

Image and Video Component Quality Measurements, Analyses, and Applications

A dissertation submitted by

Chen Gao

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

TUFTS UNIVERSITY

May 2015

© 2015, Chen Gao, All rights reserved

Adviser:

Prof. Karen Panetta

ABSTRACT

Currently, human experts provide the most reliable image quality evaluations for biomedical, industrial, and security applications. However, subjective evaluation by humans is impractical for real time applications. Therefore, it is desired to create an evaluation process that can be automated and not involve human intervention. This dissertation introduces a novel system for evaluating image quality objectively. The system utilizes multiple image measurements that evaluate one attribute aspect of an image. Each of these attribute measures is inspired by human visual system properties, and can be used separately as a standalone measure for evaluating a certain characteristic of an image. Multiple attributes are then fused together to provide an overall comprehensive evaluation for given image processing tasks. For the purpose of color image and video quality evaluation, attributes such as color, sharpness, and contrast are considered, while in applications where edge map qualities are important, edge pixel localization, corner presence, and double edge occurrence are used instead. To examine the uses of the measures for new image processing algorithm design and analysis, a set of alpha weighted quadratic filter based methods are developed for both color image enhancement and edge detection applications. The presented measures are used to automatically evaluate the image processing algorithm performances, as well as to assist in the selection of optimal operating parameters, all in accordance with the human visual system. The proposed measures are also shown to be applicable as a new method for fast database searching and retrieval. Furthermore, other new applications for the measures for monitoring thermal electrical system conditions, evaluating food quality, inspecting and exploring underwater environment, and for diagnosis and detecting of anomalies in biomedical images are presented.

To my father, mother, husband, and daughter

ACKNOWLEDGMENTS

I am grateful to my advisor, Dr. Karen Panetta for her continuous guidance and support during my research. She provided me with the great opportunity to work in her Signal and Image Laboratory at Tufts University. Her advice and ideas have been a source of great motivation for my work. I would thank her for her support through my graduate career at Tufts University, also for her efforts in helping me find jobs and internships. Especially, her consideration, understanding, and kindness for allowing me work remotely in the last year of my PhD program are always appreciated.

I am also thankful to my co-advisor Dr. Sos Aghaian at the University of Texas at San Antonio for his great advice. I am grateful to him for making the time to discuss research problems through hundreds of phone calls and directing me to innovative technologies. I am also thankful to Prof. Ronald Lasser for his advices both inside and outside the world of engineering. I thank Prof. Ethan Danahy for providing thorough review of my work. I appreciate all committee members for their valuable comments and suggestions.

I would like to thank all the members of the SIMlab, Shahan Nercassian, Arash Samani, Junjun Xia, Yicong Zhou, Debashree Mandal, Qianwen Wan, Chiamaka Chima, and Long Bao for their assistance with the instruments in the laboratory and fruitful discussions about the experimental results. I sincerely want to thank Mr. George Preble for being my mentor, encouraging and guiding me in all aspects of my life. I would like to thank the staff at the Upper Merion Township Library at King of Prussia for providing a quiet room for me where I completed the writing part of the thesis.

I would especially like to thank my parents Yanshou Gao and Juan Fang, for their constant love and care. This thesis would not have been possible without their support. I thank my husband Yansong Cheng for spending countless weekends and nights at library studying with me together. Last but not the least, I want to thank my daughter Emily Cheng for bringing endless happiness into my life and being my greatest motivation to complete my PhD program.

TABLE OF CONTENTS

ACKNOWLEDGMENTS.....	iv
LIST OF TABLES	viii
LIST OF FIGURES.....	xi
Chapter 1. Introduction.....	2
1.1. Overview.....	2
1.2. Objectives	5
1.3. Dissertation Organization.....	6
1.4. Dissertation Contributions	7
Chapter 2. Non-Reference Image and Video Attribute Measures	9
2.1. Overview.....	10
2.2. Color Measure.....	11
2.2.1 Color Models	11
2.2.2 Colorfulness Measure	14
2.2.3 Color Tone Measure	22
2.3. Contrast Measure	23
2.4. Sharpness Measure.....	29
2.5. Summary	32
Chapter 3. Non-Reference Color Image and Video Quality Measures	33
3.1. Overview.....	34
3.2. Color Quality Measure (CQM)	37
3.3. CQM for Images with the Same Source.....	38
3.4. CQM for Independent Images.....	43
3.5. Experimental Results	50
3.5.1 Parameter Selection	50
3.5.2 Benchmarking Image Processing Algorithms.....	51
3.5.3 Computational Complexity.....	54
3.6. Summary	54

Chapter 4. Reference Based Edge Measure.....	56
4.1. Overview and Problem Formulation	57
4.2. Related Works	59
4.2.1 Statistical Analysis.....	59
4.2.2 Pratt’s FOM	59
4.2.3 Pinho’s Improvement to Pratt’s FOM:.....	60
4.2.4 Boaventura’s Method.....	60
4.3. New Full-Reference Edge Measure RBEM	60
4.3.1 Edge Pixel Presence/Localization Measure	61
4.3.2 Edge Corner Presence/Localization Measure	63
4.3.3 Double Edge Occurrences Measure	66
4.3.4 Reference Based Edge Measure RBEM	69
4.4. Experimental Results	70
4.5. Summary	81
Chapter 5. Non-reference Based Edge Measure.....	82
5.1. Overview.....	83
5.2. Yitzhaky and Peli’s Edge Measure	84
5.3. New RBEM edge map measure	85
5.3.1 Grayscale Edge Map Generation	86
5.3.2 Reconstruction	86
5.3.3 Similarity Measure NREM:	89
5.4. Experimental Results	92
5.5. Summary	95
Chapter 6. Image Enhancement and Edge Detection Algorithms Using Alpha Weighted Quadratic Filter	97
6.1. Quadratic Filter	98
6.2. Application 1: Color Contrast Enhancement.....	101
6.2.1 Existing Spatial Domain Color Enhance Methods.....	102

6.2.2 Three-Dimensional Alpha Weighted Quadratic Filter (3DAWQF) For Color Contrast Enhancement	104
6.2.3 Performance Comparisons	114
6.3. Application 2: Edge Detection	118
6.3.1 Traditional Edge Detection Methods	118
6.3.2 Edge Detection Using AWQF	119
6.3.3 Experimental Results	123
6.4. Summary	128
Chapter 7. Practical Applications of the Measures	130
7.1. Color Patch Retrieval	131
7.2. Natural Color Image Retrieval	133
7.3. Electrical Component Thermography Inspection.....	136
7.4. Medical Image Quality Evaluation	137
7.5. Food Quality Evaluation	140
7.6. Underwater Inspection Assistance	143
Chapter 8. Concluding Remarks.....	145
8.1. Dissertation Contributions	145
8.2. Future Research Directions	147

LIST OF TABLES

Table	Page
Table 2-1. Existing non-reference colorfulness measures	15
Table 2-2. Image characteristics for six synthetic images in Figure 2-6	20
Table 2-3. Colorfulness measure values for four images with different colorfulness degradations in Figure 2-7. The greater measure values correspond to better image colorfulness	21
Table 2-4. Image characteristics for four Images with different colorfulness degradations in Figure 2-7.....	21
Table 2-5. Existing grayscale non-reference contrast measures	25
Table 2-6. Contrast measure values for the four images in Figure 2-12.....	28
Table 2-7. Existing non-reference sharpness measures	29
Table 3-1. Properties of good objective measures	36
Table 3-2. Recommended linear combination coefficients for CQM in evaluating the quality of images with the same source. The coefficients are obtained by applying LME on 48 training images suffering from the corresponding distortions	38
Table 3-3. Performance comparisons of qualities of images with the same source. The CQM has greater Pearson, Spearman, and Kendall correlations, and smaller RMSE and MAE for these Gaussian blurred, JPEG2000 compressed and denoised images in the TID2008 dataset.	40

Table 3-4. Comparison of measure metrics for the four JPEG2000 compressed images in Figure 3-5.....	42
Table 3-5. Comparison of measure metrics for the four JPEG2000 compressed images in Figure 3-6.....	43
Table 3-6. Recommended linear combination coefficients for CQM in evaluating independent image qualities. The coefficients are obtained by applying MLR on images suffering from mixed distortions.....	44
Table 3-7. Performance comparisons of the CQM for independent images. The CQM has greater Pearson, Spearman, and Kendall correlation coefficients, and smaller RMSE and MAE values for mixed distorted images in the TID2008 dataset	49
Table 3-8. Comparison of measure metrics for the nine enhancement results in Figure 3-11.....	53
Table 3-9. Comparisons of the training procedures for image processing and video processing applications	55
Table 4-1. Evaluation of various edge detectors for images in Figure 4-12.....	72
Table 4-2. Evaluation of various edge detectors for images in Figure 4-13.....	73
Table 4-3. Evaluation of various edge detectors for images in Figure 4-14.....	74
Table 4-4. Evaluation of various edge detectors for images in Figure 4-15.....	75
Table 4-5. Evaluation of various edge detectors for images in Figure 4-16.....	76
Table 4-6. Evaluation of various edge detectors for images in Figure 4-17.....	77
Table 4-7. Evaluation of the qualities of edge maps in Figure 4-18.....	79

Table 5-1. Basic steps in the Yitzhaky and Peli's NR edge measure	85
Table 6-1. Properties of the quadratic filter.....	100
Table 6-2. Statistics of the MOS for the Canny, Sobel, and AWQF algorithm for 20 test images. The AWQF has the greatest average score and smallest variance according to human observers.....	128
Table 7-1. Average execution time for retrieving the first 48 most similar color patches. The average retrieval time using the distance based measures increases greatly when the size of the candidate database increases. By comparison, the time complexity of the measure based retrieval method increases reasonably.	133

LIST OF FIGURES

Figure	Page
Figure 1-1. Distortions occur during image acquisition, transmission, and processing. Evaluating the qualities of the images is an important task.....	2
Figure 1-2. Non-ideal edge detection outputs. Double edges, round corners, and missing edge pixels may occur.....	3
Figure 1-3. Overview of the dissertation contributions. (Blue blocks: image processing algorithms; Orange blocks: image measures; Green blocks: applications).....	6
Figure 1-4. Illustrations of the practical applications of objective measures.....	8
Figure 2-1. A good objective contrast measure should be able to select the mammogram image with the best contrast. (a) an original mammogram image; (b) - (d) three enhanced images using the CLAHE [32], MSR [33], and 3DAWQF [8] aims to enhance the overall contrast. Radiologists use the CLAHE enhancement results for assisting diagnosis.....	11
Figure 2-2. Demonstration of opponent afterimages. Fixate upon the black spot in the center of the four colored squares in (a) for about 30 seconds then move your gaze to fixate the black spot in the uniform white area in (b). The afterimage will show colors in (c).....	13
Figure 2-3. Comparisons of color components of different color models and different parameters. From left to right: (a) RGB; (b) HSV; (c) abL in CIELab; (d) UXVL with $\alpha = \beta = \gamma = 1/3$; (e) UXVL with $\alpha = 0.299, \beta = 0.587, \gamma = 0.114$	14
Figure 2-4. Diagram of the colorfulness measure.....	17
Figure 2-5. Illustrative examples of the colorfulness measure design. (a) original image; color enhanced images using (b) WP; (c) GW; (d) GWWP; (e) ACE with original gray level preserved; (f)	

ACE without original gray level preserved. Images courtesy of Chambah et al [48]. The HS_M1 measure denotes the image in (e) has the best color quality without considering the color casting. By comparison, the UICM considers the color balance property and denotes the image in (f) is the best.	18
Figure 2-6. Synthetic image results on the sensitivity of colorfulness metrics. The synthetic images on the right column have more vivid colors.	20
Figure 2-7. Natural image results on the sensitivity of colorfulness metrics. (a)-(d) are obtained by linearly reducing the chrominance component in CIELab space.	21
Figure 2-8. (a) Yellow (RGB: 255-255-0); (b) RGB: 255-255-55 (c) RGB: 255-200-0. It is seen that (b) and (c) have the same mean value in the RGB color space, but they look quite different. Also, (b) and (c) have the same Euclidean distance to the (a), but visually (c) looks more similar to (a). For these synthetic image patches, traditional colorfulness measures fail.	22
Figure 2-9. Sensitivity of the color tone measure on synthetic image patches. The UCD can effectively distinguish the colors.	23
Figure 2-10. Buchsbaum curve: perceived contrast as a function of background illumination [50]. According to the background illumination level, the contrast sensitivity is divided to four regions: Dark region, Devries-Rose region, Weber region, and Saturation region.	24
Figure 2-11. RGB color subcube structure.	27
Figure 2-12. (a)-(d) are four images with different levels of contrast distortions. These four images are obtained from the TID2008 database.	28
Figure 2-13. Flow diagram of the new sharpness measure.	31
Figure 2-14. Image results on the sensitivity of sharpness metrics. (a)-(d) are four images with different levels of sharpness distortions. These images are obtained from the TID2008 database.	32

Figure 3-1. Four JPEG compressed images and their MOS scores. An objective quality measure that generates such measure values similar to the human vision automatically is desired.	34
Figure 3-2. Illustration of using grayscale image measures for evaluating color image quality. (a) Converting a color image to a grayscale image and then applying the grayscale measures; (b) Applying the grayscale measures in each color component and then fuse all measure values together.	35
Figure 3-3. Illustrative examples. Image (a) has more vivid color. The colors of the sky, the lawn, and the lighthouse are more saturated. Image (b) has better contrast. The details of the trees and the car are more discernable.	36
Figure 3-4. Measure metrics and MOS for 48 denoised images. The measure values obtained from the MOS and the CQM have four distinguished levels, while the CIQI measure value has a greater variance as shown in the red rectangles.	41
Figure 3-5. Four levels of JPEG2000 compressed images. These images are obtained from the TID2008 dataset.	42
Figure 3-6. Four levels of JPEG2000 compressed images.	43
Figure 3-7. Flow of the training procedure for images with different sources.	44
Figure 3-8. Sample video frames. These video frames have different contents. Frame (c) is contaminated by white Gaussian noise. Frame (g) is contaminated added motion blurring. Frame (h) is blurred by the player’s movement. Frame (j) is downsampled and then upsampled. Frame (n) is desaturated.	46
Figure 3-9. Zoom of the artificially distorted images.	49
Figure 3-10. Using CQM for assisting selecting optimal alpha power for an underilluminated image.	51

Figure 3-11. Nine enhanced images for the ‘Shoe’ image. The original image is taken in strong sunlight resulting in the significant loss of color and details. Nine image enhancement algorithms are applied on this image. Their performances need to be evaluated. Refer to the references listed in the brackets for details of the enhancement algorithms.53

Figure 4-1. Block diagram of the new RBEM evaluation system. It is composed of three edge attribute measures.....61

Figure 4-2. (a) Original image, (b) the ground truth edge map, (c) the LoG edge map, and (d) the Roberts edge map. The edge pixel detection/localization measure indicates the Roberts edge map is better than the LoG edge map.63

Figure 4-3. Basic templates used for corner detection.....64

Figure 4-4. Examples of using two basic corner templates to formulate ideal corners: (a) ‘X’ shape, (b) ‘T’ shape65

Figure 4-5. (a) Original image, (b) the ground truth edge map, (c) the Sobel edge map and (d) the LoG edge map. The corner detection/localization measure indicates the Sobel edge map is better than the LoG edge map.66

Figure 4-6. Downsampled images for the edge map in Figure 4-10 (a). (a)-(f): downsampling with y_1 to y_6 .67

Figure 4-7. An illustrative example of using box functions to expose double edge occurrence in 90 degree. (a) an local 4x4 region (b) the y_3 downsampled 2x2 region, (c) the corresponding box function in that 2x2 region. The edge pattern on the left has single pixel width, and the edge pattern on the right has double edges. Their box functions are different.67

Figure 4-8. An illustrative example of using box functions to expose double edge occurrence in 45 degree. (a) an local 4x4 region (b) the y_3 downsampled 2x2 region, (c) the corresponding box function in that 2x2 region. The edge pattern on the left has single pixel width, and the edge pattern on the right has double edges. Their box functions are different.68

Figure 4-9. An illustrative example of using box functions to expose double edge occurrence in 67.5 degree. (a) an local 4x4 region (b) the y3 downsampled 2x2 region, (c) the corresponding box function in that 2x2 region. The edge pattern on the left has single pixel width, and the edge pattern on the right has double edges. Their box functions are different.68

Figure 4-10. (a) Ground truth, (b) the Roberts edge map, (c) the LOG edge map, (d) the ground truth box function map, (e) the Roberts edge box function map, and (f) the LOG edge box function map. The Roberts output is more similar to the ground truth.....69

Figure 4-11. (a) Original image, (b) the ground truth edge map, (c) the Canny edge map, and (d) the Roberts edge map. The double edge occurrence measure indicates the Roberts edge map is better than the Canny edge map.69

Figure 4-12. Edge detection results on a synthetic image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.72

Figure 4-13. Edge detection results on a synthetic image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.73

Figure 4-14. Edge detection results on a synthetic image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.74

Figure 4-15. Edge maps for natural penguin image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.75

Figure 4-16. Edge maps for natural vase image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.76

Figure 4-17. Edge maps for natural man image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.77

Figure 4-18. (a) Original image with Gaussian noise added, (b) ground truth edge map, edge maps obtained using the Canny edge detector with (c) $\sigma = 2$, (d) $\sigma = 2.5$, (e) $\sigma = 3$, (f) $\sigma = 3.5$, (g) $\sigma = 4$, (h) $\sigma = 4.5$, (i) RBEM measure plot indicates a maximum at $\sigma = 3$79

Figure 4-19. (a) Original natural image, (b) ground truth, edge maps obtained using Canny with high level threshold (c) $Th = 0.1$, (d) $Th = 0.2$, (e) $Th = 0.3$, (f) $Th = 0.4$, (g) $Th = 0.5$, (h) $Th = 0.6$, (i) $Th = 0.7$, (j) RBEM measure plot indicates a maximum at $Th = 0.3$81

Figure 5-1. Non-reference edge measure: which edge map is the best without consulting medical experts' opinion? The presented measure NREM indicates the second edge map is the best.83

Figure 5-2. Flow and the intermediate results of the new non-reference edge measure NREM.86

Figure 5-3. Reconstruction results using different interpolation methods. (a) Original CT kidney image. (b)-(f) reconstructed results from: (b) weighted average, (c) weighted median, (d) central weighted median, (e) PDE [127], (f) alpha trimmed weighted average.89

Figure 5-4. Using measures to select optimal edge detection algorithms. (a) Original Lena image. (b)-(f) Edge detection results (right column) and the reconstructed images (middle column) from: (b) Canny, (c) Sobel, (d) Roberts, (e) LoG, (f) Prewitt edge detectors.94

Figure 5-5. Applying edge measures in assisting selecting optimal operating parameters. (a) Original X-ray chest image. (b)-(i) Edge detection results using the Sobel edge detector with the threshold T ranging from 0.01 to 0.08. (j) presented NREM measure plot indicating $T = 0.03$ as the optimal parameter value. (k) measure plot of using MGSSIM alone. (l) The Yitzhaky's method indicating $T = 0.06$ as the optimal parameter value.95

Figure 6-1. Demonstration of a 3x3 subwindow and the kernel patterns.99

Figure 6-2. Uses of color contrast enhancement. Color contrast enhancement is widely used in many applications such as consumer and biomedical applications. It is also used to correct histograms for high dynamic range images as well as low contrast images.102

Figure 6-3. Flow of the new 3DAWQF algorithm.104

Figure 6-4. Intermediate results for each of the 3DAWQF steps. (a) Original image; (b) Intensity adjusted image; (c) Intra-plane AWQF contrast enhanced image; (d) Inter-plane contrast color enhanced image; (e) Fused image. The original image is obtained from the NASA Retinex research website [149]. It can be seen that the contrasts and colors are significantly enhanced after step (c) and (d), and the final fused images achieve an overall enhancement.105

Figure 6-5. Comparisons of different intensity adjustment results using different modifications of Naka-Rushton function: (a) Adaptive Naka-Rushton in [145]. (b) Symmetric Naka-Rushton in [146]. (c) Proposed modified Naka-Rushton. Red circles denote the problem areas resulting from the corresponding Naka-Rushton function.107

Figure 6-6. Comparisons of different parameters β and γ in Eq. (6.16). The first row is the modified Naka-Rushton curve with different parameters. The second row is a representative input image which is suitable for using the corresponding parameters. The third row is the intensity adjusted image with these parameters. (a) $\beta = 3, \gamma = 1$; (b) $\beta = 1, \gamma = 1$; (c) $\beta = 1, \gamma = -1$108

Figure 6-7. Comparison of the effect of alpha power in filter design. (a) original image; (b) enhanced image without the alpha power; (c) enhanced image with the alpha power. The white stripe is clearer in the shirt image in (c). The yellow and gray color patches are more saturated in the color board image in (c). Accordingly, the CRME and MEMEE assign greater values for images in column (c).109

Figure 6-8. (a) Original image. (b) contrast enhanced image. It is seen that the PAWQF enhanced image has sharper edges compared to the original image. The details on the trees, the windows, and the boat are more discernable.110

Figure 6-9. Four pseudo inter planes.111

Figure 6-10. (a) Original image. (b) color enhanced image. It is seen that the color enhanced image has more vivid color compared to the original image.112

Figure 6-11. Comparisons of fusion results of intra plane contrast enhancement image and inter plane color image. Results are obtained using fuse toolbox [152]. According to the CQM, the PCA fusion is the best, and the FSD fusion is the worst among these fusion results.113

Figure 6-12. The intermediate results after each step in the 3DAWQF. (a) Original image; (b) Intensity adjusted image; (c) Contrast enhanced image; (d) Color enhanced image; (e) Final overall enhanced image.....114

Figure 6-13. Experimental results for applying different color image enhancement algorithms. (a) Original image; (b) Enhanced image by Wang’s algorithm; (c) Enhanced image by the Retinex algorithm; (d) Enhanced image by the proposed 3DAWQF algorithm. The CRME and the CQM indicate the performances of each algorithm and demonstrates that the 3DAWQF outperforms the other color contrast enhancement algorithms.....115

Figure 6-14. Experimental results for applying different color image enhancement algorithms on different applications: (a) Original image; (b) Enhanced image by Wang’s algorithm; (c) Enhanced image by the Retinex algorithm; (d) Enhanced image by the proposed 3DAWQF algorithm. The CRME and the CQM indicate the performances of each algorithm and demonstrates that the 3DAWQF outperforms the other color contrast enhancement algorithms.116

Figure 6-15. (a) An AirAsia plane; (b) a wreckage of the plane tail found underwater; (c) – (e) color enhancement results using the 3DAWQF with parameter M equals to 12, 20, 28, 36, respectively.118

Figure 6-16. UICM measure values for variant parameter M from 12 to 40 in the 3DAWQF filter design.118

Figure 6-17. (a) An asphalt concrete image; traditional edge detection algorithm, such as (b) the Canny algorithm and (c) the Sobel algorithm, are not able to detect the edges effectively; (d) the presented AWQF based edge detection algorithm is able to detect the crack.119

Figure 6-18. Diagram and intermediate results of the new AWQF algorithm for edge detection120

Figure 6-19. Edge detection results for a synthetic image. (a) Original image; (b) Enhanced grayscale image; Edge obtained by (c) the AWQF algorithm; (d) the Canny algorithm; (e) the Sobel algorithm. 124

Figure 6-20. Edge detection results for a medical image. (a) Original image; (b) Enhanced grayscale image; Edge obtained by (c) the AWQF algorithm; (d) the Canny algorithm; (e) the Sobel algorithm. 125

Figure 6-21. Zoom of ROI (Red Rectangle part) of the cell image in Figure 6-20. (a) ROI of the cell image; (b)-(d) are edges obtained by: (b) the AWQF algorithm; (c) the Canny algorithm; (d) the Sobel algorithm; 125

Figure 6-22. Edge detection results for a natural statue image. (a) Original image; (b) Enhanced Gray image; Edge obtained by (c) the AWQF algorithm; (d) the Canny algorithm; (e) the Sobel algorithm. 126

Figure 6-23. Edge detection results for a label image. (a) Original image; (b) Enhanced Gray image; Edge obtained by (c) the AWQF algorithm; (d) the Canny algorithm; (e) the Sobel algorithm. 127

Figure 6-24. Zoom of ROI (Red Rectangle part) of the bottle label image in Figure 6-23. (a) ROI of the bottle label; (b)-(d) are edges obtained by: (b) the AWQF algorithm; (c) the Canny algorithm; (d) the Sobel algorithm; 127

Figure 6-25. (a) Mean Opinion Scores(MOS) and (b) NREM measure values for the Canny, Sobel, AWQF1 and AWQF2 algorithms for the 20 testing images. Red: Canny, Green: Sobel; Blue: AWQF..... 128

Figure 7-1. (a) Yellow (RGB: 255-255-0); (b) RGB: 255-255-55 (c) RGB: 255-200-0. It is seen that (b) and (c) have the same mean value in the RGB color space, but they look quite different. Also, (b) and (c) have the same Euclidean distance to the (a), but visually (c) looks more similar to (a). 131

Figure 7-2. First 48 retrieved color patches for yellow (255-255-0). (a) using Euclidean distance on RGB color space. (b) using Euclidean distance on LUXV color space; (c) Using color measure. The left column shows the retrieval results from 4193 color patches, and the right column from 132651 color patches..... 132

Figure 7-3. Color matching for the painting industry133

Figure 7-4. An illustration of the database search application by comparing the image attributes measures.134

Figure 7-5. Using the color measure for color image retrieval. (a) Query image; (b)-(f) first five retrieved images. The measure based retrieval returns flowers with similar colors.135

Figure 7-6. CQM measure values for images belongs to different classes in Wang database. CQM measure values for Elephants (red), Buses (green) and Dinosaurs (blue)135

Figure 7-7. Using color measure for dinosaurs and elephants image retrieval. (a) Query image; (b) first retrieved image; (c) Second retrieved image136

Figure 7-8. Electrical system inspection using colorfulness measures on electrical thermography. When the electrical components are overheated, the thermal image is warmer. So the color measure value is greater than the one measured under normal condition. On the other hand, when the electrical components are cooling down, the thermal image is cooler. In this case, the color measure value is decreased.137

Figure 7-9. (a) Original skin cancer microscope image; Enhanced image by (b) Wang’s algorithm; (c) MSRCR algorithm; (d) 3DAWQF algorithm. The color measure CQM indicates the performances of each algorithm and the CM₃ generate the same ranking as the MOS.....138

Figure 7-10. (a) Original breast cancer mammogram image with low contrast. Enhanced images by (b) Wang’s algorithm; (c) MSRCR algorithm; (d) 3DAWQF algorithm. The CQM indicate the performances of each algorithm and the CQM generate the same ranking as the MOS.139

Figure 7-11. Using measures to select optimal edge detection algorithms. (a) Original CT abdomen image. (b)-(f) Edge detection results from: (b) Canny, (c) Sobel, (d) Roberts, (e) LoG, (f) Prewitt. 140

Figure 7-12. An illustrative banana freshness classification system using colorfulness measures without interaction with human judgement.141

Figure 7-13. CQM measure values for 35 training banana images. Image 1 to 12 are unripen banana images, the CQM measure values are smaller than 1.6; image 13 to 24 are ripen banana images, the CQM measure values are between 1.6 and 2.5; image 24 to 35 are overripen banana images, the CQM measure values are greater than 2.5.....142

Figure 7-14. Three testing banana images. The attribute measures are shown under each image. The CQM measure values for the unripen, ripen, and over ripen bananas falls in different ranges.143

Figure 7-15. (a) original image; Color enhanced images using (b) WP; (c) GW; (d) GWWP; (e) ACE with original gray level preserved; (f) ACE without original gray level preserved. Images courtesy of Chambah et al [48].....144

Figure 8-1. Summary of the contributions147

LIST OF PUBLICATIONS

Patents

[1] Karen Panetta, Sos Agaian, and Chen Gao, "Systems and Methods for Image and Video Signal Measurement," Invention disclosure form submitted to Tufts University. December 2013.

Journal Papers

[2] Karen Panetta, Chen Gao, and Sos Agaian, "No reference color image contrast and quality measures," IEEE Transactions on Consumer Electronics, vol. 59, pp. 643-651, August, 2013.

[3] Karen Panetta, Chen Gao, Sos Agaian, and Shahan Nercessian, "Nonreference Medical Image Edge Map Measure," International Journal of Biomedical Imaging, vol. 2014, 2014.

[4] Karen Panetta, Chen Gao, Sos Agaian, and Shahan Nercessian, "A New Reference Based Edge Map Measure," submitted to IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2014.

[5] Karen Panetta, Chen Gao, and Sos Agaian, "Human Visual System Inspired Underwater Image Quality Measures," submitted to IEEE Journals of Oceanic Engineering, 2015.

Conference Papers

[6] Chen Gao, Karen Panetta, and Sos Agaian, "New edge detection algorithms using alpha weighted quadratic filter," in IEEE International Conference on Systems, Man, and Cybernetics 2011, pp. 3167-3172.

[7] Chen Gao, Karen Panetta, and Sos Agaian, "A new color contrast enhancement algorithm for robotic applications," in IEEE International Conference on Technologies for Practical Robot Applications, 2012, pp. 42-47.

[8] Chen Gao, Karen Panetta, and Sos Agaian, "Three dimensional alpha weighted quadratic filter based image color contrast enhancement," in SPIE Defense, Security, and Sensing, 2013, pp. 875514-875514-12.

[9] Chen Gao, Karen Panetta, and Sos Agaian, "No reference color image quality measures," in IEEE International Conference on Cybernetics, 2013, pp. 243-248.

[10] Chen Gao, Karen Panetta, and Sos Agaian, "Color image attribute and quality measurements," in SPIE Mobile Multimedia/Image Processing, Security, and Applications, 2014, pp. 91200T-91200T-14.

- [11] Chen Gao, Karen Panetta, and Sos Aгаian, "Color image retrieval and analysis," accepted at Mobile, Multimedia, Image Processing, Security and Applications, Baltimore, 2015.
- [12] Chen Gao, Karen Panetta, and Sos Aгаian, "Robust template based corner detection algorithms for robotic vision," accepted at IEEE International Conference on Technologies for Practical Robot Applications, 2015.
- [13] Chen Gao, Karen Panetta, and Sos Aгаian, "A combined colorfulness and color tone measure and its applications," submitted to IEEE International Conference on Image Processing, 2015.
- [14] Chen Gao, Karen Panetta, and Sos Aгаian, "Coefficients training methods for linear combinations of image attribute measures," submitted to the IEEE International Conference on Imaging Systems and Techniques, Macau, China, 2015.

**Image and Video Component Quality Measurements,
Analyses, and Applications**

Chapter 1. Introduction

1.1. Overview

Digital images usually suffer from a wide variety of distortions during acquisition, transmission, storage, and reproduction [15]. Some distortions are illustrated in Figure 1-1. Such distortions may result in a visual degradation, and make the follow up processing such as pattern recognition and target tracking more difficult. There have been developed many image processing algorithms aimed to improve the visual quality of an image. Therefore, evaluating their performances and selecting the optimal algorithm for particular applications reliably and efficiently is of paramount importance.

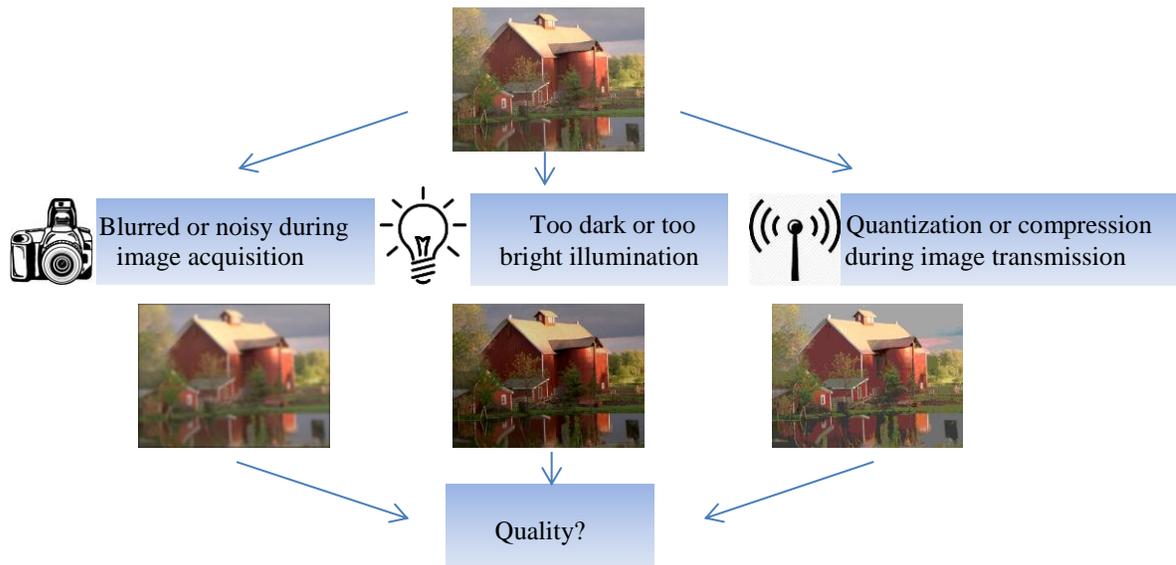


Figure 1-1. Distortions occur during image acquisition, transmission, and processing. Evaluating the qualities of the images is an important task.

Human eyes are highly sensitive to image structures. However, edge maps obtained from edge detectors may provide false information, such as double edges, round corners, missed or extra edge pixels. The false information may decrease the performance of the follow up image processing tasks, such as segmentation, object recognition, detection and tracking, which utilize information on the edge location. Therefore, a reliable measure that can effectively evaluate the edge map quality is of paramount importance.

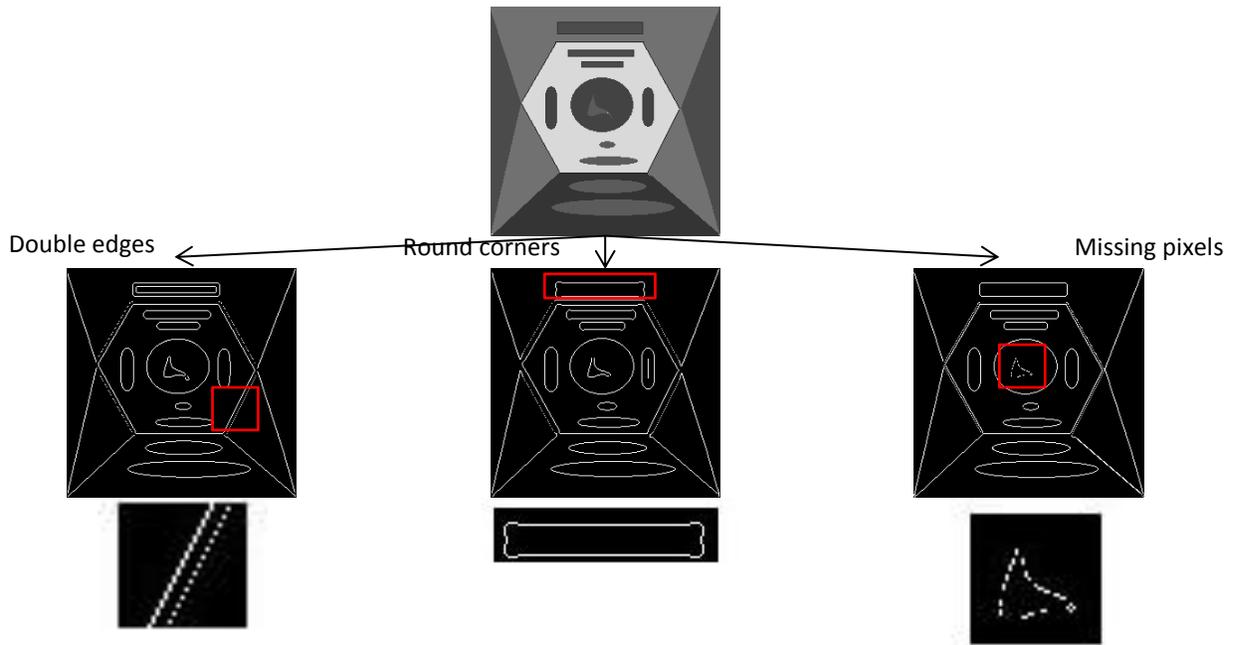


Figure 1-2. Non-ideal edge detection outputs. Double edges, round corners, and missing edge pixels may occur.

Human inspection is still the most acceptable way to evaluate image or edge qualities. However, it is impossible to remove human judgment bias, so the evaluation could be misleading. The Mean Opinion Score (MOS) suggested by the IUT-T [16] increases the subjective evaluation reliability by involving a large number of participants in the evaluation procedure, but it's impossible to remove all the bias. More importantly, the subjective evaluation is expensive with respect to time and resources, and it certainly cannot be automated for image processing and computer vision systems. In many industrial tasks, using human inspection in all quality control stages is not affordable or practical. With the ongoing development of new image processing algorithms, there should be reliable and efficient objective measures that perform similar to the human visual perceptions. All these limitations make the development of automated objective evaluations of great necessity.

In general, objective evaluations can be classified into two types: the full-reference (FR) approaches [15][17][18][19][20][21], and the non-reference (NR) approaches [22][23][24][25][26][27][28], depending on whether the reference images are available where the term 'reference' indicates the images which are considered 'true' images. For example, the real paintings in the art identification applications and the query images in the searching applications are used as the 'reference' images. Accordingly, the reference-based measures evaluate the image 'fidelity' (the similarity between a testing image and the reference image), and the non-reference measures

evaluate the ‘quality’ (such as the vividness of an image color, the degree of blurring, and the noise level) of an image.

FR approaches require an ideal image as the reference and they determine the similarity between the resultant image and the reference image. This type of approach is widely used in image database search applications such as fingerprint and face retrieval where the query image is used as the reference image. However, in many practical cases, the reference images may not always be available or accessible. For example, there is no knowledge about what the real outer space image look like. In such situations that the reference images are not available, a NR measure is necessary. NR approaches, also called “blind” approaches, attempt “blindly” to make a judgment of image quality without any prior knowledge of the reference images. This type of approach uses information only from the resultant images.

Color image enhancement and edge detection are two major tasks in many image processing applications. Therefore, it is desired to have reliable and consistent objective measures that can evaluate the image enhancement and edge detection algorithms’ performances according to the human perception. Much effort has been made in recent years to develop objective measures that correspond with the human perception. Unfortunately, only limited success has been achieved [29]. Many existing color image quality measures are based on modifications of grayscale image quality measures. Such methods either transform the color image to a grayscale image and measure the grayscale image quality, or select a color space and linearly combine the measure values on each color component. However, these simple extensions of grayscale measures are not always applicable for measuring color image qualities. For edge quality evaluation purposes, most existing edge measures generate a quantitative value for each edge map under test by utilizing the distances between the testing edge maps and the ground truth edge maps. These methods do not have strong correlations with the human perception. For example, if an edge map is shifted several pixels, it has a great distance-based edge measure value but it looks very similar to the original edge map. Besides, the distance-based measures only account for some important criteria of an edge map indirectly, such as the corner location and double edge occurrence. There are many limitations in existing color image quality measures and edge measures, so it is desirable to have new measures that have strong correlation with the human visual perception.

Before quantifying the overall image or edge quality, it is intuitive to focus on one aspect of the image or edge attributes first. One reason is that it is easy for human beings to interpret the quantitative values just focused on one image characteristic. Also, in many practical applications, only certain image attributes are affected by certain

distortions or image processing tasks. Thirdly, combining image attribute measures can provide a comprehensive evaluation for the measuring the overall image or edge quality. Since objective measures which perform similar to the human vision are desired, the Human Visual System (HVS) properties should be taken into account in the design of the objective image and edge attribute measures. For example, one fundamental characteristic of the Human Visual System (HVS) is its sensitivity to edge structures. Accordingly, edge sharpness is usually evaluated as an attribute measure. In addition, the human sensation of contrast is affected by the background illumination intensity and the neighboring pixel activity in the local region. Therefore, the luminance masking and contrast masking should be considered in the contrast measure. Such HVS properties should be incorporated in the objective measures so the measures are more reliable in the computer vision systems.

So far, human inspection is still widely involved in many practical applications. Ideally, a good measure is able to perform the evaluations similar as a human observer, but with low cost of time and resources. For example, in the food industry, computer vision technology is possible to be used as an inexpensive alternative to using manual experts in grading the food product freshness or quality. If the color, size, shape, and/or texture of the food products can be measured in the controlled conditions, and the measured values are distinguishable, these measure values can be used to classify the quality or freshness of the product [30]. The uses of the measures in painting and cosmetic color retrieval, electrical system inspection, and underwater environment exploration and search are also promising. These examples also motivate the research and implementation of the new objective measures.

1.2. Objectives

The goals of this dissertation are to develop objective image and edge attribute and quality measurements and tools for analysis, and use them in practical applications. Specifically, the tasks in this dissertation include:

(1) To develop non-reference image and video quality attribute measures, which utilize the human visual system (HVS) properties, including image colorfulness measure, color tone measure, sharpness measure, and contrast measure.

(2) To present several full-reference image edge measures, including edge presence/localization measure, corner presence/localization measure, and double edge occurrence measure.

(3) To present a novel system for evaluating the overall image and edge quality by linearly combining attribute measures. For specific applications, different subsets of attribute measures are selected and different linear combination parameters are used.

(4) To use the objective measures to benchmark image processing algorithm performances and to assist the selection of optimal operating parameters used in image processing algorithms. Accordingly, new Alpha Weighted Quadratic Filter (AWQF) based algorithms are to be developed for color contrast enhancement and edge detection applications.

(5) To examine the uses of the measures in biomedical, industrial, and security applications.

1.3. Dissertation Organization

The overview of this dissertation is shown in Figure 1-3.

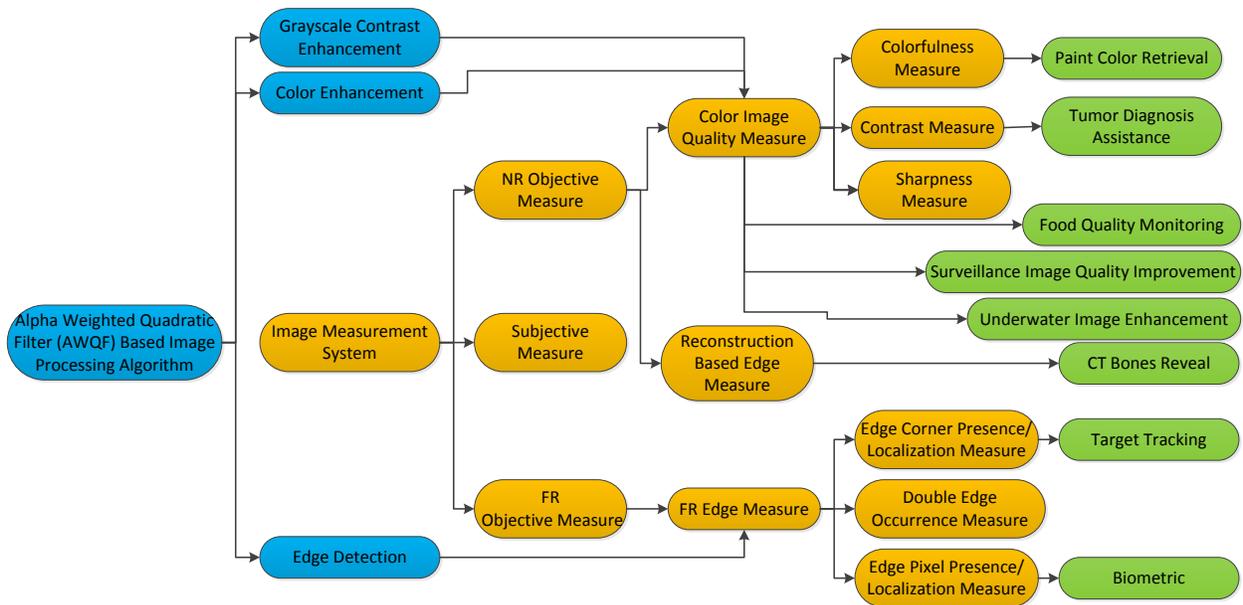


Figure 1-3. Overview of the dissertation contributions. (Blue blocks: image processing algorithms; Orange blocks: image measures; Green blocks: applications)

Chapter 2 presents a set of image and video attribute measures, namely, image colorfulness, color tone, sharpness, and contrast measures. Each measure is used to evaluate one aspect of image characteristics. In Chapter 3 these image attribute measures are combined together and formed the overall quality measures for images and

videos. Methods of training the linear combination coefficients for images with the same source and video frames with different contents are included in this chapter.

Chapter 4 presents a FR edge measure system, which measures edge presence/localization, corner presence/localization, and double edge occurrences. To show the effectiveness of the measure, experiments are conducted on both synthetic and natural images. Chapter 5 introduces a NR edge measure system which is based on comparing the similarity between a reconstructed image and the original image.

A color contrast enhancement algorithm based on the positive alpha weighted quadratic filter, and an edge detection algorithm based on the negative alpha weighted quadratic filter, are introduced in Chapter 6. The measures presented in previous chapters are used to evaluate the algorithms' performances and direct the selection of optimal operating parameters. The practical applications of the measures on fast database retrieval, medical diagnosis assistance, underwater image quality evaluation, electrical system condition inspection, and food quality evaluation are demonstrated in Chapter 7.

Chapter 8 draws the conclusions and discusses the future research directions.

1.4. Dissertation Contributions

The contributions of this dissertation are as follows:

- A new NR image and video quality measure system is presented: this image and video quality measurement system is composed of fusions of image and video attribute measures. Three image attribute measures, namely, the color measure, the sharpness measure, and the contrast measure are presented. Each attribute measure is focused on one aspect of image quality, and the formulations of attribute measures incorporate the Human Visual System (HVS) properties. This novel approach is flexible to include variant number of image attribute measures so it is applicable for many practical applications.
- For edge quality evaluation purpose, a FR measure and a NR measure are presented. The FR edge measure is composed of edge presence/localization measure, double edge occurrence measure, and edge corner presence measure, while the NR edge quality measure is based on the similarity between reconstructed image and the original image. The presented FR edge measure outperforms traditional distance based FR edge measures as it

considers more important visual features in its evaluation. The NR edge measure has better correlation with the human visual perception compared with existing NR edge measures.

- A set of new image processing algorithms which utilize the property of non-linear filters are presented. The Positive Alpha Weighted Quadratic Filter (PAWQF) is applied within each color component and across color component to enhance image structures and colors. The Negative Alpha Weighted Quadratic Filters (NAWQF) is developed for edge detection purposes. To optimize the design of the AWQF, the new measures are used to benchmark image processing algorithm performances and assist selecting optimal operating parameters.
- The new measure systems have many practical applications. For specific applications, different sets of image attribute measures and parameters are used. Practical uses on painting color and cosmetic color search, thermal and mechanical system condition monitoring, underwater image quality evaluation, and food quality evaluation are demonstrated.

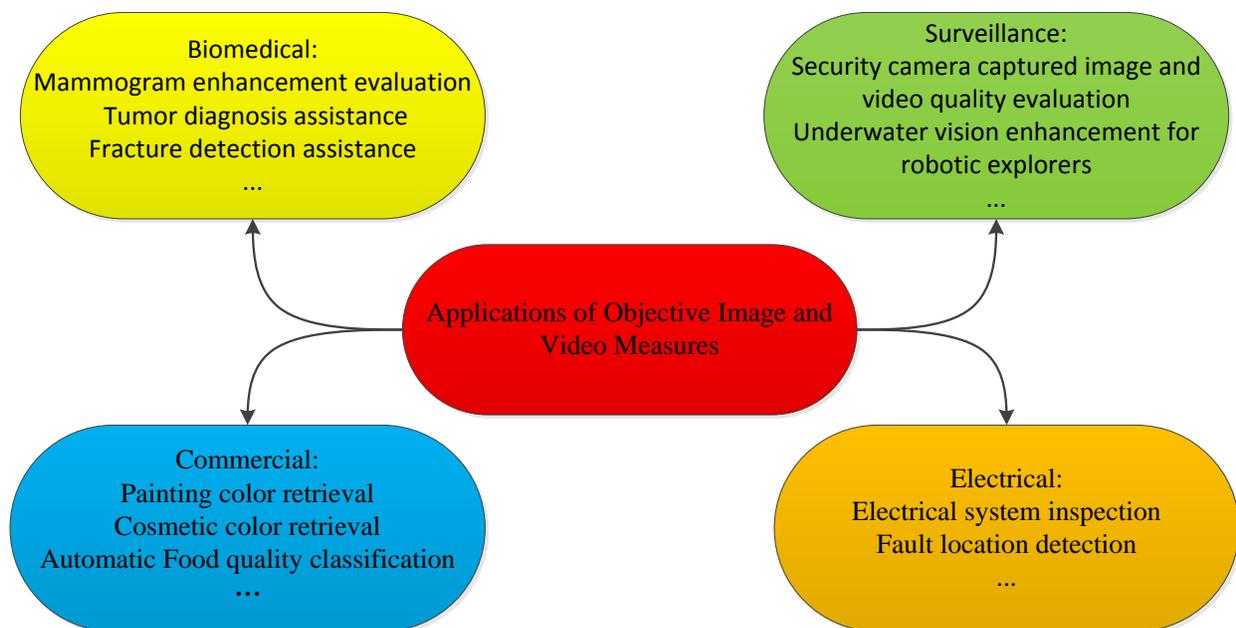


Figure 1-4. Illustrations of the practical applications of objective measures.

Chapter 2. Non-Reference Image and Video Attribute Measures

In this chapter, color image attribute measures, namely, the colorfulness measure, the color tone measure, the sharpness measure, and the contrast measure are presented. Each of these attribute measures evaluates one aspect of image quality, and is inspired by the properties of the Human Visual System (HVS).

2.1. Overview

The human eyes have the capability to distinguish, recognize and interpret different image features [31]. However, it is still a challenge for computer programs. With the development of computer vision, it is necessary to develop a reliable objective measure that evaluates image properties similar to human vision. The human visual system usually perceives images as a layout of homogeneous objects and regions with respect to visual features, such as color, edge, and texture [31]. Therefore, it is necessary to measure the quality of these image attributes. Besides, in many applications, only a certain attribute of an image draws a viewer's attention. For example, in mammogram images, the contrast (for example, the bright white spots for micro-calcifications) between the cancer tissues and the normal tissues helps radiologists diagnose the breast cancer. The mammogram images are captured under the controlled lab environment, so some image attributes are either not related with the diagnosis (such as the colorfulness) or not altered (such as the lightning condition). A typical acquired mammogram image as shown in Figure 2-1(a) usually has low contrast, due to limited exposure time to the radiation. To improve the quality of mammogram images and make the tumors more discernable for diagnosis, many contrast enhancement algorithms are applied at the post-acquisition stage. Figure 2-1 (b) - (d) demonstrate three enhancement results, and experienced radiologists choose the CLAHE [32] enhanced mammogram image to assist the breast cancer diagnosis. In this scenario, an objective measure that can evaluate image contrast in accordance with the human visual perception is desired.

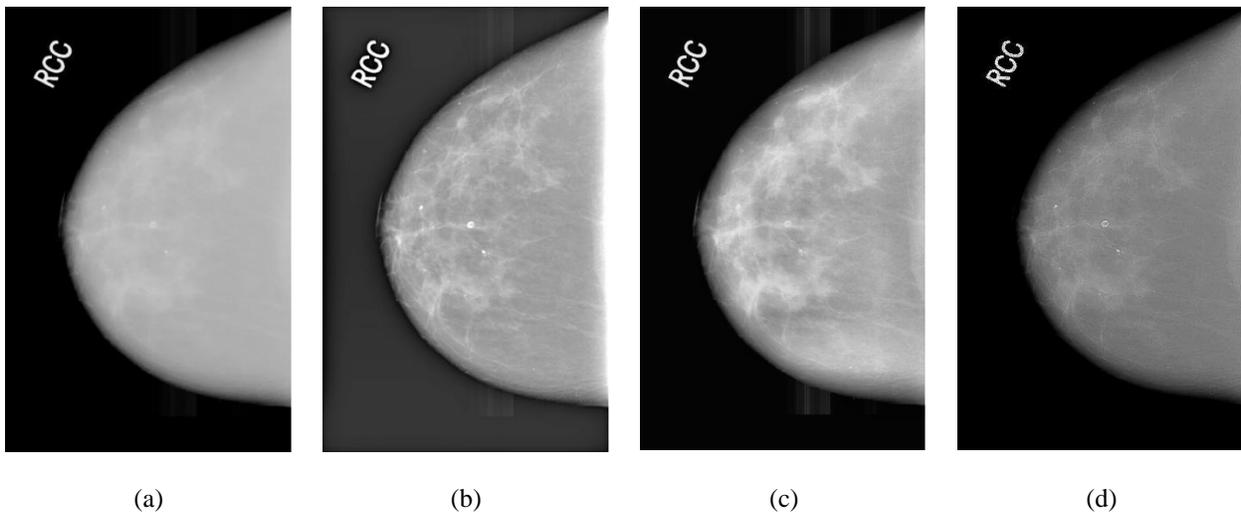


Figure 2-1. A good objective contrast measure should be able to select the mammogram image with the best contrast. (a) an original mammogram image; (b) - (d) three enhanced images using the CLAHE [32], MSR [33], and 3DAWQF [8] aims to enhance the overall contrast. Radiologists use the CLAHE enhancement results for assisting diagnosis.

This chapter focuses on the development of the non-reference image attribute measures, for which each measure evaluates one property of the image quality, and assigns an objective value for a single image without comparing it against other images. In this chapter, the color measure, sharpness measure, and the contrast measure are presented. These attribute measures could be used separately (in Chapter 2), or further combined together to evaluate the overall quality of images or video frames (In Chapter 3).

2.2. Color Measure

The non-reference color measures evaluate the colorfulness and the color tone of an image. The colorfulness is the attribute of chrominance information human perceive. It describes the vividness of an image chrominance. The color tone measures the dominant color in the image. Most existing colorfulness measures are based on the statistical values in variant color models, and the previous color tone measures are based on the distances in specific color models. Therefore, in this section, the color models are discussed first, followed by the NR colorfulness and color tone measures.

2.2.1 Color Models

When quantifying the chrominance of an image, the first task is to determine which color model to use. Generally speaking, a color model is a coordinate system to facilitate the specification of colors in some standard and thus color can be represent as a single point in this system [34]. There have been many color models and these color models interpret the color components in different ways. This section briefly reviews the color models so that it provides a general guide for selecting appropriate color spaces for measuring colors.

The RGB and CMY spaces are the basic models for display and printing. However, it is not very intuitive to get the color of mixing 30% red, 50% green and 20% blue. The human eyes capture colors in a nonlinear model and sense colors according to color properties such as brightness, hue, and saturation. The corresponding color models that closely resemble the qualities most apparent to human perception of color includes HSI (Hue, Saturation,

Intensity), HSV (Hue, Saturation, Value), HCI (Hue, Chrominance, Intensity) and TSD (Hue, Saturation, Darkness). Their advantage lies in the extremely intuitive manner of specifying color. It is easy to select a desired hue and to then modify it slightly by adjustment of the associated saturation and intensity [35]. The advantage of this type of color model leverages the separation of color and luminance and makes applying the grayscale image processing methods to color images possible and also make the individual processing of chromatic and achromatic component possible.

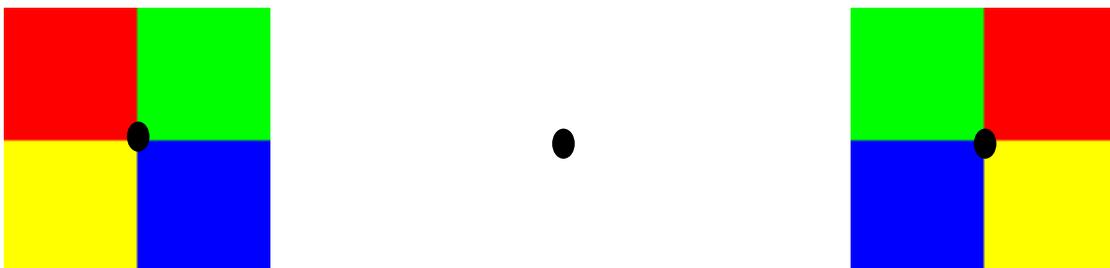
Another type of color model is generally used in the image and video transmission systems. The transmission standard is aimed to reduce the redundancy and separate luminance and chrominance. These color models include YIQ and YUV, which are analogue color models for NTSC and PAL systems respectively, as well as the YCbCr model which is a digital standard.

To better visualize colors in human visual expression, the international organization CIE (International Commission on Illumination) defines a system that classifies color according to the human visual system [36]. A set of physiological experiments have been conducted to allow observers to translate different physical responses to a universal human vision responses. Under this system of color measure, the CIE defined a series of new color models include CIELuv, CIELab, CIELhs and CIELhc [36][37][38].

Hering's opponent color is proposed according to the mechanism of the human vision. It is observed that certain hues are never perceived to occur together such as reddish-green or yellowish-blue. Figure 2-2 demonstrates the opponent natural of the afterimages [39]. According to Hering's opponent color theory, three visual quantities are opponent and paired: yellow-blue (YB), red-green (RG), and white-black (WB). The two members of each pair are opponent, in terms of the opposite nature of the physiological processes and in terms of the mutually exclusive sensory qualities [40].

$$RG = R - G \tag{2.1}$$

$$YB = \frac{R + G}{2} - B \tag{2.2}$$



(a)

(b)

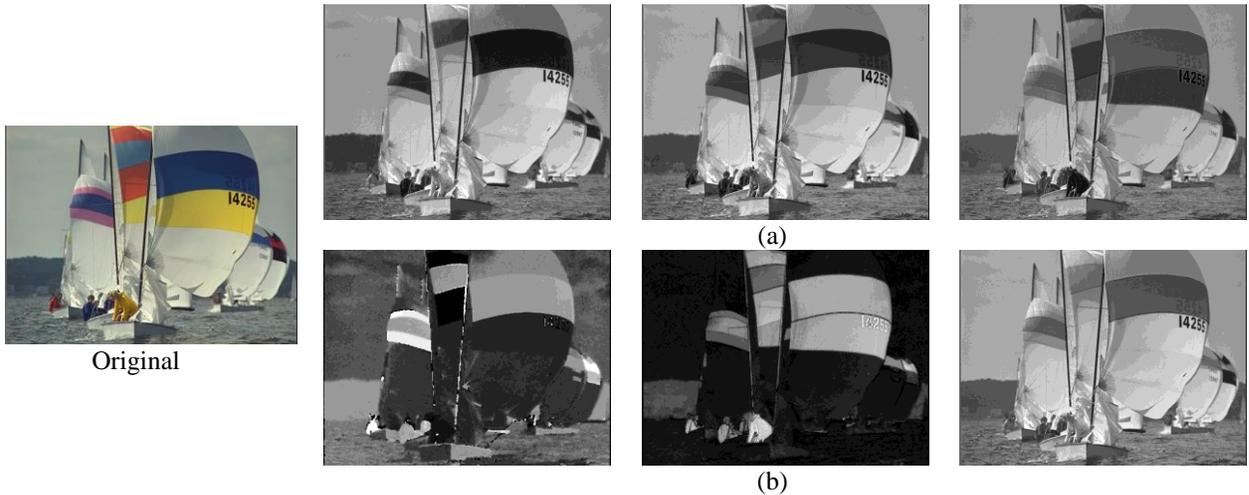
(c)

Figure 2-2. Demonstration of opponent afterimages. Fixate upon the black spot in the center of the four colored squares in (a) for about 30 seconds then move your gaze to fixate the black spot in the uniform white area in (b). The afterimage will show colors in (c).

The traditional LUX color space was proven to effectively separate chrominance [41], and it extends the application of grayscale Logarithmic Image Processing (LIP) model [42] to the color model. In the LUXV color space, an extra component V is added to incorporate all three color components R, B, and G [10]. The definition of the LUXV color plane is as Eq. (2.3):

$$\begin{aligned}
 U &= \begin{cases} 2^{M-1} \left(\frac{R+1}{L} \right) & \text{if } R < L \\ 2^M - 2^{M-1} \left(\frac{L}{R+1} \right) & \text{Otherwise} \end{cases} \\
 X &= \begin{cases} 2^{M-1} \left(\frac{B+1}{L} \right) & \text{if } B < L \\ 2^M - 2^{M-1} \left(\frac{L}{B+1} \right) & \text{Otherwise} \end{cases} \\
 V &= \begin{cases} 2^{M-1} \left(\frac{G+1}{L} \right) & \text{if } G < L \\ 2^M - 2^{M-1} \left(\frac{L}{G+1} \right) & \text{Otherwise} \end{cases} \\
 L &= (R + 1)^\alpha (G + 1)^\beta (B + 1)^\gamma
 \end{aligned} \tag{2.3}$$

where M is the number of bits used to represent an image, and α , β , and γ are the power coefficients weighting each color spectrum, and they satisfy $\alpha + \beta + \gamma = 1$. Under this concept, the UXV triplet can be seen as the ratio of RGB to the luminance. It is in accordance with the human visual property that human vision is less sensitive to image details over the brighter background.



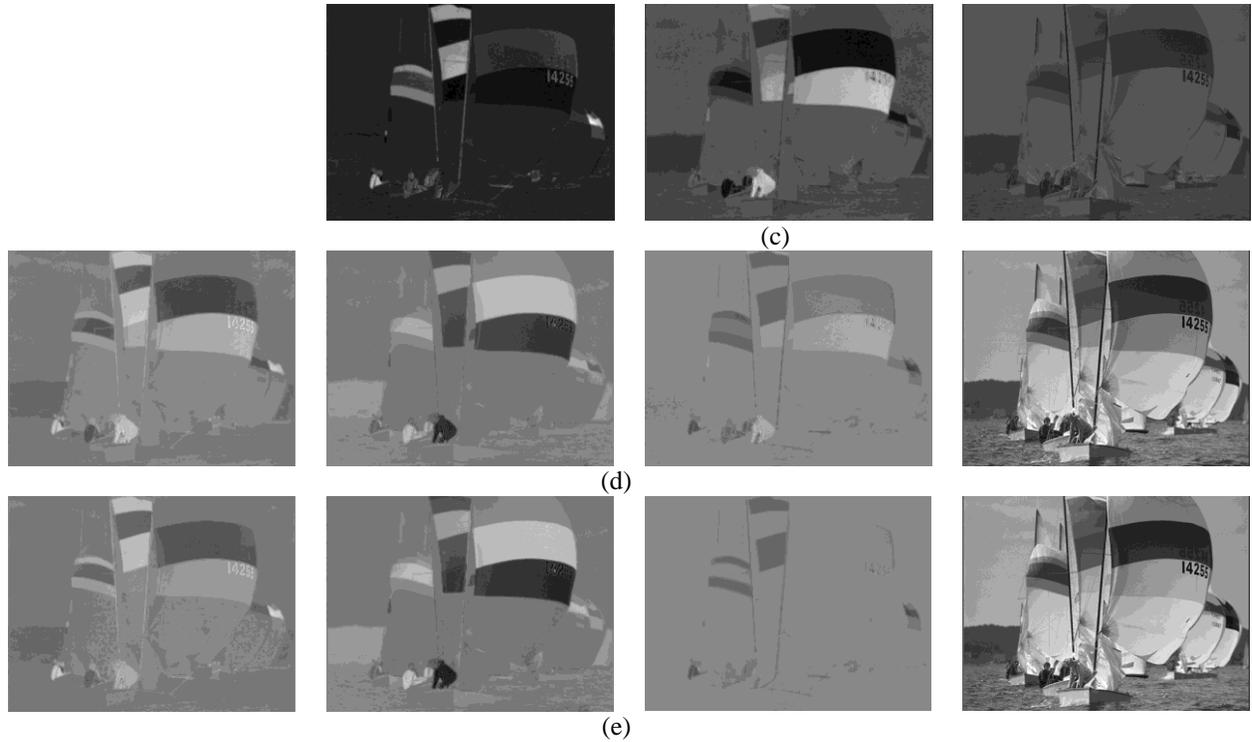


Figure 2-3. Comparisons of color components of different color models and different parameters. From left to right: (a) RGB; (b) HSV; (c) abL in CIE Lab; (d) UXVL with $\alpha = \beta = \gamma = 1/3$; (e) UXVL with $\alpha = 0.299, \beta = 0.587, \gamma = 0.114$.

Figure 2-3 shows color components in different color spaces. As seen, the RGB spaces are strongly correlated. In the HSV spaces, details such as the people on the sail are indistinguishable with the background in the chrominance components H and S. By comparison, the CIE Lab and the UVXL color space perform better at segmenting regions and maintaining details in the chrominance components. Experiments on the selection of parameters α , β , and γ in Eq. (2.3) are also conducted. Two examples, parameters $\alpha = \beta = \gamma = \frac{1}{3}$ and NTSC standard based parameters $\alpha = 0.299, \beta = 0.587, \gamma = 0.114$, are shown in Figure 2-3 (d) and (e). The experiments show parameters set equal to one another have better performance in preserving details in the color components. For example, the strips of different colors on the sail are visible in the V component when equal amount of weights are assigned.

2.2.2 Colorfulness Measure

Colorfulness is the degree of difference between a color and a grayscale image. It describes how vivid and intense the chrominance component is. So far, there is no agreement on quantifying the colorfulness in natural

images. Hasler and Susstrunk used statistical quantities such as color variance and chrominance amplitude of different color planes to evaluate image colorfulness [43]. The efficient representation on two opponent color components are then further proposed and widely adopted in many applications [43][44][45]. These methods are computationally efficient but not in accordance with the human visual perception. The definitions of some existing colorfulness measures are summarized and listed in Table 2-1.

Table 2-1. Existing non-reference colorfulness measures

Measure	Definition
Winkler [46] ^a	$\mu_C + \sigma_C$
HS_M ₁ [43] ^b	$\sigma_{ab} + 0.37\mu_{ab}$
HS_M ₂ [43] ^c	$\sigma_{ab} + 0.94\mu_C$
M_{C_CIQI} [44] [43] ^d	$\frac{\sqrt{\sigma_{rg}^2 + \sigma_{yb}^2} + 0.3\sqrt{\mu_{rg}^2 + \mu_{yb}^2}}{85.59}$

^a σ_C and μ_C : sum of mean and standard deviation of chroma values over the entire video sequence.

^{b,c} σ_{ab} and μ_{ab} are the trigonometric length of the standard deviation and mean in ab space of the CIELab color plane; μ_C is the mean of chroma. The constant coefficients 0.37 and 0.94 are obtained by training the linear models with the MOS.

^d rg and yb are two opponent color components defined as $rg = R-G$ and $yb = (R+G)/2-B$. The constant 0.3 is obtained from training, and 85.59 is the trained average dynamic range.

Instead of using the regular statistical values, the alpha-trimmed statistical values are used for measuring image colorfulness in this dissertation. For an image with size of M by N , the total number of pixels $K = M \times N$, and all pixels of the image are sorted such that $x_1 \leq x_2 \leq \dots \leq x_K$. Let $T_\alpha = \lceil \alpha K \rceil$ (the nearest integer greater than or equal to αK), be the number of the smallest and largest pixel values to be trimmed or discarded from the sorted sequence x_1, x_2, \dots, x_K . The alpha-trimmed mean is hybrid of the mean and median values. It is defined as:

$$\mu_\alpha = \frac{1}{K - 2T_\alpha} \sum_{i=T_\alpha+1}^{K-T_\alpha} X_i \quad (2.4)$$

The basic idea behind the alpha-trimmed mean is for any element of the image, look at its neighborhood, discard the most atypical elements and calculate mean value using the remaining. Therefore, it is more robust to noise than the regular mean value. The parameter alpha determines the number of trimmed elements. As the special cases, the alpha trimmed mean is equal to the regular mean value of the image for $\alpha = 0$ and the median value of the image if α is close to 0.5. Similarly, the second to the fourth order statistical values represent the alpha trimmed variance, skewness, and kurtosis values (Eq. (2.5) – Eq. (2.7)) on the color triplet.

$$\sigma_{\alpha}^2 = \frac{1}{K-2T\alpha} \sum_{i=T\alpha}^{K-T\alpha} (X_i^2 - \mu_{\alpha}^2) \quad (2.5)$$

$$\gamma_{\alpha}^2 = E\left[\left(\frac{X - \mu_{\alpha}}{\sigma_{\alpha}}\right)^2\right] \quad (2.6)$$

$$k_{\alpha}^2 = \frac{E[(X - \mu_{\alpha})^4]}{(E[(X - \mu_{\alpha})^2])^2} \quad (2.7)$$

Each parameter has its mathematical meaning: the mean values represent chrominance intensity and a greater mean value usually implies lighter illumination. The variance σ^2 demonstrates the activity within each color component. Intuitively, a greater variance corresponds to more activities in a color component. The third order skewness γ^2 is a measure of the extent to which the histogram leans to one side of the mean. The fourth order statistic kurtosis k^2 demonstrates the distribution of color components, that is to say, the maximum of a color spectrum. An image with a smaller kurtosis value has a flatter histogram distribution.

These statistical values are combined together to evaluate the image colorfulness. Generally, image colorfulness can be represented as Eq. (2.8), where, ϕ_i represents a function of the i^{th} order statistic, F is a function that combines all the statistical values and a_i is the linear weighting coefficient. Compared with existing colorfulness metrics listed in Table 2-1, the colorfulness measure in Eq. (2.8) incorporates a higher order of statistics which results in more image information. A generic diagram of the colorfulness measure is shown in Figure 2-4.

$$Colorfulness = \sum_{i=1}^N a_i F(\Phi_i) \quad (2.8)$$

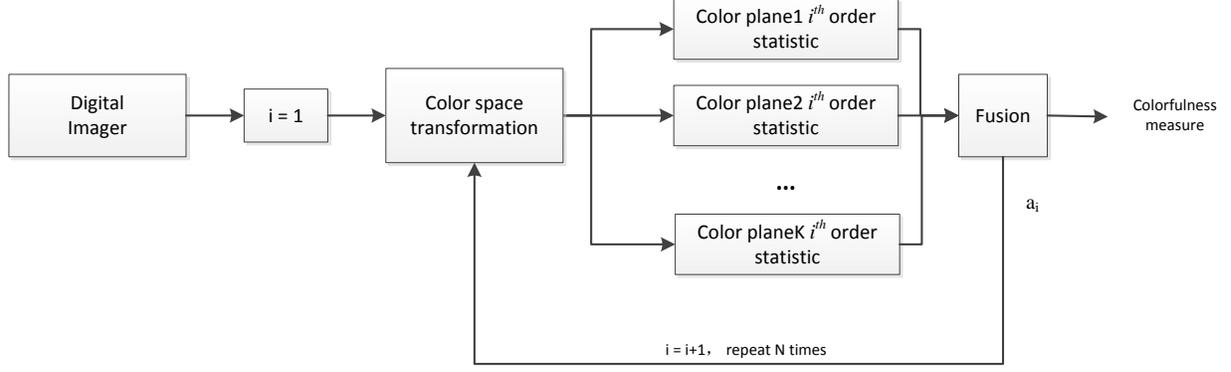


Figure 2-4. Diagram of the colorfulness measure

In the practical applications, users have the flexibility to select the combination function F , the color models, as well as the linear combination coefficients for given applications. Some instantiations are demonstrated in Eq. (2.9) to Eq. (2.12). It is seen that Eq. (2.9) and Eq. (2.10) utilize the ratio of the statistical values and Eq. (2.11) and Eq. (2.12) are formed using linear combination. Generally, the linear combination coefficients can be obtained experimentally, or by using linear regression between the training image statistical values and the subjective evaluation gathered via MOS. For example, the parameters in Eq. (2.9) and Eq. (2.10) are obtained empirically. The linear combination coefficients in Eq. (2.11) and Eq. (2.12) are obtained by using a subset of images with known subjective ratings from the TID2008 image database [47]. It is worth noting that not all four statistical values have to be used. If a certain term is not distinct in terms of a bad correlation or a large p-value for specific type of images, it should not be used in the final colorfulness measure.

$$M_{C1} = 0.02 \times \log\left(\frac{\sigma_{a, RG}^2}{|\mu_{a, RG}|^{0.2}}\right) \times \log\left(\frac{\sigma_{a, YB}^2}{|\mu_{a, YB}|^{0.2}}\right) \quad (2.9)$$

$$M_{C2} = 0.02 \times \frac{\log \sigma_{a, RG}^2 \times \log \sigma_{a, YB}^2}{\log \sigma_{a, c}^2} \times \frac{\log \mu_{a, RG}^2 \times \log \mu_{a, YB}^2}{\log \mu_{a, c}^2} \quad (2.10)$$

$$M_{C3} = \frac{1.33(\sqrt{\mu_{a, U}^2 + \mu_{a, X}^2 + \mu_{a, V}^2} - 200) + 2.39\sqrt{\sigma_{a, U}^2 + \sigma_{a, X}^2 + \sigma_{a, V}^2} - 0.49\sqrt{k_{a, U}^2 + k_{a, X}^2 + k_{a, V}^2}}{200} \quad (2.11)$$

$$M_{C4} = 0.0614\sqrt{\mu_{a, U}^2 + \mu_{a, X}^2 + \mu_{a, V}^2} + 0.1546\sqrt{\sigma_{a, U}^2 + \sigma_{a, X}^2 + \sigma_{a, V}^2} + 0.6642\sqrt{\gamma_{a, U}^2 + \gamma_{a, X}^2 + \gamma_{a, V}^2} - 0.2426\sqrt{k_{a, U}^2 + k_{a, X}^2 + k_{a, V}^2} - 13.0000 \quad (2.12)$$

The selection of the color model is also application specific. An example of designing the colorfulness measure for underwater images is shown below. Figure 2-5 shows an example of designing the colorfulness measure for the underwater images. Underwater images usually suffer from a severe color casting problem due to the absorption and scattering effect of the water media. Some underwater image enhancement results which aim to recover the colors from [48] are shown in Figure 2-5. It is seen that in certain color planes such as the CIELab and HSV, a greater

mean value usually indicate a greater color intensity, and that corresponds to a greater colorfulness measures in Table 2-1. However, the strong green casting which contributes for the great mean value should not be considered good color rendition. It is seen that the HS_M_1 colorfulness measure [43] in Table 2-1 and the M_{C3} measures consider the Fig. 5 (e) is the optimal image without taking the green color cast into consideration. In this case, the RG-YB opponent color space is desired because the smaller mean value in this color component indicates a balanced color without distinct dominant colors. The Euclidean distance to the origin is used as the combination function F and the correlation of each statistical term are tested with the MOS. In the experiment, it is seen that the third order skewness and the fourth order kurtosis terms do not have strong correlations with the MOS (p-values are greater than 0.05 so these two terms are not significant). Therefore, the first two order statistical values are used and the linear combination coefficients are obtained by linear regression. Eventually, the underwater image colorfulness measure UICM is formulated as Eq. (2.13).

$$UICM = -0.0268\sqrt{\mu_{a,RG}^2 + \mu_{a,YB}^2} + 0.1586\sqrt{\sigma_{a,RG}^2 + \sigma_{a,YB}^2} \quad (2.13)$$

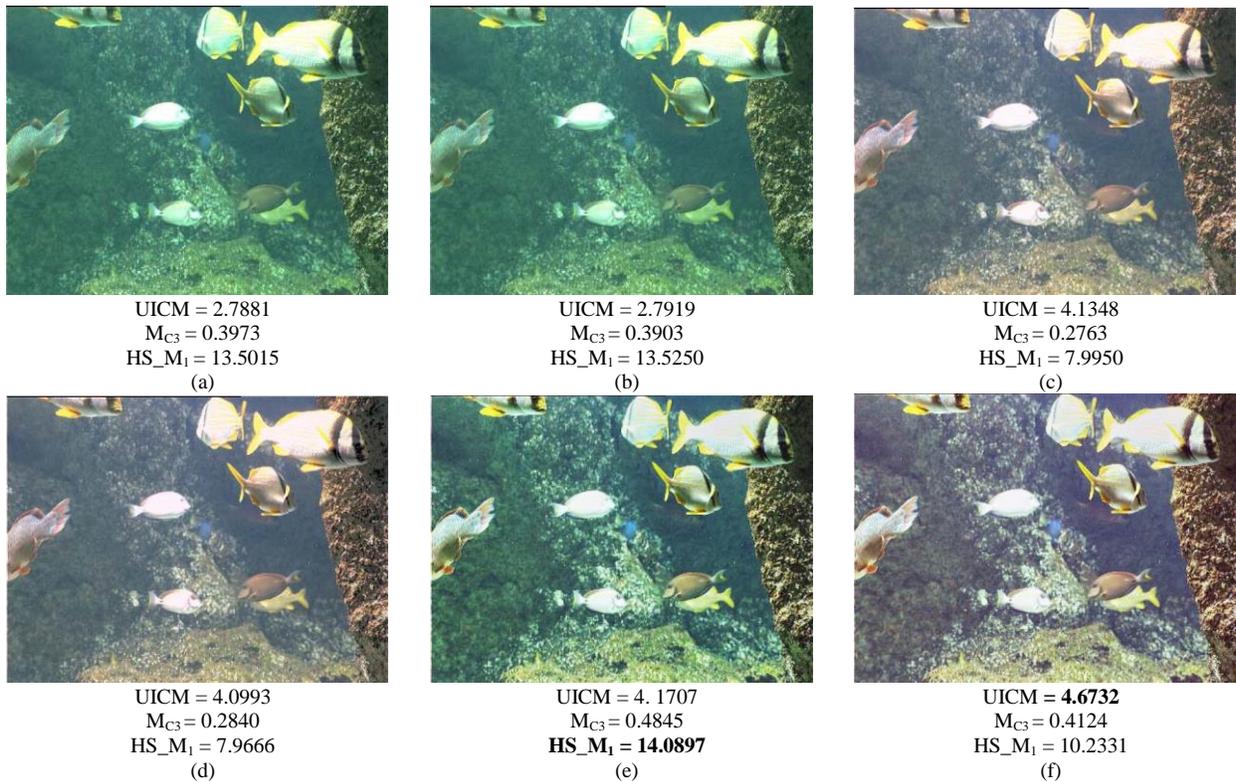
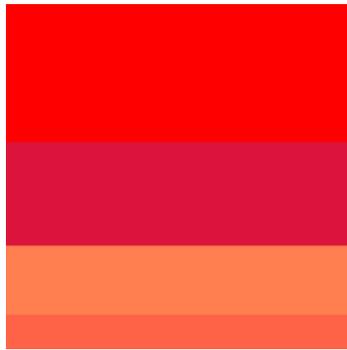


Figure 2-5. Illustrative examples of the colorfulness measure design. (a) original image; color enhanced images using (b) WP; (c) GW; (d) GWWP; (e) ACE with original gray level preserved; (f) ACE without original gray level preserved. Images courtesy of Chambah et al [48]. The HS_M_1 measure denotes the image in (e) has the best color quality without considering the color casting. By comparison, the UICM considers the color balance property and denotes the image in (f) is the best.

It is straightforward to show the colorfulness measure sensitivities on synthetic images as this could exclude the effects of structures and scenes. One example is shown in Figure 2-3, where each image has the same pattern and is comprised of four color strips. The colors are selected from HTML and CSS specification. For example, Figure 2-3 (a) is an image composed of the color ‘Red’, ‘Crimson’, ‘Coral’, and ‘Tomato’, and Figure 2-3 (b) is composed of the color ‘Light Salmon’, ‘Salmon’, ‘Dark Salmon’, and ‘Light Coral’ from top to bottom. It is observed that the colorfulness measure shows a monotonic increasing order as the visually perceived of red, blue, or green increases. The image characteristics for the six synthetic images in Figure 2-3 are shown in Table 2-2. For comparison purpose, the colorfulness measure used in CIQI [44] is also listed with the images.



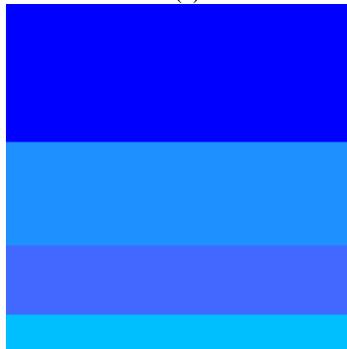
(a)

$M_{C_CIQI} = 1.4614$
 $M_{C1} = 0.7768$
 $M_{C2} = 1.2230$
 $M_{C3} = 1.0102$
 $M_{C4} = 11.3187$



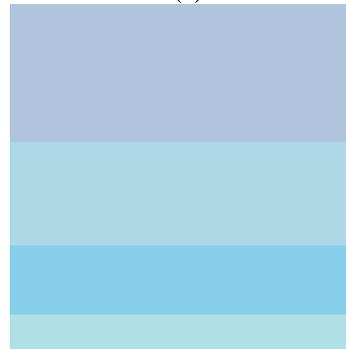
(b)

$M_{C_CIQI} = 0.6519$
 $M_{C1} = 0.3199$
 $M_{C2} = 0.7404$
 $M_{C3} = 0.2680$
 $M_{C4} = 2.0364$



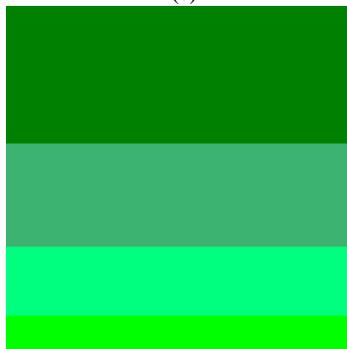
(c)

$M_{C_CIQI} = 1.6435$
 $M_{C1} = 0.9693$
 $M_{C2} = 1.2486$
 $M_{C3} = 1.2916$
 $M_{C4} = 14.5433$



(d)

$M_{C_CIQI} = 0.4624$
 $M_{C1} = 0.4339$
 $M_{C2} = 0.7808$
 $M_{C3} = 0.2734$
 $M_{C4} = 2.4856$



(e)

$M_{C_CIQI} = 1.4369$
 $M_{C1} = 0.9561$
 $M_{C2} = 1.1721$
 $M_{C3} = 1.3925$
 $M_{C4} = 16.1584$



(f)

$M_{C_CIQI} = 1.2132$
 $M_{C1} = 0.4815$
 $M_{C2} = 2.8027$
 $M_{C3} = 0.9415$
 $M_{C4} = 10.6218$

Figure 2-6. Synthetic image results on the sensitivity of colorfulness metrics. The synthetic images on the right column have more vivid colors.

Table 2-2. Image characteristics for six synthetic images in Figure 2-6

	Figure 2-6 - (a)	Figure 2-6- (b)	Figure 2-6- (c)	Figure 2-6 - (d)	Figure 2-6 - (e)	Figure 2-6 - (f)
μ_{RG}	203.2000	102.4000	-61.3000	-39.9000	-92.0000	-163.4000
μ_{YB}	101.8000	76.2000	-202.3500	-40.9500	98	40.4000
σ_{RG}^2	2430.2	221.6455	4109.3	361.2990	64.8016	3609.9
σ_{YB}^2	807.0802	84.8121	1855.7	141.6260	3969.1	1644.6
$\sqrt{\mu_U^2 + \mu_X^2 + \mu_V^2}$	239.2369	226.8383	249.6111	223.7207	249.8975	239.2550
$\sqrt{\sigma_U^2 + \sigma_X^2 + \sigma_V^2}$	63.2569	8.4942	81.1067	10.5832	89.3714	58.1534
$\sqrt{\gamma_U^2 + \gamma_X^2 + \gamma_V^2}$	0.7625	1.4766	0.6421	1.7792	1.0863	2.0657
$\sqrt{k_U^2 + k_X^2 + k_V^2}$	2.7059	4.8864	3.0850	4.4057	2.9830	5.8985

Figure 2-7 shows the sensitivity of the colorfulness metrics for natural images. Figure 2-7 (a) is a reference image from TID2008 database [47], and (b) - (d) are obtained by linearly reducing chrominance in the CIELab color space [49]. The image characteristics for the four images in Figure 2-7 are shown in Table 2-4. These results confirm that the image which is visually more colorful has a greater colorfulness measure value.



(a)



(b)



(c)



(d)

Figure 2-7. Natural image results on the sensitivity of colorfulness metrics. (a)-(d) are obtained by linearly reducing the chrominance component in CIELab space.

Table 2-3. Colorfulness measure values for four images with different colorfulness degradations in Figure 2-7. The greater measure values correspond to better image colorfulness

	Figure 2-7 – (a)	Figure 2-7- (b)	Figure 2-7- (c)	Figure 2-7 – (d)
HS_Ml	82.8105	69.9182	47.7809	30.5428
M_{C_CIQI}	0.9675	0.8169	0.5583	0.3569
M_{C1}	1.0006	0.9053	0.7123	0.5219
M_{C2}	0.9012	0.8060	0.6243	0.4186
M_{C3}	0.9336	0.7636	0.5429	0.3908
M_{C4}	10.4548	8.2998	5.4464	3.4545

Table 2-4. Image characteristics for four Images with different colorfulness degradations in Figure 2-7

	Figure 2-7 – (a)	Figure 2-7- (b)	Figure 2-7- (c)	Figure 2-7 – (d)
μ_{RG}	14.0491	12.0061	8.4765	5.4364
μ_{YB}	45.9446	39.4927	27.3220	17.5256
σ_{RG}^2	1726.1	1084.9	413.9533	148.5312
σ_{YB}^2	2952.1	2225.3	1122.6	478.3702
$\sqrt{\mu_U^2 + \mu_X^2 + \mu_V^2}$	232.3456	228.1730	224.2925	222.6767
$\sqrt{\sigma_U^2 + \sigma_X^2 + \sigma_V^2}$	61.0453	49.0728	32.7757	21.0015
$\sqrt{\gamma_U^2 + \gamma_X^2 + \gamma_V^2}$	1.2676	1.0668	0.9500	0.9320
$\sqrt{k_U^2 + k_X^2 + k_V^2}$	4.4960	4.1437	4.2179	4.4670

2.2.3 Color Tone Measure

With image statistical values, the above formula successfully measures the colorfulness for natural images where variations exist. However, using image statistical values alone is not sufficient to quantify uniform colors such as the synthetic image patches shown in Figure 2-8. For these types of images, the second and higher order statistics are zero or undefined. Simply using the mean value of color patches alone fail to distinguish different colors. For example, the color patches in Figure 2-8 (b) and Figure 2-8 (c) has the same mean value in the RGB color space but clearly they have very different visual appearance. Traditional colorfulness measures in Table 2-1 fail to distinguish these images. To overcome this limitation, a color tone measure should be added.

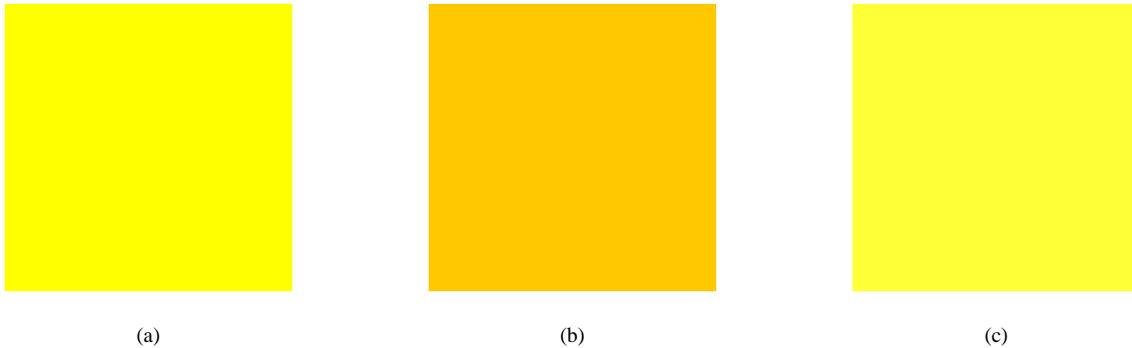


Figure 2-8. (a) Yellow (RGB: 255-255-0); (b) RGB: 255-255-55 (c) RGB: 255-200-0. It is seen that (b) and (c) have the same mean value in the RGB color space, but they look quite different. Also, (b) and (c) have the same Euclidean distance to the (a), but visually (c) looks more similar to (a). For these synthetic image patches, traditional colorfulness measures fail.

Traditionally, color tones are usually described with a color triplet in certain color spaces. However, it is not intuitive for human eyes. For example, the color patch in Figure 2-8 (b) and Figure 2-8 (c) has the same Euclidean distance to the color patch in Figure 2-8 (a), but visually, Figure 2-8 (c) is more similar to Figure 2-8 (a). The color tone measure aims to present a unique index for each color that can distinguish low activity color patches.

Instantiations of the Uni-Color Differentiate (UCD) term are shown in Eq. (2.14) to Eq. (2.17). Figure 2-9 shows the sensitivity of the UCD for synthetic image color patches. It is seen that the UCD term can effectively distinguish different levels of red. Combining the colorfulness metric and color tone metric together, the color measure $M_{C,CQM}$ can effectively measure the color for both synthetic images and natural images. Practically, if the global activity of an image is smaller than a certain threshold, the UCD is incorporated to indicate the overall color tone.

$$ColorTone(r, g, b) = \frac{|\alpha(r - g) + \beta(r - b) + \gamma(g - b)|}{|\alpha(r + g) + \beta(r + b) + \gamma(g + b)|} \quad (2.14)$$

$$ColorTone(r, g, b) = \left| \frac{r + g + b}{3} \right| \quad (2.15)$$

$$ColorTone(r, g, b) = \alpha \left| \frac{(r - g)}{(r + g)} \right| + \beta \left| \frac{(r - b)}{(r + b)} \right| + \gamma \left| \frac{(g - b)}{(g + b)} \right| \quad (2.16)$$

$$UCD = \frac{1}{\#ofnonzeros(ColorTone)} \sum ColorTone \cdot \log(ColorTone) \quad (2.17)$$

Color Patch	R-G-B value/UCD	Color Patch	R-G-B value/UCD	Color Patch	R-G-B value/UCD
	165-42-42/0.3227		220-20-60/0.3024		255-99-71/0.3470
	255-127-80/0.3617		250-128-114/0.3662		255-160-122/0.3598

Figure 2-9. Sensitivity of the color tone measure on synthetic image patches. The UCD can effectively distinguish the colors.

2.3. Contrast Measure

Contrast is defined as the ability to separate image details. The HVS is known to be sensitive to the relative, rather than the absolute changes in background luminance and the local activity. Such properties are called luminance masking (LM), and contrast masking (CM) respectively. To develop a contrast measure which is in accordance with the human visual perception, the HVS properties must be considered.

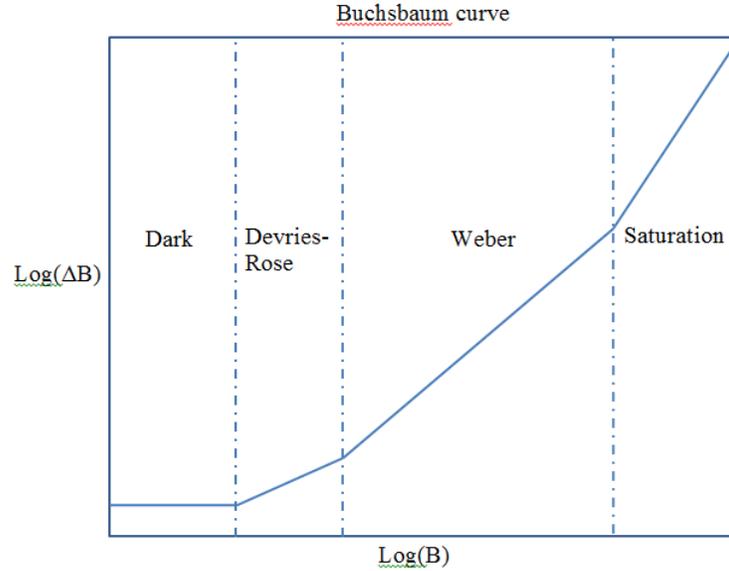


Figure 2-10. Buchsbaum curve: perceived contrast as a function of background illumination [50]. According to the background illumination level, the contrast sensitivity is divided to four regions: Dark region, Devries-Rose region, Weber region, and Saturation region.

Figure 2-10 shows the piecewise Buchsbaum curve in which the contrast sensitivity is divided into four regions according to the background intensity [51]. It is seen from Figure 2-10 that the noticeable relative luminance differences are background illumination level dependent. Generally, the ratio of the absolute difference or the gradient changes to the background intensity determines the just noticeable difference (JND) for the human eyes as shown in Eq. (2.18). The parameter γ in Eq. (2.18) indicates different sensitivities to the background at different regions. γ is usually set to 0 at the Dark region, implying that the perceived contrast is independent of the background intensity. γ is usually set as 0.5, 1, and 2 at the Devries-Rose region, Weber region, and Saturation region, respectively. That is to say, in a brighter background, a greater difference between the ROI intensity and the background intensity is required for the human eyes sense the same amount of luminance change than that in a dark background.

$$JND = \frac{\Delta B}{B^\gamma} \quad (2.18)$$

Considering that the HVS operates in the Weber region over a great range of background intensities, many existing contrast measures are either based on the Weber contrast [52][53] or the Michelson contrast [22][23] which assign $\gamma = 1$ in Eq. (2.18). These contrast measures are formulated as a function of relative changes over the

background intensity and differ in the ways of formulating the relative luminance difference and average background luminance. Some typical measures are summarized in Table 2-5.

Table 2-5. Existing grayscale non-reference contrast measures

Measure Definition	Features
$EME_{k_1 k_2} = \frac{1}{k_1 k_2} \sum_{k=1}^{k_1} \sum_{l=1}^{k_2} 20 \ln \left(\frac{I_{\max,k,l}}{I_{\min,k,l}} \right)$	Weber contrast based enhancement measure Applicable to uniform background images [52]
$EMEE_{k_1 k_2} = \frac{1}{k_1 k_2} \sum_{k=1}^{k_1} \sum_{l=1}^{k_2} \alpha \left(\frac{I_{\max,k,l}}{I_{\min,k,l}} \right)^\alpha \ln \left(\frac{I_{\max,k,l}}{I_{\min,k,l}} \right)$	Weber contrast based enhancement measure Incorporate entropy in EME [52] More randomness in image, the bigger α to choose
$Visibility = \sum_{k=1}^{k_1} \sum_{l=1}^{k_2} \frac{I_{\max,k,l} - I_{\min,k,l}}{I_{\max,k,l} + I_{\min,k,l}}$	Two dimensional Michelson contrast based measure [53] A bounded measure
$AME_{k_1 k_2} = -\frac{1}{k_1 k_2} \sum_{k=1}^{k_1} \sum_{l=1}^{k_2} 20 \ln \left(\frac{I_{\max,k,l} - I_{\min,k,l}}{I_{\max,k,l} + I_{\min,k,l}} \right)$	Michelson contrast based enhancement measure Applicable to non-uniform images/ Applicable to Periodic background [22]
$AME_{k_1 k_2}^E = -\frac{1}{k_1 k_2} \sum_{k=1}^{k_1} \sum_{l=1}^{k_2} \alpha \left(\frac{I_{\max,k,l} - I_{\min,k,l}}{I_{\max,k,l} + I_{\min,k,l}} \right)^\alpha \ln \left(\frac{I_{\max,k,l} - I_{\min,k,l}}{I_{\max,k,l} + I_{\min,k,l}} \right)$	Michelson contrast based enhancement measure Incorporate entropy in AME [22]
$\log AME_{k_1 k_2} = \frac{1}{k_1 k_2} \otimes \sum_{k=1}^{k_1} \sum_{l=1}^{k_2} \frac{1}{20} \ln \left(\frac{I_{\max,k,l} \tilde{\Theta} I_{\min,k,l}}{I_{\max,k,l} \oplus I_{\min,k,l}} \right)$	PLIP version of AME/ More appealing to human vision [23][28]
$\log AMEE_{k_1 k_2} = \frac{1}{k_1 k_2} \otimes \sum_{k=1}^{k_1} \sum_{l=1}^{k_2} \alpha \left(\frac{I_{\max,k,l} \tilde{\Theta} I_{\min,k,l}}{I_{\max,k,l} \oplus I_{\min,k,l}} \right)^\alpha \ln \left(\frac{I_{\max,k,l} \tilde{\Theta} I_{\min,k,l}}{I_{\max,k,l} \oplus I_{\min,k,l}} \right)$	PLIP version of AMEE/ More appealing to human vision [23][28]
$SDME_{k_1 k_2} = -\frac{1}{k_1 k_2} \sum_{k=1}^{k_1} \sum_{l=1}^{k_2} 20 \ln \left \frac{I_{\max,k,l} - 2I_{\text{center},k,l} + I_{\min,k,l}}{I_{\max,k,l} + 2I_{\text{center},k,l} + I_{\min,k,l}} \right $	Second derivative based measure, It is not sensitive to noise [25]
$Visibility_{HVS} = \frac{\sum_{i=1}^n a_i \tilde{\otimes} G_i \left(\frac{F_i(MF_{i1})}{F_i(MF_{i2})} \right)^{n_i} \tilde{\otimes} b_i \tilde{\otimes} G_i \left(\frac{F_i(MF_{i3})}{F_i(MF_{i4})} \right)^{n_i}}{\sum_{j=1}^n c_j \tilde{\otimes} L_j \left(\frac{F_j(MF_{j1})}{F_j(MF_{j2})} \right)^{m_j} \tilde{\otimes} d_j \tilde{\otimes} L_j \left(\frac{F_j(MF_{j3})}{F_j(MF_{j4})} \right)^{m_j}}$	A general visibility model [54]

One limitation of the existing contrast measures in Table 2-5 is that these measures are not robust to noise. Take the EMEE equation shown in Table 2-5 as an example. The $I_{\max,k,l}$ and $I_{\min,k,l}$ represent the maximum and minimum values inside a sub-block respectively. This calculation has its own limitations because image statistical features are spatially non-stationary, and image distortion is also space variant. So the maximum and minimum values may be taken on the noise pixels. Leveraging on the observation, a modified *EMEE* measure, named *MEMEE*, is presented as shown in Eq. (2.19). In the *MEMEE* measure, all values within each block are sorted, and

then they are separated into two groups based on a suitable threshold. Suppose that the sorted pixel values are $X_1 \leq X_2 \leq \dots \leq X_t \leq \dots \leq X_K$ where K is the total number of pixels in a block, and t stands for the threshold notation, the $I_{max,k,l}$ and $I_{min,k,l}$ can be redefined as Eq. (2.20). Under these definitions, the $I_{max,k,l}$ indicates the local maximum portion, while the $I_{min,k,l}$ represents the local minimum portion, and the contrast measure *MEMEE* is expressed as Eq. (2.19).

$$MEMEE_{contrast} = \frac{1}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \frac{I_{max,k,l}}{I_{min,k,l}} \log\left(\frac{I_{max,k,l}}{I_{min,k,l}}\right) \quad (2.19)$$

$$I_{min,k,l} = \sum_{i=1}^t X_i, I_{max,k,l} = \sum_{i=t+1}^K X_i \quad (2.20)$$

Another limitation of the existing contrast measures is that they are based on the Weber or Michelson contrast and these two contrast definitions are only limited to specific types of images. Usually, the Weber contrast is used to measure the local contrast of a single target seen against a uniform background, while the Michelson contrast is commonly used to measure the contrast of periodic pattern [55]. Therefore, the Weber, or Michelson contrast based measures such as EME and AME are applicable for specific image patterns only [56]. However, in complex images these uniformity or periodicity conditions are not always true. So a more general method based on the Root Mean Square (RMS) contrast which does not depend on spatial frequency content or distribution of contrast in the image is presented in this section.

Regular RMS contrast is a pixel-wised contrast as defined in Eq. (2.21). The new Root Mean Enhancement measure (RME) incorporates the idea of RMS contrast and the Human Visual System (HVS) properties. It measures the relative RMS contrast in the log domain. The RME measure is expressed in Eq. (2.22). The image is divided into

$k_1 k_2$ blocks, $I_{i,j}$ is the center pixel intensity in block (i, j) , $\frac{I_1 + I_2 + \dots + I_n}{n}$ is the average intensity in block i, j and

n is the total number of pixel within each block. It is worth noting that $\log\left|I_{i,j} - \frac{I_{c1} + I_{c2} + \dots + I_{cn}}{n}\right|$ returns a negative

value if the difference between the center and its neighbors is smaller than 1. Practically, $\max\left(1, \left|I_{i,j} - \frac{I_{c1} + I_{c2} + \dots + I_{cn}}{n}\right|\right)$

is used to make sure the RME is always positive thus is proportional to the difference. Compared with existing Weber or Michelson based contrast measures, the RME is applicable to more types of images.

$$C_{RMS} = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (I_{i,j} - \bar{I})^2} \quad (2.21)$$

$$RME = \frac{1}{k_1 k_2} \sqrt{\sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \left| \frac{\log \left| I_{i,j} - \frac{I_1 + I_2 + \dots + I_n}{n} \right|}{\log \left| I_{i,j} + \frac{I_1 + I_2 + \dots + I_n}{n} \right|} \right|^2} \quad (2.22)$$

When the center pixel has an intensity value approximately equal to the average value of the current block, the contrast is small. Thus, the RME measure is small for low contrast images and vice versa for images with high contrast. The recommended block size is 3x3 because image statistical features are spatially non-stationary and image distortion is also space variant.

A color image is a multi-dimensional signal, and therefore it is natural to expand the traditional formulation of grayscale contrast measures so that it can be used to measure the color image contrast. Some existing methods apply grayscale measures on color images by converting the color image into a grayscale image [15], or measuring the quality in each color component individually and then combining the measure values by different weights. However, it is difficult to decide which optimal color model to use and the color to grayscale conversion is a lossy procedure. Therefore, it is improper to measure color image quality using grayscale measures. To be more specific, the contrast measures can be performed not only within each color plane but also across color planes. The cross-plane contrast models the differences between color planes and reveals the color variation and structural differences raised from different color components. Based on this idea, the Color/Cube RME (CRME) is proposed to measure the relative difference of the color cube center and all the neighbors in the current color cube.

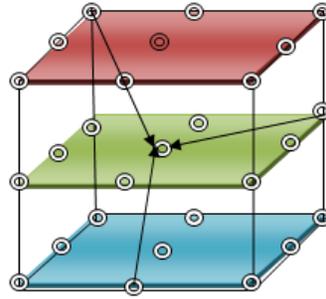


Figure 2-11. RGB color subcube structure.

$$CRME = \frac{1000}{k_1 k_2} \sqrt{\sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \left(\frac{\log \left| I_{i,j} - \sum_{c=1}^3 \lambda_c \frac{I_{c1} + I_{c2} + \dots + I_{cn}}{n} \right|}{\log \left| I_{i,j} + \sum_{c=1}^3 \lambda_c \frac{I_{c1} + I_{c2} + \dots + I_{cn}}{n} \right|} \right)^\alpha} \quad (2.23)$$

In the implementation, the intensity value of the center pixel is used as the cube center pixel intensity. λ_c represents the weights for different color components. In the experimental results shown in this dissertation,

$\lambda_r = 0.299$, $\lambda_g = 0.587$, and $\lambda_b = 0.114$ are used in accordance with the NTSC standard. According to the Human Visual System (HVS) properties, human eyes are not equally sensitive to the Just Noticeable Difference (JND). To compensate for this, an α power is assigned to the ratio based on the region to which the local background of the current cube belongs. If it belongs to the DeVries-Rose region, $\alpha = 0.2$; if it belongs to the Weber region, $\alpha = 0.4$; and if the background intensity falls into the Saturation region, $\alpha = 0.8$.



Figure 2-12. (a)-(d) are four images with different levels of contrast distortions. These four images are obtained from the TID2008 database.

Table 2-6. Contrast measure values for the four images in Figure 2-12.

Measure	Figure 2-12 (a)	Figure 2-12 (b)	Figure 2-12 (c)	Figure 2-12 (d)
EMEE	0.0271	0.0265	0.0096	0.0058
AMEE	0.0645	0.0622	0.0864	0.0977
M_{CON_CIQI}	0.5782	0.4897	0.4803	0.2203
MEMEE	0.7852	0.7111	0.2004	0.1256
CRME	1.0499	1.0422	0.9122	0.7845

Contrast enhancement plays an essential role in image processing. For example, surveillance cameras and aerospace cameras usually capture images with the haze problem. Many haze removal methods have been proposed. But it still difficult to tell which processed image is the best without human expert judgment. With reliable measures, the candidates can be ranked automatically; therefore, the best de-hazed image can be selected automatically. Figure 2-12 shows the sensitivity of some contrast metrics for contrast changed images in the TID2008 database [47]. It is seen that the AMEE does not generate correct rankings because the testing image does not show periodic patterns. The M_{CON_CQI} shows that the images in Figure 2-12 (b) and Figure 2-12 (c) have similar numerical contrast values. However, their human visual perceptions are very different. As a comparison, the MEMEE and CRME correctly rank the contrast of these images.

2.4. Sharpness Measure

Sharpness determines the amount of details an imaging system can produce. It is defined by the boundaries between zones of different colors or intensities. In many applications such as medical image processing, the original images acquired from devices usually suffer from a blurring problem. Thus it is desirable to have a reliable measure to assist the post-acquisition image enhancement procedure. The perception of sharpness is related to the clarity of details and edges of an image. Leveraging on this observation, the spatial domain sharpness measures utilize the information from the edge pixels [44], while the transform domain methods use the high frequency components in the DCT domain [57] or wavelet domain [58][59]. For example, the average of image variance shows to be effective for evaluating sharpness in low resolution images [60], and the relative area of the sharp region can represent the sharpness of a compressed video frame [61]. Spatial domain methods also include those utilizing the edge spread [62], edge width [63], or the difference of estimated edge map and actual distorted edge map [44]. In the transform domain, Shaked and Tastl proposed a measure using the high-pass to band-pass frequency ratio on local features [64]. Caviedes and Gurbus then combined the spatial and frequency domain approaches by applying the 2D kurtosis on the DCT domain [57][65][66]. Selected existing sharpness measures are shown in Table 2-7.

Table 2-7. Existing non-reference sharpness measures

Measure	Definition
Average gradient M_{avg} [60] ^a	$avg(\ \nabla\phi\)$

Relative area [61] ^b	$\frac{1}{MN} \{(m, n) B_1(m, n) > T_H\} $
Q_{NBM} [62] ^c	$\alpha + \beta s^\gamma$
Perceptual blur [63] ^d	$\frac{\sum \text{all edge widths}}{\text{number of edges}}$
M_{S_CIQI} [44] ^e	$1 - \left(\frac{tep_{Estimated} - tep_{Sobel}}{tep_{Sobel}}\right)^{0.2}$
$M_{Kurtosis}$ [66] [57] ^f	$E[(W - \mu)^T \Gamma^{-1} (W - \mu)]^2$
Shaked_Test1 [64] ^g	$\frac{HighPass^F}{BandPass^F}$
IQM [67] ^h	$\frac{1}{M^2} \sum_{\theta=-180}^{180} \sum_{\rho=0.01}^{0.5} S(\theta_1) W(\rho) A^2(T_\rho) P(\rho, \theta)$

^a $\|\nabla\phi\|$ is magnitude of the intensity gradient.

^b $B_1(m, n)$ is AC component of a DCT block; MN is the total number of blocks; T_H is the threshold.

^c s is the average edge spread; α , β and γ are constant parameters.

^e tep represents total edge pixel.

^f W is a two dimensional random vector; μ and Γ are mean vector and covariance matrix of W .

^h (ρ, θ) is the 2D power spectrum; $A^2(T_\rho)$ is the modulation transfer function of the human visual system; $S(\theta_1)$ is the directional scale; $W(\rho)$ is the Wiener noise filter.

Image sharpness is the attribute related to the ability to preserve fine details and edges. Therefore, image sharpness can be seen as the contrast of edges. Many contrast measures have been developed, so the contrast measures can be utilized to measure the property of image sharpness. In the new sharpness measure presented here, the Sobel edge detection algorithm is first applied on each RGB color component. The resultant edge map is then multiplied with the original image to obtain three grayscale edge maps. Since sharpness is proportional to the perceived steepness of slopes, it is reasonable to treat each of the grayscale edge maps as an input image and measure the Weber contrast if the window size is small enough. The Weber contrast based measure is adequate in such scenario because the Weber contrast is good for measuring the local contrast of a single target seen against a uniform background [55], and edge maps only contain high frequency components in black background. In the experimental results shown in this chapter, the Weber contrast based Measure of Enhancement EME [52] is used on each grayscale edge map with an overlapped window size 3x3.

$$M_S = \sum_{c=1}^3 \lambda_c EME_{sharpness}(edgmaps_c) \quad (2.24)$$

$$EME_{sharpness} = \frac{2}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \log\left(\frac{I_{max,k,l}}{I_{min,k,l}}\right) \quad (2.25)$$

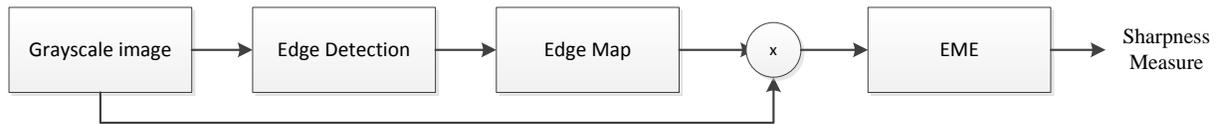


Figure 2-13. Flow diagram of the new sharpness measure.

Figure 2-14 shows the sensitivity of the sharpness metric for Gaussian blur images in TID2008 database [47]. The comparison of the CIQI sharpness measure and the M_S sharpness measure are shown in the figure. As demonstrated in Figure 2-14, the CIQI [4] measure shows that the image in Figure 2-14 (c) has better sharpness than the image in Figure 2-14 (a), which is in contrast with human observation. As a comparison, the M_S measure correctly ranks the image sharpness.



$M_{avg} = 0.0278$
 $M_{S_CIQI} = 0.3982$
 $M_S = 1.4593$
 (a)



$M_{avg} = 0.0228$
 $M_{S_CIQI} = \mathbf{0.3423}$
 $M_S = 0.9565$
 (b)



$M_{avg} = \mathbf{0.0211}$



$M_{avg} = \mathbf{0.0215}$

$$M_{S_CIQI} = 0.3457$$

$$M_S = 0.5681$$

(c)

$$M_{S_CIQI} = 0.2022$$

$$M_S = 0.3882$$

(d)

Figure 2-14. Image results on the sensitivity of sharpness metrics. (a)-(d) are four images with different levels of sharpness distortions. These images are obtained from the TID2008 database.

2.5. Summary

This chapter presents three NR measures for evaluating image and video attributes:

- An image color measure that combines the colorfulness and color tone measures. The colorfulness measures are based on combination of image statistical values. Compared with existing colorfulness measures, high order statistical values are incorporated. Therefore, more image information is included. The existing colorfulness measures are only applicable to natural images where variations exist in the image. For synthetic image patches, or natural images with low activities, these colorfulness measures fail because the high order statistical values are meaningless and the mean value alone is not able to distinguish colors. Therefore, the color tone measure is added as a uni-color differentiation term.
- Two image contrast measures. The MEMEE measure is a modification of the EMEE measure by incorporating more pixel intensities in a local region, rather than using only the maximum and minimum. This MEMEE is more stable and robust to noise. The CRME measure formulates the root mean square contrast with respect to the local region luminance in a color cube. Compared with the Weber or Michelson law based contrast measures, this CRME contrast measure does not have constraints on image patterns. The CRME compares differences in color cubes. Therefore, the image contrasts originated from the intensity differences, as well as the color differences are both accounted.
- An image sharpness measure M_S using Weber contrast law on grayscale edge map. In this measure, an edge map is treated as an image. Therefore, the contrast of the edge map indicates the sharpness property of the original grayscale image. The sharpness measure M_S can be used effectively in evaluating the blurriness conditions of images.

Chapter 3. Non-Reference Color Image and Video Quality Measures

This chapter presents a non-reference measure for evaluating color image quality by combining multiple color image attribute measures by linear addition. Different methods for training and obtaining the linear combination coefficients are used for variant applications. For image processing purpose where images before and after processing have the same source, the Linear Mixed Effect (LME) model is used. For video processing purpose where consecutive frames have completely different contents and suffer from different types of distortions, the Multiple Linear Regression (MLR) model is used. The presented Color Quality Measure (CQM) is used to benchmark image processing algorithm performances, as well as to assist in the selection of optimal operating parameters.

3.1. Overview

The objective image and video quality measures have many practical uses. For example, in communication systems, transmission of high quality images or videos requires more bandwidth. The objective measures can help to achieve a balance between transmitted image quality and the transmission cost.

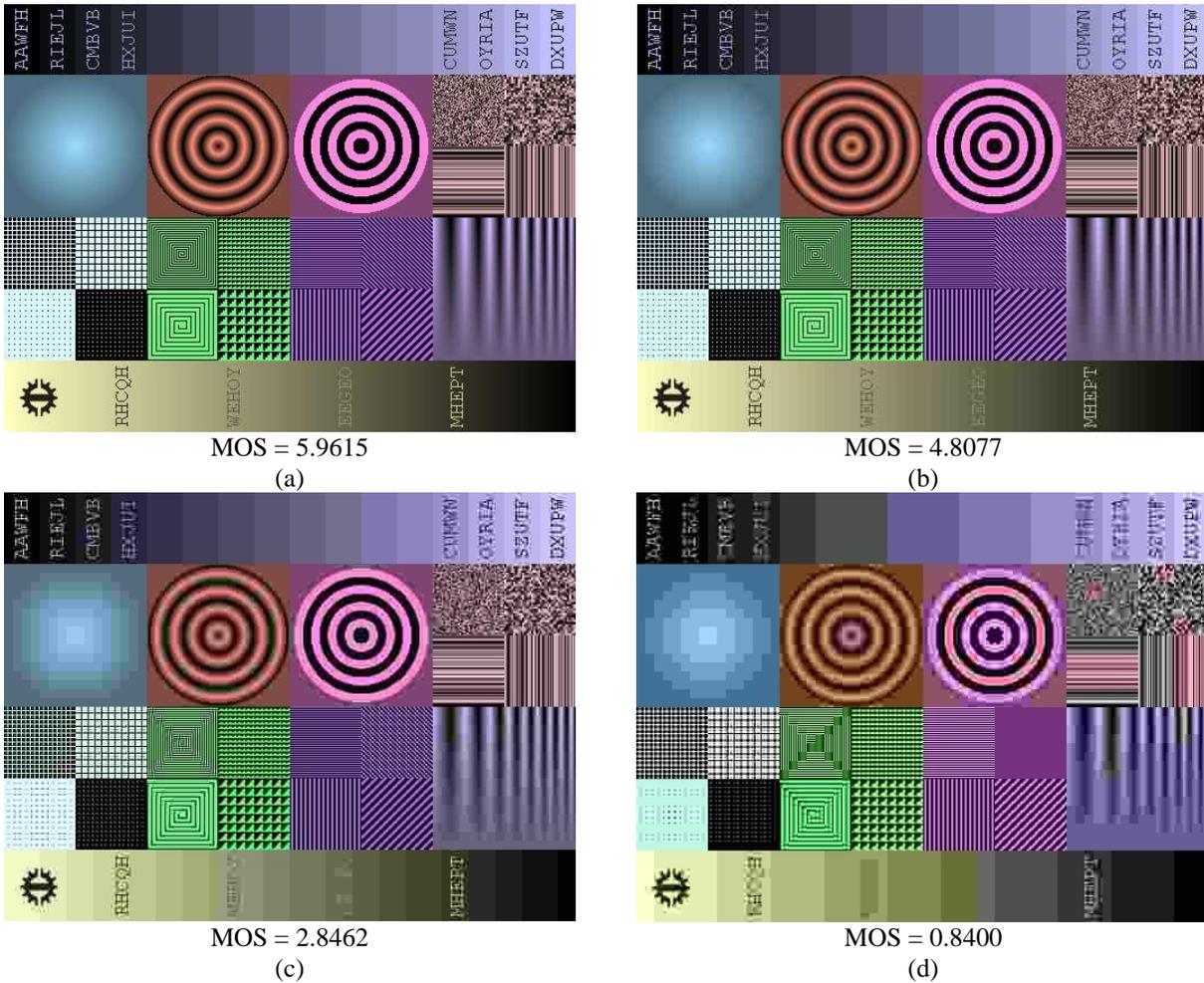


Figure 3-1. Four JPEG compressed images and their MOS scores. An objective quality measure that generates such measure values similar to the human vision automatically is desired.

The human's perception of image is a very complex phenomenon that depends on the illumination environment, the characteristics of the perceiving eye and brain, the exposure parameters, and the image acquisition and processing systems [68]. Besides, not every change in an image is perceptible, and not every region in an image receives the same consideration level [68]. In recent years, much effort has been made to develop objective color image quality metrics that perform similar to the human eyes. Some measures utilize the grayscale measures and

apply them to color images as illustrated in Figure 3-2. They either convert the color image to the grayscale, and then utilize the grayscale measures (Figure 3-2 (a)), or directly apply the grayscale measures in each color component and then fuse all measure values together (Figure 3-2 (b)). However, there are limitations with these methods. For the first type of approaches as shown in Figure 3-2 (a), the color to grayscale conversion is a lossy procedure. Therefore, some information which reveals from multiple color channels, such as image colorfulness or color contrast, are lost. For the second type of approaches as shown in Figure 3-2 (b), it is difficult to determine which the optimal color model is. For example, the RGB color model is inappropriate to extract the luminance information of an image because the three color channels in the RGB color model are highly dependent.

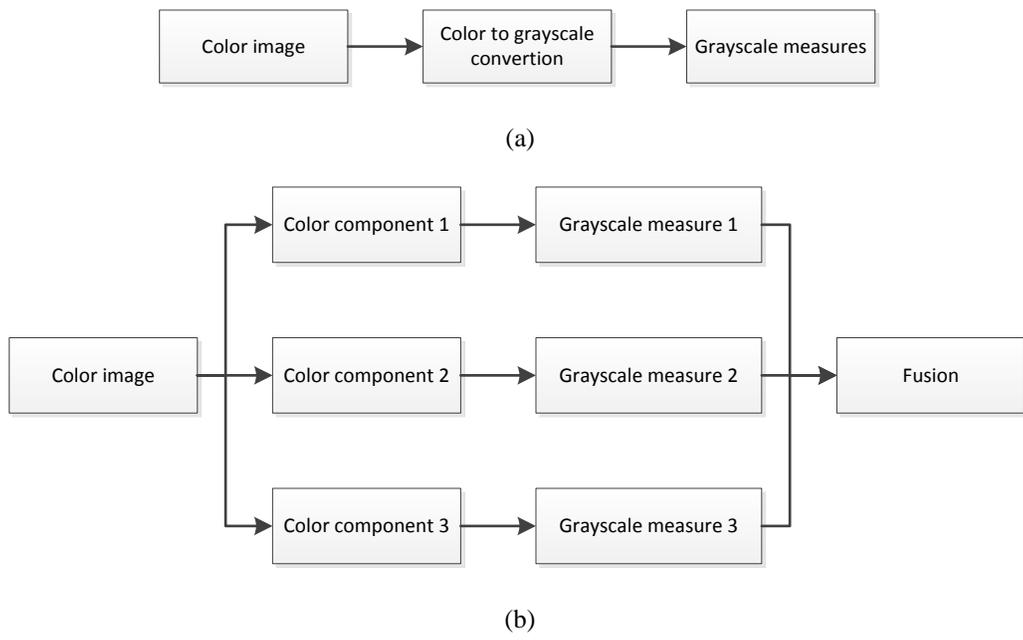


Figure 3-2. Illustration of using grayscale image measures for evaluating color image quality. (a) Converting a color image to a grayscale image and then applying the grayscale measures; (b) Applying the grayscale measures in each color component and then fuse all measure values together.

There have been developed many image attribute measures such as entropy [69][70], brightness [71][72][73], colorfulness [43][45][74], sharpness [27][57][58][59][75], and contrast [22][23][25][52][54][75][76]. However, measuring only one aspect of an image is not sufficient enough to evaluate the overall image quality because practically one image may be good at a certain image characteristics and poor at the others. One illustrated example is shown in Figure 3-3, where the image in Figure 3-3 (a) has more saturated color on the sky and the lawn, and Figure 3-3 (b) has better contrast on the trees and the car. Therefore, to evaluate the overall quality of an image, multiple image attribute qualities should be taken into consideration.



(a)



(b)

Figure 3-3. Illustrative examples. Image (a) has more vivid color. The colors of the sky, the lawn, and the lighthouse are more saturated. Image (b) has better contrast. The details of the trees and the car are more discernable.

The goal of this chapter is to create a robust Color image and video Quality Measure CQM based on linear combinations of image and video attribute measures. For different applications, variant methods of training and obtaining the linear combination coefficients are analyzed and presented. With appropriate trained parameters, the CQM is demonstrated to satisfy the properties in in Table 3-1. The competing algorithms are evaluated according to these properties in the experimental results section.

Table 3-1. Properties of good objective measures

Properties	Explanation
Accurate	A good objective measure is correlated with human perception, thus reliable for use in computer vision systems.
Applicable to Variant Types of Distortions	A good objective measure is distortion independent, thus it measures image quality consistently and reliably.
Computationally efficient	A good objective measure has low computational complexity and can be implemented in real-time.
Practical	A good objective measure can be used to assist the design of new image processing algorithms; A good objective measure can be used in practical industrial applications.

3.2. Color Quality Measure (CQM)

In Chapter 2, several image attribute measures are investigated. These attribute measures can be combined to generate an overall color image and video quality measure. The HVS is highly sensitive to changes in color, edge structures, and relative contrast. Therefore, changes on these image attributes result in completely different visual appearance according to the human eyes. Many distortions, such as unsatisfactory illumination, blurring, and display range changes that can occur during image acquisition, transmission, and display, cause degradations in image color, sharpness, and contrast. Therefore, as a generic measure for image and video qualities, the color quality measure CQM is generated by linearly combining the image color measure M_C , sharpness measure M_S , contrast measure M_{CON} , as well as the uni-color differentiation term UCD as shown in Eq. (3.1).

$$CQM = c_1 \times M_C + c_2 \times M_S + c_3 \times M_{CON} + c_4 \times UCD \quad (3.1)$$

The attribute measures presented in Chapter 2 can be used in Eq. (3.1). The linear combination coefficients used in Eq. (3.1) are obtained by training images selected from TID2008 database [47]. It's worth nothing that images in TID2008 dataset are all natural images with significant activities, and color castings do not exist in these images. So in the experiments shown in this chapter, the coefficients for the UCD term c_4 is set to zero. In this chapter, the following attribute measures are adopted and details can be referred in Chapter 2.

$$M_{C_CQM} = \frac{1.33(\sqrt{\mu_U^2 + \mu_X^2 + \mu_V^2} - 200) + 2.39\sqrt{\sigma_U^2 + \sigma_X^2 + \sigma_V^2} - 0.49\sqrt{k_U^2 + k_X^2 + k_V^2}}{200} \quad (3.2)$$

$$M_{S_CQM} = \sum_{c=1}^3 \lambda_c EME_{sharpness}(edgemap_c) \quad (3.3)$$

$$M_{CON_CQM} = MEMEE(image) \quad (3.4)$$

The TID2008 database is a published benchmark for measuring color image qualities. It features 25 reference images, 17 categories of distortions, and 4 levels of distortions for each image. Furthermore, the Mean Opinion Score (MOS) obtained from 838 subjective experiments is provided for each test image [47]. The MOS is considered as the true numerical evaluation, and is used to train the linear combination coefficients. To guarantee the training set and the testing set are independent, half of the images in the TID2008 dataset are used for training and the other half are used for testing. It is worth noting that the CQM in Eq. (3.1) has three parameters c_1 , c_2 , and c_3 . In special cases, one or more parameters could be zero. For example, the CT or X-ray images are usually captured in

grayscale so the coefficient for the color measure can be set to zero. If two of the parameters are zero, the quality measure CQM reverts to an attribute measure.

In this chapter, the color quality measure CQM is designed for two applications:

(1) For image processing applications, where the image before and after processing have the same source. In this application, the types of distortions are usually known and the CQM is aimed to differentiate the perceived degrees of distortions of the images;

(2) For video processing applications, where consecutive frames are independent. In this application, the video frames may suffer from any types of distortions. Therefore, the CQM is designed to fit independent images and distortions.

In the following, two methods of training and obtaining the linear combination coefficients c_1 , c_2 , and c_3 for these two applications are discussed.

3.3. CQM for Images with the Same Source

Usually, image processing algorithms are used for applications with knowledge of major distortion types. For example, contrast enhancement is used in medical images, de-compression is used in video transmission tasks, and color correction is used for underwater images. Therefore, for image processing applications, the focus is on discriminating different levels of distortions. Accordingly, the coefficients for each category of distortions are trained separately. Under this scenario, the Linear Mixed Effect (LME) model is appropriate to fit the data, and the effect of the picture is considered random [77]. The LME model is an extension of the linear regression model for data that are collected in groups. Recall that in the training set, there are 12 images with variant scenes and each image has four levels of distortions. Accordingly, the four images with variant levels of distortions share one picture effect, and there are 48 images in the training set for each type of distortion. Table 3-2 shows the recommended coefficients for four types of distortions: Gaussian blur, contrast change, JPEG 2000 and denoising obtained from the LME model.

Table 3-2. Recommended linear combination coefficients for CQM in evaluating the quality of images with the same source. The coefficients are obtained by applying LME on 48 training images suffering from the corresponding distortions

Distortion Type	c_1	c_2	c_3
-----------------	-------	-------	-------

Gaussian blur	1.5655	3.2981	1.9056
Contrast change	4.1707	0.1781	0.2701
JPEG2000 compression	0.1011	2.6777	0.7669
Denoising	3.3855	2.9719	-1.5440

The correlation coefficients are used to analyze the statistical relationship between the measured value and the MOS. Pearson's linear product moment correlation measures how far each measure value deviates from the MOS [78]. Spearman's rank order correlation compares the rank of image qualities, and Kendall's tau correlation is usually suggested for measuring rank of non-normal data [79]. The Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) between the prediction and the MOS are also used for comparison. The five indexes are defined as Eq. (3.5) to Eq. (3.9) where a and b represent two variables, \bar{a} and \bar{b} are the mean values of a and b , R represents the rank, and N represents the total number of variables.

$$Pearson(a, b) = \frac{\sum_{k=1}^K (a_k - \bar{a})(b_k - \bar{b})}{\sqrt{\sum_{k=1}^K (a_k - \bar{a})^2} \sqrt{\sum_{k=1}^K (b_k - \bar{b})^2}} \quad (3.5)$$

$$Spearman(a, b) = \frac{\sum_{k=1}^K (R(a_k) - \bar{R(a)}) (R(b_k) - \bar{R(b)})}{\sqrt{\sum_{k=1}^K (R(a_k) - \bar{R(a)})^2} \sqrt{\sum_{k=1}^K (R(b_k) - \bar{R(b)})^2}} \quad (3.6)$$

$$Kendall(a, b) = \frac{N(\text{concordant pairs}) - N(\text{discordant pairs})}{\frac{1}{2}K(K-1)} \quad (3.7)$$

$$RMSE(a, b) = \sqrt{\frac{1}{K} \sum_{k=1}^K (a_k - b_k)^2} \quad (3.8)$$

$$MAE = \frac{1}{K} \sum_{k=1}^K |a_k - b_k| \quad (3.9)$$

Table 3-3 summarizes the performances of the CQM measure and the CIQI [44] measure for evaluating the qualities of images with the same source. For these images, the coefficients obtained from training images that exhibit the same type of distortions in Table 3-2 are used. It is worth noting that since the measures have different dynamic ranges, all measure values are normalized for comparison purposes. The results show that the proposed CQM has a strong correlation with MOS on the order 0.90, and performance ranges from 0.65 to 47.22 percent better than CIQI.

Table 3-3. Performance comparisons of qualities of images with the same source. The CQM has greater Pearson, Spearman, and Kendall correlations, and smaller RMSE and MAE for these Gaussian blurred, JPEG2000 compressed and denoised images in the TID2008 dataset.

Distortion Type	Measure	Pearson	Spearman	Kendall	RMSE	MAE
Gaussian blur	CIQI	0.7894	0.7000	0.6389	0.4981	0.4227
	CQM	0.9740	1.0000	1.0000	0.1080	0.1733
JPEG2000	CIQI	0.8406	0.8167	0.7500	0.4108	0.3658
	CQM	0.9619	1.0000	1.0000	0.2245	0.2003
Denoise	CIQI	0.9375	0.9333	0.9167	0.2732	0.2357
	CQM	0.9933	1.0000	1.0000	0.0889	0.0812
Average	CIQI	0.8707	0.8364	0.7829	0.3762	0.3276
	CQM	0.9475	0.9656	0.9426	0.1886	0.1882

Figure 3-4 shows the measured CQM (red), CIQI (green) values and MOS (blue) for 48 denoised images. Recall that in the TID2008 dataset, 4 levels of distortions are provided for each image. The curve in Figure 3-4 shows that the CQM measure has better correspondence with the MOS, and the four levels are clearly distinguishable. In contrast, the CIQI measure values deviate from the MOS with a greater variance, resulting in possible incorrect rankings of image qualities.

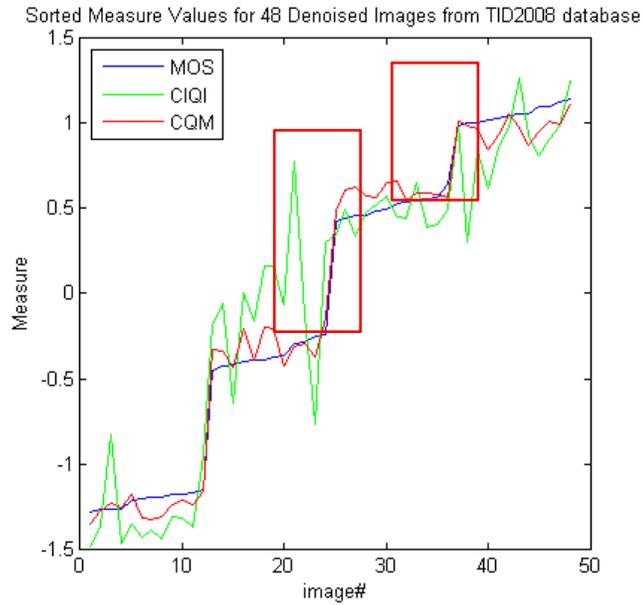
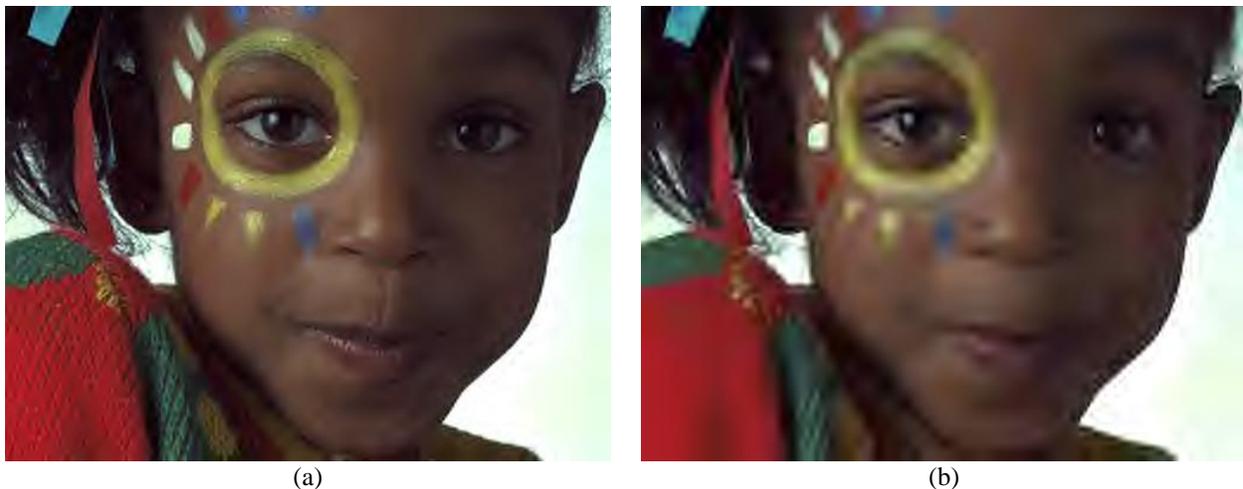


Figure 3-4. Measure metrics and MOS for 48 denoised images. The measure values obtained from the MOS and the CQM have four distinguished levels, while the CIQI measure value has a greater variance as shown in the red rectangles.

Two examples of ranking the qualities of JPEG2000 compressed images are shown in Figure 3-5 and Figure 3-6. In each figure, four images suffer from different levels of JPEG2000 compression are demonstrated. Their CQM and CIQI measure values and the MOS are listed and compared in Table 3-4 and Table 3-5. It is observed that the CQM metric generates the same rank order as the MOS. Notice that the image qualities of Figure 3-5 (c) and Figure 3-5 (d) are quite different. The CQM measure values reflect the quality degradation. On the other hand, the CIQI model assigns similar values for Figure 3-5 (c) to (d) despite the obvious visual difference of these images. In Figure 3-6, the CIQI generates the incorrect rank for image (b) and (c).



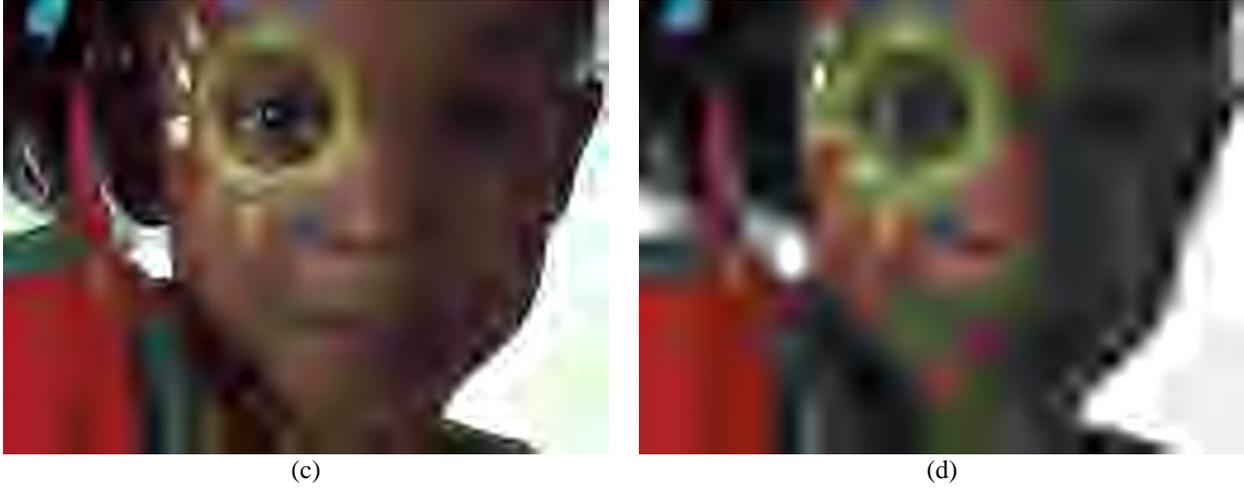


Figure 3-5. Four levels of JPEG2000 compressed images. These images are obtained from the TID2008 dataset.

Table 3-4. Comparison of measure metrics for the four JPEG2000 compressed images in Figure 3-5

	MOS	CIQI	CQM
Figure 3-5 (a)	4.7222	0.5357	12.8060
Figure 3-5 (b)	3.0000	0.5330	12.2404
Figure 3-5 (c)	1.8889	0.5244	11.1069
Figure 3-5 (d)	0.3429	0.5228	6.6886



(a)



(b)



Figure 3-6. Four levels of JPEG2000 compressed images.

Table 3-5. Comparison of measure metrics for the four JPEG2000 compressed images in Figure 3-6.

	MOS	CIQI	CQM
Figure 3-6 (a)	5.8182	4.2910	6.9480
Figure 3-6 (b)	5.0909	4.2023	6.7421
Figure 3-6 (c)	4.3333	4.2358	5.1298
Figure 3-6 (d)	2.3939	3.7088	3.3758

3.4. CQM for Independent Images

Another practical application of the CQM is video processing. Figure 3-8 shows some sampled video frames from a YouTube clip for the Super Bowl 2015. These images are obtained by sampling the video every 20 frames. It is seen that the scenes in the video are changing frame by frame. Some frames are distorted by camera movement or color desaturation due to the disturbances in video capturing or band limits in video transmission. This application differs from the image processing application described in the previous section in the means that consecutive frames may have completely different contents or suffer from different distortions. Therefore, a different training method for obtaining the linear combination coefficients is used.

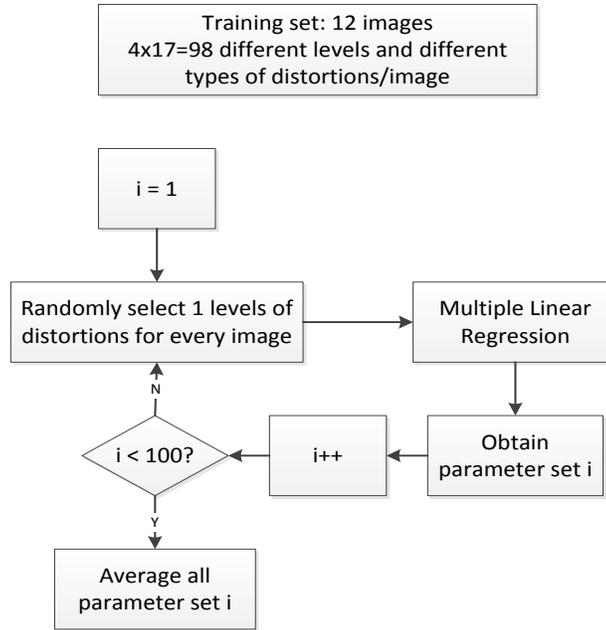


Figure 3-7. Flow of the training procedure for images with different sources.

The training for video processing tasks aims to extract content independent statistical features of natural scenes [61]. In the training procedure as shown in Figure 3-7, the training image pool contains twelve images with different contents. For each image, different types of distortions are not distinguished, and sixteen distortions with variant types and levels are mixed together. One of the sixteen distorted images is randomly selected for each of the twelve images. Then the Multiple Linear Regression (MLR) is applied between the subjective MOS and objective values of the attributes. The random selection and calculation procedure is repeated 100 times, and the final coefficients are averaged and listed in Table 3-6. Generally, a greater CQM measure indicates a better image quality. However, video frames may suffer low frequency noise such as Gaussian blurring, as well as high frequency noise, such as Salt and Pepper noise. In practical applications, usually a certain range is selected. If the CQM measure value is out of bound, the image may be distorted.

Table 3-6. Recommended linear combination coefficients for CQM in evaluating independent image qualities. The coefficients are obtained by applying MLR on images suffering from mixed distortions.

Distortion Type	c_1	c_2	c_3
Mixed distorted images	1.6256	1.5012	3.0173

With the coefficients in Table 3-6, the CQM measure values are obtained for the frames in Figure 3-8. In Figure 3-8, some frames are intended to be distorted manually for simulating the possible video quality degradations during video capturing or transmission. These artificially distorted images are zoomed and demonstrated in Figure 3-9. Frame Figure 3-8 (c) is added with white Gaussian noise to simulate the noise exist in the transmission; frame Figure 3-8 (g) is contaminated by artificial motion blur to simulate the camera movement; frame Figure 3-8 (j) is down sampled and then up sampled by bilinear interpolation to simulate the situation that the high resolution videos are temporarily lost; frame Figure 3-8 (n) suffers from de-saturated colors to simulate the color demodulation error. It is seen that these frames in Figure 3-8 show completely different contents. The frames with good quality have comparable CQM measure values around 3 and 4, while the distorted frames show deviated measure values. For such videos, the upper and lower thresholds can be set as 3.2 and 5 respectively. If certain video frames are detected have the out of bound CQM measure values, the TV station should get an alert of possible transmission errors.



CQM = 4.5486
(a)



CQM = 3.8726
(b)



CQM = 6.7065
(c) Additive white gaussian noise



CQM = 4.5039
(d)



CQM = 3.2763
(e)



CQM = 3.5424
(f)



CQM = 2.2566
(g) artifatial motion blur



CQM = 3.1556
(h) blurring from scene



CQM = 4.3059
(i)



CQM = 2.1392
(j) downsample then upsample



CQM = 3.5406
(k)



CQM = 3.3999
(l)



CQM = 3.5865
(m)



CQM = 3.1490
(n) desaturated

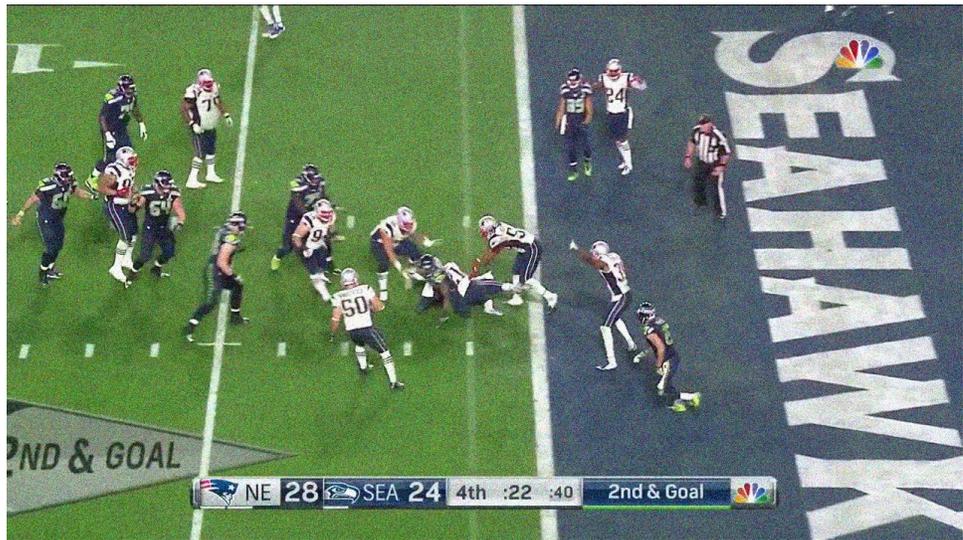


CQM = 3.5436
(o)

Figure 3-8. Sample video frames. These video frames have different contents. Frame (c) is contaminated by white Gaussian noise. Frame (g) is contaminated added motion blurring. Frame (h) is blurred by the player's movement. Frame (j) is downsampled and then upsampled. Frame (n) is desaturated.

Frame (c)

white Gaussian noise



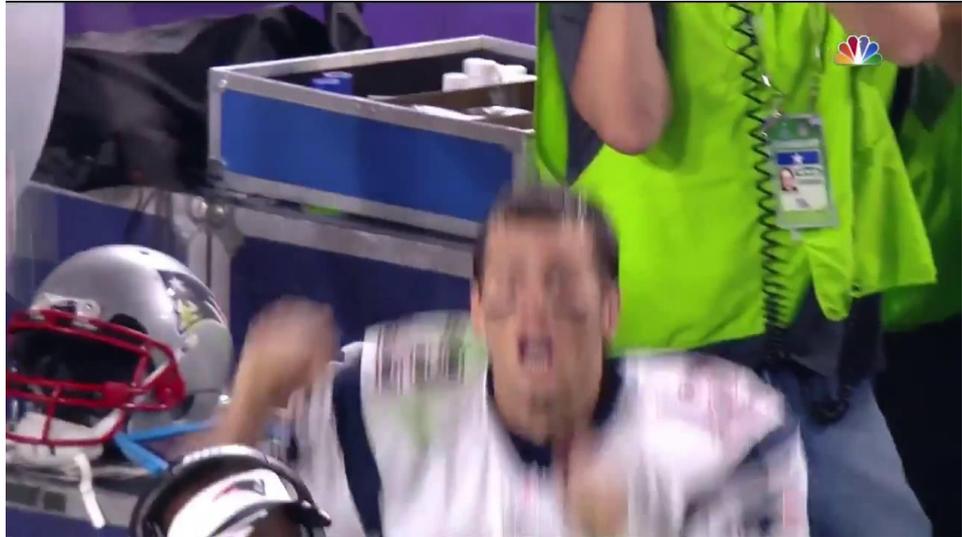
Frame (g)

motion blurring



Frame (h)

player's movement



Frame (j) downsampled
and then upsampled



Frame (n)
desaturated



Figure 3-9. Zoom of the artificially distorted images.

In the testing procedure, one of the distorted images is randomly selected for each of the images under test. The CQM calculated using the coefficients listed in Table 3-2 is then compared with the MOS to get the correlation coefficients. The sample size for each experiment is 12, and the simulation is repeated 100 times. The performance of CIQI and CQM are compared in terms of Pearson’s, Spearman’s, Kendall’s correlation, RMSE, and MAE in Table 3-7. Note that the coefficients used in the CIQI model [44] are generated using the same training images for fair comparison. It can be seen that the CQM is 15% to 20% more correlated with MOS than the CIQI measure for independent images.

Table 3-7. Performance comparisons of the CQM for independent images. The CQM has greater Pearson, Spearman, and Kendall correlation coefficients, and smaller RMSE and MAE values for mixed distorted images in the TID2008 dataset

Distortion Type	Measure	Pearson	Spearman	Kendall	RMSE	MAE
Mixed distorted images	CIQI	0.6057	0.5643	0.4323	1.4138	1.1720
	CQM	0.7485	0.7671	0.6067	1.2322	1.0195

3.5. Experimental Results

The CQM measures can be effectively used for benchmarking image processing algorithm performances and assisting selecting optimal operation parameters. In this section, some experimental results are demonstrated. The CQM measure is then compared with the CIQI [44] and the JPQM [80].

3.5.1 Parameter Selection

The measures can also be used for assisting in the selection of optimal operating parameters for image enhancement, denoising, or decompression algorithms. Figure 3-10 shows an example of applying the intensity adjustment power law on each color component of an under illuminated image from Barnard dataset [81]. For dark images, the power α should be smaller than 1. In the experiment, the power α ranges from 0.2 to 0.9. For the test image in Figure 3-10, the CQM selected the optimal alpha power at 0.5, and it achieves good enhancement results with respect to both contrast and color. Figure 3-10 also shows enhancement results with other α values. It can be seen that smaller alpha values tend to wash out the color, and greater alpha values do not effectively correct the illumination.

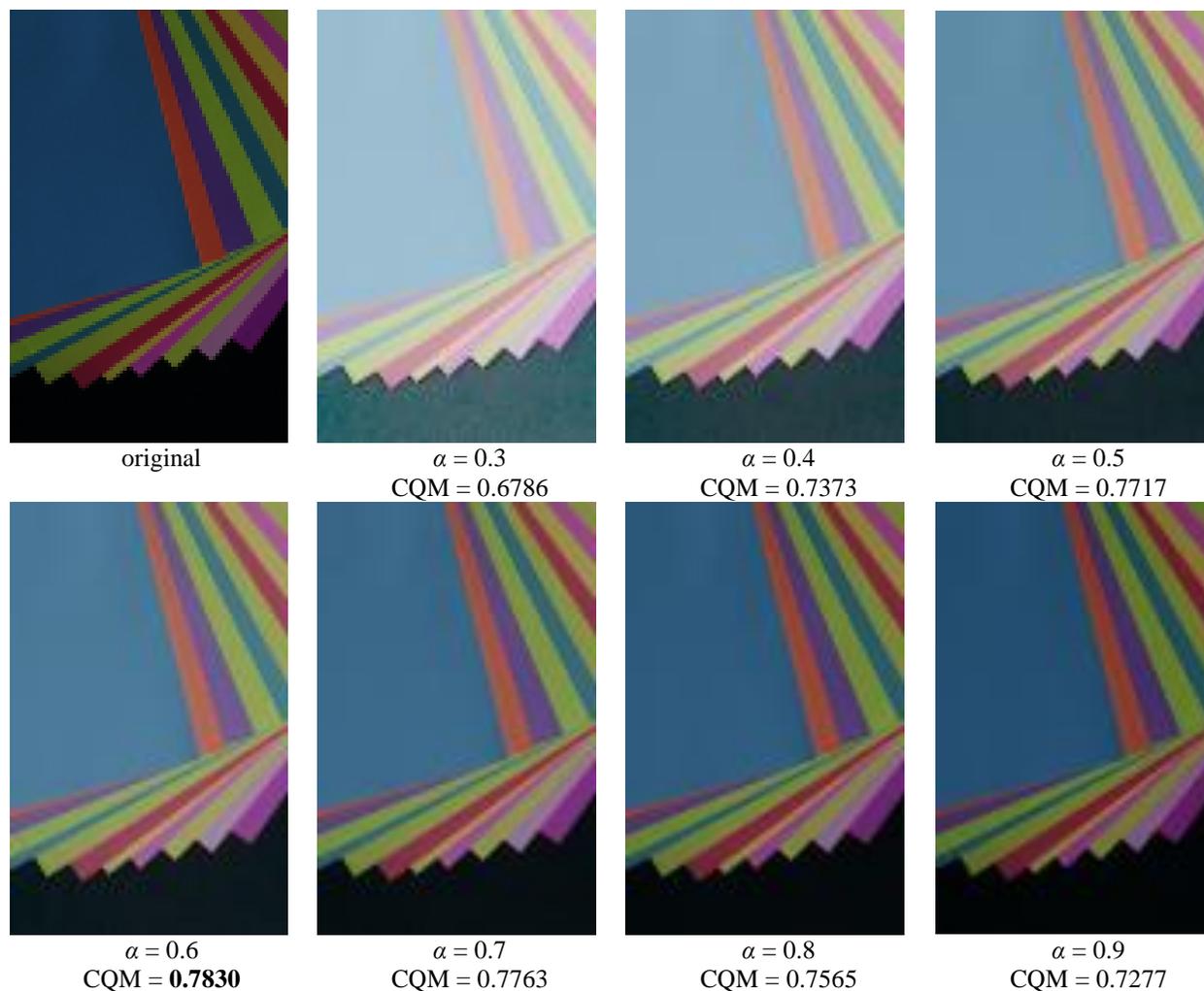


Figure 3-10. Using CQM for assisting selecting optimal alpha power for an underilluminated image

3.5.2 Benchmarking Image Processing Algorithms

A good color image quality measure can be used to benchmark image processing algorithms. An example is shown in Figure 3-11. The original test image is from the NASA Retinex database and is taken in strong sunlight resulting in the significant loss of color and details [82]. Seven color enhancement results are obtained from the package provided in [45][83]. The image enhancement algorithms include Alpha Rooting (AR) [84], Multi-Contrast Enhancement (MCE) [26], Multi-contrast Enhancement with Dynamic Range Compression (MCEDRC) [85], and varieties of Contrast Enhancement by Scaling (CES) [45]. The colorfulness measure M_{C_CQM} selects the TW-CES-BLK as the enhancement result shows the most vivid color. The CIQI and the CQM choose the MCE enhancement result as the best because it has the best overall quality. As a contrast, the JPQM [80] treats the enhancement as

introduced artifacts, thus it selects the AR enhancement result which is most similar to the original image as the best enhanced image.



Original



AR [84]



DRC-CES [45]



DRC-CES-BLK[45]

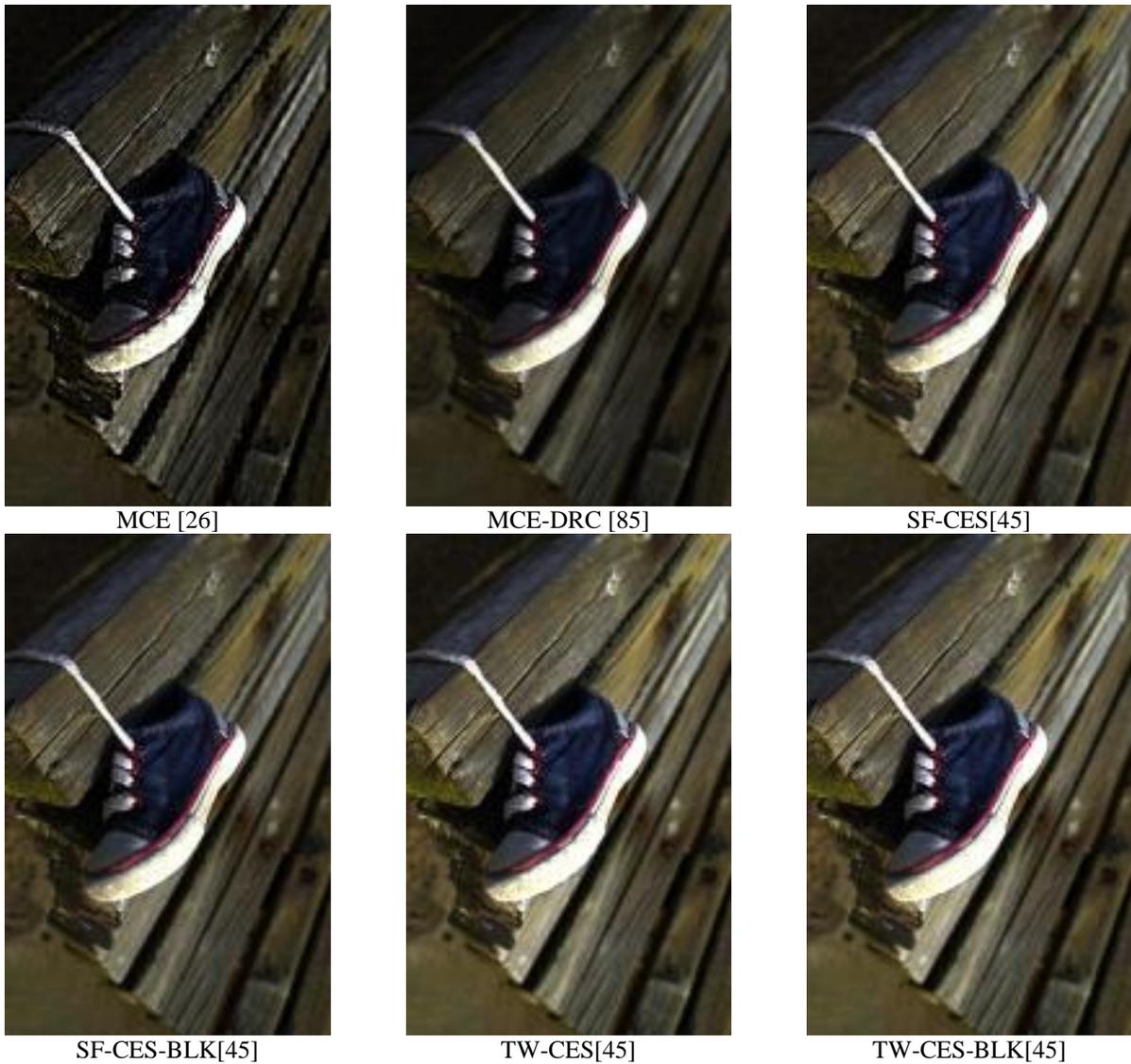


Figure 3-11. Nine enhanced images for the ‘Shoe’ image. The original image is taken in strong sunlight resulting in the significant loss of color and details. Nine image enhancement algorithms are applied on this image. Their performances need to be evaluated. Refer to the references listed in the brackets for details of the enhancement algorithms.

Table 3-8. Comparison of measure metrics for the nine enhancement results in Figure 3-11.

Techniques	JPQM	M_{C_CQM}	CIQI	CQM
AR	8.9465	0.7063	0.1950	1.6560
DRC-CES	7.6822	0.7111	0.2379	1.6449
DRC-CES-BLK	8.1790	0.7103	0.2382	1.7309
MCE	7.7206	0.7426	0.3327	2.7249
MCE-DRC	8.6363	0.6984	0.1917	1.6038
SF-CES	8.0296	0.7250	0.2519	1.6294

SF-CES-BLK	8.4431	0.7242	0.2555	1.6616
TW-CES	7.6154	0.7432	0.2964	1.6376
TW-CES-BLK	7.9287	0.7792	0.2996	1.6297

3.5.3 Computational Complexity

The CQM average execution time is tested for 48 test images. The size of the test images is 512 by 384 by 3. The testing computer features an Intel Core 2 Quad CPU, 2.66GHz frequency, and 3.25GB RAM. Each simulation is repeated 100 times using Matlab. The simulation results show that the average execution time for CQM is 0.9272 seconds per image. It is worth noting that the colorfulness, sharpness, and contrast are calculated independently, so the three attributes can be processed in parallel to further accelerate operating speed.

3.6. Summary

In this chapter, multiple image attributes are combined to form a color image quality measure CQM. CQM, a non-reference color image quality assessment, can be used in ranking the qualities of images with the same or different contents. To achieve these two tasks, different training methods for obtaining the linear combination coefficients are adopted. For images with the same content, the Linear Mixed Effect (LME) model is used with a random effect indicating the image effect. These coefficients are trained for each type of distortions considering that the applications of the CQM measure under this model is mainly used for image processing tasks with known distortions. For images with different contents, the Multiple Linear Regression (MLR) model is used, and the independent images with multiple types of distortions are trained together. The CQM measure under this model is usually used in video quality evaluations where unknown distortions may occur. Comparisons of the two training methods are demonstrated in Table 3-9.

Table 3-9. Comparisons of the training procedures for image processing and video processing applications

	Images with the same source	Independent video frames
Coefficients Training Method	Linear Mixed Effect (LME)	Multiple Linear Regression (MLR)
Images Used in Training	Images have the same source are considered the random effect	All training images are independent
Distortions Used in Training	Train coefficients for each type of distortion separately	All types of distortions are mixed together
Measure Value Interpretation	The greater CQM corresponds to better quality	CQM within a certain range corresponds to good quality

Experimental results demonstrate that the CQM is correlated with human visual perception for both image processing and video processing applications; it is not limited to specific types of distortion; it can be used to benchmark image processing algorithms; it can be embedded into image processing systems to assist selecting the optimal operating parameters; and it has low computational complexity and can be implemented in real-time.

Chapter 4. Reference Based Edge Measure

Edge detection is an important task in image processing, and the quality of recognition and tracking tasks is often reflected by the quality of edge detector outputs. Therefore, it is necessary to develop effective edge map quality measures to assist evaluating the performances of edge detectors. Objective evaluation measures are crucial in automatically determining the optimal edge map for a given image or application, as well as its parameter values. In this chapter, existing reference-based measures are briefly reviewed, their limitations are analyzed, and the reference-based evaluation standards are discussed. This chapter then introduces a new reference-based edge measure, namely RBEM, to evaluate the performance of edge detector outputs relative to a ground truth. The new measure fuses three edge attribute metrics, the edge pixel presence, edge corner localization, and double edge occurrence. Each of these metrics can be used separately as a standalone measure to evaluate edge map in terms of specific characteristics. The proposed measure is demonstrated effective on selecting the best edge detector among several edge detectors, as well as on selecting the optimal parameter values, for both synthetic images and natural images. Experimental results show that the presented RBEM measure outperforms existing methods according to subjective evaluation Mean Opinion Scores, as it considers more important visual features in its evaluation.

4.1. Overview and Problem Formulation

Edge detection is essential and used extensively in image processing and computer vision. Edges of an image capture the most important structures of an image. Therefore, the amount of information needed to be processed is substantially reduced [86]. Edge detection has also served as a basis, or preprocessing step for many other image processing tasks, such as feature extraction, object detection and recognition, image enhancement, and image segmentation [87]. Consequently, research has been dedicated in designing good edge detectors for many years [6][88][89][90]. The correct indication of boundaries between objects allows for accurate object identification and precise motion analysis for several machine vision applications. Edge detection has been widely used in military applications [91], biometrics [92], geophysical applications [93], surveillance application [94], consumer uses [95], and biomedical imaging [96][97][98]. In these computer vision applications, the edge detection procedure often lead to further calculations of area and perimeters, classification of shapes and scenes, and tracking of targets once they have been isolated from the image background [99].

As a result, accurately determining edge pixel locations in an image is of paramount importance [100]. There are some common standards for evaluating edge quality. For example, in many image analysis processes that use edge information, it is required that determined edge pixels in an image represent actual edges, and that edge fragments have uniform width and a certain degree of continuity and connectivity [101]. However, quantitative evaluation of edge map quality is still far from ideal. As edge detection is such an important task for vision systems, it is necessary to develop effective and efficient objective edge quality evaluation measures. Such a measure should be able to determine which edge detection algorithm works best for a certain application. Moreover, it should also be able to select the optimal operating parameters of an edge detection algorithm automatically.

There are three commonly accepted types of edge map evaluation approaches: human evaluation, reference-based objective measures, and non-reference objective measures. Human evaluation methods use blind subjective ratings made by experts in that area. The average rating score is used as the quantitative indication of edge quality. This approach avoids pixel level comparisons. As a result, however, it does make statements about the false positive and false negative errors [102]. Methods have been suggested to remove some of the subjective bias in visual assessments by considering the opinions of many participants over a large number of test cases and determining the consistency in participant rankings [103]. However, it is impossible to remove all of this bias and the results can

still be misleading. More importantly, the rankings certainly cannot be automated for image processing and computer vision systems. In this chapter, the reference-based objective edge measure is presented. The non-reference objective edge measure will be presented in Chapter 5.

Full-reference edge map evaluation measures require a ground truth as the reference to assess the quality of the resultant edge map. Pratt's Figure of Merit (FOM) is one of the well-known reference-based objective edge map evaluation [104]. It measures the pixel level distances between the resultant edge map and the ground truth edge map. Pratt's FOM is easy to implement and its mathematical meaning is straightforward. Suggested revisions to Pratt's FOM, such as Pinho's method, improve the measure by separately considering different edge pixel probabilities [105][106]. However, while the existing reference-based measures, such as Pratt's and Pinho's FOM, have shown to have a number of desirable properties, they fail to account for some very important features that are seen visually in assessing the quality of an edge map, or only accounts for them indirectly [107]. This is to say that there are qualities of edge maps that go beyond the edge pixel level. Moreover, the practical discrepancy of the existing measures is their limited treatment of false positive and false negative instances. Therefore, it is necessary to establish a more comprehensive set of criteria to measure edges in accordance with the human vision.

The primary goal in this chapter is to develop and evaluate a new reference-based measure and to compare its performance to traditional reference-based approaches. Additionally, some criteria of a good edge map are established, including the edge pixel presence/localization, the edge corner presence/localization, and the double edge occurrence. Then a separate measure is proposed for each of these important criteria. Finally, the overall reference-based objective edge quality metric RBEM is generated by fusing the measures according to their importance for a specific application. The effectiveness of the RBEM measure is demonstrated by comparing the performance of many traditional edge detectors. It can also be used for selecting optimal operating parameters within a single edge detector. In addition, the performance of the presented measure is evaluated based on visual scores. Experimental results via computer simulations validate that the presented method RBEM outperforms other existing reference-based edge detection measures by means of the RBEM better coincides with the subjective opinions of experts gathered via Mean Opinion Score (MOS).

4.2. Related Works

This section briefly reviews some existing reference based edge measures. These measures are used as comparisons to the new RBEM edge measure.

4.2.1 Statistical Analysis

When the ground truth is known, standard probabilistic measures compare the resultant edge map to the actual edge map, resulting in some ratio of identified or misidentified edge location to actual locations. Typical statistical analysis methods include the error-based measures such as the Mean Square Error (MSE), or the distance-based measures such as the cyclic distance measure. Receiver Operating Characteristic (ROC) analysis and the Precision Recursion (PR) curve improves the basic measures by extending the analysis to look at sensitivity vs. specificity and considering four probabilities for each match: true positive, false positive, true negative, and false negative. In their commonly used form, ROC analysis and PR curve neglects the issues of locality which occur when the edge is correctly identified, but in the wrong location. This type of methods can suffer greatly when small variations of the binary data result in large differences of the resulting ratios.

4.2.2 Pratt's FOM

Pratt's FOM provide a single quantitative index of edge detector performance. It compares the distance between the actual and the ideal edge pixels by:

$$FOM_{Pratt} = \frac{1}{\max\{N_O, N_T\}} \sum_{i=1}^{N_T} \frac{1}{1 + \alpha d_i^2} \quad (4.1)$$

where N_O is the number of edge pixels in the ground truth, N_T is the number of pixels in the testing edge map, d_i is the distance between an actual edge pixel in the resultant edge map and its nearest ideal edge pixel, and α is a penalty parameter [104]. Pratt's FOM ranges from 0 to 1, where a 1 corresponds to a perfect match between the edge map being evaluated and the ground truth edge map. In the experimental results shown in this paper, α is set to 1/9, as suggested in the literature [104][108].

4.2.3 Pinho's Improvement to Pratt's FOM:

Pinho et. al [105] suggested an improvement edge measure by bridging the gap between Pratt's FOM, where distances are considered, and ROC analysis, where multiple probabilities are considered. The Pinho's measure F is given by

$$F = \left(\frac{1}{N_0} \sum_{k=1}^{N_0} \frac{1}{1 + \alpha d_k^2} \right) \times \left(\frac{1}{1 + \beta \frac{N_{FA}}{N_0}} \right) \quad (4.2)$$

where N_0 is the number of edge pixels in the ground truth, and N_{FA} is the number of false edge elements. The first term in Eq. (4.2) is a revised version of Pratt's FOM differing only in the direction of matching, and the second term penalizes false edge pixels explicitly. The two parameters α and β are accountable for the balancing between the missing and displaced edge error accordingly. In practice $\alpha = 1/9$ and $\beta = 1$ are used, as suggested in the literature [104][108].

4.2.4 Boaventura's Method

Boaventura and Gonzaga [109] proposed a comprehensive edge measure which combines four terms: the percentage of pixels that are correctly detected P_{co} , that are not detected P_{nd} , that are erroneously detected P_{fa} , and the Pratt's FOM. As the definitions of each of the four terms indicate, P_{co} and Pratt's FOM approach 1, P_{nd} and P_{fa} approach 0 when the detected edge map is ideal. Combining these four terms, they presented a global index, which is defined by Euclidean distance in \mathbb{R}^4 to the point P (1, 1, 0, 0), where its coordinates are the optimal values achieved by these four indices respectively. The distance to the ideal point P is calculated by Eq. (4.3) and the ideal edge detector will minimize this distance [109].

$$d_2^4 = \sqrt{(P_{co} - 1)^2 + (FOM_{pratt} - 1)^2 + P_{nd}^2 + P_{fa}^2} \quad (4.3)$$

4.3. New Full-Reference Edge Measure RBEM

In this chapter, a new reference based edge measure RBEM is presented according to a set of criteria. The new measure fuses three component metrics, based on edge pixel presence, edge corner localization, and double edge

occurrence. Figure 4-1 shows a block diagram of the new RBEM evaluation system. Terms regarding edge pixel presence/localization, edge corner presence/localization, and double edge occurrences are calculated separately. Each term is an edge measure on its own, evaluating edge maps based on an important perceptual property. These attribute measures are then fused with different weights to yield the overall edge map quality evaluation for specific applications.

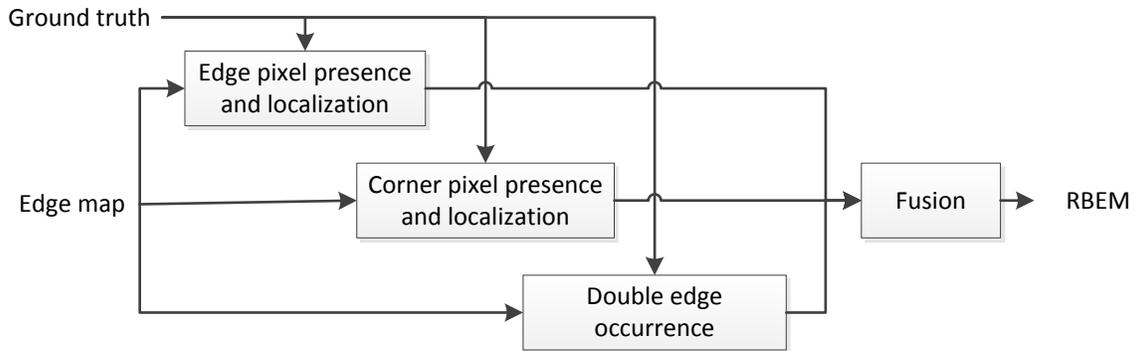


Figure 4-1. Block diagram of the new RBEM evaluation system. It is composed of three edge attribute measures.

4.3.1 Edge Pixel Presence/Localization Measure

The detection and evaluation of edge locations to pixel level accuracy is important in many computer vision applications, including the estimation of disparities in edge-based stereo matching, the calculation of optic flow, the tracking of target of interest, and the gauging in industrial inspection systems [110]. The edge pixel location errors between the resultant edge map and the ground truth edge map have been studied previously on mathematical analyses [111][112][113]. In these methods, edge locations are defined as the maximum point of a linear filter output. This definition is reasonable in the continuous domain. However, it has been noted that the error bound derived in the continuous domain will be degraded in passing to the discrete domain [114]. Therefore, the methods of extending the continuous domain method to the discrete domain suffer from accuracy problem. Accordingly, the evaluation of edge locations based on edge pixel presence and localization is considered. The edge pixel presence indicates the number of false positive and false negative edge pixels in a testing edge map, and the edge pixel localization regards the distances between the falsely detected or falsely missed edge pixels and their nearest ideal edge pixels or correctly detected pixels.

The proposed edge pixel presence/localization term includes two parts: a false positive term as shown in Eq. (4.4) and a false negative term in Eq. (4.5). Similarly as the definition used in the Pratt's FOM, the false positive term measures the average distance between the false positive edge pixels in the testing edge map and its closest ideal edge pixels in the ground truth edge map. It is formulated as follows:

$$D_{PFP} = \frac{1}{mn - P_O} \sum_{i=1}^{P_{FP}} 1 - \frac{1}{1 + \alpha d_{PFPi}^2} \quad (4.4)$$

where m and n are the size of the edge map, P_O is the number of edge pixels in the ground truth edge map, P_{FP} is the number of false positive edge pixels in the edge map being evaluated, and d_{PFPi} is the distance between the i^{th} false positive edge pixel from the closest ideal edge pixel [107].

The edge pixel presence/localization measure also contains a false negative term, which measures the false negative edge pixels and their distances from the closest correctly detected edge pixels. It is given by

$$D_{PFN} = \frac{1}{P_O} \sum_{i=1}^{P_{FN}} 1 - \frac{1}{1 + \alpha d_{PFNi}^2} \quad (4.5)$$

where P_{FN} is the number of false negative edge pixels in the edge map being evaluated, and d_{PFNi} is the distance between the i^{th} false negative edge pixel and the closest correctly detected edge pixel [107].

The two terms are then linearly fused based on the user specified weights according to a given application. For edge detections in surveillance or military applications, more weight should be applied on the false negative term, while for general consumer uses, users can avoid getting an edge map with too many fine details by applying a greater weighting coefficient on the false positive term. In the experimental results shown in this article, the two terms are fused by the uniform average:

$$D_p = \frac{D_{PFP} + D_{PFN}}{2} \quad (4.6)$$

It is worth noting that D_p in Eq. (4.6) indicates the distances of the false positive and false negative edge pixels. Therefore, the smaller D_p , the better edge pixel detection/localization performance is. In this section, $1 - D_p$ is used to show the edge detection/localization performance. The effectiveness of the D_p term on the LoG and Roberts edge detection results is shown in Figure 4-2. Visually, as seen in the red-circled part of Figure 4-2, the Roberts output detects and generates more accurate edge pixels than the LoG edge map. The new edge pixel presence/localization measure correctly labels the Roberts edge map as the better one, while the Pratt's FOM indicates the LoG edge map has better quality.

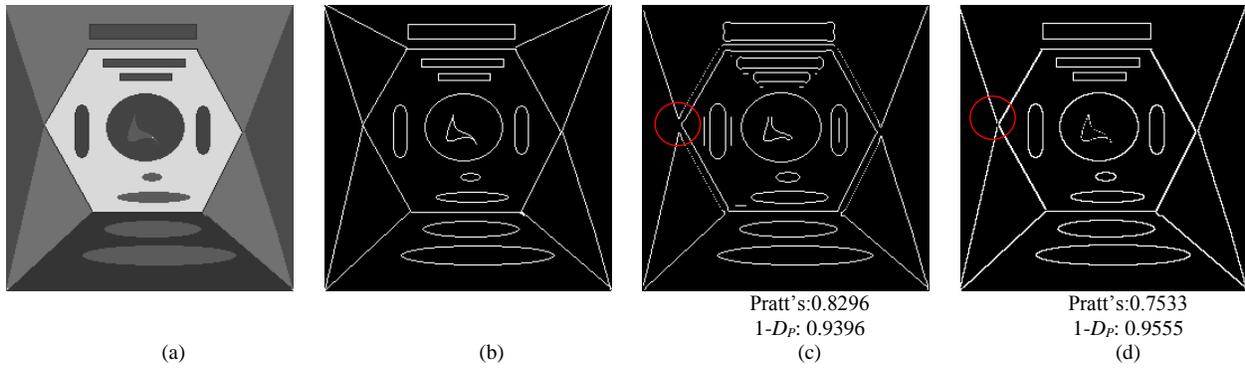


Figure 4-2. (a) Original image, (b) the ground truth edge map, (c) the LoG edge map, and (d) the Roberts edge map. The edge pixel detection/localization measure indicates the Roberts edge map is better than the LoG edge map.

4.3.2 Edge Corner Presence/Localization Measure

As a central problem in image processing and computer vision, corner detection has been an extensively researched topic in the past twenty years. Corner detection is used in the applications that require fast and efficient feature matching, such as robot localization and navigation tasks, where the use of corners for matching is preferred over edges or other larger features because corners contain gradient changes in multiple directions [115]. Consequently, corner detection has also been used extensively for tracking and motion estimation [116]. Therefore, it is required that the detected corners are close to their real locations. Accordingly, evaluation metrics for corner detection have been proposed, such as those based on accuracy and consistency measures [117]. In the context of edge detection, the edge corners produced by edge detectors are displaced relative to ground truth or can even be lost. This is particularly prevalent if a smoothing is used as a pre-processing step in an edge detection algorithm, such as the Canny edge detector, which uses a Gaussian filter to remove noise while simultaneously smoothing the sharp changes in corners [118]. Since corner is such an important visual feature seen by human observers, an objective edge evaluation measure should take such property into account. Accordingly, important criteria for the evaluation of edge map considered here are both the presence and localization of edge corners.

Generally, corners are not indicated directly from edge maps. Therefore, to compare the corner preservation and localization performance, a reliable corner detection algorithm is needed to abstract the corners from the ground truth and the testing edge map. In this article, edge corners are detected using a template matching technique. In the context of the template based algorithms, ideal corners are defined as sharp corners with at least two-pixel length at

corner directions. This definition reduces the false detections of spurious edges. Besides, the round corners are not considered as ideal corners.

Three basic templates are used in the corner detection and illustrated in Figure 4-3. 12 templates are then generated by rotating each of the masks in Figure 4-3 in steps of 90 degree. These 12 templates detect not only all possible 90 degree corners, including left hand or right hand corners, but also all possible 45 degree corners.

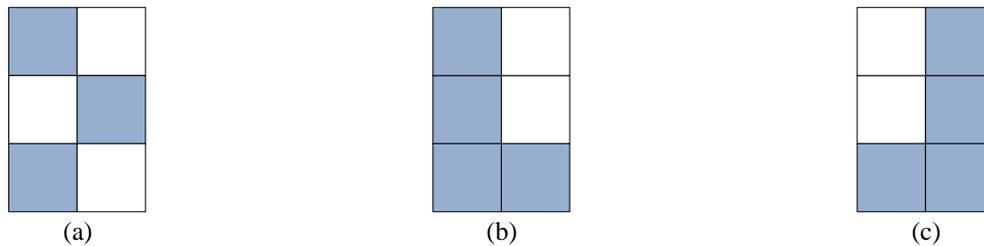
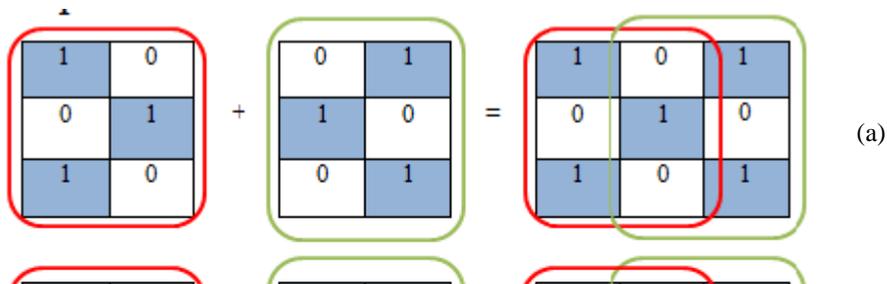


Figure 4-3. Basic templates used for corner detection.

The corner detection algorithm is image type dependent. For synthetic images, a corner is detected by directly matching the 12 templates with local regions in the testing image. If the two local patterns match, the center of the testing image local region is determined as a corner. This detection technique is practical for synthetic images because synthetic images have relatively clean textures and this method captures as much corner patterns as possible. For natural images, it is observed that human observers tend to capture major corners, while consider the fine details non-essential. If the same method is used to detect corners for natural images, many false positive corners are obtained especially when images are quantized or noise contaminated. Therefore, for natural images, we use a tighter condition to claim a pixel a corner. The tight detection method, on the other hand, recognizes a pixel as a corner only if at least two basic corner patterns preserve at that pixel location. Two examples of using the combinations of basic templates to represent ideal corner patterns are shown in Figure 4-4.



(b)

Figure 4-4. Examples of using two basic corner templates to formulate ideal corners: (a) ‘X’ shape, (b) ‘T’ shape

Similar as the edge pixel presence/localization measure, the corner presence/localization term also contains two terms: a false positive corner term which takes into account the number of false positive edge corners and their distance from the closest ideal edge pixel, as well as a false negative corner term which takes into account the number of false negative corners and their distance from the closest correctly detected corner [107]. For an $m \times n$ edge map, it is given by

$$D_{CFP} = \frac{1}{mn - C_O} \sum_{i=1}^{C_{FP}} 1 - \frac{1}{1 + \alpha d_{CFP_i}^2} \quad (4.7)$$

and

$$D_{CFN} = \frac{1}{C_O} \sum_{i=1}^{C_{FN}} 1 - \frac{1}{1 + \alpha d_{CFN_i}^2} \quad (4.8)$$

where C_O is the number of corner pixels in the ground truth, C_{FP} is the number of false positive corners in the edge map being evaluated, and d_{CFP_i} is the distance between the i^{th} false positive corner from the closest ideal corner, C_{FN} is the number of false negative corners in the edge map being evaluated, and d_{CFN_i} is the distance between the i^{th} false negative edge corner and the closest correctly detected edge corner [107].

Similarly, weighted averaging is used to fuse these two terms. Users have the flexibility to choose a specific tolerance for a given application. In this article, uniform average is used.

$$D_C = \frac{D_{CFP} + D_{CFN}}{2} \quad (4.9)$$

Figure 4-5 shows the effectiveness of the corner measure. It is seen that the corners in the LoG edge map are displaced. As a comparison, the Sobel edge map generates more precise edge corners. The new corner measure correctly selects the Sobel edge output outperforms the Log edge output with respect to the corner detection performance. However, the Pratt’s FOM for the LoG edge detection result is greater.

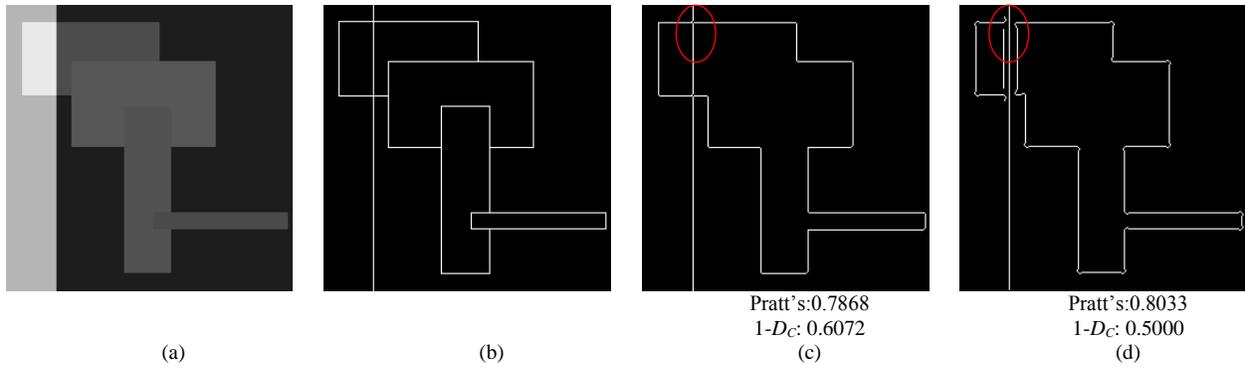


Figure 4-5. (a) Original image, (b) the ground truth edge map, (c) the Sobel edge map and (d) the LoG edge map. The corner detection/localization measure indicates the Sobel edge map is better than the LoG edge map.

4.3.3 Double Edge Occurrences Measure

The characterization of true double edges has been shown to be important for vision tasks. For example, it has recently been observed that the detection of true double edges is of prime importance in the assessment of vertebral injury or collapse that may be caused by osteoporosis and other spine pathology, and can provide a means of obtaining objective morphometric measurements on the spine [119]. The presence of false double edge occurrences, thus, degrades the quality of the edge analysis. Therefore, the presence of false double edges is an important criterion for edge map evaluation. False double edges exist when images are contaminated with low frequency noise or captured in a blurring environment. Double edge occurrences could also be introduced by the smoothing effects used in many edge detection algorithms, such as the Canny, Sobel, and Laplacian of Gaussian (LOG) [120]. It is understood that double edge occurrences are a special case of false positive edges, but due to their importance in computer vision, they are regarded as relevant and undesirable enough to be addressed separately.

It is assumed that double edges occur in adjacent locations. Therefore, downsampling can be used to expose double edge occurrences in edge maps. Four types of horizontal and vertical downsampling methods, and two types of diagonal downsampling methods are performed as shown in Eq. (4.10) to Eq. (4.15). The results of downsampling an edge map for a synthetic image by the various different methods are shown in Figure 4-6.

$$y_1[m, n] = x[2m, 2n] \quad (4.10)$$

$$y_2[m, n] = x[2m, 2n + 1] \quad (4.11)$$

$$y_3[m, n] = x[2m + 1, 2n] \quad (4.12)$$

$$y_4[m, n] = x[2m+1, 2n+1] \quad (4.13)$$

$$y_5[m, n] = x[m, 2n + m\%2] \quad (4.14)$$

$$y_6[m, n] = x[m, 2n + 1 + m\%2] \quad (4.15)$$

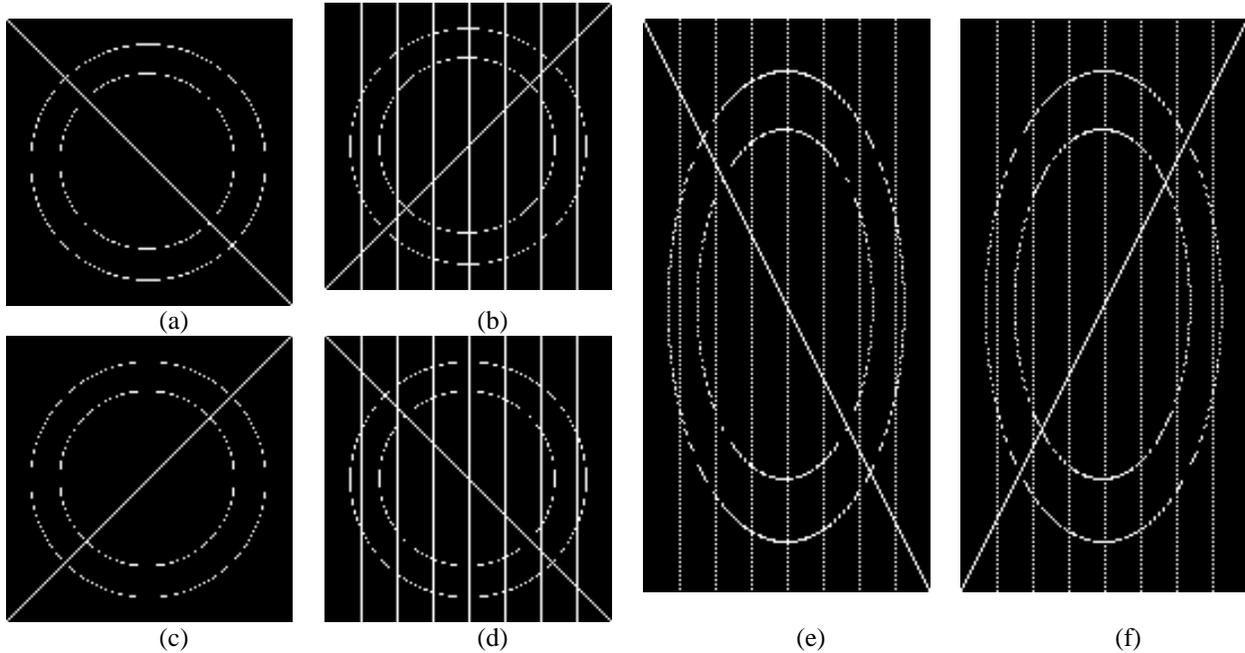


Figure 4-6. Downsampled images for the edge map in Figure 4-10 (a). (a)-(f): downsampling with y_1 to y_6 .

To guarantee the downsampled edge maps have the same size, the original edge maps are padded to ensure even row and column dimensions. The number of edge pixels in non-overlapping 2x2 windows is then determined for each downsampled image. Double edge locations are exposed by only considering the 2x2 windows in which the number of edge pixels is equal to 1 (this is referred as the box function used in Eq. (4.16) and Eq. (4.17) [107]. If double edges occur in a testing edge map, at least one of the six directional downsampled testing edge maps should have different number of edge pixels compared with the corresponding downsampled ground truth edge map. Figure 4-6 to Figure 4-9 show examples of using box functions to expose double edge occurrences when vertical edges, diagonal edges and 67.5 degree edges exist.

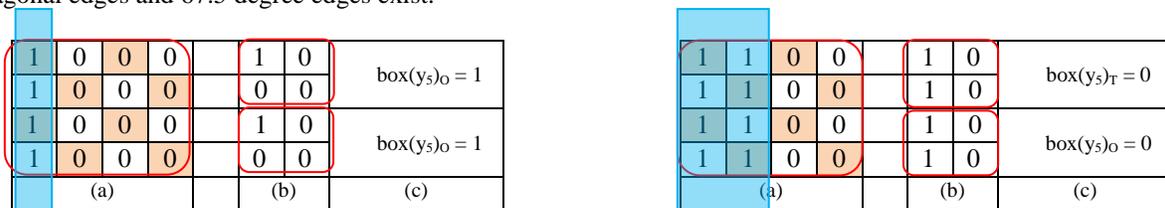


Figure 4-7. An illustrative example of using box functions to expose double edge occurrence in 90 degree. (a) a local 4x4 region (b) the y_3 downsampled 2x2 region, (c) the corresponding box function in that 2x2 region. The

edge pattern on the left has single pixel width, and the edge pattern on the right has double edges. Their box functions are different.



Figure 4-8. An illustrative example of using box functions to expose double edge occurrence in 45 degree. (a) an local 4x4 region (b) the y_3 downsampled 2x2 region, (c) the corresponding box function in that 2x2 region. The edge pattern on the left has single pixel width, and the edge pattern on the right has double edges. Their box functions are different.

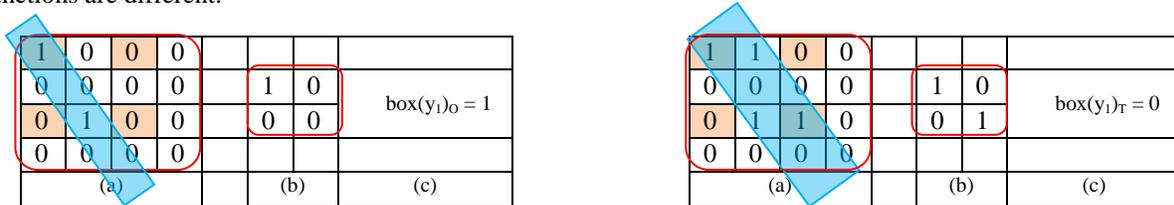


Figure 4-9. An illustrative example of using box functions to expose double edge occurrence in 67.5 degree. (a) an local 4x4 region (b) the y_3 downsampled 2x2 region, (c) the corresponding box function in that 2x2 region. The edge pattern on the left has single pixel width, and the edge pattern on the right has double edges. Their box functions are different.

Figure 4-10 shows an example of the box function map. The Roberts and the LoG edge maps are downsampled by y_1 as shown in Eq. (4.10), and the corresponding box functions of these two downsampled edge maps are also shown. Obviously, the Roberts edge map is more similar to the ground truth than the LoG edge map. This is also revealed in the downsampled box function map.

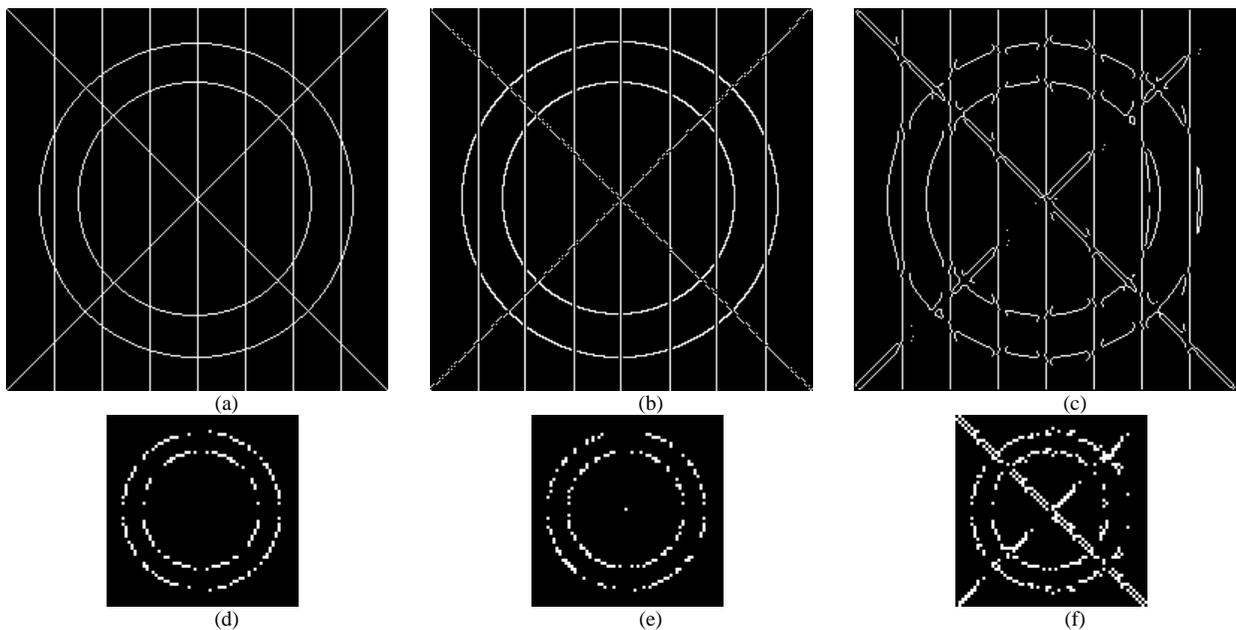


Figure 4-10. (a) Ground truth, (b) the Roberts edge map, (c) the LOG edge map, (d) the ground truth box function map, (e) the Roberts edge box function map, and (f) the LOG edge box function map. The Roberts output is more similar to the ground truth.

With these definitions, distances of each box function map d_{DE1} and d_{DE2} are defined as

$$d_{DE1} = (box(y_1)_O + box(y_2)_O + box(y_3)_O + box(y_4)_O) - (box(y_1)_T + box(y_2)_T + box(y_3)_T + box(y_4)_T) \quad (4.16)$$

$$d_{DE2} = (box(y_5)_O + box(y_6)_O) - (box(y_5)_T + box(y_6)_T) \quad (4.17)$$

where O denotes the ground truth and T denotes the testing edge map. The double edge term is then given by

$$D_{DE} = \frac{\frac{|d_{DE1}|}{4m_1n_1} + \frac{|d_{DE2}|}{2m_2n_2}}{2} \quad (4.18)$$

where m_1 and n_1 are the dimensions of an edge map after horizontal/vertical downsampling and m_2 and n_2 are the dimensions of an edge map after diagonal downsampling. For even dimensions m_1 and n_1 , $m_2n_2 = 2m_1n_1$, and in general, $m_2n_2 \approx 2m_1n_1$.

One example of using D_{DE} term measuring double edge detection performance is illustrated in Figure 4-11. Compared with the Roberts edge map, the Canny edge map contains distinct double edges due to the Gaussian smoothing used in the preprocessing step. The new double edge occurrence measure values correctly rank the performances of these two edge detection algorithms on this image.

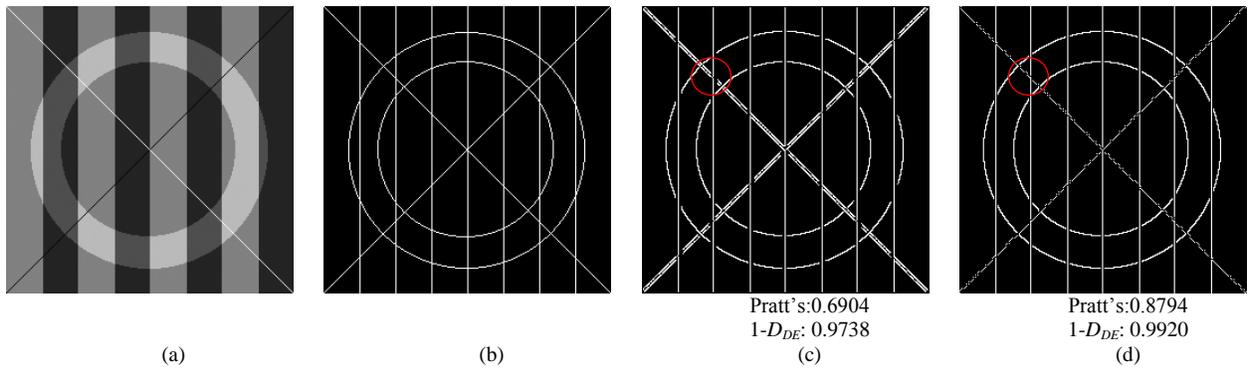


Figure 4-11. (a) Original image, (b) the ground truth edge map, (c) the Canny edge map, and (d) the Roberts edge map. The double edge occurrence measure indicates the Roberts edge map is better than the Canny edge map.

4.3.4 Reference Based Edge Measure RBEM

Finally, the new reference based edge measure RBEM is given by the weighted average of the three edge attribute measures:

$$RBEM = 1 - \frac{\beta_P D_P + \beta_C D_C + \beta_{DE} D_{DE}}{\beta_P + \beta_C + \beta_{DE}} \quad (4.19)$$

where β_P , β_C , and β_{DE} are weights for the edge pixel, edge corner, and double edge occurrence terms, respectively.

Due to its generalized nature, the new measure is applicable for a broader range of vision tasks than previously established measures. For example, an edge map used ultimately for motion detection or tracking may be more concerned with edge corner presence/localization, while an object detection system may be more concerned with edge pixel presence/localization or double edge occurrences. These inferences can be used to tailor the use of the measure for a given application. Practically, users can select these coefficients for specific applications by using training images that belong to the same type. In this paper, these coefficients are obtained using linear regression method for computer generated synthetic images and natural images randomly selected from the Berkeley segmentation database [121]. It is worth noting that the training image set and the testing image set are independent. In the experimental results presented in this paper, $\beta_P = 1.02$, $\beta_C = 0.53$, $\beta_{DE} = 6.24$ are used for synthetic images and $\beta_P = 0.27$, $\beta_C = 0.94$, $\beta_{DE} = 0.88$ are adopted for natural images in the RBEM measure.

4.4. Experimental Results

To evaluate the performance of the new RBEM edge measure and compare it with previous measures, experiments are conducted for both synthetic images and natural images. The synthetic images were designed to include straight lines oriented at various angles, curved lines such as ovals and circles, and various edge corner types [107]. The real images and the corresponding human marked ground truth edge maps are obtained from the published benchmark, the Berkeley segmentation dataset [121]. The selected natural images include varieties of scenes, such as buildings, portraits and animals. The RBEM edge measure is shown to have high correlation with human vision judgment in the applications of ranking edge detection performances of multiple edge detectors, and selecting the optimal operating parameters within one edge detector. Some of the experimental results are shown in this section.

First, the edge map evaluation methods are used to compare the Canny, Sobel, LOG and Roberts edge detector performances for a set of synthetic and natural images. In these experiments, $\sigma = 2$, threshold = 0.003 are used for the LOG edge detector. In the Canny detector, a Gaussian standard deviation $\sigma = 1$, fixed thresholds [0.005, 0.01]

are used for the synthetic images, and thresholds [0.05, 0.2] are used for natural images. The thresholds for the Sobel and the Roberts edge detector are set to 0.05. Evaluation measures using the new measure RBEM, Pratt's FOM and the Pinho's method are then calculated for each edge map and compared with subjective evaluation. As a reference, the subjective human evaluations are performed blindly by 48 researchers in the image processing field who are familiar with edge detection algorithms and performance. Namely, the edge maps are subjectively evaluated using the Mean Opinion Score (MOS). Each image is given a score from 1-10, where a 10 indicates the best quality. The average value of these scores is then taken to be the MOS for a given edge map. It should be noted that all the terms (D_P , D_C , D_{DE}) range from 0 to 1, and as in Pratt's FOM, a value of 1 corresponds to a perfect match using the new measure. However, due to the nature of the math involved, the individual terms (D_P , D_C , D_{DE} , etc.) which make up the new measure are such that a 0 corresponds to a perfect match for these terms. Figure 4-12 to Figure 4-14 show some testing synthetic images and the corresponding edge maps. Table 4-1 to Table 4-3 summarizes the results for these images, with the best performers according to a given objective measure and subjective MOS highlighted in bold. Figure 4-15 to Figure 4-17 show some edge maps for natural images. Table 4-4 to Table 4-6 summarizes the results for these images, with the best performers according to a given objective measure and subjective MOS highlighted in bold.

The results demonstrate that the RBEM measure values generally correspond better with subjective opinions using the MOS. In most cases, the proposed measure RBEM agrees with the MOS regarding the best performer, and in many cases, agrees with the relative order of performance. On the other hand, the best edge detectors for each image determined by the RBEM and the Pratt's FOM or the Pinho's method were dramatically different. One may note that Pratt's FOM and Pinho's method often times falsely select a busy edge map, or an edge map that contains double edges and smoothed corners as the best edge map. Both subjective evaluation and the presented RBEM measure indicate that the Roberts edge detector is the best edge detector for simple clean synthetic images. This selection is reasonable because there is no embedded smoothing used by the Roberts edge detector. Consequently, fewer displaced edges, smoothed corners, or double edge occurrences exist in the resultant edge map. Similarly, with proper parameter settings, the Canny edge map is considered the best for natural images by human observers and by the RBEM measure, as there are less fine details which are considered non-essential by human eyes in the Canny edge map with the parameters described previously .

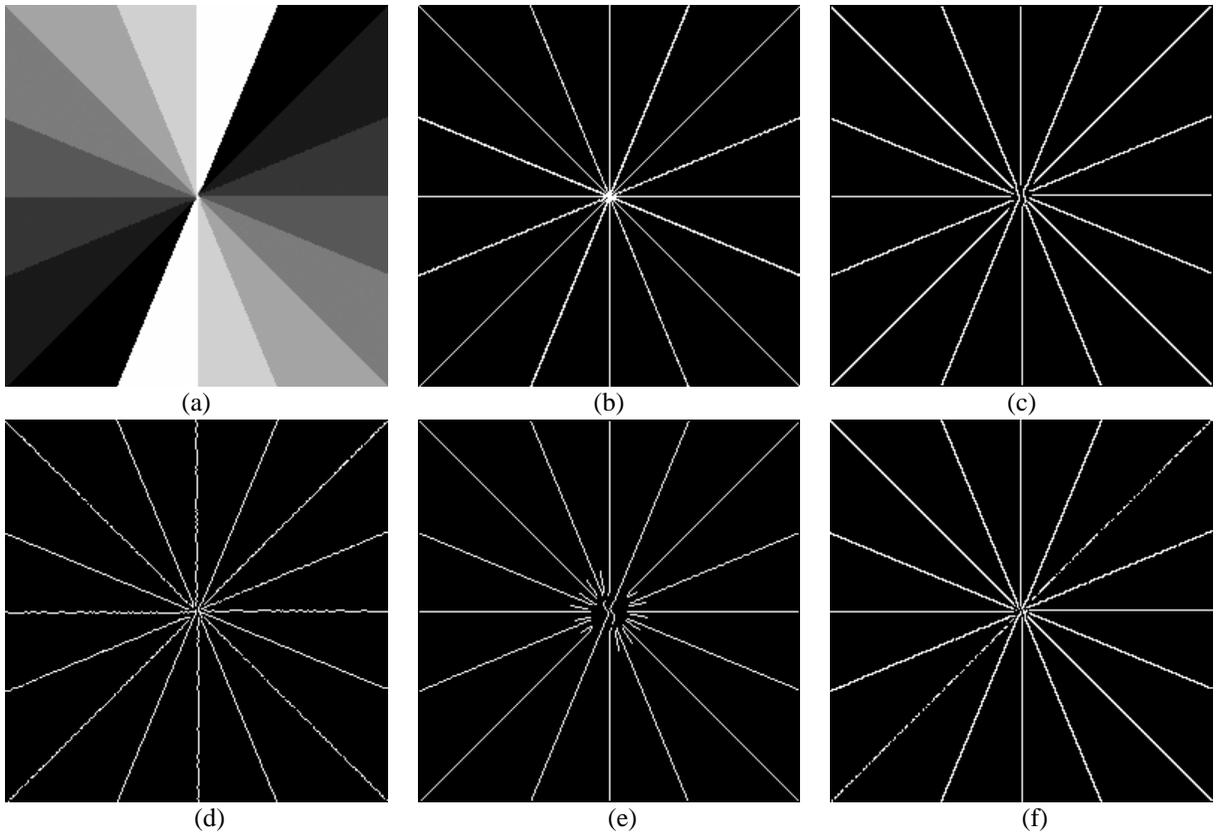
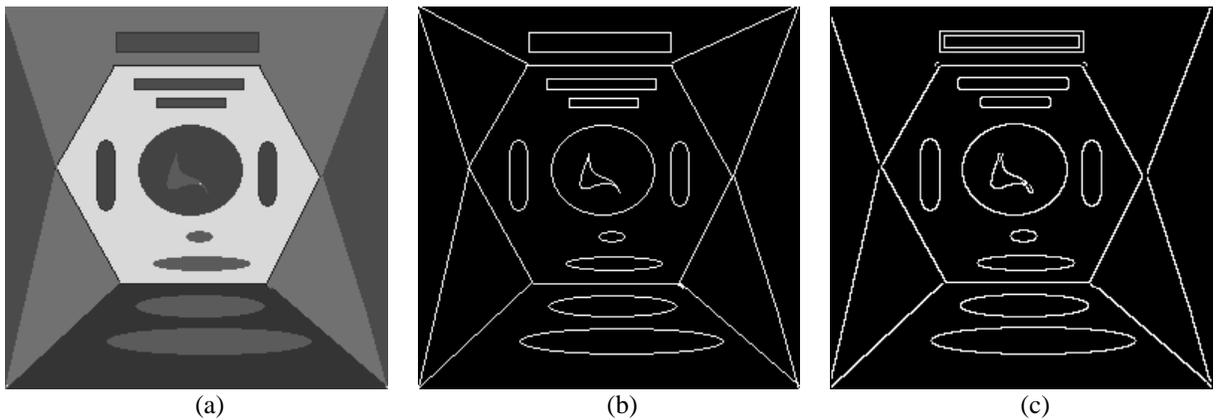


Figure 4-12. Edge detection results on a synthetic image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.

Table 4-1. Evaluation of various edge detectors for images in Figure 4-12.

Image	Pratt's	Pinho's	D_P	D_C	D_{DE}	RBEM	MOS
Figure 4-12. (c)	0.8072	0.4822	0.0260	0.0390	0.0078	0.9877	6.6190
Figure 4-12. (d)	0.9419	0.7759	0.0299	0.4620	0.0185	0.9499	8.0952
Figure 4-12. (e)	0.9660	0.5947	0.0174	0.4683	0.0076	0.9598	5.1190
Figure 4-12. (f)	0.8512	0.6361	0.0259	0.0382	0.0078	0.9878	8.6429



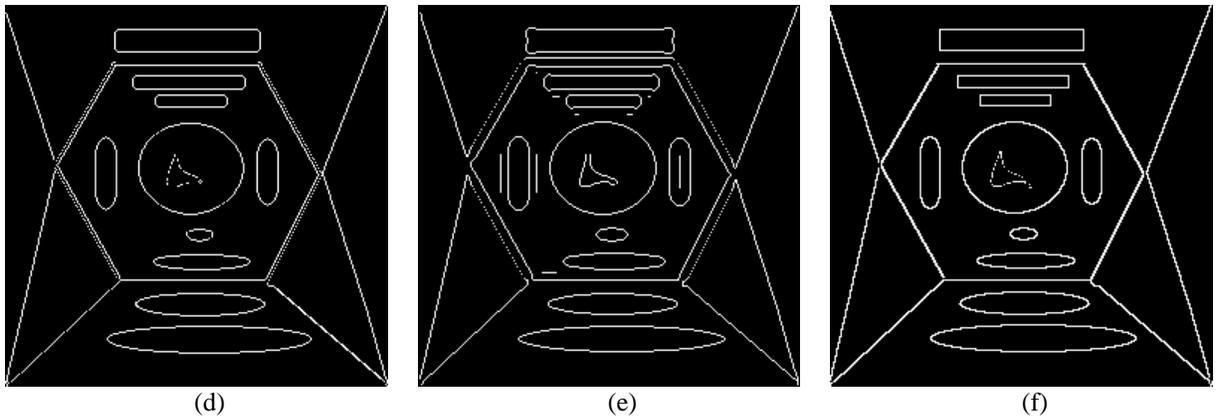


Figure 4-13. Edge detection results on a synthetic image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.

Table 4-2. Evaluation of various edge detectors for images in Figure 4-13.

Image	Pratt's	Pinho's	D_P	D_C	D_{DE}	RBEM	MOS
Figure 4-13. (c)	0.7299	0.5379	0.0475	0.0916	0.0111	0.9787	7.9286
Figure 4-13. (d)	0.8843	0.6481	0.0595	0.2965	0.0040	0.9689	7.1190
Figure 4-13. (e)	0.8296	0.6995	0.0604	0.3923	0.0029	0.9631	6.1190
Figure 4-13. (f)	0.7533	0.6880	0.0445	0.0749	0.0126	0.9790	8.4524

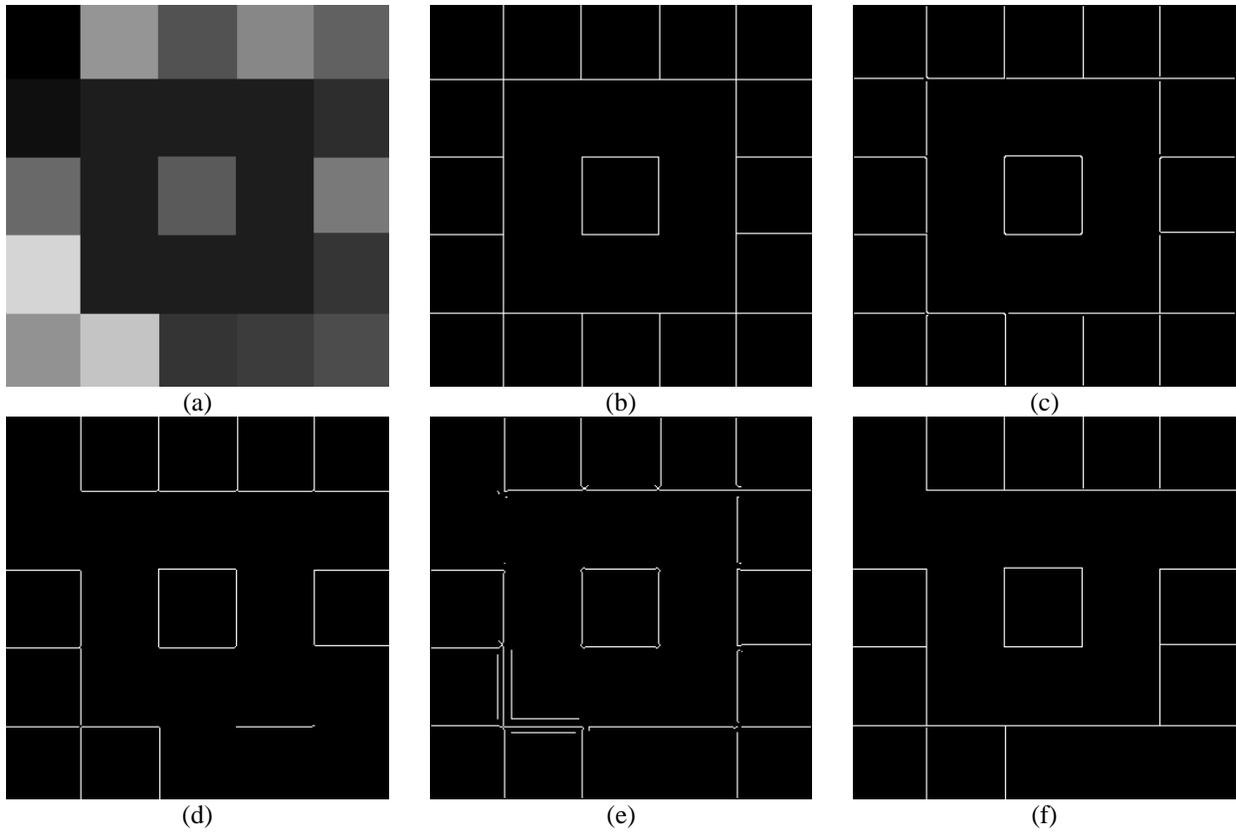
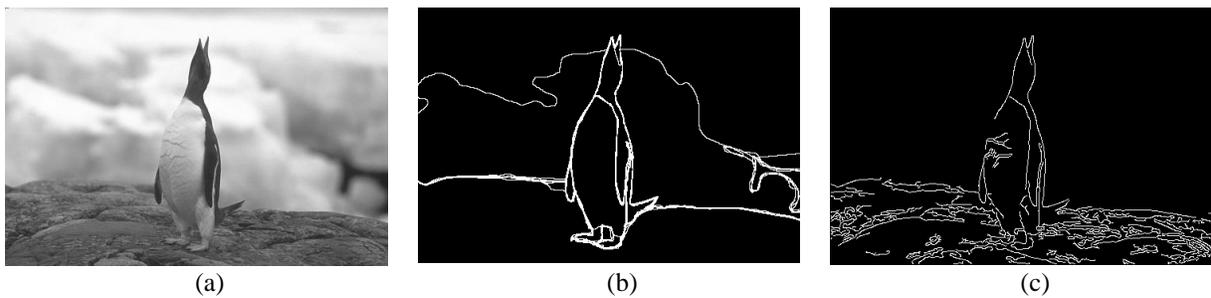


Figure 4-14. Edge detection results on a synthetic image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.

Table 4-3. Evaluation of various edge detectors for images in Figure 4-14.

Image	Pratt's	Pinho's	D_P	D_C	D_{DE}	RBEM	MOS
Figure 4-14. (c)	0.9731	0.3101	0.0137	0.1217	0.0009	0.9892	8.7143
Figure 4-14. (d)	0.7217	0.4377	0.1397	0.3276	0.0029	0.9571	4.3571
Figure 4-14. (e)	0.8557	0.8006	0.0718	0.3692	0.0018	0.9641	7.1667
Figure 4-14. (f)	0.8368	0.6238	0.0818	0.1123	0.0016	0.9803	5.5952



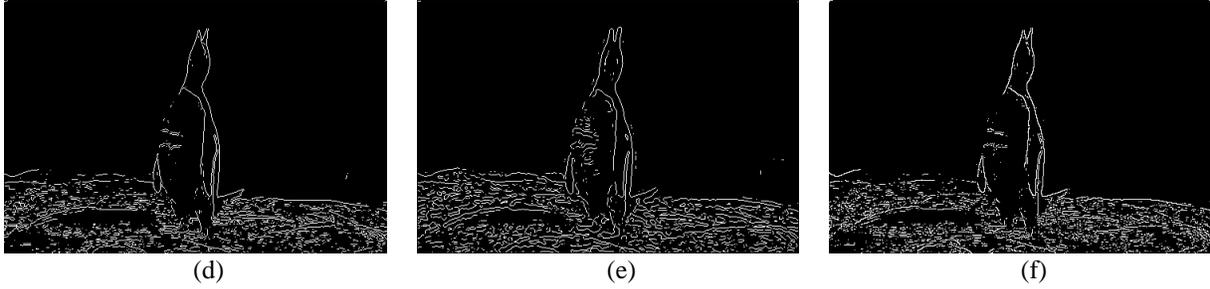
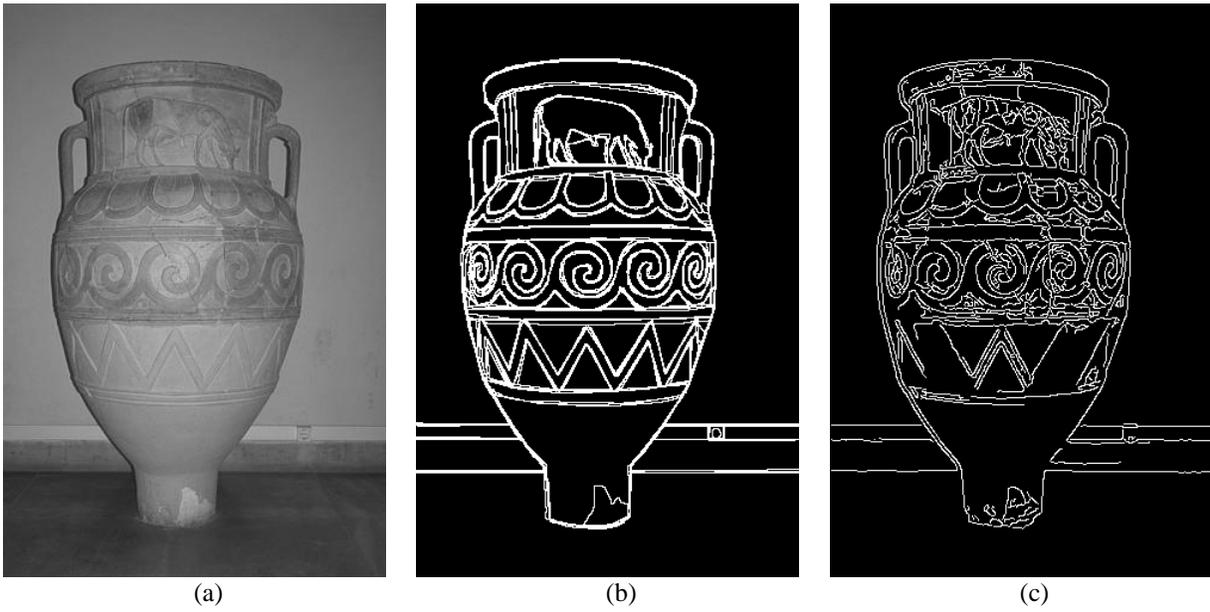


Figure 4-15. Edge maps for natural penguin image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.

Table 4-4. Evaluation of various edge detectors for images in Figure 4-15.

Image	Pratt's	Pinho's	D_P	D_C	D_{DE}	RBEM	MOS
Figure 4-15. (c)	0.5662	0.1627	0.2171	0.4424	0.0231	0.7633	6.1667
Figure 4-15. (d)	0.5523	0.1386	0.2279	0.4603	0.0246	0.7532	5.6042
Figure 4-15. (e)	0.5306	0.2149	0.2201	0.4811	0.0376	0.7394	4.6667
Figure 4-15. (f)	0.5041	0.1950	0.2322	0.4598	0.0277	0.7515	5.7500



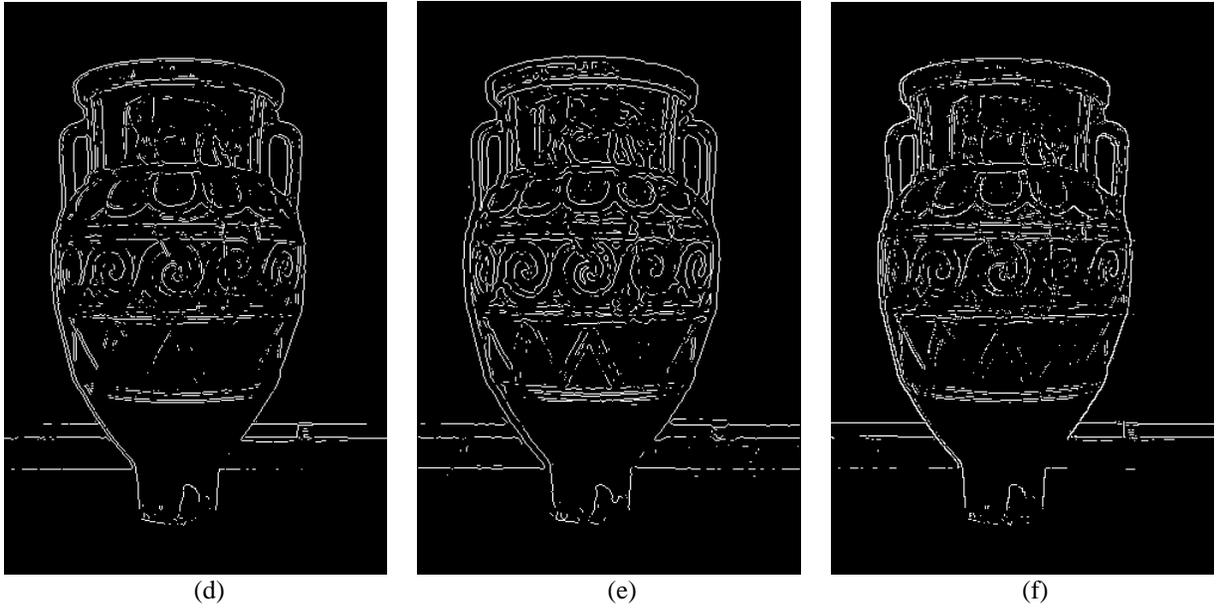
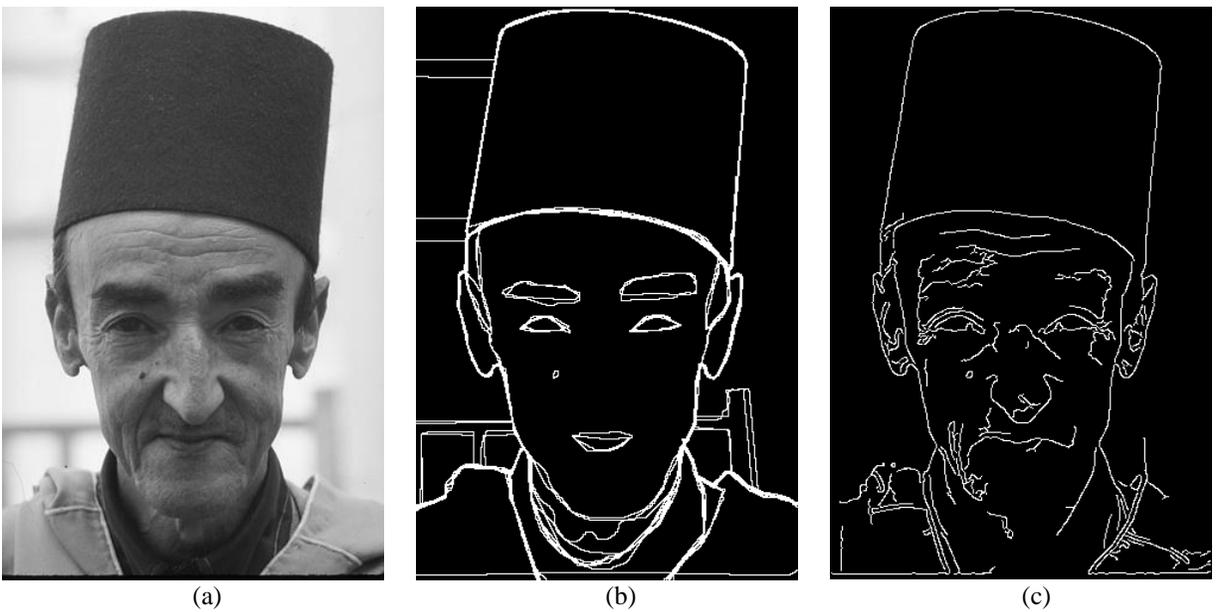


Figure 4-16. Edge maps for natural vase image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.

Table 4-5. Evaluation of various edge detectors for images in Figure 4-16.

Image	Pratt's	Pinho's	D_p	D_C	D_{DE}	RBEM	MOS
Figure 4-16. (c)	0.8597	0.1847	0.0773	0.3812	0.0546	0.7956	6.3404
Figure 4-16. (d)	0.7407	0.1859	0.1318	0.4663	0.0511	0.7517	5.2340
Figure 4-16. (e)	0.8021	0.3191	0.1032	0.4691	0.0617	0.7497	6.2553
Figure 4-16. (f)	0.7793	0.2259	0.1124	0.4531	0.0509	0.7603	5.5957



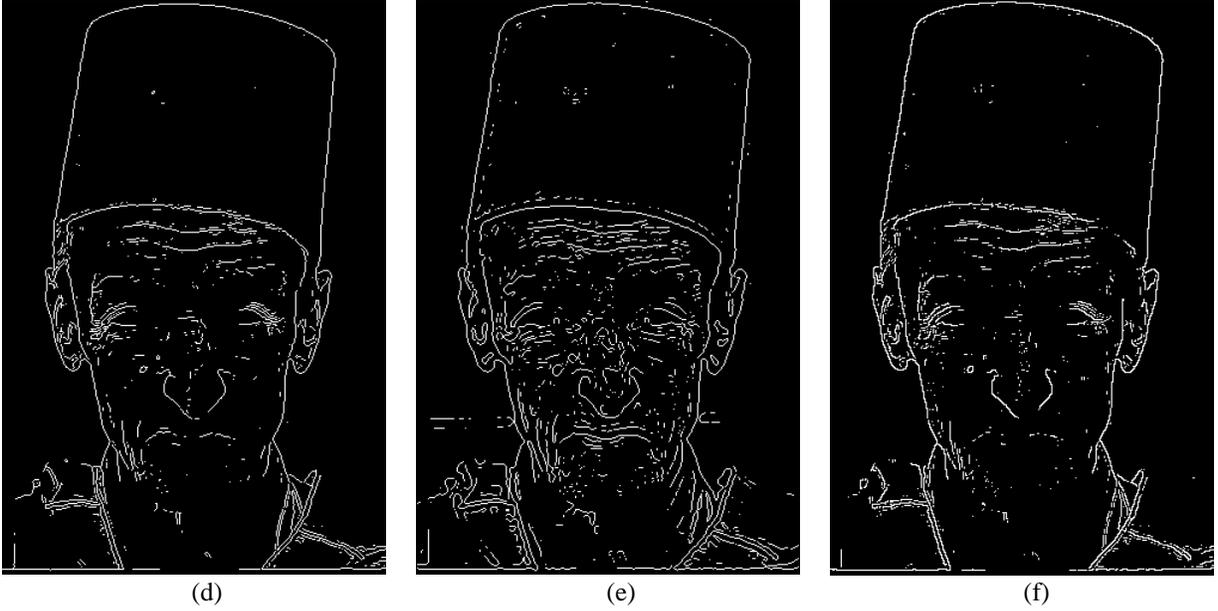


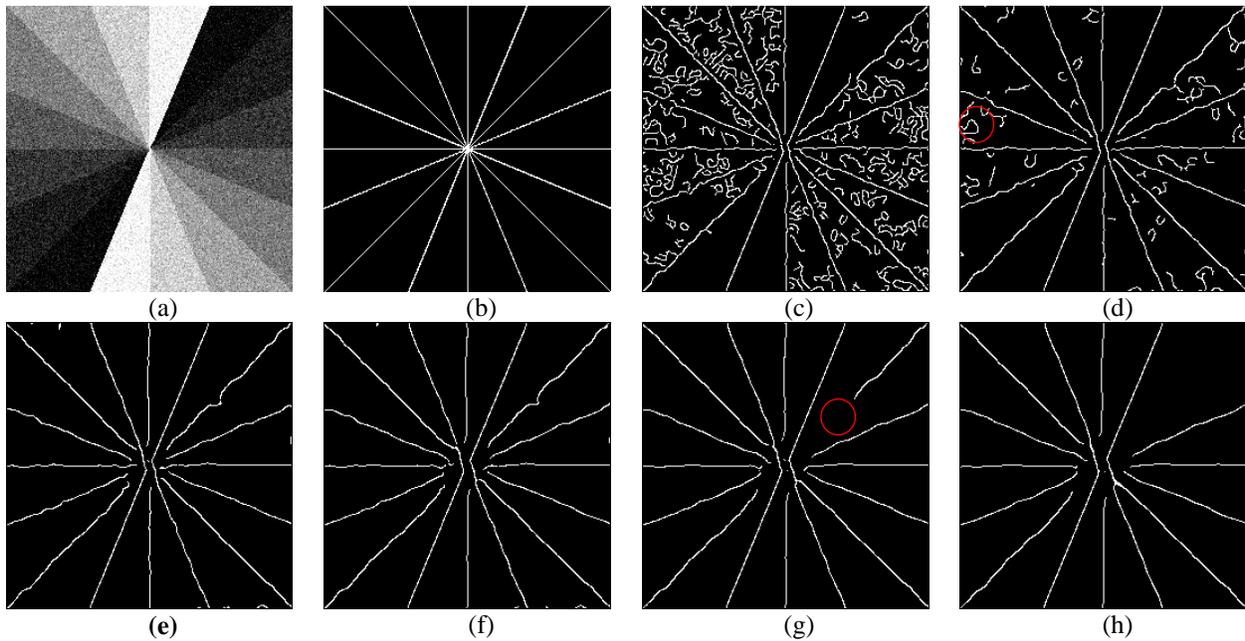
Figure 4-17. Edge maps for natural man image. (a) original image; (b) ground truth edge map; (c) Canny edge map; (d) Sobel edge map; (e) LoG edge map; (f) Roberts edge map.

Table 4-6. Evaluation of various edge detectors for images in Figure 4-17.

Image	Pratt's	Pinho's	D_P	D_C	D_{DE}	RBEM	MOS
Figure 4-17. (c)	0.6752	0.1737	0.1702	0.4194	0.0335	0.7753	7.1190
Figure 4-17. (d)	0.6144	0.1591	0.1994	0.4583	0.0322	0.7546	6.6190
Figure 4-17. (e)	0.6672	0.2353	0.1775	0.4867	0.0441	0.7396	4.0952
Figure 4-17. (f)	0.6171	0.2208	0.1980	0.4302	0.0285	0.7689	6.5952

The edge map evaluation methods could also be used to select optimal operating parameters within a single edge detector. Two illustrative experiments are shown in Figure 4-18 and Figure 4-19, performed on a synthetic image when noise is presented, and on a clean natural image respectively.

In the first experiment, Gaussian noise with zero mean and 0.01 standard deviation is added to a synthetic image. The Canny edge detector is applied on this image with various values for the Gaussian standard deviation σ in the smoothing step. It is known that the Gaussian smoothing used in the Canny edge detector can effectively suppress noise [114], but also introduce blurring effect which causes round corners and double edge occurrences. The optimal value of σ should find a balance between noise removal and edge preservation. Some Canny outputs with changing σ are shown in Figure 4-18. The best performer according to each evaluation metric is highlighted in bold in Table 4-7. It is seen that the RBEM measure outperforms other edge measures in the means that it agrees with human judgment in selecting $\sigma = 3$ as the best parameter. It is observed that a smaller σ is not able to remove noise effectively, while a greater σ losses important edge features.



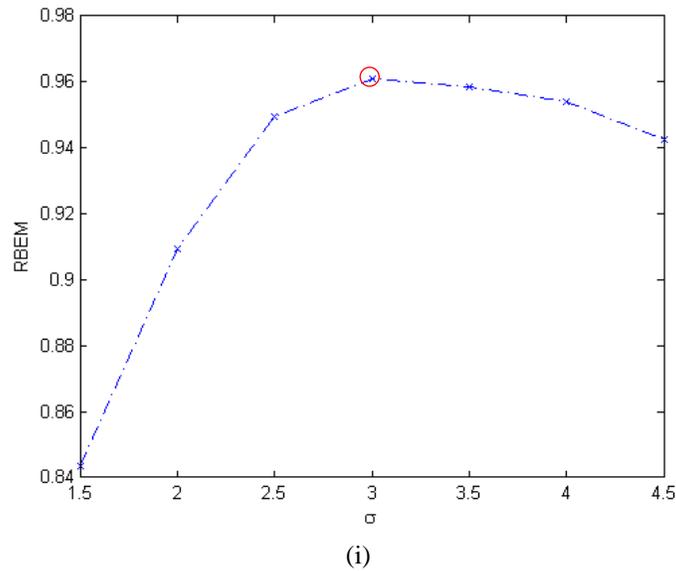


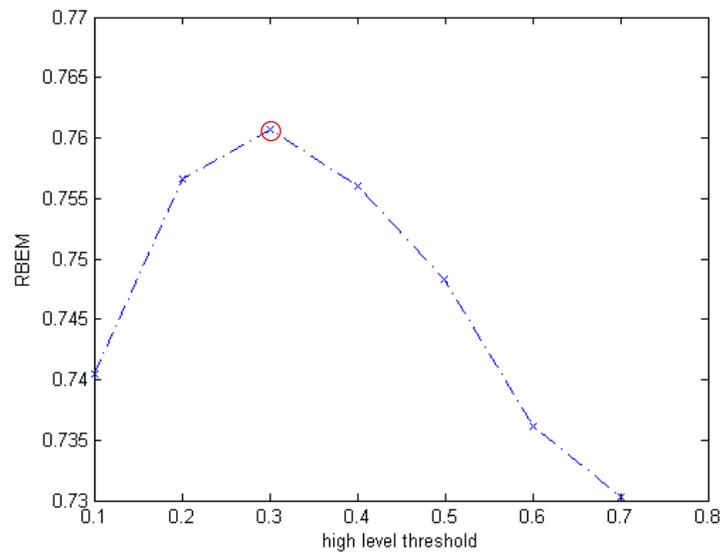
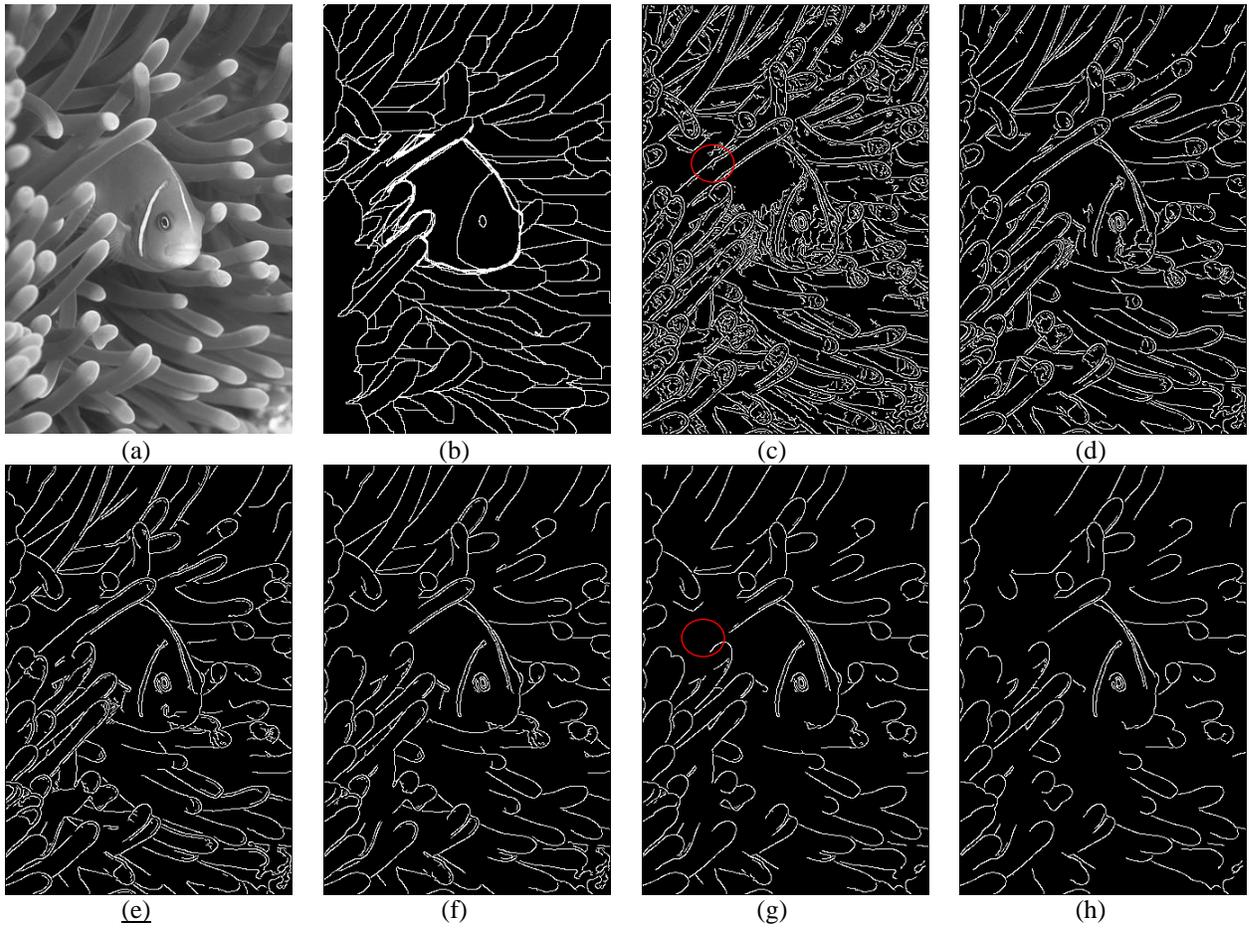
Figure 4-18. (a) Original image with Gaussian noise added, (b) ground truth edge map, edge maps obtained using the Canny edge detector with (c) $\sigma = 2$, (d) $\sigma = 2.5$, (e) $\sigma = 3$, (f) $\sigma = 3.5$, (g) $\sigma = 4$, (h) $\sigma = 4.5$, (i) RBEM measure plot indicates a maximum at $\sigma = 3$.

Table 4-7. Evaluation of the qualities of edge maps in Figure 4-18.

σ	Pratt's	Pinho's	RBEM	Boaventura	MOS
2.0 (c)	0.3134	0.2278	0.9096	1.4044	4.6957
2.5 (d)	0.6036	0.4059	0.9496	1.1643	5.9149
3.0 (e)	0.8679	0.5028	0.9611	1.0449	7.9130
3.5 (f)	0.8674	0.5031	0.9585	1.0383	7.8936
4.0 (g)	0.8737	0.5054	0.9541	1.0300	6.3404
4.5 (h)	0.8325	0.4756	0.9427	1.0429	6.2553

In the next experiment, the measures are used to select the high level threshold in the Canny edge detector for a clean natural image. The Canny edge detector uses hysteresis thresholds which require two thresholds, high and low. Pixels with intensity greater than the high level threshold are classified as genuine edge pixels, while pixels with intensity smaller than the low threshold are classified as non-edge pixels directly. Pixels with intensities between these two thresholds are traced by other conditions. Generally, using a smaller low level threshold reserves more edge pixels in the final edge map. In this experiment, low level threshold is set as 0.4 times high level threshold. From the results, it is seen that the RBEM measure selects the optimal high level threshold equals to 0.3, which is in

correspondence with the human observation. A smaller high level threshold reserves too many fine details and a greater high level threshold loses essential edges.



(j)

Figure 4-19. (a) Original natural image, (b) ground truth, edge maps obtained using Canny with high level threshold (c) $Th = 0.1$, (d) $Th = 0.2$, (e) $Th = 0.3$, (f) $Th = 0.4$, (g) $Th = 0.5$, (h) $Th = 0.6$, (i) $Th = 0.7$, (j) RBEM measure plot indicates a maximum at $Th = 0.3$.

4.5. Summary

In this chapter, a new reference-based edge map quality measure RBEM is presented. The RBEM incorporates three edge characteristic terms, namely the edge pixel presence/localization, corner pixel presence/localization, and double edge occurrences. Each term can be used as a standalone measure, which assesses one aspect of edge detection performance. The experimental results demonstrate that the RBEM measure outperformed previously established measures as a means of ranking the performance of many edge detectors, as well as selecting parameters within a single edge detector, for both synthetic images and natural images, with and without noise contamination. In the experimental test, the presented RBEM measure was more in accordance to the subjective opinions gathered via the MOS ranking test relative to other edge detection evaluation metrics.

Chapter 5. Non-reference Based Edge Measure

In this chapter, a reconstruction based non-reference edge measure NREM is presented. The theoretical basis of this method is that a good edge map captures the essential structures and details of the original images. Therefore, the reconstruction using the pixel information on a better edge map would be more similar to the original image. In this method, the edge measure is composed of two components: the first is the gradient based structural similarity measure between the original image and the reconstructed image, and the second component is the penalty factor that accounts the number of pixels in the edge map used for reconstructing the grayscale image. The NREM measure is used to compare the edge detection performances and direct the selection of optimal operating parameters used in the edge detection algorithms.

5.1. Overview

Many edge detection algorithms have been proposed. Although the performance of most of these detectors is acceptable for simple noise free images, the case is dramatically different when noise exists. Unfortunately, noise-free or distortion free is impractical for real applications. Figure 5-1 shows an example of a medical image. Medical images usually suffer from low contrast or poor resolution due to the limitation of hardware systems or the exposure time. Therefore, it is necessary to have a reliable evaluation method to measure the performance of different edge detection algorithms and help select the optimal algorithm for specific applications.

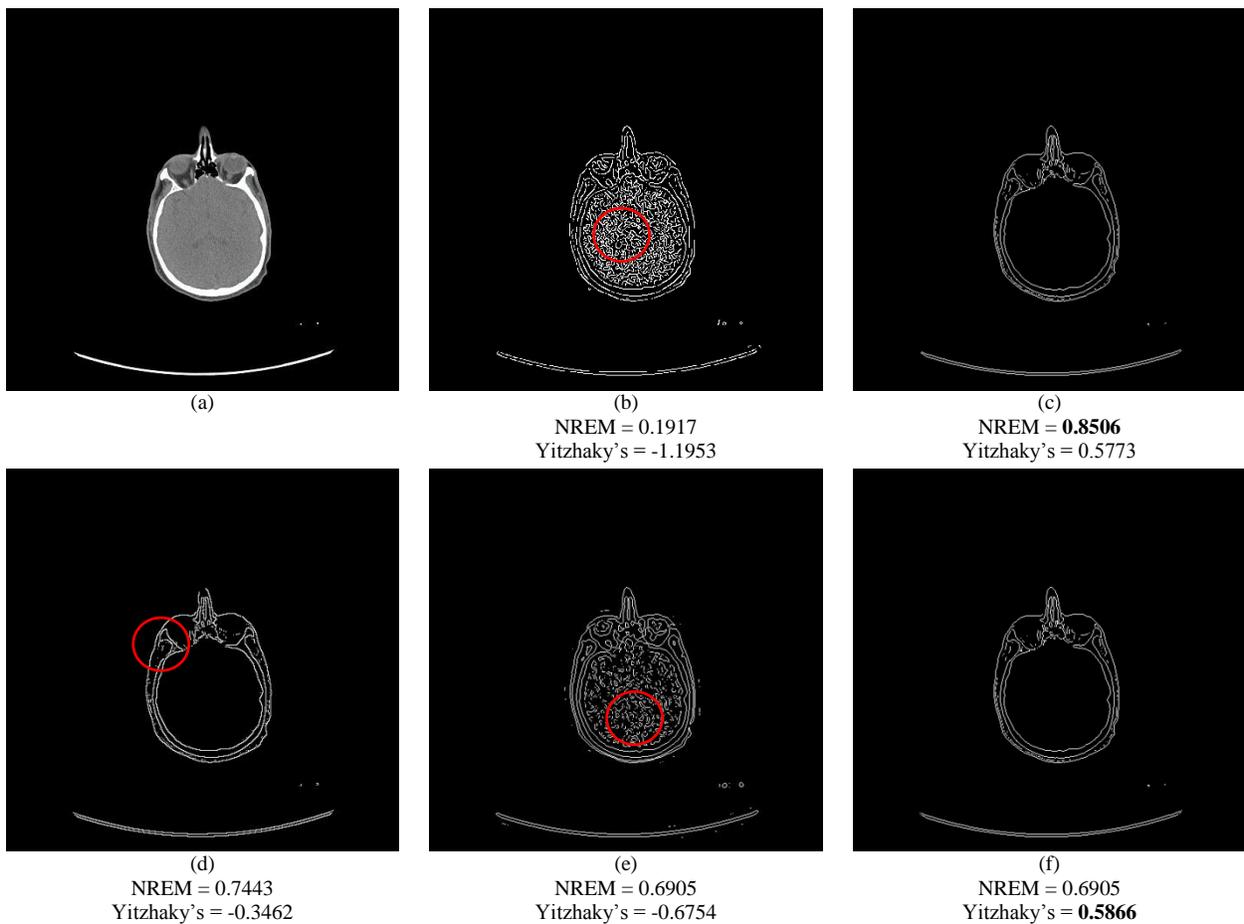


Figure 5-1. Non-reference edge measure: which edge map is the best without consulting medical experts' opinion? The presented measure NREM indicates the second edge map is the best.

Many edge measures have been proposed including the full reference objective edge measures [122], the non-reference objective measures [123][124], and the subjective evaluations. The full reference edge measure requires a ground truth image as a reference and compares the similarity between candidate edge maps and the ground truth

edge map. However, the ground truth image is not always available for many practical applications. Subjective evaluation avoids the use of ground truth edge maps. However, it is impossible to remove all the bias and the results can still be inconsistent. Furthermore, subjective evaluation is expensive with respect to time and resources, thus it is difficult to be automated. The non-reference based method does not require a ground truth and it can be automated. Unfortunately, the existing non-reference edge measures are still far from ideal. A non-reference based edge map evaluation should only use the information from the resultant edge map and the original image itself to make an evaluation. In this chapter, a non-reference edge quality measure NREM is discussed.

5.2. Yitzhaky and Peli's Edge Measure

The research in the non-reference edge measure is still a challenge because the task of edge quality evaluation is very subjective without a ground truth edge map. Yitzhaky and Peli proposed an probabilistic based non-reference objective edge detection evaluation method utilizing the receiver operating characteristics (ROC) in [124]. In this chapter, the new measure NREM is compared with this method.

Yitzhaky and Peli's method performs statistical analysis of the correspondence of detection results produced using different detector parameters. The statistical measures are used to estimate the ground truth edge map considering the tradeoff between the true and false edges, and to extract the best detector's parameter sets. In their method, first, an estimated ground truth is automatically constructed by examining the Corresponding Threshold Receiver Operating Characteristics (CT-ROC) curve, given a range of detection results obtained from different detection parameter sets. This is conducted under the assumption that the best detection of a certain edge detector in a given image is that which is most consistent with the variety of detection outputs that can be produced by the detection algorithm when different parameters are used. Then, the single parameter set that provide the most similar edge map to the estimated ground truth edge map is identified as the optimal parameter set. Similarly, if the edge detection results for estimating the ground truth are obtained from different edge detectors, this algorithm can be used to compare performances between these detection approaches. The major steps of the Yitzhaky and Peli's edge measure are summarized in Table 5-1.

Table 5-1. Basic steps in the Yitzhaky and Peli's NR edge measure

Yitzhaky's Edge Measure
1: Generate N edge detection results $D_i, i = 1, \dots, N$ using N combinations of parameters
2: Generate N potential ground truth $PGT_i, i = 1, \dots, N$
3: Calculate the average True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) rate for each potential ground truth
4: Construct the Correspondence Threshold ROC curve (CT-ROC)
5: Extract the estimated ground truth using either a diagnosis line or the chi-square estimation
6: Select the best edge map which gives the best match to the estimated ground truth

Yitzhaky and Peli's method balances specificity and sensitivity. However, this method suffers from its bias regarding the generation of the estimated ground truth, because the candidate edge maps used can directly affect the estimated ground truth. Therefore, if the majority of the edge maps used are not of adequate quality or fail to extract certain features, this will be reflected in the derived estimated ground truth. Also, since the original image data is not used, there is no way to indicate how well the best determined edge detector output from this approach corresponds to the original image.

5.3. New RBEM edge map measure

A generalized block diagram and intermediate results of the established Non-reference Reconstruction based Edge Measure (NREM) is shown in Figure 5-2. It consists of three major steps: grayscale edge map generation, reconstruction and similarity measure.

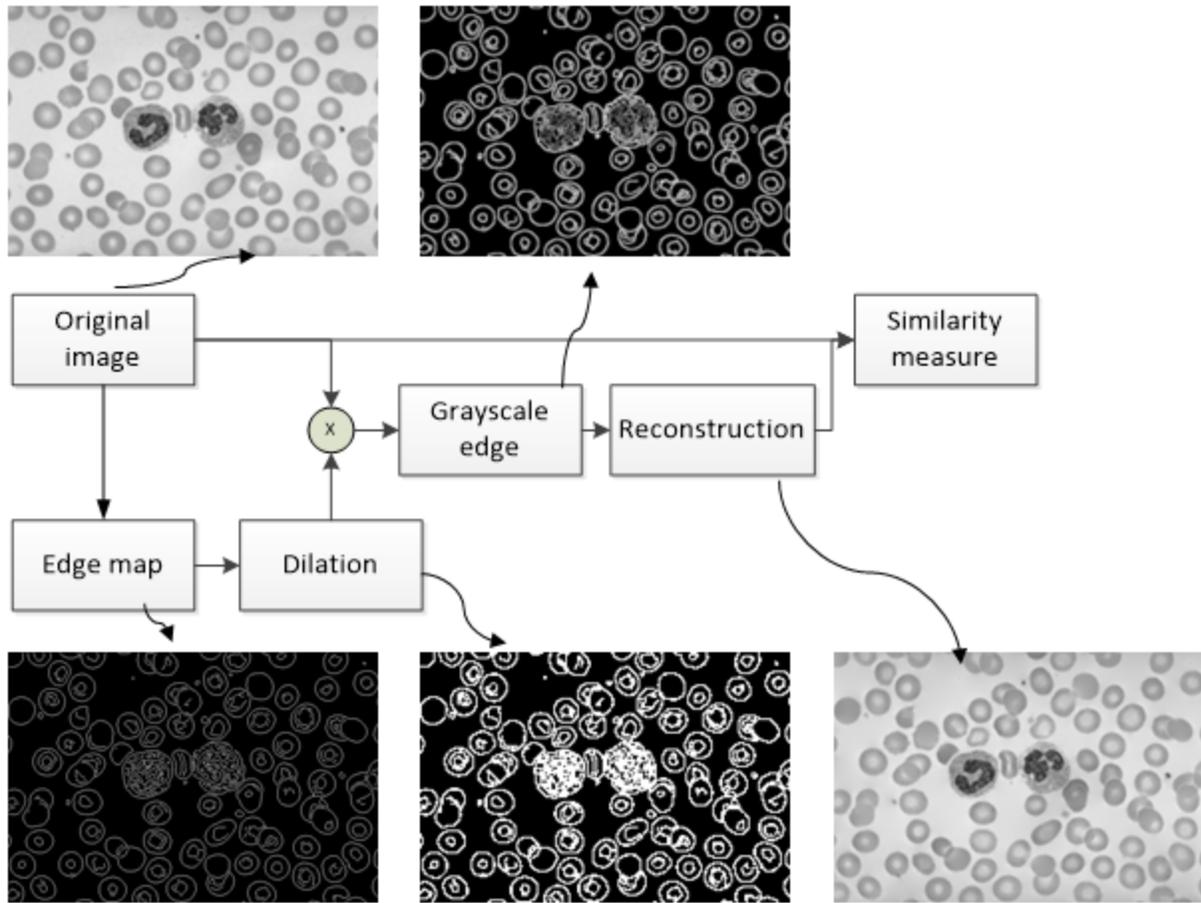


Figure 5-2. Flow and the intermediate results of the new non-reference edge measure NREM.

5.3.1 Grayscale Edge Map Generation

First, a morphological dilation is applied on the input edge map to generate a continuous edge map. The dilated edge map is then multiplied with the original image, yielding a grayscale edge map. In this way, the pixels on the dilated edge map contain information from the original image and these pixels are used to predict the pixel intensity in the smooth area.

5.3.2 Reconstruction

Interpolation has been used to obtain the missed pixels from the known pixels. In the NREM, the pixels along edges are used to predict the pixel values in the smooth areas. Several reconstruction methods can be used:

(1) Weighted average: For each pixel at location $(i, j) \notin e_D$, the algorithm searches in the four horizontal and vertical directions, and four diagonal directions for the nearest pixel in the given direction that $\in e_D$ [125]. The inverse of the distances of the first pixel encountered in each direction from the given pixel d_k are then used as the weights for the weighted average of their respective image intensity values t_k , yielding the reconstructed intensity value for the given pixel. Thus, reconstruction is carried out for each pixel location $(i, j) \notin e_D$ by Eq. (5.1):

$$r(i, j) = \frac{\sum_{k=1}^8 \frac{1}{d_k} t_k}{\sum_{k=1}^8 \frac{1}{d_k}} \quad (5.1)$$

For images with heavy noise, weighted median is usually used instead of the weighted mean to reduce the influence of noisy pixels.

(2) Central weighted median: The central weighted median of a sequence x with weights w is given by Eq. (5.2), where the weights are inversely related with the distances (Eq. (5.3)), and the \diamond is the replication operator representing the intensity value t_i repeats w_i times in the sequence of median calculation in Eq. (5.2).

$$r(i, j) = \text{median}(t_1 \diamond w_1, t_2 \diamond w_2, \dots, t_8 \diamond w_8) \quad (5.2)$$

$$w_k = \text{round}\left(\frac{100}{d_k}\right) \quad (5.3)$$

(3) Alpha-trimmed weighted mean: The weighted average (Eq. (5.1)) method utilizes all the information from the eight neighbors but is sensitive to noise. Unfortunately, noise commonly exists in many image applications. To be more robust to noise, the weighted median and central weighted median (Eq. (5.2)) can be used. In this way, only one of the neighbors is used to predict the new pixel value. This replication solves the noise problem but also results in another problem. That is, in some areas with a low gradient change, such as the breast tissue in the mammogram image, this reconstruction may mistakenly yields a large uniform region.

To get a good balance, a weighted alpha trimmed mean can be used in the reconstruction. Each of the eight neighbors, is assigned with the weighted intensity $x_k = \frac{1}{d_k} t_k$, where t_k is the actual edge pixel intensity, and d_k is the distance between the pixel to be predicted and edge pixel on a specific direction. Then sort values of all the

neighbors in ascending order such that $x_1 \leq x_2 \leq \dots \leq x_K$. Let $T_\alpha = \lceil \alpha K \rceil$ (the nearest integer greater than or equal to αK) be the number of the smallest and largest pixel values to be trimmed or discarded from the sorted sequence x_1, x_2, \dots, x_K . The alpha trimmed mean [126] is defined by

$$X_\alpha = \frac{1}{K - 2T_\alpha} \sum_{i=T_\alpha+1}^{K-T_\alpha} x_i \quad (5.4)$$

With the alpha trimmed mean, the reconstructed pixel with an average of a portion of the neighbouring pixels in eight directions can be expressed as:

$$r(i, j) = \frac{\sum_{k=T_\alpha}^{8-T_\alpha} \frac{1}{d_k} t_k}{\sum_{k=T_\alpha}^{8-T_\alpha} \frac{1}{d_k}} \quad (5.5)$$

The alpha-trimmed mean will be different when the parameter α changes. For example, it will be the mean value of the image for $\alpha = 0$ and the median value of the image if α is close to 0.5. In this way, the parameters can be tuned for different applications. In this chapter, the results are obtained by discarding the maximum and minimum neighbors in the calculation of alpha weighted mean.

These reconstruction based methods have clear advantages for effectively incorporating the original image information on the edge pixels. In non-reference measures this information is essential because no ground truth exists. Figure 5-3 (a) shows a CT kidney image obtained from the Frederick National Library for Cancer Research Database [6] and five reconstructions from the Canny edge detector using the weighted mean, weighted median, central weighted median, PDE [127] and alpha trimmed weighted mean based interpolation are demonstrated. It is seen from the results that for the medical image, the mean based reconstruction is not as sharp as the median based reconstruction. However, the median based reconstruction introduces some artificial lines. As analyzed before, the alpha trimmed mean can be converted to mean or median filter with different parameter alpha. In Figure 5-3, the alpha trimmed weighted mean reconstruction achieves a good balance. The central weighted median retrieves details but also introduces false details especially near the edge of the real tissues. The PDE based inpainting method suffers from severe blurring effects.

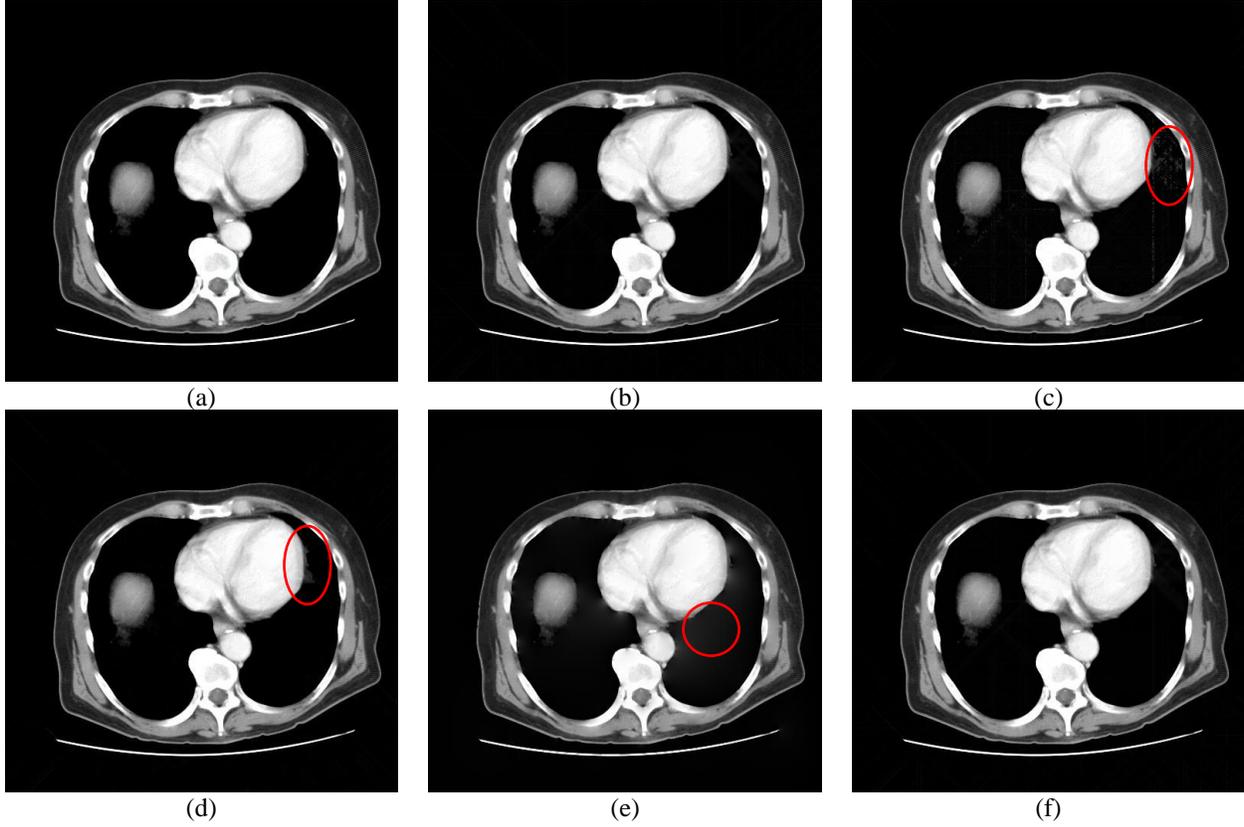


Figure 5-3. Reconstruction results using different interpolation methods. (a) Original CT kidney image. (b)-(f) reconstructed results from: (b) weighted average, (c) weighted median, (d) central weighted median, (e) PDE [127], (f) alpha trimmed weighted average.

5.3.3 Similarity Measure NREM:

The reconstructed image is then compared to the original image using a similarity measure, which is then used as an assessment of the edge map. To compare the similarity between two images, the most commonly adopted methods are the statistical methods including the pixel-wised Mean Square Error (MSE) and Mean Absolute Error (MAE). The MSE and MAE between two images x and y are defined as shown in Eq. (5.6) and Eq. (5.7). In Eq. (5.6) and Eq. (5.7), i and j represent the pixel locations. These statistical methods have clear physical meanings and are straightforward. Under these definitions, two images with more similarity have smaller MSE or MAE.

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [x(i, j) - y(i, j)]^2 \quad (5.6)$$

$$MAE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |x(i, j) - y(i, j)| \quad (5.7)$$

However, these simple distances do not take the human visual system (HVS) properties into consideration. Therefore, they are inappropriate to be used as reliable measures. The Structural Similarity Measure (SSIM) [128] is based on the hypothesis that human visual system (HVS) is highly adapted for extracting structural information. The SSIM measure defines the similarity of two images as a function of luminance, contrast, and structure. The luminance, contrast and structure are defined as:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (5.8)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (5.9)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (5.10)$$

Where given two images x and y , the μ_x and μ_y represent the means, σ_x and σ_y represent the standard deviation of the image x and y respectively, and σ_{xy} represents the covariance of x and y , C_1 , C_2 and C_3 represent constant values. SSIM is a combination of luminance, contrast and structure measure and the overall expression is shown in Eq. (5.11). The SSIM is applied on non-overlapping windows thus the mean of the SSIM values over the entire image (MSSIM, Eq. (5.12) is used to indicate the similarity between two images.

$$SSIM(x, y) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right) \left(\frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right) \quad (5.11)$$

$$MSSIM(x, y) = \frac{1}{N} \sum_{i=1}^N SSIM(x_i, y_i) \quad (5.12)$$

The SSIM [128] is appropriate to be used for clean images but it is observed that the performance of the SSIM index degraded substantially when assessing Gaussian blurred images. In many practical applications, noise is usually difficult to model. To measure the similarity between the original image and the reconstructed image, the Gradient Structural Similarity Measure (GSSIM) [129][130] is used. The GSSIM suggests that the gradients of the images to be integrated into the image similarity assessment to penalize dissimilarity in image contours and edges when noise exists. Thus, the GSSIM index makes comparisons between both x and y and the gradients of x and y . The gradients of x and y specifically indicate the similarity between edges. In the expression of GSSIM, the contrast

and structure terms of SSIM are modified as in Eq. (5.13) and Eq. (5.14), where the σ_x' and σ_y' represent the standard deviation of the gradient magnitude of image x and y respectively.

$$c(x', y') = \frac{2\sigma_x' \cdot \sigma_y' + C_2}{\sigma_x'^2 + \sigma_y'^2 + C_2} \quad (5.13)$$

$$s(x', y') = \frac{\sigma_{x'y'} + C_3}{\sigma_x' \cdot \sigma_y' + C_3} \quad (5.14)$$

Using the similar methods to fuse luminance, gradient contrast and gradient structure together, the GSSIM over image sub regions can be expressed as in Eq. (5.15):

$$GSSIM(x, y) = [I(x, y)]^\alpha [c(x', y')]^\beta [s(x', y')]^\gamma \quad (5.15)$$

Therefore, the average of GSSIM over the entire image can be used to indicate the similarity between the reconstructed image and the original image.

$$MGSSIM(x, y) = \frac{1}{N} \sum_{i=1}^N GSSIM(x_i, y_i) \quad (5.16)$$

It is worth noting that using the similarity measure itself is not accurate enough to measure the reconstruction performance. The reason is that when more edge pixels exist in an edge map, more information from the original image is used in the reconstruction which definitely yields amore similar result. Therefore, a penalty factor f_p which is formulated as a decreasing function of the total number of edge pixels is introduced in Eq. (5.17), where N_e represents the total number of edge pixels in the dilated edge map and MN represents the total number of pixels in the original image. By adding this term, the NREM measure chooses the optimal edge map as the one that shows the structural details in the image with minimal information and minimal false positives.

$$f_p = \frac{1}{1 + \frac{N_e}{MN}} \quad (5.17)$$

The final non-reference based edge measure (NREM) is comprised as the alpha weighted product of these two terms. In this chapter, the results are shown with $\alpha = 1$ and $\beta = 3$, which are obtained experimentally.

$$NREM(i, e) = (MGSSIM)^\alpha f_p^\beta \quad (5.18)$$

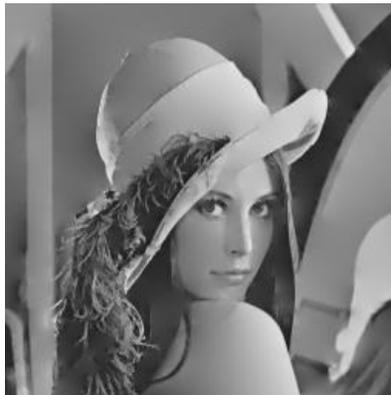
5.4. Experimental Results

In this section, some applications of the new non-reference edge measure NREM are demonstrated.

The example in Figure 5-4 is using the non-reference edge measures to select the optimal edge detection algorithm. Multiple edge detection results and the reconstructed images from the Canny, Sobel, Roberts, Log, and Prewitt for the Lena image are shown from Figure 5-4 (b) – Figure 5-4 (f). These edge detection algorithms are commonly used edge detection algorithms and each has its advantages and disadvantages. For example, gradient based edge detection algorithms such as Sobel and Prewitt are simple but sensitive to noise. The Canny edge detector improves the signal to noise ratio by smoothing the image, however, the smoothing may lead to loss of corners and detection of double edges. Therefore, it is necessary to have a reliable edge measure that can help to decide the optimal edge detection algorithm for a specific image. It is seen that the Sobel, Roberts and Prewitt outputs miss a lot of essential edges such as the hat and the background structure, while the LoG output has many isolated pixels. The NREM selects the Canny edge detection result as the optimal results which agrees with the visual assessment. The Yitzhaky's measure is used for comparison. These five edge maps are used to estimate the ground truth edge map in the Yitzhaky's measure. It is observed that most of the five edge maps do not contain details in the upper left corner so the Yitzhaky's selection of the optimal edge output, the Prewitt edge map, misses edge pixels in that portion of the image too. This observation demonstrates the limitation of the Yitzhaky's method that if the majority edge maps used for estimating the ground truth are biased, the estimated ground truth edge map is biased, so the measure values are biased.



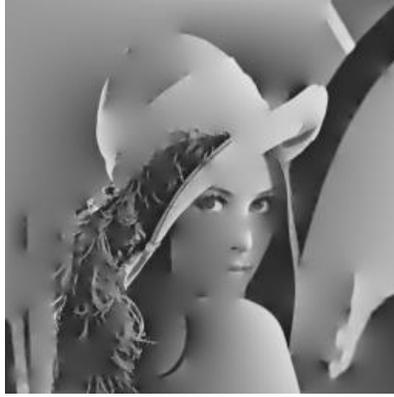
(a)



(b)



NREM = **0.7888**
Yitzhaky's = -1.1516



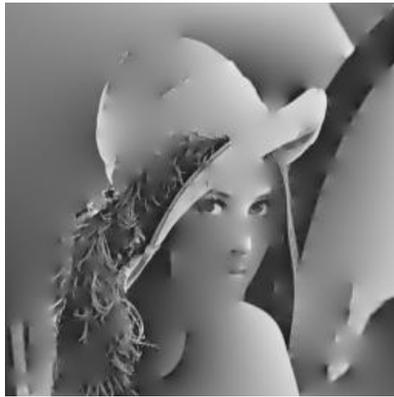
(c)

NREM = 0.7217
Yitzhaky's = 0.4983



(d)

NREM = 0.7851
Yitzhaky's = -0.7417



(e)

NREM = 0.6959
Yitzhaky's = 0.0814

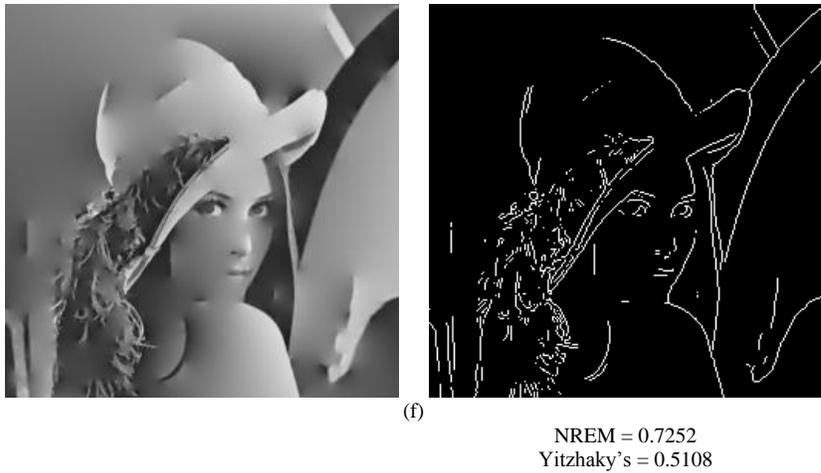
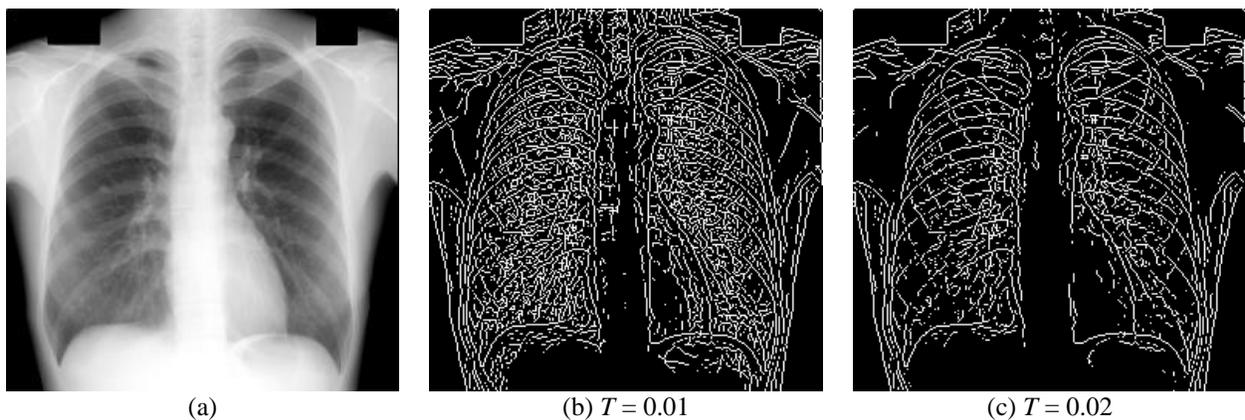


Figure 5-4. Using measures to select optimal edge detection algorithms. (a) Original Lena image. (b)-(f) Edge detection results (right column) and the reconstructed images (middle column) from: (b) Canny, (c) Sobel, (d) Roberts, (e) LoG, (f) Prewitt edge detectors.

Figure 5-5 shows an example of using the edge measure as a means of selecting optimal parameter values. In this experiment, the Sobel edge detection algorithm with different threshold T ranging from 0.01 to 0.08 is used. The testing image is an X-Ray chest image which suffers from low contrast. Therefore, lower threshold values tend to keep more soft tissue or other noise components in the edge map, while higher threshold values discard essential edges. The proposed measure selects the optimal parameter at threshold $T = 0.03$ and achieves the best tradeoff between noise removal and feature extraction. In contrast, the Yitzhaky's method selects threshold $T = 0.06$ which losses some ribs in the edge map. Figure 5-5 (k) also illustrates the need for the edge pixel density function in the formulation of NREM as the use of MGSSIM alone results in the discussed edge pixel bias.



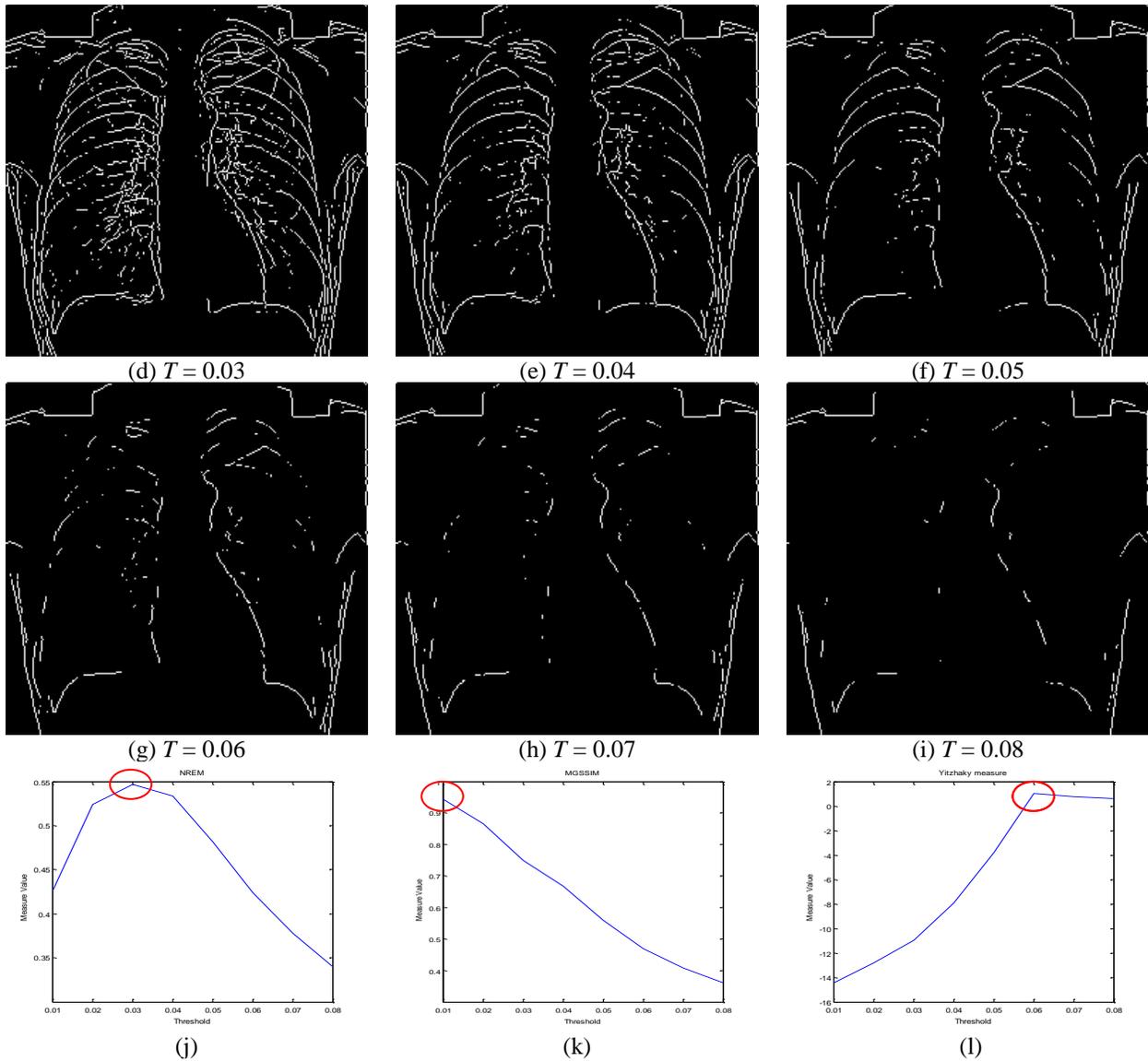


Figure 5-5. Applying edge measures in assisting selecting optimal operating parameters. (a) Original X-ray chest image. (b)-(i) Edge detection results using the Sobel edge detector with the threshold T ranging from 0.01 to 0.08. (j) presented NREM measure plot indicating $T = 0.03$ as the optimal parameter value. (k) measure plot of using MGSSIM alone. (l) The Yitzhaky's method indicating $T = 0.06$ as the optimal parameter value.

5.5. Summary

Non-reference edge measure is very useful in image segmentation, registration and reconstruction. A new non-reference edge map evaluation NREM is proposed in this chapter. This measure is based on the observation that the best edge map results consists of the least number of edge pixels at their correct locations needed to characterize all the relevant structure in the reconstruction image. Comparison with state of the art non-reference edge detection

measure shows the advantages of the new measure: The NREM utilizes the information from the original image thus can achieve better performance. Experimental results on using the NREM on selecting the optimal edge detection algorithm and optimal operating parameters show that the measure coincides with subjective evaluations, validating the usefulness of the measure.

Chapter 6. Image Enhancement and Edge Detection Algorithms Using Alpha Weighted Quadratic Filter

Several color image quality measures and edge measures have been developed in the previous chapters. In this chapter, these measures are used to direct the design of new color image enhancement algorithm and edge detection algorithm. In the new algorithms, the quadratic filter is adopted as the baseline for its nonlinear property of enhancing image details while simultaneously removing noise. Furthermore, the positive alpha power is added to introducing the global contrast adjustment for color enhancement applications, and the negative alpha power is added to revealing image structures for edge detection applications. The alpha weighted quadratic filter is parametric and the measures are used to assist the selection of optimal operating parameters. The new algorithms are compared with other existing algorithms, and the measures are used to compare their performances. All the measures presented in the previous chapters are summarized in Appendix A for references.

6.1. Quadratic Filter

The human visual system features intrinsic nonlinear behavior. It characterizes the quadratic relation between the optical intensity and the response [131]. Therefore, the nonlinear filters are desired to model real-world image processing systems in accordance with the human perception. Among the nonlinear filter categories, polynomial filters estimate new pixel intensities from geometrically related neighboring pixels. The polynomial filters are flexible to be extended from one dimensional to multiple dimensional, from first order linear to high order nonlinear, and from small local region to the entire image. Therefore, the polynomial filters are used for the image processing tasks in this chapter. Generally, the M -sample memory p th-order truncated polynomial filter is expressed as Eq. (6.1), where x represents the input pixel intensity, y represents the output pixel intensity, w_1 and w_2 are the weighting coefficients, k_i represents the distances to pixel x [132].

$$y(n) = w_0 + \sum_{k_1=0}^{M-1} w_1(k_1)x(n-k_1) + \sum_{k_1=0}^{M-1} \sum_{k_2=0}^{M-1} w_2(k_1, k_2)x(n-k_1)x(n-k_2) + \dots + \sum_{k_1=0}^{M-1} \dots \sum_{k_p=0}^{M-1} w_p(k_1, \dots, k_p)x(n-k_1) \dots x(n-k_p) \quad (6.1)$$

It is seen that the output of the polynomial filter is linear with respect to the kernel coefficients [132], and the computational complexity increases extremely high with more terms used in Eq. (6.1). To reduce the complexity without loss of the nonlinear properties, only the first order linear and second order quadratic terms are preserved in the filter design in this chapter as shown in Eq. (6.2). As a simple implementation of the polynomial filter, the quadratic filter introduces the nonlinear effects to the traditional linear analysis in an efficient and manageable way.

$$y(n) = \sum_{i=0} w_1(i)x(n-i) + \sum_{i=0} \sum_{j=0} w_2(i, j)x(n-i)x(n-j) \quad (6.2)$$

The basic quadratic filter in the image processing algorithms presented in this dissertation considers the pixels in a 3 by 3 local region as shown in Figure 6-1 (a), to ensure that fine details are captured with additions of limited complexity in the coefficients design. Besides, since the human visual system captures image details hierarchically, images can be processed in a multi-scale approach. Accordingly, image details at a large scale can be downsampled and processed locally at a 3x3 sub-window. One example of processing a larger image region using the downsample methods is illustrated in Figure 6-1 (b).

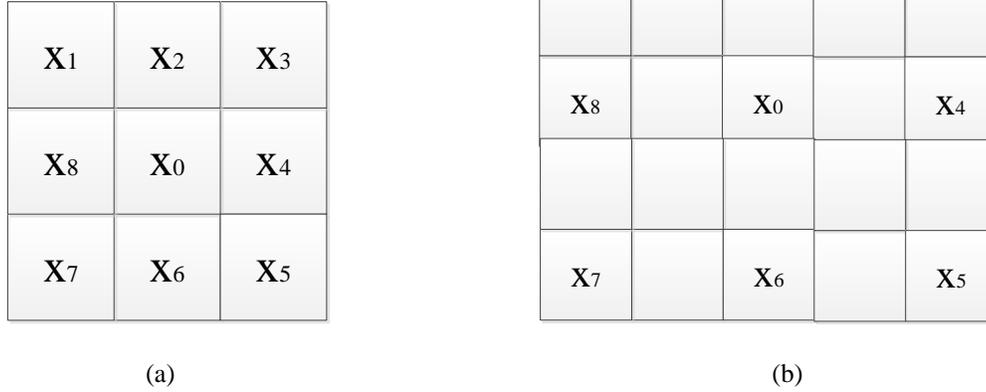


Figure 6-1. Demonstration of a 3x3 subwindow and the kernel patterns.

The basic form of the quadratic filter used in this chapter is composed of two terms: (1) a quadratic linear filter y_0 , which consists of all the first and second order terms in which the distance of two elements is zero, this occurs when two elements have the same coordinates; and (2) a quadratic operator y_1 , which includes the second order terms in which the distance of two elements is one. In other words, two elements of y_1 are two adjacent pixels.

$$y(x) = y_0(x) + y_1(x) \quad (6.3)$$

The relations between the center pixel and the neighboring pixels in the 3x3 local regions are examined. Therefore y_0 and y_1 can be expressed as:

$$y_0(x) = \sum_{i=0}^8 w_1(i)x(i) + \sum_{i=0}^8 w_2(i)x(i)^2 \quad (6.4)$$

$$y_1(x) = \sum_{i=0}^8 \sum_{j=0, |j-i|=1}^8 w_2(i, j)x(i)x(j) \quad (6.5)$$

Where x is defined as a 3 by 3 kernel with x_0 in the center, the eight neighbors are x_1 to x_8 as shown in Figure 6-1. From the definition in Eq. (6.5), it is seen that the output of a quadratic filter relies linearly on the coefficients of the filter itself [131]. Therefore, the design of the coefficients determines the performance of the filter. The non-linear filter is highly parametric, even with the linear and quadratic terms only, a great number of parameters (9 linear coefficients and 81 quadratic coefficients) are needed in the design of the quadratic filters. Therefore, the

properties of the quadratic filters used in image processing applications need to be examined to reduce the number of parameters and instruct the design of the coefficients. Some properties are tabulated in Table 6-1:

Table 6-1. Properties of the quadratic filter

Property	Reason	Implementation
Symmetry	The kernels are symmetric for equivalent representations.	The coefficients $w(i,j)$ acts on the pixels $x(i)$ and $x(j)$ even if the index i and j are exchanged.
Isotropy	The quadratic filter should be invariant to the rotation of the input image.	The coefficients are isotropic with respect to the center pixel.
Uniform response	A constant input should have the same value in the output.	Assuming the sum of the linear coefficients be one and the sum of the quadratic coefficients be zero [133].

After collecting and grouping the isotropic and symmetric terms together, the degree of freedom is reduced. y_0 can be written as:

$$y_0 = h_0x_0 + h_1(x_1 + x_3 + x_5 + x_7) + h_2(x_2 + x_4 + x_6 + x_8) + w_0x_0^2 + w_1(x_1^2 + x_3^2 + x_5^2 + x_7^2) + w_2(x_2^2 + x_4^2 + x_6^2 + x_8^2) \quad (6.6)$$

While y_1 can be written in the following form:

$$y_1 = s_1(x_0x_1 + x_0x_3 + x_0x_5 + x_0x_7) + s_2(x_0x_2 + x_0x_4 + x_0x_6 + x_0x_8) + r_1(x_1x_2 + x_3x_4 + x_5x_6 + x_7x_8 + x_1x_8 + x_3x_2 + x_5x_4 + x_7x_6) + r_2(x_2x_8 + x_4x_2 + x_6x_4 + x_8x_6) \quad (6.7)$$

Where the coefficient s_i and r_i are weights that determine whether the couple of the pixels is related or not related to the central pixel [134].

The following constraints on the coefficients meet the uniform response property:

$$4h_1 + 4h_2 + h_0 = 1 \quad (6.8)$$

$$4w_2 + 8r_1 + 4s_1 + 4w_1 + 4r_2 + 4s_2 + w_0 = 0 \quad (6.9)$$

The y_0 filter is designed as an edge-preserving smoother, where the linear terms are designed as a low pass filter to remove a noise and the quadratic terms will compensate for the blurring effect. Since the design in Eq. (6.8) can be seen as the low pass smoothing filter, the coefficients w_i should be compensate the smoothing effect and designed as a high pass linear filter:

$$4w_2 + 4w_1 + w_0 = 0 \quad (6.10)$$

Therefore, $w_0 = 1-h_0$, $w_1 = -h_1$, and $w_2 = -h_2$, satisfy the constraint in Eq. (6.8) and Eq. (6.10). Subtract Eq. (6.10) from Eq. (6.9), the following constraint is obtained:

$$8r_1 + 4s_1 + 4r_2 + 4s_2 = 0 \quad (6.11)$$

y_l is designed as the detail enhancement filter. To ensure equal responses on the directional details [134], the following constraints are imposed:

$$s_l = s_2, r_l = -s_l/2, \text{ and } r_2 = -s_2 \quad (6.12)$$

In this chapter, these values for coefficients of the linear and quadratic terms are used, which satisfy the constraints as shown in Eq. (6.8), Eq. (6.10) and Eq. (6.12):

$$\begin{aligned} h_0 = 0.2, h_1 = 0.1; h_2 = 0.1; w_0 = 0.8; w_1 = -0.1; w_2 = -0.1 \\ s_1 = 0.25; s_2 = 0.25; r_1 = -0.125; r_2 = -0.25 \end{aligned} \quad (6.13)$$

6.2. Application 1: Color Contrast Enhancement

The goal of color contrast enhancement is to recover color and enhance contrast from captured image scenes. Image color contrast enhancement is widely used in various applications such as homeland security and medical diagnosis. For example, surveillance cameras usually capture images under strong sunlight, in a dark night, or in a haze day. Biomedical images usually suffer from low contrast or low resolution due to the limited radiation exposure time. Contrast enhancement algorithms are used after image acquisition to boost the visibility of such images. Some medical applications also require an image with close-to-nature colors such as skin cancer diagnosis. Therefore, a good color enhancement algorithm should correct image histograms and make it more appealing with respect to the human vision. Figure 6-2 visualizes some practical applications of the color contrast enhancement algorithms.

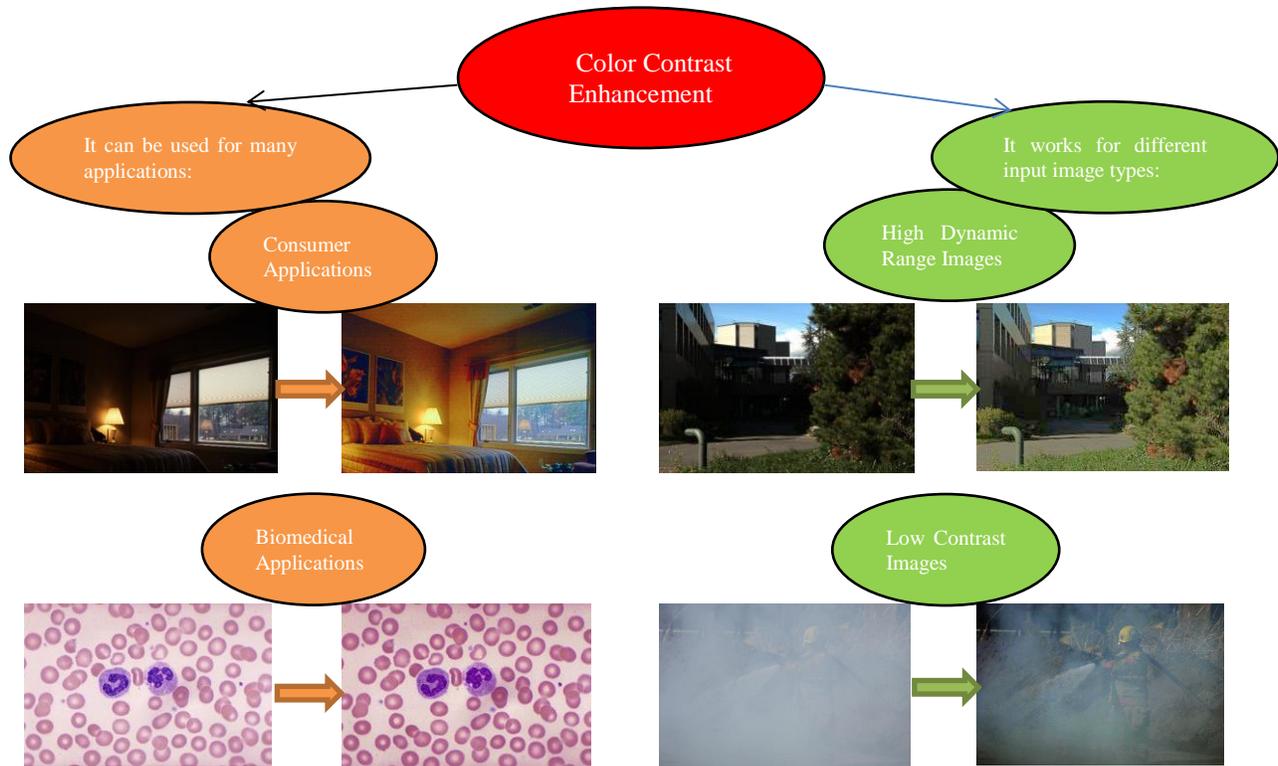


Figure 6-2. Uses of color contrast enhancement. Color contrast enhancement is widely used in many applications such as consumer and biomedical applications. It is also used to correct histograms for high dynamic range images as well as low contrast images.

6.2.1 Existing Spatial Domain Color Enhance Methods

Spatial domain enhancement algorithms are applied on the pixel value directly, such as the 3D histogram equalization [135][136][137][138] and the Retinex techniques [139][140] [141] [142]. The 3D histogram equalization algorithm is inspired by the simplicity and effectiveness of 1D histogram equalization for enhancing grayscale images. However, for multichannel images, there are numerous ways to adjust the histogram to achieve equal density of pixel numbers. Simply applying histogram equalization on each color component does not guarantee that the results yield a uniform histogram for color images. This is due to the nonlinear characteristics of multidimensional color images. As a result, this may cause a color casting problem. Recently, sophisticated histogram equalization based methods have been proposed to solve this problem [135][136][137][138], yet these algorithms still process single pixel intensities and do not take into account neighboring or background intensity relationships. For complex images, for example, when image intensities are highly variant and possess a bi-model

distribution, i.e. a sunny outdoor image with shadows, the global techniques tend to over enhance the entire image and cannot render all the details from both the bright region and the dark region.

The recent research on the Human Visual System (HVS) inspired a series of new color contrast enhancement algorithms. One typical method is the Retinex algorithm. The Retinex theory, which was first proposed by Land and McCann, achieves color consistency by computing the local ratio of a target pixel to its neighbors [139][140]. It explains how the visual system extracts information from a scene despite changes in illumination. Based on the observations, Land proposed the path version of Retinex which computes the product of ratios between pixel values along a set of paths in the image [139][140]. Then, Rizzi et al developed a similar method but the paths are randomly distributed using Brownian motion [143]. Rahman, Jobson and Woodell introduced logarithmic operations to the algorithm so that the subtraction operations are converted into division operations[82][144]. They also observed that using different sizes of neighboring windows results in retrieving different sized details. Motivated by this observation, they proposed the Multi Scale Retinex (MSR) which averaged the three images obtained from the three different surround sizes (large, medium and small scales) [141]. To better recover colors, the Multi Scale Retinex with Color Restoration (MSRCR) are further developed which corrects colors by assigning a gain factor to each pixel [142]. The Retinex algorithms have good performance on extracting details from all brightness levels of an image but they also have limitations such as the halo phenomenon when they process multiple components of the incident light for an image with strong contrast.

Wang and Luo proposed an adaptive color contrast enhancement algorithm focused on solving the non-satisfactory illumination and color casting problem [145][146]. They designed a pixel related adaptation factor in the Naka-Rushton function [147]. The adaptation factor is proportional to the local intensity and refined by the local chrominance and contrast. Then, a center-surround comparison based contrast correction is applied to compensate for the lost contrast caused by the intensity adjustment. Finally, a color enhancement process is applied to gain more vivid colors. Their color enhancement algorithm has a similar methodology to Lukac's automatic white balancing (AWB) [148] but they operate in the CIELAB color space and their method is insensitive to the parameters. The algorithm works well on underexposed images since the adaptive Naka-Rushton function changes the intensity in such a way that the output intensity is always greater than the input intensity. In their consecutive work, they set a local illuminant related threshold for classifying whether the current pixel belongs to a bright region or dark region. Then, they treat the overexposed portion as an inverse image of the underexposed image and modify the traditional

Naka-Rushton formula to a symmetric expression such that it works on both the underexposed and overexposed regions [146]. Unfortunately, the subsequent contrast and color enhancement steps have high computational complexity.

6.2.2 Three-Dimensional Alpha Weighted Quadratic Filter (3DAWQF) For Color Contrast Enhancement

The color image quality is related with three fundamental factors: intensity, contrast, and color. The three-dimensional alpha weighted quadratic filter (3DAWQF) algorithm consists of four steps: (1) intensity adjustment, (2) intra-plane AWQF contrast enhancement, (3) inter-plane contrast color enhancement, and (4) fusion of the color and contrast enhanced images. The flowchart and the corresponding intermediate results after each step are illustrated in

Figure 6-3 and Figure 6-4. Details of each step are described in the following sections.

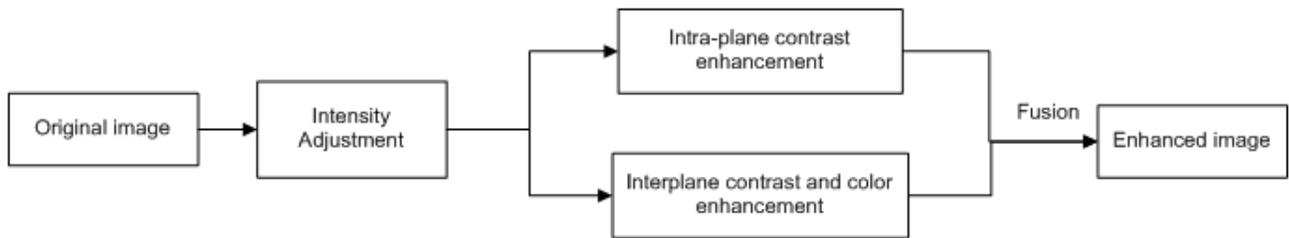
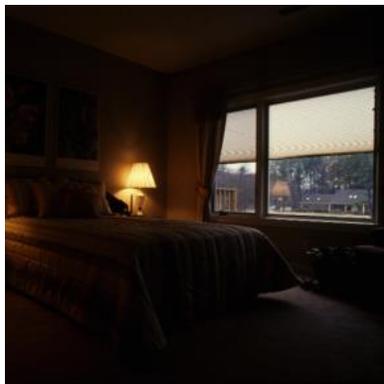


Figure 6-3. Flow of the new 3DAWQF algorithm.



(a)



(b)



(c)



Figure 6-4. Intermediate results for each of the 3DAWQF steps. (a) Original image; (b) Intensity adjusted image; (c) Intra-plane AWQF contrast enhanced image; (d) Inter-plane contrast color enhanced image; (e) Fused image. The original image is obtained from the NASA Retinex research website [149]. It can be seen that the contrasts and colors are significantly enhanced after step (c) and (d), and the final fused images achieve an overall enhancement.

Step1: Intensity Adjustment

The power law is the most commonly used method to adjust image intensities by modifying the image histogram, because of its effectiveness and simplicity. Using a fixed power through the entire image does not produce the optimal contrast enhancement results especially when the images are captured from multiple light sources or demonstrate non-uniform illuminations. To solve this problem, Wang and Luo proposed an automatic intensity adjustment method based on the uses of the Naka-Rushton function [150]. The traditional Naka-Rushton function is described in Eq. (6.14)

$$V_c(j,k) = \frac{I_c(j,k)}{I_c(j,k) + H_c^*(j,k)} \times \frac{I_{c,max} + H_c^*(j,k)}{I_{c,max}} \quad (6.14)$$

Where I_c represents the intensity value of each color component and $I_{c,max}$ represents the maximum intensity value of that color channel. H_c^* is an adaption factor. Wang and Luo proposed an automatic selection method of the adaption factor H for each local pixels [145]. The essential part of the adaption factor is proportional to the local intensity and then refined by the global and neighboring intensity to preserve chrominance and enhance contrast [145]. Their algorithm shows good performance in eliminating underexposure and compensating the channel with a low mean value to get rid of the color casting problem [145]. However, in essence, it is still comparable to the power law with a power less than 1 (Figure 6-5 (a)), which always increases pixel intensities thus is limited applicable to ‘too dark’ images only. Note that in Figure 6-5 (a), the input-output curve always shows in the upper side of the diagonal line, and the shirt of the boy is over enhanced.

Recognizing this limitation, Wang and Luo later proposed a symmetric Naka-Rushton function by setting an automatic threshold to separate the image into two portions: a "dark" region and a "bright" region [146]. The "dark" portions are processed using a similar method as the adaptive Naka-Rushton function (Eq. (6.15)). The "bright" portion of an image is converted to its grayscale inversed image, treated as an under illuminated portion, and then converted back to the original gray level. As seen in Eq. (6.15), two different expressions of H need to be developed for different illuminations. For overexposed regions, the output intensity is an increasing function of adaption factor H_{high} . In this case, the globally brighter image should lead to a smaller H thus the intensities will be decreased more. In the contrast, for the underexposed regions, the output intensity is designed to be a decreasing function of H_{low} . This means a smaller H should be generated for larger increments for darker pixels. However, the two definitions of H result in discontinuity for neighboring regions if they are partitioned to different illumination regions. As an example shown in Figure 6-5 (b), the arm of the boy demonstrates different color tone with other portions of the image.

$$V_{sym} = \begin{cases} \frac{I}{I + H_{low}}(T + H_{low}) & 0 < I \leq T \\ 1 - \frac{1 - I}{(1 - I) + H_{high}}[(1 - T) + H_{high}] & T < I \leq 1 \end{cases} \quad (6.15)$$

Inspired by the threshold idea, a modified Naka-Rushton function is proposed as in Eq. (6.16) which does not require two different representations of the adaption factor H . The input-output transition curve is demonstrated in Figure 6-5 (c).

$$V = \begin{cases} I \left(\frac{T + \beta H}{I + \beta H} \right)^\gamma & 0 < I \leq T \\ (I - T) \left(\frac{I - T + \beta(1 - H)}{1 - T + \beta(1 - H)} \right)^\gamma + T & T < I \leq 1 \end{cases} \quad (6.16)$$

This essential part of the adaption factor H is designed to be proportional to the pixel intensity in the local region. So if a pixel is partitioned to be in the dark regions, the darker the pixel, the smaller H which leads to greater increment; on the other hand, if a pixel is located in bright regions, a brighter pixel which has a greater H will result in a more decreased intensity. An example of the resulting intensity adjusted image is shown in Figure 6-5 (c). The result shows that the intensity adjusted image with modified Naka-Rushton function reveals the details in both shadow and the sun exposed parts. Note that in Figure 6-5, the objective image quality measure values CRME is shown under each image, where the greater value of the CRME corresponds to an image with better quality.

Compared with the intensity corrected images obtained from the traditional Naka-Rushton function in Figure 6-5 (a) or the symmetric Naka-Rushton function in Figure 6-5 (b), the modified Naka-Rushton function achieves the best intensity adjustment performance according to the visual perception and the CRME measure.

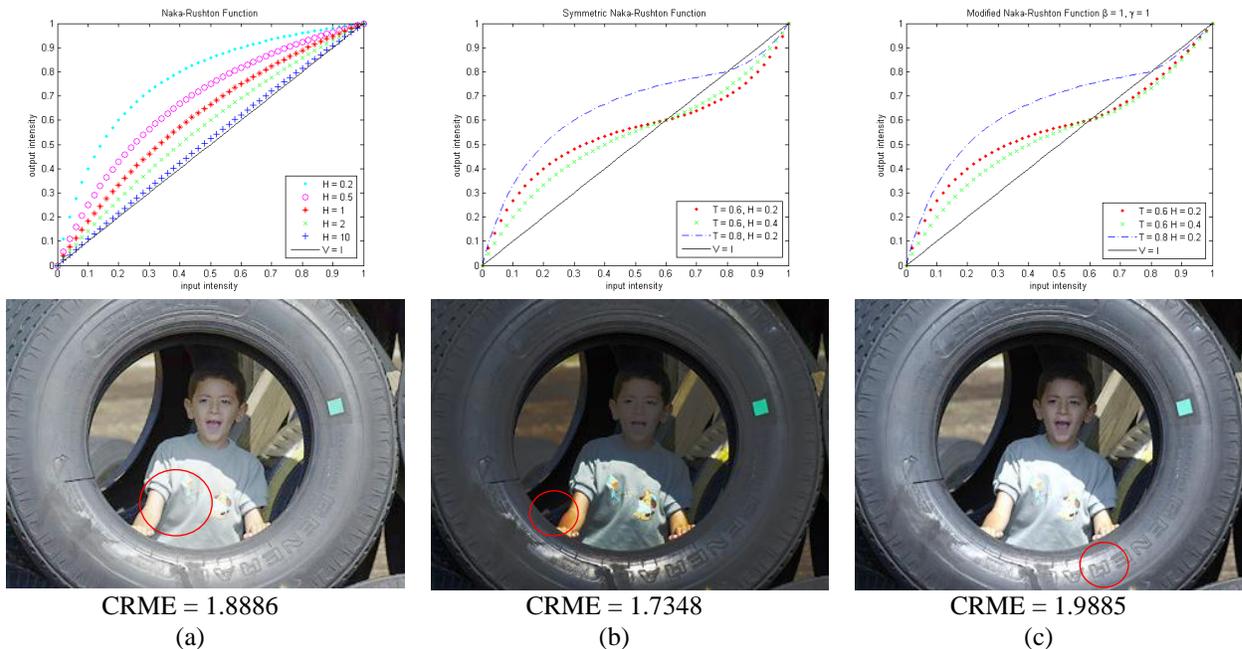


Figure 6-5. Comparisons of different intensity adjustment results using different modifications of Naka-Rushton function: (a) Adaptive Naka-Rushton in [145]. (b) Symmetric Naka-Rushton in [146]. (c) Proposed modified Naka-Rushton. Red circles denote the problem areas resulting from the corresponding Naka-Rushton function.

The two parameters β and γ used in Eq. (6.16) are selected based on different applications. β is the weight of the adaption factor H . When the image has a high dynamic range or possesses a strong bi-model distribution, users should choose a smaller β which results in a greater change of output intensities. In contrary, if the input image intensity is well distributed, a greater β should be selected to avoid over-adjusting the intensity. Parameter γ is set to be equal to 1 when it is used to process high dynamic range images and set to -1 when it is used to enhance the images centered on the middle range and suffering from the low contrast problem. Examples of the effect of these two parameters and the illustrative applicable images are shown in Figure 6-6.

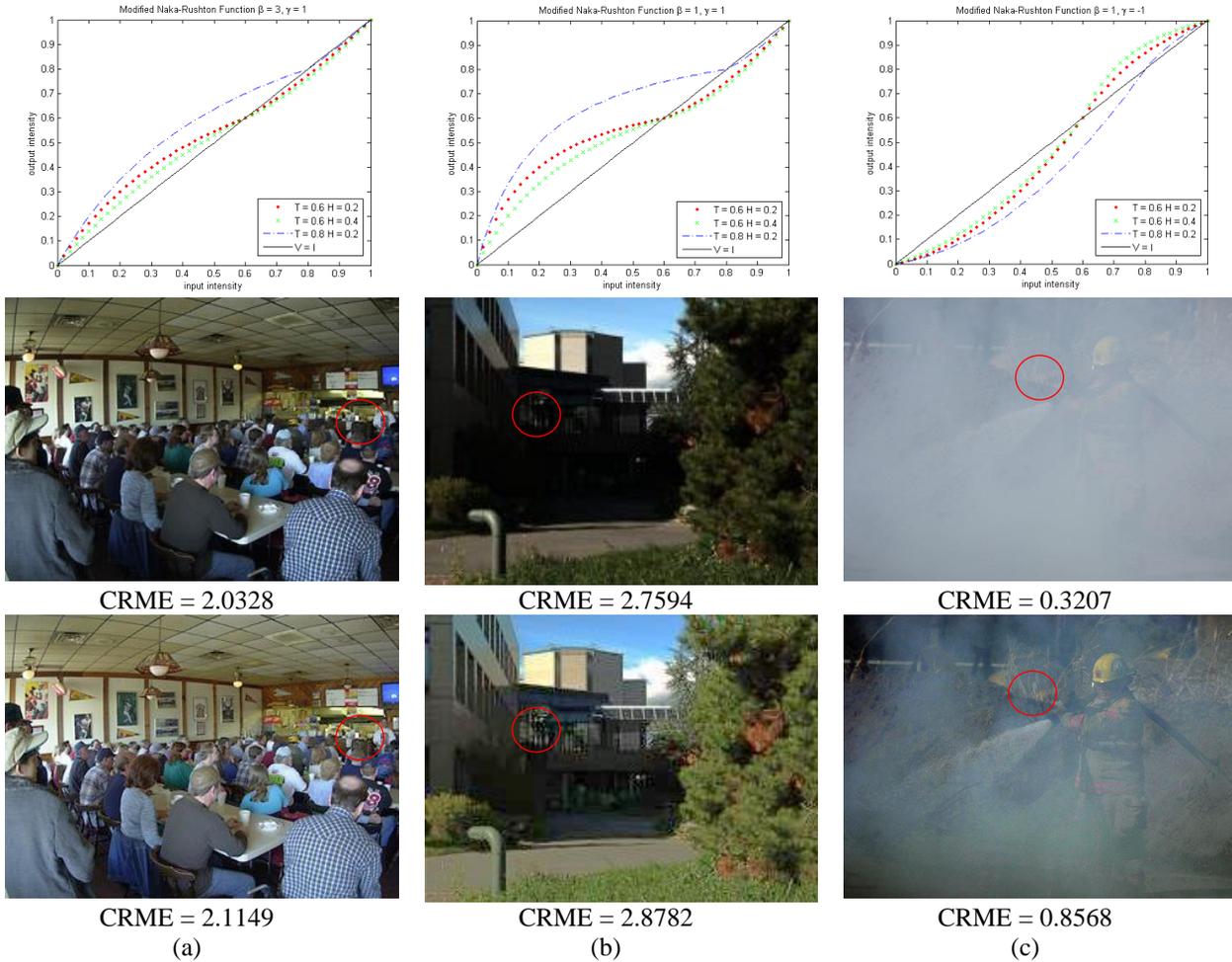


Figure 6-6. Comparisons of different parameters β and γ in Eq. (6.16). The first row is the modified Naka-Rushton curve with different parameters. The second row is a representative input image which is suitable for using the corresponding parameters. The third row is the intensity adjusted image with these parameters. (a) $\beta = 3, \gamma = 1$; (b) $\beta = 1, \gamma = 1$; (c) $\beta = 1, \gamma = -1$.

Step2: PAWQF Intra-plane Contrast Enhancement

After the intensities are adjusted, the contrast needs to be enhanced for each RGB color component. The quadratic filter works on neighboring pixels to enhance local contrast in each 3×3 local regions. Using positive values for the alpha power variable in the method introduces the global power law, which improves the global contrast. The alpha powers are added to each term so that it incorporates the power law into the local regions so the local contrast is enhanced as well as the global contrast. The Positive Alpha Weighted Quadratic Filter (PAWQF), as the name implies, is realized by raising each collected and grouped term to a power of alpha. In this case, y_0 becomes:

$$y_0 = h_0x_0 + h_1(x_1 + x_3 + x_5 + x_7)^a + h_2(x_2 + x_4 + x_6 + x_8)^b + w_0x_0^{2c} + w_1(x_1^2 + x_3^2 + x_5^2 + x_7^2)^d + w_2(x_2^2 + x_4^2 + x_6^2 + x_8^2)^e \quad (6.17)$$

And y_1 becomes:

$$y_1 = s_1(x_0x_1 + x_0x_3 + x_0x_5 + x_0x_7)^f + s_2(x_0x_2 + x_0x_4 + x_0x_6 + x_0x_8)^g + r_1(x_1x_2 + x_3x_4 + x_5x_6 + x_7x_8 + x_1x_8 + x_3x_2 + x_5x_4 + x_7x_6)^h + r_2(x_2x_8 + x_4x_2 + x_6x_4 + x_8x_6)^i \quad (6.18)$$

The benefits of using the alpha power in the quadratic filter are shown in Figure 6-7. The theoretical analysis shows that when a positive alpha is applied, based on the power law, both the global and local contrast is enhanced. The experimental results confirm this and further demonstrate that by properly choosing the alpha power, the color can be better recovered, according to the visual perception and the CRME and MEMEE measures. It is seen that with the alpha power, the white stripe is clearer in the shirt image, and the yellow and gray color patches are more saturated in the color board image.

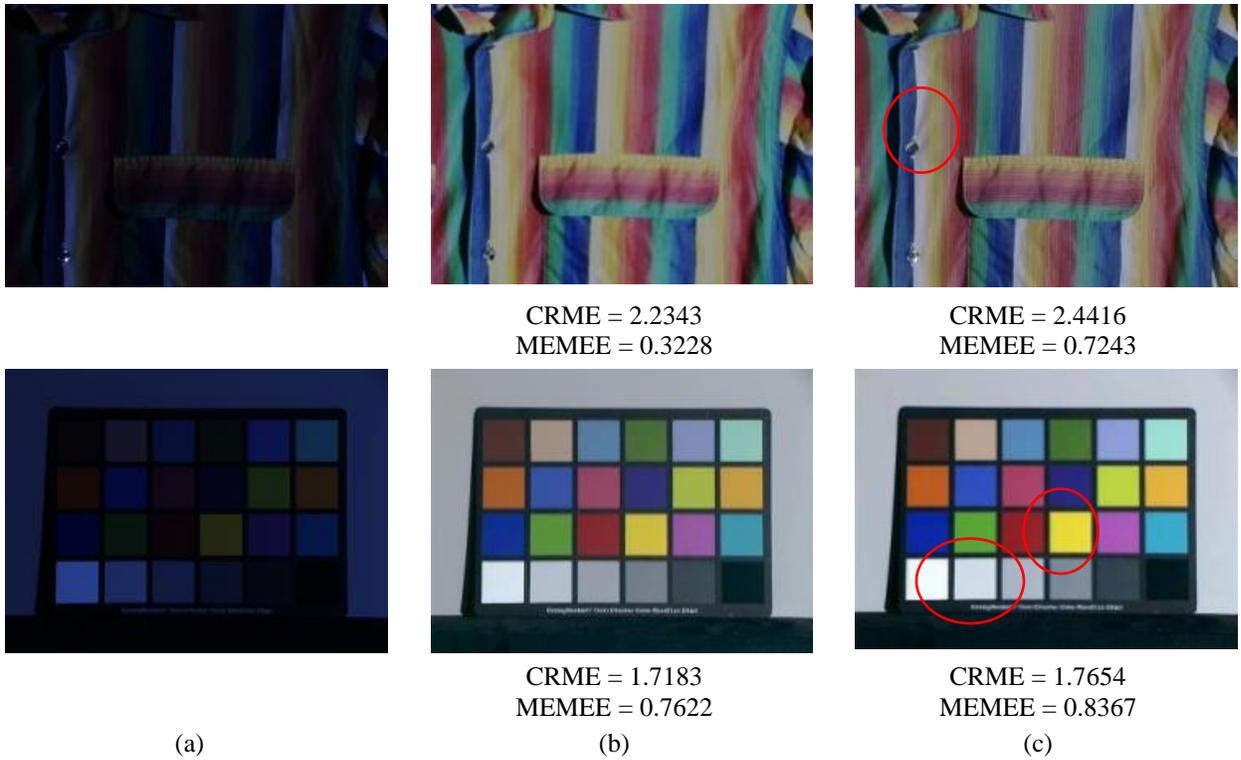


Figure 6-7. Comparison of the effect of alpha power in filter design. (a) original image; (b) enhanced image without the alpha power; (c) enhanced image with the alpha power. The white stripe is clearer in the shirt image in (c). The yellow and gray color patches are more saturated in the color board image in (c). Accordingly, the CRME and MEMEE assign greater values for images in column (c).

Some PAWQF contrast enhancement results are shown in Figure 6-8. It is seen that more image details are distinguishable after the contrast enhancement. The details on the trees, the windows, and the boat are more discernible. The CRME and MEMEE also demonstrate the effectiveness of the contrast enhancement performance of the PAWQF.



CRME = 1.3198
MEMEE = 0.4173



CRME = 1.6485
MEMEE = 0.6230



CRME = 1.6397
MEMEE = 0.2223



CRME = 1.7998
MEMEE = 0.6486

Figure 6-8. (a) Original image. (b) contrast enhanced image. It is seen that the PAWQF enhanced image has sharper edges compared to the original image. The details on the trees, the windows, and the boat are more discernible.

Step3: Inter-plane Color Contrast Enhancement

The PAWQF described in previous section operates within each color plane and does not take the information across color planes into consideration. However, processing contrast enhancement in each color component

individually may cause color casting problem. Recent research also showed that statistically about 10% high contrast details exist in color changes [151]. To capture these changes between color planes and to recover a more natural color, it is necessary to have another step that considers pixel relations across color planes. To capture such relations, four pseudo planes are constructed perpendicular to the real RGB color planes as shown in Figure 6-9. Each pseudo plane corresponds to a directional sub plane and the combinations of all four pseudo planes traverse all pixels in the RGB color cube. A high pass filter is designed on each pseudo color plane to capture the difference among color planes. In Eq. (6.19), a smaller M corresponds to a greater influence of the neighboring pixels.

$$y_{c1} = (Mg_0 - r_5 - r_0 - r_1 - g_1 - b_1 - b_0 - b_5 - g_5)/(M - 8)$$

$$y_{c2} = (Mg_0 - r_6 - r_0 - r_2 - g_2 - b_2 - b_0 - b_6 - g_6)/(M - 8)$$

$$y_{c3} = (Mg_0 - r_7 - r_0 - r_3 - g_3 - b_3 - b_0 - b_7 - g_7)/(M - 8)$$

$$y_{c4} = (Mg_0 - r_8 - r_0 - r_4 - g_4 - b_4 - b_0 - b_8 - g_8)/(M - 8)$$

$$y_c = \frac{y_{c1} + y_{c2} + y_{c3} + y_{c4}}{4} \quad (6.19)$$

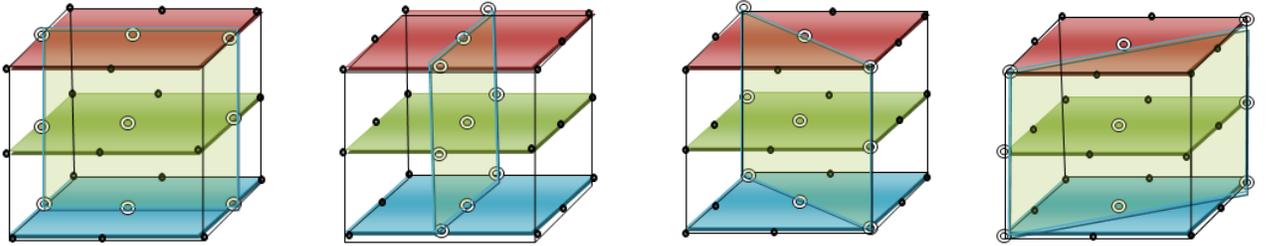


Figure 6-9. Four pseudo inter planes.

One example of using the PAWQF in the pseudo planes to enhance image colors are shown in Figure 6-10. This demonstrated result is obtained with parameter $M = 24$. It is seen that the colors are significantly enhanced. Four colorfulness measures are shown under the images and the measure values conform to the visual observations.



$M_{C1} = 0.4100$
 $M_{C2} = 0.2872$
 $M_{C3} = 0.6087$
 $M_{C4} = 5.9576$
 (a)



$M_{C1} = 1.0344$
 $M_{C2} = 0.5336$
 $M_{C3} = 1.3541$
 $M_{C4} = 15.8288$
 (b)

Figure 6-10. (a) Original image. (b) color enhanced image. It is seen that the color enhanced image has more vivid color compared to the original image.

Step4: Fusion

After the intra plane contrast enhanced image and the inter plane color contrast enhanced image are obtained, the next step is to fuse them together. Users have the option to choose the fusion methods based on the desired fusion quality. Figure 6-11 shows some examples of using existing pixel level fusion methods [152]. For example, Parameterized Logarithmic Image Processing (PLIP) [76] addition is a specific algebraic and functional operation which has the benefit of maintaining intensity values in the bounded range. The addition operation is consistent with the logarithmic transmittance image formation model and the saturation character of human eye. Principle Component Analysis (PCA) method performs fusion as a superposition of two input images and the weighting coefficients are determined with respect to the information content of the two images. The Image pyramid is initially described for multi-resolution image analysis and is motivated by the fact that human visual system captures image details hierarchically. An image pyramid usually consists of two part: a set of low-pass or band-pass copies of an image and another difference pyramid containing high frequency component [153]. Each copy represents pattern information of a variant scale. Ratio pyramid and Contrast pyramid are constructed in a nonlinear way of computing pixel by pixel division of neighboring resolutions. At each level, the pyramid is the ratio of two successive levels of Gaussian pyramid [154] and that schematic conforms to the contrast definition which denotes the ratio of the difference between a certain location luminance and local background luminance. Gradient pyramid decomposes the image into its vertical, horizontal and two diagonal edge representations using directional derivative filter and fusion

is applied on each representation. Other fusion methods include but are not limited to morphological fusion, wavelet transform fusion, neural network fusion, total probability density fusion, and optimization approaches [152][153].

The color image quality measure CQM is used to benchmark the performances of these fused results. According to the CQM, for the lighthouse image, the fused image from the PCA fusion has the best quality, while the result from the FSD fusion is the worst. The results conform to the visual perception in the means that the PCA fusion result has relatively more vivid color and discernable structures, while the FSD fusion result has a washed out appearance with blurred image details.

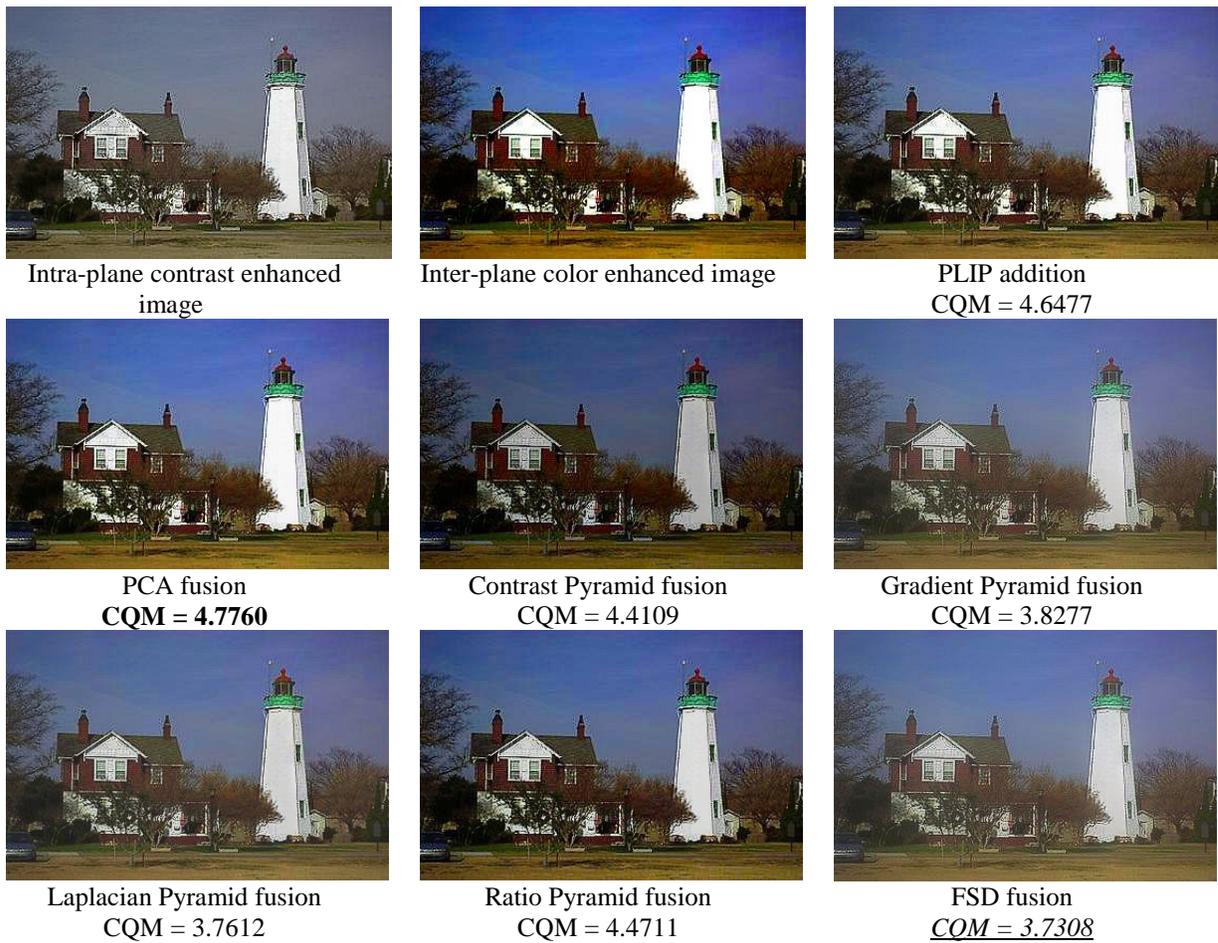


Figure 6-11. Comparisons of fusion results of intra plane contrast enhancement image and inter plane color image. Results are obtained using fuse toolbox [152]. According to the CQM, the PCA fusion is the best, and the FSD fusion is the worst among these fusion results.

One example of the entire procedure of the 3DAWQF after each step is illustrated in Figure 6-12. The effects of each processing step are demonstrated, and the overall enhancement is significant with respect to local contrast, global contrast, as well as color.

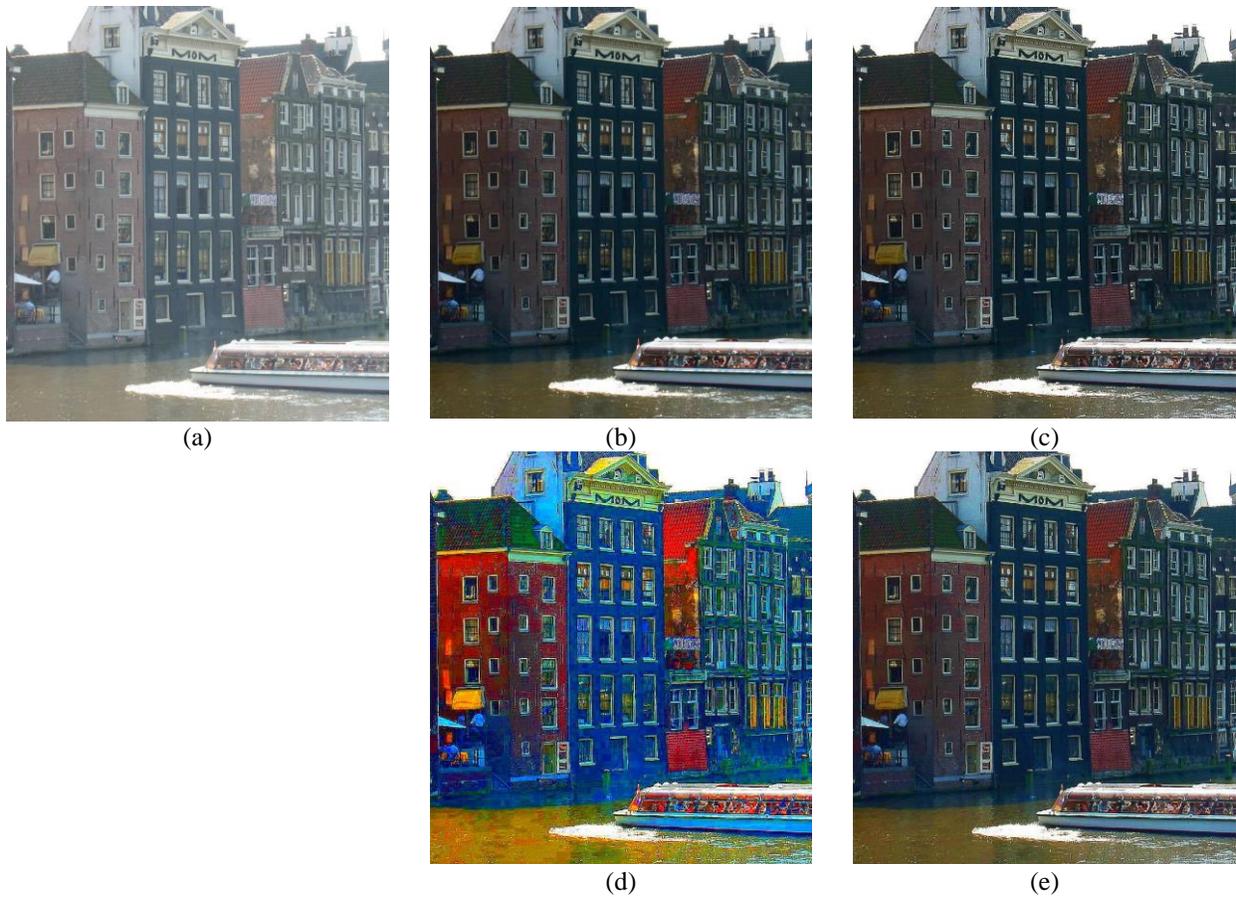


Figure 6-12. The intermediate results after each step in the 3DAWQF. (a) Original image; (b) Intensity adjusted image; (c) Contrast enhanced image; (d) Color enhanced image; (e) Final overall enhanced image.

6.2.3 Performance Comparisons

Some enhancement results from the 3DAWQF and other enhancement algorithms are presented in this section. The CRME and the CQM are used to evaluate the image qualities and to compare the enhancement results.

The “white house” image in Figure 6-13 is obtained from the NASA Retinex research website [149]. The original image is taken under strong sunlight and it shows loss of color and details in the shadows. The MSRCR [155] algorithm retains the details under the house shadows; however, it also brings halo artifacts to the bottom left trees. Compared with Wang’s algorithm [146], the 3DAWQF algorithm has better contrast especially in the railings portion of the white house and the bricks of the red house.



CRME = 2.0129
 CQM = 2.5414
 (a)



CRME = 2.2564
 CQM = 3.5047
 (b)



CRME = 2.2713
 CQM = 4.2474
 (c)



CRME = 2.7037
 CQM = 4.9496
 (d)

Figure 6-13. Experimental results for applying different color image enhancement algorithms. (a) Original image; (b) Enhanced image by Wang's algorithm; (c) Enhanced image by the Retinex algorithm; (d) Enhanced image by the proposed 3DAWQF algorithm. The CRME and the CQM indicate the performances of each algorithm and demonstrates that the 3DAWQF outperforms the other color contrast enhancement algorithms.

Figure 6-14 shows an example of processing a high dynamic range image. The original image is captured with a Canon camera in raw mode [156]. Capturing all details in the sunlight area as well as in the shadows is a challenging problem. Comparing with Wang's algorithm [146] and Meylan's retinex algorithm [156], more details are recovered in dark regions and more vivid color is shown with the 3DAWQF algorithm. The original image and the Retinex processing result is courtesy of Meylan et al [156].



CRME = 1.7561
CQM = 2.3134
(a)



CRME = 1.7480
CQM = 2.1741
(b)



CRME = 1.7522
CQM = 2.6869
(c)



CRME = 2.0501
CQM = 3.8034
(d)

Figure 6-14. Experimental results for applying different color image enhancement algorithms on different applications: (a) Original image; (b) Enhanced image by Wang's algorithm; (c) Enhanced image by the Retinex algorithm; (d) Enhanced image by the proposed 3DAWQF algorithm. The CRME and the CQM indicate the performances of each algorithm and demonstrates that the 3DAWQF outperforms the other color contrast enhancement algorithms.

It is seen that the alpha weighted quadratic filter is parametric. These parameters offer the quadratic filter more power and more design flexibility to meet the specific and complex requirements in real world applications. Therefore, a reliable objective measure that can help in selecting the optimal operating parameters is of paramount importance. In the experimental results shown in this chapter, the parameters of the alpha power are determined by both experimentation and the objective measures. Figure 6-15 shows one example of using the measures to assist in the selection of optimal operating parameters used in the 3DAWQF. The image in Figure 6-15 (b) is captured by a remotely operated underwater vehicle during the retrieval of Air Asia black box. The low visibility and strong color

casting bring extra difficulty for the underwater searching. To enhance the vision of underwater vehicles or robots, the 3DAWQF is used to enhance the contrast and recover the colors. The image in Figure 6-15 (b) is underwater wreckage from the tail part of the AirAsia flight. It is seen that the real color of the plane tail is not recognizable and the directly captured underwater image shows strong blue casting. In the 3DAWQF algorithm, a high pass filter is applied across the RGB color plane to recover colors and a parameter M is used to control the influence of neighboring pixels to the center pixel of the color cube. Figure 6-15 (c) to (f) show color enhancement results with variant parameter M used in Eq. (6.19). Recall that the UICM presented in Chapter 2 is designed for measuring the color property for underwater images. Therefore, the UICM is used to assist in the selection of the parameter M used in the color enhancement step in the 3DAWQF. It is seen that more color is restored with a smaller value of M . A greater value of M also introduces some artifacts. The color measure value UICM for the resultant fused images with parameter M ranging from 12 to 40 (Figure 6-16) confirms the observations that a smaller M corresponds to a better color quality.

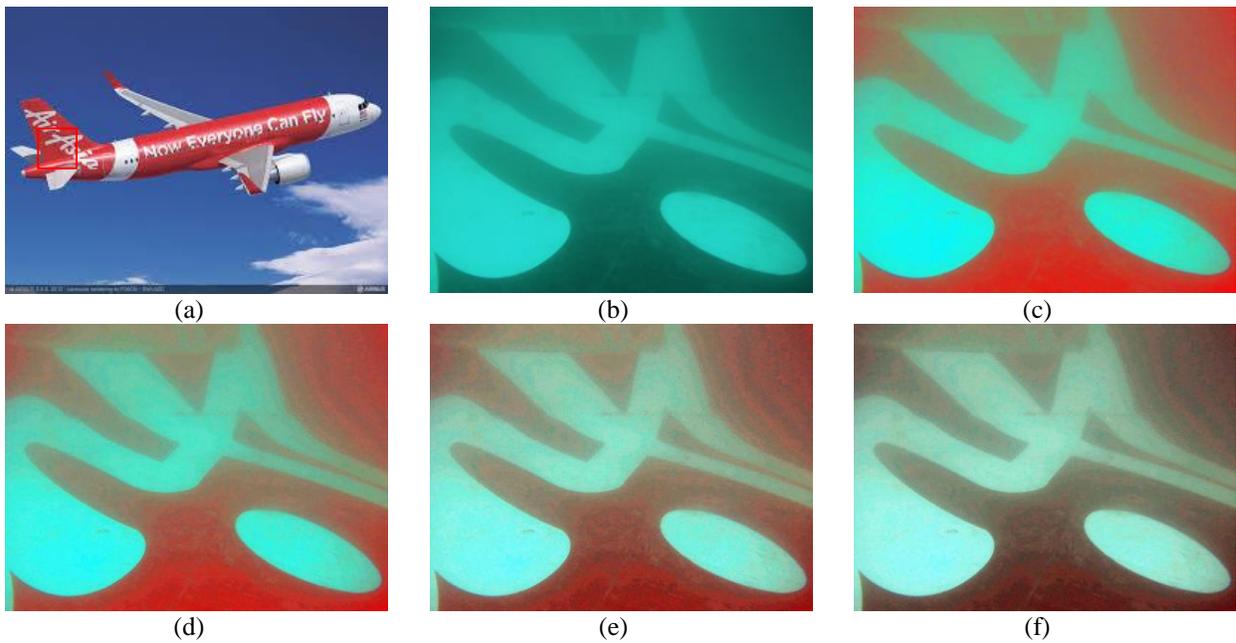


Figure 6-15. (a) An AirAsia plane; (b) a wreckage of the plane tail found underwater; (c) – (e) color enhancement results using the 3DAWQF with parameter M equals to 12, 20, 28, 36, respectively.

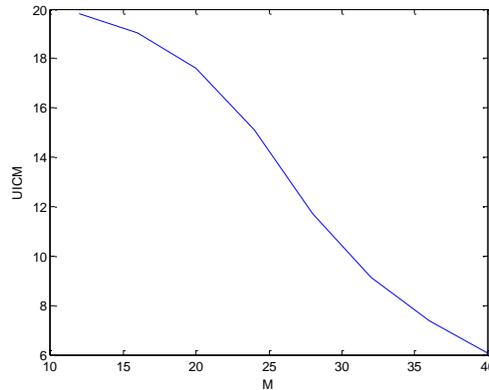


Figure 6-16. UICM measure values for variant parameter M from 12 to 40 in the 3DAWQF filter design.

6.3. Application 2: Edge Detection

6.3.1 Traditional Edge Detection Methods

Edge detection is important for many applications in computer vision and object tracking systems. As such, it is desired to develop edge detectors that can effectively and efficiently detect edges in images. Commonly used edge detection algorithms can be grouped into two categories: one utilizes the first order gradient information and the other uses second derivative zero crossing information. The Sobel edge detector is an example of a first order based edge detector. It is straightforward to implement but not effective in detecting thin or fine edges. The Laplacian of Gaussian is a typical zero crossing based edge detector, which offers better localization especially when edges are not sharp. However, it has a drawback wherein it cannot handle corner conditions well [157]. Although the performance of most of these detectors is acceptable for simple noise free images, the case is dramatically different for images of real scenes and images which are contaminated with noise [158]. One common challenge in existing edge detection algorithms is the trade-off between noise removal and edge preservation. The presence of noise may lead to false detection of edge pixels so smoothing is oftentimes used in conjunction with preprocessing the original image. The side-effects of the smoothing may result in the loss of fine edges. Thus, it is difficult to design an edge detector that performs well for both noise removal and edge preservation using the gradient based methods. An

example of using the traditional methods in detecting edges when noise exists is shown in Figure 6-17. The asphalt concrete image in Figure 6-17 (a) can be seen as naturally contaminated with heavy noise. For this image, the Canny and Sobel algorithms fail to detect the road crack. An edge detection result similar as Figure 6-17 (d), which removes a lot of asphalt concrete textures but preserve the crack, is desired.

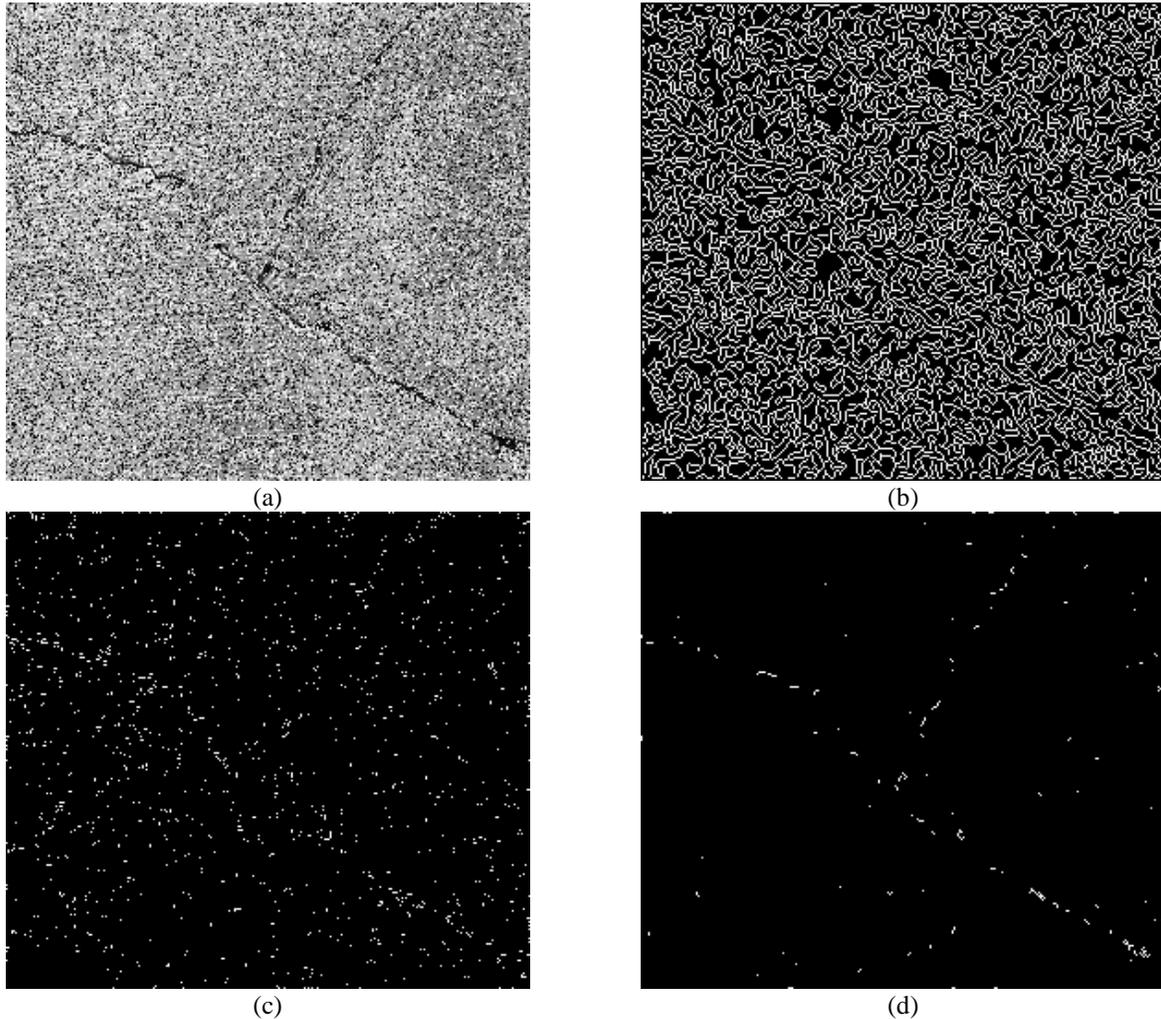


Figure 6-17. (a) An asphalt concrete image; traditional edge detection algorithm, such as (b) the Canny algorithm and (c) the Sobel algorithm, are not able to detect the edges effectively; (d) the presented AWQF based edge detection algorithm is able to detect the crack.

6.3.2 Edge Detection Using AWQF

As analyzed in the previous sections, the AWQF has the property of enhancing image details while simultaneously removing noise. Therefore, it is reasonable to utilize such characteristic to enhance the edge details and de-noise simultaneously. In this chapter, an edge detection method based on the AWQF is presented. The main

procedures and the corresponding intermediate results are demonstrated in Figure 6-18. The algorithms include the following steps:



Figure 6-18. Diagram and intermediate results of the new AWQF algorithm for edge detection

Step1: Preprocessing

For edge detection applications, the original image needs to be scaled before applying the AWQF in order to bring the background to a level zero and the image details to a level one [159]. In this way, the background can be considered as zero mean noise and thus can be smoothed by the $y0$ quadratic linear filter, which is designed as the

edge preserving smoother. Meanwhile, the image details can be enhanced mainly by the type1 quadratic term, which serves to detecting the presence of correlated details. In this chapter, Otsu's automatic threshold [160] is used to do banarization and the corresponding black regions are scaled to an average zero and the corresponding white regions are scaled to an average level one.

Step2: NAWQF Edge Enhancement

The edge information is usually extracted using the high pass filters, while noise removal is generally achieved by applying low pass filters. In this situation, the nonlinear filters can be effectively used. As a nonlinear filter which composed of the first and the second terms, the quadratic filter has the intrinsic ability to preserve contrast while simultaneously removing noise.

To recall the baseline filter designed in Chapter 6.1, the y_0 term, which is designed as the linear quadratic filter, serves as an edge preserving nonlinear smoother. It removes noise while minimizing the blurring of edges. The y_l term, which is the quadratic term, detects the presence of correlated details [161]. Leveraging on these advantages, related pixels are collected and grouped, further raised to a power of alpha. In this case, y_0 becomes:

$$y_0 = h_0 x_0 + h_1 (x_1 + x_3 + x_5 + x_7)^a + h_2 (x_2 + x_4 + x_6 + x_8)^b + w_0 x_0^{2c} + w_1 (x_1^2 + x_3^2 + x_5^2 + x_7^2)^d + w_2 (x_2^2 + x_4^2 + x_6^2 + x_8^2)^e \quad (6.20)$$

And y_l becomes:

$$y_l = s_1 (x_0 x_1 + x_0 x_3 + x_0 x_5 + x_0 x_7)^f + s_2 (x_0 x_2 + x_0 x_4 + x_0 x_6 + x_0 x_8)^g + r_1 (x_1 x_2 + x_3 x_4 + x_5 x_6 + x_7 x_8 + x_1 x_8 + x_3 x_2 + x_5 x_4 + x_7 x_6)^h + r_2 (x_2 x_8 + x_4 x_2 + x_6 x_4 + x_8 x_6)^i \quad (6.21)$$

Here, a negative power is introduced to the y_l quadratic term. In this way, both dark regions and bright regions are enhanced, and to some extent, may even result in being over enhanced. From a visual point of view, dark regions become bright while bright regions remain bright. Meanwhile, for those pixels along the edges, the neighboring pixels are of different gray scale levels; they compensate for each other, and thus, edge pixels survive the nonlinear filter processing. In this way, the original edge pixels are enhanced.

Notice all parameters in the alpha weighted quadratic filter work together to achieve the effect. Among all the power parameters, h in Eq. (6.21) works on all neighbors around the center pixel of a 3 by 3 window. Therefore, as guidelines for a user friendly interface, the recommended parameter set for the other power factors can be set as $a =$

3; $b = 1$; $c = 5$; $d = 5$; $e = 5$; $f = 1$; $g = 2$; $i = 1$ based on the contrast adjustment power law and experimental results. The user can then simply tune h to produce desired edge enhancement results. From the experiments shown in this chapter, h is tested between the range of -0.1 and -1.

The quadratic filter is applied on image local regions. When applying the quadratic filter for image contrast enhancement purpose, the mask size is usually selected to be proportional to the target size to achieve optimal results. To reduce the computational complexity, there are many ways, such as the downsampling method to reduce the mask size to a smaller one [161][162]. But for the edge detection purpose, to avoid missing thin edge information, the mask size is usually set as 3 by 3. After the NAWQF procedure, most edge pixels are already revealed in the enhanced grayscale image. The follow up steps are aimed to yield a binary edge map with continuous, single pixel width edge pixels.

Step3: Local adjustment

The enhanced gray image is processed locally by subtracting the mean or median value of each local window adaptively. By doing this, the uniform regions in the enhanced grayscale image are further removed.

Step4: HVS masking

Recall that the Buchsbaum curve in Figure 2-10 in Chapter 2 describes that the human perceived contrast is formulated as the luminance change over the background intensity. This luminance masking property is incorporated in the edge detection algorithm. Usually the HVS masking is applied on a gradient image obtained from directional kernels [163]. However, the AWQF is an isotropic design procedure, so there is no existing directional information available. Here, the masking procedure is applied on the local adjusted grayscale image. The pixels on the image is masked by

$$im_{mask} = \frac{im(x, y)}{\alpha + \left(\frac{|I(x, y) - I_0|}{I_0}\right)^\gamma} \quad (6.22)$$

Where im_{mask} denotes the processed gray image masked with local luminance, $I(x, y)$ is the local alpha-trimmed average luminance and I_0 is the average luminance of the entire image. The parameters α and γ are determined empirically.

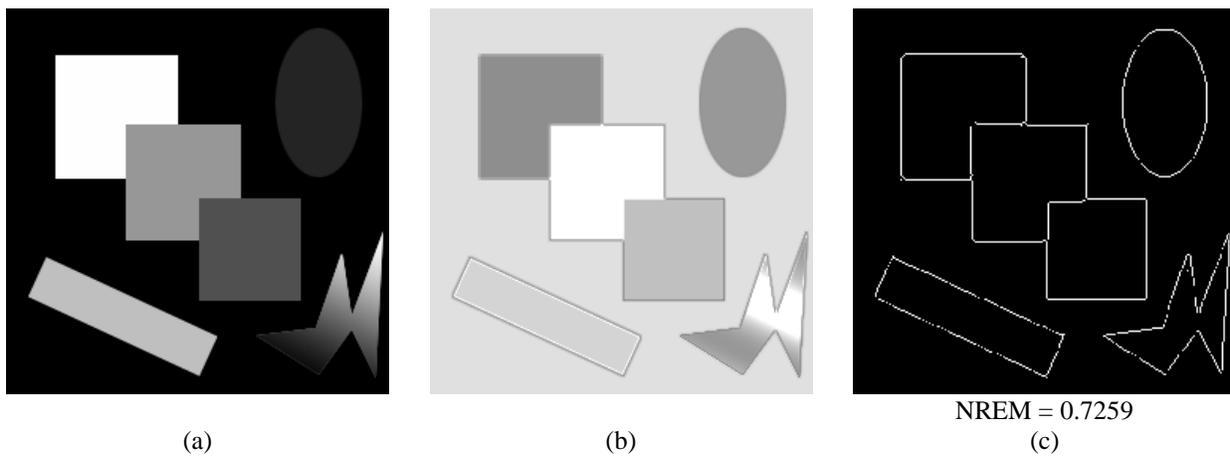
Step5: Hysteresis thresholding

To get a clean and continuous edge map, a hysteresis thresholding is used for labeling edges. A pixel is classified as a significant edge pixel if its intensity is greater than the high level threshold, and is classified as non-edge pixel if its intensity is smaller than the low level threshold. If its intensity value is in between the high and low level threshold, the connectivity with neighboring edge pixels should be considered. The pixel is classified as an edge pixel if it is connected with another significant edge pixel. Finally a morphological thinning is applied to ensure that edges in the edge map are one pixel wide.

6.3.3 Experimental Results

The AWQF algorithm is tested on 20 images including synthetic images, natural images, aerial images, medical images and cell phone camera images. These edge detection results are compared with those generated by the Canny and Sobel algorithms. In practice, users have the flexibility to manually tune the AWQF by adjusting its filter coefficients. The images in Figure 6-19 to Figure 6-24 reflect some of the challenging problems encountered by edge detection algorithms. Namely, they illustrate the challenges of detecting edges in images suffering from non-uniform illumination, low contrast, containing fine details, and non-flat surfaces.

The synthetic image in Figure 6-19 has spatially varying contrast due to non-uniform illumination. In this case, the Canny algorithm is prone to detect false edges, and the Sobel algorithm is not effective at detecting edges in the low contrast regions of the image. By comparison, the AWQF captures the essential edges in the synthetic image. The non-reference edge measure NREM also assigns the greatest measure value for the AWQF output.



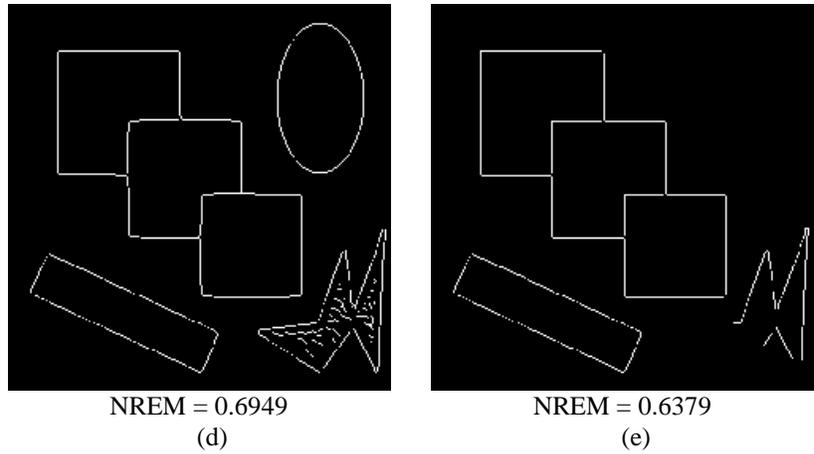


Figure 6-19. Edge detection results for a synthetic image. (a) Original image; (b) Enhanced grayscale image; Edge obtained by (c) the AWQF algorithm; (d) the Canny algorithm; (e) the Sobel algorithm.

The application of performing edge detection in medical images is shown in Figure 6-20. Medical images are always of low contrast and poor resolution, due to the limitations of hardware system or that may require noninvasive techniques. In this low contrast situation, the Canny algorithm is inadequate because it detects double lines and artifacts. The Sobel algorithm generates many noise-like textures in the final edge map. While the AWQF output yields a clean edge map and simultaneously capture the object structure due to the contrast enhancing nature of the quadratic filter. For these edge detection outputs, both the MOS and the NREM denote the AWQF edge map is the best. A zoom of the region of interests (ROI) is shown in Figure 6-21.

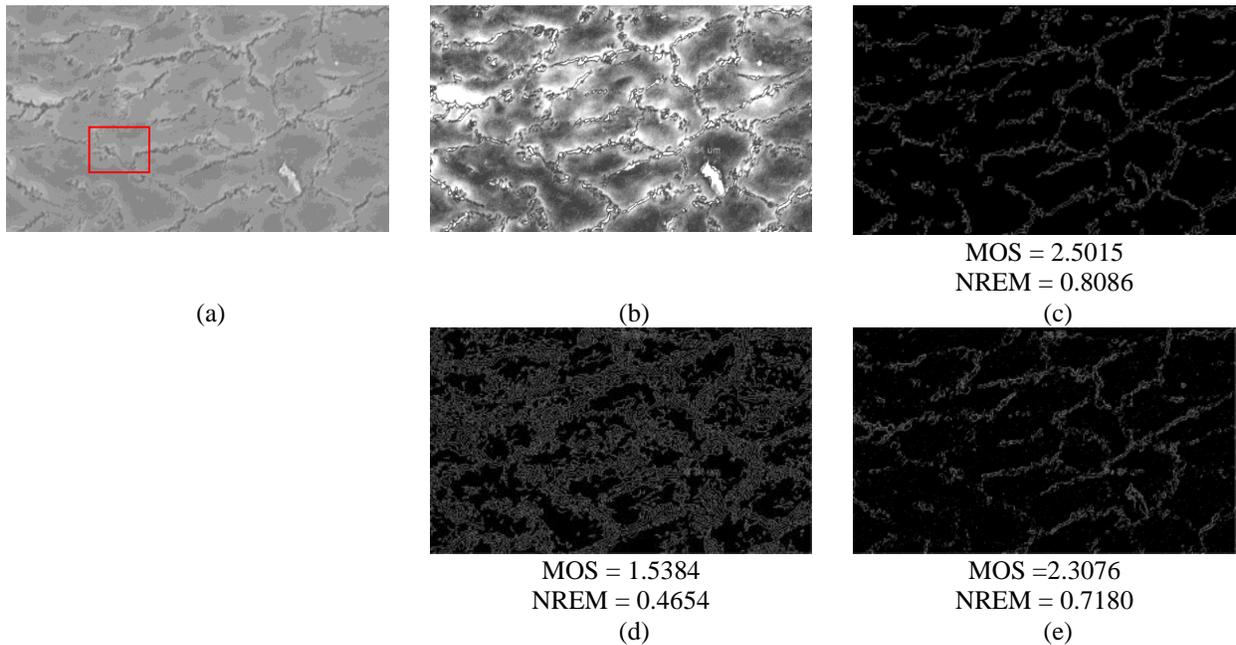


Figure 6-20. Edge detection results for a medical image. (a) Original image; (b) Enhanced grayscale image; Edge obtained by (c) the AWQF algorithm; (d) the Canny algorithm; (e) the Sobel algorithm.

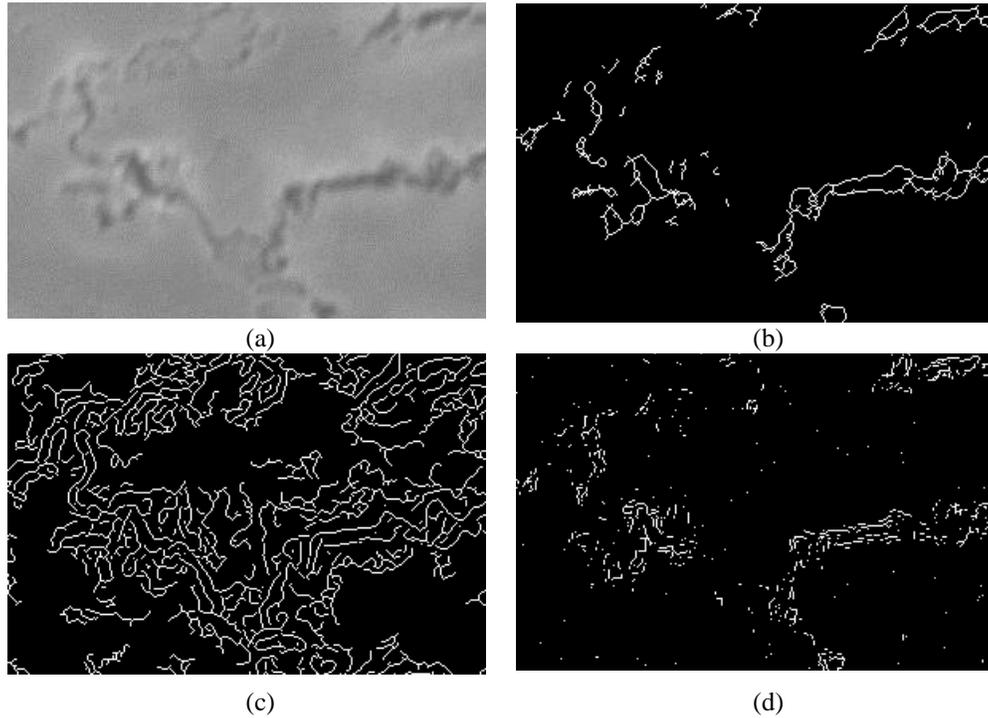


Figure 6-21. Zoom of ROI (Red Rectangle part) of the cell image in Figure 6-20. (a) ROI of the cell image; (b)-(d) are edges obtained by: (b) the AWQF algorithm; (c) the Canny algorithm; (d) the Sobel algorithm;

Figure 6-22 presents the edge maps of a statue image which contains many fine details from executing the Canny, Sobel, and two AWQF edge detectors. The Canny algorithm detects tiny details which are insignificant according to human vision. Notice that Sobel algorithm loses some essential edges. By comparison, the AWQF algorithm achieves a good tradeoff. Simulation results show that edges obtained from the AWQF are more consistent with the human vision system.



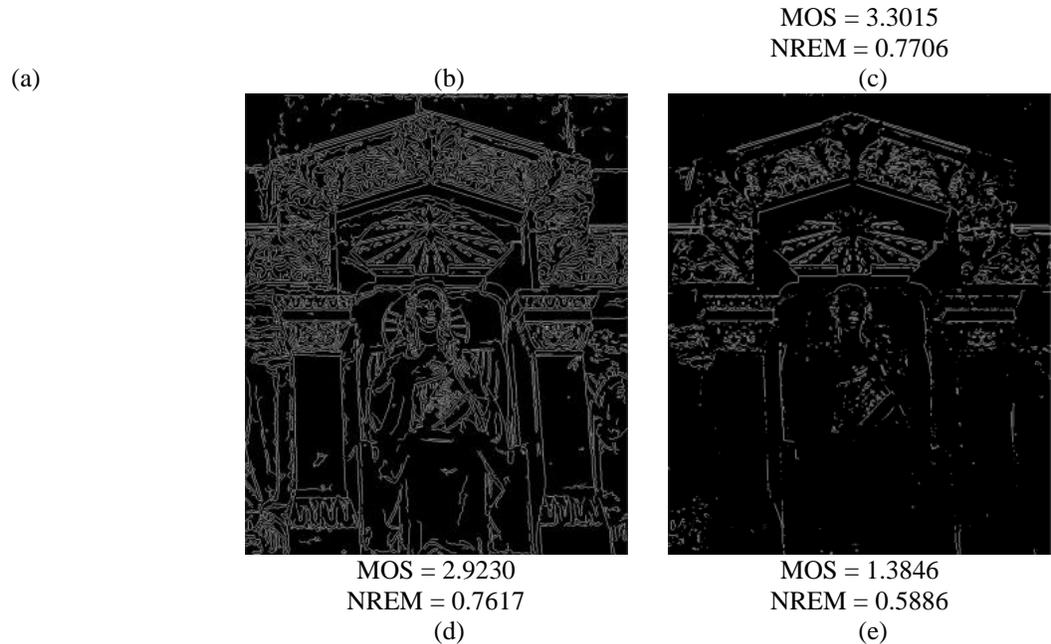
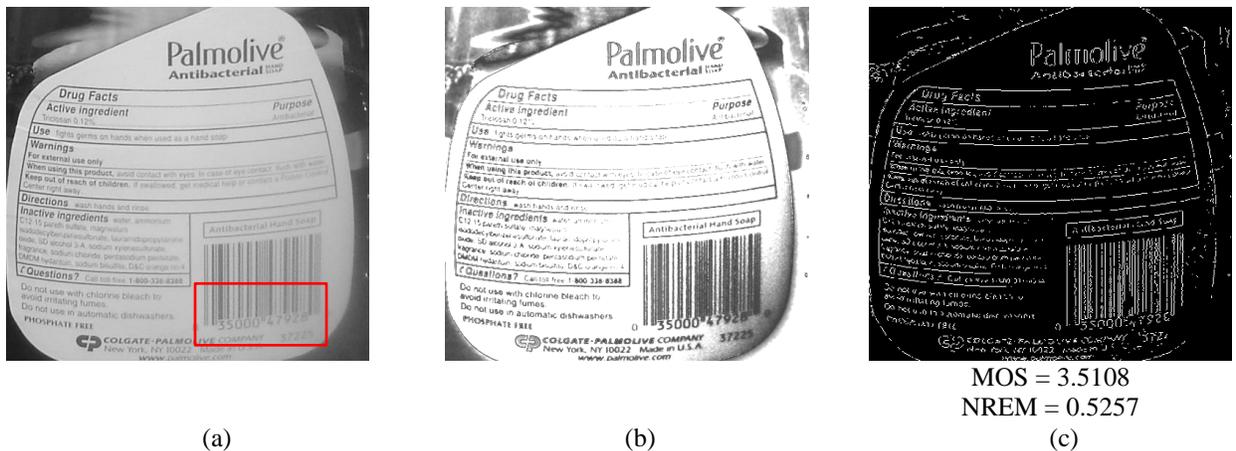


Figure 6-22. Edge detection results for a natural statue image. (a) Original image; (b) Enhanced Gray image; Edge obtained by (c) the AWQF algorithm; (d) the Canny algorithm; (e) the Sobel algorithm.

Figure 6-23 shows an example of detecting the texts and barcode on a bottle label. The non-flat surface on the bottle brings more difficulties to the edge detection. It is seen that the texts in the Canny output suffers from double edges, and the Sobel output misses the essential edges such as the barcodes. The label ‘Palmolive’ is more discernable, the barcode is more readable, and the numbers under the barcode are clearer in the AWQF output. The zoomed barcode detection results are shown in Figure 6-24.



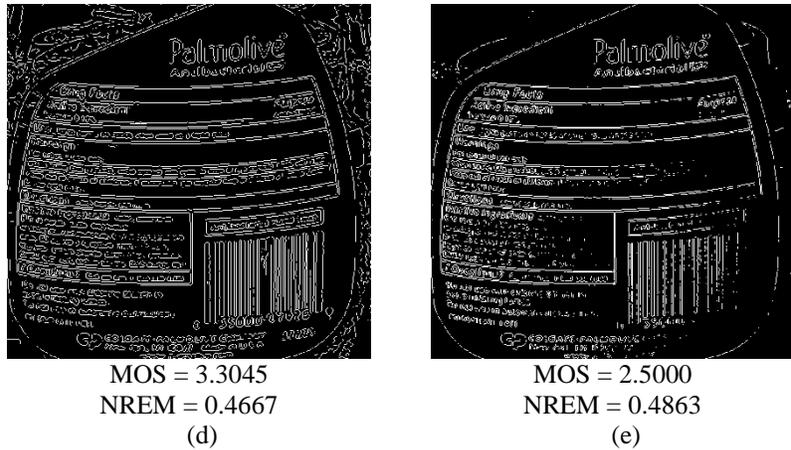


Figure 6-23. Edge detection results for a label image. (a) Original image; (b) Enhanced Gray image; Edge obtained by (c) the AWQF algorithm; (d) the Canny algorithm; (e) the Sobel algorithm.



Figure 6-24. Zoom of ROI (Red Rectangle part) of the bottle label image in Figure 6-23. (a) ROI of the bottle label; (b)-(d) are edges obtained by: (b) the AWQF algorithm; (c) the Canny algorithm; (d) the Sobel algorithm;

Since there is no ground truth for real images, the best evaluation of algorithm performance is by subjective evaluation Mean Opinion Score (MOS) suggested by ITU-T [16]. In the subjective tests, twenty six people were randomly selected. They varied in age, gender, race, and came from different disciplines. They were asked to assign a score based on visual recognition as follows: (1) bad; (2) poor; (3) fair; (4) good; (5) excellent for each of the edge maps obtained from the Canny, Sobel, and AWQF algorithms for the 20 testing images in a blind survey of results. A greater MOS value represents a better consistency with the human perception.

Figure 6-25 shows the obtained MOS and RBEM measure values for the 20 testing images. For most images, the AWQF edge detection algorithms scores higher than the Canny and the Sobel edge detection results. Table 6-2 shows that the AWQF edge maps have greater MOS values than the Canny and Sobel algorithms according to the MOS and the NREM. The AWQF algorithm also shows a small variance. These illustrate the robustness of the proposed algorithms for a wide variety of image. It's worth noting that the variance of the NREM is not used to evaluate the performances of the edge detection algorithms. This is because the NREM has an edge pixel number penalty term. For a natural image with many structures, more edge pixels exist in the edge map. For such images,

the NREM measure values are smaller but that does not imply the edge quality is bad. Accordingly, it is not reasonable to compare the NREM measure values for independent images. Instead, the NREM is only used to compare edge detection performances for images with the same content.

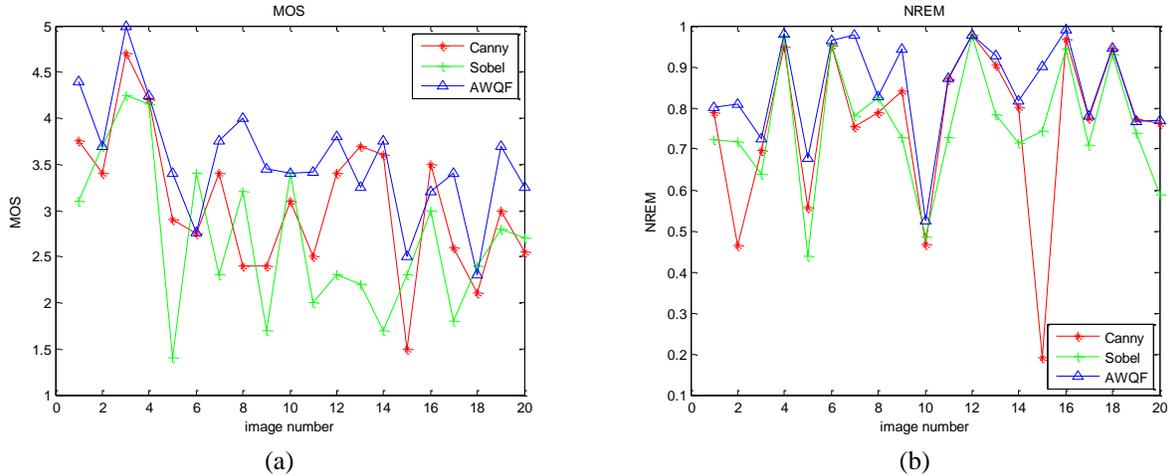


Figure 6-25. (a) Mean Opinion Scores(MOS) and (b) NREM measure values for the Canny, Sobel , AWQF1 and AWQF2 algorithms for the 20 testing images. Red: Canny, Green: Sobel; Blue: AWQF.

Table 6-2. Statistics of the MOS for the Canny, Sobel, and AWQF algorithm for 20 test images. The AWQF has the greatest average score and smallest variance according to human observers

	Canny	Sobel	AWQF
Average MOS Score	3.0686	2.6813	3.5010
Average RBEM Score	0.7608	0.7562	0.8493
MOS Variance	0.5671	0.6296	0.3821

6.4. Summary

New Alpha Weighted Quadratic Filter based color image contrast enhancement algorithm and edge detection algorithm are introduced in this chapter:

- For images with less than desired illuminations, a modified Naka-Rushton function is applied to adjust the intensities. Then a PAWQF is applied within each color component to enhance the intra-plane image details. Inter-plane filter is applied on four pseudo planes to enhance the color details. The analysis and experimental results demonstrate the good performance of the 3DAWQF in improving the visual quality in terms of illumination, color, and contrast. Experiments confirm that the image color and contrast attribute measures, the

3D contrast measure CRME, as well as the overall quality measure CQM can effectively evaluate color image quality as similar way as human observers.

- The introduction of a negative power of alpha extends the application of quadratic filter for edge detection. The negative power suppresses the uniform portions of the image while further enhance the edge portions. Comparing with the Canny and Sobel algorithms, the NAWQF edge detection algorithms have been shown to provide better visual performance for edge detection. It provides many benefits especially in detecting edges in the presence of noise. Unlike other edge detection algorithms which detect edges from the image itself, the AWQF starts from the texture of the image and enhances edge information, thus resulting in an edge map that is more akin to human vision. The non-reference edge measure NREM and the subjective evaluation MOS both agree that the new edge detection method outperforms existing Canny and Sobel edge detection algorithms over a wide range of images.

Chapter 7. Practical Applications of the Measures

The non-reference image and video quality measures can serve for many industrial applications. Recently, objective measures have been used in evaluating the visual aesthetic appeal of consumer images and videos [164][165]. A similar idea can be used to evaluate the freshness of meats and fruits because the visual aspects of the food reflect its quality. It can also be used for electrical and mechanical system inspection, where the overheating or cooling down of equipment demonstrates abnormal colors in the thermograph. The non-reference measures also provide an alternative idea of retrieving images by features. Previous chapters provide several image and edge attribute measures, as well as a general framework for evaluating image, video, and edge qualities. For specific applications, users need to determine which attributes of images to measure, as well as the values of the combination coefficients. Some practical design examples for industrial uses are shown in this chapter. The experiments are conducted under the assumption that ideal lighting conditions are satisfied. It is admitted that in real practical applications, such conditions may not always be reached, but the results in the initial tests make the potential industrial uses of the measures very promising.

7.1. Color Patch Retrieval

Color patch retrieval returns the most similar colors in the database for a given input image color patch. Traditionally, color retrieval is conducted by comparing the distances between the input color patch and all candidate color patches under a certain color model. However, many existing color spaces and color distances are not able to correctly capture color differences as human vision [31]. For example, the color yellow in Figure 7-1 (a) has the RGB value (255-255-0). The two colors in Figure 7-1 (b) and (c) has the same Euclidean distance to the color yellow in the RGB color plane but they look quite different.

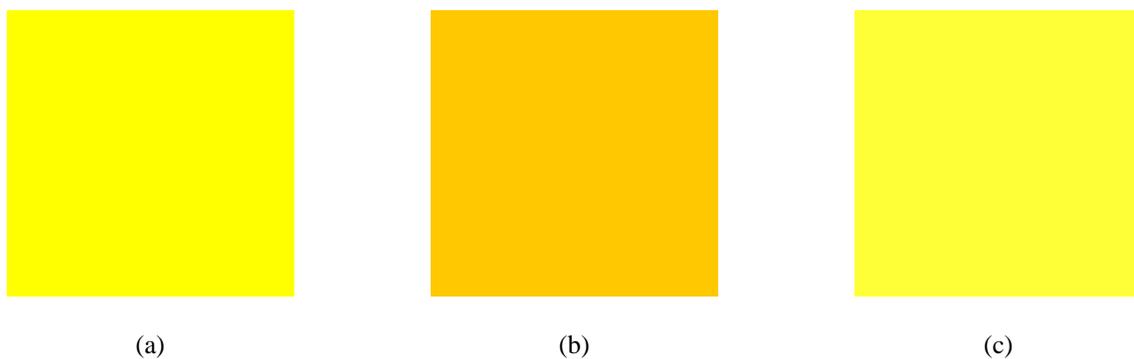


Figure 7-1. (a) Yellow (RGB: 255-255-0); (b) RGB: 255-255-55 (c) RGB: 255-200-0. It is seen that (b) and (c) have the same mean value in the RGB color space, but they look quite different. Also, (b) and (c) have the same Euclidean distance to the (a), but visually (c) looks more similar to (a).

The non-reference color measure provides an alternative method of retrieving images. For each color patch, the non-reference color measure assigns a unique index value for it. In the retrieval procedure, only the indexes are compared. Figure 7-2 shows an example of the color patch retrieval for the color yellow (RGB: 255-255-0). In this experiment, two databases with 4913 and 132651 color patches are generated by automatically by a variation of R, G and B values ($0 < R, G, B < 255$) using an increment equal to 15 and 5, respectively. The left column in Figure 7-2 shows the retrieval results on the small database, and the right column shows the retrieval results on the larger database. Figure 7-2 (a) and (b) demonstrate the first 48 retrieved colors provided by the Euclidean distance in the RGB and LUXV color spaces. Figure 7-2 (c) shows the first 48 colors with the most similar color measure values with the measure value of color yellow. The results show that the color measure performs satisfactorily in classifying and distinguishing colors. When there are more color patches in the database, the retrieval precision increases.

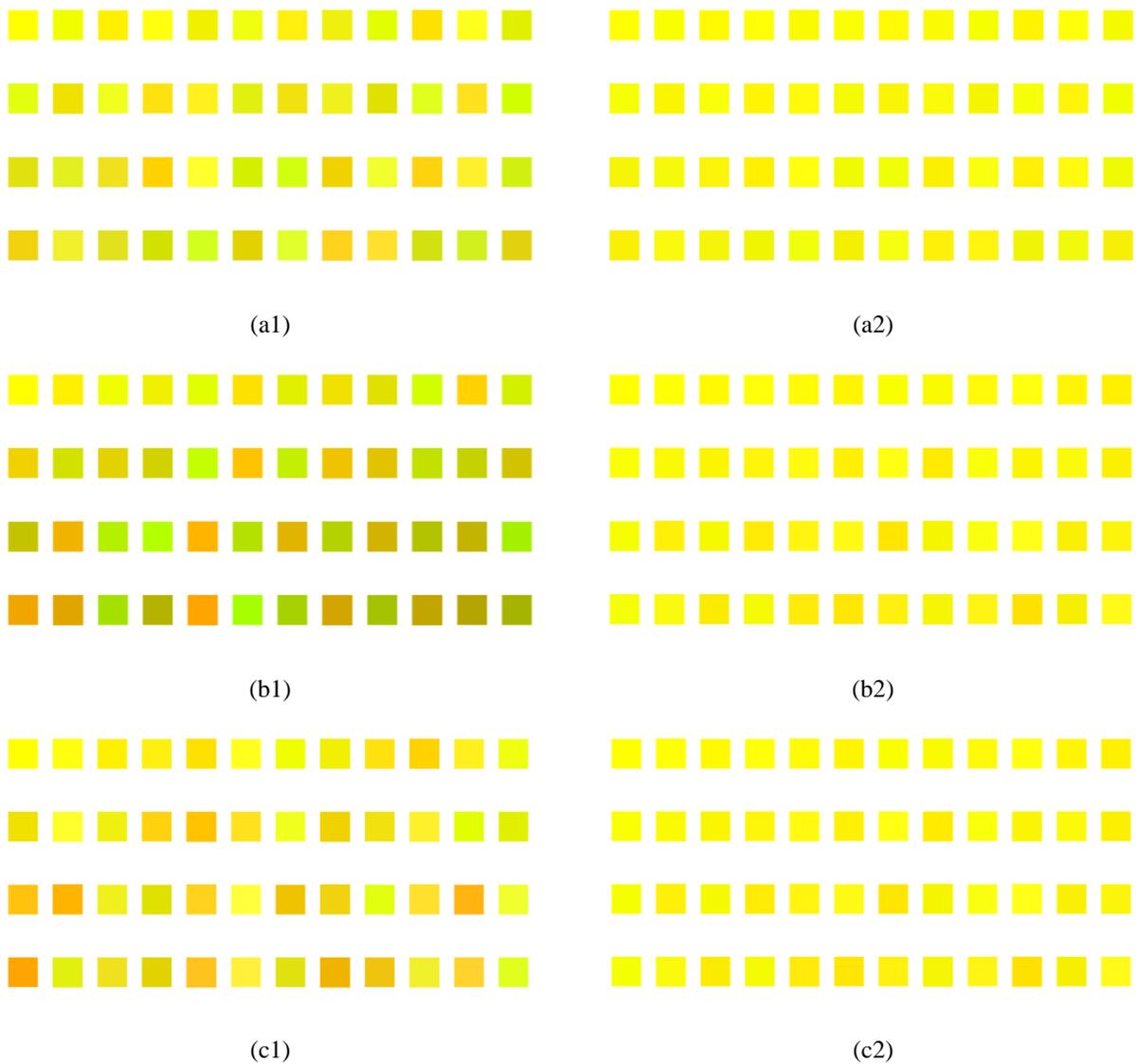


Figure 7-2. First 48 retrieved color patches for yellow (255-255-0). (a) using Euclidean distance on RGB color space. (b) using Euclidean distance on LUXV color space; (c) Using color measure. The left column shows the retrieval results from 4193 color patches, and the right column from 132651 color patches.

More importantly, unlike the distance based methods, in the color measure based retrieval, the calculations of colorfulness for all candidate colors are performed offline and the colorfulness measures for each color patch are saved. Without the time-consuming distance calculations for each candidate colors, the non-reference measure based retrieval system is much more efficient because it only needs to compare the numerical color measure values. Table 7-1 shows the average retrieval time for obtaining the first 48 images of the three methods shown in Figure 7-2. Obviously, the color measure based method has much shorter retrieval time. The time-efficiency property of the measure based method is more obvious when the size of the candidate database grows. At this scenario, the

execution time for the distance based methods increases significantly because the calculations need to be done for all candidate images during the retrieval procedure.

Table 7-1. Average execution time for retrieving the first 48 most similar color patches. The average retrieval time using the distance based measures increases greatly when the size of the candidate database increases. By comparison, the time complexity of the measure based retrieval method increases reasonably.

Candidate database size (color patch number)	Euclidean distance on RGB color plane (second)	Euclidean distance on LUXV color plane (second)	Colorfulness measure (second)
4193	0.0101	0.0628	0.0075
132651	20.2733	22.1392	0.0577

The good performance of the color patch retrieval using the non-reference color measure makes it applicable for many industrial applications. For example, in the painting industry (Figure 7-3), users can capture images of the wall at their home using the cellphone camera and send it to the painting shop. The painting shop manager can use software to calculate the color value for this picture and find the most similar color in the inventory according to the color measure value instead of comparing it with the color chart exhaustively. Similarly, it could also be used in the beauty industry, where users can simply take a picture of their lips, face, or hair, and get a color measure for it using cellphone apps, and compare it with the color measure values labeled on the products.

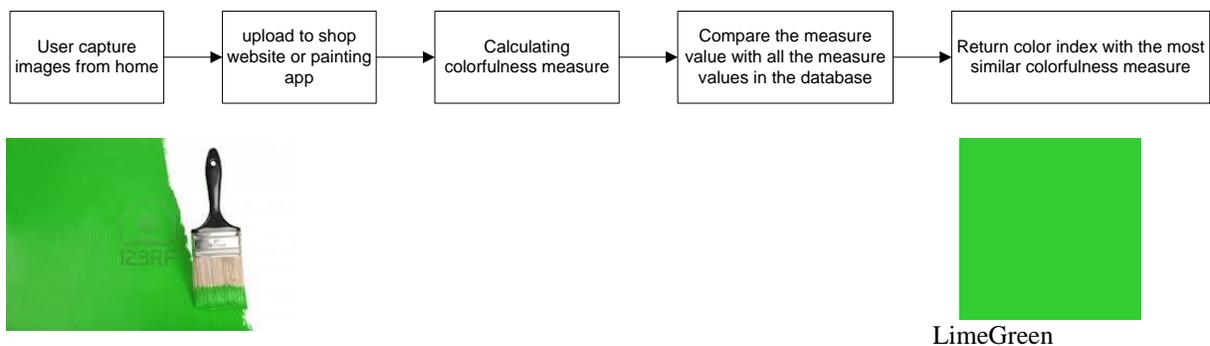


Figure 7-3. Color matching for the painting industry

7.2. Natural Color Image Retrieval

Instead of using the traditional methods, which compare the distances between the query image and all images in the database, the non-reference measure assigns a unique value for each candidate image and only compares the numbers. This new idea significantly reduces the computational cost during the retrieval procedure and it provides

potential uses in the fast database search. This section shows using the non-reference measures for searching for natural images with the most similar attributes.

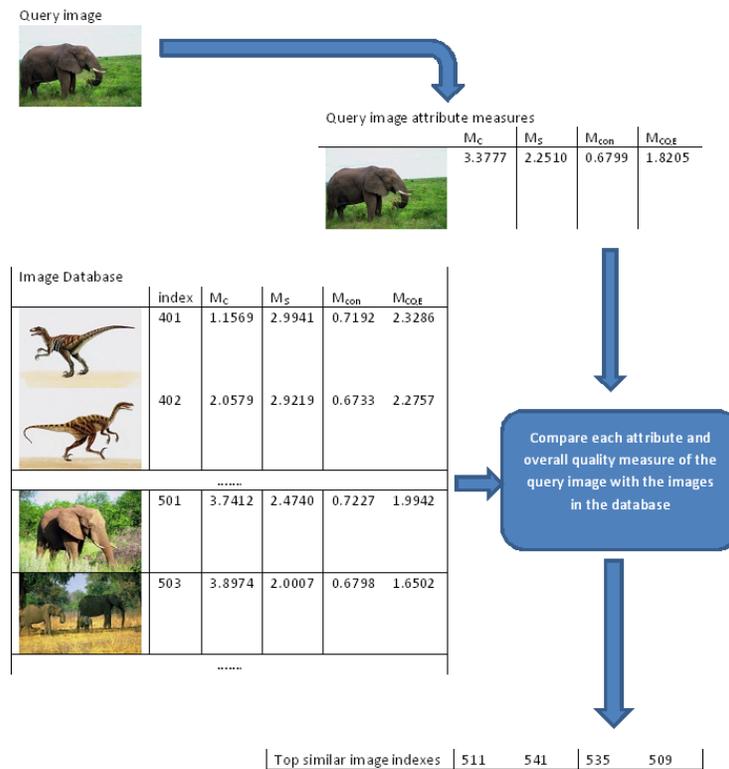
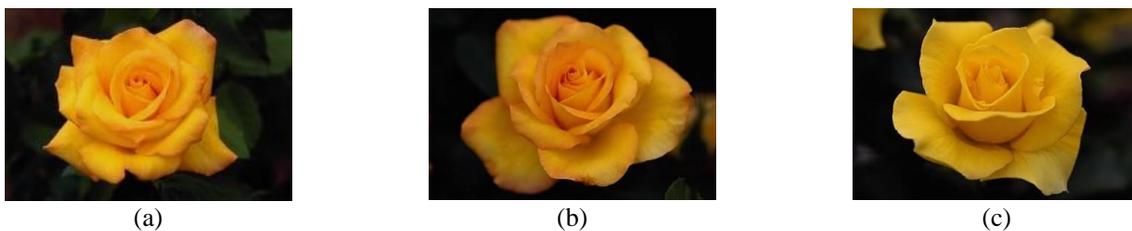


Figure 7-4. An illustration of the database search application by comparing the image attributes measures.

The flower database in [166] is a benchmark used for image retrieval and it is used in this section to test the effectiveness of using the color measure in natural image database search. This database contains 100 images including flowers with different sizes, colors and shapes. It is observed that there exist strong similarities for flowers that belong to the same type. The results of using the color measure to retrieve a yellow flower image are shown in Figure 7-5. The first five retrieved results show similar color as the query image.



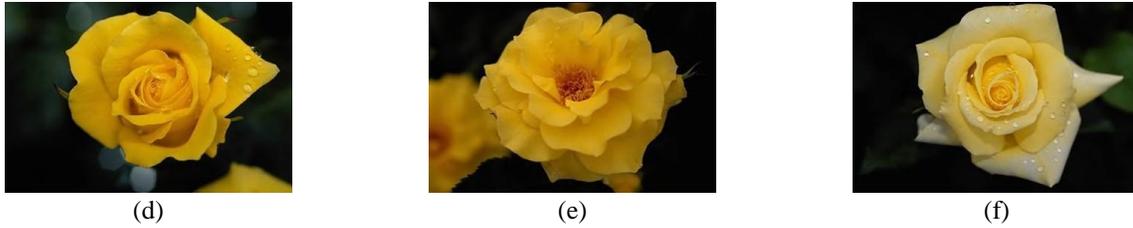


Figure 7-5. Using the color measure for color image retrieval. (a) Query image; (b)-(f) first five retrieved images. The measure based retrieval returns flowers with similar colors.

The Wang database in [166] is a benchmark used for color image retrieval and it is used in this section to test the effectiveness of using the NR color measures to perform image search and retrieval from a database. This database contains 10 image classes with 100 images each. It is observed that there exist strong color similarities for images that belong to the same category. As shown in Figure 7-6, the CQM measure values are distinguishable for the category of Buses, Dinosaurs and Elephants. Therefore, it is reasonable and promising to achieve fast database search or image classification using the proposed NR measures. The results of using NR measures to retrieve some query images are shown in Figure 7-7.

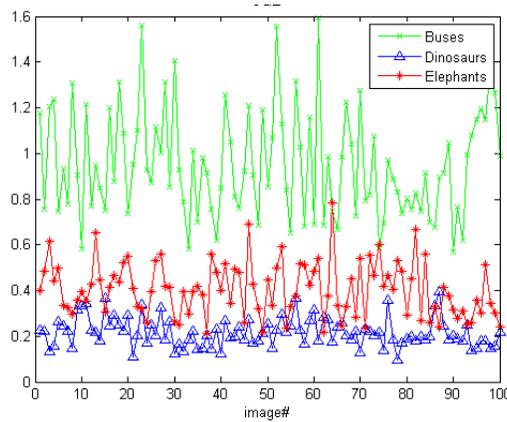
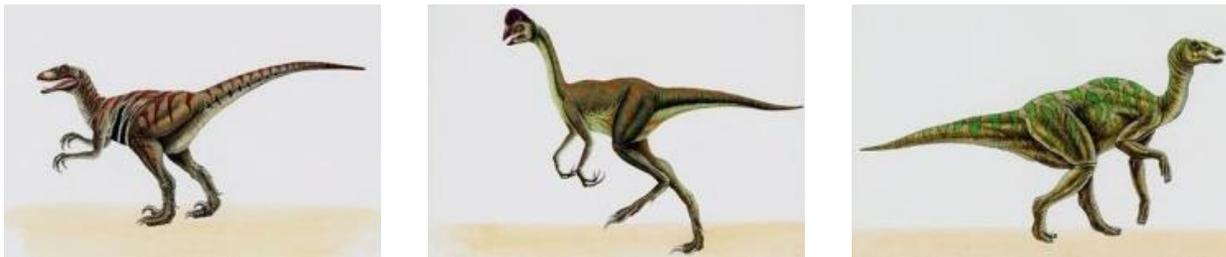


Figure 7-6. CQM measure values for images belongs to different classes in Wang database. CQM measure values for Elephants (red), Buses (green) and Dinosaurs (blue)



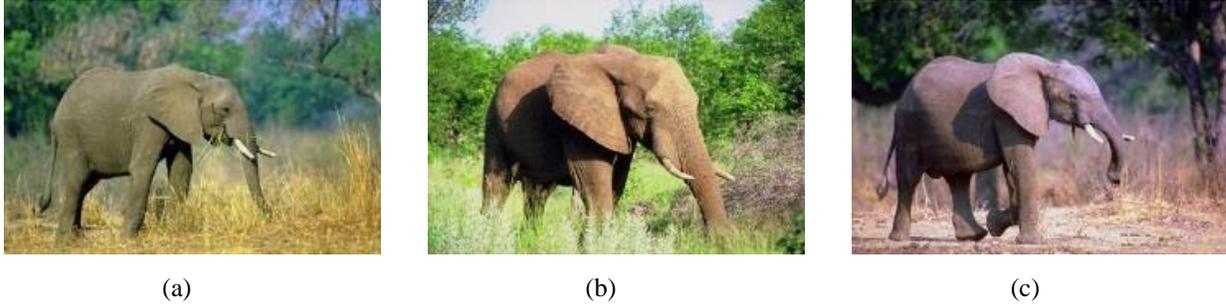


Figure 7-7. Using color measure for dinosaurs and elephants image retrieval. (a) Query image; (b) first retrieved image; (c) Second retrieved image

7.3. Electrical Component Thermography Inspection

Monitoring the thermograph of the electrical equipment is important for maintaining the reliability of the electrical system. The thermal images indicate the temperature of the electrical components. Generally, the higher the temperature, the warmer the image is. When the equipment is overheated or cooling down, the color in the thermal images changes. At this situation, the color measure can tell when this problem happens. If the thermograph is divided into regions and applies the color measure on each region, the location of the problematic component is also accessible.

Practically, a thermal image of the electrical equipment can be captured when the system is working in normal condition, and the color measure of this thermal image is defined as the reference. Then the infrared camera monitors the equipment and takes thermal images periodically. The color measure value of the periodically taken thermal images is analyzed. If the color measure is above a certain threshold, an alert of overheating should be generated, and if the colorfulness is below a certain threshold, a system restart alert should be generated. In Figure 7-8, the first column shows the original thermal images of the electrical components, the second and the third columns show the generated images with 40% less and 40% more contrast, which simulate the abnormal cases [167]. It is seen that the images with less contrast also shows less colorfulness and the images with more contrast has better color. The color measure value confirms the observations.

Original

40% less contrast

40% more contrast

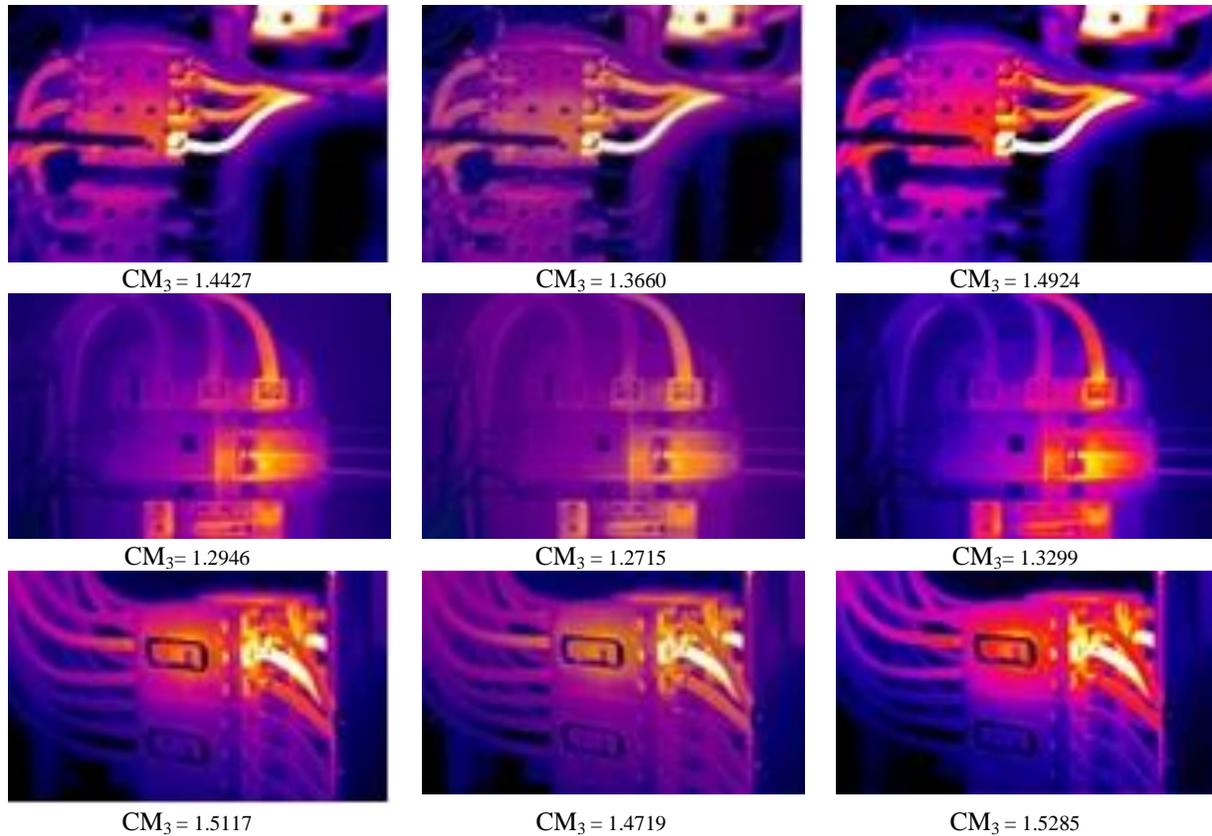


Figure 7-8. Electrical system inspection using colorfulness measures on electrical thermography. When the electrical components are overheated, the thermal image is warmer. So the color measure value is greater than the one measured under normal condition. On the other hand, when the electrical components are cooling down, the thermal image is cooler. In this case, the color measure value is decreased.

7.4. Medical Image Quality Evaluation

Medical images usually suffer from low contrast and heavy noise, due to the limitations of the hardware system or limitations on allowable exposure time, such as in the case of X-rays. Therefore, many image enhancement algorithms are applied after the image acquisition procedure to enhance the visibility of image details and further to assist doctors' diagnosis. For medical applications, the reference images are not available so the non-reference quality measures are desired.

Figure 7-9 shows an example of using variant algorithms to enhance the colors in a skin cancer image. It is seen that the Wang's algorithm (Figure 7-9 (b)) does not effectively enhance the contrast, and the Retinex algorithm (Figure 7-9 (c)) has a color casting problem. In contrary, the 3DAWQF in Figure 7-9 (d) enhances the image without color distortion and the textures on the skin and the melanoma are more discernable.

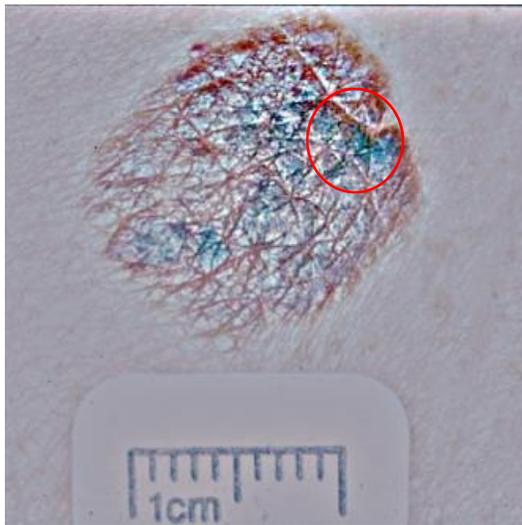


(a)



(b)

CQM = 3.0044
MOS = 5.2340



(c)

CQM = 0.1216
MOS = 5.0638



(d)

CQM = 5.7339
MOS = 6.2553

Figure 7-9. (a) Original skin cancer microscope image; Enhanced image by (b) Wang's algorithm; (c) MSRCR algorithm; (d) 3DAWQF algorithm. The color measure CQM indicates the performances of each algorithm and the CM_3 generate the same ranking as the MOS.

Figure 7-10 (a) shows a mammogram image with low contrast. For this image neither Wang's algorithm (Figure 7-10 (b)) nor the Retinex algorithm MSRCR (Figure 7-10 (c)) enhances the mass portion. The block artifacts and halo effects from these two algorithms may mislead the surgeon's diagnosis. By comparison, the 3DAWQF (Figure 7-10 (d)) enhances the breast cancer in the mammogram image, which is useful information for surgeons. The MOS

for these two images are obtained from 48 people and for both examples, the CQM generates the same ranking as the MOS.

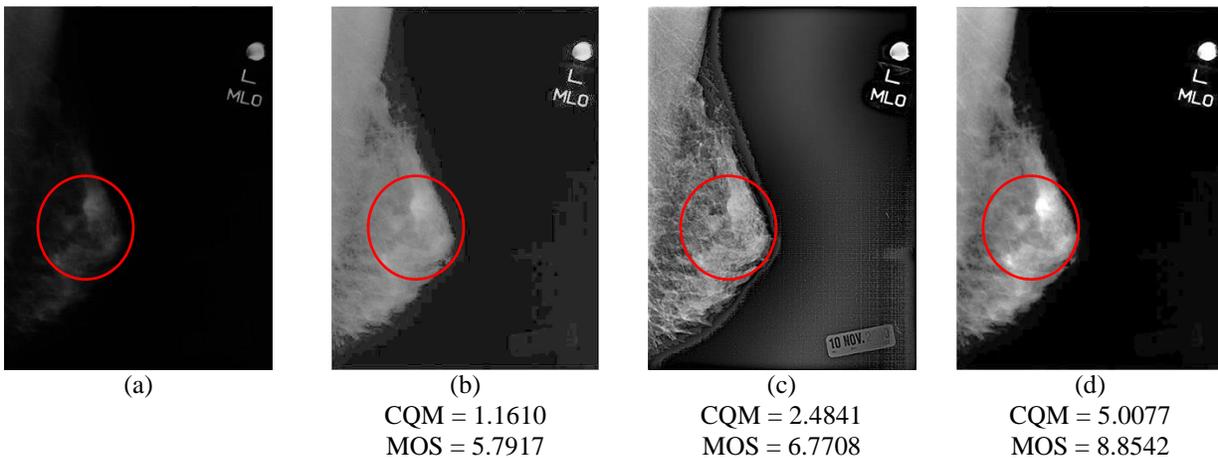


Figure 7-10. (a) Original breast cancer mammogram image with low contrast. Enhanced images by (b) Wang's algorithm; (c) MSRCR algorithm; (d) 3DAWQF algorithm. The CQM indicate the performances of each algorithm and the CQM generate the same ranking as the MOS.

Edge detection is an essential pre-processing step for early cancer detection and diagnosis in medical image processing such as medical image segmentation, registration and reconstruction. For example, accurate edge detection algorithms can be used to track the size of a tumor and this information can help to monitor whether the treatment is effective or not. In Figure 5-4, multiple edge detection results from the Canny, Sobel, Roberts, Log, and Prewitt for a CT abdomen image are shown from Figure 5-4 (b) – Figure 5-4 (f). These edge detection algorithms are commonly used edge detection algorithms and each has its advantages and disadvantages. For example, gradient based edge detection algorithms such as Sobel and Prewitt are simple but sensitive to noise. The Canny edge detector improves the signal to noise ratio by smoothing the image, however, the smoothing may lead to loss of corners and detection of double edges. Therefore, it is necessary to have a reliable edge measure that can help to decide the optimal edge detection algorithm for a specific image. The NREM selects the Sobel edge detection result as the optimal and this result agrees with the visual assessment. As a comparison, the Canny and LoG edge subtract all the soft tissues inside the abdomen, while the Roberts edge has the disconnection problem on the key edges.

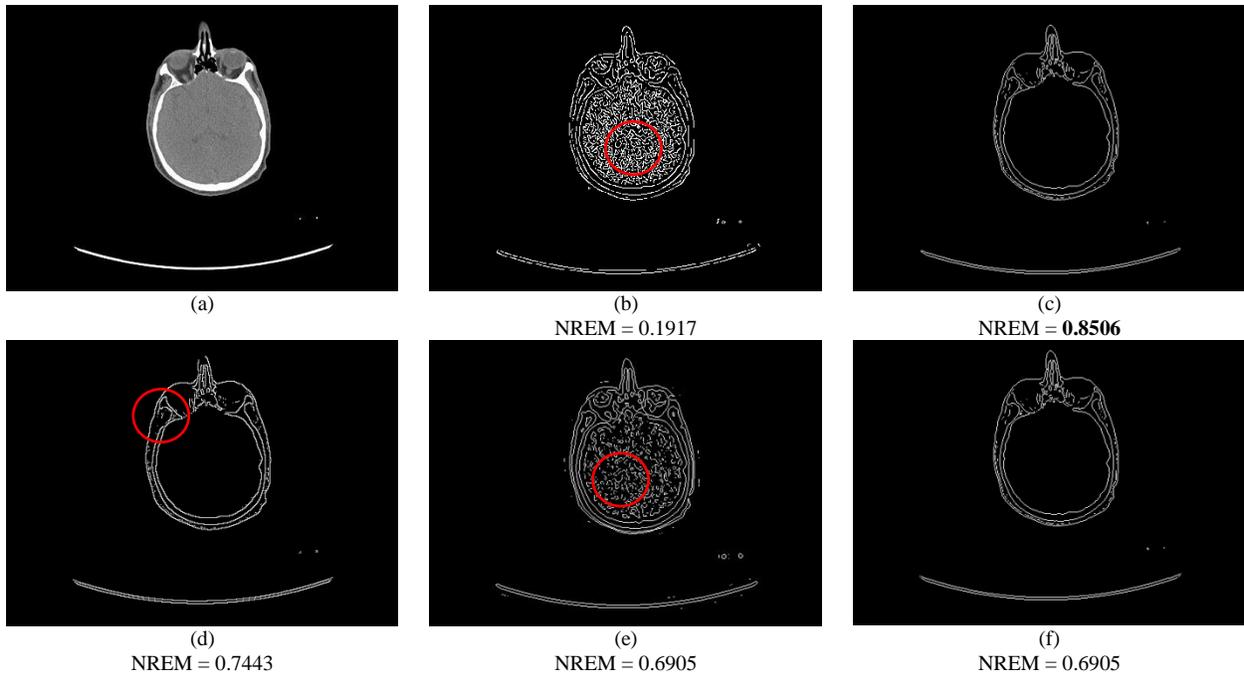


Figure 7-11. Using measures to select optimal edge detection algorithms. (a) Original CT abdomen image. (b)-(f) Edge detection results from: (b) Canny, (c) Sobel, (d) Roberts, (e) LoG, (f) Prewitt.

7.5. Food Quality Evaluation

The measure can be used to quantitatively evaluate freshness or class of a food product without a human expert. This has significant practical uses for the reason that, in food industry, some food quality features are suitable for automatic measurement such as color, size, shape, and texture. The measurement values can be saved in a database for training purpose. When a new product is put in the system, it is able to be automatically labeled what type of product it is, how fresh it is, and whether the size or shape is qualified in the market. Therefore, computer vision technology is possible to be used as an inexpensive alternative to manual expert grading of the food product in the future [30]. Figure 7-12 shows an automatic food freshness classification system.

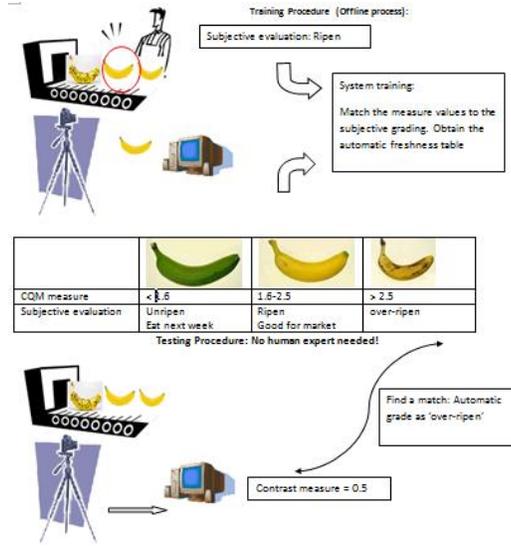


Figure 7-12. An illustrative banana freshness classification system using colorfulness measures without interaction with human judgement.

During the training procedure, the subjective evaluation is involved. The human experts manually classify the bananas into three classes: unripen, ripen, and overripen. Meanwhile, the attribute measure values are calculated. Among these attribute measures, the contrast, colorfulness, and color tone are observed to be significantly correlated with the overall banana quality. Therefore, for the banana quality evaluation system, the CQM is formed as shown in Eq. (7.1). The CQM plots for 35 training banana images are shown in Figure 7-13. Therefore, threshold [1.6,2.5] can be obtained from Figure 7-13. For a new banana image, if the CQM measure value is smaller than 1.6, it is classified as an unripen banana; if the CQM measure value is greater than 2.5, it is classified as an overripen banana; if the CQM measure value is between 1.6 and 2.5, it is classified as a ripen banana. With the measure values, the banana quality can be evaluated automatically without the intervention of human.

$$CQM = 24.03 \times MEMEE - 15.86 \times M_{C3} + 24.47 \times Colortone_3 \quad (7.1)$$

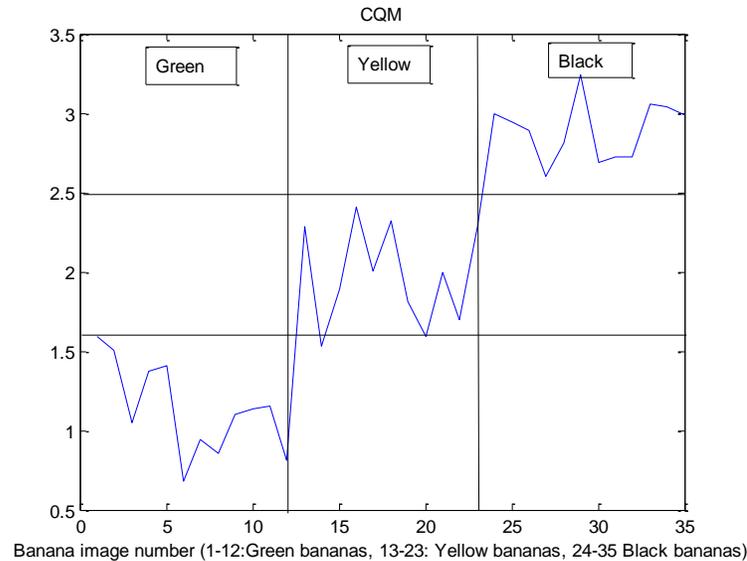


Figure 7-13. CQM measure values for 35 training banana images. Image 1 to 12 are unripen banana images, the CQM measure values are smaller than 1.6; image 13 to 24 are ripen banana images, the CQM measure values are between 1.6 and 2.5; image 24 to 35 are overripen banana images, the CQM measure values are greater than 2.5.

It is understood that in the practical banana freshness classification system, many factors affect the measure values such as the lighting conditions and the viewing angles. Some initial testings are conducted in the lab under the controlled conditions. Figure 7-14 shows three banana images. It is seen that more black dots show on the skin when the banana is overripen. The obtained obcontrast measure MEMEE increases for the banana with more black dots. It is also observed that the color of the banana skin changes from green to yellow to black so the colorfulness and color tone measure values of the images change as well. The overall CQM for the three bananas belong to different regions: the CQM for Figure 7-14 (a) is smaller than 1.6 which corresponds to unripen bananas, the CQM for Figure 7-14 (b) is greater than 1.6 but smaller than 2.5 which corresponds to ripen bananas, and the CQM for Figure 7-14 (c) is greater than 2.5 which corresponds to overripen bananas. Therefore, these measures can be used to evaluate the banana ripening stages.

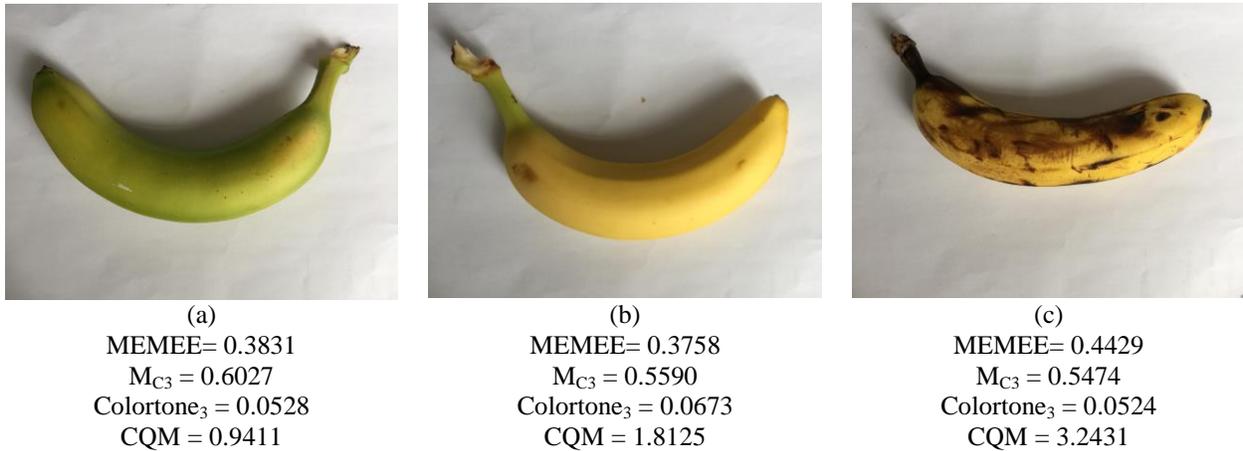


Figure 7-14. Three testing banana images. The attribute measures are shown under each image. The CQM measure values for the unripen, ripen, and over ripen bananas falls in different ranges.

7.6. Underwater Inspection Assistance

The quality of underwater image plays a pivotal role in ocean engineering, scientific research and surveillance applications, such as monitoring sea life, accessing geological environment, locating enemy submarines, explore underwater mines, and ocean rescue. Images taken by underwater vehicles or aquatic robots suffer from blurring effect, low contrast, grayed out colors and color castings due to the absorption and scattering of the water media. Many contrast enhancement and color correction algorithms have been proposed. Their performances determine how much knowledge is correctly revealed from the low visibility underwater images. Therefore, a good measure is desired to assist in the selection of the optimal enhancement algorithm.

One example is shown in Figure 7-15. Figure 7-15 (a) shows an underwater image with different fish species in an overall view (original image and the processed images by courtesy of Chambah et al [48]). This image suffers from poor color dynamic range and strong cyan-green cast. Figure 7-15 (b) – Figure 7-15 (f) demonstrate different enhanced images. Among these enhancement results, the one obtained with the ACE method without the ‘keep original gray’ feature (Figure 7-15 (f)) is visually the best. Using this algorithm, the color cast is removed and the resultant image is chromatically diverse. The CQM confirms the observation and denotes Figure 7-15 (f) as the optimal enhanced image. The Hasler’s colorfulness measure for general natural images [43] is used for comparison

purpose here. Contrarily, it is seen that the Hasler's method considers Figure 7-15 (e) is the optimal image without taking the green color cast into consideration.

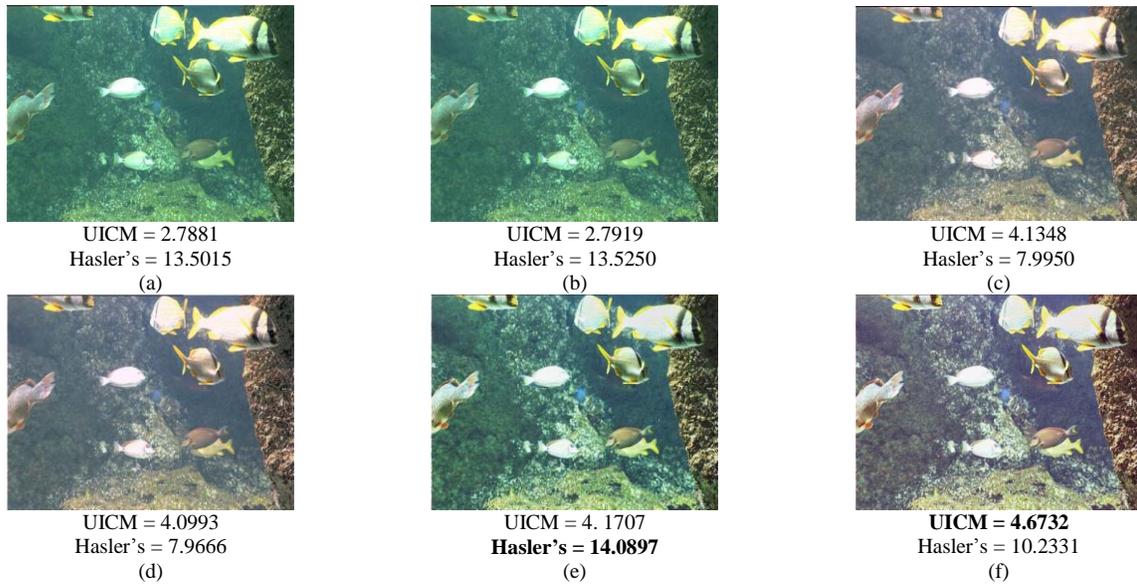


Figure 7-15. (a) original image; Color enhanced images using (b) WP; (c) GW; (d) GWWP; (e) ACE with original gray level preserved; (f) ACE without original gray level preserved. Images courtesy of Chambah et al [48].

Chapter 8. Concluding Remarks

8.1. Dissertation Contributions

Human evaluation is expensive with respect to time and resources. More importantly, it cannot be automated. Therefore, developing reliable objective measures are of paramount importance. In this dissertation, new measures for objectively evaluating image, video, and edge map qualities are presented.

A new measure system which is composed of fusions of image, video, and edge attribute measures is developed. Multiple image and edge attribute measures are presented and used in the system including: (1) generic image colorfulness measures M_{C1} , M_{C2} , M_{C3} , and M_{C4} based on combinations of image high order statistical values (Chapter 2.2.2); (2) a special colorfulness measure designed for underwater images UICM (Chapter 2.2.2); (3) an image color tone measure Unic-Color Differentiation (UCN) which is formulated as functions of the chrominance triplet (Chapter 2.2.3); (4) an image contrast measure MEMEE using Michelson contrast on local regions with new definitions of maximum and minimum (Chapter 2.3); (5) a color contrast measure CRME by evaluating the Root Mean Square in the color cube (Chapter 2.3); (6) an image sharpness measure using weber contrast law on grayscale edge maps (Chapter 2.4); (7) an edge pixel presence and localization measure D_p which compares the distances between falsely detected pixels and the corresponding ideal pixels (Chapter 4.3.1); (8) a corner presence and localization measure D_c which compares the distances between falsely detected pixels and the corresponding ideal pixels (Chapter 4.3.2); (9) a double edge occurrence measure D_{DE} by comparing downsampled edge maps (Chapter 4.3.3).

For specific applications, users have the flexibility of selecting appropriate attribute measures. In this dissertation, generic image and video quality measure framework, as well as the edge measure framework are demonstrated. For Color image Quality Measure (CQM), image colorfulness, color tone, sharpness, and contrast are selected (Chapter 3), while for Reference Based Edge Measure (RBEM), edge presence/localization, corner presence/localization, and double edge occurrence are used (Chapter 4). Users also have the flexibility to select the fusion method to combine these attribute measures and further obtain the overall image and edge quality measures. In this dissertation, linear combination is used, and two new training methods are presented for obtaining the linear combination coefficients. For image processing purpose where images before and after processing have the same

source, Linear Mixed Effect (LME) model is used to train the coefficients (Chapter 3.3). For video processing purpose where video frames are independent, the Multiple Linear Regression (MLR) model is used (Chapter 3.4).

A non-reference reconstruction based edge measure NREM is also presented (Chapter 5). This edge measure is based on the observation that the best edge map results consist of the least number of edge pixels at their correct locations needed to characterize all the relevant structure in the reconstruction image.

These newly developed measures are compared with existing state of the art measures. The presented measures are demonstrated to have strong correlations with the subjective evaluation gathered from Mean Opinion Scores (MOS).

One application of the objective measures is directing the development of new image processing algorithms. In this dissertation, several new image processing algorithms are developed including: (1) a positive alpha weighted quadratic filter (PAWQF) for color contrast enhancement (Chapter 6.2), and (2) a negative alpha weighted quadratic filter (NAWQF) for edge detection (Chapter 6.3). The developed measures CQM and NREM are used to validate the effectiveness of the new image processing algorithm, as well as assist the selection of optimal operating parameters for color enhancement and edge detection algorithms respectively. Each step of the PAWQF color contrast enhancement algorithm is aimed to enhance one aspect of image property. Therefore, image color and contrast attribute measures are used to show the effectiveness of each step. The newly developed algorithms outperform existing color enhancement and edge detection algorithms from both subjective evaluations and the objective evaluation measure values.

The measures are also demonstrated to be useful practical industrial applications. The measures are effectively used for fast painting and cosmetic color retrieval (Chapter 7.1). Natural scenes can be fast searched by colors as well (Chapter 7.2). Combining color and contrast measures, the electrical and mechanical system conditions can be monitored automatically (Chapter 7.3). Medical diagnosis assistance can be achieved using image quality and edge quality measures (Chapter 7.4). The quality of food products can be evaluated automatically by color and contrast measures (Chapter 7.5). Measuring the color desaturation and blurring effects caused by the forward and backward scattering of the water media can greatly assist the underwater research and exploration.

The contributions of the thesis are illustrated in Figure 8-1.

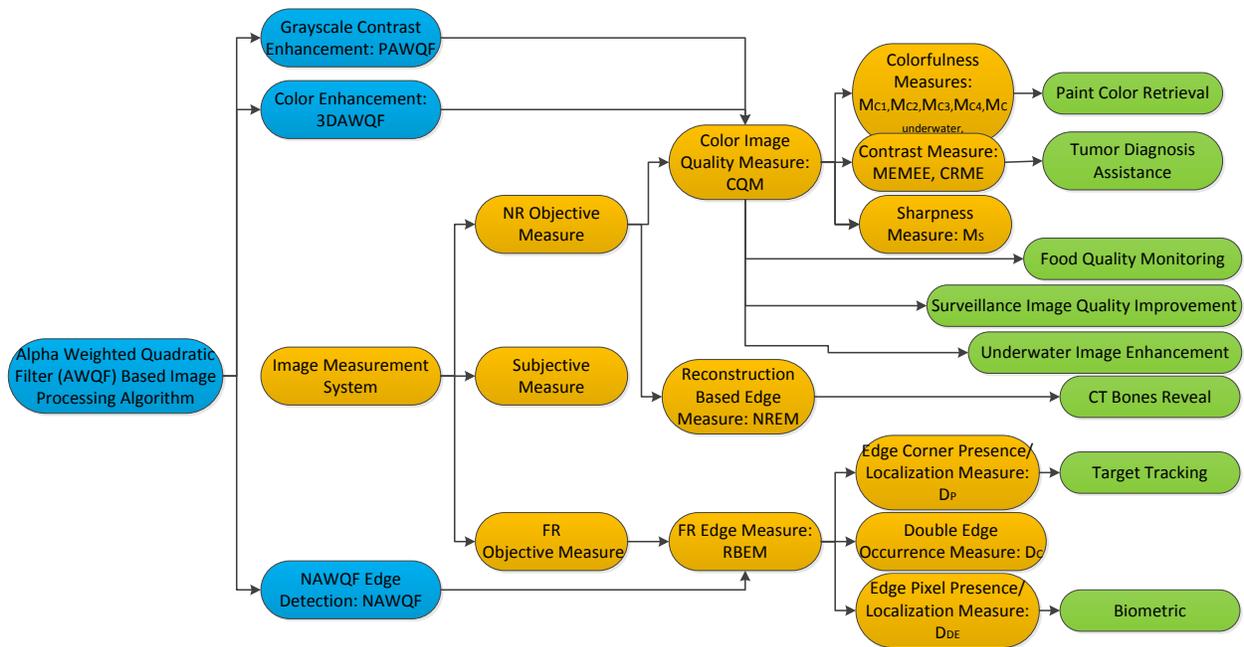


Figure 8-1. Summary of the contributions

8.2. Future Research Directions

The dissertation shows industrial applications of the measure systems under the controlled environment. Practically, such ideal situations may not always be satisfied. For example, users may capture images under colored light sources, or the captured image may be out of focus. In such situations, a certain preprocessing step that can standardize images before using the measures should be developed.

The linear combination coefficients in the CQM are trained for given types of distortions for image processing purposes. The measure values of the CRME are interpreted differently for high frequency distortions and low frequency distortions. In the future, it is desired to develop a certain method that can determine the type of distortions in an image automatically. Consequently, the coefficients can be automatically selected and the measure values can be automatically interpreted.

When training the linear combination coefficients, images and MOS from published database are used. The obtained coefficients are for general uses. For specific applications, images belong to the specific type and MOS gathered from experts in that area can be used to further improve the performance of the measures. For example, for

medical image quality evaluation, gathering the MOS by consulting doctors' and radiologists' opinions can improve the accuracy of the measures.

This dissertation presents the fast database retrieval by color. Practically, images belong to the same type also show strong similarity in size, shape, and textures. The measures on these image properties, such as a contour measure can be developed. Multiple image attribute measures can be used together in the fast database search application.

The measure system framework presented in this dissertation has many implications. Any newly developed image and edge attribute measures can be incorporated into this measure system. Image attributes such as brightness, transparency and color naturalness, as well as edge attributes such as edge pixel width and connectivity can be further developed.

Appendix A: New measures presented in the dissertation

Colorfulness measures
$M_{C1} = 0.02 \times \log\left(\frac{\sigma_{a,RG}^2}{ \mu_{a,RG} ^{0.2}}\right) \times \log\left(\frac{\sigma_{a,YB}^2}{ \mu_{a,YB} ^{0.2}}\right)$
$M_{C2} = 0.02 \times \frac{\log \sigma_{a,RG}^2 \times \log \sigma_{a,YB}^2}{\log \sigma_{a,c}^2} \times \frac{\log \mu_{a,RG}^2 \times \log \mu_{a,YB}^2}{\log \mu_{a,c}^2}$
$M_{C3} = \frac{1.33(\sqrt{\mu_{a,U}^2 + \mu_{a,X}^2 + \mu_{a,V}^2} - 200) + 2.39\sqrt{\sigma_{a,U}^2 + \sigma_{a,X}^2 + \sigma_{a,V}^2} - 0.49\sqrt{k_{a,U}^2 + k_{a,X}^2 + k_{a,V}^2}}{200}$
$M_{C4} = 0.0614\sqrt{\mu_{a,U}^2 + \mu_{a,X}^2 + \mu_{a,V}^2} + 0.1546\sqrt{\sigma_{a,U}^2 + \sigma_{a,X}^2 + \sigma_{a,V}^2} + 0.6642\sqrt{\gamma_{a,U}^2 + \gamma_{a,X}^2 + \gamma_{a,V}^2} - 0.2426\sqrt{k_{a,U}^2 + k_{a,X}^2 + k_{a,V}^2} - 13.0000$
$UICM = -0.0268\sqrt{\mu_{a,RG}^2 + \mu_{a,YB}^2} + 0.1586\sqrt{\sigma_{a,RG}^2 + \sigma_{a,YB}^2}$
Color tone measures
$ColorTone(r, g, b) = \frac{ \alpha(r-g) + \beta(r-b) + \gamma(g-b) }{ \alpha(r+g) + \beta(r+b) + \gamma(g+b) }$
$ColorTone(r, g, b) = \left \frac{r+g+b}{3} \right $
$ColorTone(r, g, b) = \alpha \left \frac{(r-g)}{(r+g)} \right + \beta \left \frac{(r-b)}{(r+b)} \right + \gamma \left \frac{(g-b)}{(g+b)} \right $
$UCD = \frac{1}{\#ofnonzeros(ColorTone)} \sum ColorTone \cdot \log(ColorTone)$
Image contrast measures
$MEMEE_{contrast} = \frac{1}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \frac{I_{max,k,l}}{I_{min,k,l}} \log\left(\frac{I_{max,k,l}}{I_{min,k,l}}\right)$
$I_{min,k,l} = \sum_{i=1}^l X_i, I_{max,k,l} = \sum_{i=t+1}^K X_i$
$CRME = \frac{1000}{k_1 k_2} \sqrt{\sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \left(\frac{\log \left I_{i,j} - \sum_{c=1}^3 \lambda_c \frac{I_{c1} + I_{c2} + \dots + I_{cn}}{n} \right }{\log \left I_{i,j} + \sum_{c=1}^3 \lambda_c \frac{I_{c1} + I_{c2} + \dots + I_{cn}}{n} \right } \right)^\alpha}$
Image sharpness measure

$M_S = \sum_{c=1}^3 \lambda_c EME_{sharpness}(edgemaps_c)$ $EME_{sharpness} = \frac{2}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \log\left(\frac{I_{max,k,l}}{I_{min,k,l}}\right)$
Color image and video quality measure
$CQM = c_1 \times colorfulness + c_2 \times sharpness + c_3 \times contrast + c_4 \times colortone$
Edge presence/localization measure
$D_{PFP} = \frac{1}{mn - P_O} \sum_{i=1}^{P_{FP}} 1 - \frac{1}{1 + \alpha d_{PFP_i}^2}$ $D_{PFN} = \frac{1}{P_O} \sum_{i=1}^{P_{FN}} 1 - \frac{1}{1 + \alpha d_{PFN_i}^2}$ $D_P = \frac{D_{PFP} + D_{PFN}}{2}$
Corner presence/localization measure
$D_{CFP} = \frac{1}{mn - C_O} \sum_{i=1}^{C_{FP}} 1 - \frac{1}{1 + \alpha d_{CFP_i}^2}$ $D_{CFN} = \frac{1}{C_O} \sum_{i=1}^{C_{FN}} 1 - \frac{1}{1 + \alpha d_{CFN_i}^2}$ $D_C = \frac{D_{CFP} + D_{CFN}}{2}$
Double edge occurrence measure
$d_{DE1} = (box(y_1)_O + box(y_2)_O + box(y_3)_O + box(y_4)_O) - (box(y_1)_T + box(y_2)_T + box(y_3)_T + box(y_4)_T)$ $d_{DE2} = (box(y_5)_O + box(y_6)_O) - (box(y_5)_T + box(y_6)_T)$ $D_{DE} = \frac{\frac{ d_{DE1} }{4m_1 n_1} + \frac{ d_{DE2} }{2m_2 n_2}}{2}$
Full reference edge measure

$$RBEM = 1 - \frac{\beta_P D_P + \beta_C D_C + \beta_{DE} D_{DE}}{\beta_P + \beta_C + \beta_{DE}}$$

Non reference edge measure

$$NREM(i, e) = (MGSSIM)^\alpha f_p^\beta$$

Appendix B. Acronyms Used in the Dissertation

FR	Full-reference
NR	Non-reference
HVS	Human Visual System
FOM	Figure of Merit
RG-YB	Red Green-Yellow Blue
PLIP	Parametric Logarithmic Image Processing
MOS	Mean Opinion Score
LM	Luminance Masking
CM	Contrast Masking
JND	Just Noticeable Difference
CQM	Color Quality Measure
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
MSE	Mean Square Error
LME	Linear Mixed Effect
MLR	Multiple Linear Regression
ROC	Receiver Operating Characteristic
SSIM	Structural Similarity Measure
GSSIM	Gradient Structural Similarity Measure
AWQF	Alpha Weighted Quadratic Filter
AWD	Automatic White Balance

- [1] K. Panetta, S. Agaian, and C. Gao, "Systems and Methods for Image and Video Signal Measurement," 2013.
- [2] K. Panetta, C. Gao, and S. Agaian, "No reference color image contrast and quality measures," *IEEE transactions on Consumer Electronics*, vol. 59, pp. 643-651, 2013.
- [3] K. Panetta, C. Gao, S. Agaian, and S. Nercessian, "Nonreference Medical Image Edge Map Measure," *International journal of biomedical imaging*, vol. 2014, 2014.
- [4] K. Panetta, C. Gao, S. Agaian, and S. Nercessian, "A New Reference Based Edge Map Measure," *submitted to IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2014.
- [5] K. Panetta, C. Gao, and S. Agaian, "Human Visual System Inspired Underwater Image Quality Measures," *submitted to IEEE Journal of Oceanic Engineering*, 2015.
- [6] C. Gao, K. Panetta, and S. Agaian, "New edge detection algorithms using alpha weighted quadratic filter," in *IEEE International Conference on Systems, Man, and Cybernetics 2011*, pp. 3167-3172.
- [7] C. Gao, K. Panetta, and S. Agaian, "A new color contrast enhancement algorithm for robotic applications," in *IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*, 2012, pp. 42-47.
- [8] C. Gao, K. Panetta, and S. Agaian, "Three dimensional alpha weighted quadratic filter based image color contrast enhancement," in *SPIE Defense, Security, and Sensing*, 2013, pp. 875514-875514-12.
- [9] C. Gao, K. Panetta, and S. Agaian, "No reference color image quality measures," in *2013 IEEE International Conference on Cybernetics (CYBCONF)*, 2013, pp. 243-248.
- [10] C. Gao, K. Panetta, and S. Agaian, "Color image attribute and quality measurements," in *SPIE Mobile Multimedia/Image Processing, Security, and Applications*, 2014, pp. 91200T-91200T-14.
- [11] C. Gao, K. Panetta, and S. Agaian, "Color image retrieval and analysis," in *Mobile, Multimedia, Image Processing, Security and Applications*, Baltimore, 2015.
- [12] C. Gao, K. Panetta, and S. Agaian, "Robust template based corner detection algorithms for robotic vision," in *IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*, 2015.
- [13] C. Gao, K. Panetta, and S. Agaian, "A combined colorfulness and color tone measure and its applications," in *IEEE International Conference on Image Processing*, 2015.
- [14] C. Gao, K. Panetta, and S. Agaian, "Coefficients training methods for linear combinations of image attribute measures," in *IEEE International Conference on Imaging Systems and Techniques*, Macau, China, 2015.
- [15] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, April, 2004.
- [16] P. ITU-T RECOMMENDATION, "Subjective video quality assessment methods for multimedia applications," 1999.
- [17] D. Chandler and S. Hemami, "VSNR: A Wavelet-Based Visual Signal-to-Noise Ratio for Natural Images," *IEEE Transactions on Image Processing*, vol. 16, September, 2007.
- [18] E. A. Silva, K. Panetta, and S. S. Agaian, "Quantifying image similarity using measure of enhancement by entropy," *Mobile Multimedia/Image Processing for Military and Security Applications*, vol. 6579, p. 65790U, 2007.
- [19] S. Nercessian, S. Agaian, and K. Panetta, "An Image Similarity Measure Using Enhanced Human Visual System Characteristics," in *Mobile Multimedia/Image Processing, Security, and Applications*, 2011.
- [20] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? A new look at signal fidelity measures," *Signal Processing Magazine, IEEE*, vol. 26, pp. 98-117, 2009.

- [21] H.-S. Han, D.-O. Kim, and R.-H. Park, "Structural information-based image quality assessment using LU factorization," *Consumer Electronics, IEEE Transactions on*, vol. 55, pp. 165-171, 2009.
- [22] S. S. Agaian, B. Silver, and K. A. Panetta, "Transform coefficient histogram-based image enhancement algorithms using contrast entropy," *Image Processing, IEEE Transactions on*, vol. 16, pp. 741-758, 2007.
- [23] K. Panetta, E. Wharton, and S. Agaian, "Human visual system-based image enhancement and logarithmic contrast measure," *IEEE Transactions on Systems, Man, Cybernetics*, vol. 38, Feb. 2008.
- [24] J. Caviedes and F. Oberti, "A new sharpness metric based on local kurtosis, edge and energy information," *Signal Processing: Image Communication*, vol. 19, pp. 147-161, 2004.
- [25] Y. Zhou, K. Panetta, S. Agaian, and H. Jia, "Nonlinear unsharp masking for mammogram enhancement," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, pp. 918-928, November, 2011.
- [26] J. Tang, E. Peli, and S. Acton, "Image enhancement using a contrast measure in the compressed domain," *Signal Processing Letters, IEEE*, vol. 10, pp. 289-292, 2003.
- [27] A. Maalouf and M.-C. Larabi, "A no reference objective color image sharpness metric," in *EUSIPCO*, 2010, pp. 1019-1022.
- [28] E. Wharton, K. Panetta, and S. Agaian, "A logarithmic measure of image enhancement," in *Mobile Multimedia/Image Processing for Military and Security Applications*, 2006.
- [29] Z. Wang, A. C. Bovik, and L. Lu, "Why is image quality assessment so difficult?," in *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, 2002, pp. IV-3313-IV-3316.
- [30] P. Jackman, D.-W. Sun, and P. Allen, "Recent advances in the use of computer vision technology in the quality assessment of fresh meats," *Trends in Food Science & Technology*, vol. 22, pp. 185-197, 2011.
- [31] M. Sarifuddin and R. Missaoui, "A new perceptually uniform color space with associated color similarity measure for content-based image and video retrieval," in *Proc. of ACM SIGIR 2005 Workshop on Multimedia Information Retrieval (MMIR 2005)*, 2005, pp. 1-8.
- [32] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *Journal of VLSI signal processing systems for signal, image and video technology*, vol. 38, pp. 35-44, 2004.
- [33] Z.-U. Rahman, D. J. Jobson, and G. A. Woodell, "Multi-scale retinex for color image enhancement," in *Image Processing, 1996. Proceedings., International Conference on*, 1996, pp. 1003-1006.
- [34] R. C. Gonzalez, *Digital image processing*: Pearson Education India, 2009.
- [35] A. Ford and A. Roberts, "Colour space conversions," *Westminster University, London*, vol. 1998, pp. 1-31, 1998.
- [36] T. Smith and J. Guild, "The CIE colorimetric standards and their use," *Transactions of the Optical Society*, vol. 33, p. 73, 1931.
- [37] W. D. Wright, "A re-determination of the trichromatic coefficients of the spectral colours," *Transactions of the Optical Society*, vol. 30, p. 141, 1929.
- [38] J. Guild, "The colorimetric properties of the spectrum," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, pp. 149-187, 1932.
- [39] M. D. Fairchild, *Color appearance models*: John Wiley & Sons, 2013.
- [40] L. M. Hurvich and D. Jameson, "An opponent-process theory of color vision," *Psychological review*, vol. 64, p. 384, 1957.

- [41] M. Liévin and F. Luthon, "Nonlinear color space and spatiotemporal MRF for hierarchical segmentation of face features in video," *Image Processing, IEEE Transactions on*, vol. 13, pp. 63-71, 2004.
- [42] G. Deng and J.-C. Pinoli, "Differentiation-based edge detection using the logarithmic image processing model," *Journal of Mathematical Imaging and Vision*, vol. 8, pp. 161-180, 1998.
- [43] D. Hasler and S. E. Suesstrunk, "Measuring colorfulness in natural images," in *Electronic Imaging 2003*, 2003, pp. 87-95.
- [44] Y.-Y. Fu, "Color image quality measures and retrieval," Ph.D, Department of Computer Science, New Jersey Institute of Technology, January 2006.
- [45] J. Mukherjee and S. K. Mitra, "Enhancement of color images by scaling the DCT coefficients," *Image Processing, IEEE Transactions on*, vol. 17, pp. 1783-1794, 2008.
- [46] S. Winkler, "Visual fidelity and perceived quality: Toward comprehensive metrics," in *Photonics West 2001-Electronic Imaging*, 2001, pp. 114-125.
- [47] N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola, M. Carli, and F. Battisti, "Color image database for evaluation of image quality metrics," in *Multimedia Signal Processing, 2008 IEEE 10th Workshop on*, 2008, pp. 403-408.
- [48] M. Chambah, D. Semani, A. Renouf, P. Courtellemont, and A. Rizzi, "Underwater color constancy: enhancement of automatic live fish recognition," in *Electronic Imaging 2004*, 2003, pp. 157-168.
- [49] S. Winkler, "Color Space Conversions," *Digital Video Quality*, pp. 155-156.
- [50] S. Nercessian, "Human visual system-based multi-scale tools with biomedical and security applications," 2012.
- [51] G. Buchsbaum, "An analytical derivation of visual nonlinearity," *Biomedical Engineering, IEEE Transactions on*, pp. 237-242, 1980.
- [52] S. S. Agaian, K. P. Lentz, and A. M. Grigoryan, "A new measure of image enhancement," in *IASTED International Conference on Signal Processing & Communication*, 2000, pp. 19-22.
- [53] S. DelMarco and S. Agaian, "The design of wavelets for image enhancement and target detection," *Mobile Multimedia/Image Processing, Security, and Applications*, pp. 735103-12, 2009.
- [54] S. S. Agaian, "Visual morphology," in *Electronic Imaging'99*, 1999, pp. 139-150.
- [55] E. Peli, "Contrast in complex images," *Journal of Optical Science of America*, vol. 7, October 1990.
- [56] A. Samani, K. Panetta, and S. Agaian, "Transform domain measure of enhancement—TDME—for security imaging applications," in *Technologies for Homeland Security (HST), 2013 IEEE International Conference on*, 2013, pp. 265-270.
- [57] J. Caviedes and S. Gurbuz, "No-reference sharpness metric based on local edge kurtosis," in *Image Processing. 2002. Proceedings. 2002 International Conference on*, 2002, pp. III-53-III-56 vol. 3.
- [58] P. V. Vu and D. M. Chandler, "A fast wavelet-based algorithm for global and local image sharpness estimation," *Signal Processing Letters, IEEE*, vol. 19, pp. 423-426, 2012.
- [59] C. T. Vu, T. D. Phan, and D. M. Chandler, "S3: A spectral and spatial measure of local perceived sharpness in natural images," *IEEE Transactions on Image Processing*, vol. 21, pp. 934-945, 2012.
- [60] J. Sanderson and Y. Liang, "No-reference image quality measurement for low-resolution images," in *SPIE Defense, Security, and Sensing*, 2013, pp. 874404-874404-17.
- [61] K. Zhu, V. Asari, and D. Saupe, "No-reference quality assessment of H. 264/AVC encoded video based on natural scene features," in *SPIE Defense, Security, and Sensing*, 2013, pp. 875505-875505-11.

- [62] E. Ong, W. Lin, Z. Lu, X. Yang, S. Yao, F. Pan, L. Jiang, and F. Moschetti, "A no-reference quality metric for measuring image blur," in *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, 2003, pp. 469-472.
- [63] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: application to JPEG2000," *Signal Processing: Image Communication*, vol. 19, pp. 163-172, 2004.
- [64] D. Shaked and I. Tastl, "Sharpness measure: Towards automatic image enhancement," in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, 2005, pp. 1-937-40.
- [65] K. V. Mardia, "Measures of multivariate skewness and kurtosis with applications," *Biometrika*, vol. 57, pp. 519-530, 1970.
- [66] N. Zhang, A. Vladar, M. Postek, and B. Larrabee, "A kurtosis-based statistical measure for two-dimensional processes and its application to image sharpness," *Proceedings of Section of Physical and Engineering Sciences of American Statistical Society*, pp. 4730-4736, 2003.
- [67] N. B. Nill and B. Bouzas, "Objective image quality measure derived from digital image power spectra," *Optical Engineering*, vol. 31, pp. 813-825, 1992.
- [68] W. Lin and C.-C. Jay Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication and Image Representation*, vol. 22, pp. 297-312, 2011.
- [69] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," *Consumer Electronics, IEEE Transactions on*, vol. 45, pp. 68-75, 1999.
- [70] M. Kim and M. Chung, "Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement," *Consumer Electronics, IEEE Transactions on*, vol. 54, pp. 1389-1397, 2008.
- [71] S.-D. Chen and A. R. Ramli, "Minimum mean brightness error bi-histogram equalization in contrast enhancement," *Consumer Electronics, IEEE Transactions on*, vol. 49, pp. 1310-1319, 2003.
- [72] C. Wang and Z. Ye, "Brightness preserving histogram equalization with maximum entropy: a variational perspective," *Consumer Electronics, IEEE Transactions on*, vol. 51, pp. 1326-1334, 2005.
- [73] C. H. Ooi, N. S. P. Kong, and H. Ibrahim, "Bi-histogram equalization with a plateau limit for digital image enhancement," *Consumer Electronics, IEEE Transactions on*, vol. 55, pp. 2072-2080, 2009.
- [74] F. Luthon, B. Beaumesnil, and N. Dubois, "LUX color transform for mosaic image rendering," in *Automation Quality and Testing Robotics (AQTR), 2010 IEEE International Conference on*, 2010, pp. 1-6.
- [75] B. Bringier, N. Richard, M. C. Larabi, and C. Fernandez-Maloigne, "No-reference perceptual quality assessment of colour image," in *Proceedings of the European Signal Processing Conference (EUSIPCO'06)*, 2006.
- [76] K. Panetta, S. Agaian, Y. Zhou, and E. Wharton, "Parameterized logarithmic framework for image enhancement," *IEEE Transactions on Systems, Man, Cybernetics*, vol. 41, pp. 460-473, 2011.
- [77] E. Demidenko, *Mixed models: theory and applications* vol. 518: John Wiley & Sons, 2005.
- [78] I. Lawrence and K. Lin, "A concordance correlation coefficient to evaluate reproducibility," *Biometrics*, pp. 255-268, 1989.
- [79] J. B. Carroll, "The nature of the data, or how to choose a correlation coefficient," *Psychometrika*, vol. 26, pp. 347-372, 1961.
- [80] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images," in *Image Processing. 2002. Proceedings. 2002 International Conference on*, 2002, pp. 1-477-1-480 vol. 1.

- [81] K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for color research," *Color Research & Application*, vol. 27, pp. 147-151, 2002.
- [82] Z.-u. Rahman, D. J. Jobson, and G. A. Woodell, "Retinex processing for automatic image enhancement," *IS&T/SPIE Electronic Imaging*, vol. 13, Jan 2004.
- [83] J. Mukhopadhyay and S. K. Mitra, "Color enhancement in the compressed domain," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2008, pp. 3144-3147.
- [84] S. Aghagolzadeh and O. K. Ersoy, "Transform image enhancement," *Optical Engineering*, vol. 31, pp. 614-626, 1992.
- [85] S. Lee, "An efficient content-based image enhancement in the compressed domain using retinex theory," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, pp. 199-213, 2007.
- [86] A. Rosenfeld and A. C. Kak, *Digital picture processing* vol. 1: Elsevier, 2014.
- [87] M. Tabb and N. Ahuja, "Multiscale image segmentation by integrated edge and region detection," *IEEE Transactions on Image Processing*, vol. 6, pp. 642-655, 1997.
- [88] R. Qian and T. Huang, "Optimal edge detection in two-dimensional images," *IEEE Transactions on Image Processing*, , vol. 5, pp. 1215-1220, 1996.
- [89] D. Demigny, "On optimal linear filtering for edge detection," *IEEE Transactions on Image Processing*, vol. 11, pp. 728-737, 2002.
- [90] H. Moon, R. Chellappa, and A. Rosenfeld, "Optimal edge-based shape detection," *IEEE Transactions on Image Processing*,, vol. 11, pp. 1209-1227, 2002.
- [91] S. Ghosal and R. Mehrotra, "Detection of composite edges," *IEEE Transactions on Image Processing*, vol. 3, pp. 14-25, 1994.
- [92] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Robust face recognition for uncontrolled pose and illumination changes," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, pp. 149-163, 2013.
- [93] M. Suresh, R. Stoleru, E. Zechman, and B. Shihada, "On event detection and localization in acyclic flow networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, pp. 708-723, 2013.
- [94] R. Labati, A. Genovese, V. Piuri, and F. Scotti, "Wildfire smoke detection using computational intelligence techniques enhanced with synthetic smoke plume generation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, pp. 1003-1012, 2013.
- [95] J. García, A. Gardel, I. Bravo, J. L. Lázaro, and M. Martínez, "Tracking People Motion Based on Extended Condensation Algorithm," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, pp. 606-618, 2013.
- [96] L. Ke, R. Wang, X. Wang, and R. Tian, "Medical Image Enhancement Based on HVC," in *The 1st International Conference on Bioinformatics and Biomedical Engineering*, 2007, pp. 995-998.
- [97] Y. Yu and S. T. Acton, "Edge detection in ultrasound imagery using the instantaneous coefficient of variation," *IEEE Transactions on Image Processing*, vol. 13, pp. 1640-1655, 2004.
- [98] Y. Zhao, W. Gui, Z. Chen, J. Tang, and L. Li, "Medical images edge detection based on mathematical morphology," in *27th Annual International Conference of the Engineering in Medicine and Biology Society*, 2006, pp. 6492-6495.
- [99] J. Parker, *Algorithms for image processing and computer vision*: John Wiley & Sons, 2010.
- [100] L. Middleton and J. Sivaswamy, "Edge detection in a hexagonal-image processing framework," *Image and Vision Computing*, vol. 19, pp. 1071-1081, 2001.
- [101] Q. Zhu, "Improving edge detection by an objective edge evaluation," in *Proceedings of the 1992 ACM/SIGAPP Symposium on Applied computing: technological challenges of the 1990's*, 1992, pp. 459-468.

- [102] T. Kanungo, M. Jaisimha, J. Palmer, and R. Haralick, "A methodology for quantitative performance evaluation of detection algorithms," *IEEE Transactions on Image Processing*, vol. 4, pp. 1667-1674, 1995.
- [103] M. Heath, S. Sarkar, T. Sanocki, and K. Bowyer, "A robust visual method for assessing the relative performance of edge-detection algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 1338-1359, 1997.
- [104] W. Pratt, *Digital image processing* vol. 242. New York: Wiley Interscience, 1978.
- [105] A. Pinho and L. Almeida, "Edge detection filters based on artificial neural networks," in *Image Analysis and Processing*, 1995, pp. 159-164.
- [106] Y. Yitzhaky and E. Peli, "A method for objective edge detection evaluation and detector parameter selection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1027-1033, 2003.
- [107] S. Nercessian, S. Agaian, and K. Panetta, "A new reference-based measure for objective edge map evaluation," in *SPIE Defense, Security, and Sensing Conferences*, 2009, pp. 73510J-73510J-12.
- [108] Q. Zhu, "Efficient evaluations of edge connectivity and width uniformity," *Image and Vision Computing*, vol. 14, pp. 21-34, 1996.
- [109] I. Boaventura and A. Gonzaga, "Method to evaluate the performance of edge detector," presented at the The XXIIInd Brazilian Symposium on Computer Graphics and Image Processing, Brazil, 2009.
- [110] P. Rockett, "The accuracy of sub-pixel localisation in the Canny edge detector," in *British Machine Vision Conference*, 1999, pp. 1-10.
- [111] R. Kakarala and A. Hero, "On achievable accuracy in edge localization," in *International Conference on Acoustics, Speech, and Signal Processing*, 1991, pp. 2545-2548.
- [112] H. Tagare and R. deFigueiredo, "On the localization performance measure and optimal edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 1186-1190, 1990.
- [113] J. Koplowitz and V. Greco, "On the edge location error for local maximum and zero-crossing edge detectors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, pp. 1207-1212, 1994.
- [114] S. Coleman, B. Scotney, and D. Kerr, "Integrated edge and corner detection," in *14th International Conference on Image Analysis and Processing 2007*, pp. 653-658.
- [115] P. Rosin, "Measuring corner properties," *Computer Vision and Image Understanding*, vol. 73, pp. 291-307, 1999.
- [116] F. Mokhtarian and F. Mohanna, "Performance evaluation of corner detectors using consistency and accuracy measures," *Computer Vision and Image Understanding*, vol. 102, pp. 81-94, 2006.
- [117] M. Prieto and A. Allen, "A similarity metric for edge images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1265-1273, 2003.
- [118] S. Kamalakannan, A. Gururajan, H. Sari-Sarraf, R. Long, and S. Antani, "Double-edge detection of radiographic lumbar vertebrae images using pressurized open DGVF snakes," *IEEE Transactions on Biomedical Engineering*, vol. 57, pp. 1325-1334, 2010.
- [119] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 679-698, 1986.
- [120] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 207, pp. 187-217, 1980.

- [121] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *8th IEEE International Conference on Computer Vision*, 2001, pp. 416-423.
- [122] S. C. Nercessian, S. S. Agaian, and K. A. Panetta, "A new reference-based measure for objective edge map evaluation," in *SPIE Defense, Security, and Sensing*, 2009, pp. 73510J-73510J-12.
- [123] S. Nercessian, K. Panetta, and S. Agaian, "A non-reference measure for objective edge map evaluation," in *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, 2009, pp. 4563-4568.
- [124] Y. Yitzhaky and E. Peli, "A method for objective edge detection evaluation and detector parameter selection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, pp. 1027-1033, 2003.
- [125] B. Govindarajan, K. A. Panetta, and S. Agaian, "Image reconstruction for quality assessment of edge detectors," in *Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on*, 2008, pp. 691-696.
- [126] J. Bednar and T. Watt, "Alpha-trimmed means and their relationship to median filters," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 32, pp. 145-153, 1984.
- [127] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *Image Processing, IEEE Transactions on*, vol. 10, pp. 1200-1211, 2001.
- [128] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, pp. 600-612, 2004.
- [129] G.-H. Chen, C.-L. Yang, and S.-L. Xie, "Gradient-based structural similarity for image quality assessment," in *Image Processing, 2006 IEEE International Conference on*, 2006, pp. 2929-2932.
- [130] S. Nercessian, S. S. Agaian, and K. A. Panetta, "An image similarity measure using enhanced human visual system characteristics," in *SPIE Defense, Security, and Sensing*, 2011, pp. 806310-806310-9.
- [131] G. Sicuranza, "Quadratic filters for signal processing," *Proceedings of the IEEE*, vol. 80, pp. 1263-1285, 1992.
- [132] T. Ogunfunmi, *Adaptive nonlinear system identification: The Volterra and Wiener model approaches*: Springer, 2007.
- [133] G. Ramponi, "Bi-impulse response design of isotropic quadratic filters," *Proceedings of the IEEE*, vol. 78, pp. 665-677, 1990.
- [134] P. Fontanot and G. Ramponi, "A polynomial filter for the preprocessing of mail address images," in *Nonlinear Digital Signal Processing, 1993. IEEE Winter Workshop on*, 1993, pp. 2.1_6. 1-2.1_6. 6.
- [135] J. Han, S. Yang, and B. Lee, "A novel 3D color histogram equalization method with 1D uniform gray scale histogram," *IEEE Transactions on Image Processing*, vol. 20, pp. 506-512, Feb. 2011.
- [136] P. E. Trahanias and A. N. Venetsanopoulos, "Color image enhancement through 3-D histogram equalization," in *Pattern Recognition, Conference C: Image, Speech and Signal Analysis, Proceedings., 11th IAPR International Conference on 1992*, pp. 545-548.
- [137] S. K. Naik and C. A. Murthy, "Hue-preserving color image enhancement without gamut problem," *IEEE Transactions on Image Processing*, vol. 12, pp. 1591-1598, 2003.
- [138] P. A. Mlsna, Z. Qiang, and J. J. Rodriguez, "3-D histogram modification of color images," in *Image Processing, Proceedings., International Conference on 1996*, pp. 1015-1018.
- [139] E. Land and J. McCann, "Lightness and retinex theory," *Journal of Optical Society of America*, vol. 61, pp. 1-11, Jan. 1971.

- [140] E. H. Land, "The retinex theory of color vision " *Scientific American*, vol. 237, pp. 108-128, 1977.
- [141] Z. Rahman, D. Jobson, and G. A. Woodell, "Multiscale retinex for color image enhancement," in *Proceedings of the IEEE International Conference on Image Processing*, 1996.
- [142] D. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing*, vol. 6, pp. 965-976, July 1997.
- [143] A. Rizzi, D. Marini, L. Rovati, and F. Docchio, "Unsupervised corrections of unknown chromatic dominants using a Brownian-path-based Retinex algorithm," *Journal of Electronic Imaging*, vol. 12, pp. 431-441, July 2003.
- [144] D. Jobson, Z. Rahman, and G. Woodell, "Properties and performance of a center/surround retinex," *IEEE Transactions on Image Processing*, vol. 6, pp. 451-462, March, 1997.
- [145] Y. Wang and Y. Luo, "Adaptive color contrast enhancement for digital images," *Optical Engineering*, vol. 50, Nov.,2011.
- [146] Y. Wang and Y. Luo, "Balanced color contrast enhancement for digital images," *Optical Engineering*, vol. 51, October 2012.
- [147] K. Naka and W. Rushton, "S-potentials from colour units in the retina of fish (Cyprinidae)," *The Journal of physiology*, vol. 185, pp. 536-555, 1966.
- [148] Rastislav.Lukac, "New framework for automatic white balancing of digital camera images," *Signal Processing*, vol. 88, pp. 582-593, 2008.
- [149] *NASA Retinex Image Processing* Available: <http://dragon.larc.nasa.gov/retinex/>
- [150] K.-I. Naka and W. A. Rushton, "S-potentials from luminosity units in the retina of fish (Cyprinidae)," *The Journal of physiology*, vol. 185, pp. 587-599, 1966.
- [151] S. S. CL Novak, "Color edge detection," in *Proc. DARPA Image Understanding Workshop*, Los Angeles, February 1987, pp. 35-37.
- [152] *Fusion Toolbox*. Available: <http://www.metapix.de/fusion.htm>
- [153] F. Sadjadi, "Comparative image fusion analysais," presented at the Computer Vision and Pattern Recognition - Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on, 2005.
- [154] A. Toet, "Merging thermal and visual images by a contrast pyramid," *Optical Engineering*, vol. 28, pp. 789-792, 1989.
- [155] Z.-u. Rahman, D. J. Jobson, and G. A. Woodell, "Multiscale retinex for color rendition and dynamic range compression," in *SPIE's 1996 International Symposium on Optical Science, Engineering, and Instrumentation*, 1996, pp. 183-191.
- [156] L. Meylan and S. Susstrunk, "High dynamic range image rendering with a retinex-based adaptive filter," *IEEE Transactions on Image Processing*, vol. 15, pp. 2820-2830, 2006.
- [157] D. Ziou and S. Tabbone, "Edge detection techniques-an overview," *Pattern Recognition And Image Analysis C/C Of Raspoznavaniye Obrazov I Analiz Izobrazhenii*, vol. 8, pp. 537-559, 1998.
- [158] M. Bennamoun, "Edge detection: Problems and solutions," in *Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on*, 1997, pp. 3164-3169.
- [159] S. K. Mitra, H. Li, S. Lin, and T.-h. Yu, "A new class of nonlinear filters for image enhancement," in *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*, 1991, pp. 2525-2528.
- [160] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, pp. 23-27, 1975.
- [161] P. Fontanot and G. Ramponi, "A polynomial filter for the preprocessing of mail address images," in *Proceeding of IEEE Winter Workshop on Nonlinear Digital Signal Processing 1993*, pp. 2.1_6.1 - 2.1_6.6

- [162] Y. Zhou, K. Panetta, and S. Agaian, "Mammogram enhancement using alpha weighted quadratic filter," presented at the Annual International Conference of IEEE Engineering in Medicine and Biology Society, Minneapolis, Minnesota, 2009.
- [163] F. Yang, Y. Chang, and S. Wan, "Gradient-threshold edge detection based on the human visual system," *Optical Engineering*, vol. 44, pp. 020505-020505-2, 2005.
- [164] A. K. Moorthy, P. Obrador, and N. Oliver, "Towards computational models of the visual aesthetic appeal of consumer videos," in *Computer Vision—ECCV 2010*, ed: Springer, 2010, pp. 1-14.
- [165] P. Obrador and N. Moroney, "Low level features for image appeal measurement," in *IS&T/SPIE Electronic Imaging*, 2009, pp. 72420T-72420T-12.
- [166] J. Li and J. Z. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, pp. 1075-1088, 2003.
- [167] M. Roopaei, S. Agaian, and M. Jamshidi, "Thermal imaging in fuzzy condition monitoring," in *World Automation Congress (WAC), 2014*, 2014, pp. 593-597.