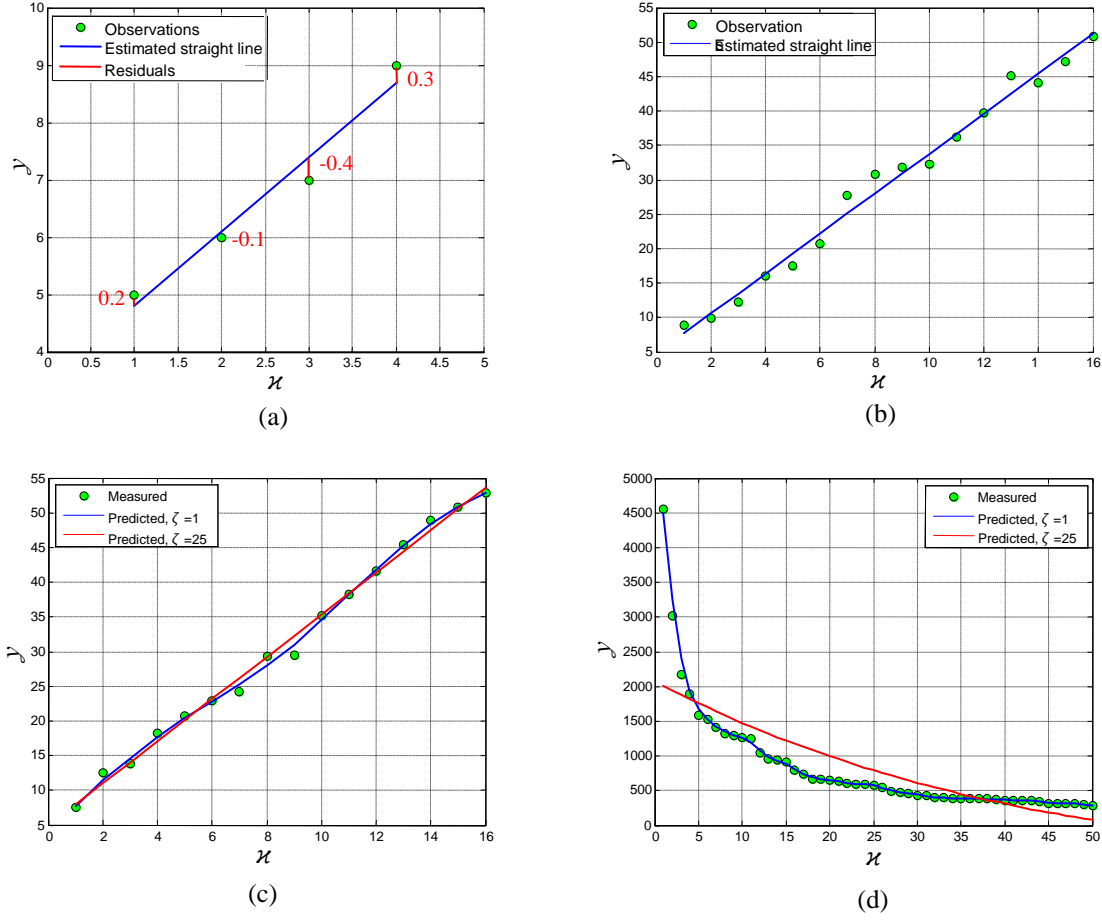


Figure 5.3: Illustration of least squares line of best fit. (a) least squares line of best fit for four data points (1,5), (2,6), (3,7), (4,9); (b) least squares line of best fit for randomly generated data points with normally distributed noise; (c) WLS line of best fit for randomly generated data points with normally distributed noise; and (d) WLS line of best fit for randomly generated data points with normally distributed noise.

Exponential weighting function is a standard choice for weighting function [183]. The formula for the exponential weighting function in locally weighted linear regression is given as

$$\mathcal{W}_i = \exp\left(-\frac{(\kappa_i - \kappa)^2}{2\zeta^2}\right) \quad (5.25)$$

Note that the weights depend on the particular query point at which κ need to be evaluated. If $|\kappa_i - \kappa|$ is small, then \mathcal{W}_i is close to 1; and if $|\kappa_i - \kappa|$ is large, then \mathcal{W}_i is small. The parameter ζ controls how quickly the weight of a training example falls off with distance of its κ_i from the query point κ ; ζ is called the bandwidth parameter. The WLS regression analysis using exponential weighting function in Equation (5.25) on two different types of data sets is shown in Figure 5.3 (c) and (d). Please note that the fitted curve passes through the data points perfectly. Notice that least squares line of best fit tends towards a classical

straight line fit as $\zeta \rightarrow \infty$, and for smaller value of $\zeta=1$, the measured and predicted values are almost on top of each other. Therefore, for a higher value of ζ , the predicted value is close to the curve obtained from the no weighting case.

5.4.2 EDGE PRESERVING PROPERTY OF WLS FILTER

Consider the basic linear transformation equation discussed in Chapter 4, and is rewritten in Equation 5.28 as follows:

$$\hat{q}_i = \hat{a}_k I_i^G + \hat{b}_k, \forall i \in \omega_k \quad (5.26)$$

WLS [133] based edge-preserving operator may be viewed as a compromise between two possible contradictory goals. Given an input image p , we seek a new image \hat{q} , which, on the one hand, is as close as possible to p , and at the same time, is as smooth as possible everywhere, except across significant gradients in p . To achieve these objectives we seek to minimize the following quadratic functional:

$$\sum_{i \in \omega_k} \left((\hat{q}_i - p_i)^2 + \psi \left(c_{x,i}(p) \left(\frac{\partial \hat{q}}{\partial x} \right)_i^2 + d_{y,i}(p) \left(\frac{\partial \hat{q}}{\partial y} \right)_i^2 \right) \right) \quad (5.27)$$

where the subscript i denotes the spatial location of a pixel. The goal of the data term $(\hat{q}_i - p_i)^2$ is to minimize the distance between \hat{q} and p , while the second (regularization) term strives to achieve smoothness by minimizing the partial derivatives of \hat{q} . The smoothness requirement is enforced in a spatially varying manner via the smoothness weights c_x and d_y , which depend on p . Finally, ψ is responsible for the balance between the two terms; increasing the value of ψ results in progressively smoother images \hat{q} .

The linear coefficients used to minimize the cost function in Equation (5.26) are determined as follows:

$$\hat{a}_k = \frac{\frac{1}{|\omega|} \sum_{i \in \omega_k} I_i^G p - \mu_k \bar{p}_k}{\sigma_k^2 + \frac{1}{n} \psi P} \quad (5.28)$$

$$\hat{b}_k = \bar{p}_k - \hat{a}_k \mu_k \quad (5.29)$$

$$\bar{p}_k = \frac{1}{|\omega|} \sum_{i \in \omega_k} p_i \quad (5.30)$$

The full derivation of the computation of \hat{a}_k and \hat{b}_k is available in Appendix B. In Equations (5.28), P is the regularization term given in Equation (5.31):

$$P = c_{x,i}(p) \left(\frac{\partial \hat{q}}{\partial x} \right)_i^2 + d_{y,i}(p) \left(\frac{\partial \hat{q}}{\partial y} \right)_i^2 \quad (5.31)$$

5.4.3 WEIGHT MAP REFINEMENT FOR BL

Using matrix notation we may rewrite Equation (5.27) in following quadratic form:

$$(\hat{q} - p)^T (\hat{q} - p) + \psi (\hat{q}^T O_x^T V_x O_x \hat{q} + \hat{q}^T O_y^T V_y O_y \hat{q}) \quad (5.32)$$

Here V_x and V_y are diagonal matrices containing the smoothness weights $c_x(v)$ and $d_y(v)$, respectively, and the matrices O_x and O_y are discrete differentiation operators.

The vector \hat{q} that minimizes Equation (5.32) is uniquely defined as the solution of the linear system

$$(I_{dent} + \psi L_v) \hat{q} = p, \quad (5.33)$$

where I_{dent} is the identity matrix and $L_v = O_x^T V_x O_x + O_y^T V_y O_y$. Modulo the difference in notation, this is exactly the linear system used in [165], where it was primarily used to drive piecewise smooth adjustment maps from a sparse set of constraints.

In the present approach, O_x and O_y are forward difference operators, and hence O_x^T and O_y^T are backward difference operators, which means that L_v is a five-point spatially inhomogeneous Laplacian matrix. As for the smoothness weights, we define in the same manner as in [165]:

$$c_{x,i}(p) = \left(\left| \frac{\partial \ell}{\partial x}(i) \right|^{\alpha_4} + \varepsilon' \right)^{-1} \text{ and } d_{y,i}(p) = \left(\left| \frac{\partial \ell}{\partial y}(i) \right|^{\alpha_4} + \varepsilon' \right)^{-1} \quad (5.34)$$

where ℓ is the log-luminance channel of the input image p , the exponent α_4 (typically between 1.2 and 2.0) determines the sensitivity to the gradients of p , while ε' is a small constant (typically 0.0001) that prevents division by zero in areas where p is constant.

Equation (5.34) tells us that \hat{q} is obtained from p by applying a nonlinear operator Z_ψ , which depends on p :

$$\hat{q} = Z_\psi(p) = (I_{dent} + \psi L_v)^{-1} p \quad (5.35)$$

The proof of Equation (5.35) is given in Appendix C. In the present approach, let us assume that $\mathbb{W}_{\psi, \alpha_4, \varepsilon'}(p)$ represents the WLS filtering operation on input p . Where ψ , α_4 , and ε' are the parameters which decide the degree of smoothness, sensitivity to the gradients and small constant of the WLS filter, respectively. In WLF method W_k^{wlf} computed in Equation

(5.4) serves as the input to WLS filter (i.e. $p = W_k^{wlf}$). More specifically, the coarser version of weight map W_k^{wlf} will serve as refined weight map for k_{th} BL $W_{BL_k}^{wlf}$:

$$W_{BL_k}^{wlf} = \mathbb{W}_{\psi, \alpha_4, \varepsilon'}(p) \quad (5.36)$$

5.4.4 WEIGHT MAP REFINEMENT FOR DETAIL LAYER

Once the resulting weight maps for BLs are obtained, sharp and edge-aligned weights are computed based on 1-D sigmoid function for fusing the DLs. As shown in Figure 5.4, the spatially smoothed weight maps of BLs are utilized to compute sharp weight mask of DL which preserve texture details in the fused image. Therefore, unlike [115], the proposed solution attempts computationally simple approach to estimate the best possible weight maps for DLs fusion.

Let $W_{DL_k}^{wlf}$ denote refined weight map for k_{th} DL and $\mathcal{S}_{\hat{a}, t_s, \theta}(W_{BL_k}^{wlf})$ is the 1-D sigmoid function applied on $W_{BL_k}^{wlf}$, where $\hat{a} \in \mathcal{R}$, $t_s \in \mathcal{R}$, and θ are the weight parameter, independent variable, and the parameter which decide the threshold to further control the degree of sharpness, respectively. Then $W_{DL_k}^{wlf}$ is computed as:

$$W_{DL_k}^{wlf} = \mathcal{S}_{\hat{a}, t_s, \theta}(W_{BL_k}^{wlf}) \quad (5.37)$$

In theory, the 1-D sigmoid is computed as:

$$\mathcal{S}_{\hat{a}, t_s, \theta}(t_s) = \frac{1}{1 + e^{(-\hat{a}t_s + \theta)}} \quad (5.38)$$

The detailed description of sigmoid function is given in Sub-section-3.5.1 (Chapter 3).

5.5 WEIGHTED FUSION OF COARSER DETAILS AND FINER DETAILS

Once the resultant weight maps $W_{DL_k}^{wlf}$ and $W_{BL_k}^{wlf}$ are obtained, the pixel-wise weighted fusion of BLs (i.e. BL_F^{wlf}), and DLs (i.e. DL_F^{wlf}) is computed as follows:

$$BL_F^{wlf}(x, y) = \sum_{k=1}^N W_{BL_k}^{wlf}(x, y) BL_k^{wlf}(x, y) \quad (5.39)$$

$$DL_F^{wlf} = \sum_{k=1}^N W_{DL_k}^{wlf}(x, y) DL_k^{wlf}(x, y) \quad (5.40)$$

In WLF approach, once the fused BL and fused DL is obtained from Equation (5.39) and Equation (5.40), respectively, the detail enhanced fused image ($I_F^{wlf}(x, y)$) is obtained by Equation (5.2).

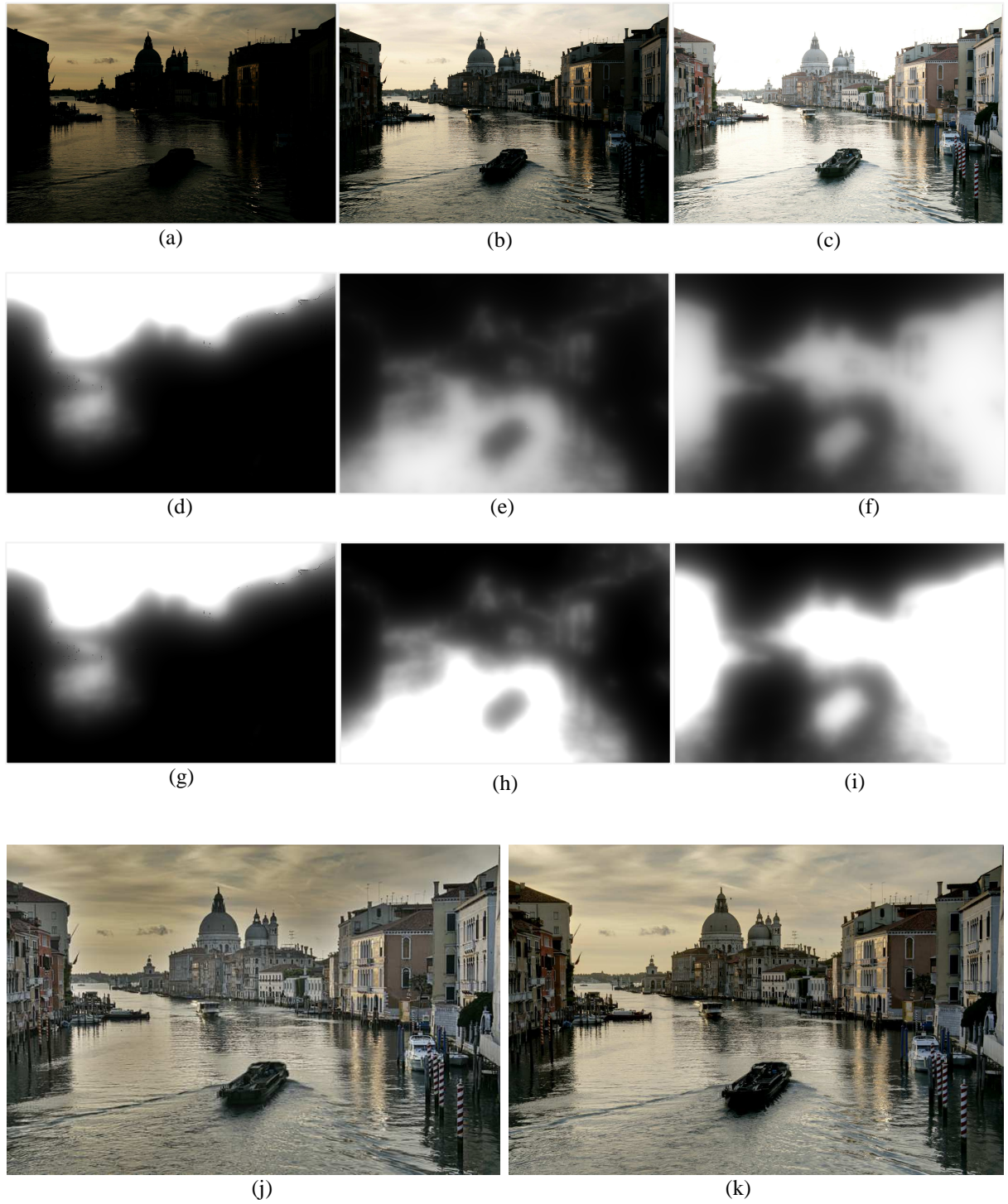


Figure 5.4: Illustration of weight map refinement and fused image using WLF approach. (a-c) Input exposures; (d-f) weight maps computed for BLs across input exposures; (g-i) weight maps computed for DLs across input exposures; (j) WLF results; and (k) results proposed by Merten's et al. [113]

5.6 SUMMARY

The approach developed in this chapter constructs a detail enhanced fused image from a set of multi-exposure images by using WLS optimization framework. In particular the utilization of the weak textures and saturation measures are investigated for weight map computation. We consider the weight map optimization criterion that yields weighted average fusion of BLs and DLs. As in Chapter 3 and 4, two scale decomposition based on AD and GF are best formulated that helps to produce detail enhanced fused image using MRP, but suffer from some problem. Common drawback of pyramidal image structure is an increase in the computational complexity of the fusion approach. In the WLF approach, since we specifically generate smooth and noise-less weight map functions that are utilized for the fusion of BLs and DLs without pyramidal decomposition.

As described in Chapter 3, the AD based edge-preserving filter is applied on the input exposures, which is utilized to extract fine textures for detail enhancement. WLF also include a sigmoid function for DLs fusion that refines and enhances the weight mask function in order to prevent the artifacts due to over-enhancement of strong edges or low contrast areas. In this way, the influence of very bright and very dark regions of the scene on a fused image is decreased, thus producing well-exposed detail enhanced image. In forthcoming chapter, when compared with the existing techniques which use multi-resolution and single resolution analysis for exposure fusion, the WLF approach perform better in terms of enhancement of texture details in the fused image.

CHAPTER-VI

EXPERIMENTAL RESULTS AND ANALYSIS

6.1 ANISOTROPIC DIFFUSION FILTER BASED EXPOSURE FUSION

6.1.1 COMPARISON WITH OTHER EXPOSURE FUSION AND TONE-MAPPING METHODS

The approach proposed in this thesis implemented ADF approach in MATLAB-7.5.0 and run on a PC with 2.2 GHz i5 processor and 2GB of RAM. As shown in Figures 6.1(b) and Figure 6.2(b), note that the fused image provides natural contrast and has no noticeable artifacts. We are testing ADF approach on a variety of bracketed sequences of multi-exposure and multi-focus image databases. Figure 3.1 (in Chapter 3) shows the block diagram of the proposed texture feature based detail enhancing exposure fusion technique.

The comparisons of experimental results of proposed ADF approach are shown in Figures 6.1, 6.2, 6.3, and 6.4. In these experiments, the optimal block size used for weight map calculation was 3-by-3. Figures 6.1(a) and 6.2(a) show image pairs of the “Igloo” and “Door” image sequence (size of $221 \times 336 \times 6$ and $223 \times 332 \times 6$, respectively). It can be seen from Figure 6.1(b) that all the light in the scene that appears to come from natural light source is optimally reproduced with crisp shadows. In Figure 6.1, one auto-exposure image captured with the digital camera and two recently proposed fusion results of “Igloo” are illustrated. It can be noticed that the ADF technique provides better texture details in highlights and shadows as compare to the results of auto-exposure (Figure 6.1(c)) and Mertens et al. [113] (Figure 6.1(d)). It may also be observed that the brightly illuminated region (i.e., sky area) is overexposed in the result proposed by Shen et al. [116] (see Figure 6.1(e)). Figure 6.1(b) shows more comparison example of ADF results for scene depicting outdoor and indoor details. The ADF technique is visually compared with the results of auto-exposure (Figure 6.2(c)), and recently proposed Mertens et al. [113] (Figure 6.2(d)) and Zhang and Cham [120] (Figure 6.2(e)). By contrast, it is seen that ADF combines the best of multiple exposures into one realistic-looking image that is much closer to what our eyes originally saw. However, both indoor and outdoor details of input LDR images (Figure 6.2(a)) are simultaneously produced in the fused image with optimal contrast and without the

introduction of artifacts. Although Mertens et al. [113] have produced comparable results; but it does not preserve all texture details from input LDR images. As shown in Figure 6.2(e), the results produced by Zhang

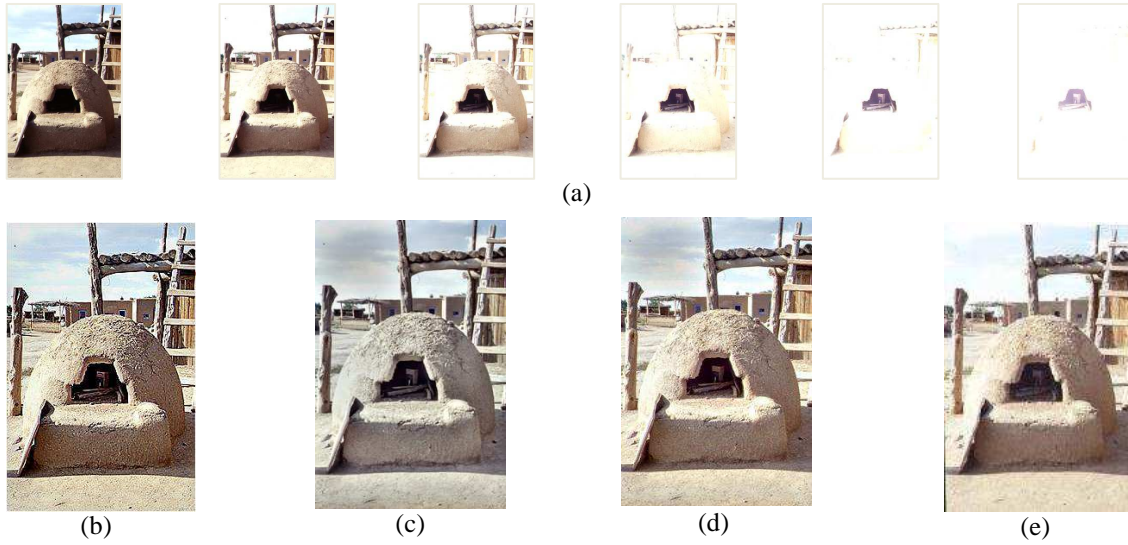


Figure 6.1: (a) Images representing multiple exposures of an outdoor scene depicting highlights from natural sun light and shadows; (b) ADF approach fuses the multiple exposures to obtain high quality image. Note the fused image yields more texture details and natural contrast without the introduction of artifacts. (c) Auto-exposure (d) Mertens et al. [113] (e) Shen et al. [116]. Input images courtesy of Shree Nayar.

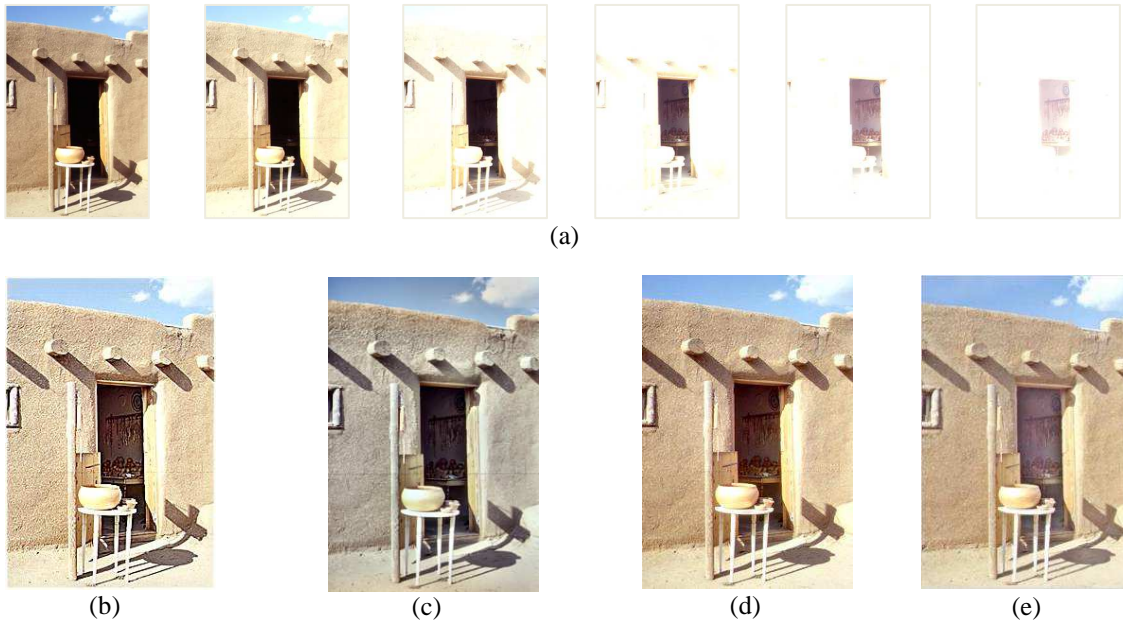


Figure 6.2: (a) Images representing multiple exposures of an indoor and outdoor scene depicting sunlit details and shadows; (b) ADF approach fuses the multiple exposures to obtain high quality image. Note the fused image yields more texture details and natural contrast without the introduction of artifacts. (c) Auto-exposure (d) Mertens et al. [113] (e) Zhang & Cham [120]. Input images courtesy of Shree Nayar.



Figure 6.3: “House” image series: Comparison of ADF results to other recently proposed exposure fusion techniques. (a) Results of ADF method; (b) Mertens et al. [113]; (c) Shen et al. [116]. Note that ADF yields enhanced texture and edge features. Input image sequence courtesy of Tom Mertens.

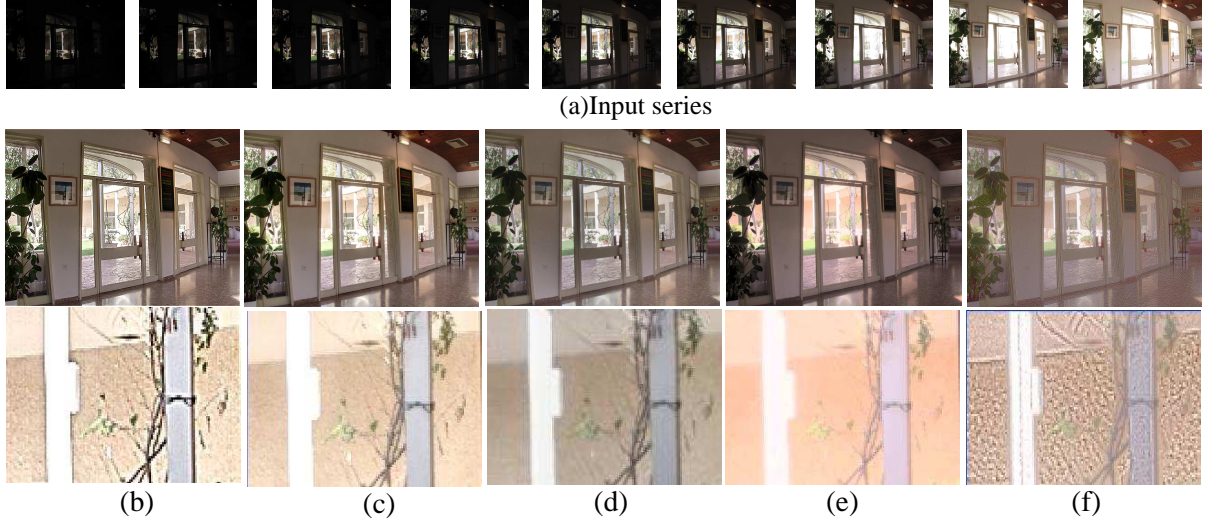


Figure 6.4: “Belgium House” image series: (a) Series of multiple exposures depicting both indoor and outdoor areas. The exposure value is varying from (1/1000 of a second) to (1/4 of a second). Comparison results to other popular tone mapping techniques. (b) Results of ADF method, window size = 3-by-3; (c) Mertens et al. [113] (d) Shen et al. [116] (e) Results of Ward Larson et al. [75]; (f) Results of LCIS method [184]. Note that ADF yields combined features in the fused image that can only be recorded using different exposures. Input images courtesy of Dani Lischinski.

and Cham [120] depict washed out details in underexposed regions which are not able to preserve texture details from input LDR shots.

To further compare ADF results visually with Mertens et al. [113] and Shen et al. [116], respectively, Figures 6.3(a), 6.3(b), and 6.3(c) depict a close-up view. Figure 1.3 depicts the “House” LDR image sequence of size $752 \times 500 \times 4$ which is provided by Mertens et al. [113]. It can be observed that the texture details (see the fine textures on the

chair and books behind the chair) are accurately preserved in the proposed fused image (see Figure 6.3(a)).

This section compares ADF results for “Belgium house” image sequence of size $1025 \times 769 \times 9$ (see Figure 6.4(a)) with the popular exposure fusion and tone-mapped HDR images, which are depicted in Figures 6.4 (b), 6.4 (c), 6.4 (d), 6.4 (e), and 6.4 (f). In particular, the results of ADF are compared with the perceptually driven works [75] and Low Curvature Image Simplifier (LCIS) hierarchical decomposition [184]. As shown in Figure 6.4 (b), ADF approach yields fine texture details in the fused image with natural contrast that is entirely free of halo artifacts. To illustrate the effectiveness of proposed ADF technique, the close-up comparison of results is illustrated in Figures 6.4 (b-f). Larson et al. [75] presented a DR compression method based on a HVS adaptation, and it was also found to suffer from halo artifacts and does not offer good color information (see Figure 6.4(e)). Tumblin and Turk [184] preserve fine details in the image, while weak halo artifacts are present around certain edges in strongly compressed areas (see Figure 6.4(f)). Experimental results have demonstrated that ADF method worked very well on a variety of multiple exposures and preserved the original scene’s relative visual contrast impression.

6.1.2 MULTI-FOCUS IMAGE FUSION

Furthermore, to check the effectiveness of the ADF algorithm for other applications, we have employed the same technique for the fusion of multi-focus image series (Figures

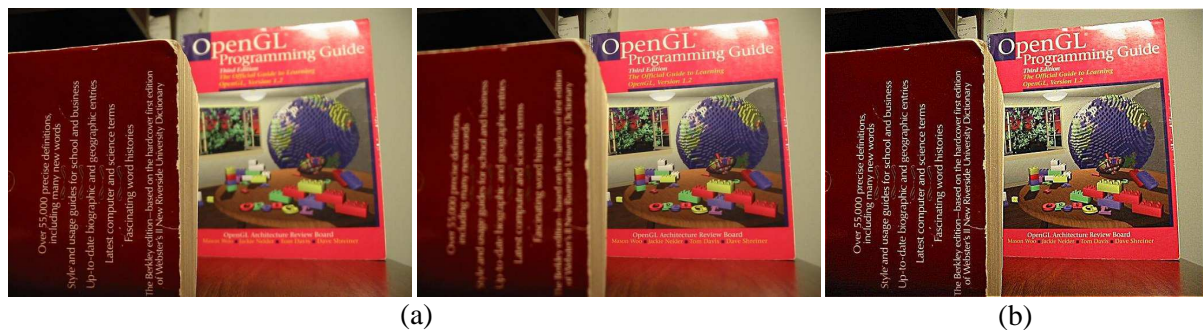


Figure 6.5: “Book” image series. (a) Two partially focused images, (focused on different targets), and (b) image generated by the ADF approach, which illustrate that the fused image extracts more color and texture details from the original input images. (Input sequence courtesy of ADu and Wang [110])

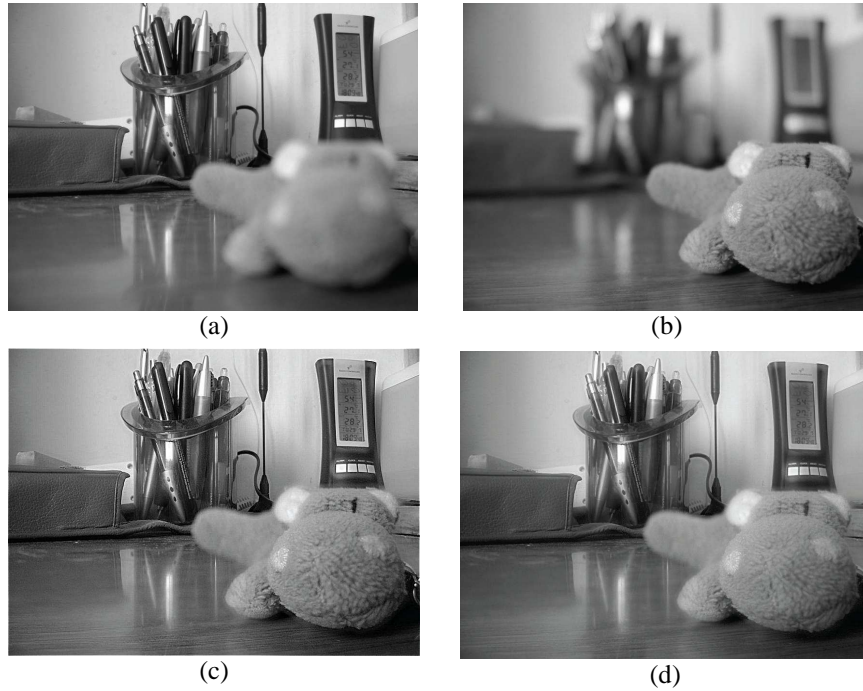


Figure 6.6: “Table” image series. (a,b) two partially focused images, (focused on different targets), (c) image generated by the ADF approach, which illustrate that the fused image extracts more information from the original images and (d) Hod’akov’a et al. [185].

6.5, 6.6, and 6.7) and images captured with flash and no-flash (see Figure 6.8). Figure 6.5(a) illustrates two partially focused RGB images (focused on two different targets). It is illustrated in Figure 6.5(b) that the color information is preserved in the fused image with better visualization of texture details. On the other hand ADF approach have been tested and compared for two sets of multi-focused gray scale images of “table” and “clock”, which are illustrated in Figures 6.6(a)–6.6(d) and Figures 6.7(a)–6.7(d), respectively. As demonstrated in Figure 6.6(c) that results of ADF approach produce pleasing image with rich texture details, the results produced by P. Hod’akov’a et al. [185] in Figure 6.6(d) do not reveal fine details present across all input images. It can easily be noticed that ADF based fused image shown in Figure 6.7(c) extracts more information from the original images. Moreover, Adu and Wang technique [110] in Figure 6.7(d) appears washed out, which is responsible for losing perception of fine texture details.

6.1.3 FLASH AND AMBIENT IMAGE FUSION

Finally, ADF technique has also been tested on two sets of images captured with flash and no-flash images (see Figures 6.8(a) and 6.8(b)). ADF approach also provides interesting solution for fusing the flash/no-flash image pair. Figure 6.8(c) illustrates present results,

which combine details from the flash/no-flash image pair. As shown in Figure 6.8(c), the ADF approach allows removal of highlights from flash images and yields high quality flash image with optimal contrast and detail enhancement. The experimental results in Figure 6.8(c) depict largest amount of information and has relatively better contrast than that of results of Mertens et al. [113] in Figure 6.8(d).

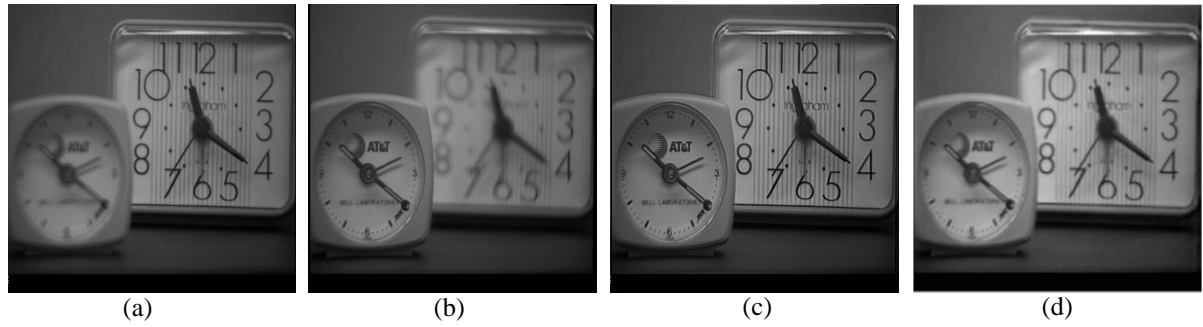


Figure 6.7: “Clock” Image series. (a,b) two partially focused images, (focused on different targets), (c) image generated by the ADF approach, which illustrate that the fused image extracts more information from the original images and (d) ADu and Wang [110]. (Input sequence courtesy of ADu and Wang [110]).

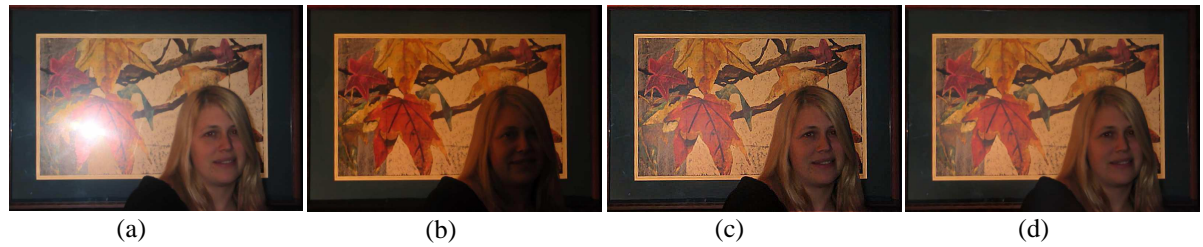


Figure 6.8: (a,b) input images photographed with and without flash; (c) Enhanced fused image by ADF algorithm which maintains the warm appearance, and the sharp details after removing strong highlight and (d) Mertens et al. [113]. Images taken from Agrawal et al. [98]

To perform visual inspection of exposure fusion results of Mertens et al. [113] shown in Figures 6.1(d), 6.2(d), 6.3(b), 6.4(c), and 6.8(d) are produced with the help of MATLAB code provided by the authors. The original results of GRW based fusion [116] in Figures 6.1(e), 6.3(c), and 6.4(d) are provided by the authors on request. All the experimental results of Zhang and Cham [120] in Figure 6.2(e), tone-mapped HDR [75,184] in Figures 6.4(e) and 6.4(f), and multi-focused fusion [110, 185] in Figures 6.6(d) and 6.7(d) are taken from its papers. It is noticed that unlike the previous work such as [116], ADF approach preserves more details with higher contrast and does not require further post-processing. Thus, this approach can be utilized in computer graphics applications.

6.1.4 ANALYSIS OF FREE PARAMETERS

To analyze the effect of iteration (t_a) parameter used in AD on quality score [186], entropy, computational time and Mean Square Error (MSE), four plots (see Figures 6.9(a), 6.9(b), 6.9(c), and 6.9(d), respectively) are illustrated that are plotted at a different value of iteration (t_a) for input image sequences of “House,” “Igloo,” and “Door.” To assess the effect of iteration on fusion performance, the quality score [186] and entropy were adopted in all experiments. To measure computational time, all the experiments were executed on a PC with 2.2 GHz i5 processor and 2 GB of RAM. The MSE is estimated as the difference between pixel values implied by different iterations (i.e., $t_a = 2, 3, 4, 5, 6, 7, 8$) and the reference image obtained with low iteration value (i.e., $t_a = 1$). ADF fixes $\lambda = 1/7$, $K = 30$, $\alpha_2 = 2$, and $\hat{a} = 27$ in all experiments and they are set as default parameters.

First, to analyze the effect of iteration on quality score, entropy, and computational time the threshold (θ) used in Equation (3.28) for scale selection was set to 0.002. As shown in Figures 6.9(a) and 6.9(b), the best fusion performance is given at $t_a = 1$. The quality score and entropy decrease as t_a increases. As shown in Figure 6.9(c), the computational time increases as t_a increases. The visual inspection of effect of t_a on image sequences (i.e., “House” and “Igloo”) is depicted in Figures 6.10 and 6.11, respectively. It can easily be noticed from the close up view (please see Figures 6.10(b), 6.10(c), 6.10(d), 6.11(b), 6.11(c), and 6.11(d)) that as t_a increases, the sharp edges get brighter and therefore lead to artifacts at sharp edges. To analyze the error (i.e., MSE) introduction against t_a the one of the image produced with $t_a = 1$, $\lambda = 1/7$, $K = 30$, $\alpha_2 = 2$, $\theta = .002$, and $\hat{a} = 27$ is considered as reference image. The error increases as the number of iterations (t_a) increases. From Figure 6.9(d), it can also be noticed that when $t_a = 8$, the total error introduced is still less than 9%.

In the analysis of threshold (θ), ADF fixes $t_a = 1$, $\lambda = 1/7$, $K = 30$, $\alpha_2 = 2$, and $\hat{a} = 27$. Four results obtained by different θ 's are shown in Figures 6.12(a-d). For the result in Figure 6.12(a), the value of θ is .002, and in Figures 6.12(b-d) the values of θ are .003, .004, and .005. Increasing the value of θ for controlling the sharpness of sigmoid function reveals more details in strongly illuminated areas (i.e., overexposed regions) and the image gets darker. In order to balance the details and contrast in ADF approach, it is found that $\theta = .002$ generates reasonably good results for all cases. Finally, from these experiments, it is concluded that the best results were obtained with $t_a = 1$, $\lambda = 1/7$, $K = 30$, $\alpha_2 = 2$, $\theta = .002$, and $\hat{a} = 27$, which yield more details and good contrast.

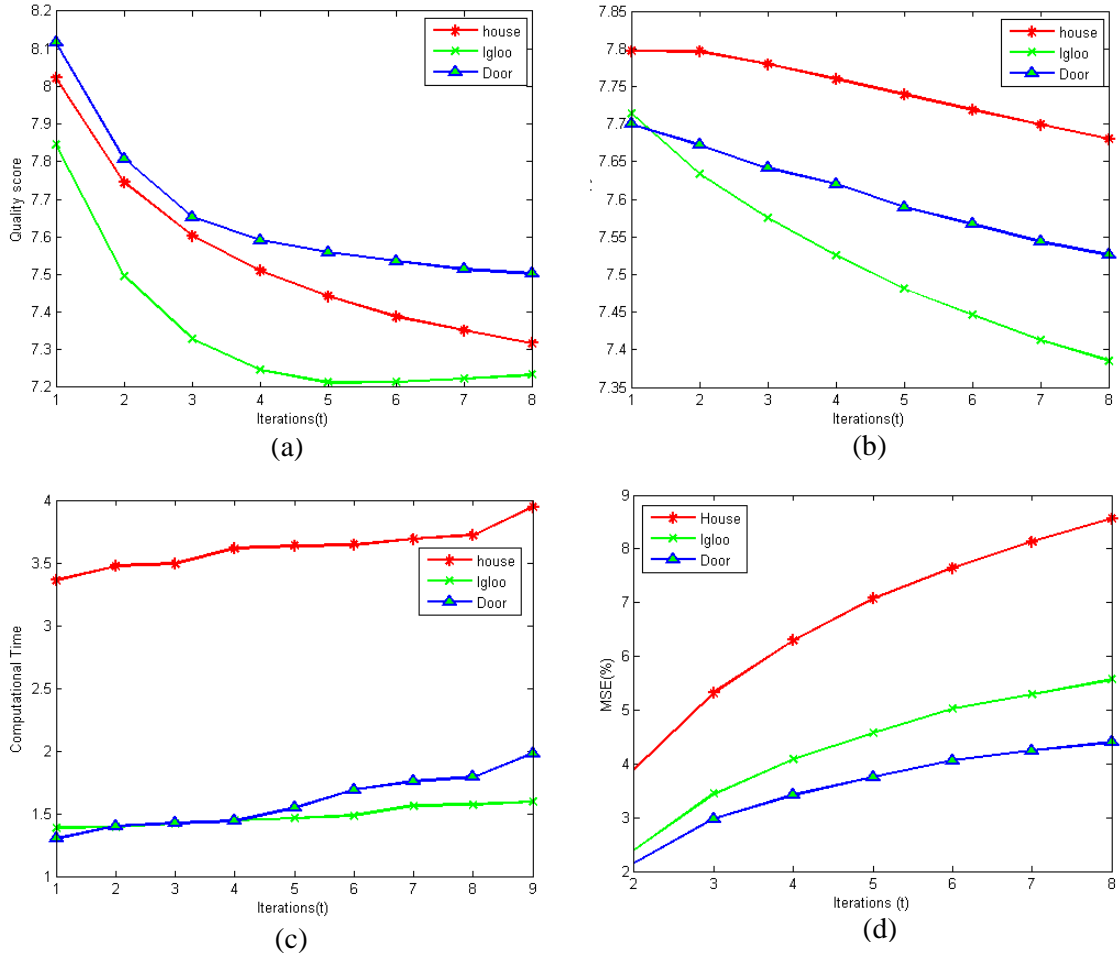


Figure 6.9: Analysis of number of iterations used for BL computation. MSE is defined as the relative difference from the results generated with $t_a = 1$. Maximum quality score and entropy is only observed when $t_a = 1$. It is observed that MSE and computational time increases as t_a increases. (a) Effectiveness of t_a on quality score; (b) Entropy; (c) Computational time; and (d) Error introduced.

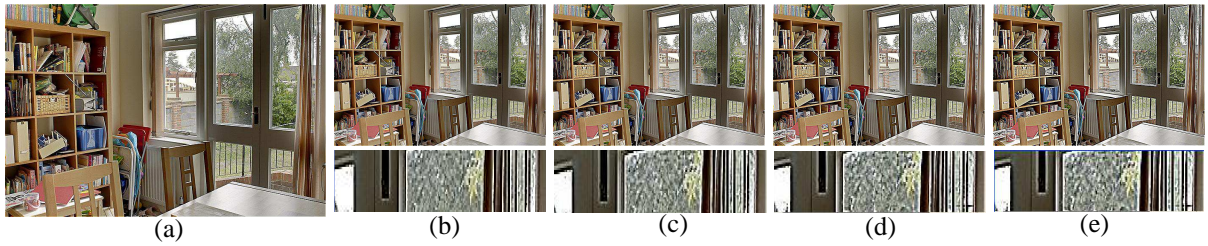


Figure 6.10: “House” image series. The free parameter t_a in Equation (3.11) is used to controls detail enhancement. It is found that $t_a=1$ is sufficient for fine details extraction and gives better results for most cases. Higher value of t_a brings in artifacts near strong edges. (a); $t_a=1$ (b) $t_a=2$; (c) $t_a=3$; (d) $t_a=4$; and (e) $t_a=5$.

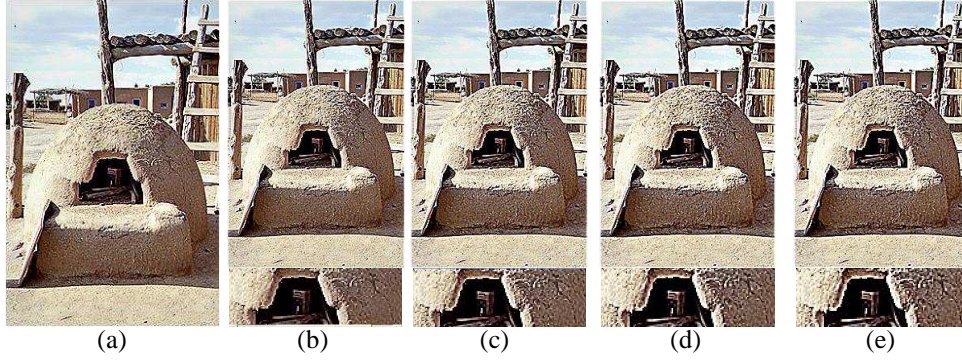


Figure 6.11: “Igloo” image series. The free parameter t_a in Equation (3.11) is used to controls detail enhancement. It is found that $t_a=1$ is sufficient for fine details extraction and gives better results for most cases. Higher value of t_a brings in more details while introducing artifacts near strong edges. (a); $t_a=1$ (b) $t_a=2$; (c) $t_a=3$; (d) $t_a=4$; and (e) $t_a=5$.

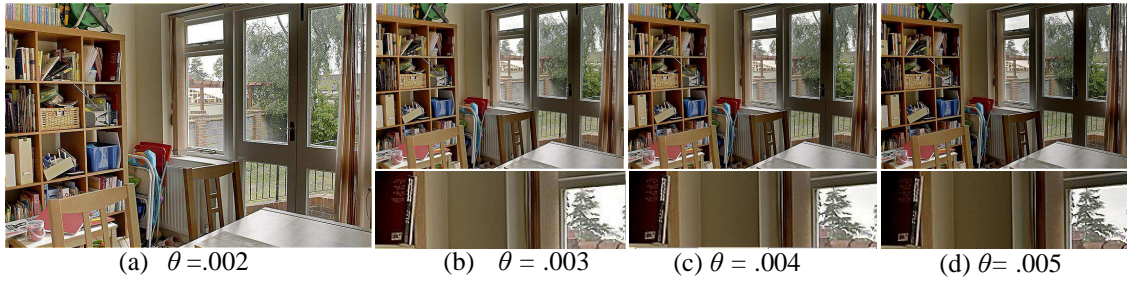


Figure 6.12: “House” image series. The free parameter θ in Equation (3.28) is used to control sharpening. It is found that $\theta=.002$ gives better results for most cases. Higher value of θ brings in more details in highly illuminated areas. (a) $\theta=.002$; (b) $\theta=.003$; (c) $\theta=.004$; and (d) $\theta=.005$.

6.2 GUIDED FILTER BASED EXPOSURE FUSION

6.2.1 COMPARISON WITH OTHER EXPOSURE FUSION

Figures 6.13, 6.14, 6.15, and 6.16 depict examples of fused images from the multi-exposure images. It is noticed that the GFF enhances texture details while preventing halos near strong edges.

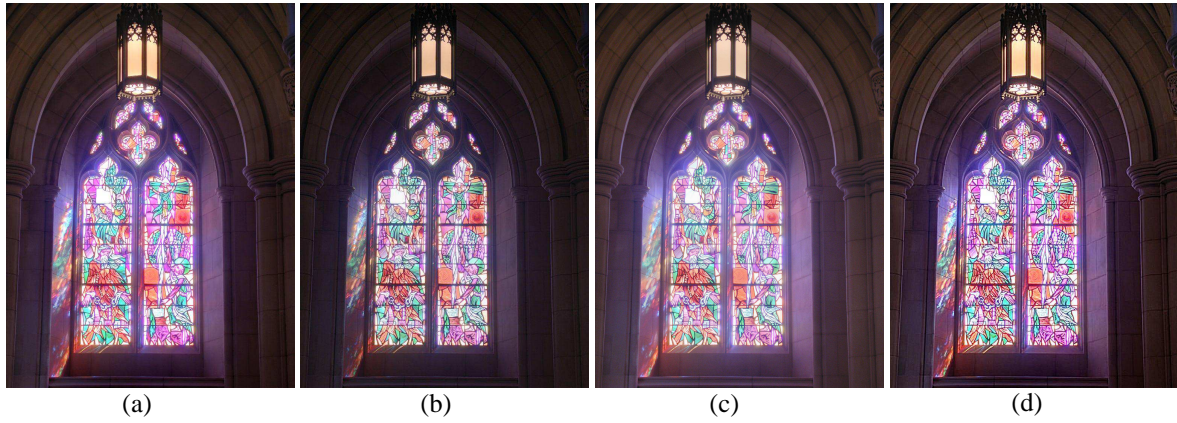


Figure 6.13: Comparison results to other recent exposure fusion techniques. (a) Mertens et al. [113], (b) Burt et al. [89], (c) Shen et al. [116] and (d) Results of GFF method. Note that GFF method yields enhanced texture and edge features. Input image sequence is courtesy of Tom Mertens.

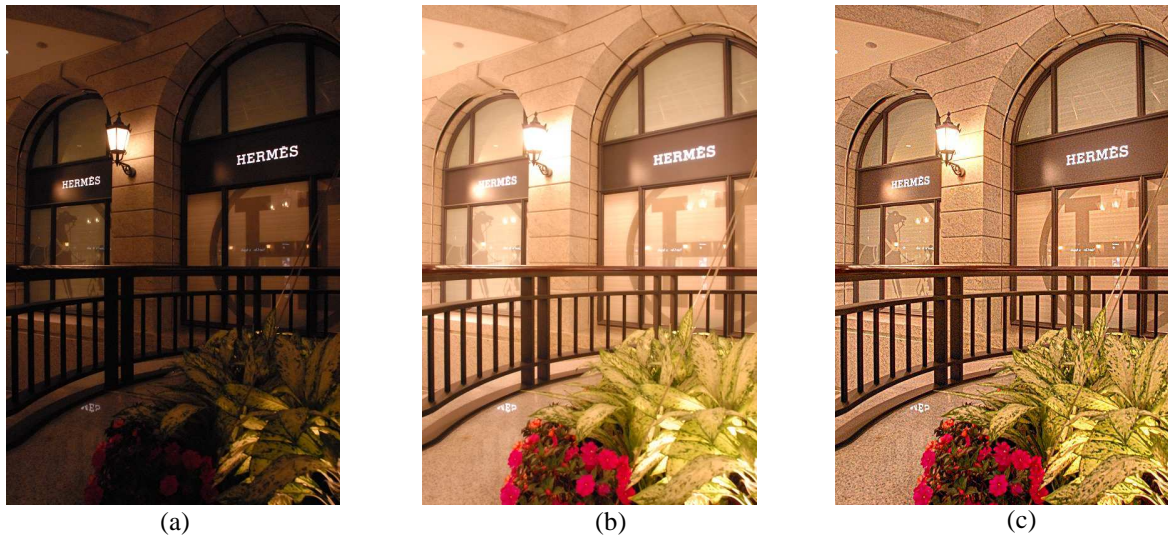


Figure 6.14: GFF's results for "Hermes" image sequence: (a, b) Different multi-exposure sequences (two input exposures), (c) Fused image by GFF.

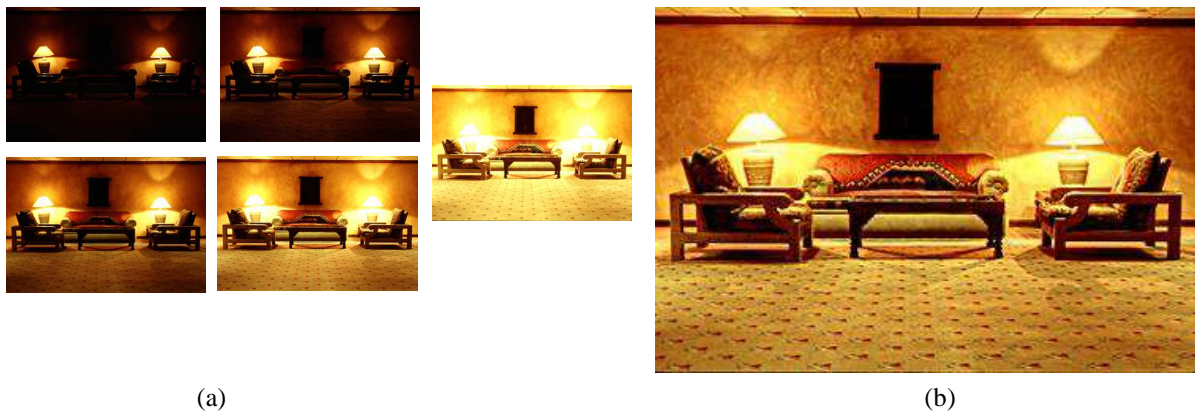


Figure 6.15: Results GFF for "Chairs" image series. (a) Different multi-exposure sequences (five exposures), (b) Fused image by GFF.



Figure 6.16: Results of GFF for “Syn” image series (a) Different multi-exposure sequences (seven input exposures), (b) Fused image by GFF.

As shown in Figures 6.13, 6.14, 6.15, and 6.16 the details from all of the input images are perfectly combined. In Figures 6.13(a-d), the results of GFF are compared to the recently proposed approaches. Figure 6.13(a) and Figure 6.13(b) shows the fusion results using the multi-resolution pyramid based approach. The result of Mertens et al. [113] (please see Figure 6.13(a)) appears blur and loses texture details while in GFF results (please see Figure 6.13(d)) the wall texture and painting on the window glass are emphasized that are difficult to visible in Figure 6.13(a). Clearly this is suboptimal as it removes pixel-to-pixel correlations by subtracting a low-pass filtered copy of the image from the image itself to generate Laplacian pyramid and result is a texture and edge details reduction in the fused image. Figure 6.13(b) shows the results using pyramid approach [89] which reveals many details but losses contrast and color information. GRW based exposure fusion is shown in Figure 6.13(c) which depicts less texture and color details in brightly illuminated regions (i.e. lamp and window glass). Note that Figure 6.13(d) retains colors, sharp edges and details while also maintaining an overall reduction in high frequency artifacts near strong edges.

Figures 6.14, 6.15, and 6.16 show the results of GFF for different image sequences captured at variable exposure settings (please see Figure 6.14(a): “Hermes” sequence, Figure 6.15(b): “Chairs” sequence, and Figure 6.16(c): “Syn” sequence. Note that, the strong edges and fine texture details are accurately preserved in the fused image without introducing halo artifacts. The halo artifacts will stand out if the DLs undergo a substantial boost.

6.2.2 IMPLEMENTATION AND COMPARISON OF VARIOUS CLASSIC EDGE-PRESERVING FILTERS

Figure 6.17(a) and 6.17(b) depicts the color-coded maps of under-exposed and over-exposed images respectively. Dark blue color indicates over-exposed region and pure white color indicates under-exposed region. Figures 6.17(c-f) illustrate the comparisons of color-coded maps of three edge preserving filters based detail enhancement and the results obtained by Mertens et al. [113] on the Cathedral sequence. It is accepted that the default parameter settings suggested by the different edge preserving filters [98,133]. Figure 6.17(e) shows the fusion results of multi-resolution pyramid based exposure fusion approach [89], which are both clearly close to the results obtained using GF (please see Figure 6.17(f)) but overall they yields less texture and edge details. The texture detail enhancement using BF [167] and WLS filter [133] shown in Figure 6.17(c) and Figure 6.17(d), respectively, depicts over-enhancement near strong edges and less color details. As shown in the close-up view in Figure 6.17(f), the GFF method based on GF can enhance the image texture details while preserving the strong edges without over enhancement.

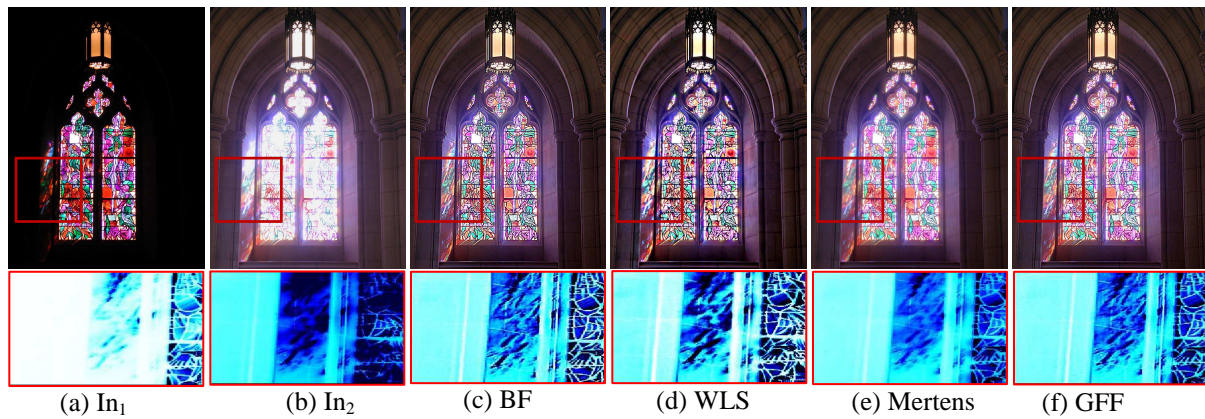


Figure 6.17: Color-coded map comparison (dark blue color indicates over-exposed region and pure white indicates under-exposed region): Comparison of GFF results to other classic edge preserving and exposure fusion techniques. (a) Source exposure 1; (b) Source exposure 2; (c) BF [167]; (d) WLS [133] (e) Mertens et al. [113]; (f) Results of GFF method. Note that GFF method yields enhanced texture and edge features. Input image sequence courtesy of Tom Mertens.

6.2.3 MULTI-FOCUS IMAGE FUSION

Moreover, in Figure 6.18, it is demonstrated that GFF method is also suitable for multi-focus image fusion to yield rich contrast. As illustrated in Figure 6.18(c), the edges and textures are relatively better than that of input images. Because GFF approach excludes fine

textures from the BLs for significantly preserving and enhancing fine details separately. However, utilizing the edge-preserving capability of GF, multi-resolution pyramid approach can be used for retaining strong edges and enhancing texture details in multi-focus image fusion problem.

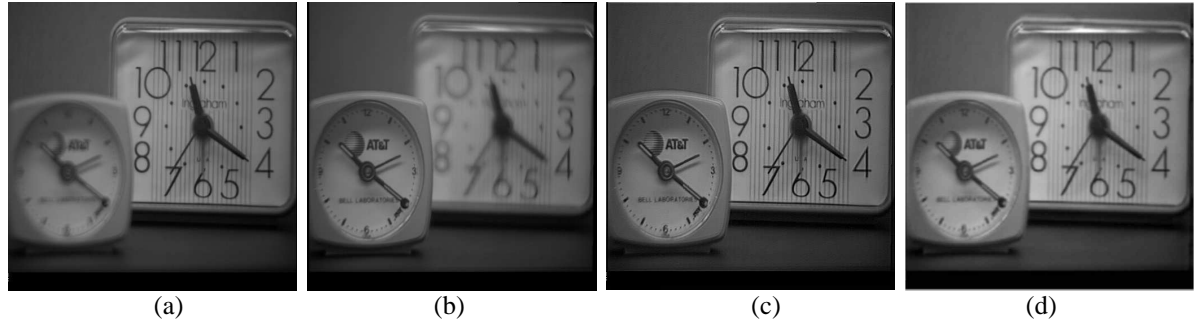


Figure 6.18: “Clock” image series. ((a), (b)) two partially focused input images (focused on different targets), (c) Image generated by the GFF approach, which illustrates that the fused image extracts more information from input images., and (d) Adu and Wang [110]. Input sequence is courtesy of Adu and Wang.

6.2.4 ANALYSIS OF FREE PARAMETERS AND FUSION PERFORMANCE METRICS

To analyze the affect of epsilon, gamma and window size used in GF on quality score ($Q^{AB/F}$) [186], entropy, and Visual Information Fidelity for Fusion (VIFF) [187], three plots (please see Figures 6.19(a-c) respectively) are illustrated for input image sequence of “Cathedral”. VIFF [187] first decomposes the source and fused images into blocks. Then, VIFF utilizes the models in VIF (Gaussian Scale Mixture (GSM) model, Distortion model and HVS model) to capture visual information from the two source-fused pairs. With the help of an effective visual information index, VIFF measures the effective visual information of the fusion in all blocks in each sub-band. Finally, the assessment result is calculated by integrating all the information in each sub-band. $Q^{AB/F}$ [186] evaluates the amount of edge information transferred from input images to the fused image. A Sobel operator is applied to yield the edge strength and orientation information for each pixel.

First, to analyze the affect of ε on $Q^{AB/F}$, Entropy and VIFF, the square window parameter (r) and texture amplification parameter (γ) were set to 2 and 5 respectively. As shown in Figure 6.19(a) the quality score and entropy decreases as ε increases, and VIFF increases as ε increases. It should be noticed in Figure 6.19(b) that the VIFF and entropy increases as r increases and $Q^{AB/F}$ decreases as r increases. It is preferred to have a small filter size (r) to reduce computational time. In the analysis of r , the other parameter are set to $\varepsilon=.01$

and $\gamma=5$. The visual inspection of affect of r 's on “Cathedral” sequence is depicted in Figure 6.20. It can easily be noticed (please see Figures 6.20(a-c) that as r increases, the strong edges and textures gets over-enhanced and therefore leads to artifacts. To analyze the influence of γ it should be noticed that entropy and $Q^{AB/F}$ decreases as γ increases and VIFF increases as γ increases. In order to obtain optimal detail enhancement and low computational time, it is concluded that the best results were obtained with $\varepsilon = 0.01$, $\gamma = 5$, and $r = 2$, which yield reasonably good results for all cases. Further, it has been analyzed that the selection of free parameters used in GF based approach will depend on the strength of textures and noise present in the input exposures.

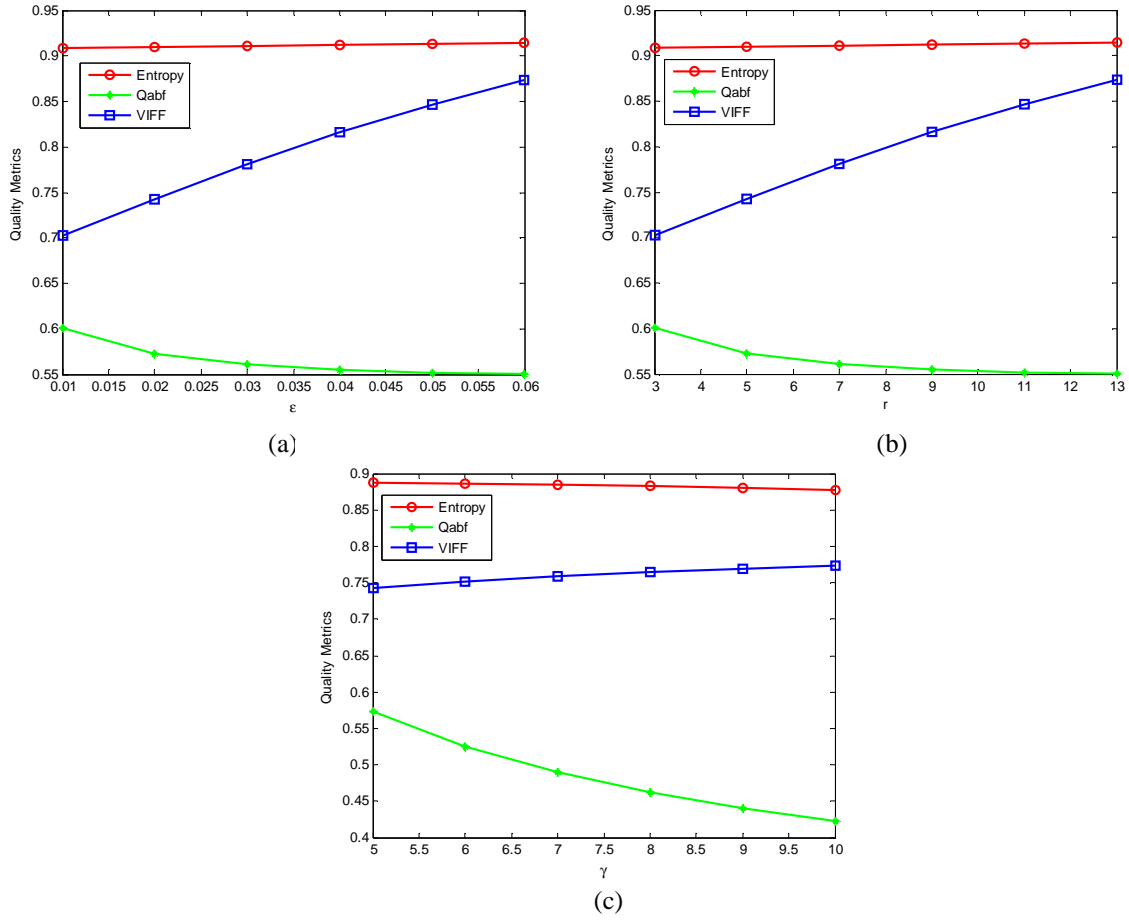


Figure 6.19: Analysis of different free parameters used in the GFF algorithm. Maximum quality score and entropy are only observed when $\varepsilon = 0.01$, $\gamma = 5$, and $r = 2$ (which are set as default parameters). It is observed that VIFF increases as ε , γ , and r increases but the larger values are responsible for over-enhancement. (a) Effectiveness of ε on metrics, (b) Effectiveness of r on metrics, (c) Effectiveness of γ on metrics.



Figure 6.20: Visual inspection: The affect of free parameter r on detail enhancement. It is found that $r=2$ is sufficient for fine details extraction and gives better results for most cases. Higher value of r brings in artifacts near strong edges. (a); $r=1$ (b) $r=3$; and (c) $r=6$.

6.3 WLS FILTER BASED EXPOSURE FUSION

In order to evaluate the performance and effectiveness of the WLF approach, the comparison of GFF approach with different exposure fusion, tone-mapping and multi-focus image fusion methods is summarized in this section. Two objective evaluation metrics (i.e. quality score $Q^{AB/F}$ [186] and VIFF [187]) are employed to evaluate the quality of fused images. Moreover these two evaluation metrics are used to analyze the affect of free parameters used in the present approach. In the WLF approach, all experimental results are generated by the MATLAB implementation. Furthermore, to measure distortion in the fused image and strengthen the evaluation capability of $Q^{AB/F}$ and VIFF, the Dynamic Range Independent Visible Difference Predictor (DRIVDP) [188] is incorporated in this section. DRIVDP metric is sensitive to three types of structural changes for distortions measurement (i.e., loss of visible contrast, amplification of invisible contrast, and reversal of visible contrast) between images under a specific viewing condition.

6.3.1 COMPARISON WITH OTHER EXPOSURE FUSION AND TONE-MAPPING METHODS

Figure 6.21, 6.22 and 6.23 depict the examples of fused images from the source multi-exposure images. It is noticed that the WLF approach enhances texture details while preventing halos near strong edges. In order to check the effectiveness and robustness of

WLF approach, the algorithm is tested on variety of multi-exposure image series. The WLF approach is computationally simple and the results are comparable to several exposure fusion and tone-mapping techniques. As shown in Figures 6.21(a-c) the details from all of the source images are perfectly combined and reveal fine textures while preserving local contrast and natural color. In Figures 6.22(a-d), the results of WLF are compared to the recently proposed approaches. Figure 6.22(d) depicts the results of optimization framework [118] and Figure 6.22(e) shows the mate-based fusion results using the edge preserving filter such as BF [111,167]. It can be observed that other fusion methods perform well in preserving image details while they fail to reconstruct texture and color details in the brightly illuminated areas. The result of Mertens et al. [113] (please see Figure 6.22(f)) appears blur and loses texture details while in WLF results (please see Figure 6.22(c)) the fine texture are emphasized that are difficult to visible in Fig. 6.22(f). This is because of utilization of Gaussian kernel for pyramid generation in [113]: as it removes pixel-to-pixel correlations by subtracting a low-pass filtered copy of the image from the image itself to generate Laplacian pyramid and result is a texture and edge details reduction in the fused image. The results produced in 6.22(d-f) lose visibility in a brightly-illuminated areas and details are lost in the tree leaves, and the texture on the wall is washed out. Although the results of Raman & Chaudhuri [111] (please see Figure 6.22(e)) exhibit better color details in tree leaves, but appear slightly blurry. In WLF results (Figure 6.22(c)) details are preserved in the brightly-illuminated areas, yet at the same time fine details are well preserved (tree leaves, wall texture, lizard).

To further compare the results of WLF visually with Mertens et al. [113], iCAM06 [189], WLS [133] and GRW [116], respectively, Figures 6.23(a-e) depict experimental results for “National Cathedral” sequence (1024×768×3). The fusion results of WLF shown in Figure 6.23(a) illustrate the ability of enhancing fine texture details. As well as having the ability to produce good color information with natural contrast. This can bring an increased illusion of depth to an image textures. Therefore, enhanced texture details in the fused image let you get everything sharp and yield an accurate exposure that is entirely free of halo artifacts. Although tone-mapped results of iCAM06 [189] and WLS [133] have produced comparable results, but they do not preserve contrast from input LDR image series. Figure 6.23(b) and Figure 6.23(e) shows the results of pyramid approach [113] and GRW optimization framework [116] respectively, which

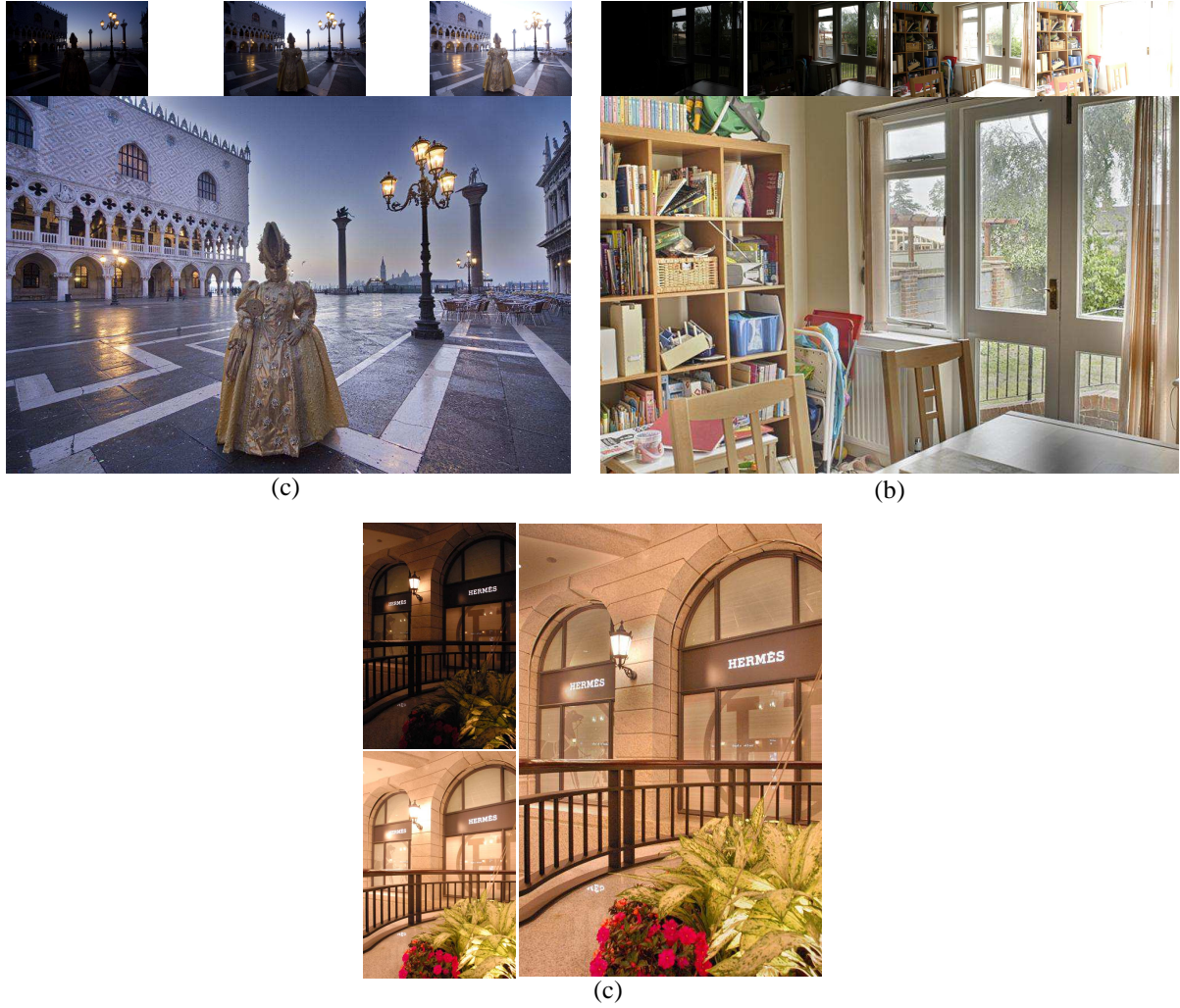


Figure 6.21: Results of WLF approach for different input multi-exposure sequences. (a) “Carnival” (top: three input exposures, bottom: fusion results), (b) “House” (top: four input exposures, bottom: fusion results) and (c) “Bellavita” (left: two input exposures, bottom: fusion results). Input image sequences are courtesy of HDRsoft.com, Tom Mertens.

preserve global contrast but losses color information. GRW [116] based exposure fusion is shown in Figure 6.23(e) which depicts less texture and color details in brightly illuminated regions (i.e. lamp and window glass). Note that Figure 6.23(a) retains colors, sharp edges and details while also maintaining an overall reduction in high frequency artifacts near strong edges. The results produced in Figure 6.23(b-d) were generated by the programs provided by their respective authors. The HDR images for iCAM06 [189] and WLS [133] were generated using HDR reconstruction [17]. The results of GRW [116] shown in Figure 6.23(e) are taken from its paper. In order to give a relatively fair comparison of WLF results, the default sets of parameters for tone-mapping [189, 133] and exposure fusion [113] methods are used.

Figure 6.24 shows the distortion maps computed from DRIVDP metric proposed by Aydın et al [188]. This quality assessment metric detects loss of visible contrast (green) and

amplification of invisible contrast (blue). The main advantage of this metric is that it yields meaningful results even if the input images have different DRs. Though the DRIVDP based quality assessment is used to compare WLF method with one exposure fusion [113] method and two tone-mapping methods [133,189]. It is assumed that the LDR images are shown in a typical LCD display with maximum luminance 100 and gamma 2.2. It is also assumed that for all LDR images, the viewing distance is 0.5 meters and the number of pixels per visual degree is 30 and peak contrast is 0.0025. Significance of the choice of these

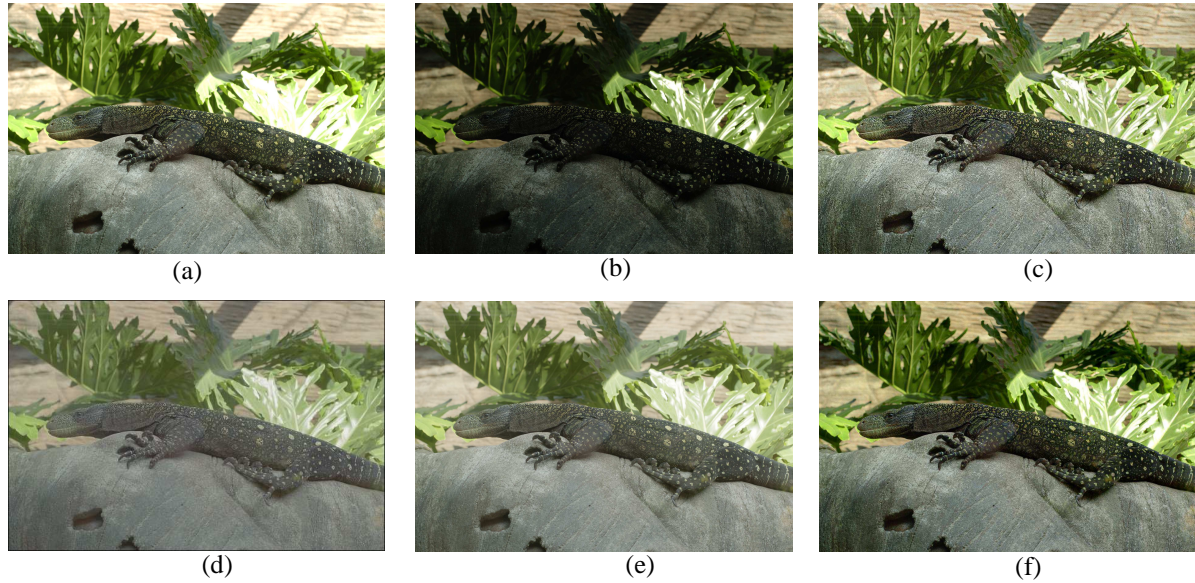


Figure 6.22: “Lizard” image series: Comparison of WLF results with popular exposure fusion approaches. (a, b) Source images, (c) Results of WLF method: tree leaves and wall texture appear over-exposed and blurry in (d) Ketan & Chaudhuri [118], (e) Raman & Chaudhuri [111] and (f) Mertens et al. [113]. In the proposed results, it becomes almost possible to preserve texture and strong edge features simultaneously. Moreover, by applying the WLF approach, the fine textures are accurately enhanced. Input image sequence is Eric Reinhard, University of Bristol.



Figure 6.23: “National Cathedral” image series: Comparison of WLF results with popular exposure fusion and tone mapping methods. (a) Results of WLF method, (b) Mertens et al. [113], (c) iCAM06 [189], (d) WLS [133] and (e) GRW [116]. Note that WLF method yields enhanced texture and edge features with better color appearance. Input image sequence is courtesy of Max Lyons.

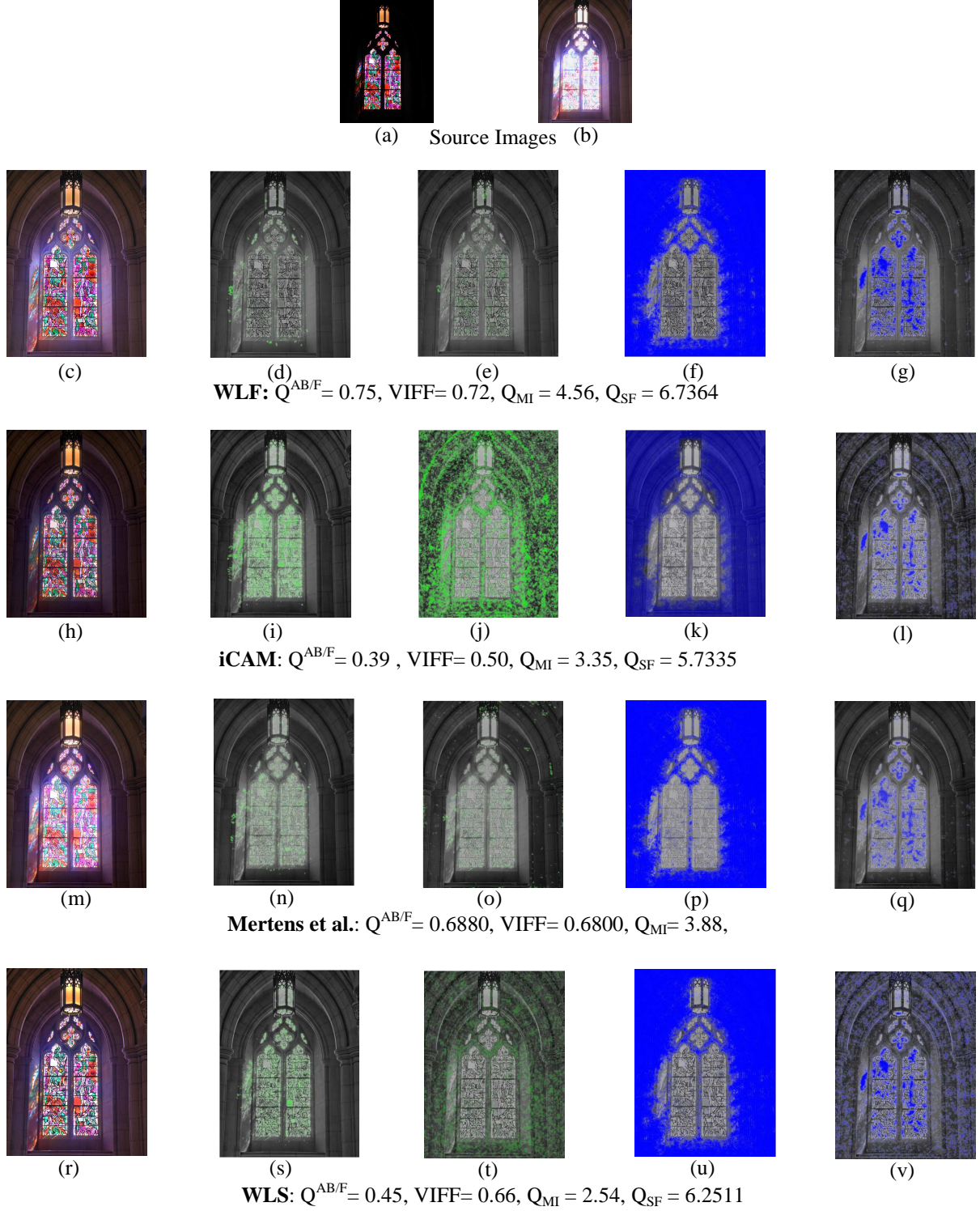


Figure 6.24: Comparison of WLF results with iCAM06, Mertens et al. and WLS on the National Cathedral sequence using DRIVDP [188]. (a), (b) The two source images give good exposures for the paintings on window glass and lamp, and the wall, respectively. In a distortion map, green, blue, and gray pixels indicate visible contrast loss, amplification, and no distortion, respectively. WLF results are more effective in preserving local details and colors than the others. For the paintings on window glass and lamp, and the wall, proposed results (see c-g) depict the least distortion, followed by iCAM06 (see h-l), Mertens et al. (see m-q), and WLS (see r-v).

parameters can be found in [188]. Figures 6.24(a-v) show a side-by-side comparison of the loss of visible contrast (green), and amplification of invisible contrast (blue) of WLF results with others methods. To compute visible contrast loss illustrated in Figure 6.24(d), 6.24(i), 6.24(n), and 6.24(s), respectively, for fused images in Figures 6.24(c), 6.24(h), 6.24(m), and 5(r), the under-exposed image (i.e. Figure 6.24(a)) is used as reference image. Similarly, to compute visible contrast loss illustrated in Figure 6.24(e), 6.24(j), 6.24(o), and 6.24(t), respectively, for fused images in Figure 6.24(c), 6.24(h), 6.24(m), and 6.24(r), the over-exposed image (i.e. Figure 6.24(b)) is used as reference image. The invisible amplification metrics is run on fused images, which are generated using a similar procedure as used for loss of visible contrast metric. The two source images with good exposures respectively for the brightly illuminated region (i.e. window) and the poorly illuminated region (i.e. wall) are given in Figure 6.24(a) and 6.24(b). The distortion maps for WLF method, iCAM06 [189], Mertens et al. [113] and WLS [133] are given in Figure 6.24(c-g), Figure 6.24(h-l), Figure 6.24(m-q), and Figure 6.24(r-v), respectively, along with the fused images. In a distortion map, green, blue, and gray pixels indicate contrast loss, amplification, and no distortion, respectively. It can be noticed that the results of WLF are more effective in preserving local contrast and color information than the other methods. Please note that visible contrast loss and distortions are the least using the WLF approach. Moreover, to compare the performance of the WLF approach, iCam06, Mertens et al. and WLS, the four fusion quality metrics i.e. $Q^{AB/F}$, VIFF, Mutual Information (MI) [190], and Spatial Frequency (SF) [191] are employed.

$Q^{AB/F}$ [186] evaluates the amount of edge information transferred from input images to the fused image. A Sobel operator is applied to yield the edge strength and orientation information for each pixel. For two input images A and B , and a resulting fused image F , the Sobel edge operator is applied to yield the edge strength $e(m,n)$ and orientation $\beta(m,n)$ information for each pixel as:

$$e_A(m,n) = \sqrt{sb_A^x(m,n)^2 + sb_A^y(m,n)^2} \quad (6.1)$$

$$\beta_A(m,n) = \tan^{-1} \left[\frac{sb_A^x(m,n)}{sb_A^y(m,n)} \right] \quad (6.2)$$

Where $sb_A^x(m,n)$ and $sb_A^y(m,n)$ are horizontal and vertical Sobel template cantered on pixel $P_A(m,n)$ and convolved with the corresponding pixels of image A . The relative strength and orientation values of an input image A with respect to F are formed as:

$$(G_{m,n}^{AF}, A_{m,n}^{AF}) = \left(\left(\frac{e_{m,n}^F}{e_{m,n}^A} \right)^\psi, 1 - \frac{|\beta_A(m,n) - \beta_F(m,n)|}{\pi/2} \right) \quad (6.3)$$

$$\text{where } \Psi = \begin{cases} 1 & \text{if } e_A(m, n) > e_F(m, n) \\ -1 & \text{otherwise} \end{cases} \quad (6.4)$$

The process of edge information preservation values is defined in [186]. Finally, the $Q^{AB/F}$ is defined as:

$$Q^{AB/F} = \frac{\sum_{\forall m,n} Q_{m,n}^{AF} w_{m,n}^A + Q_{m,n}^{BF} w_{m,n}^B}{\sum_{\forall m,n} w_{m,n}^A + w_{m,n}^B} \quad (6.5)$$

which evaluates the sum of edge information preservation values for both inputs Q^{AF} and Q^{BF} weighted by local importance perceptual factors w^A and w^B . We defined $w^A(m, n) = [e_A(m, n)]^L$ and $w^B(m, n) = [e_B(m, n)]^L$. Where L is a constant. For the ‘‘ideal fusion’’, $Q^{AB/F} = 1$.

VIFF [187] first decomposes the source and fused images into blocks. Then, VIFF utilizes the models in VIF (i.e. GSM model, Distortion model and HVS model) to capture visual information from the two source-fused pairs. With the help of an effective visual information index, VIFF measures the effective visual information of the fusion in all blocks in each sub-band. Finally, the assessment result is calculated by integrating all the information in each sub-band:

$$VIFF(I_1, \dots, I_n, I_F) = \sum_k P'_k \cdot VIFF_k(I_1, \dots, I_n, I_F) \quad (6.6)$$

where P'_k is a weighting coefficient. According to VIF theory, a high VIF yields a high quality test image. Therefore, as VIFF increases, the quality of the fused image improves.

The quality metric Q_{MI} measures how well the original information from source images is preserved in the fused image.

$$Q_{MI} = 2 \left(\frac{MI(A, F)}{H(A) + H(F)} + \frac{MI(B, F)}{H(B) + H(F)} \right) \quad (6.7)$$

where $H(A)$, $H(B)$ and $H(F)$ are the marginal entropy of A , B and F , and $MI(A, F)$ is the mutual information between the source image A and the fused image F .

$$MI(A, F) = H(A) + H(F) - H(A, F) \quad (6.8)$$

where $H(A, F)$ is the joint entropy between A and F , $H(A)$ and $H(F)$ are the marginal entropy of A and F , respectively, and $MI(B, F)$ is similar to $MI(A, F)$.

The fourth criterion is Q_{SF} . The spatial frequency, which originated from the human visual system, indicates the overall active level in an image and has led to an effective objective quality index for image fusion [191]. The total spatial frequency of the fused image is computed from row (RF) and column (CF) frequencies of the image block and Q_{SF} is defined as:

$$Q_{SF} = \sqrt{RF^2 + CF^2} \quad (6.9)$$

$$RF = \sqrt{\frac{1}{M'N'} \sum_{m=1}^{M'} \sum_{n=1}^{N'} (I_F(m, n) - I_F(m, n-1))^2} \quad (6.10)$$

where $I_F(m, n)$ is the gray value of pixel at position (m, n) of image I_F .

$$CF = \sqrt{\frac{1}{M'N'} \sum_{m=1}^{M'} \sum_{n=1}^{N'} (I_F(m, n) - I_F(m-1, n))^2} \quad (6.11)$$

The quantitative performance analysis using the aforesaid evaluation indices are shown in the caption of Figure 6.24. The WLF approach has outperformed the other methods. It can be seen that the WLF method can preserve more useful information compared with iCam06, Mertens et al. and WLS methods. In particular, evaluation results in Figure 6.24 have demonstrated that $Q^{AB/F}$, VIFF, Q_{MI} and Q_{SF} have correspondence with the DRIVDP-based evaluation.

6.3.2 MULTI-FOCUS IMAGE FUSION

Furthermore, to check the applicability of WLF approach for other image fusion applications, the experimental results for multi-focus image fusion are presented in this chapter. In Figure 6.25, Figure 6.26, Figure 6.27 and Figure 6.28, it is demonstrated that WLF method is also suitable for multi-focus image fusion to yield rich contrast and texture details. One of the key characteristics of WLF approach for multi-focus image fusion application is illustrated in Figure 6.25(c): the color details are preserved in the fused image with better visualization of texture details. It can also be noticed in Figure 6.25(f) that the edges and textures are relatively better than that of source images. Fusion results of WLF method on the four standard test scenes (please see Figures 6.26(a, b, d, e) and Figure 6.27(a, b, d, e)) are shown in Figures 6.26(c, f) and Figure 6.27(c, f), respectively. Note that, the strong edges and fine texture details are accurately preserved in the fused image without introducing halo artifacts. The halo artifacts may stand out if the DL undergoes a substantial boost. Comparisons of Adu and Wang [110], DWT [192], Tian et al. [193], and WLF approach for multi-focus image fusion are illustrated in Figures 6.28(a-d). The result produced in Figure 6.28 (b) is taken from its paper [110]. Results of DWT [192] shown in Figure 6.28(c) were generated by the MATLAB Wavelet toolbox. For the DWT-based

methods, the low-pass sub-band coefficients and the high-pass sub-band coefficients are simply merged by the averaging scheme and the choose-max selection scheme, respectively. The DWT-based fusion algorithm is performed using five-level decomposition and db3 wavelets are used in scale decomposition. The results of Tian et al. [193] shown in Figure 6.28 (d) are generated from the MATLAB code provided by the author. Note that WLF method (please see Figure 6.28 (a)) yields enhanced texture and edge features. The approach proposed in Chapter 5 can significantly preserve and enhance fine details separately because WLF approach excludes fine textures from the BLs.

However, with the WLF approach, the DOF in fused image is improved so that the foreground and background of the image is in focus. Having done so, we will then begin to make considered background and foreground simultaneously in the single image that helps to capture sharper details in various security applications.

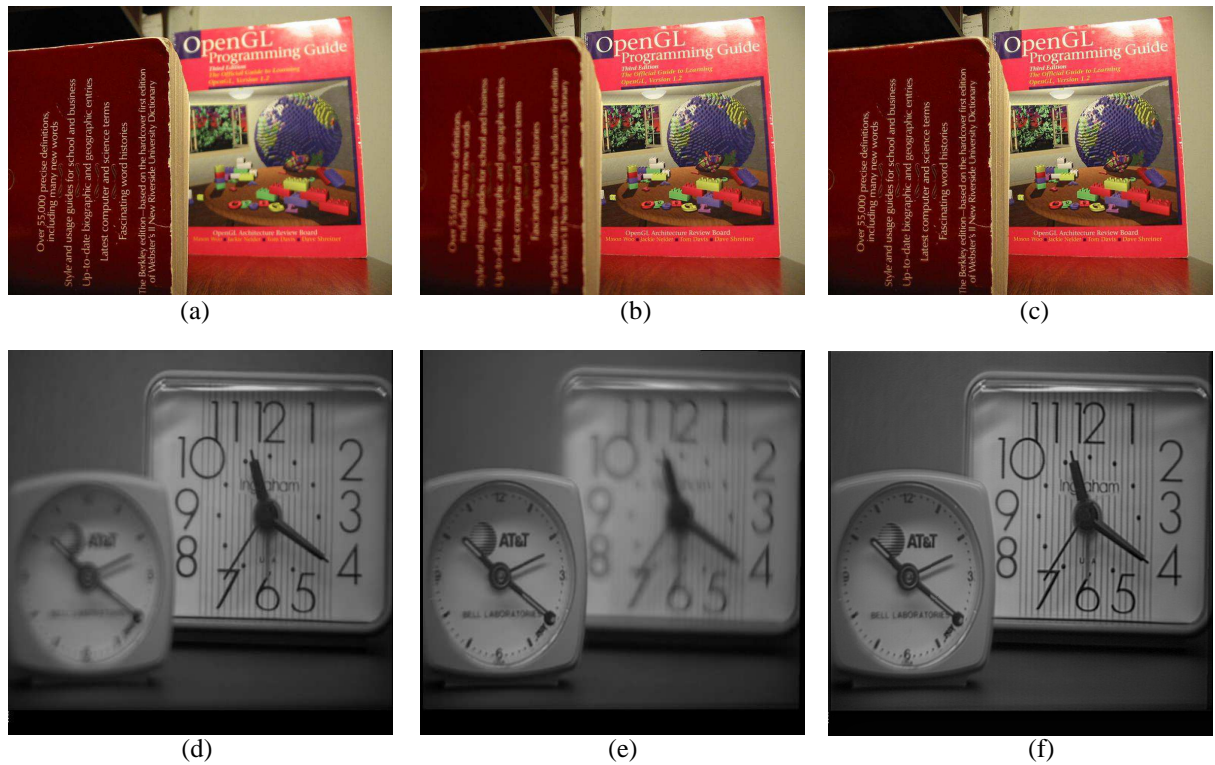


Figure 6.25: Multi-focus image fusion results: (a, b, d, e) demonstrates the effect of selective focus to capture more details from a particular part of a scene and (c, f) images generated by the WLF approach. The fused images appear all-in-focus, and the present method enhances the color and texture details present in the foreground and background. (Input sequence is courtesy of Adu and Wang [110]).

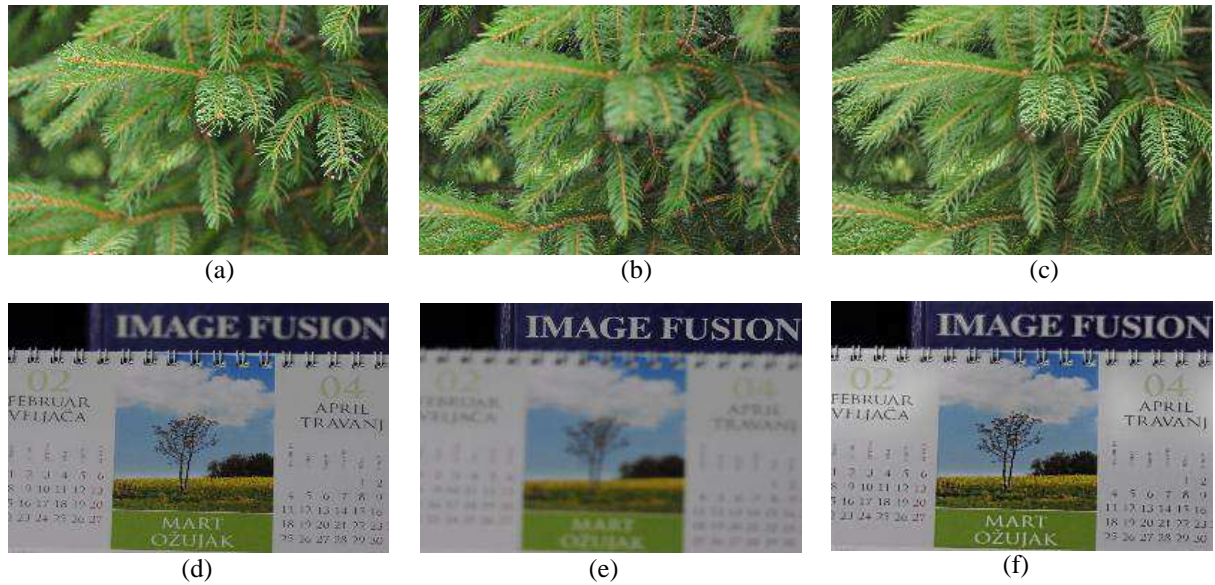


Figure 6.26: Multi-focus image fusion results: (a, b, d, e) demonstrates the effect of selective focus to capture more details from a particular part of a scene and (c, f) images generated by the WLF approach. These results demonstrate that WLF approach has helped to handle typical situations of foreground and background present in the scene (input sequence is courtesy Slavica Savić).

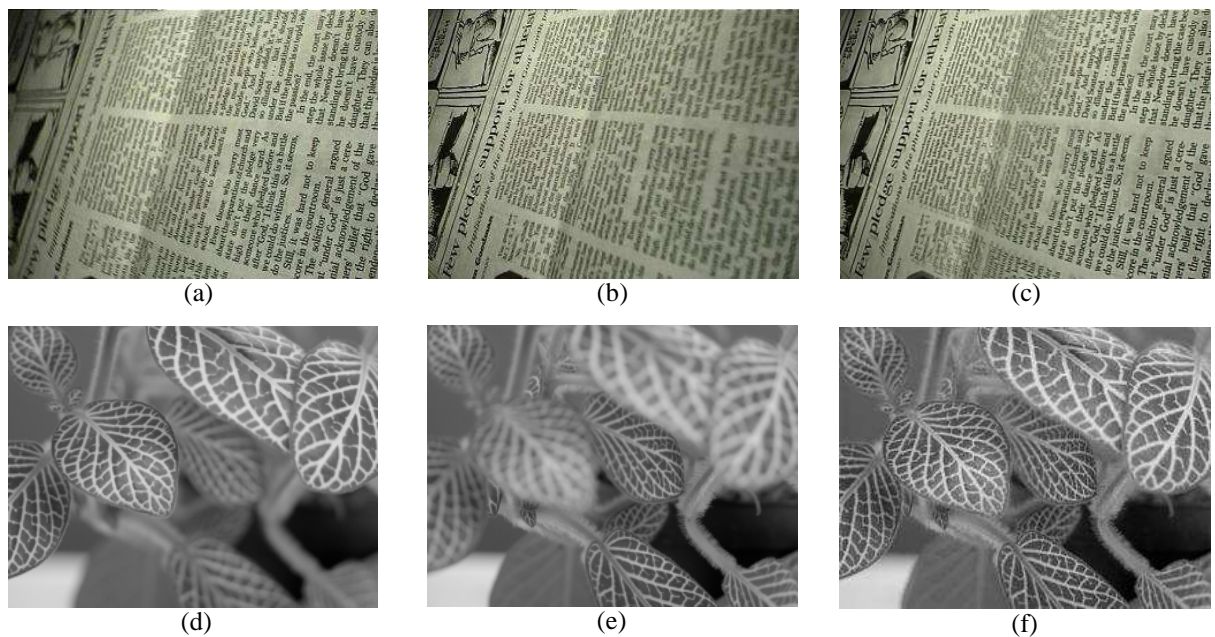


Figure 6.27: Multi-focus image fusion results for "Documents" and "Leaves" image series: (a, b, d, e) demonstrates the effect of selective focus to capture more details from a particular part of a scene and (c, f) images generated by the WLF. These results demonstrate that WLF approach has helped to handle typical situations of foreground and background present in the scene (input sequence is courtesy Slavica Savić).

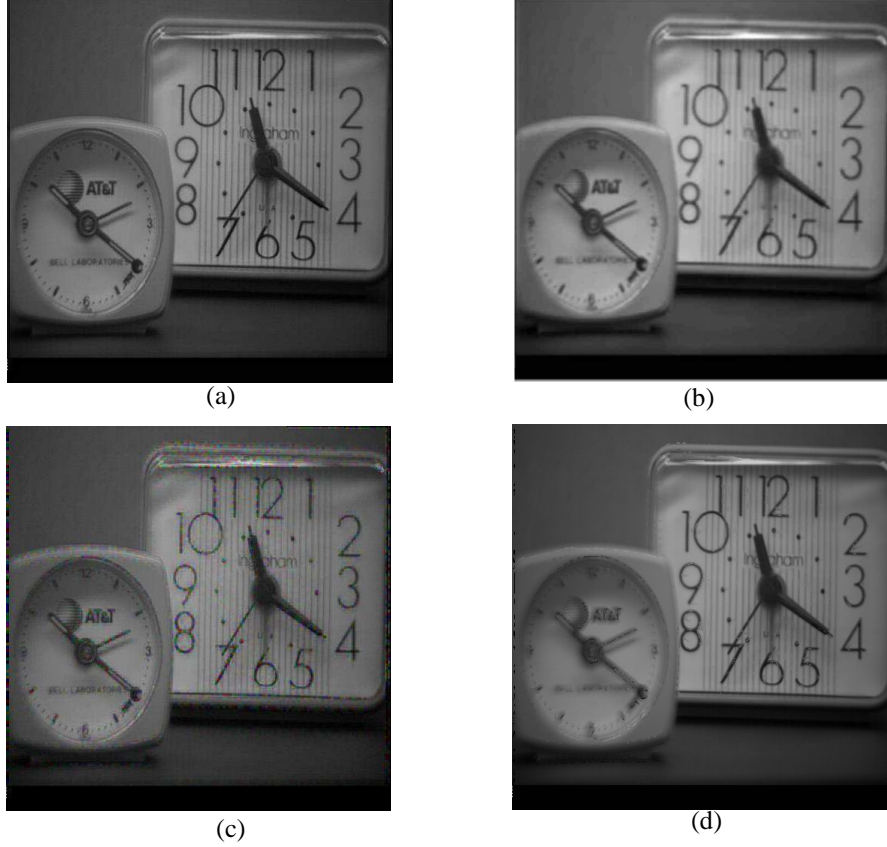


Figure 6.28: Comparison of “Clock” image results with recently proposed conventional multi-focus image fusion methods. (a) WLF results, (b) Adu and Wang [110], (c) DWT [192] and (d) Tian et al. [193]. It has been found that the WLF approach helps the viewer to observe enhanced texture and edge features simultaneously without depicting visible artifacts. Input image sequence is courtesy of Adu and Wang.

6.3.3 ANALYSIS OF FREE PARAMETERS, AND FUSION PERFORMANCE METRICS

WLF method has eight free parameters i.e. t_a , K , λ , \hat{a} , θ , ψ , α_4 , ε' . It fixes $t_a=5$, $K=30$, $\lambda=1/7$, $\hat{a}=3$, $\theta=0.001$, $\psi=0.1$, $\alpha_4=1.2$, and $\varepsilon'=0.0001$ in all experiments and they are set as default parameters. It is preferred to have a small number of iterations (t_a) to reduce computational time. The fusion performance does not affected when $t_a \leq 5$ because present method does not depend much on the exact parameter choice of t_a . The parameters selection criterion for K , and λ is given in [156]. WLF has set $K=30$, $\lambda=1/7$ and $\theta=0.001$ as default parameters for all experiments. In the WLF approach, the fusion performance is dependent on two free parameters i.e. ψ and \hat{a} . To analyze the affect of parameter (ψ) and free parameter (\hat{a}) on $Q^{AB/F}$ [186], and VIFF [187], four plots (please see Figures 6.29(a-d)) are illustrated

for input image sequence of “Cathedral” (1024×768×3), “Bellavita” (800×535×3), and “Book” (569×758×3).

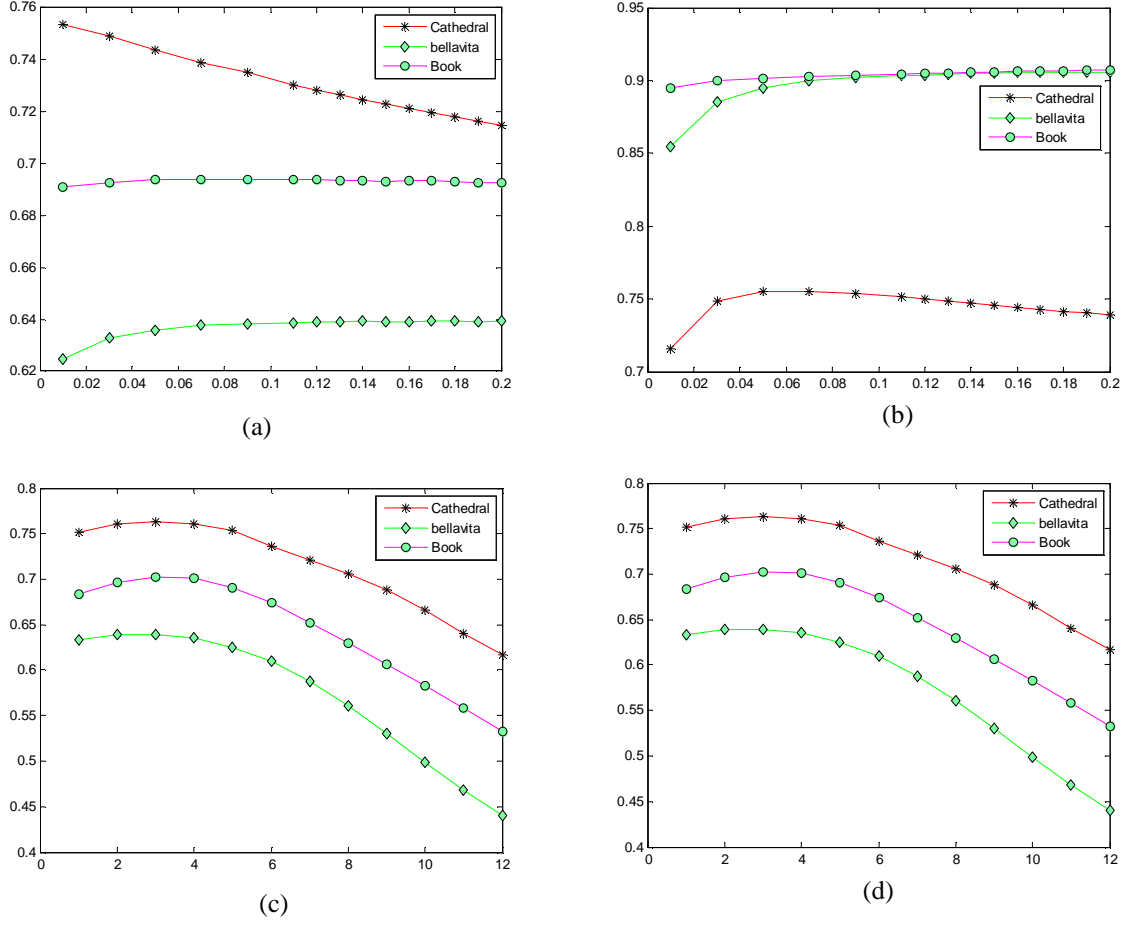


Figure 6.29: Sensitivity analysis of free parameter (ψ) and free parameter (\hat{a}) which respectively control the smoothness in weight maps computed for BLs and DLs in WLF approach. (a) Effectiveness of ψ on $Q^{AB/F}$, (b) Effectiveness of ψ on VIFF, (c) Effectiveness of \hat{a} on $Q^{AB/F}$ and (d) Effectiveness of \hat{a} on VIFF.

The detailed description of $Q^{AB/F}$ and VIFF is given in the previous Sub-section. To assess the affect of parameter (ψ) and free parameter (\hat{a}) on fusion performance, the $Q^{AB/F}$ and VIFF are experimented.

To analyze the influence of ψ and \hat{a} on $Q^{AB/F}$ and VIFF, other parameters are set to $t_a=5$, $K=30$, $\lambda=1/7$, $\theta=0.001$, $\alpha_4=2$, and $\varepsilon'=0.0001$. As shown in Figures 6.29(a) and 6.29(b), the fusion performance will be worse when the values of ψ and \hat{a} are too large or too small. It should be noticed in Figure 6.29 that the $Q^{AB/F}$ and VIFF decreases when the ψ and \hat{a} are too large or too small. The visual inspection of affect of \hat{a} 's on “Cathedral” sequence is depicted in Figures 6.30(a-c).



Figure 6.30: Visual inspection of fine detail enhancement in WLF approach in typical lighting situations: The free parameter \hat{a} in Equation (5.38) controls detail enhancement and sharpening. It is found that variation of “ \hat{a} ” between 3 and 6 is sufficient for fine details extraction and give better results for various typical situations. By choosing higher value of \hat{a} , the texture details are accurately enhanced. It also shows that selection of much higher value for detail enhancement does not introduce artifacts near object boundaries due to detail-enhancement. (a) $\hat{a} = 1$, (b) $\hat{a} = 6$, (c) $\hat{a} = 12$.

It can easily be noticed in Figures 6.30(a-c) that as \hat{a} increases, the strong edges and textures get enhanced and therefore leads to a detail preserving fusion results. In order to obtain optimal detail enhancement and low computational time, it is concluded that the best results were obtained with $t_a=5$, $K = 30$, $\lambda = 1/7$, $\hat{a} = 3$, $\theta = 0.001$, $\psi = 0.1$, $\alpha_4 = 1.2$, and $\varepsilon' = 0.0001$, which yield reasonably good results and satisfactory subjective performance for all cases.

To further demonstrate the analysis of errors introduced by the free parameter \hat{a} , four fundamental error performance metrics are adopted i.e. Root Mean Squared Error (RMSE), Normalized Absolute Error (NAE), Laplacian Mean Squared Error (LMSE), and Peak Signal to Noise Ratio (PSNR). The RMSE measure the differences between resulting image and reference image. The error in a pixel $E_e(P_i^{res})$ is calculated using Euclidean distance between a pixel in a resulting image (P_i^{res}) with $\hat{a} > 1$ and the corresponding pixel in the reference image (P_i^{ref}) with $\hat{a}=1$, i.e., $E_e(P_i^{res}) = \|P_i^{res} - P_i^{ref}\|$. The total error in a resulting fused image is computed using square root of the Mean Square Error (MSE), i.e., $\sqrt{1/M \sum_i E_e(P_i^{res})}$.

NAE is a measure of how far is the resulting fused image (when $\hat{a} > 1$) from the reference fused image (when $\hat{a}=1$) with the value of zero being the perfect fit. Large value of NAE indicates poor quality of the image [191]. NAE is computed as follows:

$$NAE = \frac{\sum_{m=1}^{M'} \sum_{n=1}^{N'} |I_F(m, n) - \hat{I}_F(m, n)|}{\sum_{m=1}^{M'} \sum_{n=1}^{N'} |I_F(m, n)|}$$

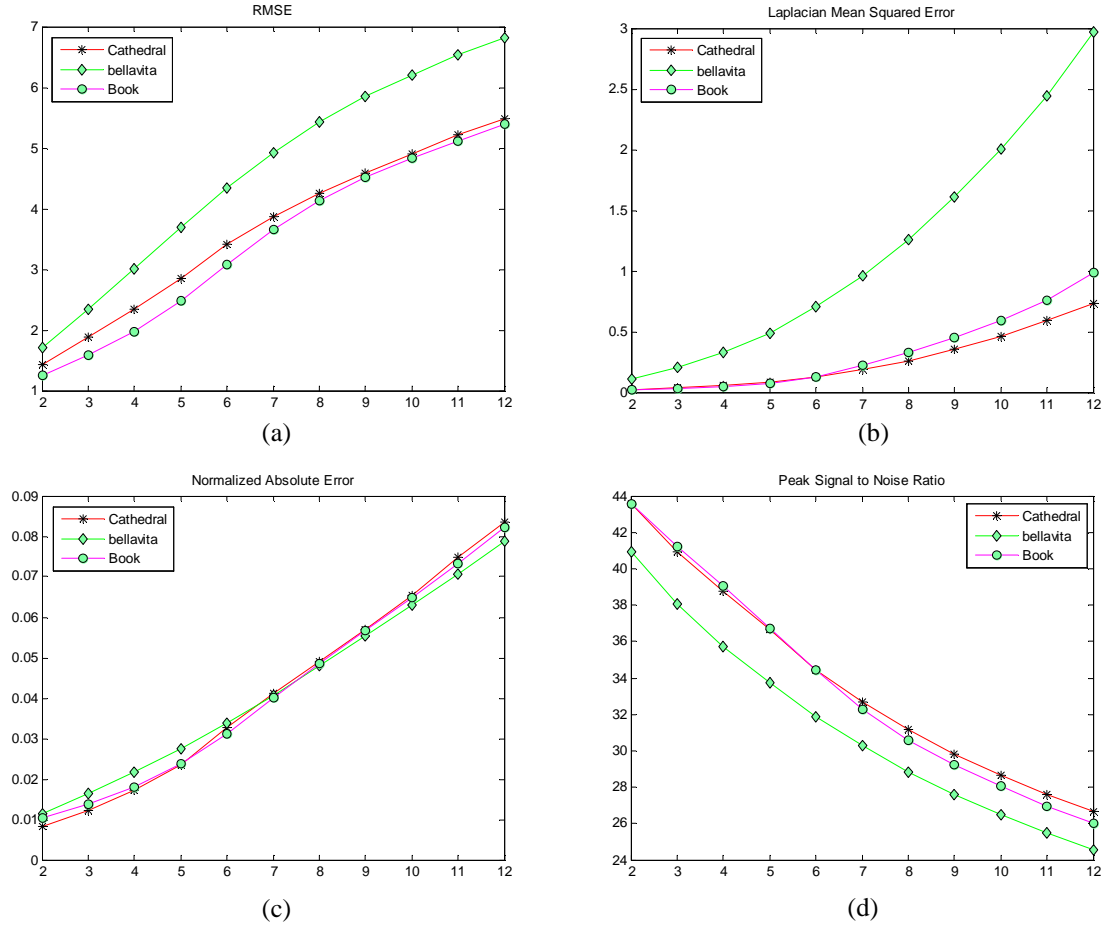


Figure 6.31: Sensitivity analysis of errors introduced by the free parameter \hat{a} in WLF approach. The errors increases dramatically as \hat{a} becomes too large, but increases slowly when $\hat{a} \leq 6$. It is observed in Figure (b) and Figure (c) that error increases as \hat{a} increases but still the deviation range is less. (a) Effectiveness of \hat{a} on RMSE, (b) Effectiveness of \hat{a} on LMSE, (c) Effectiveness of \hat{a} on NAE, and (d) Effectiveness of \hat{a} on PSNR.

LMSE is based on the importance of edges detail measurement, which is also the most critical feature for image quality assessment. The large value of LMSE means that image is poor in quality. LMSE is defined as follow:

$$LMSE = \frac{\sum_{m=1}^{M'} \sum_{n=1}^{N'} [L(I_F(m, n)) - L(\hat{I}_F(m, n))]^2}{\sum_{m=1}^{M'} \sum_{n=1}^{N'} [L(I_F(m, n))]^2}$$

where $L(I_F(m, n))$ is a Laplacian operator:

$$L(I_F(m, n)) = I_F(m+1, n) + I_F(m-1, n) + I_F(m, n+1) + I_F(m, n-1) - 4I_F(m, n) \quad (6.12)$$

PSNR is the ratio between the maximum possible pixel value of the fused image MAX_{I_F} and the MSE, which is computed as follows:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_{I_F}^2}{MSE} \right) \quad (6.13)$$

As shown in Figures 6.31(a-c) the RMSE, LMSE, and NAE increases as parameter a increases. It should be noticed in Figure 6.31(a) that when $\hat{a} = 12$, the total error introduced is still less than 7%. It is seen from the computed values of objective measures like NAE and LMSE (please see Figures 6.31(b-c)) for $\hat{a} = 2, 3, 4, 5, 6, 7, 8, 9, 10, 11$, and 12, for input image sequence of “Cathedral”, “Bellavita”, and “Book”, the errors increases dramatically as free parameter \hat{a} in Equation (5.38) becomes too large, but increases slowly when $\hat{a} \leq 6$. Using graph presented in Figure 6.31 (d), we want to illustrate what can happen if PSNR is used as distortion measure. It has been found that PSNR decreases gradually as free parameter \hat{a} increases and can be seen that the WLF approach is performing consistently for different value of free parameter (i.e. \hat{a}) proposed for detail enhancement.

6.3.4 INTERACTIVE TOOL FOR DETAIL ENHANCEMENT

Another interesting interactive tool of the WLF approach for manipulating the detail and contrast in the multi-focus image fusion has been experimented. Figures 6.32(a-f) show the results generated for “Clock” and “Leaves” image series by varying the free parameters ψ in Equation (5.32), and \hat{a} in Equation (5.38) (please see Figure 6.32(a) for $\psi = 0.5$, $\hat{a} = 7$), (please see Figure 6.32(b) for $\psi = 0.5$, $\hat{a} = 12$), (please see Figure 6.32(c) for $\psi = 0.5$, $\hat{a} = 17$), and (please see Figure 6.32(d) for $\psi = 0.5$, $\hat{a} = 7$), (please see Figure 6.32(e) for $\psi = 0.5$, $\hat{a} = 12$), (please see Figure 6.32(f) for $\psi = 0.5$, $\hat{a} = 17$), respectively. Here, it is demonstrated that highly detail enhanced fused image is generated from multi-focus image series, before objectionable artifacts appear. It is found that WLF approach is very effective for boosting the amount of local contrast and fine details. The effective manipulation range is very wide and will vary in accordance with the texture details present in the input image series: it typically takes a rather extreme manipulation to cause artifacts near strong edges to appear.

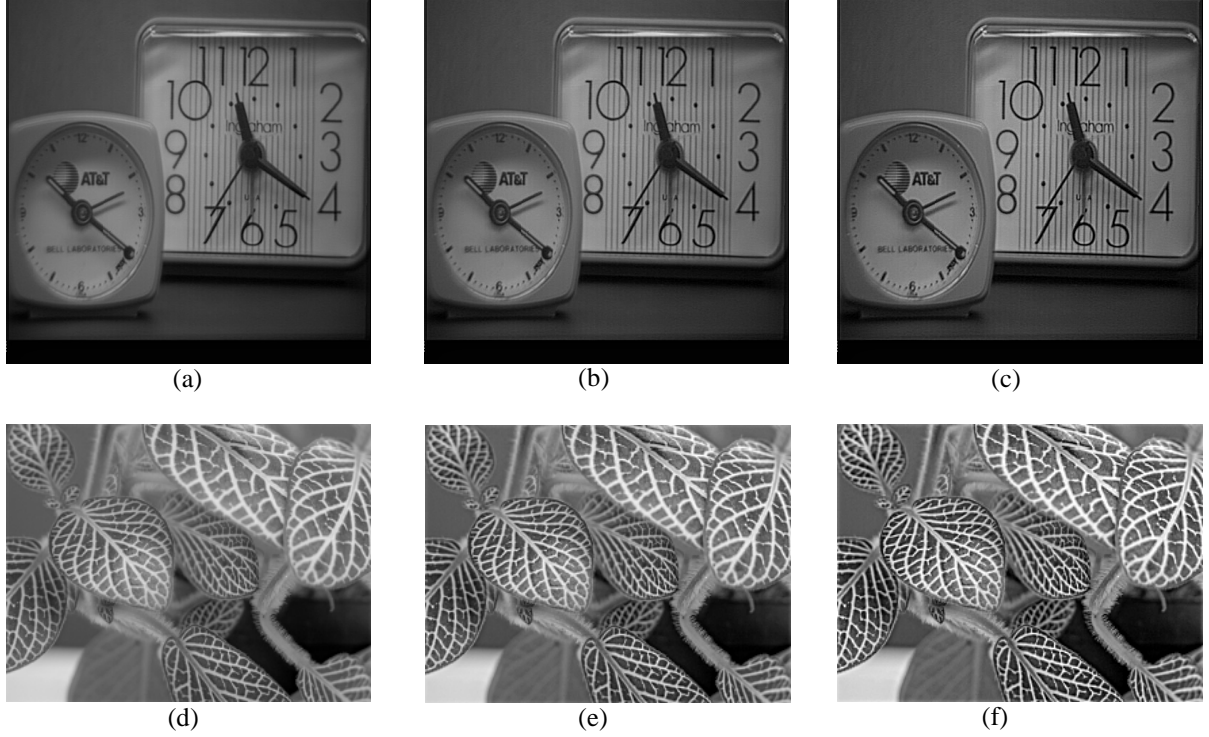


Figure 6.32: Visual inspection of contrast and detail enhancement in WLF approach in multi-focus image fusion. The free parameter ψ in Equation (5.32) and \hat{a} in Equation (5.38) controls detail enhancement and sharpening respectively. WLF is able to achieve aggressive detail enhancement and exaggeration, while avoiding artifacts. “Clock” image: (a) $\psi = 0.5$, $\hat{a} = 7$, (b) $\psi = 0.5$, $\hat{a} = 12$, (c) $\psi = 0.5$, $\hat{a} = 17$. “Leaves” image: (d) $\psi = 0.5$, $\hat{a} = 7$, (e) $\psi = 0.5$, $\hat{a} = 12$, (f) $\psi = 0.5$, $\hat{a} = 17$.

6.4 OBJECTIVE FUSION QUALITY EVALUATION

In this section, the objective performances of different methods are evaluated. To measure the performance of the proposed methods, three objective image fusion quality metrics are employed. They evaluate the fusion performance from different aspects such as the amount of edge information transferred from input images to the fused image $Q^{AB/F}$ [186], how well the original information from source images is preserved in the fused image Q_{MI} [191] and the spatial frequency which indicates the overall active level in an image Q_{SF} [190]. Table I and Table II give the values of several indicators generated by different methods for multi-exposure and multi-focus image fusion, respectively. Table I gives the values of $Q^{AB/F}$, Q_{MI} , and Q_{SF} for the fused results by Kaung et al. [189], Mertens et al. [113], Farbman et al. [133], Shen et al. [116] and the three proposed methods. As can be seen from Table I, the ADF method obtains the largest value for three quality evaluation indexes computed for “House” and “Door” image data sets, which demonstrates that the ADF method provides a better performance in terms of the larger values for all tests. These results are consistent with

subjective visual perception. It is also confirmed that, for the “House” and “Door” image data sets, the GFF and WLF methods can capture abundant information of the original images into the fused image.

Table-I Quantitative assessment of different multi-exposure image fusion and tone mapping methods

Image Dataset	Index	Methods						
		Kaung et al. [189]	Mertens et al. [113]	Farbman et al. [133]	Shen et al. [116]	ADF	GFF	WLF
House	$Q^{AB/F}$ [186]	0.27	0.39	0.35	0.32	0.44	0.41	0.38
	Q_{MI} [191]	2.45	3.23	2.33	1.97	4.38	3.51	3.45
	Q_{SF} [190]	7.24	9.46	8.13	6.38	10.16	8.74	9.64
Door	$Q^{AB/F}$	0.43	0.61	0.59	0.44	0.64	0.63	0.57
	Q_{MI}	1.94	2.76	2.01	1.48	2.80	2.48	2.79
	Q_{SF}	6.51	7.88	7.26	6.33	13.31	9.70	10.43
Igloo	$Q^{AB/F}$	0.59	0.62	0.54	0.61	0.65	0.60	0.61
	Q_{MI}	2.03	2.69	2.25	2.14	2.30	2.58	2.45
	Q_{SF}	7.06	8.96	8.43	6.27	11.02	10.11	9.39
Venice	$Q^{AB/F}$	0.41	0.56	-	0.34	0.54	0.47	0.49
	Q_{MI}	5.34	6.84	-	5.63	5.64	5.43	6.27
	Q_{SF}	6.95	8.05	-	6.20	7.95	6.33	7.82
Bellavita	$Q^{AB/F}$	-	0.27	-	-	0.63	0.43	0.44
	Q_{MI}	-	3.55	-	-	3.22	3.40	3.41
	Q_{SF}	-	5.44	-	-	10.56	7.60	9.98
Cathedral	$Q^{AB/F}$	0.39	0.54	0.45	0.58	0.579	0.572	0.75
	Q_{MI}	3.35	2.58	2.54	2.27	2.77	2.96	4.65
	Q_{SF}	5.73	6.22	6.25	5.98	7.48	6.73	7.52
Lizard	$Q^{AB/F}$	0.48	0.71	0.63	0.52	0.72	0.61	0.75
	Q_{MI}	4.56	6.56	5.75	2.85	6.69	6.82	7.55
	Q_{SF}	3.12	3.66	3.67	2.95	5.27	5.42	6.85
Belgium	$Q^{AB/F}$	-	0.19	-	0.24	0.35	0.26	0.21
	Q_{MI}	-	3.12	-	2.95	3.00	2.08	2.47
	Q_{SF}	-	8.54	-	7.41	10.72	8.02	7.75

Other than the first example, the ADF method generates fused images for the “Igloo”, “Bellavita” and “Belgium” image data sets with a significantly higher value of $Q^{AB/F}$ and Q_{SF} than those produced by the exposure fusion technique. The spatial frequency of the fused image is also often considered as one of the quality parameters during the performance evaluation of the fusion techniques. It can be seen that the proposed technique yields fused images with a higher value of spatial frequency. From this table, it can be observed that the results produced by WLF method for the “Cathedral” and “Lizard” image data sets provides the best performance in terms of the largest $Q^{AB/F}$, Q_{MI} and Q_{SF} . It can be concluded that the proposed image fusion methods can better preserve spatial frequency, the edge information of source images, and the original information from source images is preserved in the fused image.

Table-II Quantitative assessment of different multi-focus image fusion methods.

Image	Index					
Dataset		Adu & Wang [110]	Hod’akov’a et al. [185]	ADF	GFF	WLF
Clock	$Q^{AB/F}$	0.45	0.46	0.54	0.42	0.59
	Q_{MI}	6.39	5.73	6.46	4.93	6.53
	Q_{SF}	4.45	4.66	5.34	4.89	6.36
Book	$Q^{AB/F}$	0.63	0.58	0.51	0.46	0.69
	Q_{MI}	6.82	4.47	5.36	4.92	7.11
	Q_{SF}	6.93	5.46	6.85	5.82	8.08
Table	$Q^{AB/F}$	0.58	0.43	0.61	0.62	0.81
	Q_{MI}	5.82	5.71	5.98	5.44	6.47
	Q_{SF}	6.10	5.23	6.64	5.20	6.69
Documents	$Q^{AB/F}$	0.31	-	0.34	0.38	0.54
	Q_{MI}	1.75	-	1.82	1.86	2.87
	Q_{SF}	12.69	-	14.49	13.77	15.32
Leaves	$Q^{AB/F}$	0.21	-	0.28	0.27	0.35
	Q_{MI}	3.02	-	3.92	3.07	3.18
	Q_{SF}	11.63	-	11.89	10.21	12.03

6.5 SUMMARY

Above discussion makes it clear that detail-enhanced image can be produced by using the edge-preserving filters. The details present in the fused image can be controlled with the help of different free parameters. By increasing or decreasing the value of these free parameters, we can control texture details in the fused image, which in turn will produce good visual perception. Experimental results demonstrated that the present methods have applicability for capturing details from multi-exposure and multi-focus image databases.

The objective performance analysis based on three quality metrics ($Q^{AB/F}$, Q_{MI} and Q_{SF}) demonstrated that proposed image fusion methods can better preserve spatial frequency, the edge information of source images, and the original information from source images is preserved in the fused image. From Table II, it can be observed that the results produced by WLF method for all multi-focus image data sets provides the best performance in terms of the largest $Q^{AB/F}$, Q_{MI} and Q_{SF} .

CHAPTER – 7

CONCLUSION AND FUTURE SCOPE

The present thesis proposes three methods to construct a detail enhanced image from a set of multi-exposure images by using a MRP technique and WLS optimization framework. When compared with the existing techniques which use multi-resolution and single resolution analysis for exposure fusion the proposed methods perform better in terms of enhancement of texture details in the fused image. The frameworks are inspired by the edge-preserving property of AD, and GF that has better response near strong edges. The two layer decomposition based on edge-preserving filters is used to extract fine textures for detail enhancement.

In particular, experimental results demonstrated that the present approaches have applicability for capturing details from poorly illuminated and brightly illuminated areas simultaneously. One of the important applications of proposed techniques for security application is capturing details at entrance of the buildings. Conventional cameras are not able to faithfully capture the interior and exterior of a building simultaneously while present approach would be able to simultaneously record indoor as well as outdoor activities. Other important applications of proposed approaches can be the satellite, scientific, and medical imagery, in which data is analyzed and visualized to record more details.

Moreover, the results proposed in the thesis have demonstrated that the present methods can also be applied to fuse multi-focus images (i.e. images focused on different targets), and images captured with flash and no-flash. More importantly, it is interesting to note that the information in the resultant image can be controlled with the help of proposed free parameters.

At last, the future work involves improvement of these methods for adaptively choosing the parameters of the edge-preserving filter and checking the utilization of present methods for different kinds of image fusion applications. In future, work can be extended in multi-modal image fusion problem. Also the work can be done to optimize the weight maps depending upon various other optimization techniques.

REFERENCES

- [1] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg, "Model of visual adaptation for realistic image synthesis," in *Proc. of the Computer Graphics Conf. (SIGGRAPH '96)*, pp. 249-258, August 1996.
- [2] Fanping Zhou, and Jochen Lang, "Efficient spherical high dynamic range imaging for image-based virtual environments," in *Proc. of IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*, pp. 124 – 129, October 2014.
- [3] Li Zhengguo, and Jinghong Zheng, "Visual-salience-based tone mapping for high dynamic range images," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 12, pp. 7076 - 7082, December 2014.
- [4] J. M. Geigel. "Tone reproduction for computer graphics using photographic principles," *PhD thesis, The George Washington University*, UMI Order Number: AAI9961552, 2000.
- [5] Bryan Peterson, "Understanding exposure: How to shoot great photographs with any camera paperback," *Bryan Peterson*, 3rd Edition, August, 2010.
- [6] Jones Loyd Ancile, and H. R. Condit, "Sunlight and skylight as determinants of Photographic exposure. I. Luminous density as determined by solar altitude and atmospheric conditions," *Journal of the Optical Society of America*, vol. 38, no 2, pp. 123–178, February 1948.
- [7] ANSI PH2.7, "*Photography-photographic exposure guide*," *American National Standards Institute*, New York, pp. 18-26, 1986.
- [8] M. Gupta, D. Iso, and S.K. Nayar, "Fibonacci exposure bracketing for high dynamic range imaging," in *Proc. of IEEE International Conference on Computer Vision (ICCV)*, pp. 1473 – 1480, December 2013.
- [9] David Yang, and Abbas El Gamal, "Comparative analysis of SNR for image sensor with enhanced dynamic range," in *Proc. of SPIE*, vol. 3649, San Jose, CA, pp. 197-211, January 1999.
- [10] Damon M. Chandler, and Sheila S. Hemami, "Additivity models for supra-threshold distortion in quantized wavelet-coded images" in *Proc. of SPIE on Human Vision and Electronic Imaging VII*, vol. 4662, May 2002.

- [11] Tunç Aydın, Nikolce Stefanoski, Simone Croci, Markus Gross, and Aljoscha Smolic, “Temporally coherent local tone-mapping of HDR video,” *ACM Trans. on Graphics*, vol. 33, no 6, 2014.
- [12] Ahmet Akyüz, Kerem Hadimli, Merve Aydinlilar, and Christian Bloch, “Style-based tone-mapping for HDR images,” in *Proc. of the 6th ACM SIGGRAPH Conference and Exhibition on Computer Graphics and Interactive Techniques in Asia*, vol. 23, pp. 1-4, November 2013.
- [13] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, “Photographic tone-reproduction for digital images,” *ACM Trans. on Graphics*, vol. 21, no. 3, pp. 267-276, 2002.
- [14] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, “High dynamic range imaging acquisition, manipulation and, display,” *Morgan Kaufmann*, 2005.
- [15] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, “High dynamic range display system,” *ACM Trans. on Graphics*, vol. 23, no. 3, pp. 760-768, 2004.
- [16] H. Seetzen, L. A. Whitehead, and G. Ward, “A high dynamic range display using low and high resolution modulator,” in *the society for Information Display International Symposium*, vol. 34, no. 1, pp. 1450-1453, 2003.
- [17] P. E. Debevec, and J. Malik, “Recovering high dynamic range radiance maps from photographs,” in *Proc. of the 24th ACM Annual Conference on Computer Graphics and Interactive techniques (SIGGRAPH '97)*, Los Angeles, Calif. USA, pp. 369–378, August 1997.
- [18] B. J Chang., and K. Winick, “Silver-Halide Gelatin holograms,” in *Proc. of the SPIE Conferences on Recent Advances in Holography*, vol. 0215, 172, May 1980.
- [19] Kenji Naoi, Yoshihisa Ohko, and Tetsu Tatsuma “TiO₂ films loaded with silver nanoparticles: Control of multicolor photochromic behavior,” *Journal of the American Chemical Society*, vol. 126, no. 11, pp. 3664-3668, 2004.
- [20] Franz Josef Gassmann, “Camera, process for reconstructing an image information, and process for calibrating an image information,” *United States Patent*, US7346275 B2, March 2008.
- [21] Richards, and Dan, “Black & white photography,” *PSA Journal*, vol. 77 no. 12, pp. 38–40, December 2011.
- [22] Eastman Kodak Company, “Applied science fiction web site, www.asf.com/products/FPS/fpsfaqs.shtml, July 2004.

- [23] S. Boyle Willard, and George Elwood Smith, "The inception of charge-coupled devices," *IEEE Trans. on Electron devices*, vol. ED-23, no. 7, July 1976.
- [24] J. P. Carrere, S. Place, J. P. Oddou, D. Benoit, and F. Roy, "CMOS image sensor: Process impact on dark current," in *Proc. of the IEEE Conference on Reliability Physics Symposium*, Waikoloa, HI, pp. 3C.1.1 - 3C.1.6, June 2014.
- [25] J. P. Carrere, S. Place, J. P. Oddou, D. Benoit, and F. Roy, "Dictionary learning based color demosaicing for plenoptic cameras," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 455 – 460, June 2014.
- [26] Laroche, and M. Prescott, "Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients," *United States Patent*, 5, 373, 322, 1994.
- [27] J. Hamilton, and J. Adams, "Adaptive color plan interpolation in single sensor color electronic camera," *United States Patent*, 5,629,734, 1997.
- [28] L. Chang, and Y. P. Tan, "Color filter array demosaicking: new method and performance measures," *IEEE Trans. on Image Processing*, vol. 12, no. 10, pp. 1194-1210, October 2002.
- [29] R. Ramanath, W. Snyder, G. Bilbro, and W. Sander, "Demosaicking methods for the Bayer color arrays," *Journal of Electronic Imaging*, vol. 11, no. 3, pp. 306-315, July 2002.
- [30] D. D. Busch, "Mastering digital SLR photography," *Third Edition by Cengage Learning PTR*, January 2011.
- [31] Hicks, W. Roger, "The SEI photometer a legend among spot meters," *Shutterbug Magazine*, Retrieved December 2009.
- [32] M. Song, D. Tao, C. Chen, J. Bu, J. Luo, and C. Zhang, "Probabilistic exposure fusion," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 341-357, January 2012.
- [33] S. Mann, R. W. Picard, "Being "undigital" with digital cameras: extending dynamic range by combining differently exposed pictures," in *Proc. of the IS&T's 48th Annual Conference*, pp. 442-448, May 1995.
- [34] S. K. Nayar, and T. Mitsunaga, "High dynamic range imaging: Spatially varying pixel exposures," in *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2000)*, Head Island, SC, vol. 1, pp. 472-479, 2000.
- [35] Barbara Zitova, and Jan Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977-1000, October 2003.
- [36] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, pp. 325-376, 1992.

- [37] G. Ward, "Fast, robust image registration for compositing high dynamic range photographs from hand-held exposures," *journal of graphics tools*, vol. 8, no. 2, pp. 17-30, 2003.
- [38] S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High dynamic range video," *ACM Trans. on Graphics*, vol. 22, no. 3, 2003.
- [39] A. Tomaszewska, and R. Mantiuk, "Image registration for multi-exposure high dynamic range image acquisition," in *Proc. of the International Conference on Computer Graphics, Visualization and Computer Vision*, Plzen, Czech Republic, 2007.
- [40] B. D. Lucas, and T. Kanade, "An iterative image registration technique with an application in stereo vision," in *Proc. of the Seventh Int. Joint Conf. on Artificial Intelligence (IJCAI-81)*, pp. 674–679, 1981.
- [41] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91-110, 2004.
- [42] M. A. Fischler, and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381-395, 1981.
- [43] P. E. Debevec, "Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with illumination and high dynamic range photography," in *Proc. of the Annual Conf. Series, ACM SIGGRAPH*, pp. 45-50, 1998.
- [44] G. Ward-Larson, and R. A. Shakespeare, "Rendering with radiance," *San Francisco: Morgan Kaufmann*, 1998.
- [45] Adobe. Tiff 6.0 specification, 1992, <http://partners.adobe.com/asn/tech/tiff/specification.jsp>.
- [46] F. Kainz, R. Bogart, and D. Hess, "The OpenEXR image file format," in *Proc. of the Conference on SIGGRAPH Technical Sketches*, July 2003.
- [47] G. W. Larson, "Overcoming gamut and dynamic range limitations in digital images," in *Proc. of the IS&T 6th Color Imaging Conf., Scottsdale: IS&T*, 1998.
- [48] Greg Ward's website for HDR image formation and image data base: <http://www.anywhere.com/>.
- [49] V. R. Garcia-Hansen, M. Cowley, S. S. Smith, and G. Isoardi, "Testing the accuracy of luminance maps acquired by smart phone cameras," in *Proc. of the CIE Centenary Conference on Towards a New Century of Light, Commission International Eclairage*, Paris, pp. 951-955, April 2013 .
- [50] HDR image formation software: <http://www.hdrshop.com/>.

- [51] F. Mccollough, "Complete guide to high dynamic range digital photography", *China: Lark Books*, 2008.
- [52] Michael Freeman, "Mastering HDR photography: combining technology and artistry to create high dynamic range images," *Amphoto Books*, 160 pages, April 2008.
- [53] S. Nayar, and V. Branzoi, "Adaptive dynamic range imaging: Optical control of pixel exposures over space and time," in *Proc. of the Ninth International Conference on Computer Vision (ICCV)*, Nice, France, vol. 2, pp. 1168-1175, 2003.
- [54] S. Nayar, and S. Narasimhan, "Assorted pixels: multi-sampled imaging with structural models," in *Proc. of the European Conference on Computer Vision (ECCV)*, Copenhagen, Denmark, vol. 4, pp. 636-652, May 2002.
- [55] S. Narasimhan, and S. Nayar, "Enhancing resolution along multiple imaging dimensions using assorted pixels," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 518-530, 2005.
- [56] M. Aggarwal, and N. Ahuja, "Split aperture imaging for high dynamic range," *Int. J. Comput. Vision*, vol. 58, no. 1, pp. 7-17, 2004.
- [57] L. C. Gouveia, M. Waqas, and B. Choubey, "A reconfigurable CMOS pixel for applying tone-mapping on high dynamic range images," in *Proc. of IEEE Instrumentation and Measurement Technology Conference (I2MTC)*, Montevideo, Uruguay, pp. 1098 - 1101, May 2014.
- [58] D. Yang, A. El Gamal, B. Fowler, and H. Tian, "A 640x512 CMOS image sensor with ultra-wide dynamic range floating-point pixel level ADC," *IEEE Journal of Solid State Circuits*, vol. 34, no. 12, pp. 1821-1834, 1999.
- [59] E. Adelson, and J. Bergen, "The plenoptic function and the elements of early vision," in *Proc. of Conference on Computational models of visual processing*, MIT Press, vol. 1, no. 2, pp. 3-20, 1991.
- [60] T. Georgiev, and A. Lumsdaine, "Rich image capture with plenoptic cameras," in *Proc. of IEEE International Conference on Computational Photography (ICCP)*, Cambridge MA, USA, pp. 1-8, 2010.
- [61] D. Das, and S. Collins, "Fixed-pattern-noise correction for an integrating wide-dynamic-range cmos image sensor," *IEEE Trans. on Electron Devices*, vol. 60, no. 1, pp. 314 - 319, November 2012.
- [62] Z. Li, J. Zheng, and S. Rahardja, "Detail-enhanced exposure fusion," *IEEE Trans. on Image Processing*, vol. 21, no. 11, pp. 4672-4676, 2012.

- [63] A. Spivak, A. Belenky, A. Fish, and O. Yadid-Pecht, "Wide-dynamic-range CMOS image sensors comparative performance analysis," *IEEE Trans. on Electron Devices*, vol. 56, pp. 2446–2461, November 2009.
- [64] A. Spivak, A. Teman, A. Belenky, O. Yadid-Pecht, and A. Fish, "Power-performance tradeoffs in wide dynamic range image sensors with multiple reset approach," *Journal of Low Power Electronics and Applications*, vol. 1, no. 1, pp. 59-76, 2011.
- [65] G. Qiu, and J. Duan, "An optimal tone reproduction curve operator for the display of high dynamic range images," in *Proceedings of IEEE international Symposium on Circuits and Systems (ISCAS)*, vol. 6, pp. 6276-6279, 2005.
- [66] S. Palmer, "Vision science: Photons to phenomenology," *The MIT Press*, ch. 3.3, *Surface-Based Color Processing*, 1999.
- [67] M. D. Fairchild, "Color appearance models," *Addison-Wesley*, ISBN 0-201-63464-3, 1998.
- [68] B. A. Wandell, "Foundations of vision," *Sinauer Associates, Sunderland, Massachusetts*, 1995.
- [69] M. H. Kim, T. Weyrich, and J. Kautz, "Modeling human color perception under extended luminance levels," *ACM Trans. on Graphics*, vol. 28, no. 3, August 2009.
- [70] M. Fairchild, and G. Johnson, "Image appearance modeling," in *Human Vision and Electronic Imaging VIII, SPIE*, vol. 5007, January 2003.
- [71] K. Dmitriev, T. Annen, G. Krawczyk, K. Myszkowski, and H. P. Seidel, "A CAVE system for interactive modeling of global illumination in car interior," in *ACM Symposium on Virtual Reality Software and Technology (VRST 2004)*, *ACM Press*, New York, NY, USA, pp. 137-145, 2004.
- [72] Y. Li, L. Sharan, and E. H. Adelson, "Compressing and companding high dynamic range images with sub-band architectures," *ACM Trans. on Graphics*, vol. 24 no. 3, pp. 836-844, 2005.
- [73] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. on Graphics*, vol. 21, no. 3, pp. 249-256, 2002.
- [74] F. Durand, and J. Dorsey, "Fast bilateral filtering for the display of high dynamic range images," *ACM Trans. on Graphics*, vol. 21, no. 3, pp. 257-266, 2002.
- [75] G. W. Larson, H. E. Rushmeier, and C. D. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes," *IEEE Trans. on Visualization and Computer Graphics*, vol. 3, no. 4, pp. 291-306, 1997.

- [76] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Journal of Computer Graphics Forum*, vol. 22, pp. 419-426, 2003.
- [77] E. Reinhard, and K. Devlin, "Dynamic range reduction inspired by photoreceptor physiology," *IEEE Trans. on Visualization and Computer Graphics*, vol. 11, no. 1, pp. 13-24, 2005.
- [78] J. Tumblin, J. K. Hodgins, and B. K. Guenter, "Two methods for display of high contrast images," *ACM Trans. on Graphics*, vol. 18, no. 1, pp. 56-94, 1999.
- [79] W. F. Lee, T. Y. Lin, M. L. Chu, T. H. Huang, and H. H. Chen, "Perception-based high dynamic range compression in gradient domain," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '09)*, pp. 1805–1808, November 2009.
- [80] Yongqing Huo, Fan Yang, and Vincent Brost, "Inverse tone-mapping based upon retina response," *The Scientific World Journal*, vol. 2014, Article ID 168564, 5 pages, 2014. doi:10.1155/2014/168564.
- [81] G. D. Licciardo, A. D'Arienzo, and A. Rubino, "Stream processor for real-time inverse tone mapping of full-HD images," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, vol. PP, no. 99, November 2014.
- [82] A. G. Rempel, M. Trentacoste, H. Seetzen, H. D. Young, Heidrich, W. L. Whitehead, and G. Ward, "Ldr2hdr: on-the-fly reverse tone mapping of legacy video and photographs," *ACM Trans. Graph.*, vol. 26, no. 3, pp. 39, 2007.
- [83] L. Meylan, S. Daly, and S. S  ijsstrunk, "The reproduction of specular highlights on high dynamic range displays," in *IST/SID 14th Color Imaging Conference*, vol. 6, pp. 333-338, May 2006.
- [84] F. Banterle, P. Ledda, K. Debattista, and A. Chalmers, "Expanding low dynamic range videos for high dynamic range applications," in *ACM Spring Conference on Computer Graphics, (Budmerice Castle, Slovakia)*, New York, NY, USA, pp. 33-41, 2008.
- [85] F. Banterle, P. Ledda, K. Debattista, and A. Chalmers, "Inverse tone mapping," in *GRAPHITE '06: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*, (Kuala Lumpur, Malaysia), ACM, New York, NY, USA, pp. 349-356, 2006.
- [86] A. O. Aky  uz, R. Fleming, B. E. Riecke, E. Reinhard, and H. H. B  ulthoff, "Do hdr displays support ldr content?: a psychophysical evaluation," *ACM Trans. Graph.*, vol. 26, July 2007.

- [87] B. Masia, S. Agustin, R. W. Fleming, O. Sorkine, and D. Gutierrez, "Evaluation of reverse tone mapping through varying exposure conditions," *ACM Trans. Graph.*, vol. 28, pp. 160:1-160:8, December 2009.
- [88] T. O. Aydin, R. Mantiuk, K. Myszkowski, and H. P. Seidel, "Dynamic range independent image quality assessment," *ACM Trans. Graph.*, vol. 27, no. 3, 2008.
- [89] P. J. Burt, and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. on Communications*, vol. 31, no. 4, pp. 532-540, 1983.
- [90] M. I. Smith, and J. P. Heather, "Review of image fusion technology," in *Proceedings of the Defense and Security Symposium*, Orlando, Fla, USA, vol. 5782, pp. 29–45, 2005.
- [91] G. Bhatnagar, Q. M. J. Wu, and Zheng Liu, "Directive contrast based multimodal medical image fusion in NSCT domain," *IEEE Trans. on Multimedia*, vol. 15, no. 5, pp. 1014 – 1024, February 2013.
- [92] A. A. Goshtasby, "2-D and 3-D image registration for medical, remote sensing, and industrial applications," *Wiley-interscience publication*, Pages 284, February 2005.
- [93] J. M. Orgden, E. H. Adelson, J. R. Bergen, and P. J. Burt, "Pyramid based computer graphics," *RCA Engineer*, vol. 30, no. 5, pp. 4-15, 1985.
- [94] V. S. Petrovic, and C. S. Xydeas, "Gradient-based multiresolution image fusion," *IEEE Trans. on Image Processing*, vol. 13, no. 2, pp. 228-237, 2004.
- [95] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," in *Proc. of the 11th annual international conference of the Center for Nonlinear Studies on Experimental mathematics: computational issues in nonlinear science: computational issues in nonlinear science*, vol. 60, pp. 259-268, November 1992.
- [96] M. Kumar, and S. Dass, "A total variation-based algorithm for pixel level image fusion," *IEEE Trans. on Image Process.*, vol. 18, no. 9, pp. 2137–2143, September 2009.
- [97] A. Ben Hamza, Yun He, Hamid Krim, and Alan Willsky, "A multiscale approach to pixel-level image fusion," *Integr. Comput.-Aided Eng.*, vol.12, no. 2, 135-146, April 2005.
- [98] A. Agrawal, R. Raskar, S. K. Nayar, and Y. Li, "Removing photography artifacts using gradient projection and flash exposure sampling," *ACM Trans. on Graphics*, vol. 24, no. 3, pp. 828-835, 2005.
- [99] G. Petschnigg, R. Szeliski, M. Agrawala, M. F. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Trans. on Graph.*, vol. 23, no. 3, pp. 664-672, 2004.

- [100] S. Li, J. T.Y. Kwok, I. W. H. Tsang, and Y. Wang, "Fusing images with different focuses using support vector machines," *IEEE Trans. on Neural Netw.*, vol. 15, no. 6, pp. 1555–1561, Nov. 2004.
- [101] V. Vapnik, "The Nature of statistical learning theory," *New York: Springer-Verlag*, January 1995.
- [102] J. A. K. Suykens, and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Process. Lett.*, vol. 9, no. 3, pp. 293–300, 1999.
- [103] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, "Least squares support vector machines," *World Scientific*, Singapore, 2002.
- [104] S. Zheng, Y. Q. Sun, J. W. Tian, and J. Liu, "Mapped least squares support vector regression," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 19, no. 3, pp. 459–475, 2005.
- [105] S. T. Li, and B. Yang, "Multi-focus image fusion using region segmentation and spatial frequency," *Image and Vision Computing*, vol. 26, pp. 971-979, 2008.
- [106] J. Shi, and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905, 2000.
- [107] S. Li, X. Kang, J. Hu, and B. Yang, "Image matting for fusion of multi-focus images in dynamic scenes," *Information Fusion*, vol. 14, no. 2, pp. 147-162, April 2013.
- [108] J. Wang, and M. F. Cohen, "Optimized color sampling for robust matting," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, USA, pp. 1-8, June 2007.
- [109] Kim Yong-Deok, and Choi Seungjin, "Weighted nonnegative matrix factorization," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1541-1544, April 2009.
- [110] JianHua ADu, MingHui Wang, "Multi-focus image fusion based on WNMF and focal point analysis," *Journal of Convergence Information Technology*, vol. 6, no. 7, pp. 109-117, 2011.
- [111] S. Raman, and S. Chaudhuri, "Bilateral filter based compositing for variable exposure photography," in *Proceedings of Eurographics*, Munich, Germany, 2009.
- [112] A. Goshtasby, "Fusion of multi-exposure images," *Image and Vision Computing*, vol. 23, pp. 611-618, 2005.
- [113] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion: a simple and practical alternative to high dynamic range photography," *Computer Graphics Forum*, vol. 1, pp. 161-171, 2008.

- [114] S. Li, and X. Kang, "Fast Multi-exposure image fusion with median filter and recursive filter," *IEEE Trans. on Consumer Electronics*, vol. 58, No. 2, May 2012.
- [115] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. on image processing*, vol. 22, no. 7, pp. 2864-2875, July 2013.
- [116] R. Shen, I. Cheng, J. Shi, and A. Basu, "Generalized random walks for fusion of multi-exposure images," *IEEE trans. on image processing*, vol. 20, no. 12, December 2011.
- [117] T. Kartalov, Z. Ivanovski, and L. Panovski, "Optimization of the pyramid height in the pyramid-based exposure," *Infoteh-Jahorina*, 10 Ref. B-III-7, pp. 219-223, 2011.
- [118] K. Kotwal, and S. Chaudhuri, "An optimization based approach to fusion of multi-exposure, low dynamic range images," in *Proc. of the 14th International Conference on Information Fusion*, Chicago, Illinois, USA, 2011.
- [119] S. Raman, and S. Chaudhuri, "A matte-less, variational approach to automatic scene compositing," in *Proc. of the IEEE 11th Int. Conference on Computer Vision*, pp. 1-6, October 2007.
- [120] W. Zhang, W. K. Cham, "Gradient-directed multi-exposure composition," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2318-2323, April 2012.
- [121] Ying Zhao, Jianbing Shen, and Ying He, "Sub-band architecture based exposure fusion," in *Proc. of the 4th Pacific-Rim Symposium on Image and Video Technology (PSIVT)*, pp. 501 - 506 , November 2010
- [122] Marco Block, Maxim Schaubert, Fabian Wiesel, and Raúl Rojas, "Multi-exposure document fusion based on edge-intensities," in *ICDAR, IEEE Computer Society*, 136-140, 2009.
- [123] An optical character recognition engine developed at HP labs (Tesseract-OCR): <http://code.google.com/p/tesseract-ocr/>.
- [124] K. He, J. Sun, and X. Tang, "Guided image filtering," in *Proceedings of the ECCV, Lecture Notes in Computer Science, Springer*, vol. 6311, pp. 1-14, 2010.
- [125] L. Grady, "Multilabel random walker image segmentation using prior models," in *Proc. of IEEE Conf. on Comput. Vis. Pattern Recognit.*, vol. 1, pp. 763–770, 2005.
- [126] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, November 2006.
- [127] P. G. Doyle, and J. L. Snell, "Random walks and electric networks," *Mathematical Association of America*, vol. 94, no. 2, pp. 202-205, 1984.
- [128] E. S. L. Gastal, and M. M. Oliveira, "Domain transform for edge-aware image and video processing," *ACM Trans. Graphics*, vol. 30, no. 4, pp. 69:1–69:11, July 2011.

- [129] R.C. Gonzalez, R.E. Woods, and S. Eddins, “Digital image processing using MATLAB,” *Prentice Hall*, New York, USA, 2004.
- [130] R. Li, S. Yu, and X. Yang, “Efficient spatio-temporal segmentation for extracting moving objects in video sequences,” *IEEE Trans. Consumer Electronics*, vol. 53, no. 3, pp. 1161–1167, August 2007.
- [131] R. Szeliski, “System and process for improving the uniformity of the exposure and tone of a digital image,” *U. S. Patent No.*, 6687400, February 2004.
- [132] J. H. Zheng, Z. G. Li, S. Rahardja, S. S. Yao, and W. Yao, “Collaborative image processing algorithm for details refinement and enhancement via multi-light images,” in *Proc. of IEEE International Conference on Acoustics Speech Signal Processing*, pp. 1382–1385, March 2010.
- [133] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, “Edge-preserving decompositions for multi-scale tone and detail manipulation,” *ACM Trans. on Graphics*, vol. 27, no. 3, article 67, 2008.
- [134] L. Yin, R. Yang, M. Gabbouj, and Y. Neuvo, “Weighted median filters: a tutorial,” *IEEE Trans. on Circuits Syst.*, vol. 41, pp. 157–192, 1996.
- [135] A. C. Bovik, T. S. Huang, and D. C. Munson, “A generalization of median filtering using linear combinations of order statistics,” *IEEE Trans. Acoust. Speech Signal Process.*, pp. 1342–1349, 1983.
- [136] Y. Lee, and S. Kassam, “Generalized median filtering and related nonlinear filtering techniques,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 31, pp. 672–683, 1985.
- [137] H. A. David, “Order Statistics,” *New York: Wiley*, 1981.
- [138] I. Pitas, and A.N. Venetsanopoulos, “Order statistics in digital image processing,” in *Proc. of the IEEE*, vol. 80, no. 12, pp. 1893 – 1921, December 1992.
- [139] L. R. Rabiner, M. R. Sambur, and C. E. Schmidt, “Applications of a nonlinear smoothing algorithm to speech processing,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 23, pp. 552–557, December 1975.
- [140] T. M. Nodes, and N. C. Gallagher, “Median filters: Some modification and their properties,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, no. 5, pp. 739–746, October 1980.
- [141] J. W. Tukey, “Nonlinear (nonsuperposable) methods for smoothing data,” in *ConJ Rec., EASCON*, p. 673, 1974.

- [142] E. Ataman, V. K. Aatre, and K. M. Wong, "Some statistical properties of median filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 29, pp. 1073-1075, October 1981.
- [143] M. S. Pan, J. T. Tang, and X. L. Yang, "A modified adaptive median filter method and its applications in medical images," *Biomedical Engineering: Applications, Basis and Communications*, vol. 22, no. 06, pp. 489-496, December 2010.
- [144] H. Gomez-Moreno, "A modified median filter for the removal of impulse noise based on the support vector machines," *Advances in signal processing and computer technologies*, 2001.
- [145] P. Yang, and O. A. Basir, "Adaptive weighted median filter using local entropy for ultrasonic image de-noising," in *Proc. of the 3rd International Symposium on Image and Signal Processing and Analysis (ISPA)*, vol. 2, September 2003.
- [146] T. Loupas, W. N. McDicken, and P. L. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images" *IEEE Trans. on Circuits and Systems*, vol. 36, no. 1, pp. 129 – 135, January 1989.
- [147] T. Bernholt, R. Fried, U. Gather, I. Wegener, "Modified repeated median filters," *Journal of Statistics and Computing, Springer*, vol. 16, no. 2, pp. 177-192, June 2006.
- [148] Ling Shao, Ruomei Yan, Xuelong Li, and Yan Liu "From heuristic optimization to dictionary learning: a review and comprehensive comparison of image denoising algorithms," *IEEE Trans. on Cybernetics*, vol. 44, no. 7, pp. 1001-1013, August 2013.
- [149] N. Himayat, S. A. Kassam, "Approximate performance analysis of edge preserving filters," *IEEE Trans. on Signal Processing*, vol. 41, no. 9, pp. 2764 – 2777, September 1993.
- [150] Nagashettappa Biradar, M. L. Dewal, and Manoj Kumar Rohit, "Edge preserved speckle noise reduction using integrated fuzzy filters," *International Scholarly Research Notices*, vol. 2014, Article ID 876434, 11 pages, 2014. doi:10.1155/2014/876434.
- [151] B. Widrow, and S. Haykin, "Least-mean-square adaptive filters," *Wiley-Interscience, First edition*, September 2003.
- [152] Ting-qin Yan, Min Qu, and Chang-xiong Zhou, "Image denoising with bemd and edge-preserving self-snake model," *Intelligent Computing Theory Lecture Notes in Computer Science, Springer*, vol. 8588, pp. 435-442, August 2014.
- [153] Paras Jain, and Vipin Tyagi "A survey of edge-preserving image denoising methods," *Information Systems Frontiers, Springer*, July 2014.

- [154] Paras Jain, and Vipin Tyagi “An adaptive edge-preserving image denoising using epsilon-median filtering in tetrolet domain,” *Emerging ICT for Bridging the Future – Proc. of the 49th Annual Convention of the Computer Society of India (CSI)*, vol. 1, pp. 393-400, 2015.
- [155] M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger, “Robust anisotropic diffusion,” *IEEE Trans. on Image Processing*, vol. 7, no. 3, pp. 421-432, 1998.
- [156] P. Perona, and J. Malik, “Scale-space and edge detection using anisotropic diffusion,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [157] S. Paris, S. W. Hasinoff, and J. Kautz, “Local Laplacian filters: edgeaware image processing with a Laplacian pyramid,” *ACM Trans. Graph.*, vol. 30, no. 4, 2011.
- [158] S. G. Mallat, “A theory for multi-resolution signal decomposition: The wavelet representation,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, no. 4, pp. 674-693, 1989.
- [159] A. L. Da Cunha, J. Zhou, and M. N. Do, “The nonsubsamped contourlet transform: Theory, design, and applications,” *IEEE Trans. on Image Processing*, vol. 15, no. 10, pp. 3089-3101, 2006.
- [160] M. N. Do, and M. Vetterli, “The contourlet transform: An efficient directional multi-resolution image representation,” *IEEE Trans. on Image Processing*, vol. 14, no. 12, pp. 2091-2106, 2005.
- [161] C. Wen, G. Gao, and Z. Chen, “Multi-resolution model for image denoising based on total least squares,” in *Proceedings of the 4th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD '07)*, pp. 622-626, August 2007.
- [162] S. Liu, “Adaptive scalar and vector median filtering of noisy colour images based on noise estimation,” *IET Image Processing*, vol. 5, no. 6, pp. 541-553, 2011.
- [163] D. N. Vizireanu, S. Halunga, and G. Marghescu, “Morphological skeleton decomposition interframe interpolation method,” *Journal of Electronic Imaging*, vol. 19, no. 2, article ID 023018, pp. 1-3, 2010.
- [164] R. Fattal, M. Agrawala, and S. Rusinkiewicz, “Multi-scale shape and detail enhancement from multi-light image collections,” *ACM Trans. Graph.*, vol. 26, no.3, 51, 2007.
- [165] Dani Lischinski, Zeev Farbman, Matt Uyttendaele, and Richard Szeliski, “Interactive local adjustment of tonal values,” *ACM Trans. Graph.*, vol. 25, no. 3, pp. 646-653, 2006.

- [166] Soonmin Bae, Sylvain Paris, Frédo Durand, “Two-scale tone management for photographic look,” *ACM Trans. Graphics (TOG)-Proceedings of ACM*, vol. 25 no. 3, pp. 637-645, July 2006.
- [167] C. Tomasi, and R. Manduchi, “Bilateral filtering for gray and color images,” in *Proc. of the ICCV, IEEE Computer Society*, pp. 839-846, 1998.
- [168] Ahmed A. Hussein, and Xiaochun Yang “Colorization using edge-preserving smoothing filter,” *Signal, Image and Video Processing, Springer*, vol. 8, no. 8, pp. 1681-1689, November 2014.
- [169] N. Nikolaou, and N. Papamarkos, “Color reduction for complex document images,” *Int. J. Imaging Syst. Technol*, vol. 19, no. 1, pp. 14–26, 2009.
- [170] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, “Bilateral filtering: theory and applications,” *Foundations and Trends in, Computer Graphics and Vision*, vol. 4, no. 1, pp. 1-73, 2008.
- [171] J. Canny, “Finding edges and lines in images,” *Tech. Rep. 720, MIT, Artificial Intelligence Laboratory*, 1983.
- [172] P. J. Burt, "Fast filter transforms for image processing," *Computer Graphics, Image Processing*, vol. 6, pp. 20-51, 1981.
- [173] J. Shen, Y. Zhao, and Y. He, “Detail-preserving exposure fusion using sub-band architecture,” *The Visual Computer*, vol. 28, no. 5, pp. 463–473, 2012.
- [174] A. A. Minai, and R. D. Williams, “On the derivatives of the sigmoid,” *Neural Networks*, vol. 6, no. 6, pp. 845–853, 1993.
- [175] N. Draper, and H. Smith, “Applied regression analysis,” *John Wiley*, 2nd edition, 1981.
- [176] G. Juan Lu-Yin, and R. Luo Ming, “Magnitude estimation for scaling saturation,” in *Proceedings of SPIE, 9th Congress of the International Colour Association*, vol. 4421, pp. 575–578, April 2001.
- [177] Rafael C Gonzalez, “Digital image processing,” *Pearson Education India*, 2009.
- [178] Linda G. Shapiro, and George C. Stockman, “Computer vision,” *Upper Saddle River: Prentice–Hall*, 2001
- [179] K. Laws, “Textured image segmentation,” *Ph.D. Dissertation*, University of Southern California, January 1980.
- [180] D. Barash, “A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 844–847, 2002.

- [181] Armstrong J. Scott, "Illusions in regression analysis," *International Journal of Forecasting (forthcoming)*, vol. 28, no. 3, p. 689, 2012.
- [182] White Halbert, "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica*, vol. 48, no. 4, pp 817–838, doi: 10.2307/1912934, JSTOR 1912934, 1980.
- [183] Ng. Andrew, "Probabilistic interpretation," *Chapter 4 of CS229 Lecture notes1*, 2012.
- [184] J. Tumblin, and G. Turk, "LCIS: a boundary hierarchy for detail preserving contrast reduction," in *Proceedings of the ACM SIGGRAPH*, A. Rockwood, pp. 83–90, 1999.
- [185] P. Hodákov'a, I. Perfilieva, M. D'ankov'a, and M. Vajgl, "F-transform based image fusion," in *Image Fusion*, O. Ukimura, Ed., In Tech, Rijeka, Croatia, pp. 3–22, 2011.
- [186] C. S. Xydeas, and V. Petrovi'c, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.
- [187] Y. Han, Y. Cai, Y. Cao, and X. Xu, "A new image fusion performance metric based on visual information fidelity," *Information Fusion*, vol. 14, no. 2, pp. 127–135, 2013.
- [188] T. O. Aydin, R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "Dynamic range independent image quality assessment," *ACM Trans. on Graphics*, vol. 27, no. 3, article 69, 2008.
- [189] 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.
- [190] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electronics Letters*, vol. 38, no. 7, pp. 313–315, 2002.
- [191] R. Sakuldee, and S. Udomhunsakul, "Objective performance of compressed image quality assessments," in *Proceedings of World Academy of Science, Engineering and Technology*, vol. 26, pp. 434–443, 2007.
- [192] P. M. Zeeuw, "Wavelet and image fusion, CWI," Amsterdam, The Netherlands, March 1998, <http://www.cwi.nl/~pauldz/>.
- [193] J. Tian, L. Chen, L. Ma, and W. Yu, "Multi-focus image fusion using a bilateral gradient-based sharpness criterion," *Optics Communications*, vol. 284, no. 1, pp. 80–87, 2011.

APPENDIX A

Derivation of the computation of a_k and b_k used in GF:

We derive the linear coefficients of the GF used in Chapter 4, Equations (4.6) and Equation (4.7) as follows:

a_k and b_k are computed within ω_k to minimize following cost function:

$$E(a_k, b_k) = \sum_{i \in \omega_k} \left((a_k I_i^G + b_k - p_i)^2 + \varepsilon a_k^2 \right) \quad (1)$$

Differentiating Equation (1) w. r. t. two unknowns (a_k and b_k), and setting the result equal to zero.

$$\frac{\partial E}{\partial a_k} = \sum_{i \in \omega_k} (2I_i^G (a_k I_i^G + b_k - p_i) + 2\varepsilon a_k) = 0 \quad (2)$$

$$\frac{\partial E}{\partial b_k} = \sum_{i \in \omega_k} 2(a_k I_i^G + b_k - p_i) = 0 \quad (3)$$

Rearranging Equation (2) and Equation (3) to get normal equations:

$$\begin{aligned} a_k \sum_{i=1}^n (I_i^G)^2 + b_k \sum_{i=1}^n I_i^G + \varepsilon \sum_{i=1}^n a_k &= \sum_{i=1}^n I_i^G p_i \\ a_k \sum_{i=1}^n I_i^G + \sum_{i=1}^n b_k &= \sum_{i=1}^n p_i \\ a_k \sum_{i=1}^n (I_i^G)^2 + b_k \sum_{i=1}^n I_i^G + \varepsilon n a_k &= \sum_{i=1}^n I_i^G p_i \end{aligned} \quad (4)$$

$$a_k \sum_{i=1}^n I_i^G + n b_k = \sum_{i=1}^n p_i \quad (5)$$

To solve equations for a_k multiply Equation (4) by n and Equation (5) by $\sum_{i=1}^n I_i^G$, we have:

$$\begin{aligned} a_k \sum_{i=1}^n (I_i^G)^2 + b_k \sum_{i=1}^n I_i^G + \varepsilon n a_k &= \sum_{i=1}^n I_i^G p_i \Big] \times n \\ a_k \sum_{i=1}^n I_i^G + n b_k &= \sum_{i=1}^n p_i \Big] \times \sum_{i=1}^n I_i^G \end{aligned}$$

$$na_{\kappa} \sum_{i=1}^n (I_i^G)^2 + nb_{\kappa} \sum_{i=1}^n I_i^G + \varepsilon n^2 a_{\kappa} = n \sum_{i=1}^n I_i^G p_i$$

$$a_{\kappa} \left(\sum_{i=1}^n I_i^G \right)^2 + nb_{\kappa} \sum_{i=1}^n I_i^G = \sum_{i=1}^n p_i \sum_{i=1}^n I_i^G$$

$$na_{\kappa} \sum_{i=1}^n (I_i^G)^2 + \varepsilon n^2 a_{\kappa} - a_{\kappa} \left(\sum_{i=1}^n I_i^G \right)^2 = n \sum_{i=1}^n I_i^G p_i - \sum_{i=1}^n p_i \sum_{i=1}^n I_i^G$$

$$a_{\kappa} = \frac{n \sum_{i=1}^n I_i^G p_i - \sum_{i=1}^n p_i \sum_{i=1}^n I_i^G}{n \sum_{i=1}^n (I_i^G)^2 + \varepsilon n^2 - \left(\sum_{i=1}^n I_i^G \right)^2} \quad (6)$$

Dividing numerator and denominator of Equation (6) by n^2 , we have:

$$a_{\kappa} = \frac{\frac{1}{n} \sum_{i=1}^n I_i^G p_i - \frac{1}{n} \sum_{i=1}^n p_i \frac{1}{n} \sum_{i=1}^n I_i^G}{\frac{1}{n} \sum_{i=1}^n (I_i^G)^2 - \frac{1}{n} \sum_{i=1}^n I_i^G \frac{1}{n} \sum_{i=1}^n I_i^G + \varepsilon} \quad (7)$$

In GF I_i^G and p_i are identical, from Equation (7):

$$a_{\kappa} = \frac{\frac{1}{n} \sum_{i=1}^n (I_i^G)^2 - \frac{1}{n} \sum_{i=1}^n I_i^G \frac{1}{n} \sum_{i=1}^n I_i^G}{\frac{1}{n} \sum_{i=1}^n (I_i^G)^2 - \frac{1}{n} \sum_{i=1}^n I_i^G \frac{1}{n} \sum_{i=1}^n I_i^G + \varepsilon}$$

$$a_{\kappa} = \frac{\frac{1}{n} \sum_{i=1}^n (I_i^G)^2 - (\mu_{\kappa})^2}{\frac{1}{n} \sum_{i=1}^n (I_i^G)^2 - (\mu_{\kappa})^2 + \varepsilon}$$

$$a_{\kappa} = \frac{\sigma^2}{\sigma^2 + \varepsilon} \quad (8)$$

To derive b_{κ} from Equation (5):

$$a_{\kappa} \sum_{i=1}^n I_i^G + nb_{\kappa} = \sum_{i=1}^n p_i$$

$$nb_{\kappa} = \sum_{i=1}^n p_i - a_{\kappa} \sum_{i=1}^n I_i^G$$

$$b_{\kappa} = \frac{1}{n} \sum_{i=1}^n p_i - a_{\kappa} \frac{1}{n} \sum_{i=1}^n I_i^G$$

$$b_{\kappa} = \mu_{\kappa} - a_{\kappa} \mu_{\kappa} \quad (9)$$

APPENDIX B

Derivation for the computation of \hat{a}_k and \hat{b}_k used in WLS filter:

The linear coefficients of WLS filter in Chapter 5, Equations (5.28) and Equation (5.29) can be derived as follows:

Let \hat{q} be a linear transform of I^G in a window centered at the pixel k :

$$\hat{q}_i = \hat{a}_k I_i^G + \hat{b}_k, \forall i \in \omega_k \quad (10)$$

\hat{a}_k and \hat{b}_k are computed within ω_k to minimize following cost function:

$$\sum_{i \in \omega_k} \left((\hat{q}_i - p_i)^2 + \psi \left(c_{x,i}(p) \left(\frac{\partial \hat{q}}{\partial x} \right)^2 + d_{y,i}(p) \left(\frac{\partial \hat{q}}{\partial y} \right)^2 \right) \right) \quad (11)$$

$$\begin{aligned} \sum_{i \in \omega_k} \left((\hat{a}_k I_i^G + \hat{b}_k - p_i)^2 \right. \\ \left. + \psi \left(c_{x,i}(p) \left(\frac{\partial}{\partial x} (\hat{a}_k I_i^G + \hat{b}_k) \right)^2 + d_{y,i}(p) \left(\frac{\partial}{\partial y} (\hat{a}_k I_i^G + \hat{b}_k) \right)^2 \right) \right) \end{aligned} \quad (12)$$

$$\begin{aligned} \frac{\partial}{\partial x} (\hat{a}_k I_i^G + \hat{b}_k) &= \hat{a}_k \frac{\partial I_i^G}{\partial x} + \frac{\partial}{\partial x} \hat{b}_k \\ &= \hat{a}_k \frac{\partial I_i^G}{\partial x} + 0 \end{aligned} \quad (13)$$

$$\begin{aligned} \frac{\partial}{\partial y} (\hat{a}_k I_i^G + \hat{b}_k) &= \hat{a}_k \frac{\partial I_i^G}{\partial y} + \frac{\partial}{\partial y} \hat{b}_k \\ &= \hat{a}_k \frac{\partial I_i^G}{\partial y} + 0 \end{aligned} \quad (14)$$

$$\sum_{i \in \omega_k} \left((\hat{a}_k I_i^G + \hat{b}_k - p_i)^2 + \psi \left(c_{x,i}(p) \hat{a}_k^2 \left(\frac{\partial I_i^G}{\partial x} \right)^2 + d_{y,i}(p) \hat{a}_k^2 \left(\frac{\partial I_i^G}{\partial y} \right)^2 \right) \right) \quad (15)$$

Differentiating Equation (15) w. r. t. unknowns (\hat{a}_k and \hat{b}_k), and setting the result equal to zero.

$$\begin{aligned}\frac{\partial \mathbf{E}}{\partial \hat{a}_\ell} &= \sum_{i \in \omega_\ell} \left(2I_i^G (\hat{a}_\ell I_i^G + \hat{b}_\ell - p_i) + \psi \left(c_{x,i}(p) 2\hat{a}_\ell \left(\frac{\partial I_i^G}{\partial x} \right)^2 + d_{y,i}(p) 2\hat{a}_\ell \left(\frac{\partial I_i^G}{\partial y} \right)^2 \right) \right) \\ &= 0\end{aligned}$$

$$\frac{\partial \mathbf{E}}{\partial \hat{b}_\ell} = 2 \sum_{i \in \omega_\ell} (\hat{a}_\ell I_i^G + \hat{b}_\ell - p_i) = 0$$

$$\sum_{i \in \omega_\ell} \left(I_i^G (\hat{a}_\ell I_i^G + \hat{b}_\ell - p_i) + \hat{a}_\ell \psi \left(c_{x,i}(p) \left(\frac{\partial I_i^G}{\partial x} \right)^2 + d_{y,i}(p) \left(\frac{\partial I_i^G}{\partial y} \right)^2 \right) \right) = 0 \quad (16)$$

$$\sum_{i \in \omega_\ell} (\hat{a}_\ell I_i^G + \hat{b}_\ell - p_i) = 0$$

For simplicity, we assume that:

$$\sum_{i \in \omega_\ell} \left(c_{x,i}(p) \left(\frac{\partial I_i^G}{\partial x} \right)^2 + d_{y,i}(p) \left(\frac{\partial I_i^G}{\partial y} \right)^2 \right) = P \quad (17)$$

$$\sum_{i \in \omega_\ell} I_i^G (\hat{a}_\ell I_i^G + \hat{b}_\ell - p_i) + \hat{a}_\ell \psi P = 0 \quad (18)$$

$$\sum_{i \in \omega_\ell} (\hat{a}_\ell I_i^G + \hat{b}_\ell - p_i) = 0 \quad (19)$$

Rearranging Equation (18) and Equation (19) to get normal equations:

$$\begin{aligned}\hat{a}_\ell \sum_{i=1}^n (I_i^G)^2 + \hat{b}_\ell \sum_{i=1}^n I_i^G + \psi \hat{a}_\ell P &= \sum_{i=1}^n I_i^G p_i \\ \hat{a}_\ell \sum_{i=1}^n I_i^G + \sum_{i=1}^n \hat{b}_\ell &= \sum_{i=1}^n p_i \\ \hat{a}_\ell \left(\sum_{i=1}^n (I_i^G)^2 + \psi P \right) + \hat{b}_\ell \sum_{i=1}^n I_i^G &= \sum_{i=1}^n I_i^G p_i\end{aligned} \quad (20)$$

$$\hat{a}_\ell \sum_{i=1}^n I_i^G + n \hat{b}_\ell = \sum_{i=1}^n p_i \quad (21)$$

To solve equations for \hat{a}_ℓ multiply Equation (20) by n and Equation (21) by $\sum_{i=1}^n I_i^G$, we have:

$$\begin{aligned}\hat{a}_\ell \left(\sum_{i=1}^n (I_i^G)^2 + \psi P \right) + \hat{b}_\ell \sum_{i=1}^n I_i^G &= \sum_{i=1}^n I_i^G p_i \Big] \times n \\ \hat{a}_\ell \sum_{i=1}^n I_i^G + n \hat{b}_\ell &= \sum_{i=1}^n p_i \Big] \times \sum_{i=1}^n I_i^G\end{aligned}$$

$$n\hat{a}_k \left(\sum_{i=1}^n (I_i^G)^2 + \psi P \right) + n\hat{b}_k \sum_{i=1}^n I_i^G = n \sum_{i=1}^n I_i^G p_i$$

$$\hat{a}_k \left(\sum_{i=1}^n I_i^G \right)^2 + n\hat{b}_k \sum_{i=1}^n I_i^G = \sum_{i=1}^n p_i \sum_{i=1}^n I_i^G$$

$$n\hat{a}_k \left(\sum_{i=1}^n (I_i^G)^2 + \psi P \right) - \hat{a}_k \left(\sum_{i=1}^n I_i^G \right)^2 = n \sum_{i=1}^n I_i^G p_i - \sum_{i=1}^n p_i \sum_{i=1}^n I_i^G$$

$$\hat{a}_k \left(n \sum_{i=1}^n (I_i^G)^2 + n\psi P - \left(\sum_{i=1}^n I_i^G \right)^2 \right) = n \sum_{i=1}^n I_i^G p_i - \sum_{i=1}^n p_i \sum_{i=1}^n I_i^G$$

$$\hat{a}_k = \frac{n \sum_{i=1}^n I_i^G p_i - \sum_{i=1}^n p_i \sum_{i=1}^n I_i^G}{n \sum_{i=1}^n (I_i^G)^2 + n\psi P - \left(\sum_{i=1}^n I_i^G \right)^2} \quad (22)$$

Dividing numerator and denominator of Equation (21) by n^2 , we have:

$$\hat{a}_k = \frac{\frac{1}{n} \sum_{i=1}^n I_i^G p_i - \frac{1}{n} \sum_{i=1}^n p_i \frac{1}{n} \sum_{i=1}^n I_i^G}{\frac{1}{n} \sum_{i=1}^n (I_i^G)^2 + \frac{1}{n} \psi P - \frac{1}{n} \sum_{i=1}^n I_i^G \frac{1}{n} \sum_{i=1}^n I_i^G} \quad (23)$$

In GF I_i^G and p_i are identical, from Equation (23):

$$\hat{a}_k = \frac{\frac{1}{n} \sum_{i=1}^n (I_i^G)^2 - \frac{1}{n} \sum_{i=1}^n I_i^G \frac{1}{n} \sum_{i=1}^n I_i^G}{\frac{1}{n} \sum_{i=1}^n (I_i^G)^2 - \frac{1}{n} \sum_{i=1}^n I_i^G \frac{1}{n} \sum_{i=1}^n I_i^G + \frac{1}{n} \psi P}$$

$$\hat{a}_k = \frac{\frac{1}{n} \sum_{i=1}^n (I_i^G)^2 - (\mu_k)^2}{\frac{1}{n} \sum_{i=1}^n (I_i^G)^2 - (\mu_k)^2 + \frac{1}{n} \psi P}$$

$$\hat{a}_k = \frac{\sigma^2}{\sigma^2 + \frac{1}{n} \psi P} \quad (24)$$

To derive \hat{b}_k from Equation (21):

$$\hat{a}_k \sum_{i=1}^n I_i^G + n\hat{b}_k = \sum_{i=1}^n p_i$$

$$n\hat{b}_k = \sum_{i=1}^n p_i - \hat{a}_k \sum_{i=1}^n I_i^G$$

$$\hat{b}_k = \frac{1}{n} \sum_{i=1}^n p_i - \hat{a}_k \frac{1}{n} \sum_{i=1}^n I_i^G$$

$$\hat{b}_k = \mu_k - \hat{a}_k \mu_k \quad (25)$$

APPENDIX C

Derivation of matrix formulation of WLS filter:

The proof of matrix formulation of WLS filter in Chapter 5, Equation (5.35) can be derived as follows:

$$\sum_{i \in \omega_k} \left((\hat{q}_i - p_i)^2 + \psi \left(c_{x,i}(p) \left(\frac{\partial \hat{q}}{\partial x} \right)^2 + d_{y,i}(p) \left(\frac{\partial \hat{q}}{\partial y} \right)^2 \right) \right) \quad (26)$$

Let \hat{q} and p are diagonal matrices. Matrix formulation of Equation (26) is given by:

$$\begin{aligned} & (\hat{q} - p)^T (\hat{q} - p) + \psi (\hat{q}^T \hat{q} O_x^T O_x V_x + \hat{q}^T \hat{q} O_y^T O_y V_y) \\ & (\hat{q} - p)^T (\hat{q} - p) + \psi \hat{q}^T \hat{q} (O_x^T O_x V_x + O_y^T O_y V_y) \\ & (\hat{q} - p)^T (\hat{q} - p) + \psi \hat{q}^2 (O_x^T O_x V_x + O_y^T O_y V_y) \\ & \text{Let's assume that } L_{\hat{q}} = O_x^T O_x V_x + O_y^T O_y V_y \\ & (\hat{q} - p)^T (\hat{q} - p) + \psi \hat{q}^2 L_{\hat{q}} \\ & (\hat{q} - p)^2 + \psi \hat{q}^2 L_{\hat{q}} \end{aligned} \quad (27)$$

For Minimization the derivative of Equation (27) is taken w.r.t \hat{q} and put it equal to zero:

$$\begin{aligned} & 2(\hat{q} - p) + 2\psi \hat{q} L_{\hat{q}} = 0 \\ & 2[(\hat{q} - p) + \psi \hat{q} L_{\hat{q}}] = 0 \\ & \hat{q} - p + \psi \hat{q} L_{\hat{q}} = 0 \\ & \hat{q} (I_{dent} + \psi L_{\hat{q}}) = p \\ & \hat{q} = (I_{dent} + \psi L_{\hat{q}})^{-1} p \end{aligned} \quad (28)$$

Author's Publications

Following is a list of the publications that have been published during the accomplishment of the current Ph.D. thesis:

- [1] Harbinder Singh, Vinay Kumar, and Sunil Bhooshan, "Anisotropic Diffusion for Details Enhancement in Multi-exposure Image Fusion," *ISRN Signal Processing*, vol. 2013, Article ID 928971, 18 pages, 2013. doi:10.1155/2013/928971.

- [2] Harbinder Singh, Vinay Kumar, and Sunil Bhooshan, "A Novel Approach for Detail-Enhanced Exposure Fusion Using Guided Filter," *The Scientific World Journal*, vol. 2014, Article ID 659217, 8 pages, 2014. doi:10.1155/2014/659217.

- [3] Harbinder Singh, Vinay Kumar, and Sunil Bhooshan, "Weighted Least Squares Based Detail Enhanced Exposure Fusion," *ISRN Signal Processing*, vol. 2014, Article ID 498762, 18 pages, 2014. doi:10.1155/2014/498762.